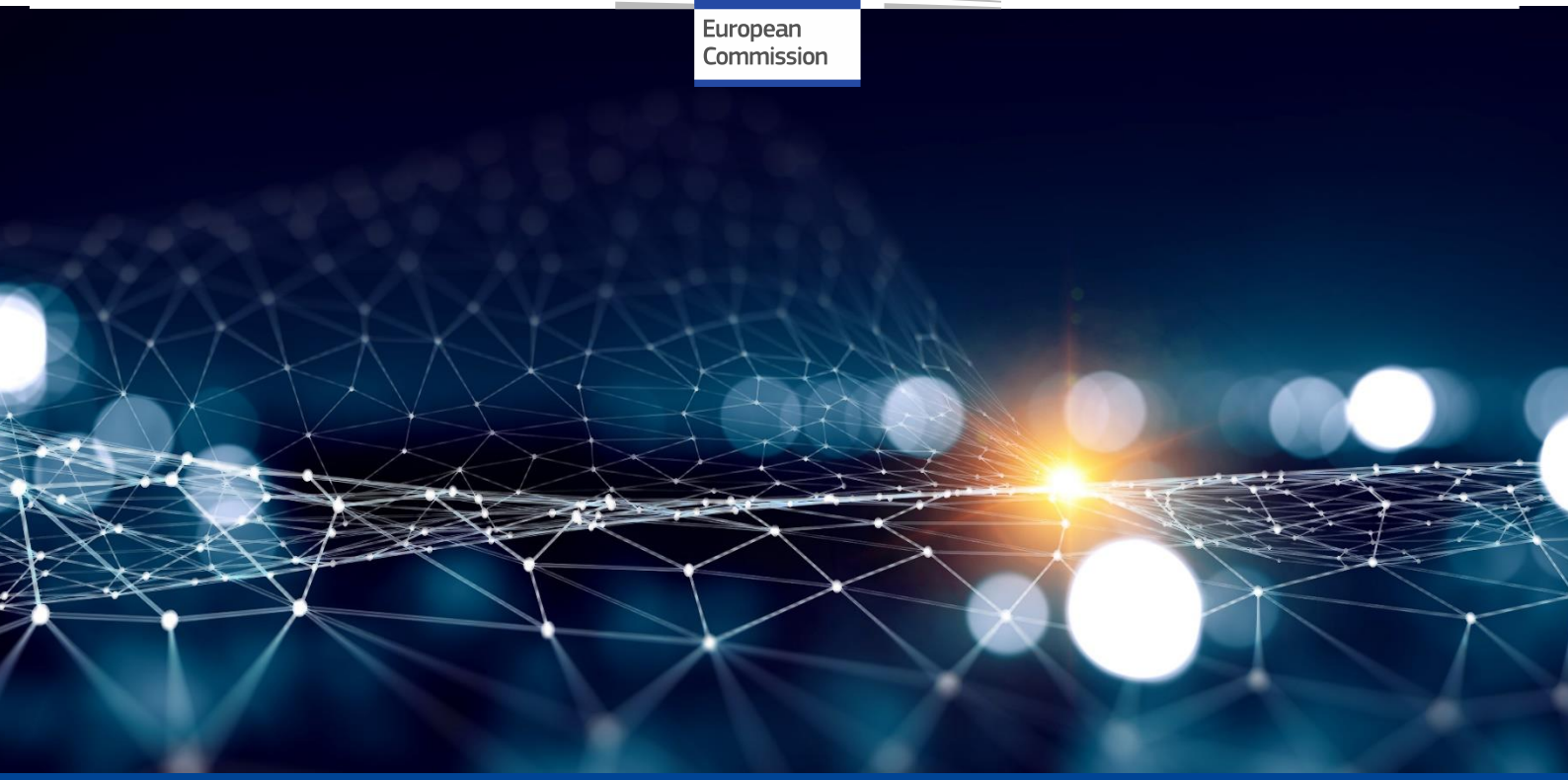
European Commission



European

DATA Market Study 2021–2023

CNECT/LUX/2020/OP/0027–VIGIE 2020-0655, contract number: LC-01568518

D2.1 FIRST REPORT ON FACTS AND FIGURES

February, 2022

IDC theLisboncouncil Lisbon Council Research

Author(s)	IDC: Mike Glennon, Marianne Kolding, Carla La Croce; Giorgio Micheletti; Nevena Raczko; Martin Sunblad. Lisbon Council: Cristina Moise; David Osimo.
Deliverable	**D2.1 First Report on Facts and Figures**
Date of delivery	31.01.2022
Version	1.2
Addressee officer	Katalin IMREI Policy Officer European Commission, DG CONNECT Unit G1 — Data Policy and Innovation EUFO 1/265, L–2557 Luxembourg/Gasperich katalin.imrei@ec.europa.eu
Contract ref.	LC-01568518

Table of Contents

Table of Tables

Table of Figures

ESSENTIAL GLOSSARY – THE KEY INDICATORS

Data professionals[1] are workers who collect, store, manage, and/or analyse, interpret, and visualise data as their primary activity or as a relevant part of their activity. Data professionals must be proficient with the use of structured and unstructured data, should be able to work with a huge amount of data, and should be familiar with emerging database technologies. For 2021–2023, the definition of data professionals was refined to differentiate the roles played by different data users: These are Data Technical Professionals, Data Business Professionals, and Data Consumers.

Data technical professionals are specialists in the collection, storage, management, modelling, and quality assurance of data, as well as the integration of various data sources, to ensure consistency, accuracy, and quality of data. A data technical professional can, given the question that needs to be answered, ensure that the data supply chain is provided and that it is accurate.

Data business professionals have as a primary or significant focus the task of performing predictive analysis, qualitative analysis, data modelling, data extraction, and data summaries with the purpose of creating new insights and knowledge from available data. They have thorough industry and/or process understanding and can put data analysis into context and relate to existing trends within the industry or line of business they are in. They typically leave collection, management, and quality of data to the data technical professional but, using analysis tools such as Excel, Tableau, and Power BI, are able to summarise large amounts of data and to visualise and present trends and insights to a wider audience of key stakeholders in the business in order to drive the strategic decision-making process in the organisation. Data scientists predominantly reside within the data business professional group.

Data consumers are product, process, human-resource, asset, or department employees and managers responsible for driving change or maintaining a position whereby decision making is heavily reliant on the supply of data and insights based on large amounts of data. They work directly with data only part of the time. They are decision makers or stakeholders in a decision process whereby the data and insights provided determine the quality of the decisions made. A data consumer guides the business based on the data and insights provided through the data supply chain.

Data companies are organisations that are directly involved in the production, delivery, and/or usage of data in the form of digital products, services, and technologies. They can be both data suppliers' and data users' organisations:

- **Data suppliers** have as their main activity the production and delivery of digital data-related products, services, and technologies. They represent the supply side of the data market.
- **Data users** are organisations that generate, exploit, collect, and analyse digital data intensively and use what they learn to improve the business. They represent the demand side of the data market.

Data companies' revenues correspond to the aggregated value of all the data-related products and services generated by Europe-based data suppliers, including exports outside the EU. This indicator measures the revenues of the data suppliers identified and classified by Indicator 2 (see the products

[1] The European Data Market Study (SMART 2013/0063) included an indicator measuring "Data Workers", which was based on a similar, but slightly more restrictive, definition. In the subsequent European Data Market Study Update (SMART2016/0063) we measured "Data Professionals" – that is, workers with a wider range of data-related roles. In this context, data professionals are not only data technicians, but also users who, based on sophisticated tools, take decisions about their business or activities after having analysed and interpreted the available data.

and services specified in our definition of the data market). Data companies' revenues do not include data monetisation as part of the data market.

The **data market** is the marketplace where digital data is exchanged as “products” or “services” as a result of the elaboration of raw data. The data market captures the aggregate value of the demand of digital data without measuring the direct, indirect, or induced impacts of data in the economy as a whole. The value of the data market is not exactly equal to the aggregated revenues of European data companies because it includes imports (data products and services bought on the global digital market from suppliers not based in Europe) and excludes the exports of the European data companies. In this report, we add to the data market an estimate for the value of data monetisation.

The **data economy** measures the overall impacts of the data market on the economy as a whole. It involves the generation, collection, storage, processing, distribution, analysis elaboration, delivery, and exploitation of data enabled by digital technologies. The data economy captures a wider concept than the data market only, as it considers the value and wealth generated in the economy as a whole (not just across businesses) by the exploitation of data.

EXECUTIVE SUMMARY

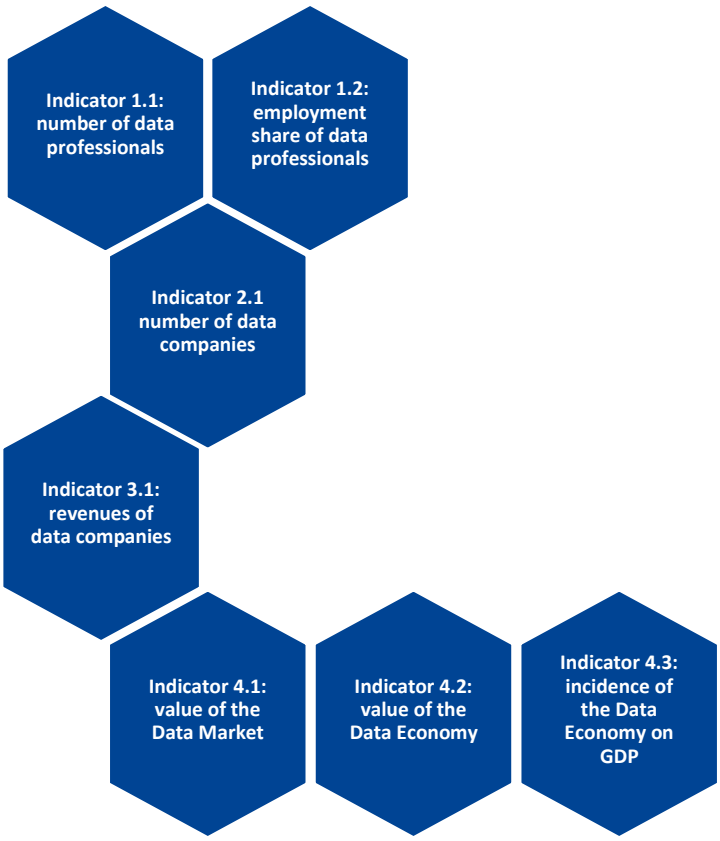
This First Report on Facts & Figures (Deliverable D2.1 of the European Data Market Study; 2021–2023, SMART VIGIE 2020-0655) presents the results obtained through the first round of measurements of the European Data Market Monitoring Tool. It was designed and developed under the previous European Data Market Studies (SMART 2013/0063 and SMART 2016/0063) — in 2013 and 2016, respectively — and covered the period 2014–2019 with forecasts to the years 2020 and 2025 under three distinct scenarios. In this update, the Monitoring Tool has been revised and extended to cover the years 2019–2021, with forecasts to the year 2030 under the same alternative scenarios. When presenting the forecasts of the indicators measured in this report, this document includes not only the results for 2030 but also those for 2025 under the Baseline scenario. The objective of this is to show the whole timeframe encompassed by the European Data Market Monitoring Tool, as well as to ensure a high level of mutual intelligibility between the 2013 study, the 2016 study, and the 2021 study.

This First Report on Facts & Figures focuses on the following set of indicators:

Each indicator has been measured for the total EU27 and for all the EU27 Member States when available and applicable; industry-specific and company-size views are also offered, with indicators provided by industry sector and company size band, when possible. The UK and Switzerland were measured separately, as was the EEA (Norway, Iceland, and Lichtenstein), in an aggregated way.

As in the previous European Data Market Studies (SMART 2013/0063 and SMART 2016/0063), a select number of indicators has been developed and updated but, this time, for four non-European countries — namely, Brazil, Japan, the United States and China.

For each of these countries, this report presents the following indicators:



Methodological Note

The EDM study requires a complex mix of qualitative and quantitative methodologies to allow it to reach its objectives. While quantitative methodologies represent the most relevant part of the study, qualitative methodologies are indispensable to balance the statistical approach and provide the market and social intelligence needed to lead to policy insights and the development of sound scenarios.

Updating the European Data Market Monitoring Tool

In order to ensure the study's continuity, the methodological approach builds upon the previous EDM Monitoring Tool releases, including necessary improvements. The definition of the main indicators is revised and updated according to the maturing of the market and data economy. It is important to emphasise that the EDM Monitoring Tool's modular structure will not change in its core substance, as it has already proven to be sufficiently adaptable to the evolution of the data market and data economy in recent years.

The main methodological changes concern four of the main indicators: 1. measuring the value of data; 2. measuring data supplier and data user companies; 3. measuring of business impacts; 4. measuring data professionals.

Field Research: Ad-Hoc Survey

In July–September 2021, IDC undertook a field survey of a representative sample of 1,191 data users and data suppliers to collect data about the uptake of data technologies, products, and services across Europe to provide up-to-date and comprehensive base data for the models and to provide insight into European companies' views, usage, and plans for the deployment of big data, analytics and AI systems.

It surveyed objectives and barriers and, in particular, the issues and intentions around the availability and development of skills to develop and use analytics systems. It questioned users on their adoption of advanced analytics in a series of recently defined use cases. This analysis will further shed light on the use of data for strategic decision-making processes.

Importantly, the survey investigated the business impacts seen by respondents on organisation performance, time, and cost reduction, correlated with efficiency improvements resulting from the use of big data analytics and AI.

Three Developments Paths to 2030: Baseline Scenario, Challenge Scenario, and High Growth Scenario

This chapter focuses on the updated scenario methodology used in the project to develop three scenarios of the evolution of the European data market and economy to 2030, based on alternative development paths driven by different macroeconomic and framework conditions. This methodology builds on the successful approach applied in previous studies, with some improvements concerning the development of assumptions and their validation.

The three scenarios provide the storylines, the contextual framework, and the main assumptions that are used to model and forecast the EDM Monitoring Tool indicators, with a particular emphasis on the role of policies.

It is important to note that scenarios are not predictions but potential development paths: Their value-added lies especially in thinking through the potential consequences of different market trajectories and therefore providing a guide to action particularly for policy makers.

This analysis highlights the critical turning points to be faced in the coming years by governments, businesses, and social actors in the development of the European data economy. The combination of alternative social and economic trends results in the following scenarios:

- The **Baseline scenario** is characterised by a substantial expansion of data innovation, a moderate concentration of power by dominant data owners with a data governance framework safeguarding personal data rights, and an uneven but relatively wide distribution of data innovation benefits in society.

- The **High Growth scenario** is characterised by a high level of data innovation, low data power concentration, an open and transparent data governance model with high data sharing, and a wide distribution of data innovation benefits in society.

- The **Challenge scenario** is characterised by a low level of data innovation, a moderate level of data power concentration due to digital markets fragmentation, and an uneven distribution of data innovation benefits in society.

Measuring Data Professionals

Data professionals[2] *are workers who collect, store, manage, and/or analyse, interpret, and visualise data as their primary activity or as a relevant part of their activity. Data professionals must be proficient with the use of structured and unstructured data, should be able to work with a huge amount of data, and should be familiar with emerging database technologies. For 2021–2023, the definition of data professionals was refined to differentiate the roles played by different data users: These are Data Technical Professionals, Data Business Professionals, and Data Consumers. The measure of data professionals includes data technical professionals and data business professionals only.*

Data Professionals Forecasts: 2025 and Three 2030 Scenarios

We anticipate strong growth in the number of data professionals over the period of the forecast, from 2020 to 2030. We show three scenarios for 2030 to accommodate potential upsides and downsides to our forecast, as 2030 is 9 years away and as 2019 and 2020 showed significant changes to social, economic, technical, and political factors that can have dramatic impacts on the number of data professionals available. In particular, current demand for data professionals is high, and supply is not yet meeting this demand. The forecast shows fulfilled demand.

Data Professionals Forecast: 2025; 2030 Challenge, Baseline, and High Growth Scenarios (000's); and CAGRs (%)

	2025	2030, Challenge	2030, Baseline	2030, High Growth	CAGR: 2020–2025	CAGR: '25–'30, Challenge	CAGR: '25–'30, Baseline	CAGR: '25–'30, High Growth
EU27	8,158	8,770	9,630	11,437	4.6%	1.5%	3.4%	7.0%
EEA (NO, LI, IS) +CH	436	458	507	634	6.0%	1.0%	3.1%	7.8%
Total, all countries	10,806	11,490	12,701	14,997	4.6%	1.2%	3.3%	6.8%

Measuring Data Companies

Data companies *are organisations that are directly involved in the production, delivery, and/or usage of data in the form of digital products, services, and technologies. They can be both data supplier and data user organisations:*

- ***Data suppliers*** *have as their main activity the production and delivery of digital data-related products, services, and technologies. They represent the supply side of the data market.*
- ***Data users*** *are organisations that generate, exploit collect, and analyse digital data intensively and use what they learn to improve their business. They represent the demand side of the data market.*

[2] The previous European Data Market Study (SMART 2013/0063) included an indicator measuring "data workers", which was based on a similar, but slightly more restrictive, definition. In line with the First Report on Facts & Figures (D2.1), in this document, we measure "data professionals" – that is, workers with a wider range of data-related roles. Indeed, data professionals are not only data technicians, but also users who, based on sophisticated tools, take decisions about their business or activities after having analysed and interpreted the available data.

Forecasting Data Supplier Companies: 2025 and Three 2030 Scenarios

The table below shows the forecast out to 2025 and 2030 for the number of data supplier companies in the 27 Member States and on the greater European continent. We anticipate this market for the EU27 will grow at 8.8% compound growth to 2025 and slow to 3.1% compound growth by 2030 (Baseline). However, this growth to 2030 could be as low as 2.3% or as high as 4.3%, depending on the scenario. Growth in the number of data suppliers will depend significantly on the rise in the number of companies that monetise data in data markets. This is an early and emerging market but shows great promise out to 2030, as organisations understand and realise the value their data has.

Data Supplier Companies Forecasts: 2025, Three 2030 Scenarios, and Growth (%)

	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR 2020–2025	CAGR '25–30, Challenge	CAGR '25–30, Baseline	CAGR '25–30, High Growth
EU27	252,791	283,084	295,043	311,397	8.8%	2.3%	3.1%	4.3%
EEA (NO, LI, IS) + CH	15,247	17,080	17,964	19,080	8.6%	2.3%	3.3%	4.6%
Total, all countries	505,562	564,976	584,542	614,069	8.7%	2.2%	2.9%	4.0%

Forecasting Data User Companies: 2025 and Three 2030 Scenarios

Data user companies forecast shows higher growth over the period of the forecast when compared with data supplier companies as the data economy begins to drive its way into all business. Many more companies will embrace data as a basis for business, but there will still be latency from those organisations that use data to fundamentally change the way business is conducted. Most will use data to reduce costs and/or raise efficiency, although there are signs that more companies are using data to develop new products and to change how business is conducted. Those companies that can fundamentally change how their business is conducted are more likely to be successful, although often it is the pioneers that are taken over by others. The table below shows the forecast for data user companies for 2025 and three scenarios for 2030.

Data User Companies Forecasts: 2025, Three 2030 Scenarios, and Growth (%)

	2025	2030, Challenge Scenario	2030, Baseline Scenario	2030, High Growth Scenario	CAGR: 2020–2025	CAGR: '25–'30, Challenge	CAGR: '25–'30, Baseline	CAGR: '25–'30, High Growth
EU27	633,359	753,920	898,220	1,086,306	3.1%	3.5%	7.2%	11.4%
EEA (NO, LI, IS) + CH	27,174	32,334	38,832	47,197	3.0%	3.5%	7.4%	11.7%
Total, all countries	875,394	1,041,537	1,237,203	1,492,761	3.2%	3.5%	7.2%	11.3%

Measuring Data Companies' Revenues

Data companies' revenues *correspond to the aggregated value of all the data-related products and services generated by Europe-based data suppliers, including exports outside the EU. This indicator measures the revenues of the data suppliers identified and classified by Indicator 2 (see the products and services specified in our definition of the data market). Data companies' revenues do not include data monetisation as part of the data market.*

Forecasting Data Companies' Revenues

Data companies' revenues will continue to grow as the market evolves, as detailed in the table below. Overall growth among the Member States will be 3.4% out to 2030 (Baseline), although, from 2020 to 2025, growth will be higher, at a compound rate of 7.9%. The data market is healthy and robust but will slow a little as organisations consolidate their spending following early years' expense. The key industries that benefit from spending in the data market are – as always – information and communication, and professional services.

Data Companies Revenues Forecasts: 2025 (€M), Three 2030 Scenarios (€M), and Compound Growth (%)

	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR 2020–2025	CAGR 2025–2030, Challenge	CAGR 2025–2030, Baseline	CAGR 2025–2030, High Growth
EU27	104,086	108,964	123,294	152,372	7.9%	0.9%	3.4%	7.9%
EEA (NO, LI, IS) + CH	9,483	10,251	12,958	15,194	5.9%	1.6%	6.4%	9.9%
Total, all countries	140,015	152,625	174,987	213,405	7.3%	1.7%	4.6%	8.8%

Measuring the Data Market

*The **data market** is the marketplace where digital data is exchanged as "products" or "services" as a result of the elaboration of raw data.*

The market grew by 3.8% in 2020 to a value of €60.6 million and is expected to reach €63.6 million in 2021. Years 2020 and 2021 were difficult across Europe, but the data market remained healthy, benefitting from an increased focus on the digital economy for both work and services. Those companies and countries that were able to invest or utilise existing investments and expertise benefited from them.

Value and Growth (%) of the Data Market (€M), 2019–2021

N.	Market	Name	Description	2019	2020	2021	Growth Rate 2020–2021
4.1	EU27	Value of the data market	Estimate of the overall value of the data market	58,427	60,635	63,627	4.9%

Forecasting the Data Market

The data market for the EU27 will grow to €105.6 million in 2030 (Baseline), representing a compound growth of 3.2% over the 2025–2030 period. The below table shows the size and growth for the market for the three scenarios for 2030.

Data Market Forecast: 2025 (€ '000s), Three 2030 Scenarios (€ '000s), and Compound Growth (%)

	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR 2020–2025	CAGR 2025–2030, Challenge	CAGR 2025–2030, Baseline	CAGR 2025–2030, High Growth
EU27	90,121	94,218	105,619	125,238	8.2%	0.9%	3.2%	6.8%
EEA (NO, LI, IS) + CH	8,453	8,822	10,327	12,316	7.5%	0.9%	4.1%	7.8%
Total, all countries	125,221	130,640	145,335	171,087	8.2%	0.9%	3.0%	6.4%

Measuring the Data Economy

The **data economy** measures the overall impacts of the data market on the economy as a whole. It involves the generation, collection, storage, processing, distribution, analysis elaboration, delivery, and exploitation of data enabled by digital technologies.

The data economy captures a wider concept than the data market only, as it considers the value and wealth generated in the economy as a whole (not just across businesses) by the exploitation of data.

Data Monetisation

The definition of data supplier companies has been extended to accommodate the inclusion of the sale and purchase of data, **data monetisation**: Data monetisation is the revenue that data suppliers get from selling data. In the data economy, we consider data monetisation as an additional direct impact, generated at the level of supplier companies

The Data Economy, 2019–2021

The value of the data economy for the EU27 has been estimated to have reached almost €400 billion in 2019 and €440 billion in 2021, with a year-on-year growth rate of 4.9% in 2021. The estimated share of overall impacts on GDP in the EU27 rose from 3.1% in 2019 to 3.6% in 2021. A similar trend can be seen for the EU27 + the UK.

The UK's share of the data economy as a part of GDP is significant, making the overall share of data economy in the EU27 + the UK of GDP in 2021 of around 4%. The flow of data is vital for the EU economies. Indeed, with its large service sector, the UK has the largest internet economy as a proportion of GDP within the G20, reflecting the centrality of data to most goods and services trade. An interruption in data flows would therefore be costly. Despite Brexit uncertainties existing so far, the Commission has recently adopted two adequacy decisions[3] for the UK that allow for a free flow of personal data from the European Union to the United Kingdom, where data benefits from an essentially equivalent level of protection to that guaranteed under EU law. This is reflected in the data economy value for the UK, which remains strong throughout the period.

[3] Commission adopts adequacy decisions for the UK (europa.eu), June 2021

Data Economy, 2019–2021: Value (€M), Growth (%), and Impact on GDP (%)

N.	Market	Name	Description	2019	2020	2021	Growth Rate, 2020–2021	Impact on GDP, 2019	Impact on GDP, 2020	Impact on GDP, 2021
5.1 5.2	EU27	Value of the data economy and impacts on EU GDP	Value of total impacts on the EU economy and % of EU GDP	386,415	421,911	442,572	4.9%	3.1%	3.6%	3.6%
5.1 5.2	EU27 + UK	Value of the data economy and impacts on EU GDP	Value of total impacts on the EU economy and % of EU GDP	507,452	550,892	579,119	5.1%	3.4%	4.0%	4.0%
5.1 5.2	Total, all countries	Value of the data economy and impacts on EU GDP	Value of total impacts on the EU economy and % of EU GDP	553,690	599,624	630,410	5.1%	3.5%	4.1%	4.1%

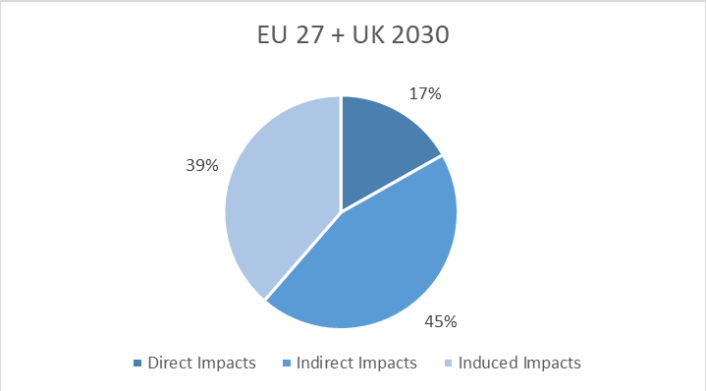
Forecasting the Data Economy, 2030

In 2030, the data economy for the EU27 is expected to reach the €1 trillion-threshold, with a 2025–2030 CAGR of 5.7%, slightly lower than the average increase of 7.1% in the period 2020–2025, as the data economy will face the highest increase in the post-COVID period as a result of the exceptional measures coming from the Next Generation EU[4] and from the urgency to adapt to the New Normal scenario.

- One of the main results is that, despite the slower growth, the share of the data economy of GDP in the EU27 will increase from 4.9% in 2025 to 5.9% in 2030.
- Another important result is the change in the composition of impacts: Indeed, from 2021 to 2030, the induced impacts share will increase from 34% to 39%, at the expense of indirect impacts, which will decrease from 50% to 45%.

[4] *NextGenerationEU is a **more than €800 billion temporary recovery instrument** to help repair the immediate economic and social damage brought about by the coronavirus pandemic. Post-COVID-19 Europe will be greener, more digital, more resilient, and better fit for the current and forthcoming challenges.*

Data Economy by Impact Type, EU27+ UK, 2030 (%)



Source: European Data Market Monitoring Tool, IDC 2021

The two alternatives show that, in the case of the Challenge scenario, the average growth for the EU27 (2020–2025 CAGR) will be half that expected in baseline scenario (3.1% in the Challenge scenario and 5.7% in the Baseline scenario), while growth will be nearly double according to the High Growth scenario (9.9% in the Challenge scenario and 5.7% in the Baseline one).

Data Economy Value (€M); 2030 Challenge, Baseline, and High Growth Scenarios (€M); and Impacts on GDP (%)

N.	Market	Name	Description	2030, Challenge Scenario	2030, Baseline Scenario	2030, High Growth Scenario	Impacts on GDP, 2025: Challenge Scenario	Impacts on GDP, 2025: Baseline Scenario	Impacts on GDP, 2025: High Growth Scenario
5.1 5.2	EU27	Value of the data economy and impacts on EU GDP	Value of total impacts on the EU economy and % of EU GDP	687,478	786,843	954,658	4.9%	5.3%	6.2%
5.1 5.2	EU27 + UK	Value of the data economy and impacts on EU GDP	Value of total impacts on the EU economy and % of EU GDP	904,567	1,028,013	1,248,004	5.5%	5.9%	6.9%
5.1 5.2	Total, all countries	Value of the data economy and impacts on EU GDP	Value of total impacts on the EU economy and % of EU GDP	976,091	1,115,486	1,353,445	5.5%	5.9%	7.0%

Source: European Data Market Monitoring Tool, IDC 2021

Measuring the Data Professionals Skills Gap

The **Data Professionals Skills Gap** indicator captures the potential gap between the demand and supply of data professionals in Europe.

The measurement of this indicator is based on a model combining the separate estimates and forecasts for the demand for data technical and business professionals and the supply of corresponding data skills by the inflow from the education system and the upskilling and reskilling of the existing workforce. This includes balancing the main sources of data skills (from the education

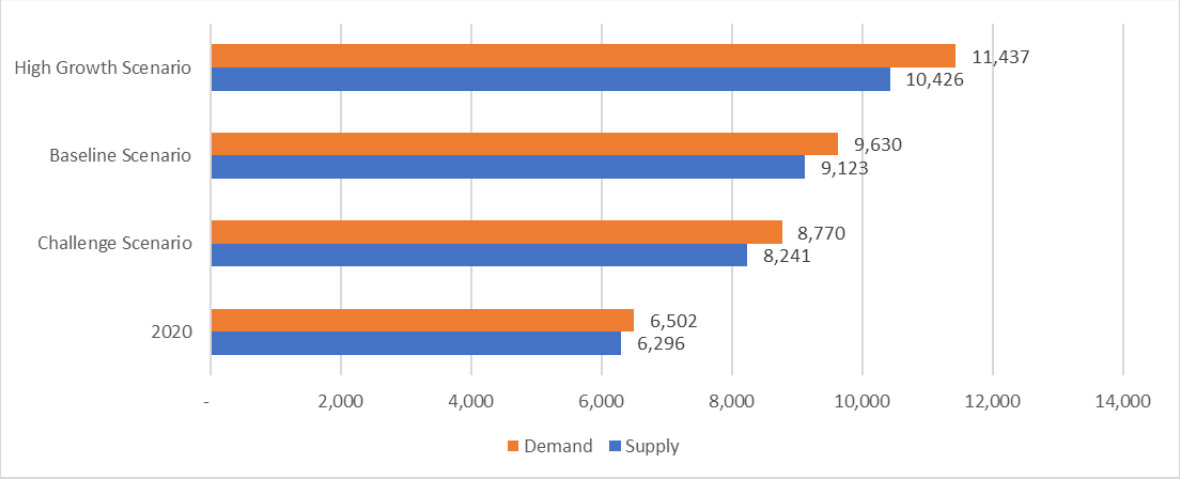
system and retraining to provision from other careers) with the estimated demand for data skills (by all data companies).

The skills gap between the demand and supply of data professionals measured as a percentage of the demand for data technical and business professionals is growing rapidly, and it will grow most significantly in the 2021–2025 period. Demand has already outgrown supply, and it is only beyond 2025 that the gap starts to decrease in any of the scenarios and countries modelled.

Regarding the skills gap for the EU27 in 2020 as a base year and in 2030 for the three scenarios, it is clear that, in all scenarios, a skills gap can be expected:

- In 2020, the data professionals skills gap is estimated at 206,000 across the EU27, growing to 507,000 in 2030 in the Baseline scenario. This means that the gap will grow from 3.2% in 2020 to 5.3% in 2030.
- For the Challenge scenario, the gap will reach 529,000 in 2030, or 6%, as graduates look for alternative careers and there is a lower number of entrants from other careers.
- In the High Growth scenario, the gap will reach 1,011,000 in 2030 as the education system, reskilling, and upskilling programs will be unable to keep up with the accelerated demand.

The Data Professionals Skills Gap for the EU27: 2020 and Three 2030 Scenarios ('000s)



Source: European Data Market Monitoring Tool, IDC 2021

Data Professionals Skills Demand and Gap for the EU by Member State: 2019–2021, 2025, and Three 2030 Scenarios ('000s)

						Baseline Scenario:		Challenge Scenario:		High-Growth Scenario:	
Member State		**2019**	**2020**	**2021**	**2025**	**2030**	**2025–2030 CAGR**	**2030**	**2025–2030 CAGR**	**2030**	**2025–2030 CAGR**
France	Numbers	18	33	46	61	74	4.1%	81	5.9%	160	21.3%
	% Gap	1.2%	3.4%	4.4%	5.2%	5.4%		6.2%		9.6%	
Germany	Numbers	42	68	92	116	137	3.4%	187	10.1%	259	17.5%
	% Gap	2.9%	4.3%	5.4%	5.5%	5.3%		7.7%		8.4%	
Italy	Numbers	12	22	37	40	45	2.3%	50	4.6%	105	21.4%
	% Gap	2.1%	3.3%	5.4%	5.0%	4.8%		6.0%		9.8%	
Poland	Numbers	14	18	22	37	38	0.5%	34	-1.4%	87	18.9%
	% Gap	2.8%	3.2%	3.8%	5.6%	5.1%		5.2%		9.8%	

Spain	Numbers	14	16	23	32	34	1.2%	33	0.7%	67	15.7%
	% Gap	3.2%	3.4%	4.6%	5.5%	5.0%		5.6%		8.3%	
EU22	Numbers	42	48	118	166	180	1.7%	143	–2.9%	334	15.1%
	% Gap	2.0%	2.2%	5.0%	5.9%	5.4%		4.8%		8.5%	
EU27	Numbers	135	206	338	451	507	2.4%	529	3.2%	1,011	17.5%
	% Gap	2.2%	3.2%	4.9%	5.5%	5.3%		6.0%		8.8%	

Source: European Data Market

The Data Economy Beyond the EU – the US, Brazil, Japan, and China

This report extends the Data Economy beyond the European Union and includes a specific section on four additional non-European countries – the US, Brazil, Japan and China. The indicators for the four selected EU international partners leverage IDC databases available at worldwide level. Data such as ICT spending is available for most countries worldwide and is gathered with the same approach across the board: These data series are perfectly comparable at an international level.

The US

The US is the strongest of the data economies when considering the internationals, which is unsurprising considering the foundation of data in business lies within the US. It is strong particularly in tools and software, but the size and strength of the economy supports a robust data economy. Many data suppliers are US based, particularly the leading proponents in the market.

The US is considered the world leader in areas such as artificial intelligence, and companies' investment in this technology is growing dramatically according to IDC's Artificial Intelligence Spending Guide. The US also has a strong focus on big data and analytics, and many, if not most, of the companies active in this technology are US based. The US is also strong in cloud technology, a foundational technology for data, with the leading cloud suppliers (AWS, Microsoft, and Google) being based in the US. As these organisations currently dictate the development of data tools and technologies, the US has a notable lead in the development of the market. However, China is emerging as a significant player in the global digital arena, particularly in artificial intelligence, 5G telecommunication networks, next generation internet, and high performance & quantum computing, to name a few key data-driven technology areas. This could pose a threat to the current US leadership in the data market and data economy fields. Future developments will depend a lot on the level of investment seen by the Chinese government, which is continuing to make significant progress in a number of data-driven technologies, which will be pivotal for the evolution of the data market and its effects on the economy as a whole.

US Indicators – 2019–2021 Overview

Name	Metrics	2019	2020	2021	Growth Rate: 2020–2021
Number of data professionals	Total number of data professionals ('000s)	14,350	15,275	16,169	5.9%
Data professionals' employment share	Percentage of data professionals on total employment	6.16%	6.44%	6.63%	3.0%
Number of data suppliers	Total number of data supplier companies ('000s)	312,215	315,857	321,847	1.9%
Revenues of data companies	Total revenues generated by companies specialised in the supply of data-related products and services (€M)	184,873	213,463	239,958	12.4%
Value of the data market	Estimate of the overall a value of the data market (€M)	184,873	213,463	239,958	12.4%
Value of the data economy (only direct and backward indirect impacts)	Direct impacts (€M)	184,873	213,463	239,958	12.4%
	Backward indirect impacts (€M)	123,480	163,296	232,101	42.1%
Impact of the data economy on GDP (only direct and backward indirect impacts)	Ratio between value of the data economy and GDP (%)	1.19%	1.26%	1.31%	4.3%

China

The size of the market and the population in China dominate any estimates of the size of the data market, and the high growth seen in the country's economy – with GDP growth remaining at around 6% compared with the EU growth rate of around 2% – ensures it will remain a strong market for the foreseeable future. Close to 28% of the labour force is in agriculture, compared with only 5% of the EU27 and 0.7% of the US labour force. These impacts estimate the share of the data force associated with data and data services. Only 43% of the Chinese labour force is employed in service industries, compared with 73% for the European Union.

China Indicators – 2019–2021 Overview

Name	Metrics	2019	2020	2021	Growth Rate: 2020–2021
Number of data professionals	Total number of data professionals ('000s)	8,717	9,184	9,815	6.9%
Data professionals' employment share	Percentage of data professionals on total employment	1.11%	1.19%	1.27%	6.7%
Number of data suppliers	Total number of data supplier companies ('000s)	756,002	858,509	952,566	11.0%
Revenues of data companies	Total revenues generated by companies specialised in the supply of data-related products and services (€M)	24,225	27,470	31,651	15.2%
Value of the data market	Estimate of the overall a value of the data market (€M)	24,225	27,470	31,651	15.2%
Value of the data economy (only direct and backward indirect impacts)	Direct Impacts (€M)	24,225	27,470	31,651	15.2%
	Backward indirect impacts (€M)	25,171	31,062	42,561	37.0%
Impact of the data economy on GDP (only direct and backward indirect impacts)	Ratio between value of the data economy and GDP (%)	0.76%	0.82%	0.84%	2.6%

Brazil

Brazil's economy: GDP growth has substantially declined in the second half of 2020, and inflation is on the rise. All in all, the country's infrastructure is not showing the growth required to support a vibrant data economy. While the revenues currently forecast growth for 2021, the investment in data professionals and their share of total employment are not as strong as in some of the other internationals, and the impact of the data economy on total GDP is declining slightly. Brazil's economy was hit badly by COVID-19 and declined by 4% in 2020.

Brazil Indicators – 2019–2021 Overview

Name	Metrics	2019	2020	2021	Growth Rate: 2020–2021
Number of data professionals	Total number of data professionals ('000s)	1,211	1,244	1,272	2.2%
Data professionals' employment share	Percentage of data professionals on total employment	7.56%	7.73%	7.79%	0.8%
Number of data suppliers	Total number of data supplier companies ('000s)	38,192	39,606	40,518	2.3%
Revenues of data companies	Total revenues generated by companies specialised in the supply of data-related products and services (€M)	7,905	8,374	8,865	5.9%
Value of the data market	Estimate of the overall a value of the data market (€M)	7,905	8,374	8,865	5.9%
Value of the data economy (only direct and backward indirect impacts)	Direct impacts (€M)	7,905	8,374	8,865	5.9%
	Backward indirect impacts (€M)	6,370	7,812	10,841	38.8%
Impact of the data economy on GDP (only direct and backward indirect impacts)	Ratio between value of the data economy and GDP (%)	0.23%	0.20%	0.20%	–2.5%

Japan

Japan represents the closest match to the EU data economy. Its data economy is reasonably robust, showing growth in the number of data professionals, although its growth in the share of total employment is limited when compared with others (excluding China), especially the EU27. Japan has a reasonably strong supply infrastructure, with the number of data suppliers about two-thirds that of the EU27, but forecasts show low growth for 2021. The infrastructure, in terms of data suppliers and data professionals, suggests that Japan will maintain its role as a key data-supplying nation in the world market.

Japan Indicators – 2019–2021 Overview

Name	Metrics	2019	2020	2021	Growth Rate: 2020–2021
Number of data professionals	Total number of data professionals ('000s)	4,236	4,398	4,567	3.8%
Data professionals' employment share	Percentage of data professionals on total employment	5.38%	5.51%	5.59%	1.5%
Number of data suppliers	Total number of data supplier companies ('000s)	106,983	106,214	106,786	0.5%
Revenues of data companies	Total revenues generated by companies specialised in the supply of data-related products and services (€M)	32,929	36,649	39,970	9.1%
Value of the data market	Estimate of the overall a value of the data market (€M)	32,929	36,649	39,970	9.1%
Value of the data economy (only direct and backward indirect impacts)	Direct impacts (€M)	32,929	36,649	39,970	9.1%
	Backward indirect impacts (€M)	26,985	30,960	40,145	29.7%
Impact of the data economy on GDP (only direct and backward indirect impacts)	Ratio between value of the data economy and GDP (%)	1.09%	1.23%	1.27%	3.9%

1. Introduction

Since 2013, the European Data Market (EDM) Monitoring Tool has monitored the evolution of the European data market and data economy, providing insights and quantitative evidence about its size and trends by industry and by region, contributing substantially to the development of European policy strategies in this field.

The European Commission contracted two European data market studies (SMART 2013/0063 and SMART 2016/0063) – in 2013 and 2016, respectively – providing facts and figures on the size and trends of the EU data economy through the development and continuation of a European data market monitoring tool. The third edition of the EDM leverages the work done in these past years and builds upon the proven conceptual framework that has been recognised by a broad group of stakeholders. In order to guarantee continuity and comparability with the previous figures produced, this study provides up-to-date indicators with forecast data for 2025 and under three alternative scenarios (Baseline, High Growth, and Challenge) for the year 2030.

The third edition of the European Data Market Study (VIGIE 2020-0655) intends to continue to deliver credible and thorough factual evidence on the growth of the EU data economy and market. Especially in light of the COVID-19 pandemic and the implications for the economy and for policy priorities at national and European level, the facts and figures on the current state of play and outlook of the data market and the data economy are essential in support of the EU Data Strategy and its related initiatives in terms of data governance and for the creation of a single market for data.

The First Report on Facts and Figures (D2.1) delivers the first round of measurement of the updated EDM Monitoring Tool, with the time horizon extended to 2030. With a broadened scope and several essential updates and innovative elements, this report focuses on the following six indicators: 1) the number of data professionals (including the number of data scientists); 2) the number of data companies and data user companies (including the number of start-ups and SMEs); 3) the revenue generated by data companies (including start-ups and SMEs); 4) the size of the data market; 5) the value of the data economy; and 6) the data skills gap (including the data scientists' skills gap).

1.1 Objectives

The main goal of this study is to update and develop the EDM Monitoring Tool, continuing to provide the market intelligence, factual evidence, and insights on the future developments needed to support European policies in this field in the next three years. A balance between continuity (to build on the value of the data and understanding accumulated in the past seven years) and innovation (to monitor and measure new developments and emerging trends as the Data Economy matures) is a key objective of the methodology presented in this report.

As for the previous editions, the study will be conducted through three correlated streams of work aimed at collecting and delivering factual evidence, data, and qualitative insights, which together will feed into the overarching goal to provide insights on progress towards policy targets.

1. Updating and implementing the EDM Monitoring Tool indicators, revising the taxonomy and methodology approach developed in the last years to adapt to the market evolution, measuring annually the main EDM indicators, providing facts and figures on all the key features of the European data market and economy, with a strong focus on the measurement of growth and job creation.

2. Producing descriptive stories analysing the main critical issues of the development of the data economy and society, complementing the EDM statistical indicators with qualitative and quantitative evidence based on case studies and expert analysis. The stories have proven a flexible and valuable tool to investigate issues and questions arising from evolving data policy priorities during the life of the project.
3. Further developing the data stakeholders' landscape and community built in the past years, improving the mapping and visualisation activities published online. In this edition of the EDM study, the EU data landscape will be considerably improved thanks to collaboration with Dealroom, a leading provider of data on tech companies and partners of the EuropeanStartups.co EU initiative. We also plan to leverage the data stakeholders community to collect information and evidence for the ongoing measurement of EDM indicators.

1.2 Overview of Indicators

This document focuses on the following indicator sets:

- Indicator 1.1: the number of data professionals
- Indicator 1.2: the employment share of data professionals
- Indicator 1.3: the intensity share of data professionals
- Indicator 2.1: the number of data supplier companies
- Indicator 2.2: the share of data supplier companies
- Indicator 2.3: the number of data user companies
- Indicator 2.4: the share of data user companies
- Indicator 2.5: the share of data user and data supplier companies that offer data for re-use
- Indicator 3.1: the total revenues of data supplier companies
- Indicator 3.2: the share of data supplier companies' revenues
- Indicator 4: the value of the data market
- Indicator 4.1: data monetisation
- Indicator 5: the value of the data economy
- Indicator 5.1: the impact of the data economy on GDP
- Indicator 6: the data professionals skills gaps

What is new?

In comparison with the previous editions of the EDM study (2013 and 2016), the listed indicators have been updated and revised, and their scope has been broadened.

Revised taxonomy:

- The definition and quantitative model for calculating the number of **data professionals** and the **data skills gap** has been substantially revised and improved – also thanks to new models developed by IDC in recent years on estimating the demand of ICT professionals, including data professionals, worldwide.

- The definition of the **data supplier companies** has been updated and broadened following the evolution of the European data industry.

Methodology approach:

- The **data market** is also extended to **include data monetisation**, the sale and purchase of data. In previous phases, this was excluded due to the tiny size of the market at the inception of the project and difficulty in obtaining meaningful data relating to the data marketplace. However, this is now a significant revenue stream in several industries, so this will be included as a component of the data market.
- **Agriculture** has been added to the industry segmentation for a total of 12 sectors modelled.

Wider geographic scope:

- **New measurements** have been added beyond the EU and its Member States, with Switzerland and the EEA countries of Iceland, Liechtenstein, and Norway. The United Kingdom will continue to be assessed separately.
- Because of China's established role in shaping socioeconomic and technology trends, the **international indicators** coverage has been expanded from the United States, Brazil, and Japan to the People's Republic of China.

The table below offers an overview of the full set of indicators that have been developed in this First Report on Facts and Figures.

Table 1 D2.1: First Report on Facts and Figures – Full Set of Indicators

#	Name of Indicator	Year	Industry	Member State	Company Size	EU27	UK+ Switzerland	EEA (NO, LI, IS)
1.1	Number of data professionals	2019	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2020	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2021 estimated data	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2025 Baseline	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2030 Baseline	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2030 High Growth	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2030 Challenge	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
1.2	Employment share	2019	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable

#	Name of Indicator	Year	Industry	Member State	Company Size	EU27	UK+ Switzerland	EEA (NO, LI, IS)
		2020	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2021 estimated data	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2025 Baseline	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
		2030 Baseline	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
		2030 High Growth	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
		2030 Challenge	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
1.3	**Intensity share**	2019	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2020	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2021 estimated data	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2025 Baseline	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
		2030 Baseline	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
		2030 High Growth	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
		2030 Challenge	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
2.1	**Number of data supplier companies**	2019	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2020	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2021 estimated data	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2025 Baseline	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2030 Baseline	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2030 High Growth	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2030 Challenge	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable

#	Name of Indicator	Year	Industry	Member State	Company Size	EU27	UK+ Switzerland	EEA (NO, LI, IS)
2.2	**Share of data supplier companies**	2019	Applicable	Applicable	Not Applicable	Applicable	Applicable	Applicable
		2020	Applicable	Applicable	Not Applicable	Applicable	Applicable	Applicable
		2021 estimated data	Applicable	Applicable	Not Applicable	Applicable	Applicable	Applicable
		2025 Baseline	Applicable	Applicable	Not Applicable	Applicable	Applicable	Applicable
		2030 Baseline	Applicable	Applicable	Not Applicable	Applicable	Applicable	Applicable
		2030 High Growth	Applicable	Applicable	Not Applicable	Applicable	Applicable	Applicable
		2030 Challenge	Applicable	Applicable	Not Applicable	Applicable	Applicable	Applicable
2.3	**Number of data user companies**	2019	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2020	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2021 estimated data	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2025 Baseline	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2030 Baseline	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2030 High Growth	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2030 Challenge	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
2.4	**Share of data user companies**	2019	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2020	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2021 estimated data	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2025 Baseline	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2030 Baseline	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2030 High Growth	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2030 Challenge	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
2.5	**Share of data user and data supplier companies that**	2019	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
		2020	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2021	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable

#	Name of Indicator	Year	Industry	Member State	Company Size	EU27	UK+ Switzerland	EEA (NO, LI, IS)
	offer data for re-use	estimated data						
		2025 Baseline	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
		2030 Baseline	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
		2030 High Growth	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
		2030 Challenge	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
3.1	**Total revenues of data supplier companies**	2019	Not applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2020	Not applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2021 estimated data	Not applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2025 Baseline	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
		2030 Baseline	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
		2030 High Growth	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
		2030 Challenge	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
3.2	**Share of data supplier companies' revenues**	2019	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2020	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2021 estimated data	Applicable	Applicable	Applicable	Applicable	Applicable	Applicable
		2025 Baseline	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
		2030 Baseline	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
		2030 High Growth	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
		2030 Challenge	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
4	**Value of the data market**	2019	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2020	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable

#	Name of Indicator	Year	Industry	Member State	Company Size	EU27	UK+ Switzerland	EEA (NO, LI, IS)
		2021 estimated data	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2025 Baseline	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2030 Baseline	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2030 High Growth	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2030 Challenge	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
4.1	**Data monetisation**	2019	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2020	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2021 estimated data	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2025 Baseline	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2030 Baseline	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2030 High Growth	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2030 Challenge	Applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
5	**Value of the data economy**	2019	Not applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2020	Not applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2021 estimated data	Not applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2025 Baseline	Not applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2030 Baseline	Not applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2030 High Growth	Not applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2030 Challenge	Not applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
5.1		2019	Not applicable	Applicable	Not applicable	Applicable	Applicable	Applicable

#	Name of Indicator	Year	Industry	Member State	Company Size	EU27	UK+ Switzerland	EEA (NO, LI, IS)
	Impact of the data economy on GDP	2020	Not applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2021 estimated data	Not applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2025 Baseline	Not applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2030 Baseline	Not applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2030 High Growth	Not applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
		2030 Challenge	Not applicable	Applicable	Not applicable	Applicable	Applicable	Applicable
6	**Data professionals skills gap**	2019	Not applicable	Applicable (only for DE, ES, FR, IT PL)	Not applicable	Applicable	Applicable	Applicable
		2020	Not applicable	Applicable (only for DE, ES, FR, IT PL)	Not applicable	Applicable	Applicable	Applicable
		2021 estimated data	Not applicable	Applicable (only for DE, ES, FR, IT PL)	Not applicable	Applicable	Applicable	Applicable
		2025 Baseline	Not applicable	Applicable (only for DE, ES, FR, IT PL)	Not applicable	Applicable	Applicable	Applicable
		2030 Baseline	Not applicable	Applicable (only for DE, ES, FR, IT PL)	Not applicable	Applicable	Applicable	Applicable
		2030 High Growth	Not applicable	Applicable (only for DE, ES, FR, IT PL)	Not applicable	Applicable	Applicable	Applicable
		2030 Challenge	Not applicable	Applicable (only for DE, ES, FR, IT PL)	Not applicable	Applicable	Applicable	Applicable

1.3 Structure of the Report

This ***First Report on Facts & Figures*** focuses on the results obtained through the first round of measurements of the European Data Market Monitoring Tool for the 2019–2021 period, with forecasts for 2025 and for 2030 under three distinct scenarios.

The report is organised in the following chapters:

- **Chapters 1 and 2** include a brief introduction and a short reminder of the overall study's goals and objectives, as well as a summary of the European Data Market Monitoring Tool and its functioning.
- **Chapter 3** is devoted to the measurement of the data professionals, including their main values in absolute terms, their share in terms of total employment, and their forecast to 2025 and 2030 according to the three scenarios under consideration.
- **Chapters 4 and 5** provide the values for the indicators measuring the data companies (both suppliers and users of data) in terms of absolute numbers and produced revenues. They include an updated forecast of the indicators to 2025 and 2030.
- **Chapter 6** presents the indicators measuring the size of the data market in Europe based on the value generated by pure data players developing BDA technologies and the value created by data-related research, businesses, information, and IT services and its contribution to the data market. In this chapter, as a new feature, an estimate for the value of data monetisation is added to the data market.
- **Chapter 7** focuses on the data economy and measures the overall impacts of the data market on the economy as a whole, including three sets of impacts – data companies' revenues, indirect impacts, and induced impacts.
- **Chapter 8** is devoted to the update of the data professionals skills gap indicator in the EU.
- **Chapter 9,** International Dimension of the Data Economy, presents a select number of indicators for the United States, China, Brazil, and Japan.
- **Chapter 10** provides a set of concluding remarks of the report.
- The **Methodological Annex** summarises the key methodological steps that we have undertaken to measure the indicators covered in both the previous reports and in the current report.

2. European Data Market Monitoring Tool: 2021–2023

2.1 A Novel Methodological Approach

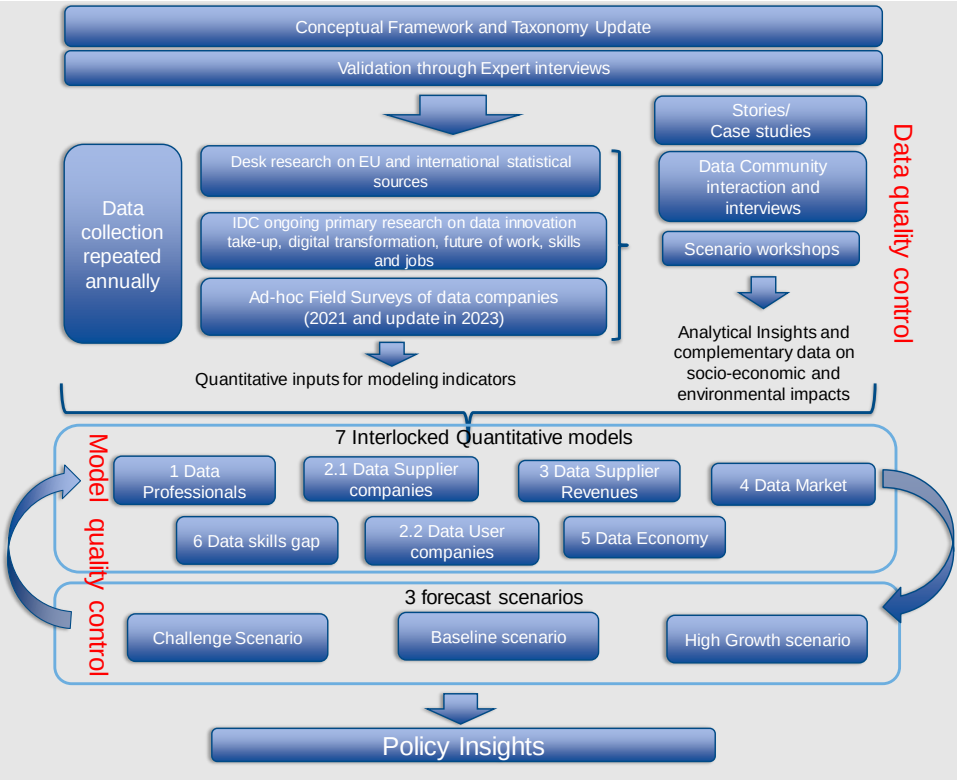
The EDM study requires a complex mix of qualitative and quantitative methodologies, allowing to reach the objectives presented in the previous chapter. While quantitative methodologies represent the most relevant part of the study, qualitative methodologies are indispensable to balance the statistical approach and provide the market and social intelligence needed to lead to policy insights and the development of sound scenarios.

In order to ensure the study's continuity, the methodological approach builds upon previous EDM Monitoring Tool releases, including necessary improvements. The definition of the main indicators is revised and updated according to the maturing of the market and data economy. It is important to emphasise that the EDM Monitoring Tool's modular structure does not change in its core substance, as it has already proven to be sufficiently adaptable to the evolution of the data market and data economy in past years.

The complexity of the methodologies used for providing the facts and figures on the European data market and European data economy is presented on the figure below, with the main methodological steps being as follows:

- The revision and update of the conceptual framework and taxonomy consist of the following elements: a) an update of definitions of data professionals, data user companies, and data market, including data monetisation value; b) the definition and management of data sharing and data interoperability issues, including the role of common European data spaces and the concept of data sovereignty; c) the assessment of the social and environmental impacts of data-driven innovation.
- The validation of the revision/update through expert interviews.
- The organisation and implementation of data collection (including desk research and field research), which will be repeated annually to feed into the measurement of indicators.
- In parallel, activities in case studies and stories and the polling of the stakeholder community will provide complementary data and analytical insights, particularly concerning socioeconomic and environmental impacts.
- The outputs of data collection and qualitative analysis will feed into the seven interlocked quantitative models used to measure the main indicators (the data companies indicator has two models, one for user companies and one for supplier companies).
- In parallel with the calculation of indicators, the scenario forecasting methodology will be implemented, developing the main assumptions driving the three alternative scenarios to 2030 and the forecasting of all indicators.
- The quantitative models and the scenarios methodology interact closely and provide reciprocal feedback.
- Quality control will accompany each step of the process, with a focus on data quality control in the data collection phase and on model quality control in the phase of the measurement of indicators and the development of scenarios.

Figure 1. A Complex Methodology



The main methodological changes introduced by this third edition of the European Data Market Monitoring Tool revolve primarily around four indicators.

1. **Measuring the Value of Data**

Measuring the value of data, and/or measuring the value of the data sharing market, is becoming increasingly relevant to assess its contribution to economic growth and assess the benefits and returns on investments in the data economy. However, measuring the value of data is also a relevant challenge because of the nature of data and the complexity of its role in the economic system.

To measure the value of data, we have improved the methodology by:

- Adding a sub-indicator on data monetisation to the data market value indicator, which can be measured by industry/country and company size, as with all of the other indicators
- Deepening the analysis and estimates of the indirect benefits of data sharing and data reuse within the data economy indicator
- Investigating other aspects of the value of data through stories and interaction with the data community

2. **Measuring Data Supplier and Data User Companies**

The approach to measuring data supplier and data user companies has been improved based on market developments. The definition of data companies (which is divided into data suppliers and data users) will remain unchanged.

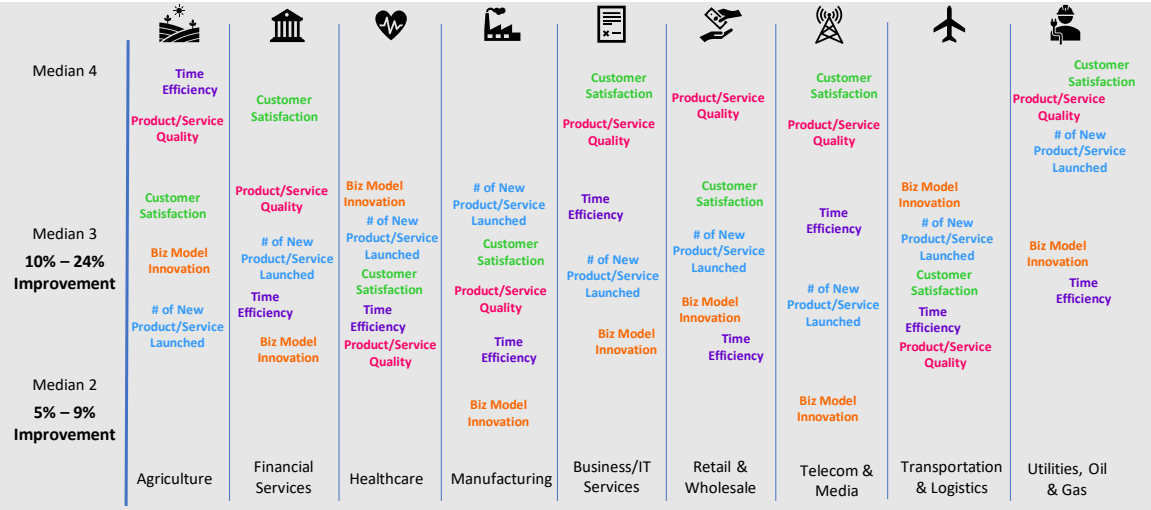
Data-supplier companies: Data innovation has been rapidly spreading from technology-driven sectors to a larger part of the economy and is now diffused in many industries. As a result, we believe that a congruous number of specialised data companies have emerged in other sectors (such as finance and retail), and these companies can now be measured. In order to gauge this evolution of the data industry, in the quantitative survey, we have investigated both data suppliers and data users in all sectors of the economy and identified them based on questions about their activities in the data market. Desk research and IDC data have complemented the survey results.

Data-user companies: In the past editions of the study, we identified data users based on of their use of big data technologies for decision-making. We are able now to use a more sophisticated method by investigating in the survey the main business goals and the range of ways in which companies use data based on data-driven use cases. Use cases are specific instances of the implementation of technologies to achieve business goals, and IDC has accumulated a rich library of use cases for all the main technologies and specifically for digital transformation. In this report, we identify data users based on their level of adoption of use cases and their main business goals.

3. Measuring the Business Impacts

To improve the estimate of the business impacts of data-driven innovation included in the data economy (indirect impacts), we extended the methodology developed in the Horizon 2020 DataBench[5] project coordinated by IDC to the survey for this first measurement of the European Data Market Monitoring Tool. Within the conceptual framework of the Databench project, we developed seven business KPIs – i.e. seven business factor categories (Figure 2 below), selected based on business literature and IDC research of technology vendors and users as the most relevant for measuring the impacts of innovative technology investments on business performance. We simplified and operationalised the business KPI definitions to allow measurement through business surveys. This approach is one of several options for the measurement of technology business impacts and was chosen in order to measure impacts that are valid for European industry and differentiated by sector and company size.

Figure 2. Data-Driven Use Cases: Business Impact Benchmarks Based on KPIs



Source: IDC for DataBench project, 2019

[5] DataBench, Evidence Based Big Data Benchmarking to Improve Business Performance: "Industry Requirements with benchmark metrics and KPIs", D.1.1, https://www.databench.eu/wp-content/uploads/2020/02/databench-d1.1-ver.2.0.pdf

4. Measuring Data Professionals

The definition of data professionals has been established in the previous editions of the EDM Monitoring Tool. As digital transformation and data-driven innovation penetrate more deeply in European organisations, we have seen an increase and differentiation of the population of data professionals with different mixes of skills. The definition of data professionals has been updated and articulated into three main types: data ***technical*** professionals, data ***business*** professionals, and data ***consumers***.

This new definition has been validated with expert interviews with its measurement approach relying on estimating the share of professionals falling under the three main categories identified from professional ISCO codes. The measurement of data professionals is one of the most complex and difficult models of this study because it builds on assumptions about the connection between skills, job roles, and statistical codes.

2.2 Field Research: Ad-Hoc Survey

The first round of measurement of the European Data Market Monitoring Tool of this third edition of the project was supported by quantitative field research. The survey was conducted in July–September 2021 among a representative sample of 1,191 data user and data supplier companies to collect data about the uptake of data technologies and data-related products and services across Europe to provide up-to-date and comprehensive base results for the models underpinning the Monitoring Tool.

The survey investigated the benefits of the adoption of data technologies, their way to sustain and enable the development of the data market and the data economy, and the issues and intentions around the availability and development of skills to develop and use analytical systems. It questioned users on their adoption of advanced analytics in a series of recently defined use cases. This analysis will further shed light on the use of data for strategic decision-making processes. The analysis of data-driven use cases and of the business impacts of big data and analytics was already developed and adopted by the IDC DataBench[6] project and further refined in the Advanced Technologies for Industry study, which IDC conducted for DG GROW[7]. A use case is a conceptual framework that provides a view of business value that is created when a set of technologies come together. Use cases are not defined by the technology itself; rather, the parameters of a use case are defined by the value being created and recognised by an organisation. In the DataBench conceptual framework, we define a use case as:

> *a discretely funded effort designed to accomplish a particular business goal or objective through the application of big data technology to particular business processes and/or application domains, employing line-of-business and IT resources.*

IDC classifies use cases as horizontal (applicable to all sectors) or vertical (applicable to one sector or a small group of sectors).

Finally, and importantly, the survey investigated business impacts seen by respondents on organisation performance and time and cost reduction, correlated with efficiency improvements resulting from the use of big data, analytics, and AI.

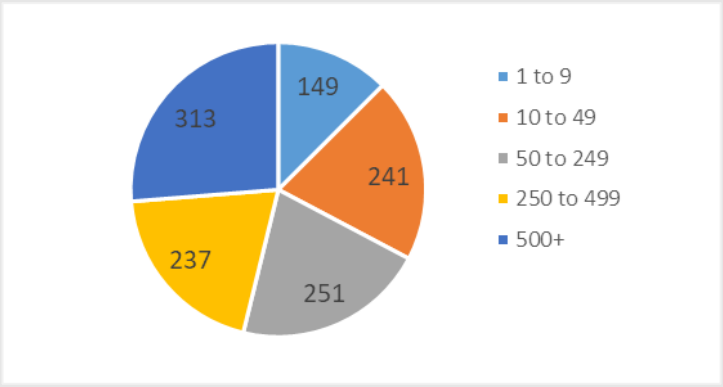
[6] DataBench, Evidence Based Big Data Benchmarking to Improve Business Performance: "Industry Requirements with benchmark metrics and KPIs", D.1.1, https://www.databench.eu/wp-content/uploads/2020/02/databench-d1.1-ver.2.0.pdf

[7] https://ati.ec.europa.eu/

- **Survey Methodology and Respondent Demographics**

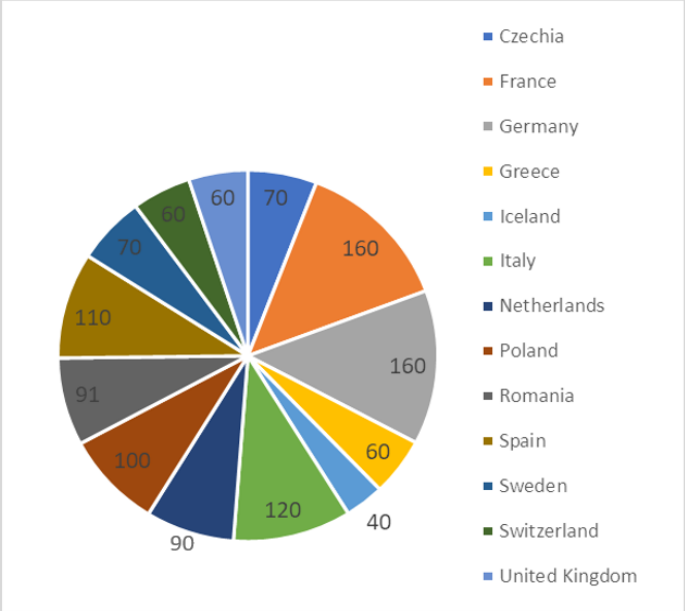
The survey comprised 1,191 respondents, who were interviewed using the computer-aided telephony (verbal) technique in the respondent's native language. The respondents are broken down as shown in the charts below. Target respondents were decision makers and influencers for the use of data and for data management in their respective organisations. The companies were filtered (screened) to ensure that they are using data at some level to make business decisions.

Figure 3. Surveyed Companies, Company Size (Employees)



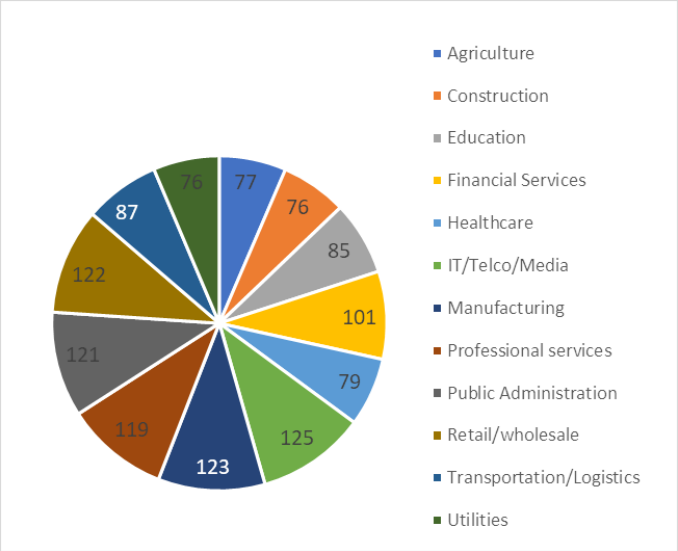
Source: European Data Market Study Survey, 2021

Figure 4. Surveyed Companies



Source: European Data Market Study Survey, 2021

Figure 5. Surveyed Companies, Industry Segments



Source: European Data Market Study Survey, 2021

- **Confidence Levels**

Each survey data point is tested for confidence levels using either a chi-square or T-test, as appropriate. A chi-square test is run to determine whether there is a significant difference between the two nominal (categorical) variables. A low chi-square value (below 0.05) means there is a high association. A pairwise T-test comparison is run to determine whether there is a significant difference between the means of two data sets.

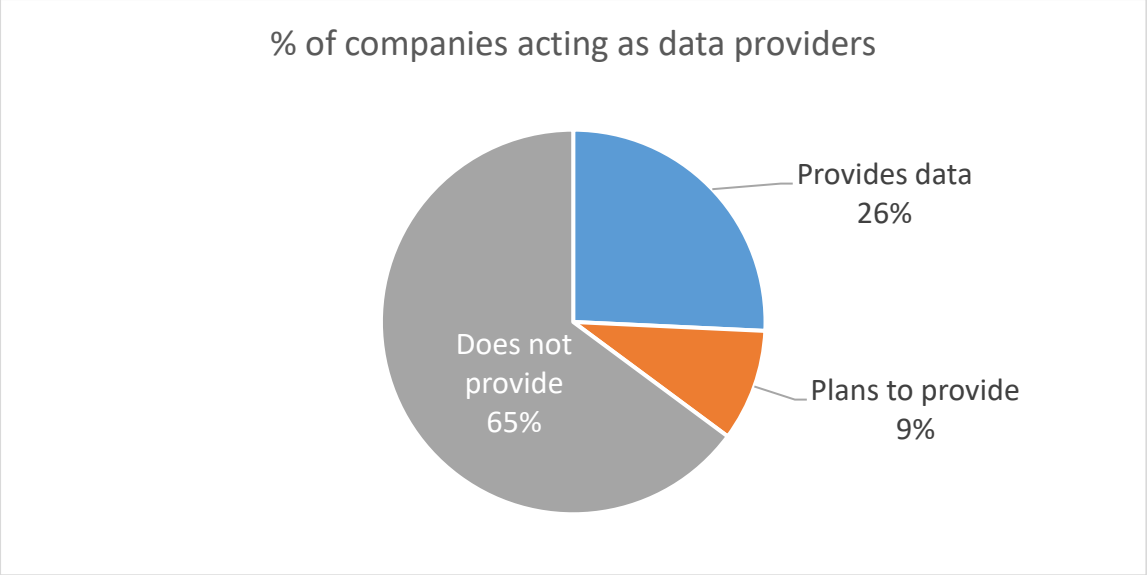
Survey Highlights

Analysis of the IDC European data market survey results shows important evidence regarding analytics and big data usage and their effects and benefits across data companies in Europe. Data usage and its applications are generally more widespread in the largest companies: the bigger the size, the more likely the adoption of any of the inquired technologies, tools, or use-cases. Regarding the industries, the same pattern applies to some of the most analytics-leaning industries (finance, manufacturing, IT/telecom/media) compared with some less-digitalised sectors – namely, agriculture, construction, and logistic.

These constant patterns also unveil interesting evidence on the external usage of data and on data supply.

- *External Data Usage and Data Supply*
  All of the interviewed companies use some external data for business decisions, but only a portion of them are also data providers to other companies.

Figure 6. Data Providers



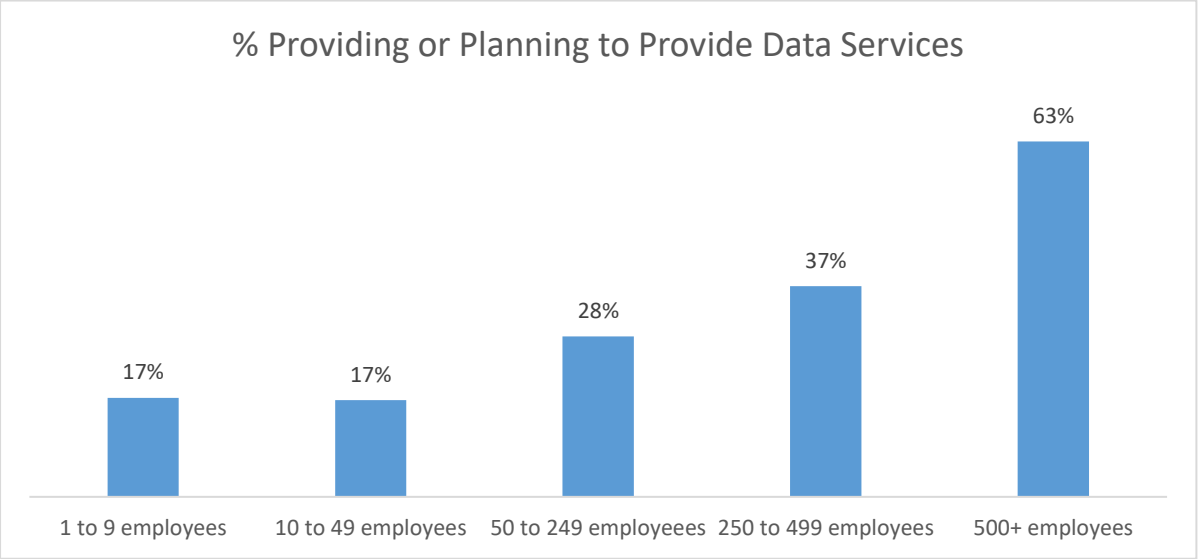
Source: European Data Market Study Survey, 2021

A surprisingly large portion of the interviewed companies – 26% – are active providers of data or analytics products, services, tools, or technologies to other organisations, while a further 9% plan to become providers in the future. However, differences across industries are wide: While 80% of IT/telecom/media organisations are data providers or planning to be, this percentage shrinks to less than 15% for agriculture, construction, public administration, and logistics.

The most widespread data service types provided are data-based products and services to end users in specific vertical markets (47%) and access to premium datasets/sources (32%). Around a quarter of the providers furnish marketing/advertising services data (28%) and software/consulting for big data tools (24%).

Also, there is dramatic variation depending on company size, as shown in the below chart.

Figure 7. Percentage Providing or Planning to Provide Data Services



Source: European Data Market Study Survey, 2021

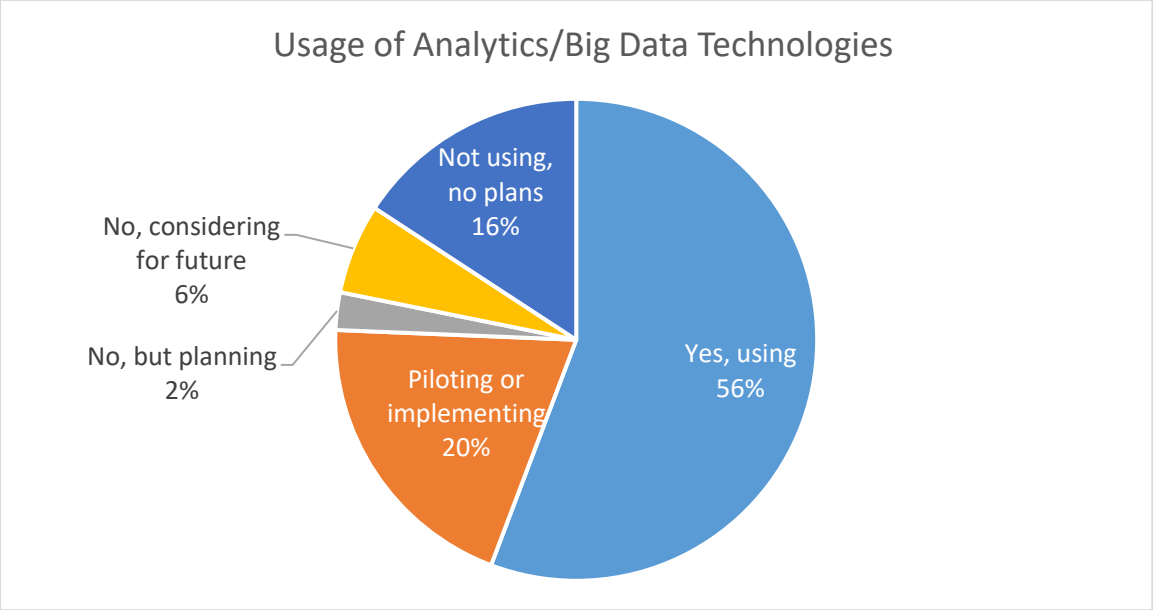
A high quota of those interviewed – 83% – declare to be users or providers of open data.

Finally, more than 50% of the respondents participate in ***data marketplaces***, with 15% more planning to participate in the near future. B2B platforms and industrial data platforms are much less used (26% and 14%, respectively). The most data-heavy industries are, again, those with the highest shares.

- *Analytics and Big Data Usage*

The usage of analytics and big data technologies in enterprises in Europe is undoubtedly on the rise. While 16% of the respondents are neither using nor planning to use BDA technologies in their workplaces, 56% are already actively using them. The most interesting number comes from companies in a transition phase: 30% of those interviewed are in one of these phases, progressing from no use of BDA to a pilot/new use of BDA. Once again, the more traditional industries and smaller companies show lower usage. Nevertheless, a high number of smaller companies are currently in a transition phase: almost 50% of companies with fewer than 50 employees are either planning to use BDA technologies or have just started piloting them.

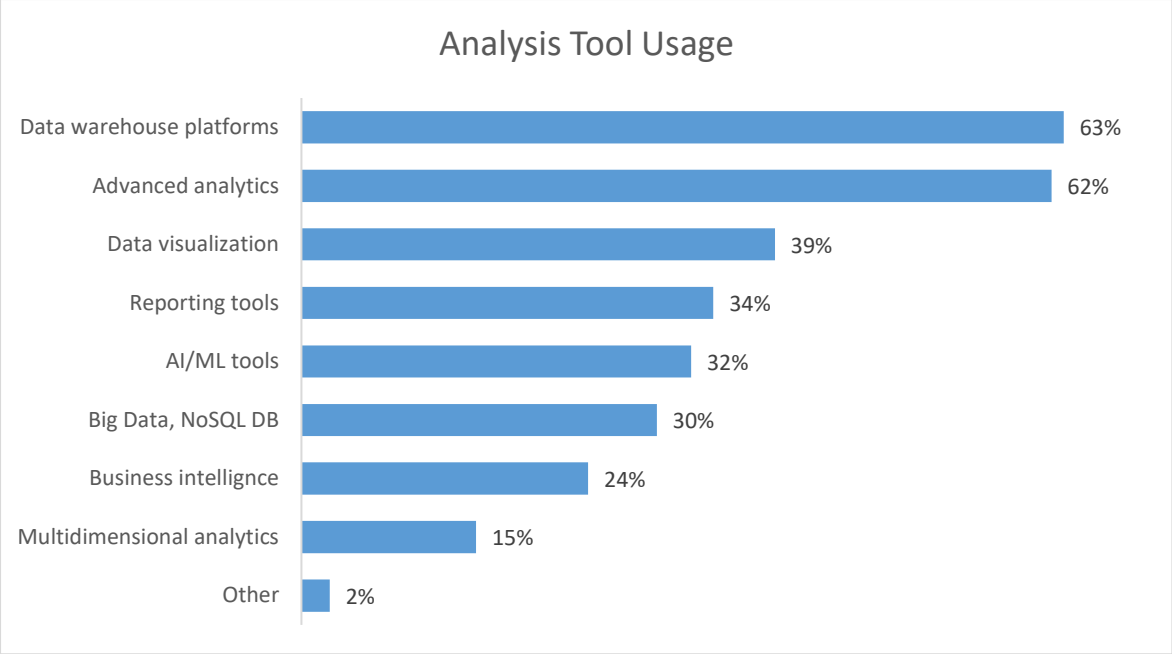
Figure 8. Usage of Big Data/Analytics Technologies



Source: European Data Market Study Survey, 2021

Regarding which technologies are used for data analysis, data warehouses platforms and advanced analytics are used by almost two-thirds of enterprises. Data visualisation, reporting, and artificial intelligence/machine learning tools are also widespread, with more than 30% of the respondents declaring their usage.

Figure 9. Analysis Tool Usage



Source: European Data Market Study Survey, 2021

- *Analytics and Big Data Usage – Use Cases*

When it comes to the most relevant use cases related to big data and analytics technologies, there is a rich vein of information for different industries across Europe. Digging deeper into some use cases, we can divide them into generic ones (applicable to most industries) and industry-specific ones.

The most widespread use cases are "cyber-security intelligence, threat detection, and prevention" and "customer experience management", both showing around 76% of the companies declaring current usage or plans for future usage of them. Moreover, some specific use cases regarding the finance industry are very pervasive: More than 90% of respondents declare the current or future use of analytics for fraud detection and digital banking experience. On the opposite side, other AI-related use cases, while rapidly emerging, are still at a nascent stage: More than 70% of respondents, for example, declare they are neither using nor currently planning to use drone inspection, digital simulation and digital twins, or intelligent-building energy management.

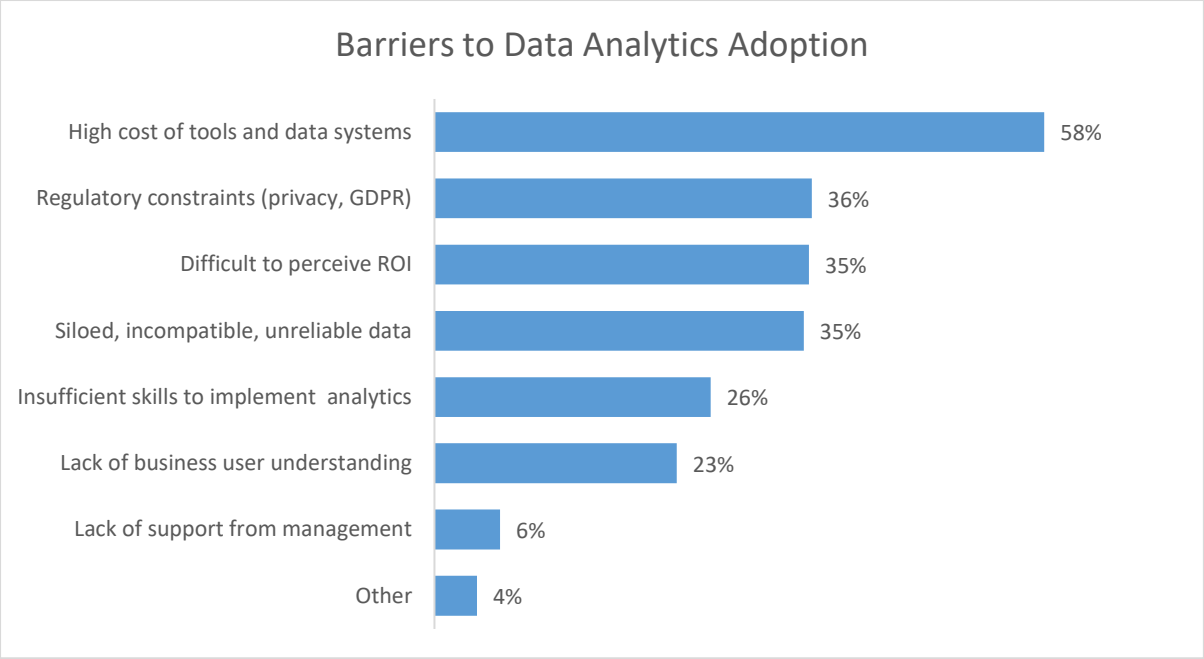
- *Barriers and Obstacles*

One of the most important topics covered is how barriers and obstacles can hinder the process of developing effective analytics and data-oriented decisions, providing obvious pathways for intervention to improve uptake. These barriers can be obstacles regarding the everyday use of the platform, or they can be structural or cultural.

In the survey, three different barriers were very close in incidence, highlighting an evenly distributed pattern. Regulatory constraints, difficulties in perceiving ROI deriving from the use of analytics, and problems with unreliable, inefficient, or siloed data – each of these was selected by more than a third of the enterprises surveyed. Other IDC research has shown that data siloes are a widespread problem in implementing efficient analytics throughout Europe.

Lack of skills and lack of understanding from the business users are also major barriers for the respondent companies. The data culture problem – a lack of support from management – is not a widespread issue according to the survey. Nevertheless, it must be stated that a significant proportion of responses are from management.

Figure 10. Barriers to Data Analytics Adoption



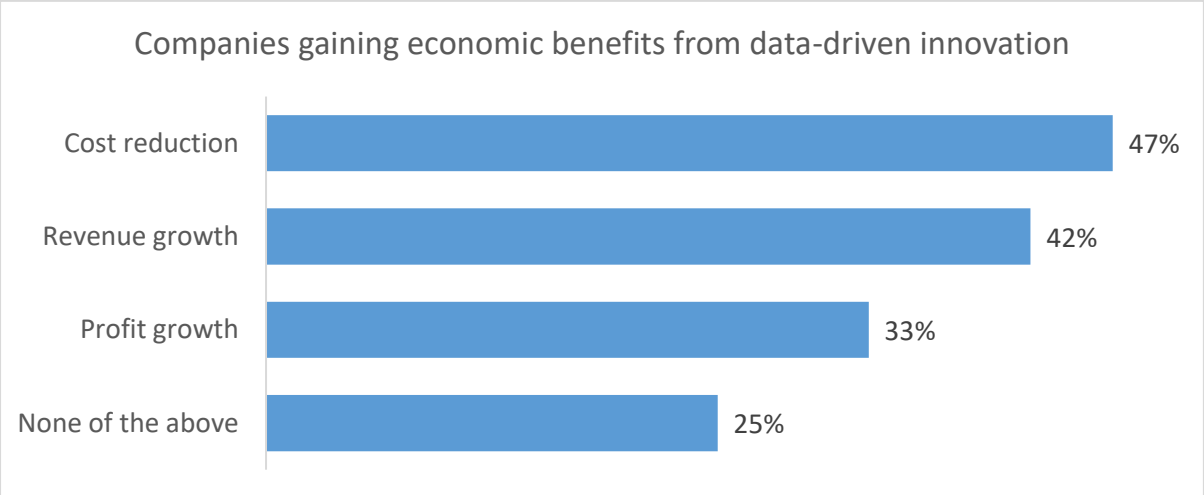
Source: European Data Market Study Survey, 2021

- *Economic and Non-Economic Benefits*

When considering benefits gained from the use of Analytics, it is crucial to underline that only a small proportion of the interviewed enterprises see data-driven innovation as having no value or cannot identify clear benefits from it.

Almost half identified cost reductions, over 40% identified revenue growth, and one-third identified profit growth from the use of analytics.

Figure 11. Achieving Economic Benefit from Data-Driven Innovation

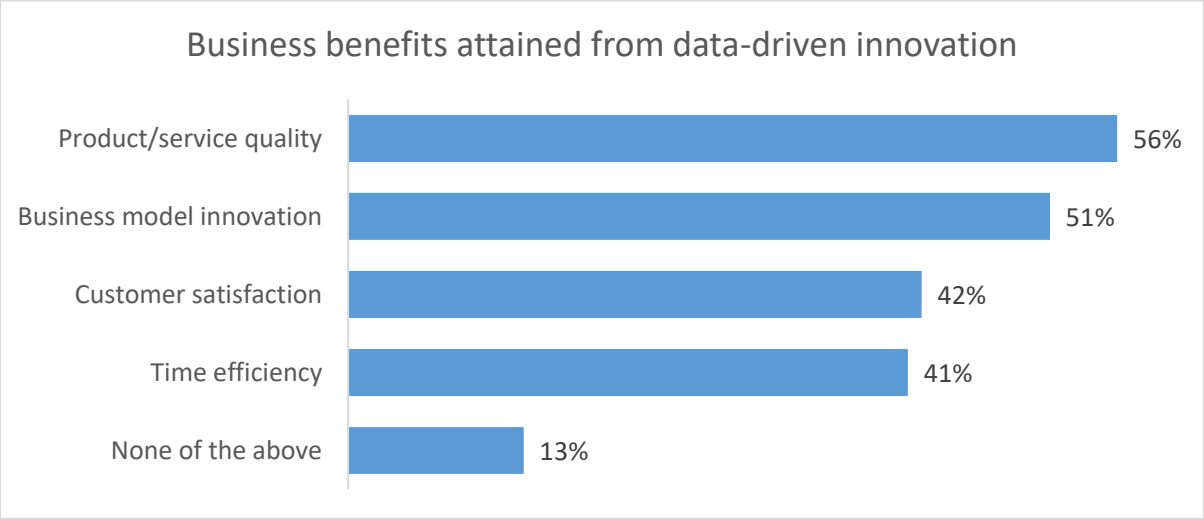


Source: European Data Market Study Survey, 2021

Furthermore, an even fewer share of the interviewees sees no business benefits – namely, less than 13%.

At the same time, other (non-financial) data-driven improvements, such as in product/service quality or business model innovation, are significant. Customer satisfaction and time efficiency enhancement are widespread: More than 40% see data as improving these fundamental company outcomes.

Figure 12. Achieving Business Benefit from Data-Driven Innovation



Source: European Data Market Study Survey, 2021

- *Data Professionals*

Lately, there has been a surge in discussions regarding the lack of skilled people in data professional roles. Data technical professionals are data engineers, data analysts, and data administrators, while data business professionals are data scientists and business data analysts.

Their numbers are highly variable in the survey sample: It very much depends on the number of employees and the industries of the respondent, varying from zero to thousands. Nevertheless, it can be estimated broadly that data professionals are about 10–20% of the workforce in small companies and 5–10% of the workforce in bigger companies. Far from hinting at a lesser need for data professionals in large companies, this percentage is likely to be the result of a statistical effect whereby, in micro or small enterprises, a limited number of data professionals represents a higher portion of the workforce. Indeed, Table 2, below, displays a direct proportionality between the average number of data professionals employed in a company and the size of the company in terms of total number of employees.

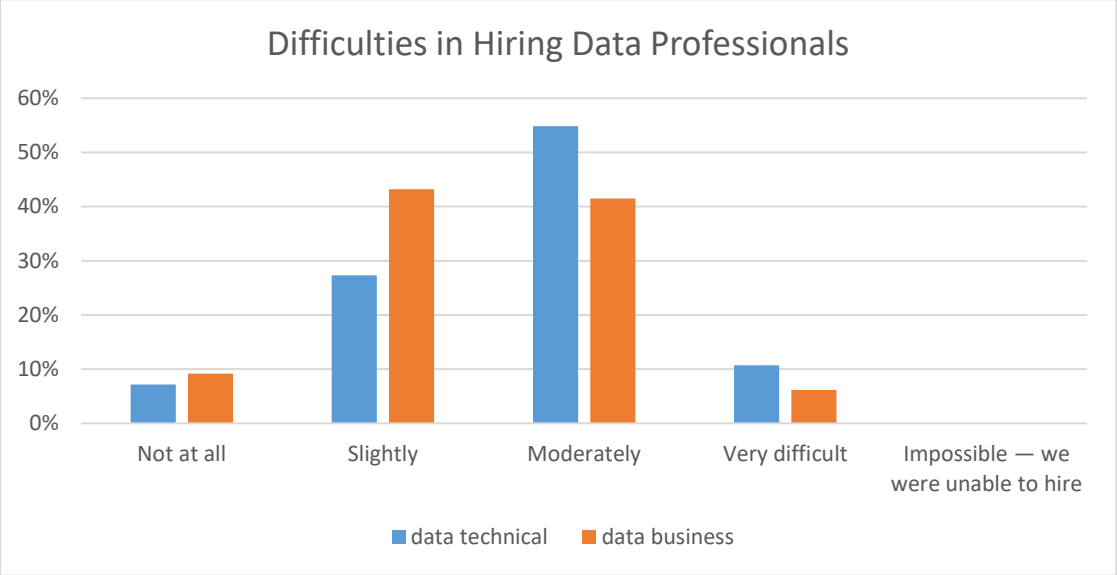
Table 2. Average Number of Data Professionals

Company Size				
1 to 9 employees	**10 to 49 employees**	**50 to 249 employees**	**250 to 499 employees**	**500+ employees**
0.85	4.2	6.2	15.1	115.5

About a third of the respondents affirm that they had hired a data professional in the last 12 months. Moreover, 30% of the enterprises plan to do so soon.

The results show that a hiring problem exists in this domain. The survey asked if there is a problem in gaining the desired number of professionals in order to fill the "skill gap." While none declared they could not hire the needed figures at all, a very notable 65% and 50% find it moderately difficult or very difficult to hire technical/data business professionals, respectively. The problem is slightly more evident for data business professionals than for data technical professionals.

Figure 13. Difficulties in Hiring Data Professionals

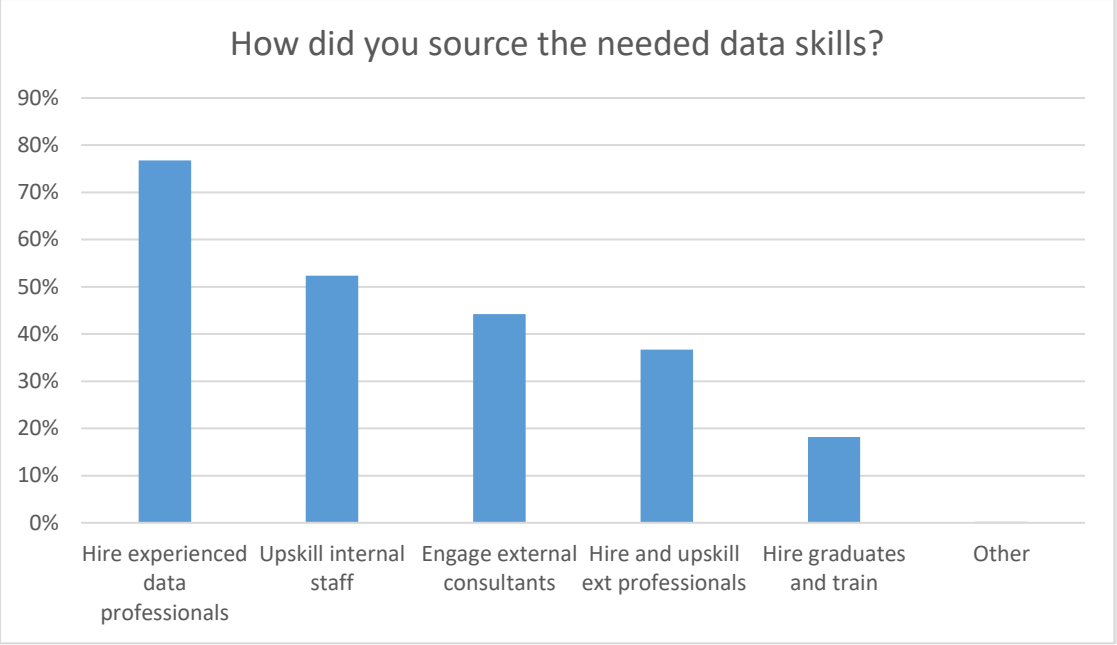


Source: European Data Market Study Survey, 2021

The most widespread way of "filling the skills gap" from this survey is hiring experienced staff. It is a different result from a similar recent IDC survey about AI/data science, in which it appeared that many people had decided to train internally or to cross-train employees from other areas.

Still, the prevailing approach is likely mixed: training existing employees or using external consultants is also a chosen path for several enterprises.

Figure 14. Sourcing the Needed Data Skills



Source: European Data Market Study Survey, 2021

2.3 Forecast Scenarios Methodology

In this paragraph, we present the updated scenario methodology, which will be used in the project to develop three scenarios for the evolution of the European data market and data economy to 2030, based on alternative development paths and driven by different macroeconomic and framework conditions. This methodology builds on the successful approach applied in the study in the past two editions, with some improvements concerning the development of assumptions and their validation through brainstorming workshops with external and internal experts.

The three scenarios provide the storylines, the contextual framework, and the main assumptions used to model and forecast the EDM Monitoring Tool indicators, with a specific focus on the role of policies. Therefore, the scenarios and the forecast models will be developed in parallel, testing their relative coherence and fine tuning their results. The time horizon to 2030 is likely to require a wider potential variation of social and business dynamics. Scenarios are not predictions but potential development paths: Their value-added lies especially in thinking through the potential consequences of different market trajectories and therefore providing a guide to action, particularly for policy makers.

The ultimate objective of the scenarios is to analyse which combination of framework conditions and policy actions may maximise the growth potential of the European data market and economy and, by feeding into the quantitative models, estimate the actual size and depth of their potential benefits. In this way, the scenarios provide a realistic approach to the forecast estimates – since we project a range of values (not a single estimate, which may be widely off the mark) – and provide guidance on the potential consequences of different external events or alternative policy choices.

Given the COVID-19 pandemic crisis, as well other macroeconomic uncertainties, such as the recent trade war between US and China, we will make improvements to the methodology by organising additional brainstorming workshops to deal with these uncertainties.

The scenarios are structured as follows:

- **Baseline scenario**, with the main assumptions based on the continuation of current growth trends and the evolution of current framework conditions
- **High Growth scenario**, whereby the data market enters a faster growth trajectory, thanks to more favourable framework conditions
- **Challenge scenario**, whereby the data market grows more slowly than in the Baseline scenario because of less favourable framework conditions and a less positive macroeconomic context

The general methodology approach is organised with the following main steps:

- Desk research to revise/update assumptions
- Definition of assumptions by scenario
- Brainstorming and validation of assumptions
- Development of scenario storylines and integration of data
- Finalisation of scenarios and development of policy insights

The scenario model used in this study is based on the definition of alternative assumptions about four main groups of key factors. This model was developed and further implemented in various projects and studies published by IDC.

The four main groups of factors are:

1. Macroeconomic factors
2. Policy/regulatory conditions
3. Data market dynamics factors
4. Global megatrends affecting all technology markets

Macroeconomic Factors

We have developed a baseline scenario to 2025 and alternative scenarios from 2025 to 2030. As inputs to the scenarios, we have estimated the value of ICT spending, as well as of GDP for the European Union countries (EU 27), the rest of the European Economic Area (EEA, NO, LI, IS), and for Switzerland and the UK (as shown in the below table).

Table 3. Macroeconomic Data Forecast for Scenarios

	2019	2020	2021	2025 Baseline	2030 Challenge	2030 Baseline	2030 High Growth
	ICT Spending						
EU27	528,969	531,181	558,396	638,747	665,610	770,537	870,062
UK	141,148	144,237	150,063	172,988	187,192	216,470	243,581
Switzerland	30,311	30,223	31,167	36,854	39,454	43,667	48,160
EEA (NO, LI, IS)	13,347	13,593	14,146	16,090	17,527	19,690	21,894
	GDP						
EU27	12,515,078	11,748,729	12,339,393	13,600,101	14,076,599	14,958,559	15,465,050
UK	2,198,318	1,981,850	2,116,615	2,341,233	2,402,358	2,561,455	2,649,620
Switzerland	653,733	637,350	660,932	711,296	729,866	778,202	804,987
EEA (NO, LI, IS)	384,443	378,280	389,774	434,630	448,630	475,055	489,638

Source: EDM Monitoring Tool 2021–2023, November 2021

- The forecast EU27 GDP cumulative growth rate for the 2025–2030 period ranges from 0.7% for the Challenge scenario, 1.9% for the Baseline scenario, and 2.6% for the High Growth scenario.

Policy and Regulatory Conditions

The potential role of policies in addressing the primary challenges associated with the development of the emerging data market and data economy has been extensively discussed in recent years, both within and outside of the EU.

The **Data Strategy**[8] is the cornerstone of the EU's digital strategy. It outlines the ambition for Europe to become a leading role model for a society empowered by data to make better decisions in business and the public sector and a global leader in the data-agile economy. This stems from concerns that Europe is already lagging behind its competitors (the US and China), not only because they have access to more data, but also because they have more effective extraction methods and extract more value from it.

In the EU, specific legislations on data besides GDPR have been adopted – mainly, in verticals and specific fields, like the automotive industry and the finance sector. However, the strategy states that the EU should have a fit-for-purpose legislation and policy environment to enable more use of data in Europe.

The first outcome of the data strategy was the adoption of a Commission proposal for a **Data Governance Act (DGA**[9]) back in November 2020. The DGA aims to strengthen mechanisms that increase data availability. The proposal also aims to enable the safe reuse of certain categories of public-sector data that are subject to the rights of others. In a way, it complements the *Open Data Directive* from 2019, which does not cover the types of data that DGA aims to address.

One other key element complementing the DGA is the so-called "**Data Act**". With this initiative, the Commission is planning to present horizontal legislation to address difficulties in accessing and using data in specific situations, such as in the B2B context. The Data Act's core element is fairness. With this initiative, the Commission aims to ensure the increasing fairness of the data economy.

The COVID-19 crisis has pushed leaders and governments around the world to rethink and revise priorities and strategies. As a result, the European Commission has updated its digital and industrial strategy. In March 2021, a new vision for Europe's digital transformation by 2030 was presented. The **2030 Digital Compass: the European way for the Digital Decade** sets out concrete 10-year targets for the EU's digital ambitions and priorities. Four areas of focus are outlined in the Compass document: 1. digitally skilled citizens and highly skilled professionals, 2. secure, performant, and sustainable digital infrastructures, 3. the digital transformation of businesses, 4. the digitalisation of public services. Compass was enshrined with a Policy Programme, under "Path to the Digital Decade", which now needs to be agreed with the European Parliament and the Council.

[8] EUR-Lex - 52020DC0066 - EN - EUR-Lex (europa.eu)
[9] EUR-Lex - 52020PC0767 - EN - EUR-Lex (europa.eu)

Table 4 EC Legislative Priorities and Initiatives 2021–2022

EC Legislative Priorities 2021–2022: A Europe Fit for the Digital Age
Proposal for a REGULATION on European Data Governance (Data Governance Act)
Proposal for a REVISION of Directive 2003/98/EC on the reuse of public sector information (Open Data Act)
Proposal for a REGULATION on the Data Act (legislative, including impact assessment, Article 114 TFEU, Q3 2021)
Proposal for the REVISION of Directive 96/9/EC on the legal protection of databases (Database Directive)
Proposal for a Regulation for Digital Services (Digital Services Act) and amending Directive 2000/31/EC
Proposal for a REGULATION on contestable and fair markets in the digital sector (Digital Markets Act)
Proposal for a REGULATION laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts
Digital levy and a proposal for a digital levy as own resource (legislative, including an impact assessment, Q2 2021)
European Cyber Resilience Act (legislative, including an impact assessment, Q3 2022)

Source: European Commission Work Programme 2021–2022

Data Market Dynamics Factors

The data market dynamics factors correlate with the main drivers of and barriers to data market growth and the interaction between supply and demand stakeholders.

As shown by Table 5 below, there is a range of relevant assumptions to be considered and that needs to be further examined.

Table 5. Selection of Critical Factors for Data Market Dynamics Assumptions

DATA MARKET ASSUMPTIONS	Level of Impact on Data Market	Level of Uncertainty
Data technologies supply-demand dynamics	Very High	Medium
Development of the data ecosystem in Europe	Very high	High
Managing data ethics and AI business risks	High	Very high
Deployment of 5G infrastructures	Very high	High

Source: extract from D2.7, Final Report on Facts and Figures – April 2020, Uptake of the EDM Monitoring Tool

Global Megatrends Affecting All Technology Markets

The fourth group of factors influencing the scenarios includes global megatrends – technological, but not only, since the digital market is global. IDC's global FutureScape predictions have proven to be a highly valuable guide for selecting megatrends.

IDC FutureScape: Worldwide IT Industry 2022 Predictions[10] presents the top 10 predictions and key drivers of the IT industry for the next five years. It highlights the medium-term and long-term

[10] IDC FutureScape: Worldwide IT Industry 2022 Predictions

challenges that enterprise IT teams face as they define, build, and govern the technologies required to thrive in a digital-first world. According to this document, by 2022, more than half the global economy is based on or influenced by digital; by 2023, 90% of worldwide organisations will be prioritising investments in digital tools to augment physical spaces and assets; and, by 2024, 55% of all ICT investment will be linked to digital transformation[11].

Three Development Paths to 2025: The Baseline Scenario, Challenge Scenario, and High-Growth Scenario

In the last EDM study, the High Growth scenario was characterised by a combination of a high pace of diffusion of data-driven innovation and a distributed model of data governance, while the Challenge scenario showed a low pace of diffusion of data-driven innovation and an uneven distribution of benefits, with a lack of trust in data sharing and a weak data ecosystem. The Baseline scenario was a mix of the two, but closer to the High Growth scenario, with some shortcomings and weaknesses.

International competitiveness is likely to play an important role in the 2030 scenarios. The development of competition in the data world is characterised by very high first-player advantage mechanisms and economies of scale for the leading actors. The leading big tech players, also known as GAFAM, have been able to take the lion's share of their markets concentrating the main data assets. Until recently, these actors have been able to take a disproportionately large part of the advantages of data, hence increasing societal and market inequality. However, Europe seems to be leading the way when it comes regulating the tech giants. Almost a year ago, Margarete Vestager, Executive Vice-President for Europe Fit for the Digital Age, presented an ambitious proposal[12] to reform the European digital space with a set of comprehensive new rules for digital services (social media, online marketplaces, and online platforms). The Digital Services Act[13] introduces a series of new obligations for digital services to ensure the fundamental rights of users protected and to establish a level playing field for business. The Digital Markets Act addresses the negative impact and the dominance of certain platforms acting as “gatekeepers”.

It is not surprising that high-tech companies have continued to thrive even during the economic recession caused by the COVID-19 pandemic. Right after the outbreak of the crisis, the largest players announced enormous quarterly earnings. While delivering its revenues report, Microsoft’s CEO, Satya Nadella, noted that the world has seen two years’ worth digital transformation in only two months[14]. However, parliaments and governments are becoming more conscious of the benefits that big multinational corporations receive, particularly eluding the tax burden weighing on traditional companies, which is focusing the attention of anti-trust authorities on them. The introduction of a global digital tax has been high on the agenda in these past years, despite a pushback from several governments. (Following years of negotiations led by the OECD on digital tax, an agreement was reached in October 2021[15].) Endorsed by the G20 leaders, the international tax system will ensure that the world’s 100 biggest companies will be subject to a minimum 15% tax rate from 2023. The introduction of a global digital tax is a breakthrough in reforming of the digital economy.

In the time period from 2021 to 2030, it is likely that the European digital policy and regulatory environment will have to be updated and revised.

[11] Ibidem
[12] Europe fit for the Digital Age: Digital Platforms (europa.eu)
[13] EUR-Lex - 52020PC0825 - EN - EUR-Lex (europa.eu)
[14] 2 years of digital transformation in 2 months (microsoft.com)
[15] International community strikes a ground-breaking tax deal for the digital age - OECD

On the other side, the emergence of small digital native companies may represent a positive and innovative factor, creating a new competition level, with a society characterised by an open, transparent, and participatory approach to data governance, whereby both citizens and organisations are able to control and extract value from their data.

This analysis highlights the critical turning points to be faced in the coming years by governments, businesses, and social actors in the development of the EU data economy. The combination of alternative social and economic trends results in the following scenarios:

- The Baseline scenario is characterised by a substantial expansion of data innovation, a moderate concentration of power by dominant data owners, with a data governance framework safeguarding personal data rights and an uneven but relatively wide distribution of data innovation benefits in society.

- The High Growth scenario is shaped by a high level of data innovation, low data power concentration, an open and transparent data governance model with high data sharing, and a wide distribution of data innovation benefits in society.

- The Challenge scenario is characterised by a low level of data innovation, a moderate level of data power concentration due to digital markets fragmentation, and an uneven distribution of data innovation benefits in society.

Baseline Scenario

The Baseline scenario envisions a healthy growth of the European data industry, an ongoing improvement in providing data products and services, and a corresponding gradual growth in demand, particularly among the most advanced, competitive, and inventive firms, large and small. Overall, the economic impact of legislative initiatives such as the Digital Governance Act and the Data Act is expected to materialise in more data driven services and products. Creating an equitable environment for all the players of the Digital Single Market.

With Next Generation EU, the Commission deployed a €750 billion temporary extraordinary recovery mechanism to assist governments in repairing the social and economic damage caused by the epidemic. In this scenario, we expect that a great proportion – between 20% and 35%, depending on the country – of national recovery and resilience plans will be spent on ICT by 2025[16].

The scenario also envisions that the EU's digital ambitions for 2030 laid down in Digital Compass would most likely be revised once the mandate of the current Commission expires. Moreover, as achieving these targets requires a joint effort by the EU Member States, this scenario foresees a slight delay in their delivery. Strong lobbying effort and the interdependencies of the dominant tech players would also lead to a compromised, less ambitious, policy environment.

This Baseline scenario is positioned between the two extremes of a high and a low concentration of power and data control.

High Growth Scenario

In the High Growth scenario, Europe's economic growth will be faster than in the Baseline scenario over the next few years, driven by a bigger role of digital innovation and higher overall ICT investments

[16] Recovery Plans in Europe and Investments in Traditional and Emerging Technologies: An IDC Overview

as a share of GDP. Before 2030, more solutions incorporating breakthrough digital technologies (such as IoT, cloud, and big data) will be implemented, and more European businesses will embark on a journey of digital transformation.

The data market will accelerate its expansion, and data technology will extend beyond a small group of pioneers to a larger audience of mainstream users. We foresee that the goal set up in the Data Strategy would be fulfilled, meaning that, by 2030, an attractive policy environment is created whereby the EU's share of the data economy corresponds to its economic weight. The supply-demand dynamics will shift from technology-driven to demand-driven, with a fully formed ecosystem creating positive feedback loops between data firms and users. This is a common virtuous cycle process that might occur if data technology adoption accelerates sufficiently to build momentum. A rapid diffusion, a characteristic of ICTs, doubles the advantages to users in their interactions and makes it simpler to consolidate standards and interoperability, further lowering adoption barriers.

To enable this scenario, we must presume a set of favourable framework conditions. If all policies are implemented and the Commission succeeds in reaching its Digital Decade target, we assume an ideal pace of digital transformation and the diffusion of technologies such as IoT, cloud, 5G, 6G, etc.

Challenge Scenario

The Challenge scenario is a combination of a less favourable macroeconomic environment than in the Baseline scenario, less favourable framework circumstances, and the slower spread of digital innovation, resulting in a low growth path for the data market. This is a fragmented scenario in which the digital single market does not materialise until after 2030. Because demand pull will be minimal, the supply-demand dynamics will be dominated by the technological push. As a result, this scenario examines the possible risks and implications of failing to remove impediments to the growth of Europe's data economy.

Both supply-side and demand-side measures will have a smaller impact and will be implemented more slowly in this scenario. Policy measures will continue to succeed in growing the data industry through R&D investments and the encouragement of digital entrepreneurship, although to a lesser extent due to the private sector's weaker investment tendency.

Policies aimed towards enabling circumstances, such as the elimination of regulatory barriers to the free cross-border flow of data, will be implemented more slowly and with less effectiveness than in the Baseline scenario. As a result, the value of the data market and data economy will be significantly lower in 2030 than in the Baseline scenario.

3. Data Professionals

3.1 Definition

Data professionals[17] *are workers who collect, store, manage, and/or analyse, interpret, and visualise data as their primary activity or as a relevant part of their activity. Data professionals must be proficient with the use of structured and unstructured data, should be able to work with a huge amount of data, and should be familiar with emerging database technologies. For 2021–2023, the definition of data professionals was refined to differentiate the roles played by different data users: These are data technical professionals, data business professionals, and data consumers – defined below. The measure of data professionals includes data technical professionals and data business professionals only.*

Data technical professionals are specialists in the collection, storage, management, modelling, and quality assurance of data, as well as the integration of various data sources to ensure consistency, accuracy, and quality of data. A data technical professional can, given the question that needs to be answered, ensure that the data supply chain is provided and that it is accurate.

Data business professionals have as a primary or significant focus the task of performing predictive analysis, qualitative analysis, data modelling, data extraction, and data summaries with the purpose of creating new insights and knowledge from available data. They have thorough industry and/or process understanding and can put data analysis into context and relate to existing trends within the industry or line of business they are in. They typically leave collection, management, and quality of data to the data technical professional but, using analysis tools such as Excel, Tableau, and Power BI, are able to summarise large amounts of data and visualise and present trends and insights to a wider audience of key stakeholders in the business in order to drive the strategic decision-making process in the organisation. Data scientists predominantly reside within the data business professional group.

Data consumers are product, process, human-resource, asset, or department employees and managers responsible for driving change or maintaining a position whereby decision making is heavily reliant on the supply of data and insights based on large amounts of data. They work directly with data only part of the time. They are decision makers or stakeholders in a decision process whereby the data and insights provided determine the quality of the decisions made. A data consumer guides the business based on the data and insights provided through the data supply chain.

3.2 Measuring Data Professionals

For 2021–2023, data professionals are defined as above and are measured using the EU Labour Force survey, with the data built from the International Labour Organisations ILOSTAT database and Eurostat's Structural Business Statistics and Business Demography. This data allows the modelling of data professionals by Member State, NACE II industry, and company size band. The data sources used are as follows:

[17] The previous European Data Market Study (SMART 2013/0063) included an indicator measuring "data workers", which was based on a similar, but slightly more restrictive, definition. In line with the First Report on Facts & Figures (D2.1), in this document, we measure "data professionals" – that is, workers with a wider range of data-related roles. Indeed, data professionals are not only data technicians, but also users who, based on sophisticated tools, take decisions about their business or activities after having analysed and interpreted the available data.

Table 6. Data Professionals – Data Sources

Data Source	Updated
Eurostat Business Demographic Statistics	Sep 2021
Eurostat Annual Structural Business Statistics	Sep 2021
IDC Worldwide Black Book (standard edition)	Aug 2021
ILOSTAT statistics and databases	Mar 2021

Prior data for 2019 and 2020 is adjusted to reflect the refined definitions, and the prior 2020 data was delivered as a forecast, which is superseded by measured data. In addition, prior publications segmented by size band and Member State included the UK as a member of the EU28. Since January 2020, the UK is no longer a member of the European Union, so EU27 totals exclude the United Kingdom. However, the UK is still a dynamic user and supplier of data alongside the EU: Together with Switzerland, the EEA countries of Iceland, Liechtenstein, and Norway are included in the statistics.

An overview of Data Professionals is included below.

Table 7. Data Professionals – Summary Statistics, 2019–2021

N.	Region	Name	Description	2019	2020	2021	Growth Rate 2020–2021
1.1	EU27	Number of data professionals	Total number of data professionals in EU (000s)	6,026	6,502	6,853	5.4%
1.2	EU27	Data professionals share of total employment	Share of data professionals on total employment in EU (%)	3.4%	3.6%	3.8%	3.0%
1.3	EU27	Intensity share of data professionals	Average number of data professionals per user company (units)	11.3	12.0	12.4	3.4%

Brexit and COVID-19 had little impact on the number of data professionals and, in fact, might well have accelerated demand for data-focused workers that can be employed from home, as organisations realised the benefits and capabilities of home working for independent data workers. In addition, the data economy grew as many people moved from face-to-face business to remote business, easing the transition to a more data-oriented economy.

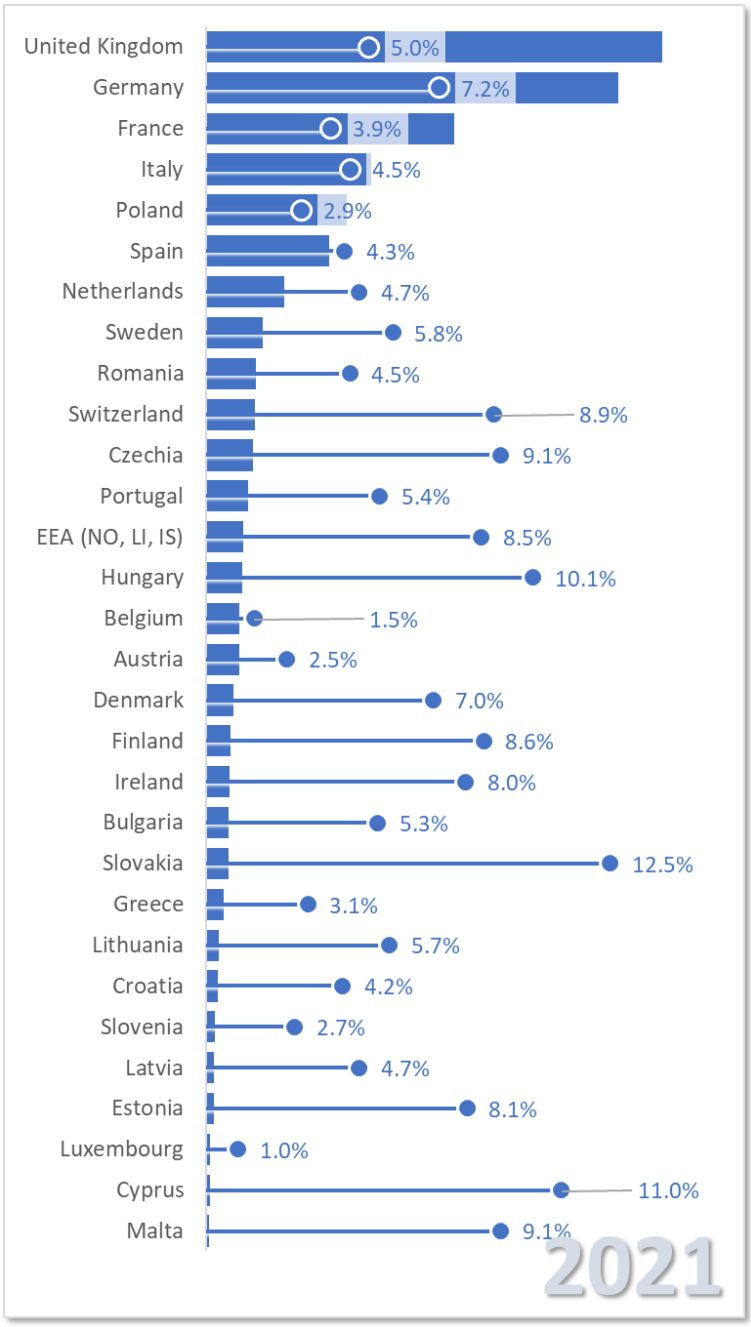
Data Professionals by Country

Table 8 shows the size and growth of the number of data professionals by Member State for 2019 and 2020 and the forecast for 2021. Growth overall was high in 2020 but falls in 2021 as the supply of data professionals is used up. Demand is strong, but there is a real shortfall of data professionals across the Member States, leaving demand unfulfilled. Figure 15The largest Member States, of course, have the biggest demand. Growth in the Member States (EU27) is marginally lower than across all countries, but only by 0.1 percentage point (5.4% for EU27 versus 5.5% for all countries). The EEA states (NO, LI, IS) show high growth in 2021, but this is from a lower base than for other countries and is more an artefact of the differing stages of the economy than any fundamental benefits the EEA states see in 2021.

Table 8. Data Professionals by Country: 2019–2021 and 2020 Growth Rate

Country	2019	2020	2021	CAGR
Austria	124	134	138	2.5%
Belgium	134	135	137	1.5%
Bulgaria	78	86	90	5.3%
Croatia	41	45	47	4.2%
Cyprus	14	14	15	11.0%
Czechia	157	177	193	9.1%
Denmark	98	105	113	7.0%
Estonia	26	28	30	8.1%
Finland	88	93	101	8.6%
France	948	986	1,024	3.9%
Germany	1,466	1,589	1,704	7.2%
Greece	69	69	71	3.1%
Hungary	129	136	150	10.1%
Ireland	81	90	97	8.0%
Italy	592	653	682	4.5%
Latvia	30	31	32	4.7%
Lithuania	44	47	50	5.7%
Luxembourg	14	14	15	1.0%
Malta	8	8	9	9.1%
Netherlands	301	309	323	4.7%
Poland	477	564	581	2.9%
Portugal	154	165	174	5.4%
EEA (NO, LI, IS)	131	141	153	8.5%
Romania	185	197	206	4.5%
Slovakia	79	82	93	12.5%
Slovenia	33	35	36	2.7%
Spain	438	487	508	4.3%
Sweden	218	220	233	5.8%
Switzerland	183	184	200	8.9%
United Kingdom	1,517	1,795	1,886	5.0%
EU27	6,026	6,502	6,853	5.4%
EEA (NO, LI, IS) + CH	314	325	353	8.8%
Total, all countries	7,858	8,622	9,093	5.5%

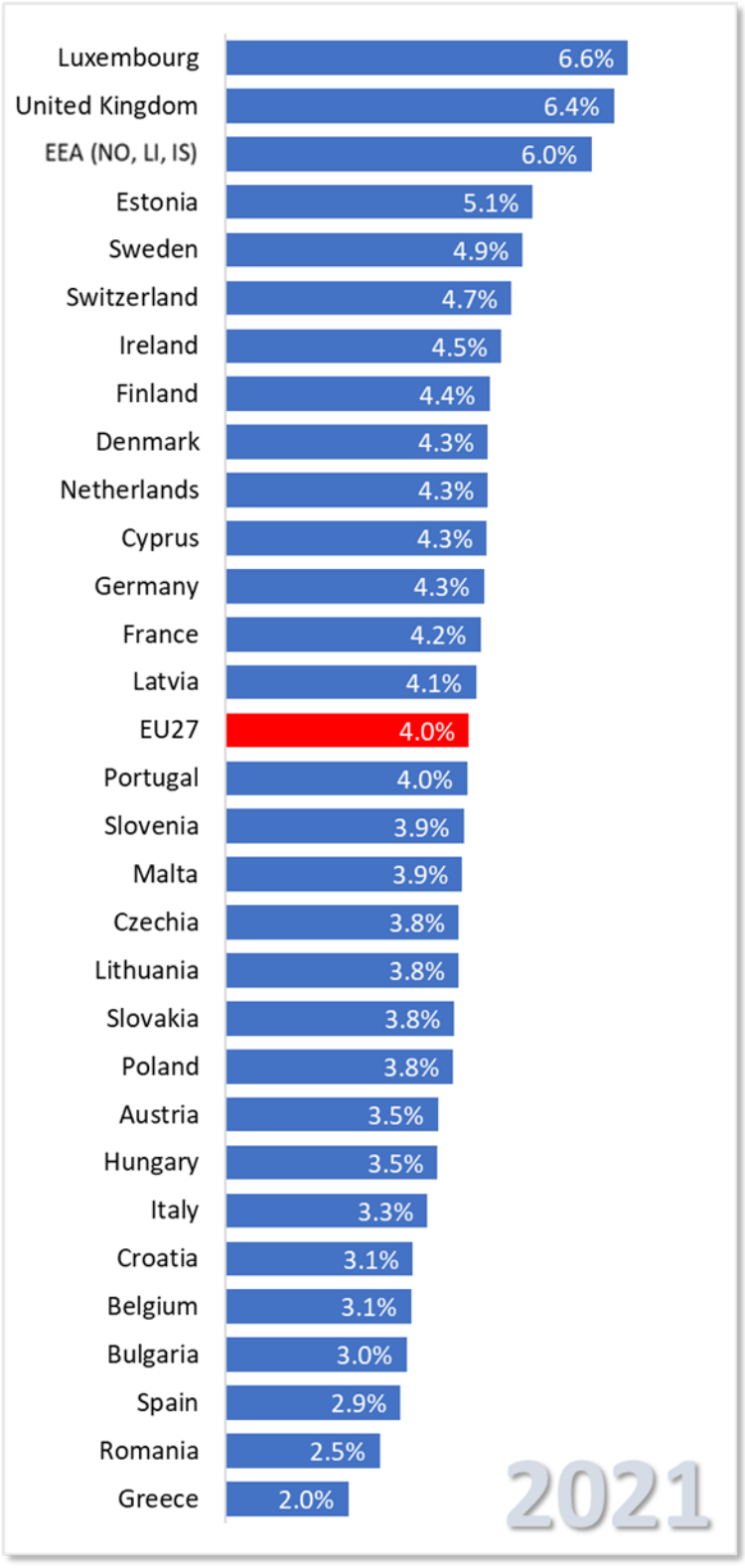
Figure 15. Data Professionals by Country: 2021 and 2020–2021 Growth



The larger Member States dominate the number of data professionals – unsurprisingly – but a view of the share of total employment taken by each Member State shows which Member States have already invested in industry and education to support the anticipated growth in the need for data experts of differing capabilities – data business professionals, data technical professionals, and data consumers. These shares have changed little since previous reports covering this area, reflecting the difficulty anticipated in growing the number of professionals. Data skills need a combination of programming, logic, and statistics, the combination of which is difficult to find. On top of this are the business skills needed to use the data analysis that data skills provide, making the development of data professionals a difficult and time-consuming task. In 2021, Luxembourg leads all the countries in terms of data professionals' share of total employment, aided to a certain extend by the strength of data-oriented industries (e.g., information & communications and professional services), but this strength is also a

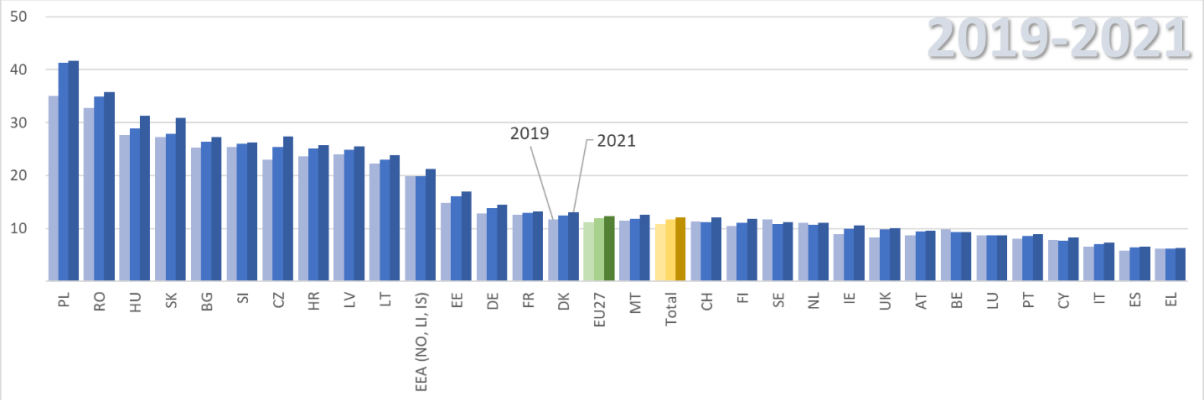
little exaggerated because of the relatively small size of the working population of Luxembourg. The United Kingdom maintains its strength as the second highest share of total employment built more on its industry predilection for data-intensive industries.

Figure 16. Data Professionals – Share of Total Employment by Country, 2021



There is a wider disparity across Member States in the average number of data professionals per data user company, and this again reflects the industries associated with specific Member States or other countries. Those that focus on data-oriented industries have a greater propensity towards data professionals, and a larger proportion of their companies will be data oriented compared to those in industries with a lower level of data involvement – thus leading to a higher average number of data professionals per data company. Despite this though, strong data-oriented countries have a lower average number of data professionals per data company, which reflects the higher volume of user companies in the country rather than a lower number of data professionals.

Figure 17. Average Number of Data Professionals Per User Company by Country, 2019–2021



Data Professionals by Industry

Figure 18. Data Professionals by Industry: 2021 and 2020–2021 Growth

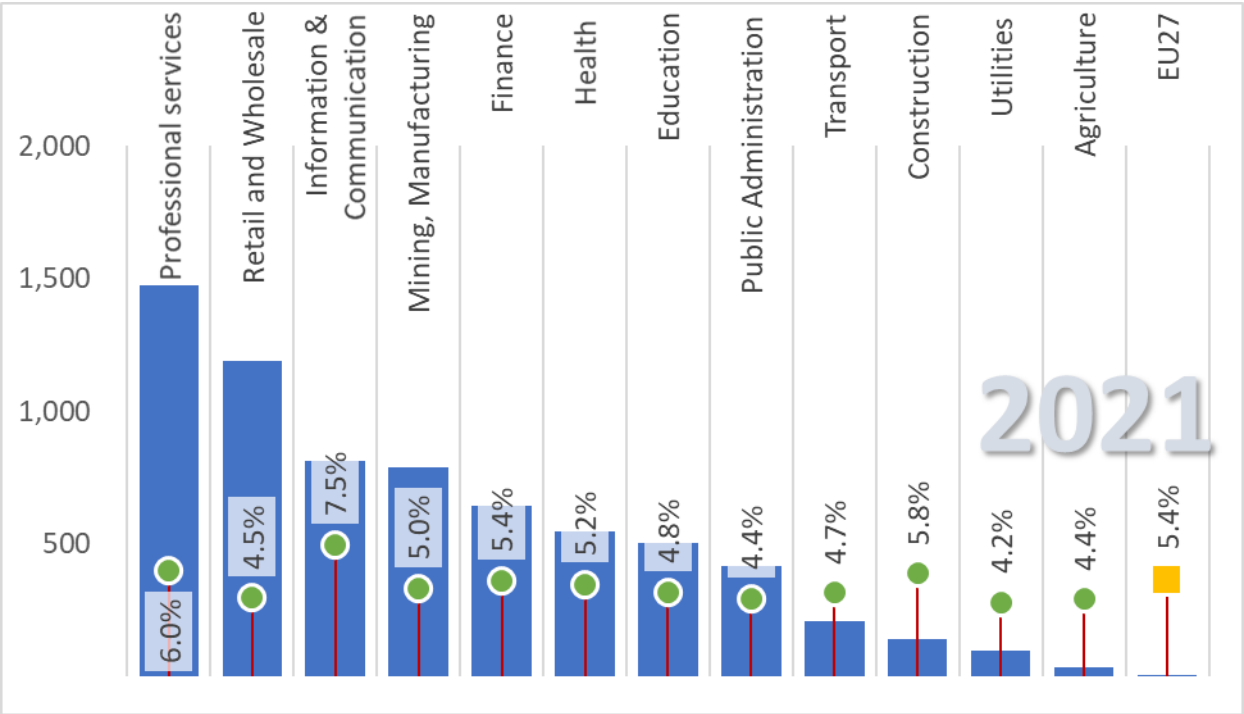


Figure 18 and Table 9 reiterate the previous comments with data – showing the levels of data professionals by industry for the specific NACE II industries tracked in this report. Added to this report is Agriculture, which shows the lowest level of data professionals of all the 12 industries tracked. In

spite of this, the growth of data professionals across all industries is reasonable in 2021, at 5.4% for the EU27.

Table 9. Data Professionals by Industry: 2019–2021 and 2021 Growth Rate

Industry	2019	2020	2021	Growth Rate 2020-2021
Agriculture	32	35	37	4.4%
Construction	124	132	139	5.8%
Education	443	482	505	4.8%
Finance	571	612	645	5.4%
Health	479	521	548	5.2%
Information & communications	692	757	813	7.5%
Mining & manufacturing	704	749	786	5.0%
Professional services	1,277	1,391	1,475	6.0%
Public administration	368	397	415	4.4%
Retail & wholesale	1,064	1,137	1,187	4.5%
Transport	185	198	207	4.7%
Utilities	88	93	97	4.2%
EU27	6,026	6,502	6,853	5.4%

3.3 Data Professionals Forecasts 2025 and 2030

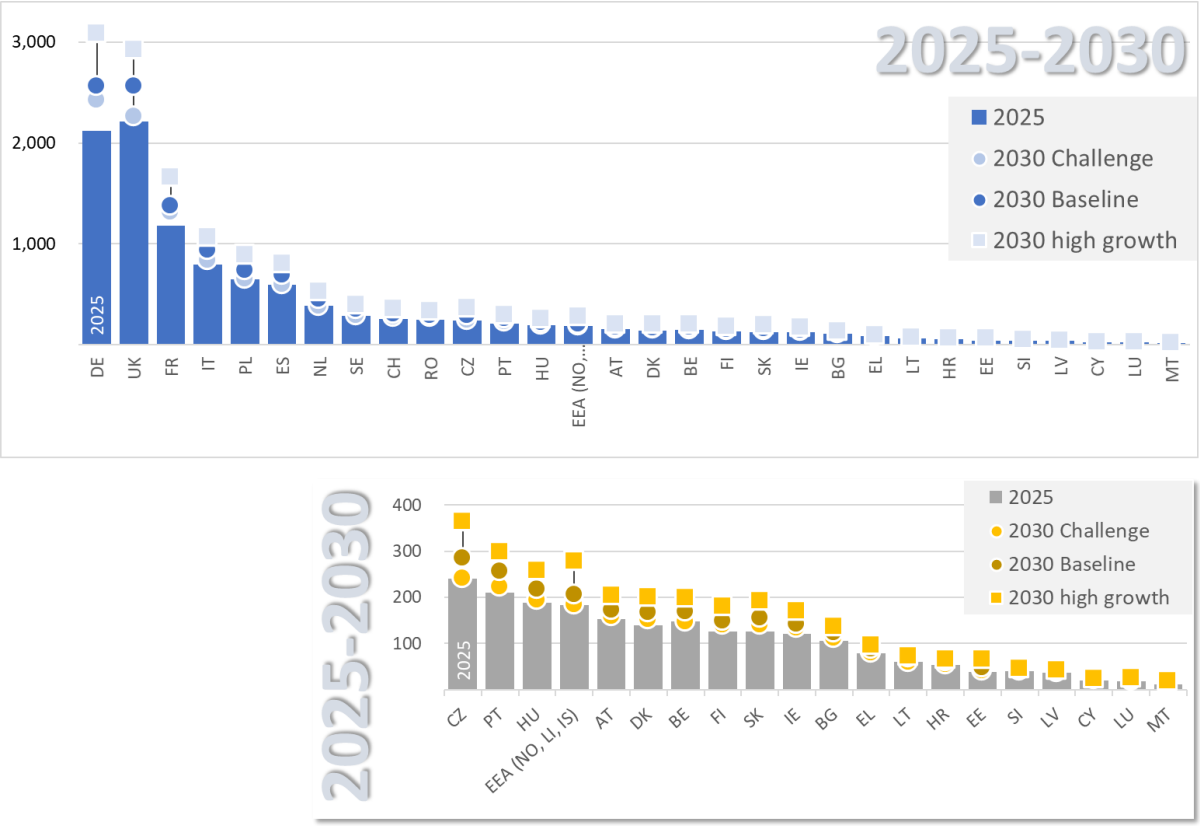
We anticipate strong growth in the number of data professionals over the period of the forecast, from 2020 to 2030. We show three scenarios for 2030 to accommodate potential upsides and downsides to our forecast, as 2030 is 9 years away and as 2019 and 2020 showed there are significant changes to social, economic, technical, and political factors that can have dramatic impacts on the number of data professionals available. In particular, the current demand for data professionals is high, and the supply is not yet meeting this demand. The forecast shows fulfilled demand.

Table 10. Data Professionals Forecast: 2025; 2030 Challenge, Baseline, and High Growth Scenarios ('000s); and CAGR's (%)

	2025	2030, Challenge	2030, Baseline	2030, High Growth	CAGR: 2020–2025	CAGR: '25–'30, Challenge	CAGR: '25–'30, Baseline	CAGR: '25–'30, High Growth
EU27	8,158	8,770	9,630	11,437	4.6%	1.5%	3.4%	7.0%
EEA (NO, LI, IS) + CH	436	458	507	634	6.0%	1.0%	3.1%	7.8%
Total, all countries	10,806	11,490	12,701	14,997	4.6%	1.2%	3.3%	6.8%

Data Professionals by Country: 2025–2030 Forecast

Figure 19. Data Professionals by Country: 2025–2030 Forecast by Scenario



Germany becomes the largest Member State for data professionals by the end of this forecast (2030), accounting for 20% of data professionals in the EU27. Germany will grow above average to 2030 – the only one of the top five economies to do so. Growth in data professionals, although reasonable, will be constrained though, with demand expected to exceed supply over the forecast period because of the difficulty and latency involved in educating data professionals.

Table 11. Data Professionals Forecast by Country: 2025; 2030 Challenge, Baseline, and High Growth Scenario (000's); and CAGR (%)

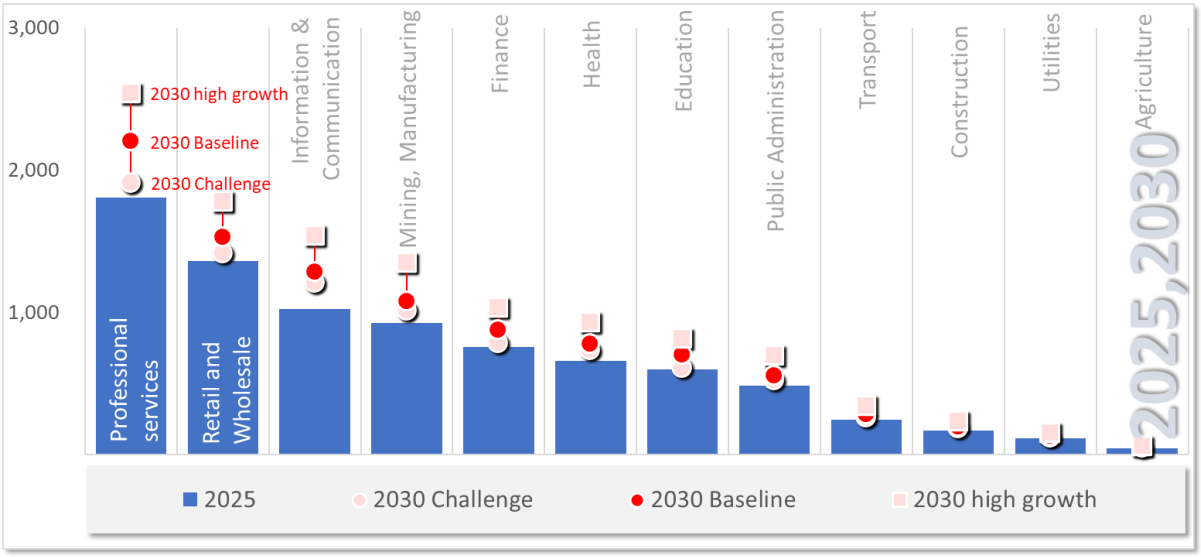
Country	2025	2030, Challenge	2030, Baseline	2030, High Growth	CAGR: 2020–2025	CAGR: '25–'30, Challenge	CAGR: '25–'30, Baseline	CAGR: '25–'30, High Growth
Austria	152	158	172	204	2.5%	0.8%	2.5%	6.1%
Belgium	147	148	168	199	1.7%	0.1%	2.7%	6.2%
Bulgaria	106	112	124	136	4.4%	1.0%	3.1%	5.1%
Croatia	54	54	61	66	3.5%	0.0%	2.6%	4.1%
Cyprus	20	22	24	24	7.5%	2.1%	3.5%	4.2%
Czechia	241	242	285	364	6.4%	0.1%	3.4%	8.6%
Denmark	140	153	168	201	5.8%	1.8%	3.7%	7.5%
Estonia	39	40	48	67	6.7%	0.6%	4.4%	11.3%
Finland	126	141	149	181	6.3%	2.2%	3.3%	7.5%

Country	2025	2030, Challenge	2030, Baseline	2030, High Growth	CAGR: 2020–2025	CAGR: '25–'30, Challenge	CAGR: '25–'30, Baseline	CAGR: '25–'30, High Growth
France	1,178	1,311	1,373	1,661	3.6%	2.2%	3.1%	7.1%
Germany	2,123	2,422	2,563	3,086	6.0%	2.7%	3.8%	7.8%
Greece	80	80	89	96	2.8%	0.1%	2.2%	3.9%
Hungary	188	194	218	258	6.7%	0.6%	3.0%	6.6%
Ireland	120	134	142	171	6.0%	2.2%	3.3%	7.3%
Italy	797	835	933	1,065	4.1%	0.9%	3.2%	6.0%
Latvia	37	37	41	43	3.7%	0.0%	2.0%	3.0%
Lithuania	60	60	70	73	4.8%	0.0%	3.2%	4.0%
Luxembourg	17	19	21	26	3.9%	1.8%	3.9%	8.3%
Malta	12	15	16	19	8.0%	4.3%	5.9%	9.6%
Netherlands	381	384	451	529	4.3%	0.1%	3.4%	6.8%
Poland	648	651	733	889	2.8%	0.1%	2.5%	6.5%
Portugal	210	223	256	299	4.9%	1.2%	4.1%	7.3%
EEA (NO, LI, IS)	182	183	206	278	5.3%	0.1%	2.4%	8.8%
Romania	242	271	287	337	4.1%	2.3%	3.5%	6.9%
Slovakia	125	140	155	192	8.7%	2.2%	4.4%	9.0%
Slovenia	40	40	45	47	2.6%	0.1%	2.3%	3.2%
Spain	590	596	686	806	3.9%	0.2%	3.1%	6.5%
Sweden	285	289	353	397	5.3%	0.3%	4.3%	6.8%
Switzerland	253	275	301	355	6.6%	1.7%	3.5%	7.0%
United Kingdom	2,213	2,262	2,564	2,926	4.3%	0.4%	3.0%	5.7%
EU27	8,158	8,770	9,630	11,437	4.6%	1.5%	3.4%	7.0%
EEA (NO, LI, IS) + CH	436	458	507	634	6.0%	1.0%	3.1%	7.8%
Total, all countries	10,806	11,490	12,701	14,997	4.6%	1.2%	3.3%	6.8%

Data Professionals by Industry: Forecasts 2025, 2030

The main drivers of the consumption of data professionals to 2025 and 2030 will be the professional services and retail & wholesale industries – together, accounting for nearly 40% of data professionals by 2030 (Baseline forecast). The top four industries account for nearly two-thirds (63%) of data professionals by this time, leaving the eight industries to compete for the remaining 37%. Demand will exceed supply unless universities significantly increase the pool of data-literate graduates, supported by industry education in the business skills needed to make valued use of the newly acquired data skills.

Figure 20. Data Professionals Forecast by Industry: 2025 and 2030 by Scenario ('000s)



3.4 Key Findings

Data professionals were redefined for this publication – clarifying the two key types of data technical professionals and data business professionals. A third type – data consumers – is not included in this dataset. As a result, the data for 2019 and 2020 is restated. Data technical professionals focus on building data models for use by data business professionals to make business decisions based on what these models show.

There will be 6.8 million data professionals in the EU27 Member States by the end of 2021, continuing the healthy growth (5.4%) seen in 2019 (5.0%).

Data professionals' share of employment has also increased in 2021 – to 4.0%, up from 3.3% in 2019 – for the EU27 Member States. For another year, this again confirms the positive trend of the share of the workforce involved in data-related professions.

According to the latest estimate, **the average number of data professionals per company is estimated at 12.4 in the EU27 in 2021, up from 11.3 in 2019.**

Data professionals in the EU27 will account for 8.1 million people in 2025 or 9.6 million data professionals according to the Baseline forecast. The number is expected to rise by a compound rate of 3.4% in the EU27 but is likely to be constrained by the limited supply of professionals.

Looking at the penetration of data professionals in total employment across all Member States in 2019, it is clear that the balance is even, aside from some unsurprising outliers, which are data-intense economies, such as Luxembourg and the UK, or economies still struggling to make the most of ongoing digitalisation (e.g.: Greece).

The number of data professionals is highest in the professional services industry, followed by retail & wholesale and information & communications. The spread of the number of data

professionals by industry is notably wider, although the correlation is fair between the number of companies and employees in those industries and the number of data user companies in the same industries.

4. Data Companies

4.1 Definition

Data companies *are organisations that are directly involved in the production, delivery, and/or usage of data in the form of digital products, services, and technologies. They can be both data supplier and data user organisations:*

- ***Data suppliers*** *have as their main activity the production and delivery of digital data-related products, services, and technologies. They represent the supply side of the data market.*
- ***Data users*** *are organisations that generate, exploit, collect, and analyse digital data intensively and use what they learn to improve their business. They represent the demand side of the data market.*

The Final Report on Facts and Figures (D2.7) of the SMART 2016/0063 European data market study applied the same definition of data supplier and data user companies as the one used in the original European Data Market study (SMART 2013/0063), as well as the First Report on Facts and Figures (D2.1) in February 2018. However, the definition of data suppliers is extended in this report to include a wider range of industries. As a result, indicator 2 measures:

- European data suppliers, counted as legal entities based in one EU Member State, as a share of the total number of enterprises included in the industry classifications of A, C, E, G, H, J, K, M, P, and Q from NACE rev2. This extends the previous definition of data supplier companies to accommodate the inclusion of the sale and purchase of data (data monetisation), as well as data software tools, hardware, and data services.
- European data users, counted as legal entities based in one EU Member State, as a share of the total number of private enterprises in the EU. This definition is unchanged from the prior definition.

4.2 Measuring Data Companies

The number of data suppliers is expanded in this cycle, as we now include industries of A, C, E, G, H, J, K, M, P, and Q from NACE rev2. Previously, we included only industry segments of J and M. We also include these industries in the total pool for data suppliers, so the share of data suppliers in this pool drops significantly, as the additional industry segments have very few data supplies included. The expansion of the industries included in data suppliers reflects the inclusion of those industries that offer data in data marketplaces and an extension of the data skills across these industries. The primary source for these is the survey of data users and data suppliers, which gives guidance on the number of companies that offer data services.

Table 12 gives an overview of the number of data supplier companies and data user companies across the Member States. It shows the number of data supplier companies in 2020 was 175,600, an increase of 5.7% from 2019. The number is forecast to rise by 5.8% in 2021 to 185,900. The extended definition of data suppliers adds to the total number. The previous definition included 148,900 in 2019, rising to the new definition of 166,033 for that year. The data users term retained its previous definition and grew to 542,510 companies in 2020. It is forecast to rise to 553,046 in 2021.

Table 12. Indicator 2: Data Companies: 2019–2021 and 2020–2021 Growth

N.	Name	Description	Market	2019	2020	2021	Growth Rate 2020–2021
2.1	Number of data suppliers	Total number of data suppliers measured as legal entities based in the EU ('000s)	EU27	166,063	175,605	185,866	5.8%
2.2	Share of data suppliers	Percentage share of data companies of total companies in the NACE II industries of A, C, E, G, H, J, K, M, P, and Q	EU27	1.6%	1.7%	1.8%	4.9%
2.3	Number of data users	Total number of data users in the EU, measured as legal entities based in one EU country	EU27	534,840	542,510	553,046	1.9%
2.4	Share of data users	Percentage share of data users of total companies in the EU industry	EU27	2.1%	2.1%	2.1%	1.4%

4.3 Measuring Data Supplier Companies

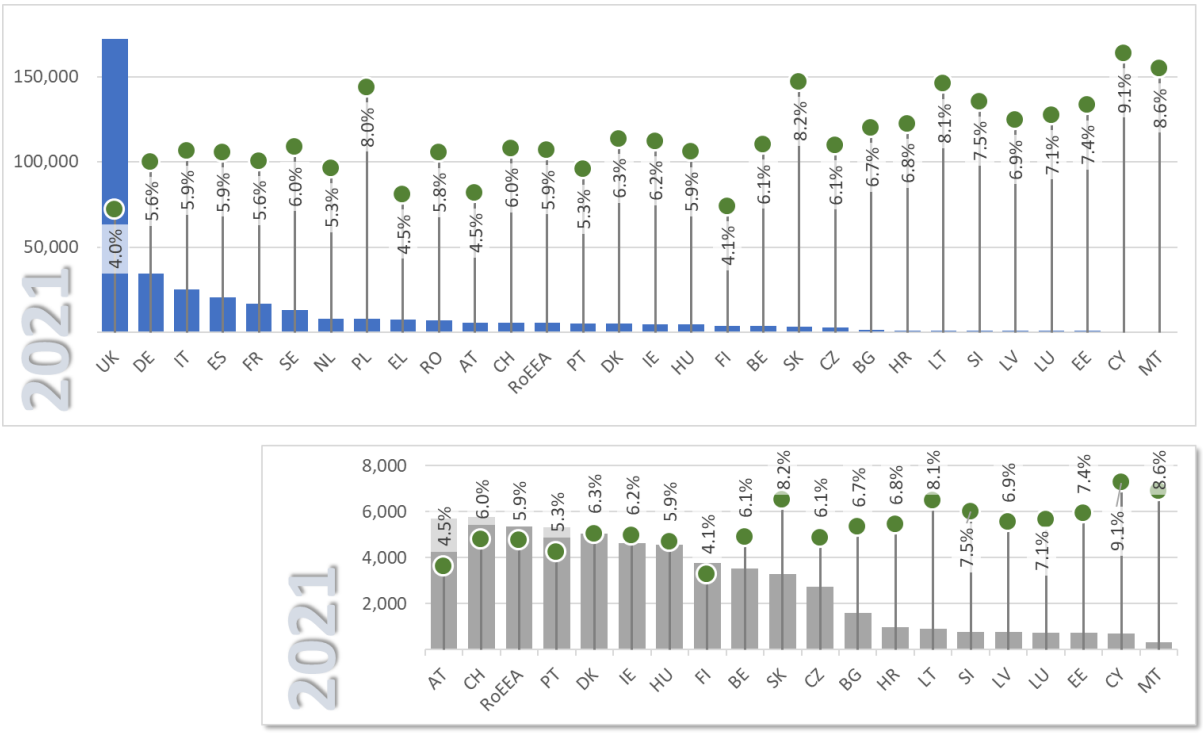
Data Suppliers by Country

In 2021, Germany accounts for the largest source of data supplier companies, representing 18.5% of all Member States, followed by Italy (at 13.5%) and Spain (with 10.9%). The United Kingdom remains the largest source of data supplier companies across the European continent, accounting for 46.6% of all data supplier companies. The United Kingdom already prior to Brexit attracted inward investment from both inside and outside of Europe. Its strategy of closing inefficient manufacturing and resource industries and replacing them with data-oriented services industries (such as finance and professional services) contributed to its success in attracting numerous data providers.

Table 13. Data Supplier Companies by Country: 2019–2021 ('000s) and 2020–2021 Growth (%)

Country	2019	2020	2021	Growth Rate 2020–2021
Austria	5,242	5,448	5,695	4.5%
Belgium	2,900	3,326	3,530	6.1%
Bulgaria	1,394	1,491	1,591	6.7%
Croatia	859	906	968	6.8%
Cyprus	580	631	688	9.1%
Czechia	2,454	2,581	2,738	6.1%
Denmark	4,573	4,760	5,060	6.3%
Estonia	625	665	714	7.4%
Finland	3,513	3,613	3,761	4.1%
France	15,391	15,829	16,712	5.6%
Germany	31,618	32,632	34,445	5.6%
Greece	6,580	6,939	7,251	4.5%
Hungary	4,071	4,301	4,554	5.9%
Ireland	4,238	4,359	4,630	6.2%
Italy	22,528	23,685	25,086	5.9%
Latvia	680	711	760	6.9%
Lithuania	803	837	905	8.1%
Luxembourg	636	677	725	7.1%
Malta	279	298	324	8.6%
Netherlands	6,636	7,570	7,975	5.3%
Poland	6,970	7,226	7,802	8.0%
Portugal	4,851	5,047	5,314	5.3%
EEA (NO, LI, IS)	4,851	5,058	5,358	5.9%
Romania	6,413	6,653	7,041	5.8%
Slovakia	2,788	3,040	3,288	8.2%
Slovenia	669	723	777	7.5%
Spain	18,513	19,212	20,338	5.9%
Sweden	10,260	12,445	13,195	6.0%
Switzerland	5,242	5,448	5,775	6.0%
United Kingdom	157,364	165,454	172,075	4.0%
EU27	166,063	175,605	185,866	5.8%
EEA (NO, LI, IS) + CH	10,093	10,506	11,133	6.0%
Total, all countries	333,519	351,565	369,075	5.0%

Figure 21. Data Supplier Companies by Country: 2021 and 2020–2021 Growth



Data Supplier Companies by Industry.

Agriculture has been added to the 11 previously tracked industries but it currently contributes to a limited extent to the number of data supplier companies. However, the industry shows strong growth, albeit from a small base, so is considerably higher than the average for all EU27 Member States. In addition, agriculture has only recently joined the digital transformation process and is therefore showing strong development trends. The biggest industries reflect the choice made previously to include only J and M NACE industries, and these two industries account for nearly 92% of data supplier companies across the 27 Member States (See Figure 23.) They also show the lowest growth, but this is unsurprising considering the overall growth will be dominated by J and M growth. The smaller industries of agriculture and construction show apparent meteoric growth in 2020, but this is attributable to the low base from which they start.

The below table shows the relative size and growth of each industry for 2020.

Table 14. Data Suppliers by Industry: 2019–2021 and 2020–2021 Growth

Industry	2019	2020	2021	Growth Rate 2020–2021
Agriculture	14	27	41	54.6%
Construction	28	53	84	58.1%
Education	106	147	199	35.3%
Finance	3,377	4,621	6,024	30.4%
Health	2,320	3,101	4,040	30.3%

Industry	2019	2020	2021	Growth Rate 2020–2021
Information & communications	92,118	93,244	94,750	1.6%
Mining & manufacturing	13	21	36	68.6%
Professional services	62,910	67,958	72,759	7.1%
Public administration				
Retail & wholesale	4,898	6,055	7,426	22.6%
Transport	89	117	154	31.0%
Utilities	190	261	352	34.8%

Figure 22. Data Supplier Companies by Industry: 2021 and 2020–2021 Growth

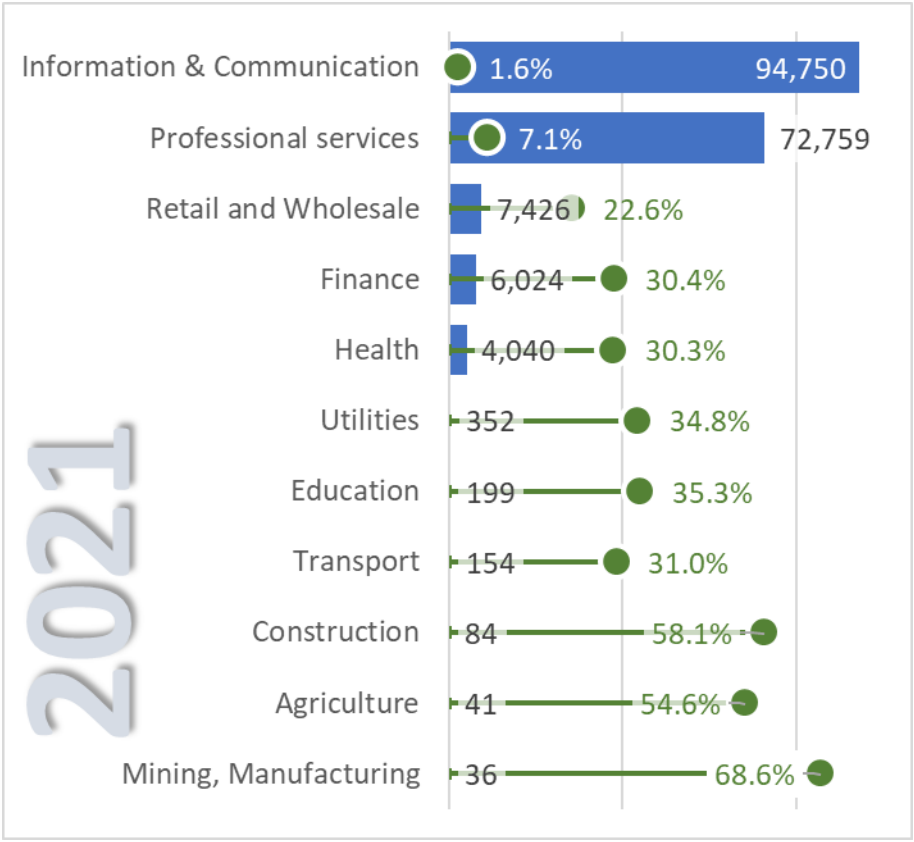
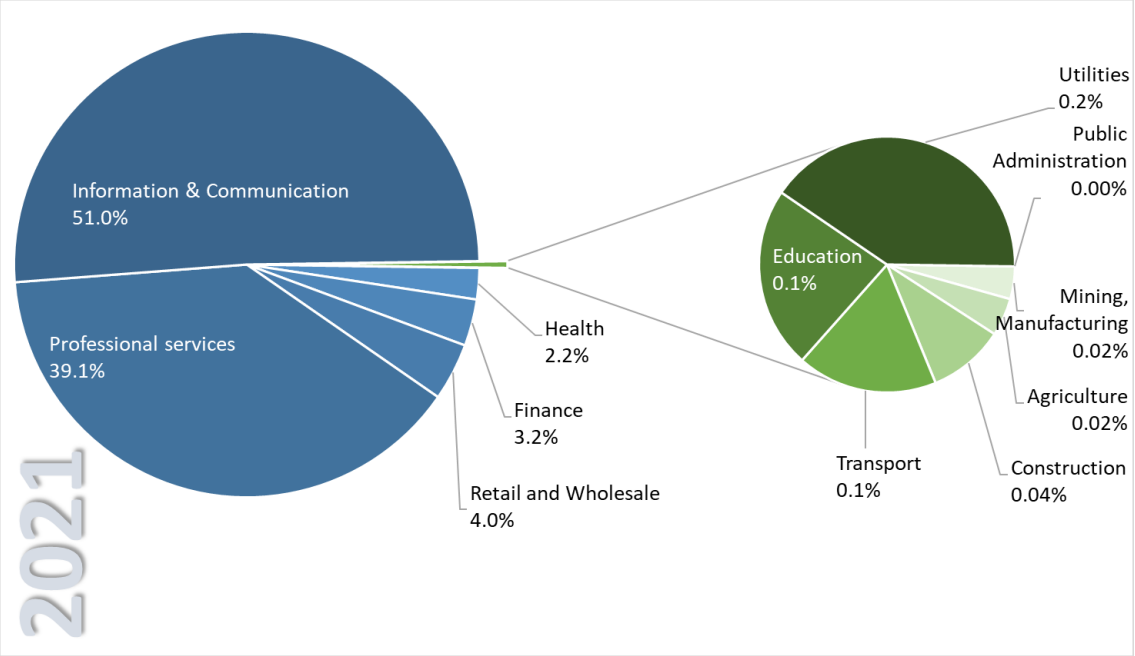


Figure 23. Data Supplier Companies: 2020 Share by Industry



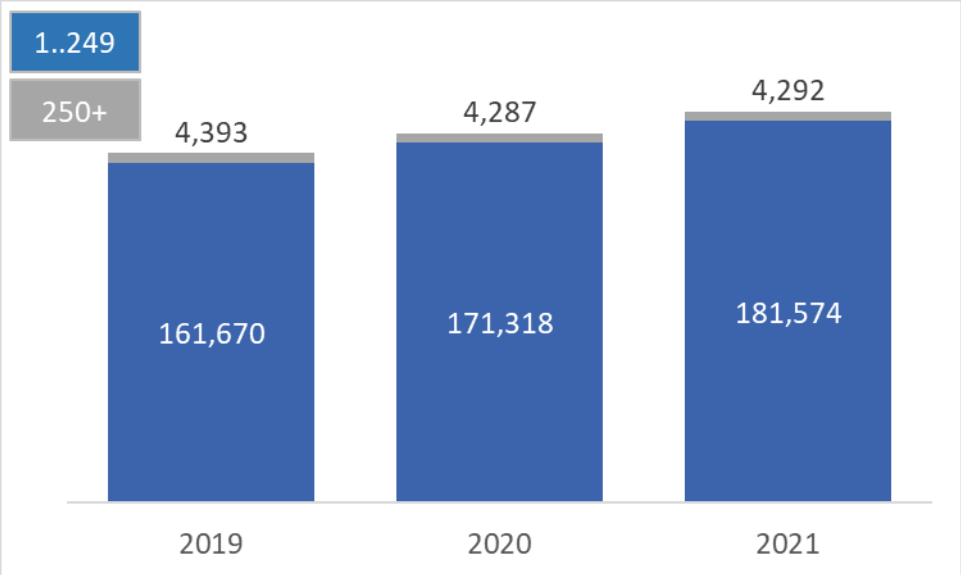
Data Supplier Companies by Company Size Band

Data suppliers by size band shows that the number of companies in the larger size band represents only 2.3% of all data supplier companies in 2021. However, this reflects the large number of small and medium sized businesses overall. More than 95% of businesses are sized below 10 employees, so the apparent poor showing of larger companies is no surprise. In terms of revenues, larger companies account for a much larger share. Table 15 shows the relative sizes of companies with fewer than 250 employees and companies with 250 or more employees for 2019–2021. There are no notable changes in the ratios of these two size bands.

Table 15. Data Suppliers by Company Size Band with Growth, 2019–2021

	2019	2020	2021	AGR 2021
250+ employees	4,393	4,287	4,292	0.1%
1–249 employees	161,670	171,318	181,574	6.0%

Figure 24. Data Supplier Companies by Company Size Band, 2019–2021



4.4 Measuring Data User Companies

There are significantly more data user companies than data supplier companies – with the EU27 including 553,046 data user companies in 2021 compared with 185,866 data supplier companies. With 66% of all data user companies in the European Union in 2021, Germany, Italy, Spain, and France account for the vast majority of data user companies. The United Kingdom has the highest number of data user companies due to its service-based economy, which embraces the impact of data on the economy. Overall growth in the number of data companies is slow, at 2.1% in 2021, but this is higher than the growth in the total number of companies in the European Union (1.9%). Figure 26 shows the share of total companies that data companies take in 2021, accounting for only 2.1% of all companies. Whitin the EU, only Italy and Spain grew their data user companies at a higher rate than the European Union as a whole.

Data Users by Country

Table 16. Data User Companies by Country: 2019–2021 and 2020–2021 Growth

Country	2019	2020	2021	Growth Rate, 2020–2021
Austria	14,200	14,250	14,394	1.0%
Belgium	13,650	14,450	14,680	1.6%
Bulgaria	3,100	3,250	3,318	2.1%
Croatia	1,750	1,800	1,835	1.9%
Cyprus	1,750	1,800	1,835	1.9%
Czechia	6,850	6,950	7,025	1.1%
Denmark	8,350	8,400	8,600	2.4%
Estonia	1,750	1,750	1,791	2.4%
Finland	8,400	8,400	8,546	1.7%
France	75,600	75,800	77,759	2.6%
Germany	114,550	115,000	117,471	2.1%
Greece	11,100	11,250	11,405	1.4%
Hungary	4,650	4,700	4,789	1.9%
Ireland	9,000	9,050	9,184	1.5%
Italy	90,850	92,050	93,822	1.9%
Latvia	1,250	1,250	1,274	1.9%
Lithuania	2,000	2,050	2,089	1.9%
Luxembourg	1,650	1,650	1,682	1.9%
Malta	690	710	723	1.9%
Netherlands	27,050	28,700	29,181	1.7%
Poland	13,600	13,650	13,950	2.2%
Portugal	18,950	19,150	19,382	1.2%
EEA (NO, LI, IS)	6,601	7,089	7,212	1.7%
Romania	5,650	5,650	5,768	2.1%
Slovakia	2,900	2,950	3,007	1.9%
Slovenia	1,300	1,350	1,376	1.9%
Spain	75,650	76,050	77,399	1.8%
Sweden	18,600	20,450	20,759	1.5%
Switzerland	16,214	16,306	16,615	1.9%
United Kingdom	181,050	183,600	188,116	2.5%
EU27	534,840	542,510	553,046	1.9%
EEA (NO, LI, IS) +CH	22,816	23,395	23,827	1.8%
Total, all countries	738,706	749,505	764,989	2.1%

Figure 25. Data User Companies by Country: 2021 and 2020–2021 Growth

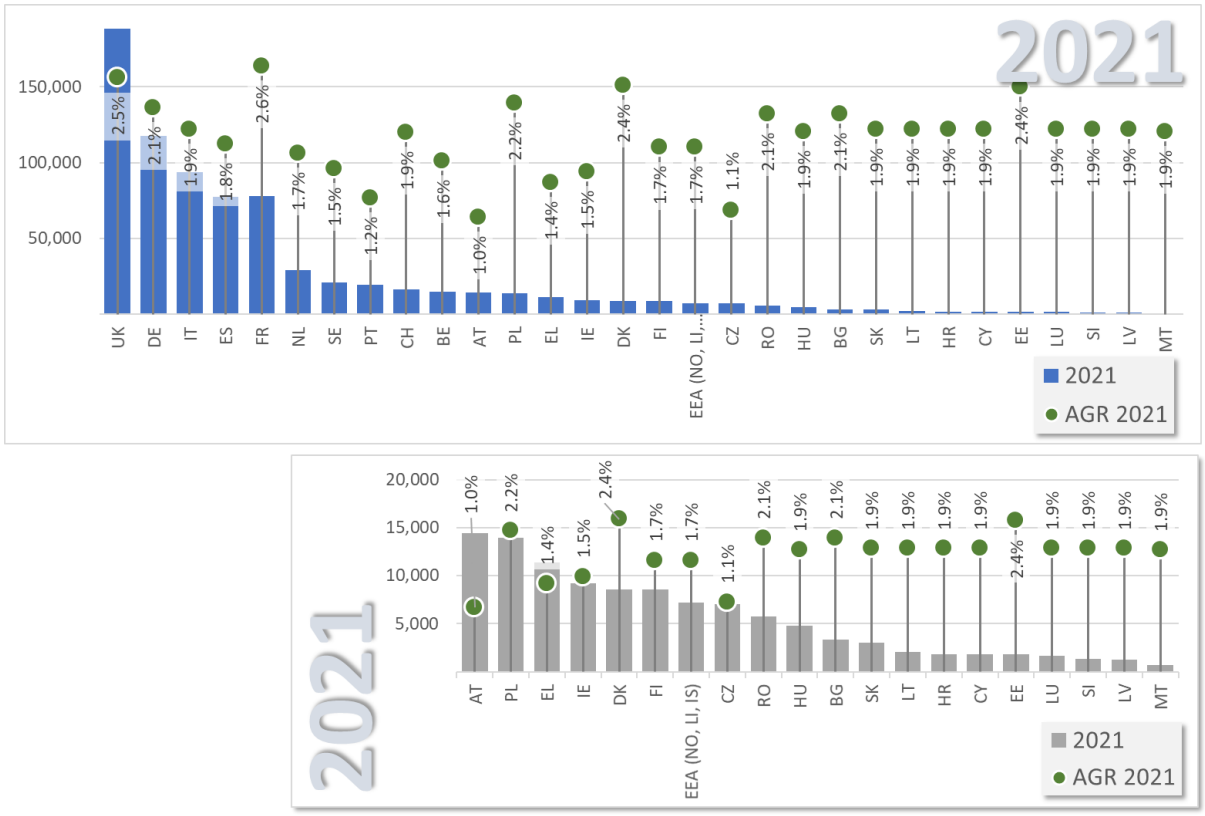
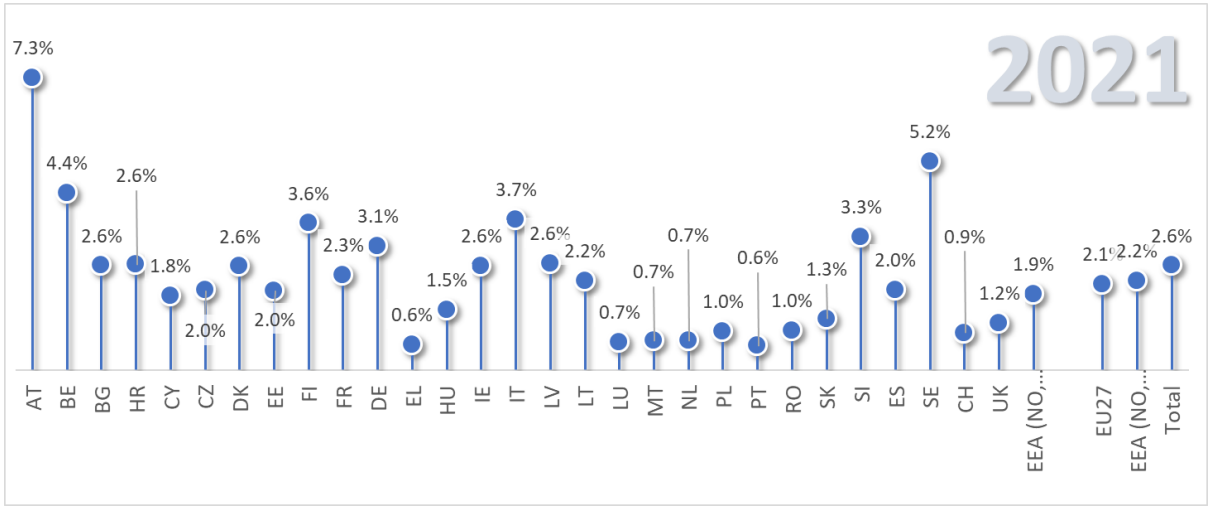


Figure 26. Data Users' Share of All Companies by Country, 2021



Data User Companies by Industry

As with data professionals, the more service-oriented industries show the largest share of data companies, with professional services, manufacturing, and retail & wholesale taking the largest share, with 57% of all data user companies. These companies were better placed following the outbreak of the COVID-19 pandemic to take advantage of the data economy; the crisis simply accelerated the adoption of data within organisations' decision making and operations. As with data suppliers,

agriculture has been added to the industries for data user companies, but it only accounts for 0.8% of all data user companies. Its growth in 2021 was high, but from a small base.

The below table presents the number of data user companies by industry and Figure 27 shows the size and growth of the industries among data user companies in 2021.

Table 17. Data User Companies by Industry: 2019–2021 and 2020–2021 Growth

Industry	2019	2020	2021	Growth Rate 2020–2021
Agriculture	3,167	3,758	4,255	13.2%
Construction	33,957	34,235	34,735	1.5%
Education	11,590	12,154	12,507	2.9%
Finance	30,013	29,960	30,264	1.0%
Health	24,743	25,955	26,705	2.9%
Information & communications	58,405	58,507	59,175	1.1%
Mining & manufacturing	85,448	83,428	83,697	0.3%
Professional services	142,922	146,795	150,743	2.7%
Retail & wholesale	72,853	77,638	80,589	3.8%
Transport	61,467	60,077	60,316	0.4%
Utilities	10,276	10,004	10,060	0.6%

Figure 27. Data User Companies by Industry: 2021 and 2020–2021 Growth

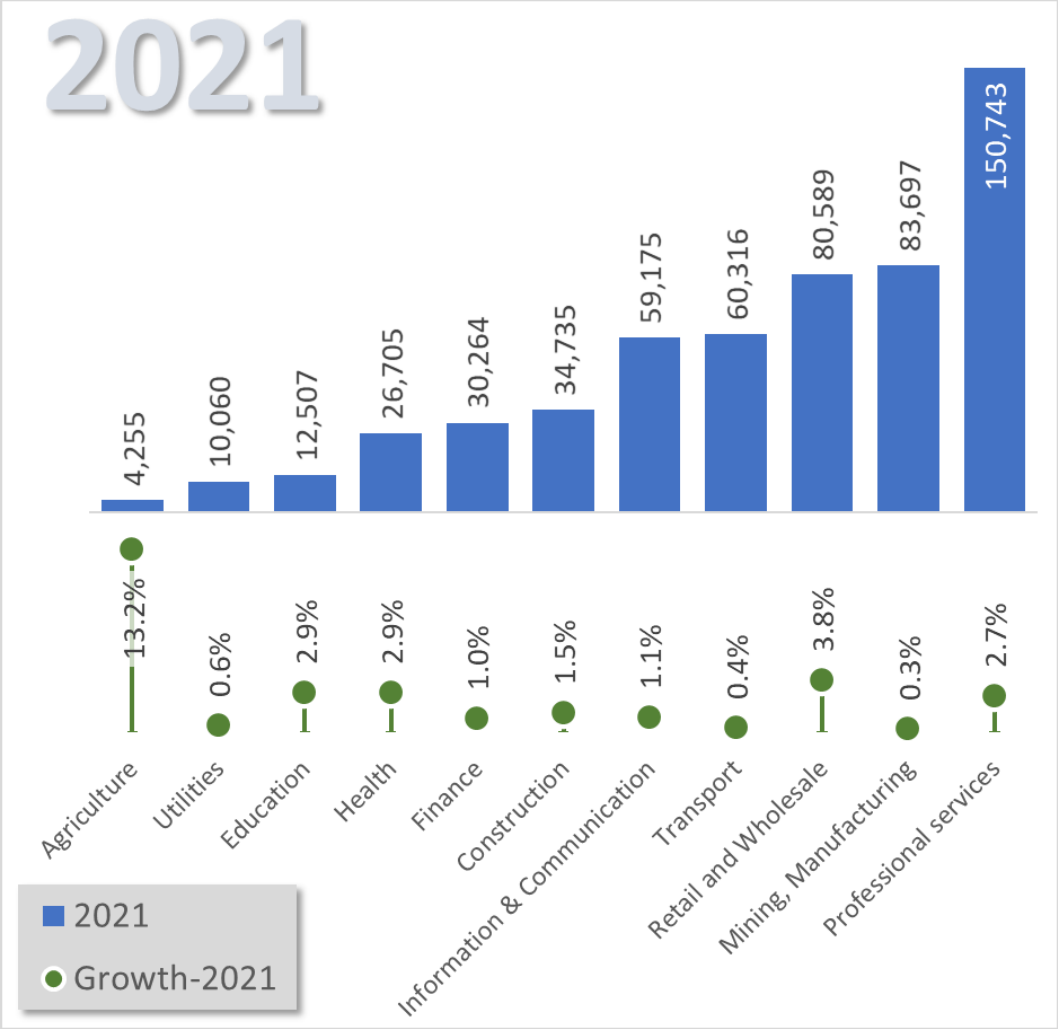
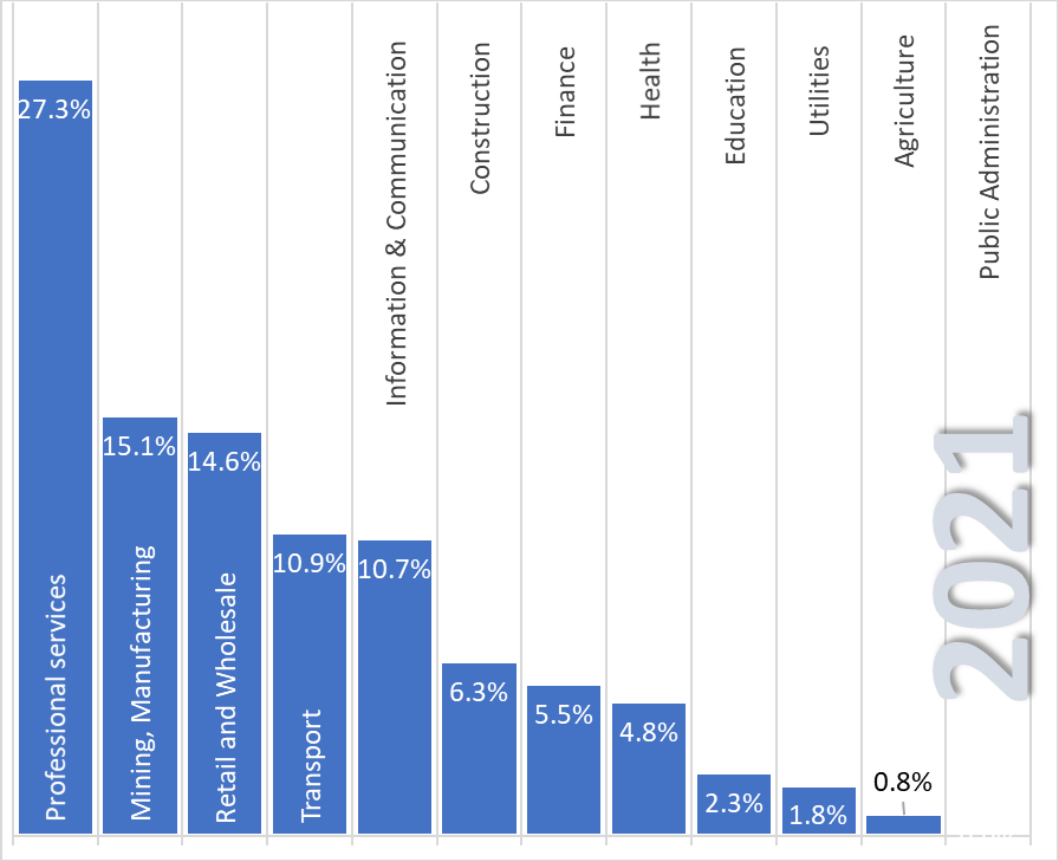


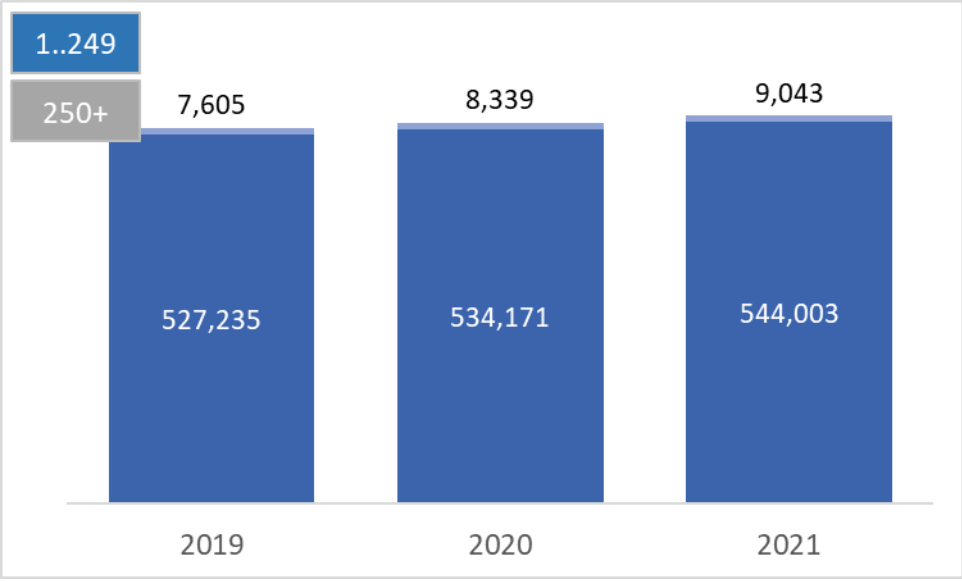
Figure 28. Data User Companies' Share of Total by Industry, 2021



Data User Companies by Size Band

Data users are dominated even more by smaller companies, with companies of 250 or more employees accounting for a mere 1.6% of data user. As with data supplier companies, it is the relative number of companies overall that is reflected in this mix.

Figure 29 Data User Companies by Size Band, 2019–2021



4.5 Forecasting Data Companies

Forecasting Data Supplier Companies, 2025 and 2030 Scenarios

The below table presents the forecast to 2025 and 2030 for the number of data supplier companies in the 27 Member States and on the greater European continent. We anticipate the market for the EU27 will grow at an 8.8% compound growth to 2025 and slow to a 3.1% compound growth to 2030 (Baseline). However, this growth to 2030 could be as low as 2.3% or as high as 4.3%, depending on the scenario. Growth in the number of data suppliers will depend significantly on the rise in the number of companies that monetise data in data markets. This is an early and emerging market, but it shows great promise to 2030 ,as organisations understand and realise the value their data has.

Table 18. Data Supplier Companies Forecast: 2025 and Three 2030 Scenarios

	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR 2020–2025	CAGR '25–30, Challenge	CAGR '25–30, Baseline	CAGR '25–30, High Growth
EU27	252,791	283,084	295,043	311,397	8.8%	2.3%	3.1%	4.3%
EEA (NO, LI, IS) +CH	15,247	17,080	17,964	19,080	8.6%	2.3%	3.3%	4.6%
Total, all countries	505,562	564,976	584,542	614,069	8.7%	2.2%	2.9%	4.0%

Forecasting Data Supplier Companies by Country

Among the five largest Member States , only Spain grows above the average for the EU27 to 2030 (Baseline) (as well as the rest of the EEA and Switzerland). Most of the largest Member States have already established a data supply market and will not grow as much as smaller countries, which can catch up in a market in which they lag slightly – hence, the higher growth for the smaller countries.

Table 19. Data Supplier Companies Forecast by Country: 2025 and Three 2030 Scenarios

Country	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR: 2020–2025	CAGR: '25–'30, Challenge	CAGR: '25–'30, Baseline	CAGR: '25–'30, High Growth
Austria	7,918	8,854	8,996	9,485	8.6%	2.3%	2.6%	3.7%
Belgium	4,691	5,369	5,686	6,040	10.1%	2.7%	3.9%	5.2%
Bulgaria	2,200	2,551	2,679	2,783	9.6%	3.0%	4.0%	4.8%
Croatia	1,294	1,455	1,531	1,579	8.5%	2.4%	3.4%	4.1%
Cyprus	913	1,022	1,062	1,084	9.5%	2.3%	3.1%	3.5%
Czechia	3,797	4,167	4,305	4,626	9.1%	1.9%	2.5%	4.0%
Denmark	6,850	7,613	7,933	8,405	8.4%	2.1%	3.0%	4.2%
Estonia	985	1,110	1,186	1,324	9.5%	2.4%	3.8%	6.1%
Finland	5,281	5,878	5,967	6,325	8.5%	2.2%	2.5%	3.7%
France	22,658	25,278	25,755	27,287	8.0%	2.2%	2.6%	3.8%
Germany	46,555	51,671	53,065	56,176	8.0%	2.1%	2.7%	3.8%
Greece	9,979	11,009	11,644	12,030	8.7%	2.0%	3.1%	3.8%
Hungary	6,478	7,370	7,747	8,223	9.7%	2.6%	3.6%	4.9%
Ireland	6,988	8,179	8,482	9,067	10.5%	3.2%	4.0%	5.3%
Italy	32,723	36,434	37,455	39,124	7.8%	2.2%	2.7%	3.6%
Latvia	1,040	1,164	1,235	1,266	8.9%	2.3%	3.5%	4.0%
Lithuania	1,221	1,368	1,451	1,484	8.7%	2.3%	3.5%	4.0%
Luxembourg	986	1,124	1,177	1,264	9.2%	2.7%	3.6%	5.1%
Malta	434	505	526	559	9.2%	3.1%	3.9%	5.2%
Netherlands	10,973	12,256	12,997	13,738	10.6%	2.2%	3.4%	4.6%
Poland	10,663	12,063	13,090	14,037	8.9%	2.5%	4.2%	5.7%
Portugal	7,306	8,129	8,565	9,022	8.5%	2.2%	3.2%	4.3%
EEA (NO, LI, IS)	7,420	8,376	8,897	9,513	8.9%	2.5%	3.7%	5.1%
Romania	10,218	11,701	12,115	12,807	9.8%	2.7%	3.5%	4.6%
Slovakia	4,321	5,017	5,268	5,685	9.2%	3.0%	4.0%	5.6%
Slovenia	1,040	1,162	1,255	1,284	9.2%	2.3%	3.8%	4.3%
Spain	27,368	30,449	32,217	34,036	8.1%	2.2%	3.3%	4.5%
Sweden	17,911	20,187	21,654	22,658	11.8%	2.4%	3.9%	4.8%
Switzerland	7,827	8,704	9,067	9,566	8.3%	2.1%	3.0%	4.1%
United Kingdom	237,524	264,811	271,536	283,593	8.6%	2.2%	2.7%	3.6%
EU27	252,791	283,084	295,043	311,397	8.8%	2.3%	3.1%	4.3%
EEA (NO, LI, IS) + CH	15,247	17,080	17,964	19,080	8.6%	2.3%	3.3%	4.6%
Total, all countries	505,562	564,976	584,542	614,069	8.7%	2.2%	2.9%	4.0%

Forecasting Data Supplier Companies by Industry

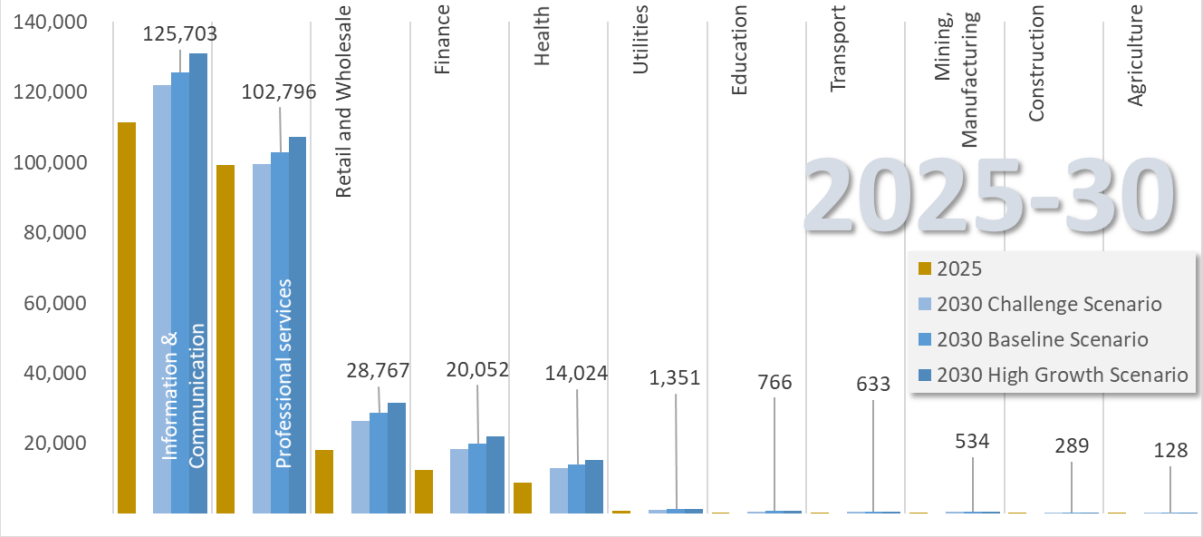
Across the data supplier industries, the key industries of information & communications and professional services – the J and M industries selected in previous reports – continue to dominate,

although their growth will be below the average across all industries. Most of the other industries show growth of around 9–10% to 2030 (Baseline), as can be seen in Table 20, although growing from a small base. Retail & finance organisations in particular are able to make a contribution to the number of data suppliers, as these are industries that have embraced the data industry and have data sets that are more amenable for sale across markets. However, public services such as health and education can also grow, as much of the data available is public domain data or government owned. As of today, this data is more likely to be made available for free though, undermining the potential for data monetisation.

Table 20. Data Supplier Companies Forecast by Industry: 2025 and Three 2030 Scenarios

	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR 2020–2025	CAGR '25–30, Challenge	CAGR '25–30, Baseline	CAGR '25–30, High Growth
Agriculture	80	118	128	140	42.1%	8.0%	9.8%	11.8%
Construction	182	267	289	317	45.4%	8.0%	9.7%	11.8%
Education	480	705	766	842	35.3%	8.0%	9.8%	11.9%
Finance	12,597	18,480	20,052	22,007	30.1%	8.0%	9.7%	11.8%
Health	8,841	12,959	14,024	15,397	30.7%	7.9%	9.7%	11.7%
Information & communications	111,539	122,013	125,703	131,126	3.9%	1.8%	2.4%	3.3%
Mining & manufacturing	336	493	534	586	92.6%	8.0%	9.7%	11.7%
Professional services	99,383	99,677	102,796	107,245	9.6%	0.1%	0.7%	1.5%
Retail & wholesale	18,099	26,538	28,767	31,557	29.9%	8.0%	9.7%	11.8%
Transport	398	584	633	695	34.8%	8.0%	9.8%	11.8%
Utilities	855	1,251	1,351	1,484	35.2%	7.9%	9.6%	11.7%

Figure 30. Data Supplier Companies Forecast: 2025 and Three 2030 Scenarios



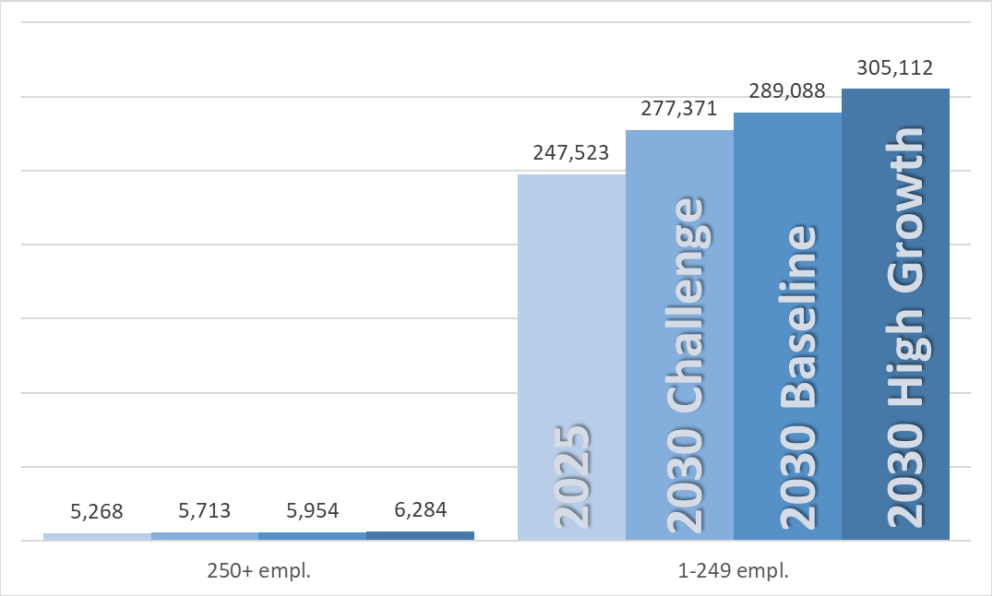
Forecasting Data Supplier Companies by Company Size

The number of data suppliers by company size behaves as expected, with the majority of companies remaining in the 1–249 employee size band. The relative sizes of the two bands means little will change, even though the larger size band shows below average growth as smaller companies catch up with the monetisation of data.

Table 21. Data Supplier Companies Forecast by Company Size Band: 2025 and Three 2030 Scenarios

	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR 2020–2025	CAGR '25–30, Challenge	CAGR '25–30, Baseline	CAGR '25–30, High Growth
250+ employees.	5,268	5,713	5,954	6,284	3.7%	1.6%	2.5%	3.6%
1–249 employees	247,523	277,371	289,088	305,112	8.9%	2.3%	3.2%	4.3%

Figure 31. Data Supplier Companies Forecast by Company Size Band: 2025 and Three 2030 Scenarios



Forecasting Data User Companies: 2025 and 2030 Scenarios

Data user companies show higher growth over the period of the forecast when compared with data supplier companies as the data economy begins to drive its way into all businesses. Many more companies will embrace data as a basis for business, but there will still be latency from those organisations that use data to fundamentally change the way business is conducted. Most will use data to reduce cost or raise efficiency, although there are signs that more companies are using data to develop new products and to change how business is conducted. Those companies that can fundamentally change how business is conducted are more likely to be successful, although often it is the pioneers that are overtaken by others.

The below table shows the forecast for data user companies for 2025 and the three scenarios for 2030.

Table 22. Data User Companies Forecast: 2025 and Three 2030 Scenarios

	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR: 2020–2025	CAGR: '25–'30, Challenge	CAGR: '25–'30, Baseline	CAGR: '25–'30, High Growth
EU27	633,359	753,920	898,220	1,086,306	3.1%	3.5%	7.2%	11.4%
EEA (NO, LI, IS) +CH	27,174	32,334	38,832	47,197	3.0%	3.5%	7.4%	11.7%
Total, all countries	875,394	1,041,537	1,237,203	1,492,761	3.2%	3.5%	7.2%	11.3%

Forecasting Data User Companies by Country

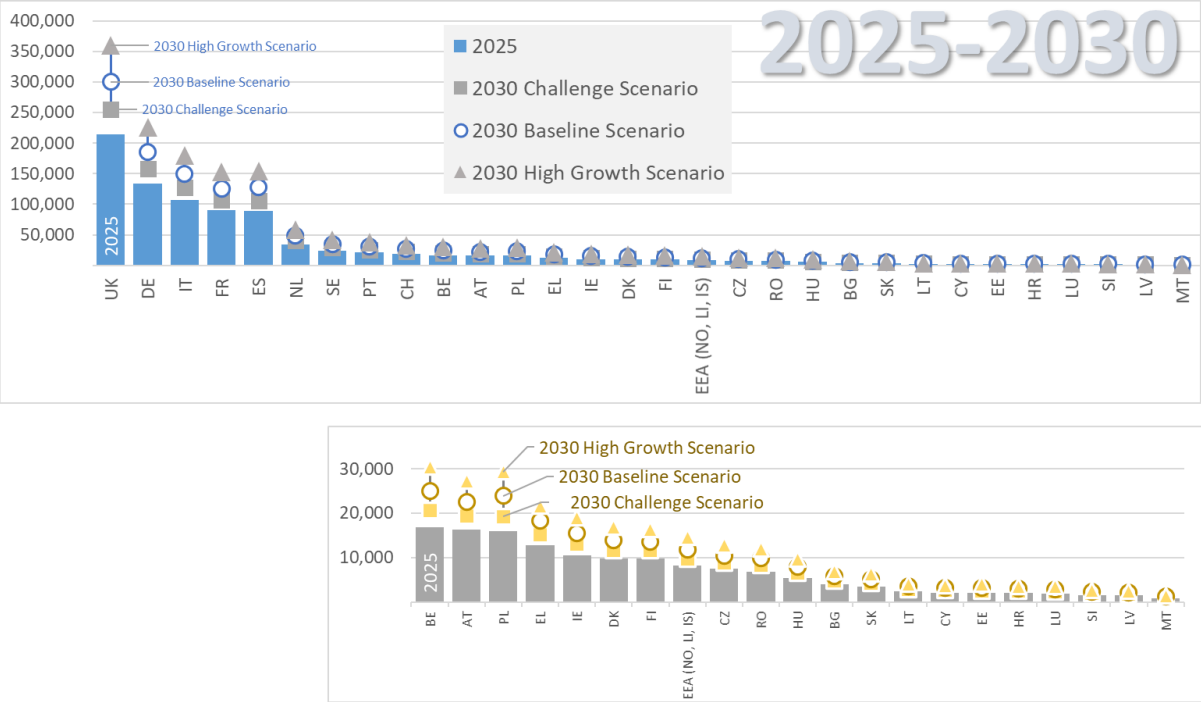
The below table shows the forecast for data user companies by Member State for 2025 and the three scenarios for 2030. We expect the numbers of these companies in the EU27 Member States to grow by 3.1% per year from 2020 to 2025 and by 7.2% per year from 2025 to 2030 (Baseline). Among the Member States, Poland shows the highest growth, at 8.4% for the Baseline forecast, followed by Bulgaria and Slovakia, both at 8.3%. There is little difference across the Member States, as we expect all states to transition into data economies, led by the companies.

Table 23. Data User Companies Forecast by Country: 2025 and Three 2030 Scenarios

Country	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR: 2020–2025	CAGR: '25–'30, Challenge	CAGR '25–'30, Baseline	CAGR: '25–'30, High Growth
Austria	16,267	19,384	22,584	27,289	2.7%	3.6%	6.8%	10.9%
Belgium	16,919	20,634	25,057	30,504	3.2%	4.0%	8.2%	12.5%
Bulgaria	3,906	4,827	5,811	6,918	3.7%	4.3%	8.3%	12.1%
Croatia	2,126	2,548	3,074	3,633	3.4%	3.7%	7.6%	11.3%
Cyprus	2,126	2,616	3,118	3,644	3.4%	4.2%	8.0%	11.4%
Czechia	7,511	8,784	10,406	12,815	1.6%	3.2%	6.7%	11.3%
Denmark	9,869	11,688	13,966	16,957	3.3%	3.4%	7.2%	11.4%
Estonia	2,122	2,549	3,124	3,994	3.9%	3.7%	8.0%	13.5%
Finland	9,824	11,653	13,565	16,478	3.2%	3.5%	6.7%	10.9%
France	90,432	107,518	125,610	152,517	3.6%	3.5%	6.8%	11.0%
Germany	133,543	157,957	186,006	225,667	3.0%	3.4%	6.9%	11.1%
Greece	12,862	15,121	18,339	21,714	2.7%	3.3%	7.4%	11.0%
Hungary	5,471	6,634	7,996	9,726	3.1%	3.9%	7.9%	12.2%
Ireland	10,473	13,064	15,535	19,030	3.0%	4.5%	8.2%	12.7%
Italy	107,408	127,445	150,231	179,844	3.1%	3.5%	6.9%	10.9%
Latvia	1,477	1,762	2,143	2,518	3.4%	3.6%	7.7%	11.3%
Lithuania	2,422	2,891	3,518	4,121	3.4%	3.6%	7.8%	11.2%
Luxembourg	1,949	2,369	2,843	3,499	3.4%	4.0%	7.8%	12.4%
Malta	827	1,026	1,224	1,493	3.1%	4.4%	8.2%	12.6%
Netherlands	33,657	40,062	48,711	59,010	3.2%	3.5%	7.7%	11.9%
Poland	15,961	19,242	23,942	29,425	3.2%	3.8%	8.4%	13.0%
Portugal	21,746	25,784	31,149	37,605	2.6%	3.5%	7.5%	11.6%
EEA (NO, LI, IS)	8,151	9,792	11,905	14,639	2.8%	3.7%	7.9%	12.4%
Romania	6,791	8,287	9,839	11,919	3.7%	4.1%	7.7%	11.9%
Slovakia	3,485	4,313	5,193	6,422	3.4%	4.4%	8.3%	13.0%
Slovenia	1,595	1,900	2,352	2,757	3.4%	3.6%	8.1%	11.6%
Spain	88,835	105,332	127,792	154,721	3.2%	3.5%	7.5%	11.7%
Sweden	23,754	28,532	35,093	42,084	3.0%	3.7%	8.1%	12.1%
Switzerland	19,022	22,542	26,927	32,559	3.1%	3.5%	7.2%	11.3%
United Kingdom	214,861	255,283	300,151	359,258	3.2%	3.5%	6.9%	10.8%
EU27	633,359	753,920	898,220	1,086,306	3.1%	3.5%	7.2%	11.4%

Country	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR: 2020–2025	CAGR: '25–'30, Challenge	CAGR '25–'30, Baseline	CAGR: '25–'30, High Growth
EEA (NO, LI, IS) + CH	27,174	32,334	38,832	47,197	3.0%	3.5%	7.4%	11.7%
Total, all countries	875,394	1,041,537	1,237,203	1,492,761	3.2%	3.5%	7.2%	11.3%

Figure 32. Data User Companies Forecast by Country: 2025 and Three 2030 Scenarios



Forecasting Data User Companies by Industry

Data user companies by industries behave in a similar way to data supplier companies by industry, with those in more established data service-oriented industries already well on the path towards contributing to the data economy. Professional services and retail & wholesale are the largest data user company industries, and these two alone will account for 47% of all data user companies in 2030 (Baseline). Growth across industries is reasonably flat, although agriculture shows the highest growth to 2030 (Baseline), followed by retail & wholesale, health, and education.

Table 24. Data User Companies Forecast by Industry: 2025 and Three 2030 Scenarios

Industry	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR: 2020–2025	CAGR: '25–'30, Challenge	CAGR '25–'30, Baseline	CAGR: '25–'30, High Growth
Agriculture	11,396	14,777	17,567	21,319	24.8%	5.3%	9.0%	13.3%
Construction	38,591	45,595	54,089	65,371	2.4%	3.4%	7.0%	11.1%
Education	16,394	19,904	23,809	28,724	6.2%	4.0%	7.7%	11.9%
Finance	31,590	36,988	44,112	53,416	1.1%	3.2%	6.9%	11.1%
Health	34,199	41,990	49,787	60,382	5.7%	4.2%	7.8%	12.0%
Information & communications	62,495	73,761	87,631	106,291	1.3%	3.4%	7.0%	11.2%
Mining & manufacturing	79,013	89,702	107,117	130,045	-1.1%	2.6%	6.3%	10.5%

Industry	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR: 2020–2025	CAGR: '25–'30, Challenge	CAGR '25–'30, Baseline	CAGR: '25–'30, High Growth
Professional services	183,430	218,740	261,404	315,110	4.6%	3.6%	7.3%	11.4%
Retail & wholesale	110,387	137,633	163,510	197,575	7.3%	4.5%	8.2%	12.3%
Transport	56,191	63,777	75,992	92,093	-1.3%	2.6%	6.2%	10.4%
Utilities	9,674	11,052	13,202	15,978	-0.7%	2.7%	6.4%	10.6%

Figure 33. Data User Companies by Industry Share of Total: 2025 and Three 2030 Scenarios

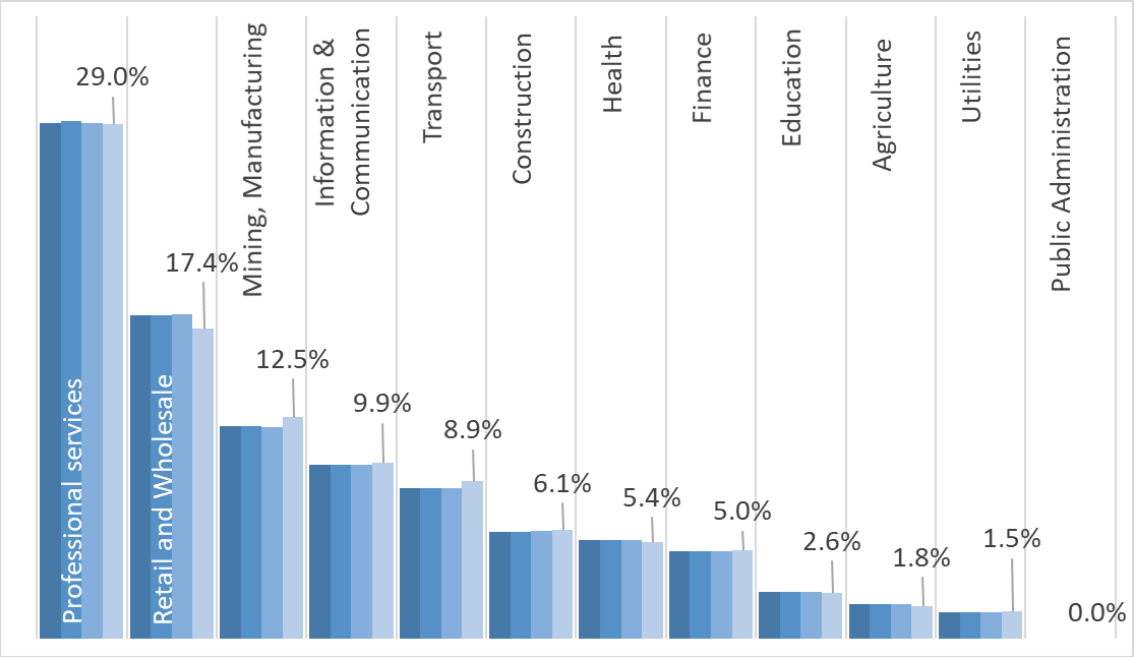
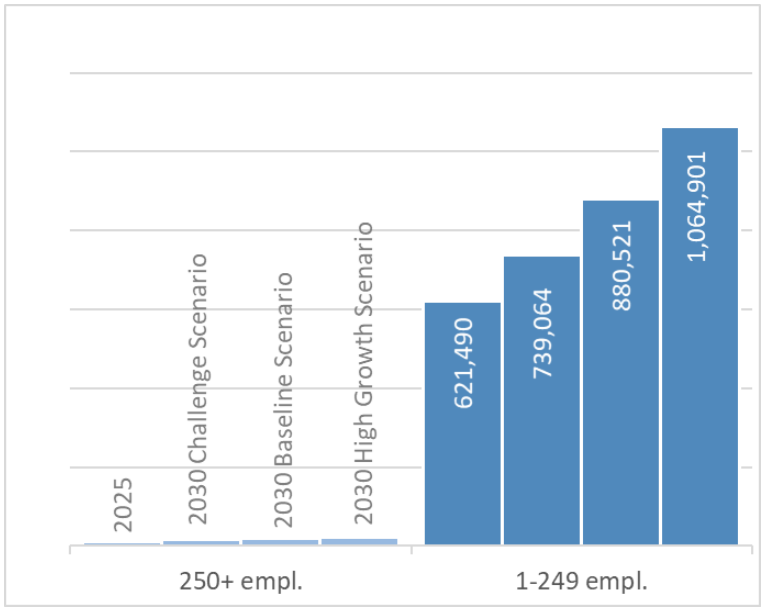


Table 25. Data User Companies Forecast by Company Size Band: 2025 and Three 2030 Scenarios

	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR: 2020–2025	CAGR: '25–'30, Challenge	CAGR: '25–'30, Baseline	CAGR: '25–'30, High Growth
250+ employees	11,869	14,856	17,699	21,405	7.3%	4.6%	8.3%	12.5%
1–249 employees	621,490	739,064	880,521	1,064,901	3.1%	3.5%	7.2%	11.4%

Figure 34. Data User Companies Forecast by Company Size Band: 2025 and Three 2030 Scenarios



4.6 Key Findings

By 2025, data user companies will account for 633,000 units – growing from 2020 at a compound rate of 3.1%. This is about the same as the growth in the number of all companies. But, in the longer-term (2025–2030 Baseline), data user companies continue grow at a compound rate that is between 5 and 10 times that of all companies. The total number of companies is expected to increase long term at between 4% and 8% per year between 2025 and 2030, depending on the scenario.

The data user penetration rates (i.e. the number of data user companies as a proportion of total companies) vary across Member States, with the larger economies dominating the scene in terms of size but not overall penetration. It is the smaller Member States that tend to show a higher share of data users per total companies. The EU27 average penetration rate is 2.1% in 2021.

Data user penetration is as low as 0.6% for Slovakia and Poland, and as high as 7.3% for Austria. Adoption rates of data technologies are higher in industries such as professional services, retail, and financial services.

Looking at the share by industry, professional services is easily the sector with the greatest number of data user companies, accounting for 27% of all data user companies in 2021. Data is core to a large share of professional services activities. Manufacturing and retail & wholesale follow professional services, with 15.1% and 14.6%, respectively, but are significantly behind the lead industry.

Small and medium-sized companies account for 98% of all companies in the European Union in 2021. Unsurprisingly, they also account for the majority of data user companies. Larger companies, however, are those that invest more heavily in digital resources. Investment in data technologies requires considerable investment and expertise, and companies with fewer than 10 employees – a large proportion small and medium businesses – are unable or unwilling to make this investment. **Larger companies spend close to 150 times the expenditure of smaller companies on data among the Member States**, although this will drop to "only" 100 times by 2030.

5. Data Companies' Revenues

5.1 Definition

Data companies' revenues *correspond to the aggregated value of all the data-related products and services generated by Europe-based data suppliers, including exports outside of the EU. This indicator measures the revenues of the data suppliers identified and classified by Indicator 2 (see the products and services specified in our definition of the data market). Data companies' revenues do not include data monetisation as part of the data market.*

The overall value of data revenues is very close, but not identical, to the overall value of the data market for the following reasons:

- The value of the data market corresponds to the aggregated value of all the data-related products and services bought by European users (demand), including imports from foreign suppliers.
- The value of revenues corresponds to the aggregated value of all the revenues generated by Europe-based enterprises (supply) through the production, distribution, and sale of data-related products and services, including exports outside of the EU.

Table 26. Main Data Sources for Data Companies' Revenues

Data Source	Updated
Consensus forecasts – Consensus Economics	July–2021
Eurostat chain-linked volumes (GDP)	Aug 2021
IDC Core IT Spending Guide 2H2021	Jun 2021
IDC Big Data Spending Guide 2H2021	Sep 2021
IDC Worldwide Black Book v3.2 (standard edition)	Sep 2021
IMF World Economic Outlook	Aug 2021

5.2 Measuring Data Companies' Revenues

Data companies' revenues grew by 2.9% in 2021, adding nearly €2.1 million. In 2020, data companies' revenues grew faster than all company revenues, as this is an emerging market, and general mitigating factors such as Brexit and COVID had little impact on growth, with organisations worldwide embracing remote working and online business during this period.

Table 27, Data Companies' Revenues and Share, 2019–2021

N.	Region	Name	Description	2019	2020	2021	Growth 2020–2021
3.1	EU27	Total revenues of data companies in the EU	Total revenues of the data suppliers calculated by Indicator 2	64,262	71,050	73,116	2.9%
3.2	EU27	Share of data companies' revenues	Ratio of data suppliers' revenues to total companies' revenues – NACE II industries A, C, E, G, H, J, K, M, P, and Q	0.3%	0.3%		

5.3 Data Companies' Revenues by Country

The data market is the single biggest influence on the size of data companies' revenues, although we now estimate data monetisation as well as data revenues. However, monetisation is not included in data companies' revenues for consistency with prior reports.

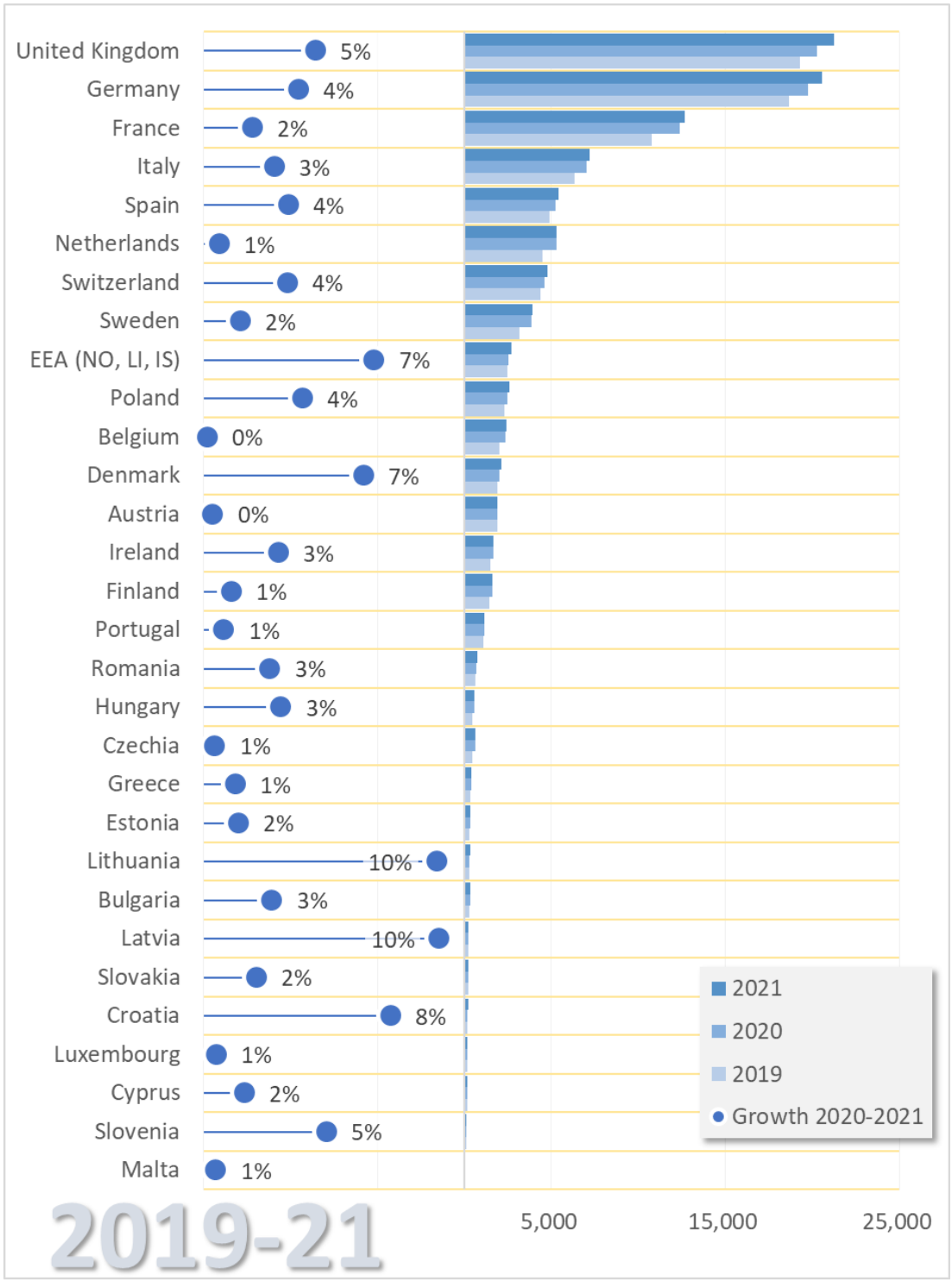
In the European Union, Germany showed the highest revenues, with 28.1% of EU27 revenues. Those countries that were able to adapt more easily to COVID 19 were the ones that grew the most in 2021 and are expected to continue this growth. The following table shows the revenues by country for 2019, 2020, and 2021, along with growth for 2020–2021.

The mix of industries that represent each country will impact future growth in revenues, with those countries that have a stronger representation of data-oriented businesses such as information & communications and professional services – strong suppliers of data tools and services – most likely to see rising revenues associated with the data market.

Table 28. Data Companies' Revenues by Country: 2019–2021 and 2020–2021 Growth

Countries	2019	2020	2021	Growth Rate 2020–2021
Austria	1,878	1,898	1,906	0.4%
Belgium	2,031	2,381	2,385	0.2%
Bulgaria	284	320	330	3.0%
Croatia	164	188	203	8.1%
Cyprus	148	178	181	1.8%
Czechia	467	641	644	0.5%
Denmark	1,886	1,992	2,130	6.9%
Estonia	308	351	356	1.5%
Finland	1,445	1,599	1,619	1.2%
France	10,783	12,398	12,660	2.1%
Germany	18,629	19,748	20,562	4.1%
Greece	362	400	405	1.4%
Hungary	481	543	561	3.4%
Ireland	1,505	1,636	1,689	3.2%
Italy	6,349	6,999	7,214	3.1%
Latvia	205	213	235	10.2%
Lithuania	289	290	320	10.1%
Luxembourg	153	185	186	0.6%
Malta	57	70	70	0.6%
Netherlands	4,512	5,266	5,304	0.7%
Poland	2,302	2,463	2,569	4.3%
Portugal	1,062	1,155	1,166	0.9%
EEA (NO, LI, IS)	2,464	2,499	2,684	7.4%
Romania	601	698	719	2.9%
Slovakia	200	231	237	2.3%
Slovenia	107	121	128	5.3%
Spain	4,875	5,233	5,427	3.7%
Sweden	3,179	3,851	3,913	1.6%
Switzerland	4,383	4,624	4,791	3.6%
United Kingdom	19,283	20,268	21,250	4.8%
EU27	64,262	71,050	73,116	2.9%
EEA (NO, LI,IS)+ CH	6,847	7,123	7,475	4.9%
Total, all countries	90,392	98,441	101,842	3.5%

Figure 35. Data Revenues by Country: 2019–2021 with Growth

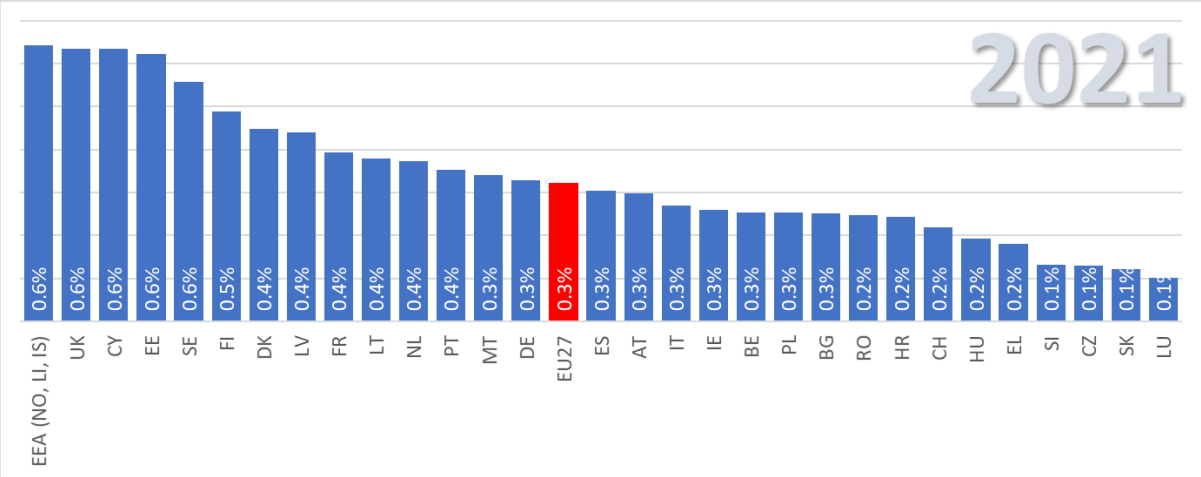


5.4 Shares of All Companies' Revenues by Country

The share of all companies' revenues accounted for by data companies drops significantly when compared with prior reports because the list of industries used to define data supplier companies has expanded considerably to accommodate potential future sales of data and the expanded roles of

companies in providing data services and tools across a wider range of industries. The industries included in the definition of the data market include the NACE II segments of A, C, E, G, H, J, K, M, P, and Q. These are: agriculture, manufacturing, utilities, retail & wholesale, transportation, information & communications, finance, professional services, education, and healthcare. The shares across the Member States are wide ranging, as those Member States with a stronger focus on industries with greater preponderance for data, such as information & communications and professional services, are those with the greater share of all companies. These shares reflect a historical focus, too, as Member States such as Italy have a stronger manufacturing and retail focus; Germany's strong manufacturing base is also reflected in the shares.

Figure 36 Data Companies' Revenue Share, 2021 – All Companies in A, C, E, G, H, J, K, M, P, and Q



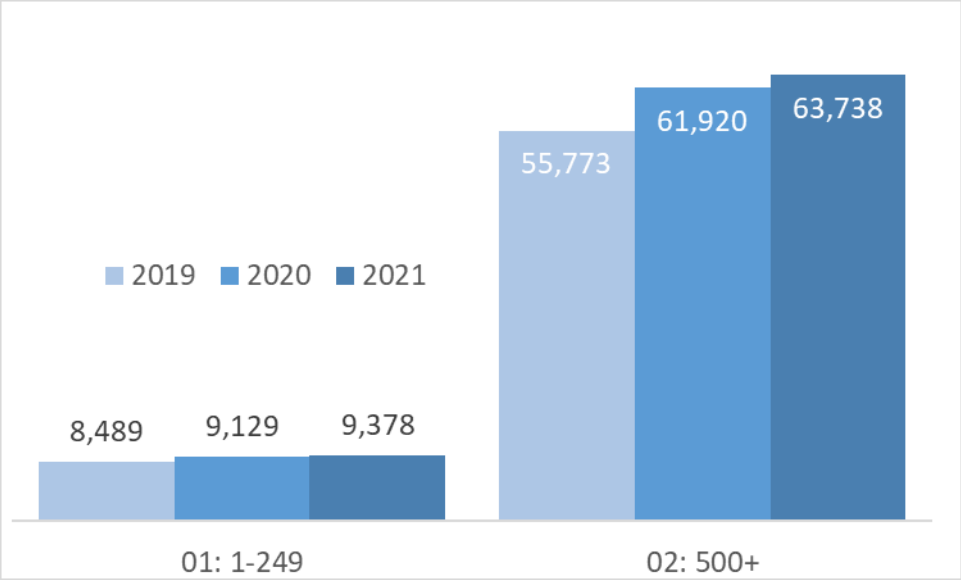
5.5 Data Companies' Revenues by Company Size

Unlike the number of data companies, data revenues are associated more with larger companies, as these have a greater focus on data, can supply the data tools and services, and have the expertise to develop and use data within the organisation. So, while the number of data companies with 250 or more employees only accounts for 2.4% of all data supplier companies, they account for 87% of data companies' revenues.

Table 29. Data Companies' Revenues by Company Size Band, 2019–2021

	2019	2020	2021	Growth Rate 2020–2021
01: 1–249	8,489	9,129	9,378	2.7%
02: 500+	55,773	61,920	63,738	2.9%

Figure 37. Data Companies' Revenues by Company Size Band, 2019–2021



5.6 Data Companies Revenue Forecast

Data companies' revenues will continue to grow as the market evolves, as detailed in the table below. The overall compound growth among the Member States will be 3.4% to 2030 (Baseline), although growth will be higher from 2020 to 2025, at a compound rate of 7.9%. The data market is healthy and robust but will slow to some extent as organisations consolidate their spending following the heavy investments that were poured into the data market in the early years. The key industries that benefit from spending in the data market are – as always – information & communications and professional services.

Table 30. Data Companies Revenue Forecast: 2025, Three 2030 Scenarios (€M), and Compound Growth (%)

	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR 2020–2025	CAGR 2025–2030, Challenge	CAGR 2025–2030, Baseline	CAGR 2025–2030, High Growth
EU27	104,086	108,964	123,294	152,372	7.9%	0.9%	3.4%	7.9%
EEA (NO, LI, IS) + CH	9,483	10,251	12,958	15,194	5.9%	1.6%	6.4%	9.9%
Total, all countries	140,015	152,625	174,987	213,405	7.3%	1.7%	4.6%	8.8%

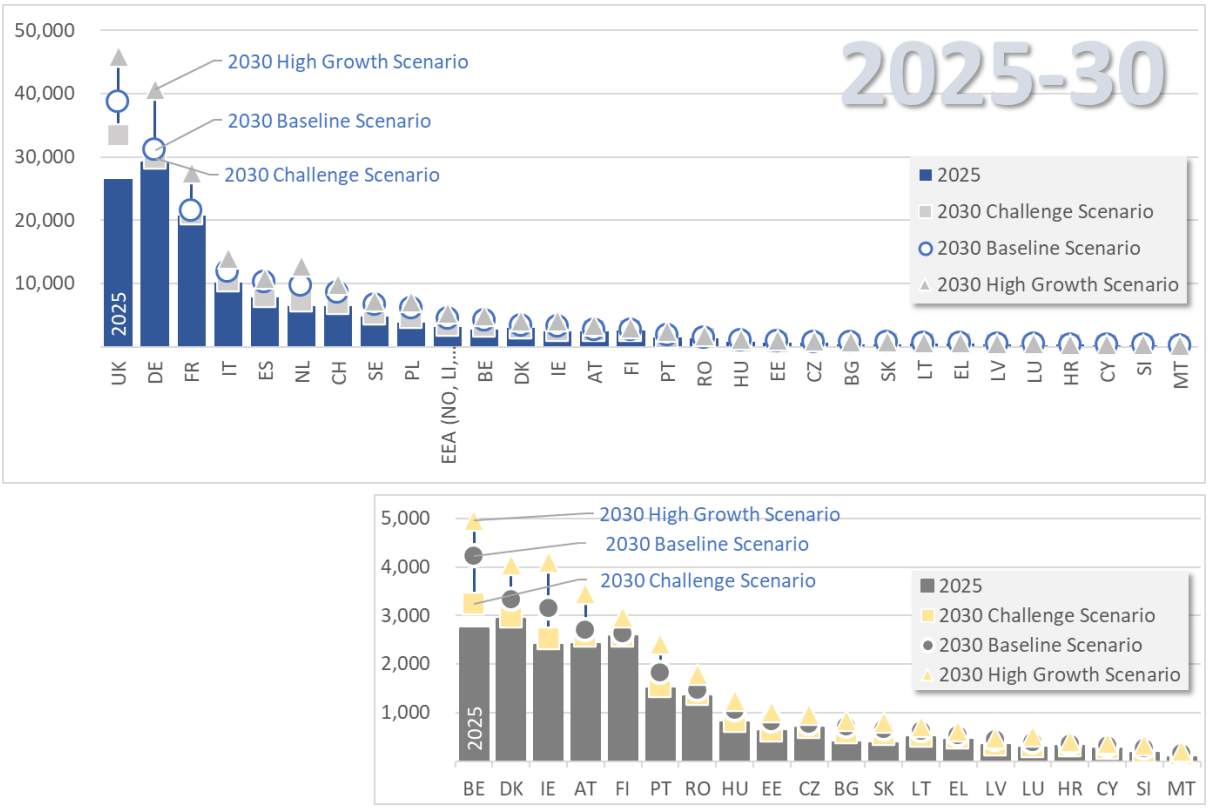
Data Companies' Revenues by Country

Among the Member States, Germany, while easily the largest recipient of data revenues, shows lower growth to 2030 (Baseline), as the country focuses more on manufacturing than on information technology and professional services. The below tableTable 31 presents for each country on the continent the revenues for 2025 and the forecasts, including growth rates, for the three scenarios for 2030.

Table 31. Data Companies Revenue Forecast: 2025, Three 2030 Scenarios, and Compound Growth (%)

Country	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR: 2020–2025	CAGR: 2025–2030, Challenge	CAGR: 2025–2030, Baseline	CAGR: 2025–2030, High Growth
Austria	2,411	2,579	2,703	3,449	4.9%	1.4%	2.3%	7.4%
Belgium	2,732	3,247	4,221	4,970	2.8%	3.5%	9.1%	12.7%
Bulgaria	380	578	712	836	3.5%	8.8%	13.4%	17.1%
Croatia	320	322	359	411	11.2%	0.1%	2.3%	5.1%
Cyprus	264	267	308	368	8.2%	0.3%	3.2%	6.9%
Czechia	687	696	762	953	1.4%	0.2%	2.1%	6.8%
Denmark	2,929	2,975	3,334	4,050	8.0%	0.3%	2.6%	6.7%
Estonia	613	622	809	1,014	11.8%	0.3%	5.7%	10.6%
Finland	2,566	2,582	2,631	2,968	9.9%	0.1%	0.5%	3.0%
France	20,790	21,061	21,475	27,403	10.9%	0.3%	0.7%	5.7%
Germany	29,271	29,807	31,055	40,718	8.2%	0.4%	1.2%	6.8%
Greece	447	481	536	627	2.3%	1.5%	3.7%	7.0%
Hungary	800	819	1,053	1,240	8.1%	0.5%	5.6%	9.2%
Ireland	2,390	2,521	3,152	4,102	7.9%	1.1%	5.7%	11.4%
Italy	10,124	10,379	11,870	14,009	7.7%	0.5%	3.2%	6.7%
Latvia	338	341	434	485	9.7%	0.2%	5.1%	7.5%
Lithuania	491	508	611	700	11.1%	0.7%	4.4%	7.4%
Luxembourg	273	302	383	509	8.1%	2.1%	7.0%	13.3%
Malta	83	130	161	204	3.5%	9.3%	14.1%	19.6%
Netherlands	6,442	7,319	9,664	12,804	4.1%	2.6%	8.4%	14.7%
Poland	3,797	4,625	6,063	7,114	9.0%	4.0%	9.8%	13.4%
Portugal	1,505	1,526	1,815	2,409	5.4%	0.3%	3.8%	9.9%
EEA (NO, LI, IS)	3,076	3,358	4,381	5,285	4.2%	1.8%	7.3%	11.4%
Romania	1,330	1,367	1,456	1,786	13.8%	0.5%	1.8%	6.1%
Slovakia	366	542	655	797	9.6%	8.2%	12.3%	16.8%
Slovenia	176	195	261	333	7.7%	2.1%	8.2%	13.6%
Spain	7,741	7,963	10,176	10,813	8.1%	0.6%	5.6%	6.9%
Sweden	4,816	5,209	6,635	7,299	4.6%	1.6%	6.6%	8.7%
Switzerland	6,407	6,893	8,577	9,908	6.7%	1.5%	6.0%	9.1%
United Kingdom	26,446	33,410	38,735	45,839	5.5%	4.8%	7.9%	11.6%
EU27	104,086	108,964	123,294	152,372	7.9%	0.9%	3.4%	7.9%
EEA (NO, LI, IS) + CH	9,483	10,251	12,958	15,194	5.9%	1.6%	6.4%	9.9%
Total, all countries	140,015	152,625	174,987	213,405	10.9%	1.7%	4.6%	8.8%

Figure 38. Data Companies Revenue Forecast by Member State: 2025 and Three 2030 Scenarios



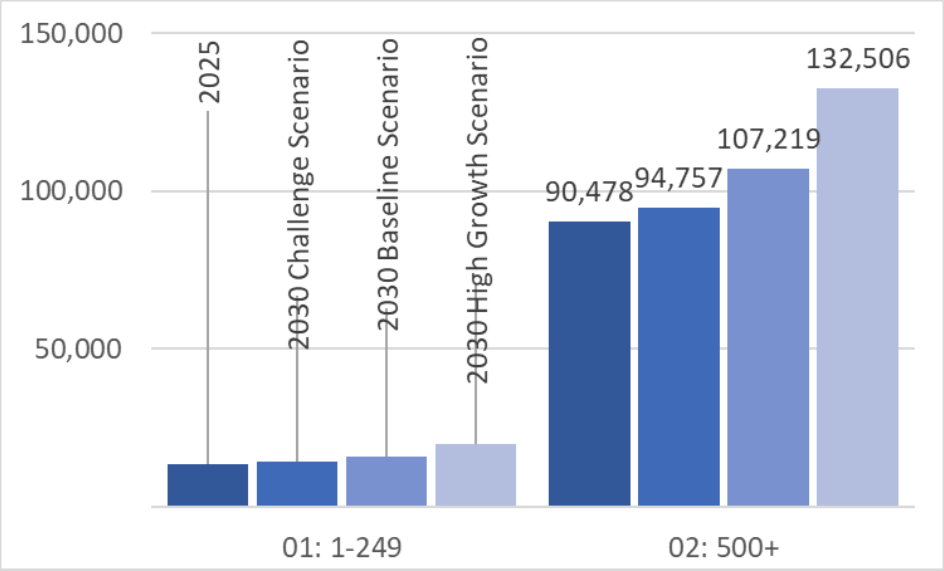
Data Companies' Revenues by Company Size

Data companies' revenues are expected to continue to be higher for larger companies, with those companies with 250 or more employees accounting for 87% of revenues in 2030 (Baseline). Figure 39 shows the relative size of revenues by company size band for the three scenarios for 2030. The compound growth for both small and larger companies' revenues is similar over the forecast period.

Table 32. Data Companies Revenue Forecast by Size Band: 2025 and Three 2030 Scenarios

	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR: 2020–2025	CAGR: 2025–2030, Challenge	CAGR: 2025–2030, Baseline	CAGR: 2025–2030, High Growth
01: 1–249	13,608	14,207	16,075	19,866	8.3%	0.9%	3.4%	7.9%
02: 500+	90,478	94,757	107,219	132,506	7.9%	0.9%	3.5%	7.9%

Figure 39. Data Companies' Revenue Forecast by Size Band: 2025 and Three 2030 Scenarios



5.6 Key Findings

Revenues generated by data suppliers have registered a constant increase over recent years to reach nearly €71 billion in the EU27 in 2020. Data companies' revenues account for 0.3% of total company revenues in 2020. This seems significantly lower than previously reported, but this is because the total for data revenues is significantly expanded, from including only the NACE II categories of J and M, to including A, C, E, G, H, J, K, M, P, and Q – a significant increase that includes manufacturing, hence the apparent dramatic drop in share taken by data revenues of total data revenues. The spread by Member State of data revenues as a share of total revenues is consistent with previous reports and reflects the differing industry focuses shown by Member States. Those with a greater focus on data intensive industries such as finance and professional services will show a greater share of total revenues associated with data.

Larger companies show greater revenues, but this is again not a surprise. Being active in the data market as a data supplier requires significant expertise in data products and services, and these tend to defer to larger organisations. Smaller data user companies too are less inclined to invest in data tools and services because of the cost and expertise needed.

The concept of data revenue is different from the one of data market, but the imports and exports of data-driven products and services are expected to track each over time.

6. The Data Market

6.1 Definition

*The **data market** is the marketplace where digital data is exchanged as "products" and "services" as a result of the elaboration of raw data.*

The data market captures the aggregate value of the demand of digital data without measuring the direct, indirect, and induced impacts of data in the economy as a whole. Further, the data market represents a wider concept than the market of big data & analytics (BDA), as it includes not only the value generated by pure data players developing BDA technologies, but also the value created by data-related research, businesses, information, and IT services. The digital data exchanged as "products" and "services" in the data market refers exclusively to data that is collected, processed, stored, and transmitted over digital information infrastructures and/or elaborated with digital technologies. This definition includes multimedia objects that are collected, stored, processed, elaborated, and delivered for exploitation through digital technologies (for example, image databases). The value of the Data market is not exactly equal to the aggregated revenues of the European data companies because it includes imports (data products and services bought on the global digital market from suppliers not based in Europe) and excludes the exports of the European data companies.

In this report, we have added to the data market an estimate for the value of data monetisation. In previous reports, we were not able to estimate this market because the data to enable us to complete this estimate was sparse and unreliable. In our latest survey, we asked the respondents about data monetisation, as well as conducting informal research among data companies, resulting in an initial estimate for the size of data monetisation, which is reported here for the first time.

6.2 Measuring the Data Market

Table 33. The Data Market – Main Data Sources

Data Source	Updated
Consensus Forecasts – Consensus economics	Jul 2021
Eurostat chain linked Volumes (GDP)	Aug 2021
IDC Core IT Spending guide 2H2018	Jun 2021
IDC European Vertical Markets survey (2019)	Sep–2019
IDC Worldwide Black Book v3.2 (standard edition)	Sep 2021
IMF World Economic Outlook (Oct 2019)	Aug 2021
IT Big Data and Analytics spending Guide 2H2021	Sep 2021

Table 34 The table below gives an overview of the latest estimates for the value and growth of the data market. The market grew by 3.8% in 2020 to a value of €60.6 million and is expected to reach €63.6 million in 2021. While 2020 and 2021 were difficult years, with Europe being hit by the COVID-19 pandemic and subsequent social and economic turmoil, the data market managed to remain healthy, benefitting from an increased focus on the digital economy for both work and services. Those companies and countries that were able to invest or utilise existing investments and expertise benefited from them. During 2020 and 2021 many organisations closed their activities or reduced business for a long period of time, following instructions from their governments, resulting in shortages of many key items and reduced output. Education, retail, transportation, hotels and catering

had significant impacts with restrictions on workers. Reductions in transportation also delayed delivery of materials which further restricted organisations' ability to operate – especially retail and manufacturing. However, data-oriented organisations were able to offer home or remote working for many of their employees and suffered less from a reduction in productivity. In addition, this working from home increased demand for many data products and services and encouraged organisations to accelerate transition plans for moving to a data oriented way of working.

Table 34. Value and Growth of the Data Market: 2019–2021 Revenue (€M) and 2020–2021 Growth (%)

N.	Market	Name	Description	2019	2020	2021	Growth Rate 2020–2021
4.1	EU27	Value of the Data Market	Estimate of the overall value of the Data Market	58,427	60,635	63,627	4.9%

6.3 Measuring the Value of Data

For the first time, this report estimates the size of data monetisation in the European Union – i.e., the value assigned to the sale and purchase of data among the Member States. This is an emerging market that shows great potential as the data economy develops. Measurement of the size of the market is preliminary because of uncertainty from survey respondents about what constitutes revenue associated with the sale of data. We estimate the value of data monetisation at €11.6 million in 2020 and €14.4 million in 2021 across the Member States. More detail about the value of data is given in section 6.6 – Data monetisation, the below table summarises the value of data monetisation among the Member States between 2019 and 2021.

There are limited channels for data, as organisations are unlikely to sell proprietary data to their competitors, and the internal data they are prepared to sell is unlikely to be of interest to many other data consumers. Much of the data that is available on open markets is mobility and location data, which has significant value in a few key industries. The survey associated with this report also identified what data vendors had success in selling, and this included retail data, customer activity data, and (anonymised, pseudo-anonymised, and/or aggregated) personal data; 64%, 51%, and 23% of companies surveyed had success in selling such data.

Table 35. Data Monetisation for the Member States: 2019–2021 and 2020 Growth (%)

N.	Market	Name	Description	2019	2020	2021	Growth Rate 2020–2021
4.1	EU27	Data Monetisation value	Estimate of the overall value of Data Monetisation	8,711	11,611	14,530	25.1%

6.4 Data Market by Country

The data market by country remains correlated with the economic strength of each of the Member States or countries under consideration. Those organisations that are active in the data market, generating data revenues, are also those organisations that are active monetising their internal data.

Across the countries analysed in this report, the United Kingdom continues to exhibit the strongest data market as a result of its focus in attracting service industries to the country. These include banking

& finance and professional services. In addition, a strong retail presence boosts the size of the market, as retail is a significant user of data tools and services.

Among the Member States, the data market is strongest in Germany, which is close in size to the United Kingdom. The country accounts for 28% of the EU27 data market in 2021, reflecting the strong manufacturing and finance base to its economy. In addition, the data market in Germany grew above average in 2021, consolidating its position as the EU27's largest market. Overall, most of the countries grew well in 2021, as each build on existing investments in the data economy. The below table presents the details by Member State and Figure 40 ranks the country markets in order, showing growth in 2021 for each country. It is easy to reflect on the strength of Germany in the data market due to its size, but its above-average growth is a better reflection of its focus on higher growth industries with a more data intensive focus, such as manufacturing and finance. Countries with larger data economies find it more difficult to achieve above average growth simply because of the value that needs to be added. In addition, smaller data economies can often be late in adopting data technologies, which means it is easier to show high growth in the early stages of a developing market such as applies to data technologies.

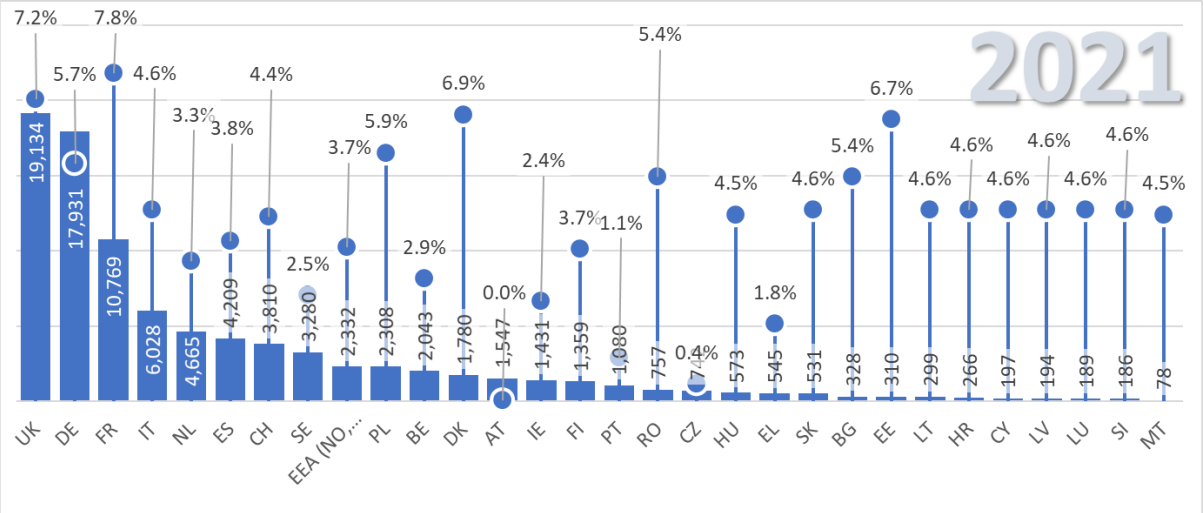
Across all the countries – including non-Member States - analysed in this report, the United Kingdom continues to exhibit the strongest data market as a result of its focus in attracting service industries to the country. These include banking and finance, and professional services. In addition, a strong retail presence boosts the size of the market as retail is a significant use of data tools and services.

Table 36. European Data Market by Country: 2019–2021 Revenue (€M) and 2020–2021 Growth (%)

Country	2019	2020	2021	Growth Rate 2020–2021
Austria	1,539	1,547	1,547	0.0%
Belgium	1,910	1,984	2,043	2.9%
Bulgaria	289	312	328	5.4%
Croatia	240	254	266	4.6%
Cyprus	178	189	197	4.6%
Czechia	732	739	742	0.4%
Denmark	1,591	1,666	1,780	6.9%
Estonia	269	291	310	6.7%
Finland	1,243	1,311	1,359	3.7%
France	9,494	9,986	10,769	7.8%
Germany	16,367	16,966	17,931	5.7%
Greece	517	535	545	1.8%
Hungary	527	549	573	4.5%
Ireland	1,388	1,397	1,431	2.4%
Italy	5,613	5,764	6,028	4.6%
Latvia	175	185	194	4.6%
Lithuania	270	286	299	4.6%
Luxembourg	171	181	189	4.6%
Malta	71	74	78	4.5%

Country	2019	2020	2021	Growth Rate 2020–2021
Netherlands	4,378	4,514	4,665	3.3%
Poland	2,086	2,179	2,308	5.9%
Portugal	1,062	1,069	1,080	1.1%
Romania	667	719	757	5.4%
Slovakia	479	508	531	4.6%
Slovenia	168	178	186	4.6%
Spain	3,996	4,053	4,209	3.8%
Sweden	3,007	3,199	3,280	2.5%
Switzerland	3,592	3,648	3,810	4.4%
United Kingdom	17,123	17,845	19,134	7.2%
EEA (NO, LI, IS)	2,228	2,249	2,332	3.7%
EU27	58,427	60,635	63,627	4.9%
EEA (NO, LI, IS) + CH	5,821	5,897	6,142	4.1%
Total, all countries	81,371	84,377	88,903	5.4%

Figure 40. European Data Market: 2020 Revenue (€M) and 2020–2021 Growth (%)



6.5 Data Market by Industry

This First Report on Facts & Figures of the European Data Market 2021–2023 introduces new industry sectors in the measurement of the data market by industry. While only contributing 0.4% to the data market in 2021, agriculture has been added to the list of sectors covered by this indicator, as it is currently undergoing a fast and profound process of digital transformation. Indeed, agriculture has potential for growth with a range of data applications relating to crop yield improvements and data management, among others.

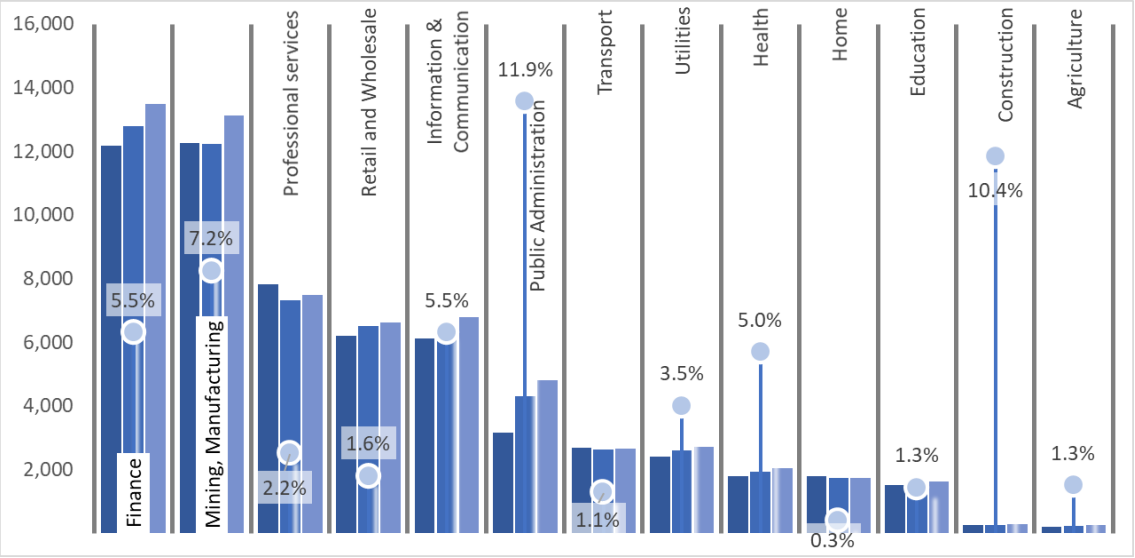
The below table sorts the data market industries by size for 2019 to 2021 and shows that the share taken by finance and manufacturing accounts for nearly 42% of the data market. These are traditional

data markets, and adding professional services, retail & wholesale, and information & communications pushes it to nearly three-quarters of all spending on data tools and services.

Table 37. Data Market by Industry: 2019–2021 (€M) and 2020–2021 Growth (%)

Industry	2019	2020	2021	Growth Rate 2020–2021
Agriculture	214	242	245	1.3%
Construction	253	256	283	10.4%
Education	1,522	1,614	1,634	1.3%
Finance	12,176	12,782	13,487	5.5%
Health	1,794	1,943	2,040	5.0%
Information & communications	6,131	6,446	6,801	5.5%
Mining & manufacturing	12,252	12,240	13,124	7.2%
Professional services	7,819	7,326	7,486	2.2%
Public administration	3,158	4,307	4,818	11.9%
Retail & wholesale	6,202	6,507	6,610	1.6%
Transport	2,697	2,632	2,661	1.1%
Utilities	2,415	2,610	2,702	3.5%
Home	1,794	1,730	1,736	0.3%
EU27	58,427	60,635	63,627	4.9%

Figure 41. Data Market by Industry: 2019–2021 Revenue (€M) and 2020–2021 Growth (%)



6.6 Data Monetisation

Data monetisation is an emerging and somewhat hidden part of the data market, which involves the sale of data between organisations rather than the sale of data software, hardware, or services. Many organisations have a wealth of data and have recently discovered there is a market for this. Typically, this includes location data – giving information about where goods or people are located in space and time, purchase data – what people buy, equipment performance data – information about the how equipment is performing and indication of when equipment is operating out of tolerance. Other data types include customer activity data and personal data, but the need to

comply with GDPR in the EU means much of this data needs to be processed to aggregate and anonymise it.

Information about the size of the market for data monetization is limited and of poor quality so until now estimating the size and potential for this market has been difficult, with widely varying estimates of the market size. This study uses survey responses to estimate the provisional size of data monetization, asking respondents about their use of data markets to buy and sell data, and the type of data sold. While this data is relatively weak (due to the newness of the market and the lack of clarity among respondents of what constitutes data sale and purchase) This is used to guide a simple penetration model, using a share of the total data market to estimate data monetization. As more reliable data becomes available this model will be refined to reflect better the supply and demand for data among the different countries and industries.

Data monetisation is estimated for the first time in this report and is expected to account for as much as 30% of additional spending on the data market to 2030 (Baseline). As shown in Figure 42, data monetisation is difficult to assess well because of the different definitions of what is included. As part of the research for this report, the survey identified the key areas where companies spend on data sale and purchase, and the key areas are retail data, customer activity data, and personal data. Location data accounts for only 11% of total data monetisation, and streaming data for even less, at 3.6%. Mobility and location data was significant in the past, but the easy availability of this undervalues the data significantly.

Data monetisation is expected to grow rapidly over the period 2020 to 2030, with a compound growth rate of close to 12 percent from 2020 to 2030 (Baseline). Figure 43 and Table 38 show data monetisation by Member State for 2019-2021 with Germany, and France taking the top two positions among the member states in terms of monetisation revenues.

Table 38. Data Monetisation by Country: 2019–2021 Revenue (€M) and 2020–2021 Growth (%)

Country	2019	2020	2021	Growth Rate 2020–2021
Austria	244	333	361	8.5%
Belgium	175	218	270	23.9%
Bulgaria	24	29	36	24.9%
Croatia	21	34	39	14.8%
Cyprus	9	12	14	24.5%
Czechia	56	80	103	29.2%
Denmark	275	349	440	26.1%
Estonia	49	67	85	26.9%
Finland	114	162	213	31.0%
France	1,431	2,104	2,744	30.4%
Germany	3,022	3,857	4,798	24.4%
Greece	34	47	57	21.5%
Hungary	38	50	59	19.5%
Ireland	202	252	306	21.3%
Italy	994	1,295	1,609	24.2%
Latvia	12	18	22	23.4%
Lithuania	22	29	35	19.7%
Luxembourg	13	19	25	31.6%
Malta	6	6	7	5.6%
Netherlands	691	902	1,120	24.1%
Poland	155	223	294	31.4%
Portugal	114	165	182	10.5%
Romania	53	70	88	27.2%
Slovakia	38	44	53	21.2%
Slovenia	13	18	21	15.5%
Spain	614	740	908	22.7%
Sweden	351	491	643	30.9%
Switzerland	434	551	634	15.0%
United Kingdom	3,110	4,228	5,151	21.8%
EEA (NO, LI, IS)	187	223	306	37.3%
EU27	8,771	11,611	14,530	25.1%
EEA (NO, LI, IS) + CH	621	774	940	21.4%
Total, all countries	12,502	16,612	20,620	24.1%

Figure 42. Data Monetisation Share of the European Data Market, 2013–2030 (Baseline) (%)

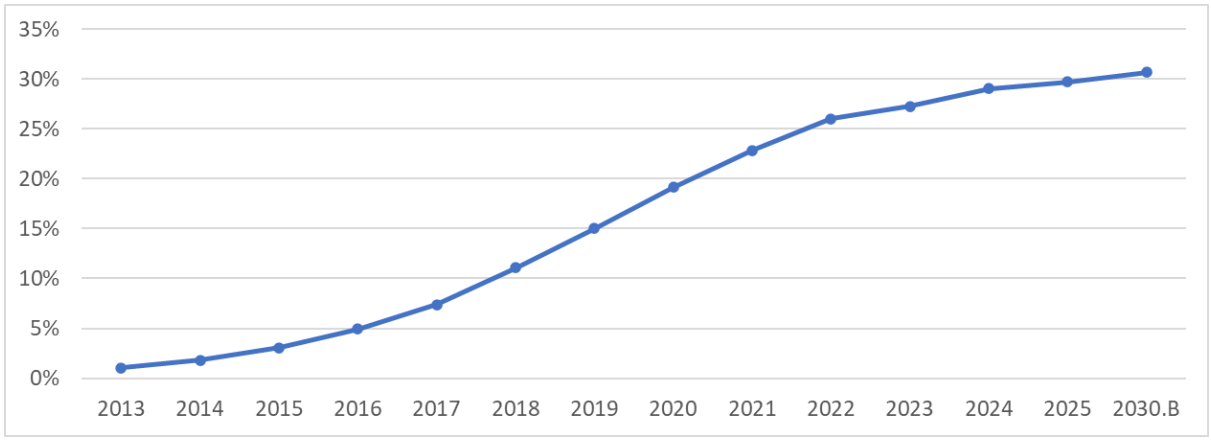
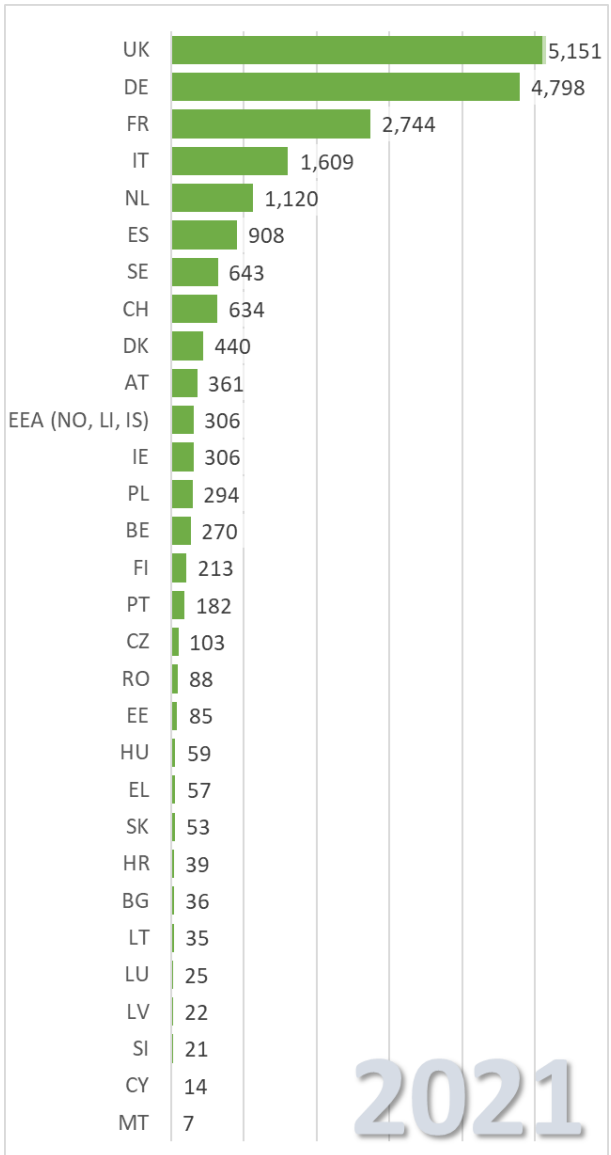


Figure 43. Data Monetisation by Country, 2021 (€M)

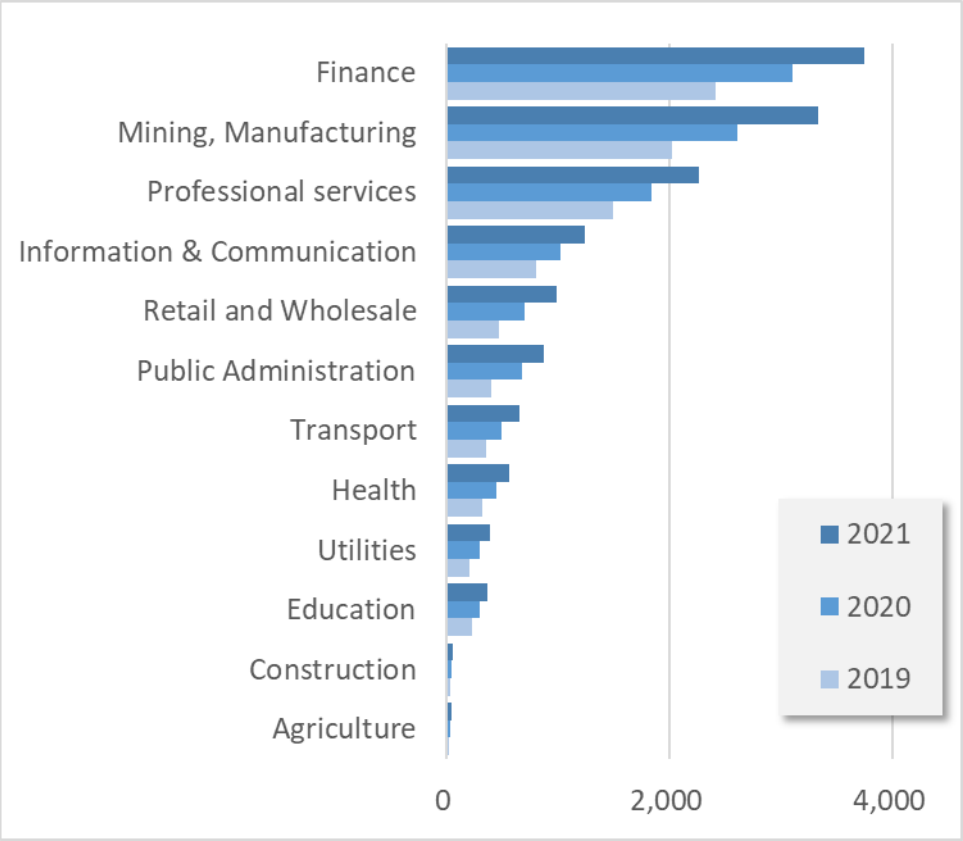


Those industries most likely to show significant revenues from data monetization are those with the strongest presence in the data market – primarily because companies in these industries already

have tools and processes in place to gather data relevant to their business. These businesses have the business processes in place for acquisition of data tools, services, and hardware so will be prepared for the sale of appropriate data sets too. However, as mentioned before, channels for the sale of data are likely to be limited because of the sensitivity of some of the data relevant to the business.

Industries most likely to be successful in the sale of data through data markets are finance, manufacturing, and professional services – as seen in Figure 44.

Figure 44. Data Monetisation by industry - 2019-2021 (Eur. M)



6.7 Forecasting the Data Market

The data market for the EU27 will grow to 105.6 million Euro in 2030 (Baseline), representing a compound growth of 3.2% over the 2025–2030 period. Table 39 shows the size and growth for the market for the three scenarios for 2030. It also shows that the EEA (NO, LI, IS) and Switzerland have stronger growth than the Member States during the period 2025-2030 for both the baseline and high growth scenarios, recovering from lower growth during the period 2020-2025. Although they represent only 6.7 percent of all countries' spend on data technologies these countries invested at a lower rate than the larger EU member states, and also have an industry focus less geared towards early investment in data. Norway dominates the GDP of the non-EU countries, accounting for close to 80 percent of the combined GDP and its biggest industries are oil and gas, shipping, and fishing. Oil and gas has the potential for significant investment in data technologies in areas such as transportation and monitoring of performance, while shipping can benefit from investment in data technologies such as location data and logistics. Luxembourg's strong finance and banking industries contribute to stronger investment in data technologies between 2025 and 2030.

Table 39. Data Market Forecast: 2025, Three 2030 Scenarios (€'000s), and Compound Growth (%)

	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR: 2020–2025	CAGR: 2025–2030, Challenge	CAGR: 2025–2030, Baseline	CAGR: 2025–2030, High Growth
EU27	90,121	94,218	105,619	125,238	8.2%	0.9%	3.2%	6.8%
EEA (NO, LI, IS) + CH	8,453	8,822	10,327	12,316	7.5%	0.9%	4.1%	7.8%
Total, All Countries	125,221	130,640	145,335	171,087	8.2%	0.9%	3.0%	6.4%

The Data Market by Country

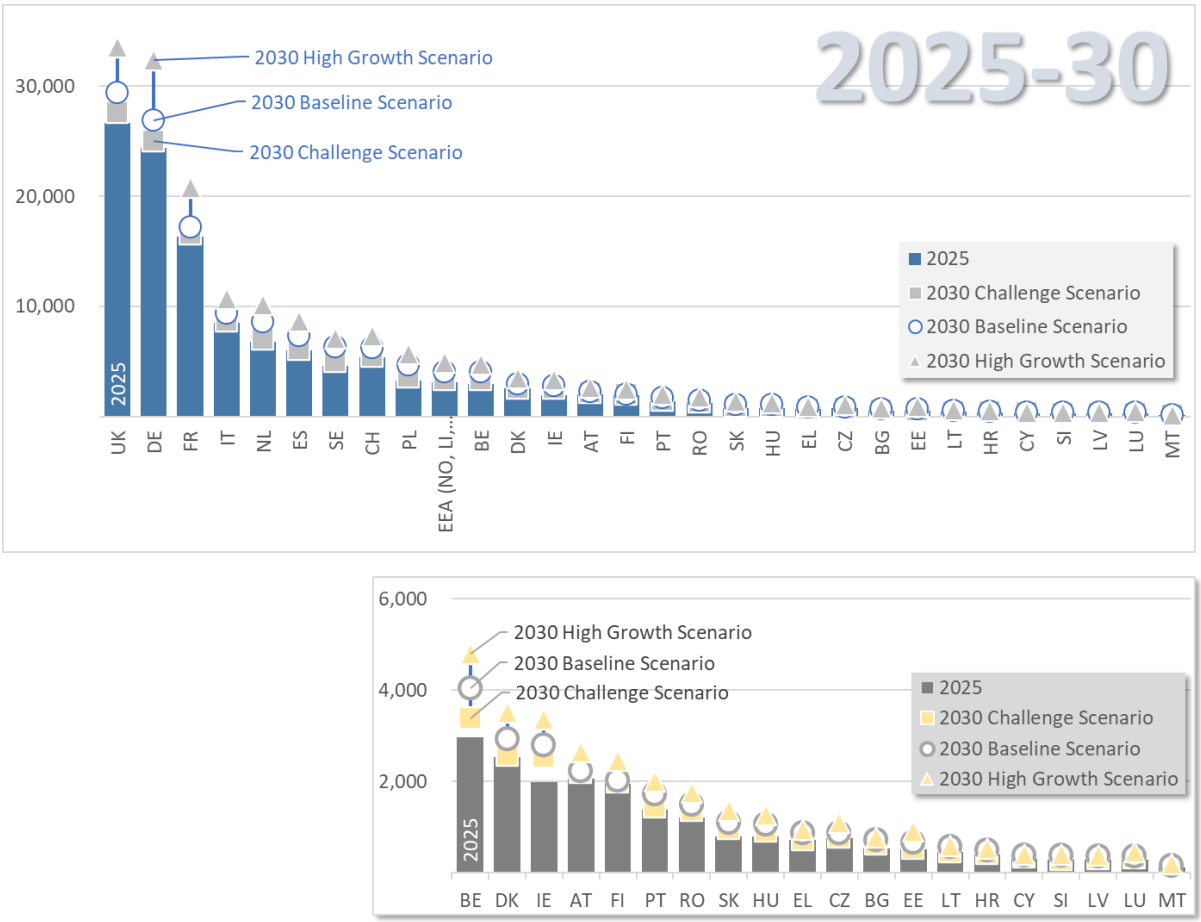
Among the Member States, Germany shows the largest market, growing to 25.4% of the EU27 total by 2030 (Baseline). However, the market in Germany will grow at a below average compound growth rate of just 1.9% from 2025 to 2030, compared with 3.2% for the EU27 as a whole. Germany is easily the largest Member State, followed by France, with nearly 10 percentage points less, representing just 16.3% of the EU27 total. Germany's strength in the larger industries of manufacturing and finance ensures it will maintain its leading position among the Member States.

Outside the Member States, the United Kingdom leads Germany in the size of its data market, with greater investment in the leading data market industries of finance, public administration, information & communications, and professional services. Table 40 shows the forecast for all countries and Member States, while Figure 45 shows the forecast, in order, and the three scenarios.

Table 40. European Data Market Forecast by Country: 2025, Three 2030 Scenarios (€M), and CAGR (%)

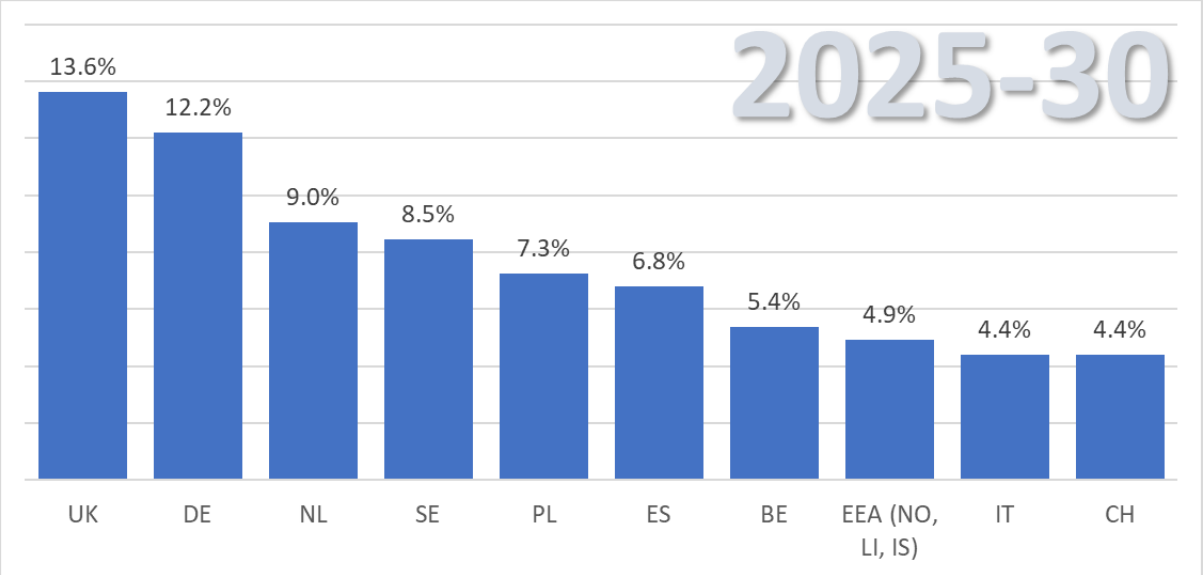
Country	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR: 2020–2025	CAGR: 2025–2030, Challenge	CAGR: 2025–2030, Baseline	CAGR: 2025–2030, High Growth
Austria	2,050	2,184	2,234	2,652	5.8%	1.3%	1.7%	5.3%
Belgium	2,967	3,380	4,051	4,792	8.4%	2.6%	6.4%	10.1%
Bulgaria	526	618	714	783	11.0%	3.3%	6.3%	8.3%
Croatia	397	419	494	531	9.3%	1.1%	4.5%	6.0%
Cyprus	294	347	387	399	9.3%	3.4%	5.6%	6.3%
Czechia	751	782	864	1,105	0.3%	0.8%	2.8%	8.0%
Denmark	2,533	2,565	2,929	3,506	8.7%	0.2%	2.9%	6.7%
Estonia	511	523	653	902	11.9%	0.5%	5.0%	12.1%
Finland	1,950	1,986	2,024	2,461	8.3%	0.4%	0.7%	4.8%
France	16,309	16,583	17,168	20,768	10.3%	0.3%	1.0%	5.0%
Germany	24,426	25,048	26,879	32,363	7.6%	0.5%	1.9%	5.8%
Greece	715	718	875	951	6.0%	0.1%	4.1%	5.9%
Hungary	799	910	1,064	1,262	7.8%	2.6%	5.9%	9.6%
Ireland	1,976	2,534	2,788	3,362	7.2%	5.1%	7.1%	11.2%
Italy	8,486	8,703	9,374	10,697	8.0%	0.5%	2.0%	4.7%
Latvia	289	299	365	384	9.3%	0.7%	4.8%	5.9%
Lithuania	446	462	565	588	9.3%	0.7%	4.9%	5.7%
Luxembourg	282	317	364	449	9.3%	2.3%	5.2%	9.7%
Malta	108	135	151	180	7.8%	4.6%	7.0%	10.7%
Netherlands	6,799	7,056	8,617	10,125	8.5%	0.7%	4.9%	8.3%
Poland	3,240	3,567	4,701	5,699	8.3%	1.9%	7.7%	12.0%
Portugal	1,382	1,440	1,718	2,003	5.3%	0.8%	4.4%	7.7%
EEA (NO, LI, IS)	3,088	3,340	4,078	4,949	6.5%	1.6%	5.7%	9.9%
Romania	1,213	1,354	1,490	1,750	11.0%	2.2%	4.2%	7.6%
Slovakia	792	957	1,105	1,369	9.3%	3.9%	6.9%	11.6%
Slovenia	277	285	372	387	9.3%	0.6%	6.1%	6.9%
Spain	5,996	6,070	7,368	8,662	8.1%	0.2%	4.2%	7.6%
Sweden	4,607	4,977	6,308	7,106	7.6%	1.6%	6.5%	9.1%
Switzerland	5,365	5,482	6,249	7,368	8.0%	0.4%	3.1%	6.5%
United Kingdom	26,647	27,601	29,390	33,533	8.3%	0.7%	2.0%	4.7%
EU27	90,121	94,218	105,619	125,238	8.2%	0.9%	3.2%	6.8%
EEA (NO, LI, IS) + CH	8,453	8,822	10,327	12,316	7.5%	0.9%	4.1%	7.8%
Total, all countries	125,221	130,640	145,335	171,087	8.2%	0.9%	3.0%	6.4%

Figure 45. European Data Market Forecast by Country: 2025, Three 2030 Scenarios (€M), and Compound Growth (%)



While the markets for Germany, France, Italy and United Kingdom will all grow at below the average for either the EU27 or all countries, the size of these markets means their lower growth will still add more to the data market over the forecast period, in spite of this lower growth. Figure 46 shows the contribution made by each country to total growth between 2025 and 2030 (Baseline). The United Kingdom will account for 13.6% of the total growth seen over this period, in spite of its low CAGR of 2.0%. The UK will be followed closely by Germany, which will account for 12.2% of the overall growth. The Netherlands is the third largest contributor to growth, despite being only the fifth largest country as a result of its stronger growth over the period, rather than its size. Focussing on the contributions shows the importance of size, as well as growth, when looking to see where the majority of the increased spending will occur.

Figure 46. Contribution to Overall Growth by Country: 2025–2030 Baseline, Top 10 Contributors



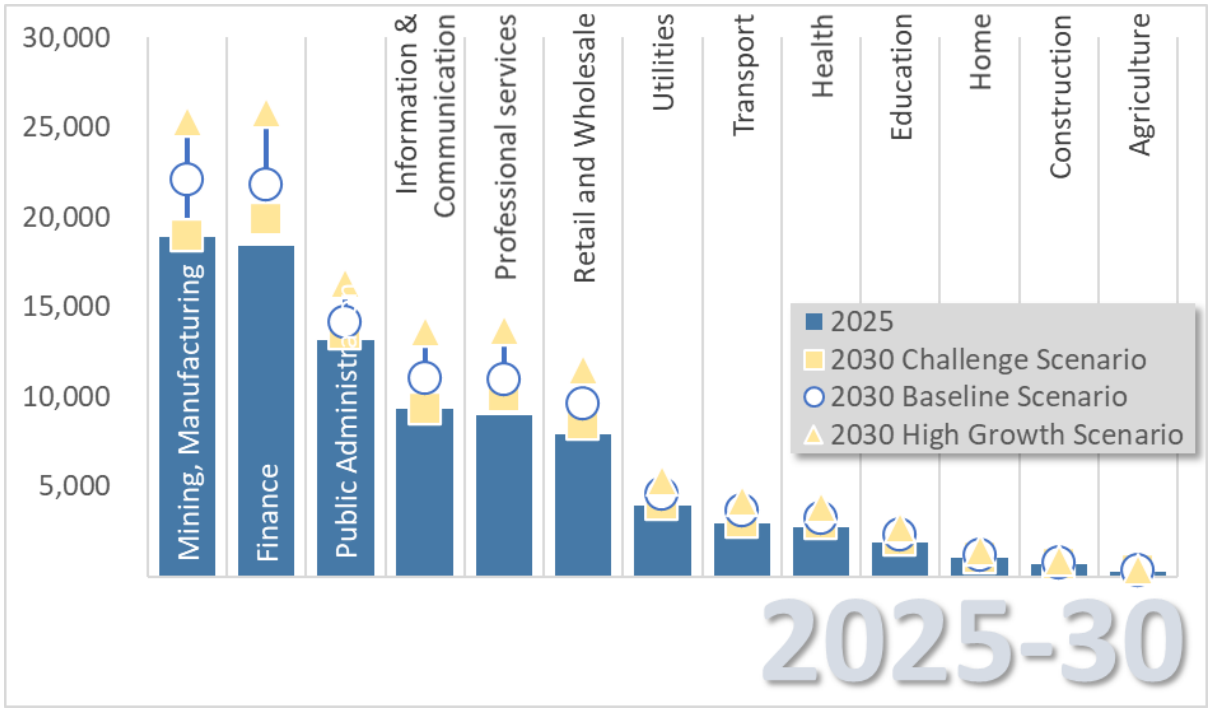
The Data Market by Industry

The largest industries within the data market in 2030 (Baseline) will be manufacturing and finance, together accounting for 41% of spending. Public administration makes a notable contribution, too, ahead of information & communications and professional services. These five industries will make up 75% of spending in the data market. Those countries with a strong presence in these industries will benefit substantially from growth in the data market. Table 41 and Figure 47 summarise the information for each of the industries in the EU27.

Table 41. European Data Market Forecast by Industry: 2025, Three 2030 Scenarios (€M), and CAGR (%)

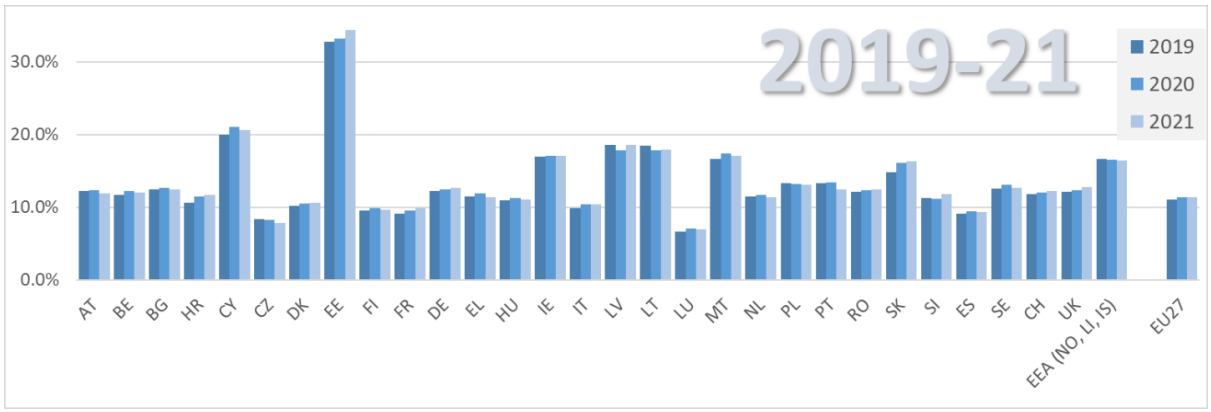
Industry	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR 2020–2025	CAGR 2025–2030, Challenge	CAGR 2025–2030, Baseline	CAGR 2025–2030, High Growth
Agriculture	271	307	337	393	2.3%	2.6%	4.4%	7.7%
Construction	687	704	744	859	21.8%	0.5%	1.6%	4.5%
Education	1,892	2,027	2,330	2,738	3.2%	1.4%	4.2%	7.7%
Finance	18,414	19,855	21,784	25,866	7.6%	1.5%	3.4%	7.0%
Health	2,729	2,962	3,244	3,868	7.0%	1.7%	3.5%	7.2%
Home	1,013	1,046	1,195	1,426	-10.2%	0.7%	3.4%	7.1%
Information & communications	9,288	9,340	10,987	13,611	7.6%	0.1%	3.4%	7.9%
Mining & manufacturing	18,917	18,972	22,090	25,306	9.1%	0.1%	3.1%	6.0%
Professional services	8,992	10,008	10,970	13,681	4.2%	2.2%	4.1%	8.8%
Public administration	13,181	13,502	14,135	16,328	25.1%	0.5%	1.4%	4.4%
Retail & wholesale	7,902	8,467	9,604	11,553	4.0%	1.4%	4.0%	7.9%
Transport	2,915	3,047	3,630	4,253	2.1%	0.9%	4.5%	7.9%
Utilities	3,919	3,978	4,571	5,357	8.5%	0.3%	3.1%	6.5%

Figure 47. European Data Market Forecast by Industry: 2025 and Three 2030 Scenarios (€M)



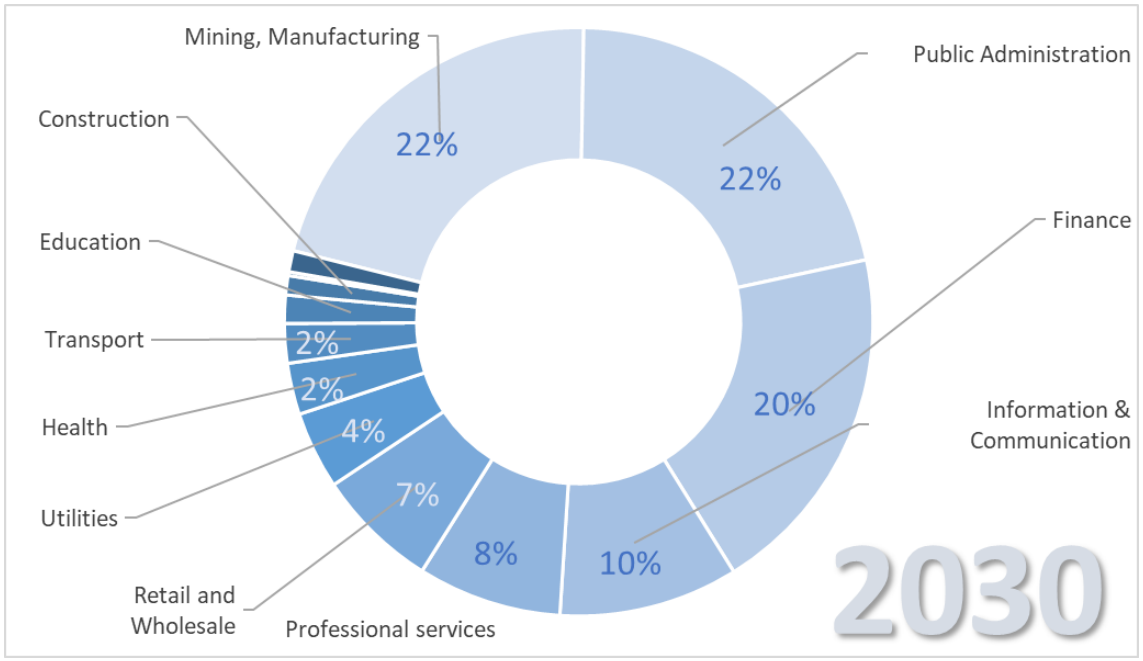
The data market's share of total ICT spending by country in continental Europe shows a reasonable balance, although Estonia has a stronger than average focus on the data market, with this country accounting for more than 33% of total ICT spending in the Member States. Across the Member States, the average share of total ICT spending accounted for by the data market is 11.4% in 2021.

Figure 48. Data Market Share of Total ICT Spending, 2019–2021 (%)



Unlike with Member States, the industry contribution to overall growth from 2020 to 2030 (Baseline) shows a more even balance across industries. Manufacturing and public administration will contribute the most to the growth seen between 2020 and 2030 (as shown in Figure 49). Public administration is only the third largest, but its above average growth means it will add more to the market overall. The order of the industries, though, reflects closely the size of each industry. But, again, this shows where the focus lies in looking for those industries that add the most to the data market over the 10 years from 2020 to 2030.

Figure 49. Industry Contribution to Overall Growth, 2020–2030 (%)



6.8 Key Findings

The value of the European data market will reach €63.6 billion for the EU27, with a growth rate of 4.9% in 2021. Most Member States show strong growth, slightly slower than the growth for the total ICT market, which is forecast to grow 5.1% in 2021. The data market's share of total ICT spending was 11.4% in 2020.

The size of the data market by country still correlates closely with the overall economic strength of each country, as well as with national spending on ICT: Germany, France, Italy, the Netherlands, and Spain accounted for approximately two-thirds of the EU27 data market in 2020. Some smaller economies, however, continue to display higher-than-average data market shares of ICT spending. This is notably the case of Estonia and, to a lesser extent, Cyprus, Latvia, and Lithuania. Aside from the outliers, the spread of share of ICT spending taken by the data market is fairly narrow across most Member States.

The larger industries, accounting for the greatest number of companies, represent the largest share of the data market. In terms of adoption by industry, the highest rates of data technology tend to be in manufacturing, finance, and public administration. Thanks to the size of these markets, organizations in these industries are the biggest consumers of data technologies – partly, because of the significant number of companies and their early investment in data tools and services.

The forecast for the data market shows which industries make the biggest contribution to overall market growth, and the key industries of manufacturing, public administration, finance, information technology, and professional services account for close to 82% of the total market growth from 2020 to 2030 (Baseline), with more than 60% coming from the three main industries of manufacturing, public administration, and finance.

Data monetisation will contribute an increasing share to the total data market; it is forecast to add as much as 30% to the European data market by 2030, according to the Baseline forecast. Even though, businesses are struggling to understand how to monetise data and how to capitalize on the data they have to sell and what data they can buy, **we expect that this market will play pivotal role.**

7. The Data Economy

7.1 Measuring the Data Economy: Definitions

The **data economy** measures the overall impacts of the data market on the economy as a whole. It involves the generation, collection, storage, processing, distribution, analysis, elaboration, delivery, and exploitation of data enabled by digital technologies.

The data economy captures a wider concept than the data market only, as it considers the value and wealth generated in the economy as a whole (not just across businesses) by the exploitation of data.

The data market captures the value of the market in which data-driven products and services (apps, IoT-based products, and all sorts of services relying on heavy use of data) are exchanged. It captures the aggregate value of the demand of digital data. It includes not only the value generated by pure data players developing BDA technologies, but also the value created by data-related research, businesses, information, and IT services. The digital data exchanged as "products" and "services" in the data market refers exclusively to data that is collected, processed, stored, and transmitted over digital information infrastructures and/or elaborated with digital technologies.

The data economy includes three sets of impacts on the economy: data companies' revenues in the form of direct impacts on the economy, indirect impacts (backward and forward) on the economy, and the induced impacts of the data market on the economy.

1. The **direct impacts** are the initial and immediate effects generated by the data supplier companies. The quantitative direct impacts will then be measured as the revenues from data products and services sold, i.e. the value of the data companies' revenues.
2. The **indirect impacts** are the economic activities generated along a company's supply chain by data supplier companies, considering input providers as well as customers of data supplier companies. And, for this reason, there are two different types of indirect impact: backward indirect impacts and forward indirect impacts.
    - **Backward indirect impacts** represent the revenues resulting from changes in sales from input providers to the data suppliers. In order to produce and deliver data products and services, data suppliers need inputs from other stakeholders.
    - The **forward indirect impacts** include the economic growth generated through the use of data products and services by the downstream industries, i.e. the data users.
3. **Induced impacts** include the economic activity generated in the whole economy as a secondary effect. Induced additional spending is generated by both new workers, who receive a new wage, and the increased wages of existing jobs. This spending induces new revenue creation in nearly all sectors of the economy. The additional consumption will support economic activity in various industries, such as retail, consumer goods, banks, and entertainment.

The difference between the data market (the exchange of data) and the broader Data Economy is represented in Figure 50. A concrete example is the following: any company either headquartered in EU or outside, globally provides an IoT solution for shipping data. Any company (in EU or outside) using the solution provides location information (data provider) of its ships at sea, so ports (data users) can schedule the arrivals. This exchange of data generates two streams of revenues:

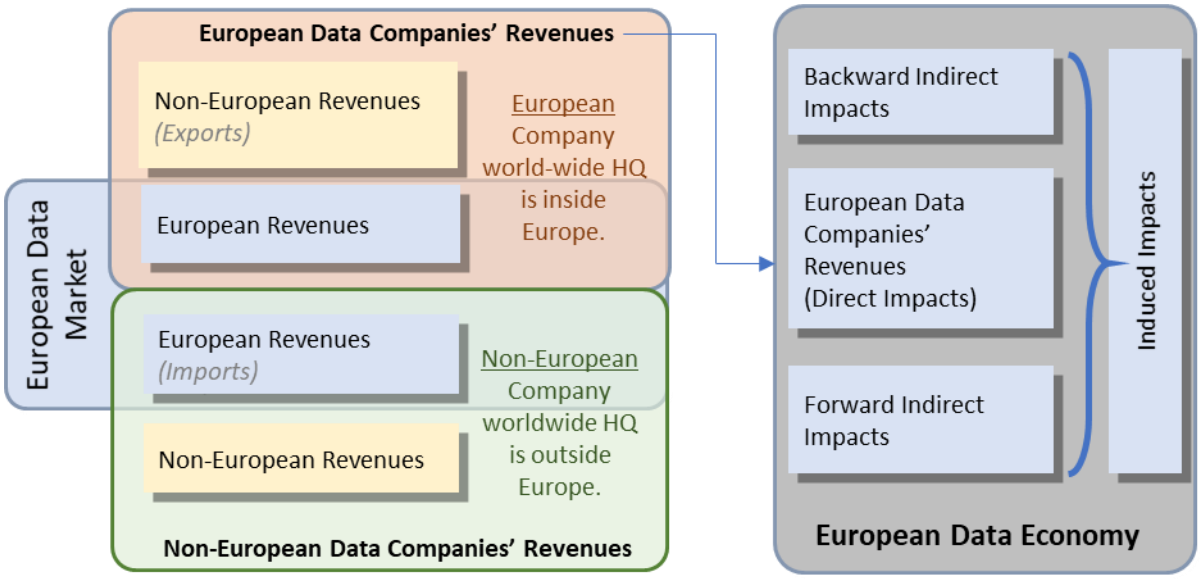
- The revenues generated at the European level (also through data imports) by this exchange of data represent the Data Market.
- The revenues generated by European companies (also through data exports) represent the European Data Companies' Revenues, which in turn represent the Direct Impacts of the broader Data Economy.

The Data Economy measures:

- the revenues generated by European data companies that sell and produce data products and services (measured by data companies' revenues), but it includes also
- the value/revenues generated in their supply chain by data companies' clients who will increase their revenues and employees thanks to the benefits deriving from the use of a specific solution
- the value/revenues generated by data companies' providers who provide all the components such as hardware, software, services, but also all procurement that enable the company to run and to build a specific solution.
- and finally, as a secondary effect, the value/revenues generated in the whole economy as new jobs will be created thanks to the benefits deriving from the use of this data products and services, but not directly generated by the data exchange market

Figure 50 shows the relationship between the Data Suppliers Companies' Revenues, the European Data Market, and the European Data Economy.

Figure 50. Relationship of Data Market, Data Revenues, Data Economy



7.2 Changes from Prior Deliverable

The most relevant change with respect to prior deliverable is due to the change in the size of indirect backward impacts, as we enlarged the pool of data suppliers companies (please refer to Data Companies chapter). Backward impacts represent the business growth resulting from changes in sales from suppliers to the data industry. In order to produce and deliver data products and services, the data companies need inputs from other stakeholders. Revenues from those sales to data companies are the backward indirect impacts. This change results in higher indirect impacts for 2020 and forecast.

7.2 Measuring the Data Economy

The Data Economy: 2019–2021

The value of the data economy for the EU27 has been estimated to have reached almost €400 billion in 2019 and €440 billion in 2021, with a year-on-year growth rate of 4.9% in 2021. The estimated share of overall impacts on GDP in the EU27 ranges from 3.1% in 2019 to 3.6% in 2021. The growth rate in 2021 is slightly lower than in previous years due to the slower growth of indirect impacts, as indirect impacts are closely linked with data companies' revenues. Indeed, in 2020, data companies' revenues grew faster than all company revenues, as explained in Chapter 5, as this is an emerging market and general mitigating factors such as Brexit and COVID-19 had little impact on growth, with organisations worldwide embracing remote working and online business during this period. This high growth rate simply makes 2021 growth lower, as we can consider 2020 as a peak in growth, which will stabilise in the coming years.

The UK share of the data economy as a part of GDP is significant, bringing the overall share of the data economy in the EU27 plus the UK as a part of GDP in 2021 to around 4%. The flow of data is vital for the UK and EU economies. Indeed, with its large service sector, the UK has the largest internet economy as a proportion of GDP within the G20, reflecting the centrality of data to most goods and services trade. An interruption in data flows would therefore be costly. Despite Brexit uncertainties existing so far, the European Commission has recently adopted two adequacy decisions that allow for the free flow of personal data from the European Union to the United Kingdom in the case that the exchanged data is granted an essentially equivalent level of protection to the one guaranteed under EU law. This is reflected in the data economy value for the UK, which will remain strong throughout the period.

EEA countries (NO, IS, LI) alone show a 2021 growth of 5.3%, as 2020 growth did not reach the same peak as in the EU countries. Norway shows the lowest growth among the EEA countries. Overall, the growth rate in 2021 for the three different markets is similar, ranging from 4.9% to 5.1%.

Table 42. Data Economy Value and Growth, 2019-2020-2021 and Impacts on GDP 2019-2020-2021 (€, Million; %)

N.	Market	Name	Description	2019	2020	2021	Growth Rate 2020–2021	Impact on GDP, 2019	Impact on GDP, 2020	Impact on GDP, 2021
5.1 5.2	EU27	Value of the data economy and impacts on EU GDP	Value of total impacts on the EU economy and % of EU GDP	386,415	421,911	442,572	4.9%	3.1%	3.6%	3.6%
5.1 5.2	EU27+ UK	Value of the data economy and impacts on EU GDP	Value of total impacts on the EU economy and % of EU GDP	507,452	550,892	579,119	5.1%	3.4%	4.0%	4.0%
5.1 5.2	Total, all countries	Value of the data economy and impacts on EU GDP	Value of total impacts on the EU economy and % of EU GDP	553,690	599,624	630,410	5.1%	3.5%	4.1%	4.1%

Source: European Data Market Monitoring Tool, IDC 2021

The Data Economy in 2025

The European strategy for data aims to create a single market for data that will ensure Europe's global competitiveness and data sovereignty. The EU is now pursuing the creation of a European data infrastructure – as well as other initiatives, such as enhancing data security – to ensure the EU's leadership in the global data economy (Proposal for a Regulation on European data governance [*EU Data Governance Act*]) and support the growth of the European data economy in the future.

In 2025, European initiatives on data and data regulation are expected to produce results: A strong increase of the data economy's value as a part of the overall economy is expected, with the data economy's impact in terms of GDP increasing by almost 1 percentage point in both the EU27 and the EU27 + UK, as well as in the overall considered region (EU27 + UK + Switzerland + EEA: Iceland, Lichtenstein, and Norway).

Another important contribution to this growth of the data economy is the investment measures imbedded in the Next Generation EU (NGEU) fund and national Recovery and Resilience reforms. Indeed, the NGEU and national funds will play a crucial role in the next three to five years, by enabling countries to put in place a series on investments around key technologies, from cloud to 5G, IoT and smart devices, edge, artificial intelligence, and digital infrastructure, among others. Behind the increasing adoption of innovative technologies, the value of data is the common denominator. European companies will focus on tailored services for consumer satisfaction, will change towards competitive tech and data-driven business models. Companies are investing in technology to transition from traditional workplace to hybrid models; they must become more agile and adjust their operations and culture by addressing employee mobility, secure remote collaboration, and virtualization needs.

These results are also in line with the previous results of the European Data Market Monitoring Tool in 2020, when IDC estimated that the value of the data economy in a Baseline scenario for 2025 would amount at approximately €555 billion for the EU27. In this new round of measurement, we estimate that the data economy for the EU27 in 2025 will reach €600 billion, which is in line with the changes

made to the data suppliers' perimeter, which, as said, increases the revenues generated at an indirect level.

Table 43. Data Economy Value (€M), Growth (%), and Impact on GDP (%), 2025

N.	Market	Name	Description	2020	Impacts on GDP 2020	2021	Impacts on GDP 2021	2025	CAGR 2025/2020	Impacts on GDP 2025
5.1 5.2	EU27	Value of the Data Economy and Impacts on EU GDP	Value of total impacts on the EU economy and % of EU GDP	421,911	3.6%	442,572	3.6%	604,099	7.4%	4.4%
5.1 5.2	EU27+ UK	Value of the Data Economy and Impacts on EU GDP	Value of total impacts on the EU economy and % of EU GDP	550,892	4.0%	579,119	4.0%	777,742	7.1%	4.9%
5.1 5.2	Total, all countries	Value of the Data Economy and Impacts on EU GDP	Value of total impacts on the EU economy and % of EU GDP	599,624	4.1%	630,410	4.1%	843,538	7.1%	4.9%

Source: European Data Market Monitoring Tool, IDC 2021

The Data Economy by Impact Type: 2021 and 2025

Analysis of the data economy also provides an overview of the total effects by type of impact (direct, indirect, and induced). It is worth mentioning how the composition of impacts changes over time, from 2021 to 2025, in favour of induced impacts, thus revealing the effects of data access, data products and services exchange, and data value distribution in the economy.

The shift toward induced impacts is slightly offset by the change in data supplier companies, which have increased the revenues generated at the input provider level, the backward impacts. These revenues have increased significantly since the previous deliverable. Indeed, as the number of data suppliers increases, requests for inputs in the supply chain also increase, thus helping the input provider companies to further improve their revenues.

Figure 51. Data Economy by Impact Type: EU27 + the UK, 2021 and 2025 (%)

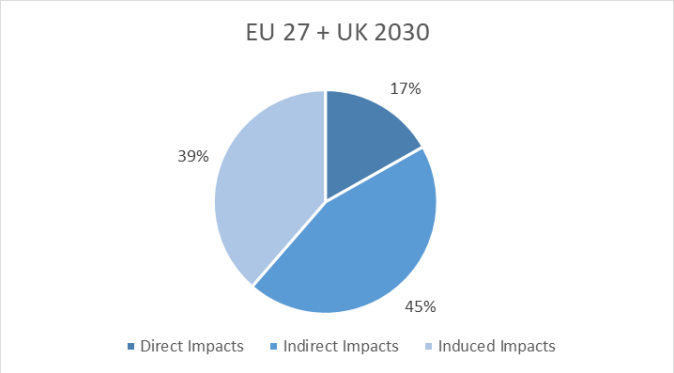
Source: European Data Market Monitoring Tool, IDC 2021

7.4 Forecasting the Data Economy in 2030

In 2030, the data economy for the EU27 is expected to stay below the €1 trillion threshold, with a 5.3% 2025–2030 CAGR. When we consider the UK, Switzerland and the rest of the EEA countries, the value of the data economy in 2030 is expected to reach the €1 trillion threshold, with a 2025–2030 CAGR of 5.7%, slightly lower than the average increase of 7.4% (7.1% with UK, CH and EEA) in the period 2020–2025. The data economy will record the highest growth in the post-COVID period as a result of the exceptional measures coming from the NGEU and from the urgency to adapt to the new normal scenario.

- One of the main results is that, despite the slower growth, the share of the data economy as a part of GDP in the EU27 plus the UK will increase from 4.9% in 2025 to 5.9% in 2030.
- Another important result is the change in the composition of impacts: Indeed, from 2021 to 2030, the induced impacts' share will increase from 34% to 39% at the expense of indirect impacts, which will diminish from 50% to 45%.

Figure 52 Data Economy by Impact Type: EU27 + the UK, 2030 (%)



Source: European Data Market Monitoring Tool, IDC 2021

The two alternatives show that, in case of the Challenge scenario, the average growth for the EU27 (2030–2025 CAGR) will be half that expected in a baseline scenario (2.6% in the Challenge scenario and 5.4% in the Baseline scenario), while it will be nearly double in High Growth scenario (9.6%).

Table 44. Data Economy Value (€M); 2030 Challenge, Baseline, and High Growth Scenarios (€M); and Impact on GDP (%)

N.	Market	Name	Description	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	Impacts on GDP: 2030 Challenge Scenario	Impacts on GDP: 2030 Baseline Scenario	Impacts on GDP: 2030 High Growth Scenario	2025-2030 CAGR: Challenge	Baseline	High Growth
5.1 5.2	EU27	Value of the data economy and impacts on EU GDP	Value of total impacts on the EU economy and % of EU GDP	687,478	786,843	954,658	4.9%	5.3%	6.2%	2.6%	5.4 %	9.6 %
5.1 5.2	EU27+ UK	Value of the data economy and impacts on EU GDP	Value of total impacts on the EU economy and % of EU GDP	904,567	1,028,013	1,248,004	5.5%	5.9%	6.9%	3.1%	5.7 %	9.9 %
5.1 5.2	Total, all countries	Value of the data economy and impacts on EU GDP	Value of total impacts on the EU economy and % of EU GDP	976,091	1,115,486	1,353,445	5.5%	5.9%	7.0%	3.0%	5.7 %	9.9 %

Source: European Data Market Monitoring Tool, IDC 2021

7.5 Data Economy by Industry

The industry split of the data economy over the 2019–2025 period follows that of the data revenues generated from the supply side and the adoption of technologies such as big data analytics, artificial intelligence, and IoT, with revenues potentially being generated through data products and services at the different impact layers.

Table 45 Data Economy Value by Industry, 2019–2021, 2025 (€M)

	2020	2021	2025	2020 Industry Distribution	2021 Industry Distribution	2025 Industry Distribution
Industry	Euro (M)	Euro (M)	Euro (M)	%	%	%
Agriculture	1,943	2,013	2,325	0.5%	0.5%	0.4%
Construction	3,210	3,282	4,069	0.8%	0.7%	0.7%
Education	9,027	9,408	11,627	2.1%	2.1%	1.9%
Finance	98,404	102,399	134,580	23.3%	23.1%	22.3%
Health	15,415	16,410	23,544	3.7%	3.7%	3.9%
Information & communications	42,022	43,541	57,843	10.0%	9.8%	9.6%
Mining & manufacturing	87,592	90,321	116,849	20.8%	20.4%	19.3%
Professional services	36,532	38,728	57,553	8.7%	8.8%	9.5%
Public administration	48,653	53,475	82,318	11.5%	12.1%	13.6%

Retail & wholesale	47,014	48,407	56,274	11.1%	10.9%	9.3%
Transport	16,049	16,629	23,544	3.8%	3.8%	3.9%
Utilities	16,049	17,959	33,572	3.8%	4.1%	5.6%
Total EU27	421,911	442,572	604,099	100%	100%	100%

Source: European Data Market Monitoring Tool, IDC 2021

The highest share of the data economy's value is generated in the finance sector (23% in 2021), where great focus is placed on open banking, security solutions, digital payments, and biometrics. The second largest share is represented by the manufacturing and mining industry (around 20% in 2021): Here, the data economy value is driven by a focus on automation, remote collaboration, and agility needs. Public administration is the third largest industry (12% in 2021), generating value in the data economy, which is an important result with respect to the previous publication. Indeed, thanks to NGEU's strong focus on the modernisation of public administrations, public administrations will invest in modernising digital services, digital connectivity, cybersecurity, and digital identity. This will increase the opportunity to generate value through the efficient use of data to deliver public services. Finally, the retail & wholesale industry accounts for 10%: Retailers will focus on hyperconnected and omnichannel customer experiences, breaking data silos, and integrating offline and online customer data to maintain healthy revenue streams and ensure customer loyalty.

7.6 Data Economy by Country

In terms of countries' contributions to the data economy, France, Germany, Italy, Spain, and the Netherlands tend to contribute the most. The NGEU again plays a significant role, as around 50% of total resources will be distributed across the four biggest countries in the EU27, thus making a significant difference in the next five years. Indeed, France and Germany are also among the fastest growing data economies, together with Croatia, Estonia, Finland, the Baltics, Romania, Slovakia, and Poland, among others.

All sectors are affected in the context of the currently ongoing digital transformation of the economy and society under the influence of innovative technologies and the global megatrends of the digital era. From smart transportation initiatives to the modernisation of the public sector, companies will be asked to fulfil the need for digital skills and digital workplaces in a more digitalised world.

Here are some of the initiatives under the national Digital Agenda strategies, as well as investments in digitalising national and local governments and healthcare for the abovementioned countries. All the initiatives listed below are strongly linked to the development of the skills and capacities of the labour force, as the increasing adoption of high-level technologies will require a step ahead – both on an infrastructure level and on a workforce level.

- In Croatia, the strategy focuses on supporting the development of digital skills and jobs for all citizens, and more advanced skills for digital experts. Its first strategic objective is to enhance the digital transition of society and the economy. On the public sector level, the digital transition of the public administration aims at the digitalisation of the justice system, the interoperability of the government's information systems, the deployment of the Digital Identity Card, and the creation of a one-stop-shop for all public administration' online services.
- One key focus of the Estonian Digital Agenda is to promote innovation in the field of online governance, cybersecurity, and information society. The focus is to enhance the response of small and medium-sized enterprises and public institutions to malicious attacks, while sharing good practices in information security.

- In France, the digitalisation of public administration aims to improve the efficiency of public administration and the quality of the working environment of public officials. Moreover, key investments are planned under the NGEU for the modernisation and renovation of the healthcare system to ensure the daily functioning of the healthcare offering, as well as modernisation through the digital transformation of services. Investment areas include e-health services, data interoperability, data exchange between national and European administrations, and the creation of European common data spaces (including the European Health Data Space).
- The Lithuanian Industry Digitization Roadmap 2019–2030 focuses on digital competences and skills to be developed to assist companies in the creation, adoption, and implementation of digitalisation solutions. The strategy includes the provision of digital and high-tech skills for the labour force, especially in the context of the industry sector, to increase the number of employees working in high-tech companies.
- In Slovakia, the 2030 Strategy sets the goal of modernising the Slovak economy with innovative and environmentally friendly industrial solutions, enabled through a knowledge-based data economy by 2030. This includes actions to improve the efficiency of public administration services, ensuring the smart use of existing infrastructure, as well as enhancing citizens' potential to live active digital lives and fully reap the benefits of technology.

7.7 Key Findings

The value of the data economy for the EU27 has been estimated to have reached almost €400 billion in 2020 and €440 billion in 2021. The estimated share of overall impacts as a part of GDP in the EU27 rose from 3.1% in 2019 to 3.6% in 2021. We estimate that the data economy for the EU27 in 2025 will reach €600 billion and €1 trillion in 2030.

Another important result is the change in the composition of impacts: From 2021 to 2030, the induced impacts share will increase from 34% to 39% at the expenses of indirect impacts, which will diminish from 50% to 45%.

Thanks to the NGEU's strong focus on the modernisation of public administrations, public administrations will invest in modernising digital services, digital connectivity, cybersecurity, and digital identity.

In terms of countries' contributions to the data economy, France, Germany, Italy, Spain, and the Netherlands tend to contribute the most to the data economy in the EU27. The NGEU again plays a significant role, as around 50% of total resources will be distributed across the four biggest countries in the EU27, making a significant difference in the next five years. Indeed, France and Germany are also among the fastest growing countries, together with Croatia, Estonia, Finland, the Baltics, Romania, Slovakia, and Poland, among others.

8. Data Professionals Skills Gap

The **data professionals skills gap** indicator captures the potential gap between the demand and supply of data professionals in Europe.

Monitoring the skills gap is of paramount importance, since the lack of skills may become a barrier to the development of the data industry and the rapid adoption of data-driven innovation. It is based on a model that balances the main sources of data skills (the education system, retraining, and other carriers) with the estimated demand (from all data companies).

For the data skills gap the data is provided as always for the 5 largest EU countries and the rest of EU 27 in an aggregated way, mainly because of the difficulty to measure data skills job vacancies for each individual Member State.

Table 46 Indicator 6 – Data Skills Gap

Indicator 6 – Description				
N.	**Name**	**Description**	**Type and Time**	**Segmentation**
6	Data Professionals Skills Gap	Gap between demand for and supply of data technical and data business professionals (not segmented)	Absolute number and % of total demand, 2019–2021. Forecast to 2025 (Baseline scenario). Forecast to 2030, 3 scenarios.	By Geography: 5 EU MS: DE, ES, FR, IT, PL Rest of EU27 Total EU 27 UK Switzerland EEA (NO, LI, IS)

8.2 Measuring the Data Professionals Skills Gap

Measurement Approach

The measurement of this indicator is based on a model that combines the separate estimates and forecasts for the demand for data technical and business professionals and the supply of corresponding data skills by the inflow from the education system and the upskilling and reskilling of the existing workforce. This includes balancing the main sources of data skills (from the education system and retraining to the provision from other careers) with the estimated demand for data skills (from all data companies).

More specifically, we use the following definitions:

- **The supply of data professionals** is equal to the data skills supply stock (the sum of employed data professionals and unemployed data professionals).
- **The demand for data professionals** is the sum of existing and open positions for data technical professionals and data business professionals – that is, the number of currently employed data professionals (indicator 1 in this study) plus the unfilled vacancies.
- **The indicator measures the difference between total demand and supply**; if demand is higher than supply, there is a data skills gap (excess demand). If supply is higher than demand, there is over supply and potentially unemployment.

Data Sources

As for the other indicators, the study team will carry out annually ad-hoc desk research on data skills supply and demand dynamics. The main sources that will be considered are (but not limited to):

- ILOSTAT (International Labour Organization) Statistics and Databases (2020)
- EUROSTAT Educational enrolment statistics (last update: 2021).
- IDC's Technology Employment Impact Guide – updated on a semi-annual basis with forecasts of employment across 40 technology job roles, including seven data management and analytics roles (last update: June 2021)
- Cedefop – Skills-OVATE data for vacancy estimations (most recent data collected between July 2018 and September 2020)
- Cedefop – Skills Index and Skills forecast (last update: 2020)

Measurement of Demand

The total demand for data technical and business professionals is calculated for the years 2019, 2020, and 2021 in the first cycle and will move forward one year for each measurement cycle. For the current year of the indicator (in the first measurement, it will be 2021) we have added to the number of data professionals sourced from Indicator 1 an estimate of existing unfilled positions (vacancies). The labour market is a dynamic environment characterised by inflows and outflows of human resources, and there are companies looking to hire at any given moment, as well as unemployed people looking for work. Our model includes estimates of these inflows and outflows due to retirements, sickness, deaths, graduations, career changes between companies, industries and job roles, and people entering or exiting the market for training or education activities.

To estimate the current vacancies, we have carried out additional data collection on job search portals such as LinkedIn to calculate the level of demand for data skills jobs, defined on the basis of the desks research and analysis calculated for Indicator 1 on data professionals. IDC's ongoing research on the demand for advanced ICT and data skills has been leveraged to support the forecasts. In addition, the survey has provided data about companies' difficulty in filling specific data professional positions. This has helped to model the demand forecast and to understand the level of the potential data skills gap.

The forecast demand for data professionals to 2025 (Baseline scenario) and 2030 under the three scenarios calculated by Indicator 1 is considered the total potential demand (as it incorporates future potential vacancies).

Measurement of Supply

Supply has been estimated by aggregating the number of graduates in the relevant disciplines corresponding to the data skills identified in Indicator 1 and the level of inflows from other careers and upskilling. The model considers inflows and outflows in the data skills market, such as retirements and unemployment.

Since we have changed the definition of data professionals from that used in previous years, in this first phase of the study, we have also updated the types of data skills to be monitored and the types of fields of study providing them. To do so, we have leveraged desk research, but also expert interviews.

The relationship between skills demand and supply and the resulting skills gap or over-supply is illustrated in Figure 53.

Figure 53. The Data Skills Demand-Supply Balance Model

Source: European Data Market Monitoring Tool, IDC 2021

The supply and recruitment of data professionals come from a wider variety of backgrounds than traditionally within technology-related roles. IDC has interviewed a number of organisations throughout Europe, and these bear witness to how the roles of business analyst, data scientist and even data analyst have become more related to business than to technology. Business professionals tend to develop skills in data science to manage emerging business models and products. The recruitment of data professionals from universities are therefore from a broader fields of study than the traditional STEM programs, and organisations recruit data professionals into a wider range of roles.

8.3 Data Professionals Skills Gap

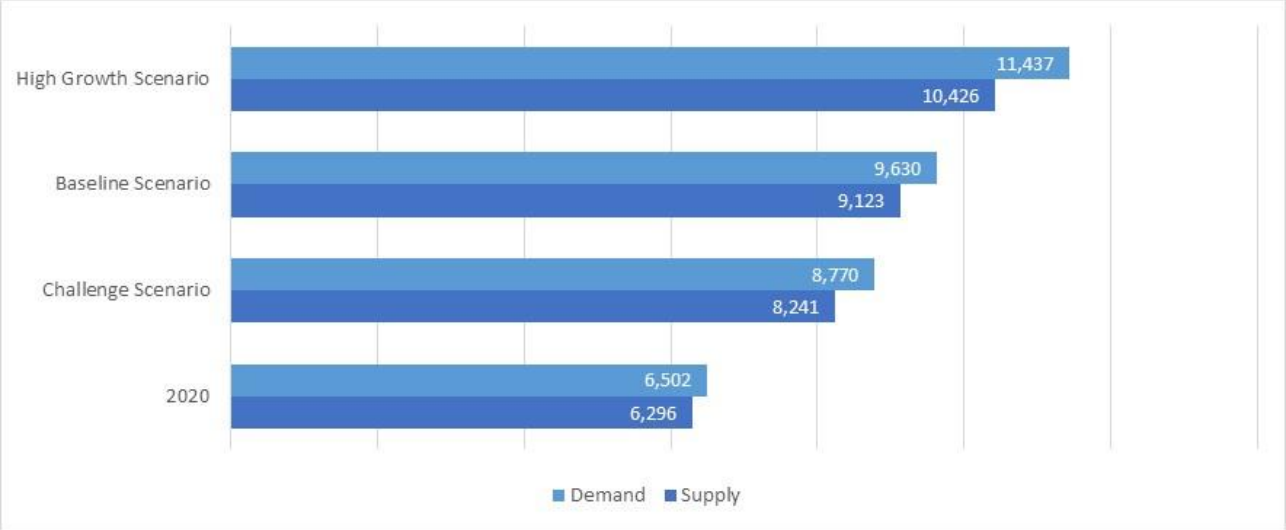
The skills gap for data professionals is growing rapidly, and it will expand most significantly in the 2021–2025 period (see Figure 8.2). Demand has already outgrown supply, and it is only beyond 2025 that the gap should start to decrease in any of the scenarios and countries modelled. The following sections provide more details for the European Union and the rest of the European countries.

8.3.1 Data Professionals Skills Gap by Member State

Figure 55 shows the skills gap for the EU27 in 2020 as a base year and in 2030 for the three scenarios. It is clear that in all scenarios, there is expected to be a skills gap:

- In 2020, the data professionals skills gap is estimated at 206,000 across the EU27, growing to 507,000 by 2030 in the Baseline scenario. This means that the gap will grow from 3.2% in 2020 to 5.3% in 2030.
- For the Challenge scenario, the gap will reach 529,000 in 2030, or 6%, as graduates look for alternative careers and there is a lower number of entrants from other careers.
- In the High Growth scenario, the gap will reach 1,011,000 in 2030 as the education system, reskilling, and upskilling programs will be unable to keep up with accelerating demand.

Figure 54 The Data Professionals Skills Gap for EU27: 2020 and Three 2030 Scenarios (Thousands)



Source: European Data Market Monitoring Tool, IDC 2021

Table 47 provides further details by Member State.

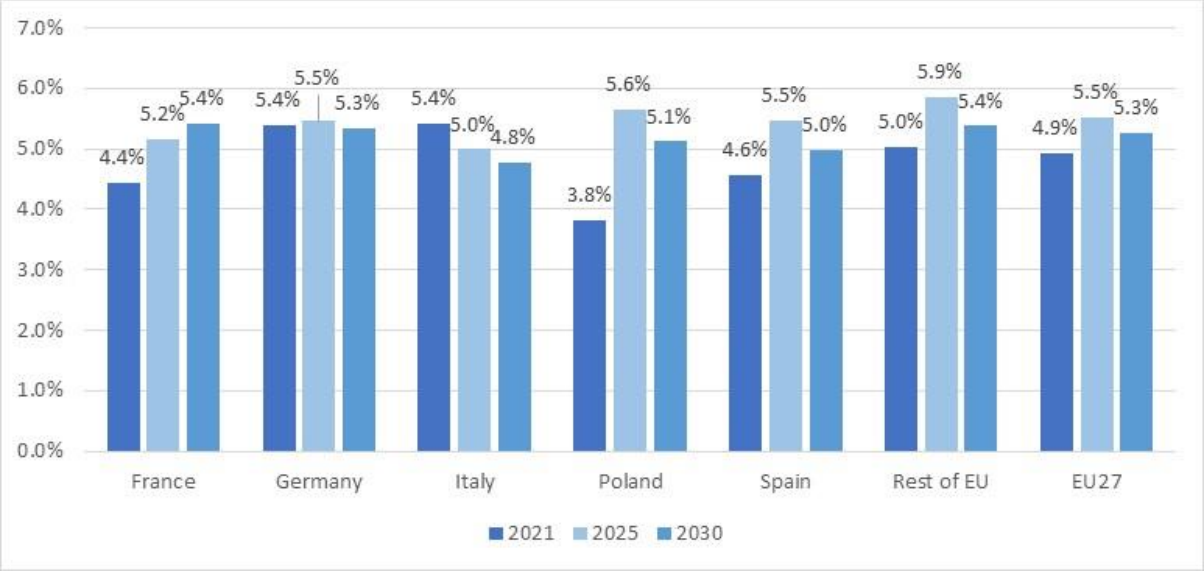
Table 47. The Data Professionals Skills Demand and Gap for the EU by Member State: 2019–2021, 2025, and Three 2030 Scenarios (Thousands)

Member State		2019	2020	2021	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR 2025-2030 Challenge	CAGR 2025-2030 Baseline	CAGR 2025-2030 High Growth
France	Numbers	12	33	46	61	81	74	160	5.9%	4.1%	21.3%
	% Gap	1.2%	3.4%	4.4%	5.2%	6.2%	5.4%	9.6%			
Germany	Numbers	42	68	92	116	187	137	259	10.1%	3.4%	17.5%
	% Gap	2.9%	4.3%	5.4%	5.5%	7.7%	5.3%	8.4%			
Italy	Numbers	12	22	37	40	50	44	105	4.6%	2.3%	21.4%
	% Gap	2.1%	3.3%	5.4%	5.0%	6.0%	4.8%	9.8%			
Poland	Numbers	13	18	22	37	34	38	87	-1.4%	0.5%	18.9%
	% Gap	2.8%	3.2%	3.8%	5.6%	5.2%	5.1%	9.8%			
Spain	Numbers	14	16	23	32	33	34	67	0.7%	1.2%	15.7%
	% Gap	3.2%	3.4%	4.6%	5.5%	5.6%	5.0%	8.3%			
Rest of EU	Numbers	42	48	118	166	143	180	334	-2.9%	1.7%	15.1%
	% Gap	2.0%	2.2%	5.0%	5.9%	4.8%	5.4%	8.5%			
EU27	**Numbers**	**135**	**206**	**338**	**451**	**529**	**507**	**1,011**	**3.2%**	**2.4%**	**17.5%**
	% Gap	**2.2%**	**3.2%**	**4.9%**	**5.5%**	**6.0%**	**5.3%**	**8.8%**			

Source: European Data Market Monitoring Tool, IDC 2021

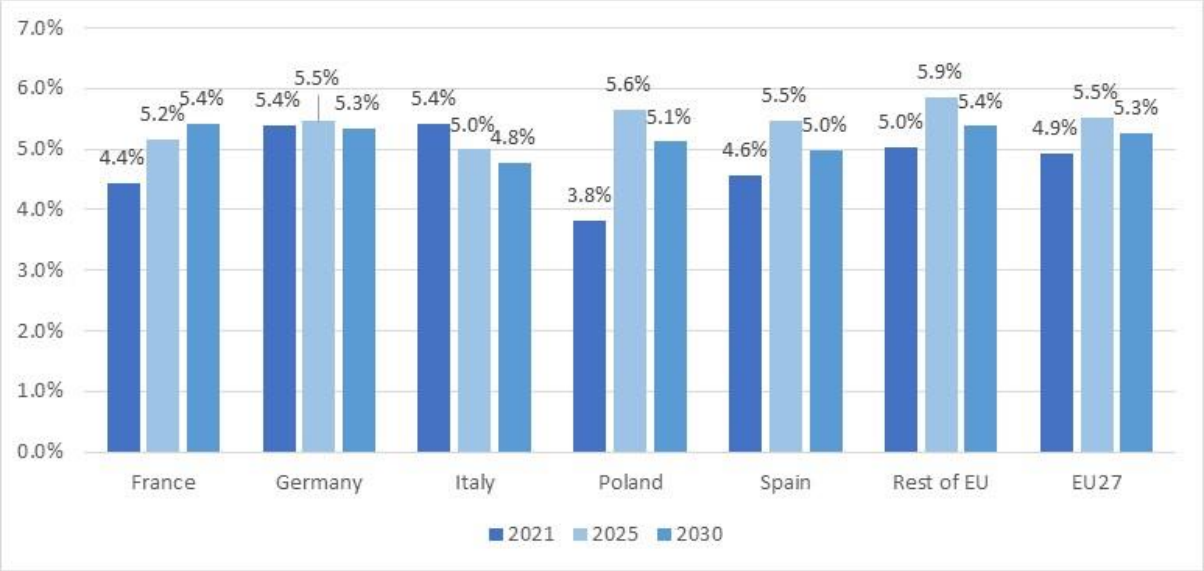
Figure 55, Figure 56 and Figure 57 provide graphic representations of the data professionals skills gap for the three scenarios by Member State.

Figure 55. Data Professionals Skills Gap for the EU by Member State: 2021, 2025, and 2030 Baseline Scenario



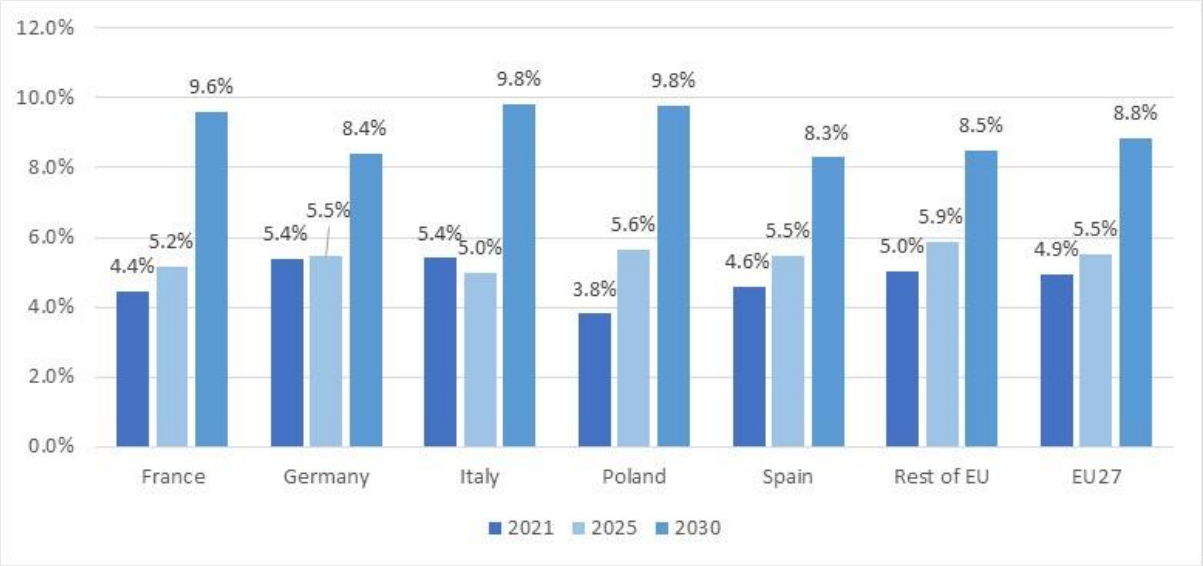
Source: European Data Market Monitoring Tool, IDC 2021

Figure 56. Data Professionals Skills Gap for the EU by Member State: 2021, 2025, and 2030 Challenge Scenario



Source: European Data Market Monitoring Tool, IDC 2021

Figure 57. Data Professionals Skills Gap for the EU by Member State: 2021, 2025, and 2030 High Growth Scenario (%)



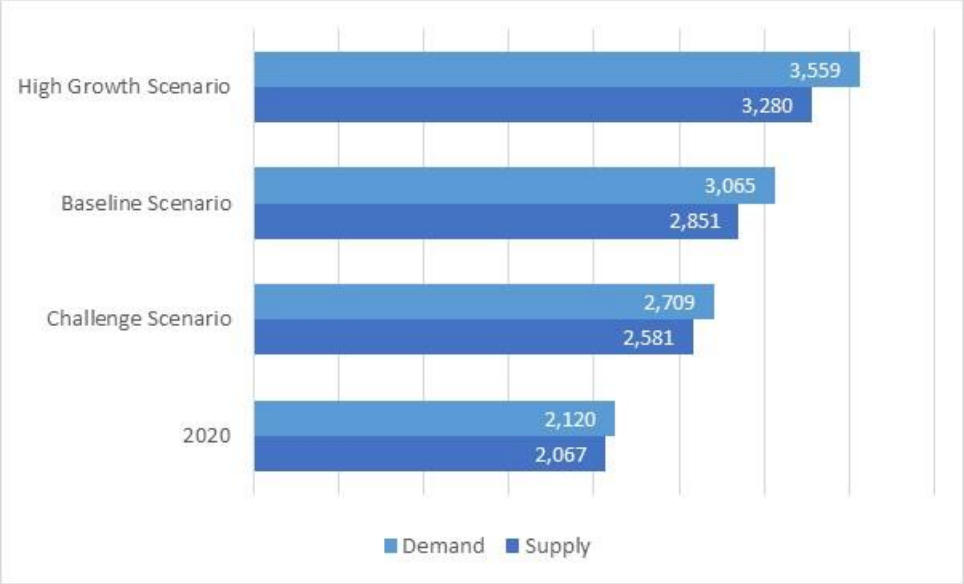
Source: European Data Market Monitoring Tool, IDC 2021

8.3.1 Data Professionals Skills Gap – the Rest of Europe

Similar to the observations for the EU Member States, the other European countries (the UK, Switzerland, and the rest of the EEA) are also experiencing a skills gap amongst data professionals and will continue to do so in all scenarios (as shown in Figure 58) – specifically:

- In 2020, the data professionals skills gap is estimated at 53,000 across the rest of Europe (outside the EU), growing to 214,000 in 2030 in the Baseline scenario. This means that the gap will grow from 2.5% in 2020 to 7% in 2030.
- For the Challenge scenario, the gap will reach 128,000 in 2030 or a gap of 4.7%. The UK represents the largest share of this country group, with a gap of 108,000 data professionals in 2030 (a gap of 4.8%).
- In the High Growth scenario, the gap will reach 279,000 in 2030, representing a gap of 7.8%. Again, this is highly influenced by the gap in the UK, which will amount to 219,000 data professionals.

Figure 58. The Data Professionals Skills Gap for the Rest of Europe: 2020–2030 in Three Scenarios (Thousands)



Source: European Data Market Monitoring Tool, IDC 2021

Table 48 provides a more detailed view of the skills gap by the other two countries and one region in the rest of Europe.

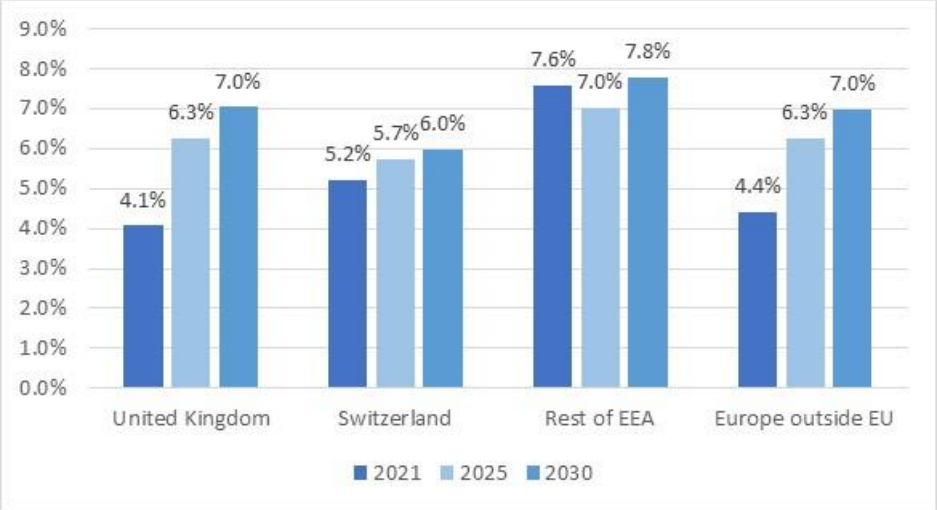
Table 48. Data Professionals Skills Demand and Gap for the Rest of Europe by Country/Region: 2019–2021, 2025, and the Three 2030 Scenarios (Thousands)

Country/ Region		2019	2020	2021	2025	2030 Challenge Scenario	2030 Baseline Scenario	2030 High Growth Scenario	CAGR 2025-2030 Challenge	CAGR 2025-2030 Baseline	CAGR 2025-2030 High Growth
United Kingdom	Numbers	39	39	77	138	108	181	219	-4.8%	4.1%	9.6%
	% Gap	2.6%	2.2%	4.1%	6.3%	4.8%	7.0%	9.6%			
Switzerland	Numbers	6	5	10	15	9	18	33	-10.2%	3.4%	17.8%
	% Gap	3.5%	2.7%	5.2%	5.7%	3.2%	6.0%	8.4%			
EEA (NO, LI, IS)	Numbers	3	9	12	13	12	16	27	-2.1%	2.3%	16.3%
	% Gap	2.6%	6.5%	7.6%	7.0%	6.3%	7.8%	9.8%			
Europe outside EU	**Numbers**	**49**	**53**	**99**	**166**	**128**	**214**	**279**	**-5.0%**	**5.3%**	**11.0%**
	% Gap	**2.7%**	**2.5%**	**4.4%**	**6.3%**	**4.7%**	**7.0%**	**7.8%**			

Source: European Data Market Monitoring Tool, IDC 2021

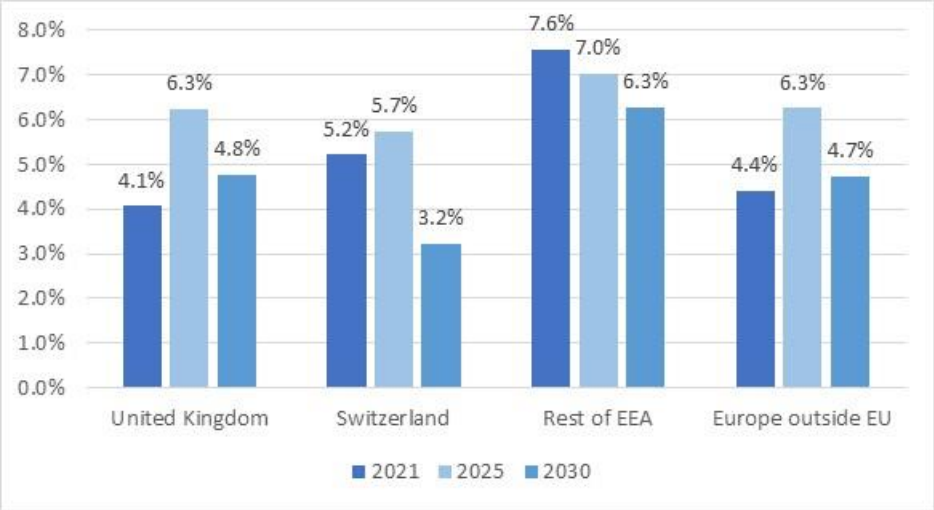
Figure 59 , Figure 60 and Figure 61 provide graphic representations of the data professionals skills gap for the three scenarios by European country/region outside of the EU.

Figure 59. Data Professionals Skills Gap for the Rest of Europe by Country/Region: 2021, 2025, and 2030 – Baseline Scenario



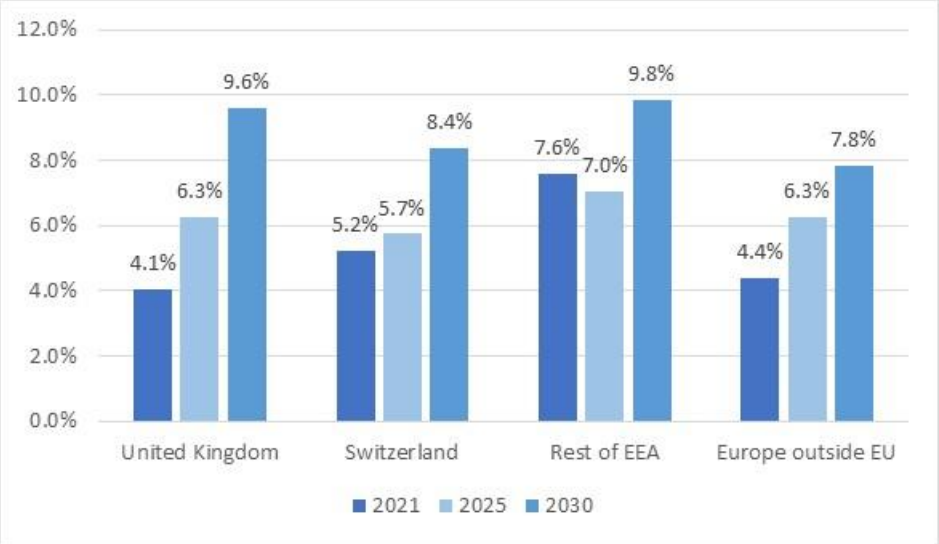
Source: European Data Market Monitoring Tool, IDC 2021

Figure 60. Data Professionals Skills Gap for the Rest of Europe by Country/Region: 2021, 2025, and 2030 – Challenge Scenario



Source: European Data Market Monitoring Tool, IDC 2021

Figure 61 Data Professionals Skills Gap for the Rest of Europe by Country/Region: 2021, 2025, and 2030 – High-Growth Scenario



Source: European Data Market Monitoring Tool, IDC 2021

8.3.2 Measuring Skills Demand

In the different forecasts that are included in this study, the investments in products and technology, which feed the demand for data professionals, continue to grow with 3-8% annually in the period up to 2025, depending on industry and technology. The 2020/2021 pandemic also showed that the underlying strength of digital transformation was such that growth continued in that period, also in terms of demand for data professionals.

Growth in demand for data professionals do not follow directly the growth in demand for data-related technology; the model for data professionals are different and factor in the maturity of the technology investments and the relation between technology investment and the human resources required to implement and operate it.

Therefore, this study shows a continued significant growth in demand for data professionals, and this growth is robust even in the challenging scenario with a weakening world economy.

8.3.3 Measuring Skills Supply

The scarcity of data professionals is a reality for most organisations throughout the EU member countries, and IDC research has clearly shown that it is already impacting the digital transformation projects across EU. IDC measured the impact of skills shortages in larger European digital transformation programs in April 2021, and among the KPIs measured was the delay due to lack of skills and experience – on average this delay was 8.1 months. According to the same survey, the perceived impact on business and customer satisfaction is significant, and this has, in turn, led to significant investments in up-skilling and re-skilling of technology and business professionals to meet the demand. In 2020, it is estimated that reskilling added around 4% to the supply of data professionals, especially in business analyst and data scientist roles, and comes from re-skilling of process-savvy business professionals. However, as organisations become more data driven, this is expected to increase.

As the baseline scenarios show, additional recruitment and re-skilling does not compensate for the growth in demand in the next four years. Organisations will therefore experience a shortage that will

impact efficiency and the ability to generate value, and these organisations will instigate additional measures to overcome the shortage. For organisations with less ability to recruit outside the traditional target groups, and with less ability to re-skill and leverage on skills outside the traditional ICT skills, this shortage will become a bottleneck in digitalisation and operational efficiency programs. As with any resource in shortage, it is anticipated that there will be investments in assessment of data proficiency skills throughout organisations, in matching supply of skills vs demand from projects and lines of business and of programs for continuous development of data proficiency skills. In addition, as with any resource in shortage, it is anticipated that there will be an increase in re-use of skills across organisational silos, thereby increasing lateral movements across organisations and cross-fertilisations across departments.

8.3.4 The Nature of the Scenarios

It is important to note how the skills gap has been calculated in the three scenarios above. The challenge scenario models negative impact on both supply and demand side. Factors impacting the supply side negatively are, for example and in this model, less recruitment and/or graduation into university programs leading up to a career as data professional, higher attrition in the industry due to lesser employee engagement or higher competition from other careers or roles, or lesser investments in reskilling programs for business professionals seeking a career as data professional. Factors impacting the demand side negatively are similarly macro-economic factors impacting the investment levels in data science, decision support and operational process automation.

Similarly, the high growth scenario models positive impact simultaneously on both the supply and the demand sides.

8.4 Key Findings

The gap between supply and demand for data professionals continue to grow. In the baseline scenario, the gap for EU27 is estimated at 2.5% of the total number of data professionals in 2020, growing to 5.3% in 2025 and 7% in 2030. The situation is similar in the European countries outside EU. This growth is robust even in the challenging scenario with a weakening world economy.

The underlying driver for the shortage is the significant growth in investments in technology, in turn driven by the digital transformation. Organisations throughout Europe will remedy the situation by a combination of continuous training, re-skilling and recruitment, but the estimates show that the gap will remain and that other measures, including re-skilling of additional business roles outside data analysis and ICT, and adding data analysis skills to the curriculum of additional university programs, will be suggested.

Key Findings – Skills Gap

The skills gap for data professionals already has a significant business impact on organisations in Europe. Business and technology are in rapid development, largely due to ongoing digital transformation of enterprise and society, but also due to a very strong pace of development in technology. A transformation of this magnitude requires large amounts of human capital and skills before the process changes stabilize and become mainstream. The reported skills gap in data professionals is therefore impacting both existing and coming

transformation initiatives and becomes one of the more serious challenges facing the European industries.

9. The International Dimension of the Data Economy

9.1 Measuring the Data Economy Beyond the EU

This report covers the data economy beyond Europe; it includes a specific section on four additional non-European countries:

- **The United States**: The top trading partner of the EU27 and other European countries.
- **Brazil**: An upper-middle-income economy and a country with among the highest ICT Development Index (IDI) scores in Latin America.
- **Japan:** The largest high-income economy in the Asia/Pacific region, and the main Asia-EU trading partner after China. Japan is a mature ICT market with many similarities to the European market.
- **China:** An emerging world force in ICT, which extends to the data economy. The second largest economy in the world and the main Asia-EU trading partner. China represents a very large market opportunity in terms of data consumption, data tools, and data monetisation.

For the international dimension, we kept the international focus on a restricted set of core indicators due to the wide disparity of the available statistical sources for these EU partners.

Specifically:

- Indicator 1.1: the number of data professionals
- Indicator 1.2: the employment share of data professionals
- Indicator 2.1: the number of data companies
- Indicator 3.1: the revenues of data companies
- Indicator 4.1: the value of the data market
- Indicator 4.2: the value of the data economy
- Indicator 4.3: the impact of the data economy on GDP

Definitions

The indicators for the four selected EU international partners leverage IDC databases available at worldwide level. Data such as ICT spending is available for most countries worldwide and is gathered with the same approach across the board. These data series are perfectly comparable at international level.

The economic model used to define the direct and indirect impacts of the data market was refreshed in 2021, resulting in a significantly higher value for indirect impacts. This is partially due to the redefined model but also the extension of the perimeter for data suppliers. Increasing industry coverage from J & M segments to include A, C, E, G, H, J, K, M, and P added significantly to the number of companies included, even if the size of the data revenues was not notably increased. The indirect impacts were accelerated by the addition of these companies in the model.

IDC also used existing data and desk research to estimate the key metrics for the four countries, accepting that the categorisation of data professionals, data companies, and revenues of data companies might not be as current as for the EU Member States. We used IMF forecast data and available statistics to validate our estimates. This is coupled with IDC's existing data sources of IT spending in the IDC Black Book and IDC Spending Guides, both of which include spending for the four

four countries beyond the EU. The table below outlines the main sources used to estimate the international indicators in this report.

Table 49. Internationals – Main Data Sources

Data Source	Updated
Consensus Forecasts – Consensus economics	Jul 2021
IDC Core IT Spending Guide 2H2021	Jun 2021
IDC Worldwide Black Book v3.2 (standard edition)	Sep 2021
IMF World Economic Outlook (Oct 2021)	Aug 2021
ILOSTAT statistics and databases	Mar 2021
IT Big Data and Analytics Spending Guide 2H2021	Sep 2021
US Census Data – number employed by industry	Oct 2021
China Statistical Yearbook 2019	Jan 2021
CIA World Factbook	Ongoing
UHY	Dec 2017

9.2 The US

Table 50.US Indicators – 2019–2021 Overview

Name	Metrics	2019	2020	2021	Growth Rate, 2020–2021
Number of data professionals	Total number of data professionals ('000s)	14,350	15,275	16,169	5.9%
Data professionals' employment share	% of data professionals on total employment	6.16%	6.44%	6.63%	3.0%
Number of data suppliers	Total number of data supplier companies ('000s)	312,215	315,857	321,847	1.9%
Revenues of data companies	Total revenues generated by companies specialised in the supply of data-related products and services (€M)	184,873	213,463	239,958	12.4%
Value of the data market	Estimate of the overall a value of the data market (€M)	184,873	213,463	239,958	12.4%
Value of the data economy (only direct and backward indirect impacts)	Direct impacts (€M)	184,873	213,463	239,958	12.4%
	Backward indirect impacts (€M)	123,480	163,296	232,101	42.1%
Impact of the data economy on GDP (only direct and backward indirect impacts)	Ratio between value of the data economy and GDP (%)	1.19%	1.26%	1.31%	4.3%

The US is the strongest of the data economies when considering the internationals, which is unsurprising considering the foundation of data in business lies within the US. It is strong particularly in tools and software, but the size and strength of the economy supports a robust data economy. Many data suppliers are US based, particularly the leading proponents in the market. The US is considered the world leader in areas such as artificial intelligence, and companies' investments in this are growing dramatically according to IDC's Artificial Intelligence Spending Guide. The US also has a strong focus on big data and analytics, and many, if not most, of the companies active in this technology are US based. The US is also strong in cloud technology, a foundational technology for data, with the leading cloud suppliers (AWS, Microsoft, and Google) being based in the US. As these organisations currently dictate the development of data tools and technologies, the US has a notable lead in the development of the market. However, China is emerging as a significant player in the global

digital arena, particularly in artificial intelligence, 5G telecommunication networks, next generation internet, and high performance & quantum computing, to name a few key data-driven technology areas. This could pose a threat to the current US leadership in the data market and data economy fields. Future developments will depend a lot on the level of investment seen by the Chinese government, which is continuing to make significant progress in a number of data-driven technologies, which will be pivotal for the evolution of the data market and its effects on the economy as a whole.

Figure 62. US Annual Growth Rates 2021–2020 – Key Metrics

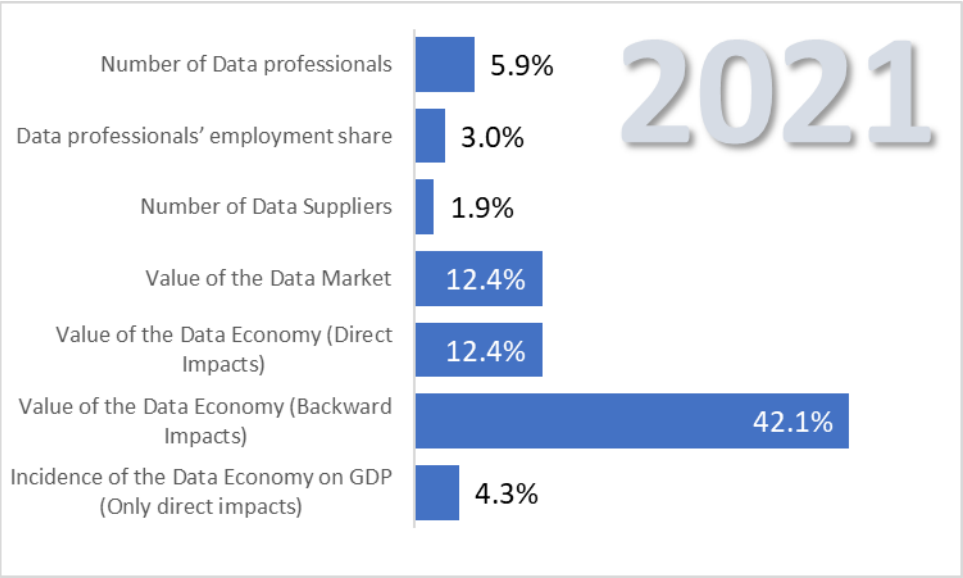


Figure 62 shows the impact of data on the US economy in 2021. The direct and backward impacts are forecast to record notable growth for 2021, at 42.1% year on year. This indicates that the data economy in the country has achieved a notable level of maturity, with effects now spreading not only to data companies and their associated products and services but also to the entire value chain associated with the functioning of the data market. The infrastructure needed to support the data economy is also growing, with the number of data professionals and the number of data suppliers forecast to record growth for 2021.

9.3 China

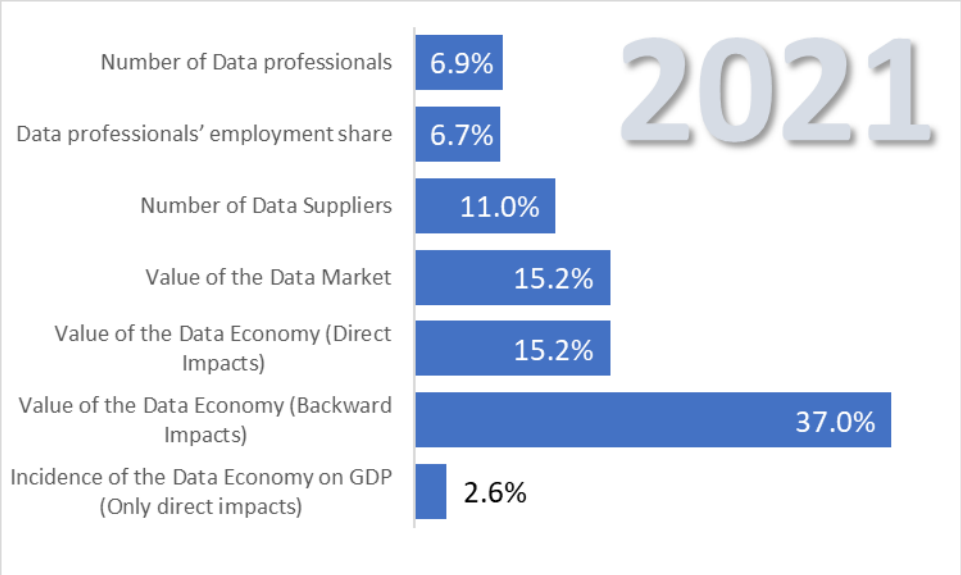
Table 51. China Indicators – 2019–2021 Overview

Name	Metrics	2019	2020	2021	Growth Rate: 2021–2020
Number of data professionals	Total number of data professionals ('000s)	8,717	9,184	9,815	6.9%
Data professionals' employment share	% of Data professionals on total employment	1.11%	1.19%	1.27%	6.7%
Number of data suppliers	Total number of data supplier companies ('000s)	756,002	858,509	952,566	11.0%
Revenues of data companies	Total revenues generated by companies specialised in the supply of data-related products and services (€M)	24,225	27,470	31,651	15.2%
Value of the data market	Estimate of the overall a value of the data market (€M)	24,225	27,470	31,651	15.2%
Value of the data economy (only direct and backward indirect impacts)	Direct impacts (€M)	24,225	27,470	31,651	15.2%
	Backward indirect impacts (€M)	25,171	31,062	42,561	37.0%
Impact of the data economy on GDP (only direct and backward indirect impacts)	Ratio between value of the data economy and GDP (%)	0.76%	0.82%	0.84%	2.6%

The size of the market and the size of the population in China dominate any estimates of the size of the data market, and the high growth seen in the country's economy – with GDP growth remaining around 6%, compared with EU growth rate around 2% – ensures it will remain a strong market in the foreseeable future. Close to 28% of the labour force is in agriculture, compared with only 5% of the EU27 and 0.7% of the US labour force. This clearly has a direct impact on the share of the labour force associated with data and data services. Only 43% of the Chinese labour force is employed in service industries, compared with 73% in the European Union.

China is still undergoing a transition from an agricultural economy to an industrial one, which affects the estimates of the number of data suppliers in the country. Although the total number of companies is large, many of them are small organisations and are unlikely or unable to invest in full digital transformation. In spite of this, China's position in the world is a strong one, so it still manages to dominate the market in terms of the number of data supplier companies. Indeed, while China's total number of companies fell in 2018 and 2019, the share of companies that are to be classified as data companies grew over the same period. Table 51 and Figure 63 summarise the data for China. In particular, we anticipate high growth for the data economy in China because of its overall size, its global position, and its potential considering the relatively low share of the economy that is geared towards business and data services.

Figure 63. China Annual Growth Rates, 2020–2021 – Key Metrics



9.4 Brazil

Table 52. Brazil Indicators – 2019–2021 Overview

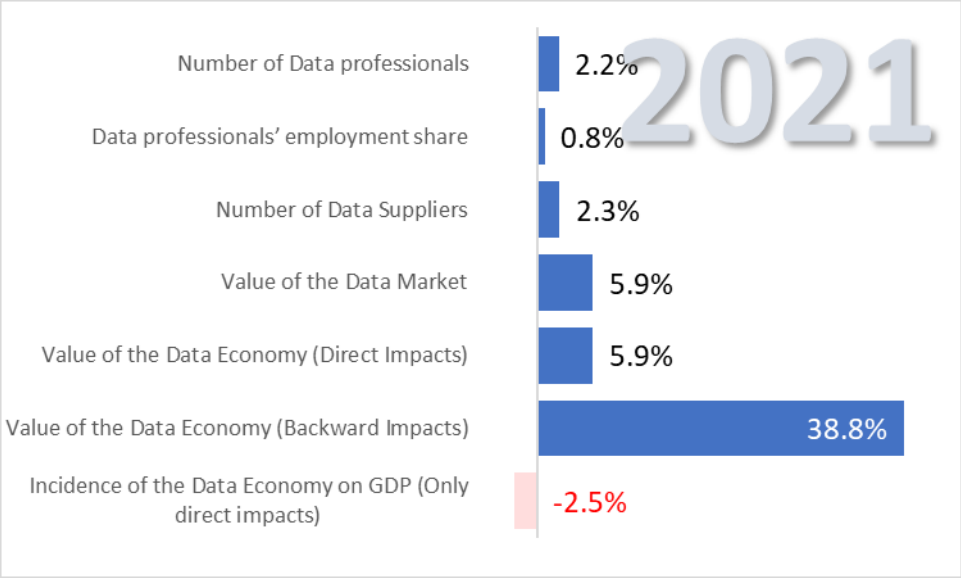
Name	Metrics	2019	2020	2021	Growth Rate: 2021–2020
Number of data professionals	Total number of data professionals ('000s)	1,211	1,244	1,272	2.2%
Data professionals' employment share	% of data professionals on total employment	7.56%	7.73%	7.79%	0.8%
Number of data suppliers	Total number of data supplier companies ('000s)	38,192	39,606	40,518	2.3%
Revenues of data companies	Total revenues generated by companies specialised in the supply of data-related products and services (€M)	7,905	8,374	8,865	5.9%
Value of the data market	Estimate of the overall a value of the data market (€M)	7,905	8,374	8,865	5.9%
Value of the data economy (only direct and backward indirect impacts)	Direct impacts (€M)	7,905	8,374	8,865	5.9%
	Backward indirect impacts (€M)	6,370	7,812	10,841	38.8%
Impact of the data economy on GDP (only direct and backward indirect impacts)	Ratio between value of the data economy and GDP (%)	0.23%	0.20%	0.20%	–2.5%

Brazil's economy: GDP growth substantially declined in the second half of 2020, and inflation is on the rise. All in all, the country's position in the world is weakening, with its infrastructure not showing the growth required to support a vibrant growing data economy.

While revenue growth is currently forecast for 2021, investments in data professionals and their share of total employment are not as strong as in the other internationals, and the impact of the data economy on total GDP is declining slightly. Brazil's economy was hit badly by COVID-19, declining 4% in 2020. When factoring in exchange-rate changes to the US dollar, the Brazilian economy fell by more than 20% in US dollar terms. However, 2021 shows signs of recovery, as is the case for many of

economies around the world. Brazil's data economy is still mostly growing, with most of the indicators either growing or close to flat; so, there is still opportunity for the economy to expand its data focus. However, the country is the smallest data market among the internationals evaluated.

Figure 64. Brazil Annual Growth Rates 2021–2020 – Key Metrics



9.5 Japan

Table 53. Japan Indicators – 2019–2021 Overview

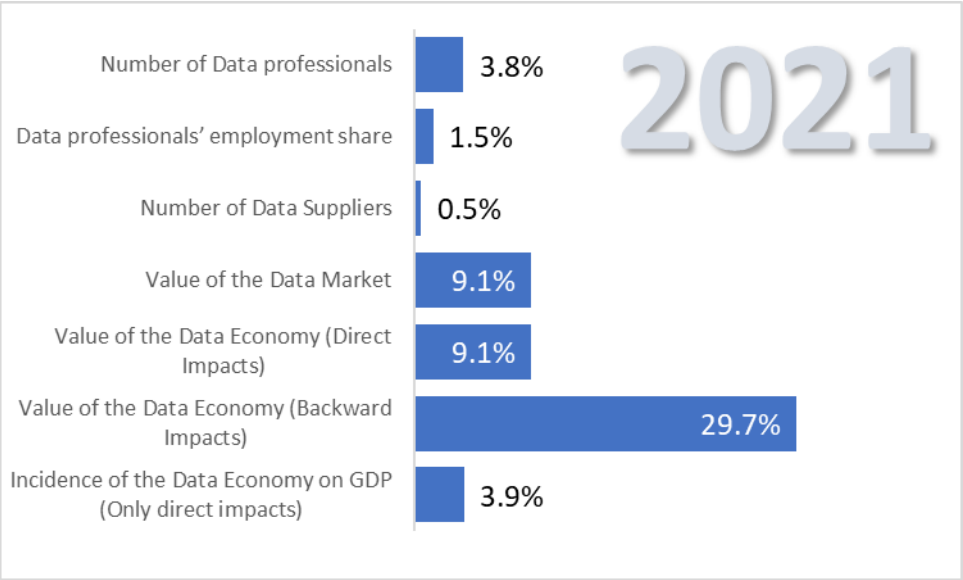
Name	Metrics	2019	2020	2021	Growth Rate: 2021–2020
Number of data professionals	Total number of data professionals ('000s)	4,236	4,398	4,567	3.8%
Data professionals' employment share	% of data professionals on total employment	5.38%	5.51%	5.59%	1.5%
Number of data suppliers	Total number of data supplier companies ('000s)	106,983	106,214	106,786	0.5%
Revenues of data companies	Total revenues generated by companies specialised in the supply of data-related products and services (€M)	32,929	36,649	39,970	9.1%
Value of the data market	Estimate of the overall a value of the data market (€M)	32,929	36,649	39,970	9.1%
Value of the data economy (only direct and backward indirect impacts)	Direct impacts (€M)	32,929	36,649	39,970	9.1%
	Backward indirect impacts (€M)	26,985	30,960	40,145	29.7%
Impact of the data economy on GDP (only direct and backward indirect impacts)	Ratio between value of the data economy and GDP (%)	1.09%	1.23%	1.27%	3.9%

Japan represents the data economy closest to the EU data economy. Its data economy is reasonably robust, showing growth in the number of data professionals, although its growth in the share of total employment is limited when compared with others (excluding China), especially the EU27.

Japan has a reasonably strong supply infrastructure, with the number of data suppliers about two thirds that of the EU27, but these are forecast to show low growth in 2021. Its infrastructure – in terms

of data suppliers and data professionals – suggests Japan will maintain its role as a key data supplier nation in the world market.

Figure 65. Japan Annual Growth Rates 2021–2020 – Key Metrics



9.5 Comparing the Internationals to the EU27

Table 54. EU27 Indicators – 2019–2021 Overview

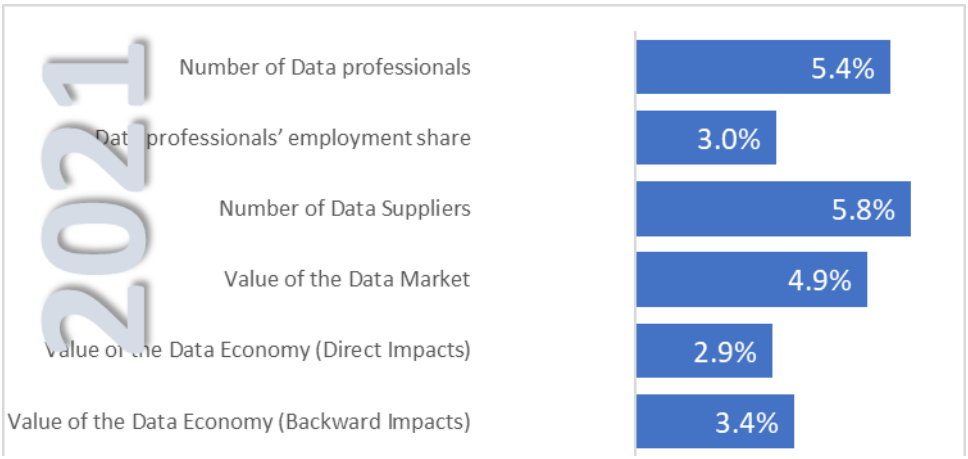
Name	Metrics	2019	2020	2021	Growth Rate: 2020–2021
Number of data professionals	Total number of data professionals ('000s)	6,026	6,502	6,853	5.4%
Data professionals' employment share	% of data professionals on total employment	3.4%	3.6%	3.8%	3.0%
Number of data suppliers	Total number of data supplier companies ('000s)	166,063	175,605	185,866	5.8%
Revenues of data companies	Total revenues generated by companies specialised in the supply of data-related products and services (€M)	58,427	60,635	63,627	4.9%
Value of the data market	Estimate of the overall a value of the data market (€M)	64,262	71,050	73,116	2.9%
Value of the data economy (only direct and backward indirect impacts)	Direct impacts (€M)	98,668	108,546	112,221	3.4%
	Backward indirect impacts (€M)	0.5%	0.6%	0.6%	-
Impact of the data economy on GDP (only direct and backward indirect impacts)	Ratio between value of the data economy and GDP (%)	6,026	6,502	6,853	5.4%

The relevant EU27 indicators are repeated in Table 54**Error! Reference source not found.** for a c omparison with the data from internationals, and the growth is summarised in Figure 66. Solid growth across all major data market indicators continues to characterise the EU27, with the number of data professionals and data companies (both suppliers and users) steadily on the rise. The data market indicator is also projected to increase in 2021, albeit at a rather slower pace. This is because Europe is slowly but progressively showing certain signs of maturity, with entire industry sectors (such as

manufacturing and healthcare) and a significant portion of its businesses (including SMEs) rapidly embracing the process of digitalisation.

The EU27 has is the third largest actor in terms of units of data professionals after the US and China, but the size of the difference to the US and the anticipated growth for China suggest that further action is needed in the EU to reduce the gap. Despite this, the EU needs to ensure it invests in sufficient infrastructure (data centres, storage facilities, and security solutions) to create an ecosystem of technology, processes, and actors/organisations that enables the effective collection, storage, maintenance, distribution, and (re)use of data by the different public and private users across all sectors of the economy. This will allow for the steady growth of the data market, thus ensuring that possible limits on the number of data professionals do not act as a brake on future growth in the data skills that organisations need to embrace the adoption of data as a basic tenet for conducting business. Figure 66 shows that growth among the EU Member States is weak in 2021 when compared with most of the other internationals under consideration. This is a sign of maturity and is evidence that the European Union has assured its position among the leading data markets and that its market is now consolidating.

Figure 66. EU27 Annual Growth Rates 2021–2020 – Key Metrics

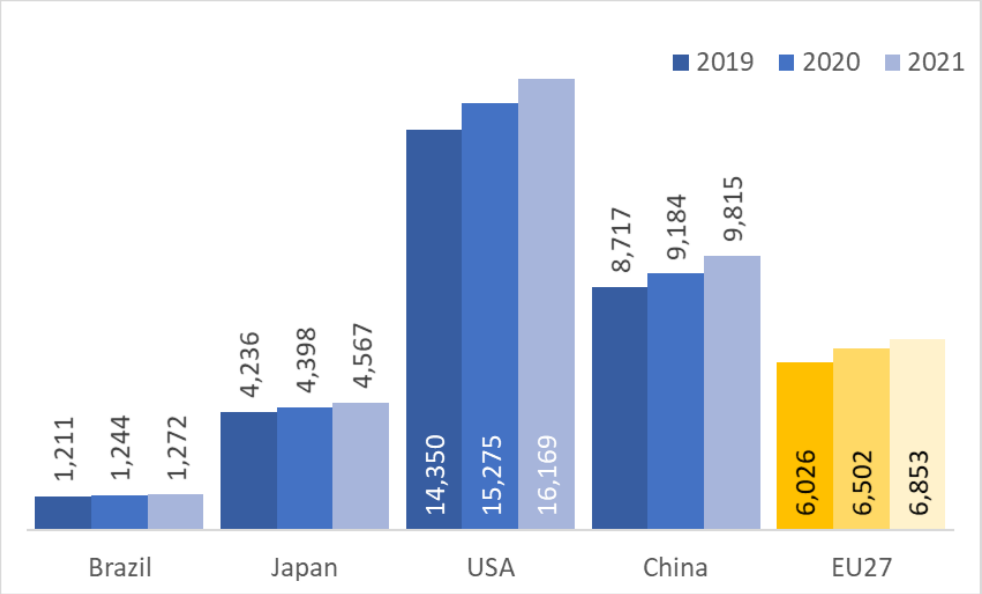


Data professionals

Table 55. Internationals – Number of Data Professionals, 2019–2021

	2019	2020	2021	Growth Rate: 2021–2020
Brazil	1,211	1,244	1,272	2.2%
Japan	4,236	4,398	4,567	3.8%
US	14,350	15,275	16,169	5.9%
China	8,717	9,184	9,815	6.9%
EU27	6,026	6,502	6,853	5.4%

Figure 67. Internationals – Number of Data Professionals, 2019–2021



The Member States' employment share of data professionals as a proportion of total employment is still relatively low – the second smallest after China (see Table 56). This suggests that, although employment growth in the European data market is high, Europe as a whole still finds it difficult to train, employ, and retain an employment base with the necessary data skills. This could also indicate that mainstream companies in the EU27 are lagging those in other regions in adopting data tools and methods for business, which could result in all businesses in Europe falling behind their international competitors.

Table 56. Internationals – Share of Total Employment, 2019–2021

	2019	2020	2021	Growth Rate: 2021–2020
Brazil	7.6%	7.7%	7.8%	0.8%
Japan	5.4%	5.5%	5.6%	1.5%
US	6.2%	6.4%	6.6%	3.0%
China	1.1%	1.2%	1.3%	6.7%
EU27	3.4%	3.6%	3.8%	3.0%

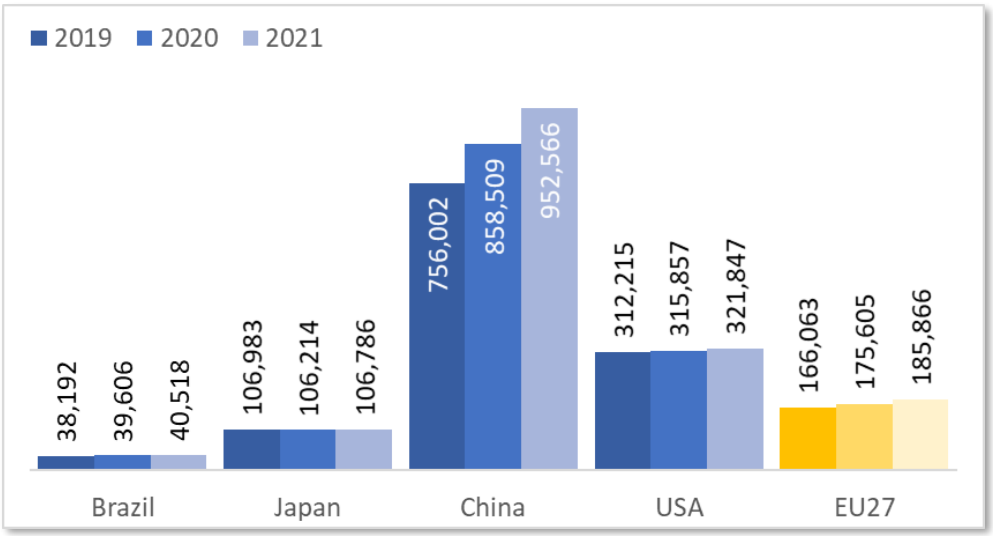
Data Supplier Companies

EU27 has a moderate share of data supplier companies (see Table 57 and Figure 68) – the third largest number after China and the US. The region should retain or even grow its presence on the world stage, as long as it is not held back by companies unwilling to invest in data transformation or by a shortage of data professionals. If insufficient companies are prepared to transform and contribute to the data economy, the risk is that new data graduates and PhD students will move outside the region for better work opportunities than the Member States can offer.

Table 57. Internationals – Number of Data Supplier Companies, 2019–2021

	2019	2020	2021	Growth Rate 2021–2020
Brazil	38,192	39,606	40,518	2.3%
Japan	106,983	106,214	106,786	0.5%
China	756,002	858,509	952,566	11.0%
US	312,215	315,857	321,847	1.9%
EU27	166,063	175,605	185,866	5.8%

Figure 68. Internationals – Number of Data Supplier Companies, 2019–2021



Data Market

The market in Europe is strong though – the second largest after the US and two thirds larger again than the next nearest competitor, Japan. This should ensure sufficient investment in data by companies in Europe and provide sufficient jobs of interest to retain locally grown data skills and encourage immigration of skilled data professionals from outside the region.

Table 58. Internationals – Value of the Data Market, 2019–2021

	2019	2020	2021	Growth Rate 2020–2021
Brazil	7,905	8,374	8,865	5.9%
Japan	32,929	36,649	39,970	9.1%
China	24,225	27,470	31,651	15.2%
US	184,873	213,463	239,958	12.4%
EU27	58,427	60,635	63,627	4.9%

The Data Economy

The overall impact on the economy is significant in the EU27 (see Table 59), and the extension of the perimeter to include a wider range of industries within the data market makes a notable contribution to growth in indirect impacts. However, the growth forecast for 2021 is among the lowest over the past few years for the EU27. Even so, the size of the impacts is among the largest among internationals. The incidence of the data economy will stumble slightly in the EU in 2021 when compared to 2020, but this is more of an artefact of different cycles combining to drop the incidence, and by 2021 the incidence will be above 0.6% again.

Table 59. Internationals – Economic Impacts (Direct and Indirect), 2019–2021 (€M)

		2019	2020	2021	Growth Rate 2021–2020
Brazil	Internationals – direct impacts	7,905	8,374	8,865	5.9%
Brazil	Internationals – indirect impacts	6,370	7,812	10,841	38.8%
Japan	Internationals – direct impacts	32,929	36,649	39,970	9.1%
Japan	Internationals – indirect impacts	26,985	30,960	40,145	29.7%
China	Internationals – direct impacts	24,225	27,470	31,651	15.2%
China	Internationals – indirect impacts	25,171	31,062	42,561	37.0%
US	Internationals – direct impacts	184,873	213,463	239,958	12.4%
US	Internationals – indirect impacts	123,480	163,296	232,101	42.1%
EU27	Internationals – direct impacts	64,262	71,050	73,116	2.9%
EU27	Internationals – indirect impacts	98,668	108,546	112,221	3.4%

Figure 69. Internationals – Economic Impacts, 2019–2021

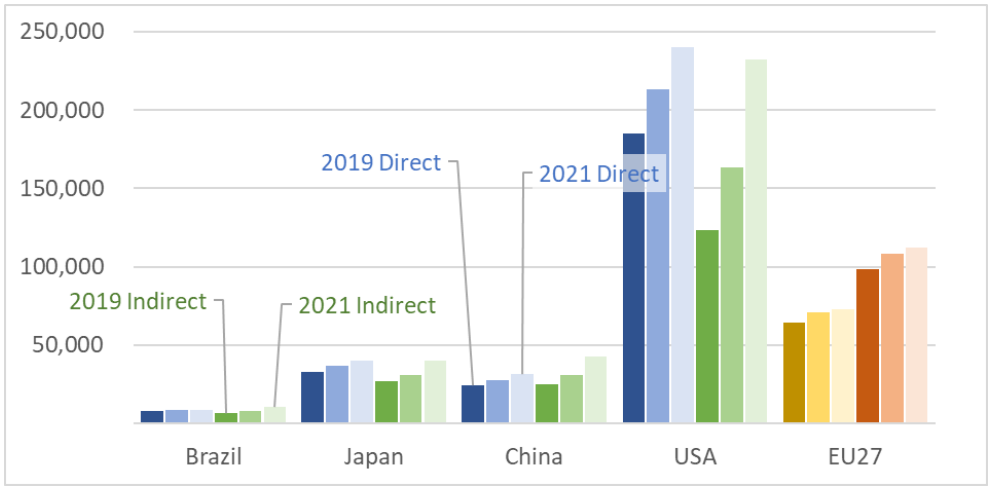
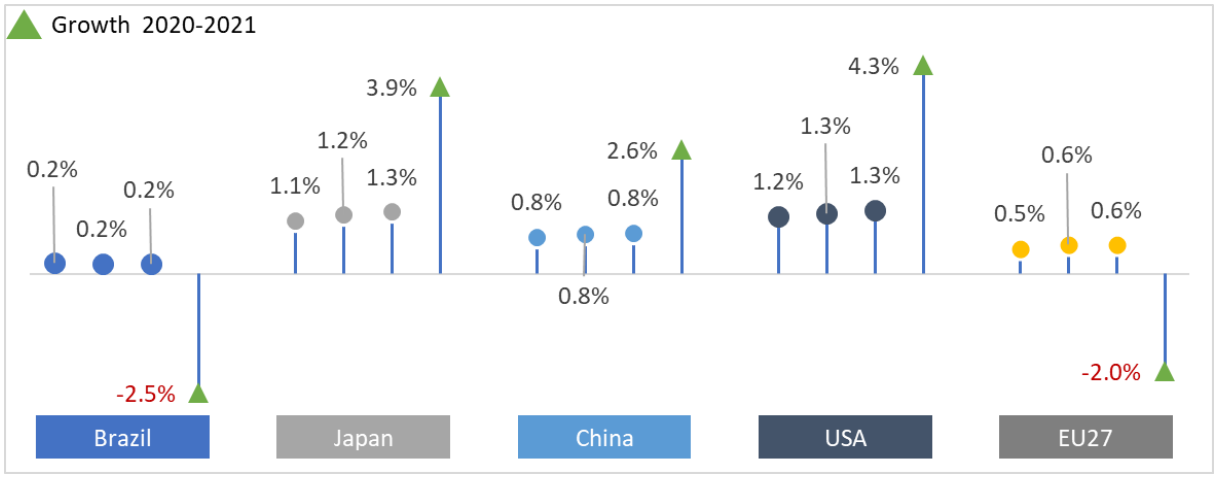


Table 60. Internationals – Impact of the Data Economy on GDP, 2019–2021 (Direct and Indirect Impacts Only)

	2019	2020	2021	Growth Rate 2021–2020
Brazil	0.23%	0.20%	0.20%	–2.5%
Japan	1.09%	1.23%	1.27%	3.9%
China	0.76%	0.82%	0.84%	2.6%
US	1.19%	1.26%	1.31%	4.3%
EU27	0.51%	0.60%	0.59%	–2.0%

Figure 70. Internationals – Impact of the Data Economy on GDP 2019–2021 (%) and 2020–2021 Growth (%)



9.5 Key Findings

The growth of the US data economy continued despite 2020 and 2021 being difficult years in the country and globally due to the COVID-19 pandemic and the subsequent socioeconomic downturn. Nevertheless, in the US, the data economy exhibited a solid 2020–2021 year-on-year growth in all the main indicators monitored, including the number of data professionals and companies and the overall data market.

China was added to the international countries monitored by the European Data Market studies for the first time and shows an already well-developed data market and data economy (comparable with those of Japan), which are expected to grow over the forecast period. Yet the country's data economy has certainly not reached its peak, and we anticipate continued growth as the country's influence spreads globally. At the moment, we believe the monetisation of data in China is weak but could strengthen as data increases in volume and value.

The US currently boasts the largest number of data professionals among the countries under consideration. Data professionals in the United States in 2021 are estimated at more than 16 million, while the nearest competitor, China, has only 9.8 million data professionals. Growth in the number of data professionals in China is higher than in the US, but the United States continues to display solid growth in the employment of data professional in the near-term, thus maintaining its advantage for the time being. Brazil's number of data professionals has slowed as the economy has weakened, while Japan showed a modicum of growth in the employment of data professionals in 2021. Brazil has the potential to play a more significant role in the data market and data economy worldwide, but the overall economy is stifling growth in the country. The Brazilian economy runs the risk of losing its skilled data professionals to other countries or regions (such as the US, the UK, or indeed the EU27), as these countries show more potential at the present stage and for the near future.

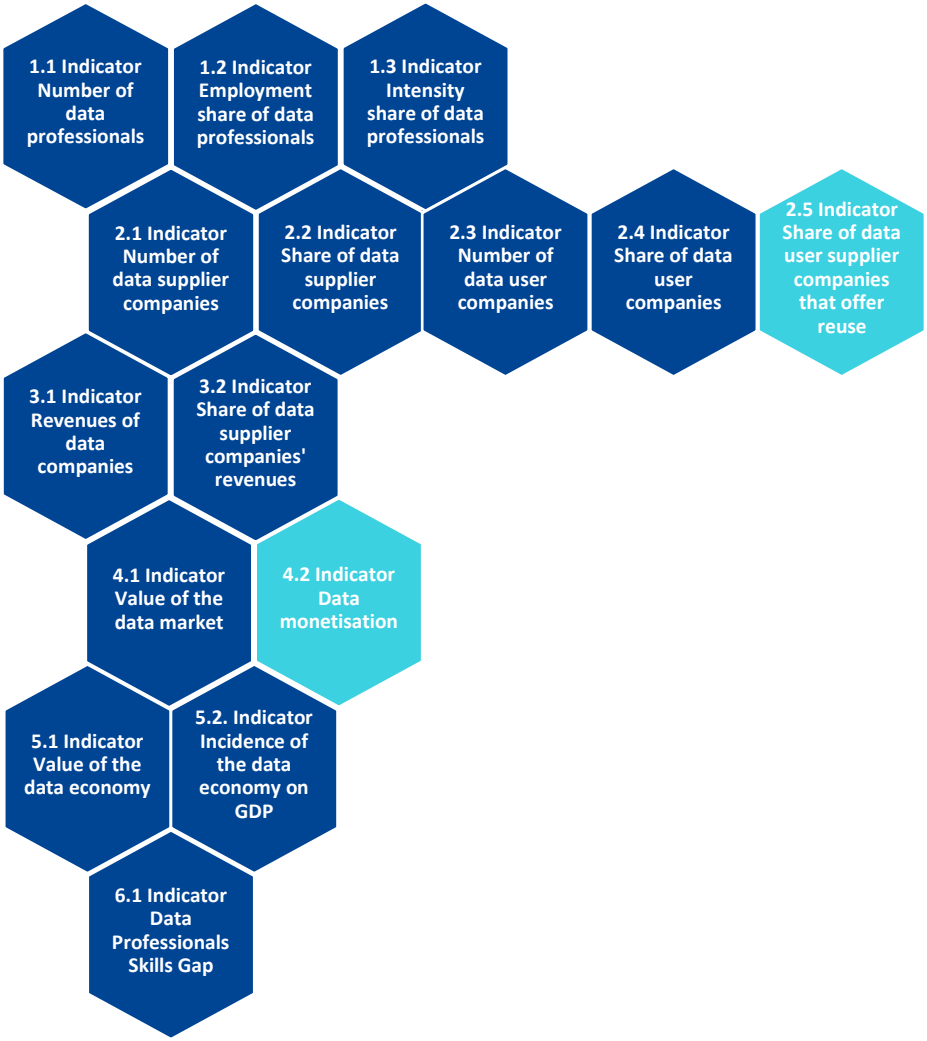
The data market in Japan is closest to the EU27 in growth and size. However, while the data market in the EU27 is slightly larger than the Japanese one, Japan nearly doubled the size of its data market in 2021. Japan also retains a larger share of employees involved in the data market, close to twice the share in the EU27. Japan's share growth (data professionals as a part of total employment) lags that of the EU27. Nevertheless, in 2021, Japan reversed the negative GDP trend of 2020, posting a significant recovery for the year. If this trend is confirmed, a positive effect on the Japanese data market and the overall data economy can be expected.

Most of economies are maintaining or increasing the impact of the data economy on GDP. Although the EU27 recorded a minor decrease in 2021, down from 0.60% in 2020 to 0.59%, the impact is expected to recover in 2022. Brazil still has a low share of its GDP attributable to the data economy – at only 0.2% compared with close to 1.0% for the remaining international economies. The US has confirmed its primacy in the data economy worldwide and shows the strongest growth in the impact of the data economy on GDP for the time being. The outlook for China is uncertain, but its investment potential in the data economy is considerable. Whether the US leadership is maintained over the medium term will depend on the strategic focus that China exerts on digital technologies and the capacity of the US to counteract that effectively.

10. Conclusions

This First Report on Facts & Figures (Deliverable D2.1) of the Update of the European Data Market Study, VIGIE 2020-0655, has presented the results obtained through the first round of measurements of the European Data Market Monitoring Tool for the period 2019–2021, with forecasts for 2030 under three distinct scenarios.

The results pertained to the following set of indicators, as per the updated version of the European Data Market Monitoring Tool:



Each indicator has been measured for the total EU27 and for all the EU27 Member States when available and applicable; industry-specific and company-size views are also offered, with indicators provided by industry sector and company size band when possible. The UK and Switzerland were measured separately in an aggregated way, as were the countries of the EEA (Norway, Iceland, and Lichtenstein). As in the previous European Data Market Studies (SMART 2013/0063 and SMART 2016/0093), a select number of indicators has been developed and updated – this time, for four non-European countries, Brazil, Japan, the United States, and China.

10.1 Data Professionals

Data professionals are workers who collect, store, manage, and/or analyse, interpret, and visualise data as their primary activity or as a relevant part of their activity. Data professionals must be proficient with the use of structured and unstructured data, should be able to work with a huge amount of data, and should be familiar with emerging database technologies. For 2021–2023, the definition of data professionals was refined to differentiate the roles played by different data users: These are data technical professionals, data business professionals, and data consumers – as defined below. The measurement of data professionals includes data technical professionals and data business professionals only.

Data professionals were redefined for this publication – clarifying the two key types of data technical professionals and data business professionals. A third type – data consumers – is not included in this dataset. As a result, the data for 2019 and 2020 is restated. Data technical professionals focus on building data models for use by data business professionals to make business decisions based on what these models show.

There will be 6.8 million data professionals in the EU27 Member States by the end of 2021, continuing the healthy growth (5.4%) seen in 2019 (5.0%).

Data professionals' share of employment has also increased in 2021 – to 4.0%, up from 3.3% in 2019 – for the EU27 Member States. For another year, this again confirms the positive trend of the share of the workforce involved in data-related professions.

According to the latest estimate, **the average number of data professionals per company is estimated at 12.4 in the EU27 in 2021, up from 11.3 in 2019.**

Data professionals in the EU27 will account for 8.1 million people in 2025 or 9.6 million data professionals according to the Baseline forecast. The number is expected to rise by a compound rate of 3.4% in the EU27 but is likely to be constrained by the limited supply of professionals.

Looking at the penetration of data professionals in total employment across all Member States in 2019, it is clear that the balance is even, aside from some unsurprising outliers, which are data-intense economies, such as Luxembourg and the UK, or economies still struggling to make the most of ongoing digitalisation (e.g.: Greece).

The number of data professionals is highest in the professional services industry, followed by retail & wholesale and information & communications. The spread of the number of data professionals by industry is notably wider, although the correlation is fair between the number of companies and employees in those industries and the number of data user companies in the same industries.

10.2 Data Companies

Data companies are organisations that are directly involved in the production, delivery, and/or usage of data in the form of digital products, services, and technologies. They can be both data supplier and data user organisations.

By 2025, data user companies will account for 633,000 units – growing from 2020 at a compound rate of 3.1%. This is about the same as the growth in the number of all companies. But, in the longer-term (2025–2030 Baseline), data user companies continue grow at a compound rate that is between 5 and 10 times that of all companies. The total number of companies is expected to increase long term at between 4% and 8% per year between 2025 and 2030, depending on the scenario.

The data user penetration rates (i.e. the number of data user companies as a proportion of total companies) vary across Member States, with the larger economies dominating the scene in terms of size but not overall penetration. It is the smaller Member States that tend to show a higher share of data users per total companies. The EU27 average penetration rate is 2.1% in 2021.

Data user penetration is as low as 0.6% for Slovakia and Poland, and as high as 7.3% for Austria. Adoption rates of data technologies are higher in industries such as professional services, retail, and financial services.

Looking at the share by industry, professional services is easily the sector with the greatest number of data user companies, accounting for 27% of all data user companies in 2021. Data is core to a large share of professional services activities. Manufacturing and retail & wholesale follow professional services, with 15.1% and 14.6%, respectively, but are significantly behind the lead industry.

Small and medium-sized companies account for 98% of all companies in the European Union in 2021. Unsurprisingly, they also account for the majority of data user companies. Larger companies, however, are those that invest more heavily in digital resources. Investment in data technologies requires considerable investment and expertise, and companies with fewer than 10 employees – a large proportion small and medium businesses – are unable or unwilling to make this investment. **Larger companies spend close to 150 times the expenditure of smaller companies on data among the Member States**, although this will drop to "only" 100 times by 2030.

10.3 Data Supplier Companies

Data companies' revenues correspond to the aggregated value of all the data-related products and services generated by Europe-based data suppliers, including exports outside the EU. This indicator measures the revenues of the data suppliers identified and classified under Indicator 2 for the products and services specified in our definition of the data market. Data companies' revenues do not include data monetisation as part of the data market.

Revenues generated by data suppliers have registered a constant increase over recent years to reach nearly €71 billion in the EU27 in 2020. Data companies' revenues account for 0.3% of total company revenues in 2020. This seems significantly lower than previously reported, but this is because the total for data revenues is significantly expanded, from including only the NACE II categories of J and M, to including A, C, E, G, H, J, K, M, P, and Q – a significant increase that includes manufacturing, hence the apparent dramatic drop in share taken by data revenues of total data revenues. The spread by Member State of data revenues as a

share of total revenues is consistent with previous reports and reflects the differing industry focuses shown by Member States. Those with a greater focus on data intensive industries such as finance and professional services will show a greater share of total revenues associated with data.

Larger companies show greater revenues, but this is again not a surprise. Being active in the data market as a data supplier requires significant expertise in data products and services, and these tend to defer to larger organisations. Smaller data user companies too are less inclined to invest in data tools and services because of the cost and expertise needed.

The concept of data revenue is different from the one of data market, but the imports and exports of data-driven products and services are expected to track each over time.

10.4 Data Market Value

The data market is the marketplace where digital data is exchanged as products and services as a result of the elaboration of raw data.

The value of the European data market will reach €63.6 billion for the EU27, with a growth rate of 4.9% in 2021. Most Member States show strong growth, slightly slower than the growth for the total ICT market, which is forecast to grow 5.1% in 2021. The data market's share of total ICT spending was 11.4% in 2020.

The size of the data market by country still correlates closely with the overall economic strength of each country, as well as with national spending on ICT: Germany, France, Italy, the Netherlands, and Spain accounted for approximately two-thirds of the EU27 data market in 2020. Some smaller economies, however, continue to display higher-than-average data market shares of ICT spending. This is notably the case of Estonia and, to a lesser extent, Cyprus, Latvia, and Lithuania. Aside from the outliers, the spread of share of ICT spending taken by the data market is fairly narrow across most Member States.

The larger industries, accounting for the greatest number of companies, represent the largest share of the data market. In terms of adoption by industry, the highest rates of data technology tend to be in manufacturing, finance, and public administration. Thanks to the size of these markets, organizations in these industries are the biggest consumers of data technologies – partly, because of the significant number of companies and their early investment in data tools and services.

The forecast for the data market shows which industries make the biggest contribution to overall market growth, and the key industries of manufacturing, public administration, finance, information technology, and professional services account for close to 82% of the total market growth from 2020 to 2030 (Baseline), with more than 60% coming from the three main industries of manufacturing, public administration, and finance.

Data monetisation will contribute an increasing share to the total data market; it is forecast to add as much as 30% to the European data market by 2030, according to the Baseline forecast. Even though, businesses are struggling to understand how to monetise data and how to capitalize on the data they have to sell and what data they can buy, **we expect that this market will play pivotal role.**

10.5 Data Economy

The data economy measures the overall impacts of the data market on the economy as a whole. It involves the generation, collection, storage, processing, distribution, analysis, elaboration, delivery, and exploitation of data enabled by digital technologies.

The value of the data economy for the EU27 has been estimated to have reached almost €400 billion in 2020 and €440 billion in 2021. The estimated share of overall impacts as a part of GDP in the EU27 rose from 3.1% in 2019 to 3.6% in 2021. We estimate that the data economy for the EU27 in 2025 will reach €600 billion and €1 trillion in 2030.

Another important result is the change in the composition of impacts: From 2021 to 2030, the induced impacts share will increase from 34% to 39% at the expenses of indirect impacts, which will diminish from 50% to 45%.

Thanks to the NGEU's strong focus on the modernisation of public administrations, public administrations will invest in modernising digital services, digital connectivity, cybersecurity, and digital identity.

In terms of countries' contributions to the data economy, France, Germany, Italy, Spain, and the Netherlands tend to contribute the most to the data economy in the EU27. The NGEU again plays a significant role, as around 50% of total resources will be distributed across the four biggest countries in the EU27, making a significant difference in the next five years. Indeed, France and Germany are also among the fastest growing countries, together with Croatia, Estonia, Finland, the Baltics, Romania, Slovakia, and Poland, among others.

10.6 Data Skills Gap

The data professionals skills gap indicator captures the potential gap between demand for and the supply of data professionals in Europe.

The gap between supply and demand for data professionals continue to grow. In the baseline scenario, the gap for EU27 is estimated at 2.5% of the total number of data professionals in 2020, growing to 5.3% in 2025 and 7% in 2030. The situation is similar in the European countries outside EU. This growth is robust even in the challenging scenario with a weakening world economy.

The underlying driver for the shortage is the significant growth in investments in technology, in turn driven by the digital transformation. Organisations throughout Europe will remedy the situation by a combination of continuous training, re-skilling and recruitment, but the estimates show that the gap will remain and that other measures, including re-skilling of additional business roles outside data analysis and ICT, and adding data analysis skills to the curriculum of additional university programs, will be suggested.

Key Findings – Skills Gap

The skills gap for data professionals already has a significant business impact on organisations in Europe. Business and technology are in rapid development, largely due to ongoing digital transformation of enterprise and society, but also due to a very strong pace of development in technology. A transformation of this magnitude requires large amounts of human capital and skills before the process changes stabilize and become mainstream. The reported skills gap in data professionals is therefore impacting both existing and coming

transformation initiatives and becomes one of the more serious challenges facing the European industries.

10.7 Data Economy Beyond the EU: the US, Brazil, Japan, and China

This report extends the analysis of the data market and the data economy to four main non-European countries: the United States, Brazil, Japan, and China.

The growth of the US data economy continued despite 2020 and 2021 being difficult years in the country and globally due to the COVID-19 pandemic and the subsequent socioeconomic downturn. Nevertheless, in the US, the data economy exhibited a solid 2020–2021 year-on-year growth in all the main indicators monitored, including the number of data professionals and companies and the overall data market.

China was added to the international countries monitored by the European Data Market studies for the first time and shows an already well-developed data market and data economy (comparable with those of Japan), which are expected to grow over the forecast period. Yet the country's data economy has certainly not reached its peak, and we anticipate continued growth as the country's influence spreads globally. At the moment, we believe the monetisation of data in China is weak but could strengthen as data increases in volume and value.

The US currently boasts the largest number of data professionals among the countries under consideration. Data professionals in the United States in 2021 are estimated at more than 16 million, while the nearest competitor, China, has only 9.8 million data professionals. Growth in the number of data professionals in China is higher than in the US, but the United States continues to display solid growth in the employment of data professional in the near-term, thus maintaining its advantage for the time being. Brazil's number of data professionals has slowed as the economy has weakened, while Japan showed a modicum of growth in the employment of data professionals in 2021. Brazil has the potential to play a more significant role in the data market and data economy worldwide, but the overall economy is stifling growth in the country. The Brazilian economy runs the risk of losing its skilled data professionals to other countries or regions (such as the US, the UK, or indeed the EU27), as these countries show more potential at the present stage and for the near future.

The data market in Japan is closest to the EU27 in growth and size. However, while the data market in the EU27 is slightly larger than the Japanese one, Japan nearly doubled the size of its data market in 2021. Japan also retains a larger share of employees involved in the data market, close to twice the share in the EU27. Japan's share growth (data professionals as a part of total employment) lags that of the EU27. Nevertheless, in 2021, Japan reversed the negative GDP trend of 2020, posting a significant recovery for the year. If this trend is confirmed, a positive effect on the Japanese data market and the overall data economy can be expected.

Most of economies are maintaining or increasing the impact of the data economy on GDP. Although the EU27 recorded a minor decrease in 2021, down from 0.60% in 2020 to 0.59%, the impact is expected to recover in 2022. Brazil still has a low share of its GDP attributable to the data economy – at only 0.2% compared with close to 1.0% for the remaining international economies. The US has confirmed its primacy in the data economy worldwide and shows the strongest growth in the impact of the data economy on GDP for the time being. The outlook for China is uncertain, but its investment potential in the data economy

is considerable. Whether the US leadership is maintained over the medium term will depend on the strategic focus that China exerts on digital technologies and the capacity of the US to counteract that effectively.

Annex I – Methodological Annex

Overview

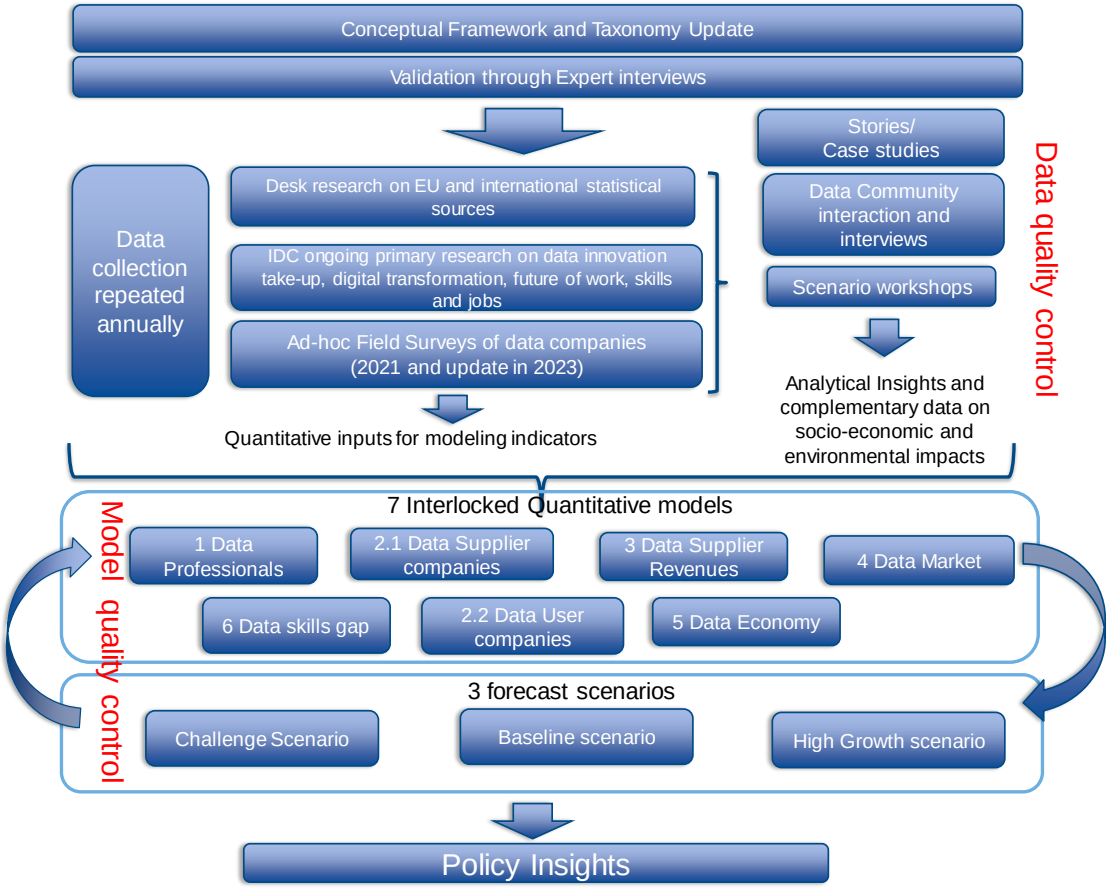
The study required a complex mix of quali-quantitative methodologies allowing to reach the interlocked objectives. While quantitative methodologies represent the most relevant part of the study, qualitative methodologies are indispensable to balance the statistical approach and provide the market and social intelligence needed to lead to policy insights and the development of sound scenarios.

In order to guarantee the continuity of the study, the methodological approach is similar to the previous releases of the EDM Monitoring Tool with the following improvements:

- Updated and revisions of the main indicators' definitions aligned with the maturing of the market and the Data Economy
- Strengthening of the quali-quantitative analysis of socio-economic and environmental impacts
- Additional data collection on data skills and jobs
- Stronger focus on start-ups, thanks to access to data from Dealroom, leading data and intelligence provider on start-ups and scaleups in Europe, and partner of EuropeanStartups.co

The main steps of the methodology did include:

- Revision and update of the Conceptual framework and taxonomy, focusing on:
    - Update of definitions of data professionals, data user companies, Data Market including data monetisation value
    - Definition and management of data sharing and data interoperability issues, including the role of Common European data spaces and the concept of data sovereignty
    - Assessment of social and environmental impacts of data-driven innovation
- Validation of the revision/update through the expert interviews.
- Organisation and implementation of data collection (including desk research and field research) which will be repeated annually to feed into the measurement of indicators.
- The outputs of data collection and qualitative analysis of the 7 interlocked quantitative models used to measure the main indicators (the data companies indicator has 2 models, one for user companies and one for supplier companies).
- In parallel with the calculation of indicators, the scenario forecasting methodology is implemented developing the main assumptions driving the 3 alternative scenarios to 2030 and the forecast of all indicators.
- The quantitative models and the scenarios methodology interact closely and provide reciprocal feedback.
- Quality control accompanied each step of the process, with a focus on data quality control in the data collection phase and on model quality control in the phase of measurement of indicators and development of scenarios. Quality control of deliverables is under project management.



Desk Research

As in the previous editions of the European Data Market study, the study team revised the list of relevant and available public sources, integrated it if necessary and collected the data that are necessary for the indicators' update.

- Concerning the indicators on data market, data companies, data companies' revenues, and the data economy the main sources are:
    - Eurostat business demography statistics in the European Union, treating aspects such as the total number of active enterprises in the business economy, their birth rates, death rates, and the survival rate (last update: March 2020)[18];
    - Eurostat annual structural business statistics with a breakdown by size-class are the main source of data for an analysis of SMEs (latest update: March 2020)[19];
    - IDC Worldwide Black Book Live Edition, monthly updates form the years 2019 through 2024[20]. The Black Book represents IDC's live analysis of the status and projected growth of the worldwide ICT industry in 89 countries.
    - IDC's spending guides[21]. Spending Guides are multi-dimensional, all-in-one data products that present technology forecast data segmented by any or all of the following views: region, country, industry, company size, line of business or use case:

[18] https://ec.europa.eu/eurostat/statistics-explained/index.php/Business_demography_statistics
[19] https://ec.europa.eu/eurostat/web/structural-business-statistics/overview
[20] https://www.idc.com/getdoc.jsp?containerId=IDC_P336
[21] As an example: Worldwide ICT Spending Guide: Industry and Company Size, IDC 2020, Worldwide ICT Spending Guide: Industry and Company Size, IDC 2020 https://www.idc.com/getdoc.jsp?containerId=IDC_P33207

        - IDC Worldwide ICT Spending Guide Industry and Company Size, semi-annual updates for IT Hardware, Software, IT, business and telecom Services from 2019 to 2024 by 20 Industries and 5 size-classes in 53 countries;
        - IDC Worldwide Big Data and Analytics Spending Guide, semi-annual updates for Big Data and Analytics spending from 2019 to 2024 by 20 Industries and 5 size-classes in 53 countries;
        - IDC Worldwide Digital Transformation Spending Guide, semi-annual updates for Digital Transformation spending from 2019 to 2024 by 20 Industries and 278 use cases in 9 regions;
        - IDC Worldwide Artificial Intelligence Spending Guide, semi-annual updates for Artificial Intelligence spending from 2019 to 2024 by 20 Industries and 194 use cases in 9 regions and 32 countries.
    - IDC European Tech and Industry Pulse Survey 2019 2020[22].
    - IDC Big Data and Analytics in the COVID-19 Era: Adoption and Spending Trends Across Vertical in Europe, Jun 2020, IDC # EUR145280920.
    - IDC FutureScape: Worldwide Data and Analytics 2021 Predictions, October 2020, IDC #US46920420
    - IMF World Economic Outlook (WEO) Database, April 2020.
    - Consensus Forecasts, Consensus Economics, monthly updates to July 2020.
    - Review of data on social networks about new and emerging companies through a thorough research of annual reports of the most relevant companies, where available.

- For the data professionals we will use in addition the following sources:
    - OECD publications about the digital economy[23]. As an example, "A roadmap toward a common framework for measuring the Digital Economy", OECD 2020 and "Going digital: Making the transformation work for growth and well-being: Measuring the Digital Transformation. A Roadmap for the Future" OECD 2020.
    - ILOSTAT (International Labour Organization) Statistics and Databases (2020)
    - EUROSTAT Educational attainment statistics (Last update: 2019)[24].
    - European Data Science Academy (EDSA) project deliverables and publications (2018).
    - IDC's Technology Employment Impact Guide – updated on a semi-annual basis with forecast of employment across 40 technology job roles, including eight data management and analytics roles (Last update December 2020)
    - Cedefop – Skills-OVATE data for vacancy estimations (Most recent data collected between July 2018 and September 2020)
    - Cedefop – Skills Index and Skills forecast (Last update 2019)[25]

- Other sources from which relevant data for the indicators' measurement and for the three updates of the indicators during the study duration will be:
    - The Digital Economy and Society Index (DESI), Human Capital Dimension, (2a Basic Skills and Usage; 2b Advanced skills and Development), last update, 2018.
    - IDC Worldwide Augmented and Virtual Reality Spending Guide.
    - IDC Quarterly Wearable Device Tracker.
    - IDC FutureScape: Worldwide Future of Work 2021 Predictions, Oct 2020, IDC #US46248920.
    - Practices to Make AR and VR a Reality for Enterprises, Jun 2020. IDC #EUR146541720.

[22] https://www.idc.com/getdoc.jsp?containerId=EUR145717319
[23] https://www.oecd.org/sti/ieconomy/
[24] https://ec.europa.eu/eurostat/statistics-explained/index.php/Educational_attainment_statistics
[25] https://ec.europa.eu/eurostat/statistics-explained/index.php/Educational_attainment_statistics

In addition, IDC has established a LinkedIn community for European start-ups and scale-ups, which we will be able to poll for specific insights on the topics above. As of March 2021, the community counts 251 members.

Scenarios Desk Research

We used a combination of external sources and IDC sources from its ongoing research. The most relevant sources were the following:

- IDC's European IT spending forecast and key digital trends across European industries and Worldwide Economic and Industry Assumptions. The most recent versions are dated September/October 2020)[26]. These documents are updated quarterly to feed into IDC's ongoing forecasting.
- IDC's FutureScape predictions by technology and industry are delivered once a year: IDC analysts deliver 10 main predictions for the next 2 to 5 years for each vertical market (for example, government or retail) and main technology area (for example Digital transformation). The predictions are developed in a global interactive process between analysts and then are specialised by world region (of which one is Europe). These predictions have proven very useful for the development of scenarios assumptions and storylines.
- The most recent research on Big Data and Analytics, Digital Transformation and Innovation accelerators and emerging technologies.[27]
- The 2025 emerging technologies landscape developed for the ATI (Advanced Technology for Industry) study for EASME-COSME by a consortium led by IDC with Technopolis Group, IDEA, Fraunhofer, Capgemini Consulting, NESTA[28].
- The most recent public studies about digital markets and big data forecasts and trends, by well renowned international bodies and organisations such as Accenture, OECD, Mc Kinsey[29].
- Updated forecasts to 2030 of EU GDP and ICT spending, under 3 alternative scenarios, leveraging the market insights and forecasts of the Economist Intelligence Unit (EIU[30]), the International Monetary Fund, and the OECD.
- Collection and review of all useful data from IDC's extensive databases to estimate the future value of the data market, including for example size and forecast of data stored, size and forecast of data analytics and Big data software, primary research on companies' plans of adoption for data analytics and data-driven applications and services, etc.
- Historical trends emerging from the EDM Monitoring Tool indicators in the period 2019–2025 as per the European Data Market study update (SMART 2016/0063).

[26] As an example: European IT Spending Forecast, 2019–2023: Key Digital Trends Across European Industries, IDC, Nov. 2019, https://www.idc.com/getdoc.jsp?containerId=EUR145632419; IDC Worldwide BlackBook, 2020 https://www.idc.com/getdoc.jsp?containerId=IDC_P336

[27] IDC FutureScape: Worldwide IT Industry 2020 Predictions, IDC, Nov. 2019, https://www.idc.com/getdoc.jsp?containerId=US45599219

[28] https://ati.ec.europa.eu/

[29] A roadmap toward a common framework for measuring the Digital Economy, OECD, 2020 http://www.oecd.org/sti/roadmap-toward-a-common-framework-for-measuring-the-digital-economy.pdf

[30] https://store.eiu.com/product/market-indicators-and-forecasts

Measuring Data Professionals

Definition and Scope

Data professionals *are workers who collect, store, manage, and/or analyse, interpret, and visualise data as their primary or as a relevant part of their activity. Data professionals must be proficient with the use of structured and unstructured data, should be able to work with a huge amount of data and be familiar with emerging database technologies. For 2021–2023 the definition of data professionals was refined to differentiate the roles played by different data users: these are Data Technical Professionals, Data Business Professionals, and Data Consumers. The measure of data professionals includes data technical professionals and data business professionals only.*

Indicator 1 – Data Professionals				
N.	**Name**	**Description**	**Type and Time**	**Segmentation**
1.1	Number of data professionals	Total number of data professionals in the EU	Number, 2019–20–21 Forecast 2025. Forecast 2030, 3 Scenarios	By Geography: EU27 MS (by country and total) EEA (NO, LI, IS) Switzerland UK
				By Industry: 12 industry sectors NACE rev.2
				By size: not applicable
1.2	Employment share	Total number as a share of total employment in the EU	% of total employment, 2019–20–21	By Geography: EU27 MS (by country and total) EEA (NO, LI, IS) Switzerland UK
				By Industry: 12 industry sectors NACE rev.2
				By Size: not applicable
1.3	Intensity share	Average number of data professionals per company (only for private sector)	Number, 2019–20–21	By Geography: EU27 MS (by country and total) EEA (NO, LI, IS) Switzerland UK
				By Industry: 12 industry sectors NACE rev.2
				By Size: not applicable

Methodology Approach:

The methodology approach is based on an iterative process and on a calibration process of the final estimates.

Statistical Identification

Data professionals are not classified as such in any of the labour and occupation statistics. In order to define them statistically, we adopted the International Standard Classification of Occupations (ISCO-08), selecting categories where data professionals may be included. The criteria adopted for the selection of the ISCO-08 codes are the following:

- We have selected the occupations where data professionals can be involved either as data providers or as data users;
- We have selected the occupations from 1 to 4-digit disaggregation;
- The occupation codes selected are those where the presence of data professionals can be detected because they fit into the definitions above:
- Data Technical Professional are a smaller subset of Data Professionals as a result of their increased expertise and focus. Out of the four digit ISCO codes only 10 categories are included in the definition of Data Technical Professionals.
- Data Business Professionals are identified as one of the 40 4-digit ISCO categories based on their management responsibilities or lower technical experience or expertise.
- The selected codes relate to the roles and responsibilities highlighted for the Data Technical Professionals and the Data Business Professionals in the preceding paragraphs and are where a significant part of the workers in these categories perform specific responsibilities relating to these roles.
- We excluded all the data professionals which are not included into the knowledge economy perimeter because their occupation is a low skilled one, i.e. with high routine level (as an example, call centre workers are in theory data professionals but since their activity is a routine one and as such excluded from the knowledge economy, they are not considered data professionals).
- Table below shows the detail of the number of codes included in each of the definitions for the 1,2-,3-, and 4-digit categories.

ISCO-08 Structure and Data Professionals	ISCO-08 structured Classification			
	Major Groups (1 digit)	Subgroups (2 digits)	Minor Groups (3 digits)	Units (4 digits)
Number of codes ISCO-08 structure	10	43	131	436
Number of selected codes including data professionals	8	23	52	245
Of which data business professionals	4	7	12	41
Of which data technical professionals	2	3	4	8
Of which data consumers	8	23	51	121
Share of data professionals' codes in the ISCO-08 structure	80%	53%	39%	28%

Calculation of the Quantitative Perimeter

The quantitative perimeter of employment where data professionals are trackable is based on the selected ISCO codes crossed with the NACE classification of economic activities, for each one of the 27 Member States, Switzerland, the EEA countries, the UK and the EU as a whole and has been updated based on the source's updates.

Estimate and Calibration of the Penetration of Data Professionals

The next step is to estimate the percentage of data professionals within the perimeter of data professional candidates. The sets of assumptions will be revised and updated for each release of the

study and applied to the model to calculate the share of data professionals by Member State and by industry. The survey of data professional companies and data user companies includes a question relating to the share of workers in each of the categories, and this is one of the prime components of the share estimate.

Forecasting Data Professionals

The same model was applied to forecast data professionals to 2030, by developing specific assumptions by scenario, even though the level of uncertainty is higher, and the reliability of the forecasts is lower.

Measuring Data Companies

Definition and Scope

***Data companies** are organisations that are directly involved in the production, delivery and/or usage of data in the form of digital products, services and technologies. They can be both data suppliers' and data users' organisations:*

- ***Data suppliers** have as their main activity the production and delivery of digital data-related products, services, and technologies. They represent the supply side of the Data Market.*
- ***Data users** are organisations that generate, exploit collect and analyse digital data intensively and use what they learn to improve their business. They represent the demand side of the Data Market.*

Indicator 2 – Description				
N.	**Name**	**Description**	**Type and Time**	**Segmentation**
2.1	Number of data supplier companies	Total number of data supplier companies in the EU & EEA & UK, measured as legal entities based in one country	Number, 2019–20–21 Forecast 2025. Forecast 2030 (3 Scenarios)	By Geography: EU27 MS (by country and total) EEA (NO, LI, IS) Switzerland UK
				By Industry: Sectors A, C, D, E, G, H, J, K, M, P, Q
				By company size: below 250 employees above 250 employees
2.2	Share of data supplier companies	Total data supplier companies on total companies in industry sectors A, C, D, E, G, H, J, K, M, P, Q	%, 2019–20–21 Forecast 2025, Forecast 2030 (3 Scenarios)	By Geography: EU27 MS (by country and total) EEA (NO, LI, IS) Switzerland UK
				By Industry: Sectors A, C, D, E, G, H, J, K, M, P, Q
2.3	Number of data user companies	Total number of data user companies in the EU, measured as legal entities based in one country	Number, 2019–20–21 Forecast 2025, Forecast 2030 (3 Scenarios)	By Geography: EU27 MS (by country and total) EEA (NO, LI, IS) Switzerland UK
				By Industry: 12 industry sectors NACE rev.2
				By company size: below 250 employees

Indicator 2 – Description				
N.	**Name**	**Description**	**Type and Time**	**Segmentation**
				above 250 employees
2.4	Share of data user companies	Total data user companies as share of total private companies	%, 2019–20–21 Forecast 2025, Forecast 2030 (3 Scenarios)	By Industry: 12 industry sectors NACE rev.2
2.5	Share of data user and data supplier companies that offer data for re-use.	Percentage of data companies that offer data reuse as a percentage of total data supplier and data user companies	2020, 2021	By industry: 12 industry sectors NACE rev. 2 By company size band: Below 250 employees Above 250 employees

Methodology Approach

The indicators on Data Supplier Companies and Data User Companies is measured by updating the same model used in the previous EDM study which leverages both IDC and public sources.

Measuring the Revenues of Data Companies

Definition and Scope

Data companies' revenues *correspond to the aggregated value of all the data-related products and services generated by Europe-based data suppliers, including exports outside the EU. This indicator measures the revenues of the data suppliers identified and classified by Indicator 2, for the products and services specified in our definition of the data market. Data companies' revenues do not include data monetisation as part of the data market.*

Indicator 3 – Description				
N.	**Name**	**Description**	**Type and Time**	**Segmentation**
3.1	Total revenues of Data Supplier Companies	Total data supplier companies' revenues	Billion €, 2019–20–21	By Geography: EU27 MS (by country and total) EEA (NO, LI, IS) Switzerland UK
				By company size: below 250 employees above 250 employees
3.2	Share of Data Supplier companies' revenues	Ratio between data supplier companies' revenues and total companies revenues in the sectors J and M	% of revenues on total, 2019–20–21	By Geography: EU27 MS (by country and total) EEA (NO, LI, IS) Switzerland UK

Methodology Approach

The indicator is measured by applying the same model used in the previous EDM study, which calculated the revenues by feeding on:

Data Source	Used in
Data Supplier Companies (Indicator 2)	Data Company Revenues
IDC Core IT Spending guide	Data Company Revenues

Data Source	Used in
IT Big Data and Analytics spending Guide	Data Company Revenues
IDC Worldwide Black Book (standard edition)	Data Company Revenues
IMF World Economic Outlook	Data Company Revenues

Measuring the Data Market

Definition and Scope

*The **Data Market** is the marketplace where digital data is exchanged as "products" or "services" as a result of the elaboration of raw data.*

Indicator 4 – Description				
N.	**Name**	**Description**	**Type and Time**	**Segmentation**
4	Value of the data market	Estimate of the overall value of the data market (including data monetisation)	Billion €, 2019–20–21 Forecast 2025 – Baseline scenario, Forecast 2030 (3 Scenarios)	By Geography: EU27 MS (by country and total) EEA (NO, LI, IS) Switzerland UK By Industry: 12 industry sectors NACE rev.2
4.1	Data monetisation	Sub-indicator Estimate of the value of data monetisation	Billion €, 2019–20–21 Forecast 2025 – Baseline scenario, Forecast 2030 (3 Scenarios)	By Geography: EU27 MS (by country and total) EEA (NO, LI, IS) Switzerland UK By Industry: 12 industry sectors NACE rev.2

Methodology Approach

The data market indicator is updated every year for the duration of the study. The model is built on data from IDC databases concerning the components of hardware, software, and services, spending which fall in the definition of the data market. The value of data monetisation is added to this and is estimated from desk research and the results of the ad-hoc data companies survey. The IDC data is already segmented by country and by industry and this is mapped to the industry segments used in this study using already established mapping tables build from detailed matches of NACE II segments. The respective shares for the software, hardware, services, and data monetisation spending are derived from IDC surveys covering Big Data, IT spending patterns and intentions in the European market, and a survey of data supplier companies and data user companies in key Member States, together with analyst expertise and alignment with IDC's European and worldwide forecasts for the business analytics and Big Data market.

This model updates the data market value shares by MS and by industry and uses the following data sources:

Data Source	Used in
New ad-hoc survey	Data Market
Data Companies' revenues (Indicator 3)	Data Market
Eurostat Business Demographic Statistics	Data Market

Data Source	Used in
Eurostat annual Structural Business Statistics	Data Market
Eurostat chain linked Volumes (GDP)	Data Market
IDC Core IT Spending guide	Data Market
IT Big Data and Analytics spending Guide	Data Market
IDC Worldwide Black Book (standard edition)	Data Market
IMF World Economic Outlook	Data Market
Consensus Forecasts – Consensus economics	Data Market

Measuring the Data Economy

Definition and Scope

The **Data Economy** measures the overall impacts of the Data Market on the economy as a whole. It involves the generation, collection, storage, processing, distribution, analysis elaboration, delivery, and exploitation of data enabled by digital technologies.

The Data Economy captures a wider concept than the Data Market only, as it considers the value and wealth generated in the economy as a whole (not just across businesses) by the exploitation of data.

The Data Economy includes three sets of impacts in the economy: the Data Companies Revenues in the form of direct impacts on the economy, the indirect impacts (as backward and forward) and the induced impacts effects of the Data Market on the economy.

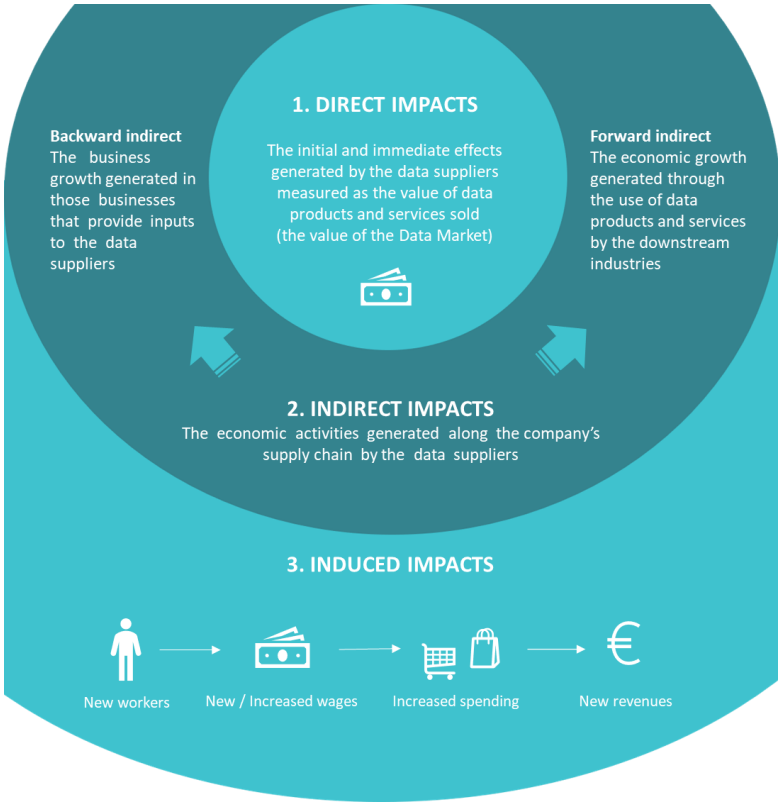
- The **direct impacts** are the initial and immediate effects generated by the data supplier companies; they represent the activity potentially engendered by all businesses active in the data production. The quantitative direct impacts will then be measured as the revenues from data products and services sold, i.e. the value of the Data Market. We consider the Data Market value as a good proxy of the direct impacts. Therefore, for the sake of simplicity, direct impacts will coincide with the value of the Data Market.
- The **indirect impacts** are the economic activities generated along the company's supply chain by the data supplier companies, considering input providers and customers of data supplier companies. Indeed, there are two different types of indirect impacts, the backward indirect impacts and the forward indirect impacts (Richardson, 1985):
    - the **backward indirect impacts**: such impacts represent the business growth resulting from changes in sales from suppliers to the data industry. In order to produce and deliver data products and services, the data suppliers need inputs from other stakeholders. Revenues generated among the providers side from those sales to data suppliers companies are the backward indirect impacts.
    - the **forward indirect impacts**: such impacts include the economic growth generated through the use of data products and services by the downstream industries, i.e. the data user companies as a selected number of industries. For the user companies, data is a relevant factor of production; the adoption of data products and services by the downstream industries provides different types of competitive advantage and

productivity gains to the user industries. Data users are engaged in digital transformation, able to make a strategic use of data and reap its benefits. The main benefits that the exploitation of data can provide to downstream industries are (OECD, 2013, Mc Kinsey, 2011):

- Optimising production and delivery processes: data-driven processes (data-driven production);
- Improving marketing by providing targeted advertisements and personalised marketing practices (data-driven marketing);
- Improving existing organisation and management practices (data-driven organisation).

- The **induced impacts** include the economic activity generated in the whole economy as a secondary effect. Induced additional spending is generated both by new workers, who receive a new wage, and by the increased wage of existing jobs. This spending induces new revenues creation in nearly all sectors of the economy. The additional consumption will support economic activity in various industries such as retail, consumer goods, banks, entertainment, etc.

The Data Economy: Direct, Indirect, and Induced Impacts



Source: European Data Market Monitoring Tool, IDC 2021

This indicator is measured according to the scope detailed in the following table.

Indicator 5 – Value of the Data Economy				
N.	**Name**	**Description**	**Type and Time**	**Segmentation**
5	Value of the data economy	Value of the direct, indirect and induced impacts of data-driven innovation on the EU economy	Billion €, 2019, 2020, 2021 Forecast to 2025. Forecast to 2030, 3 Scenarios	By Geography: EU27 MS (by country and total) EEA (NO, LI, IS) Switzerland UK
5.1	Impact of the data economy on GDP	Ratio between value of the data economy and EU GDP	Billion €, 2019, 2020, 2021 Forecast to 2025. Forecast to 2030, 3 Scenarios	By Geography: EU27 MS (by country and total) EEA (NO, LI, IS) Switzerland UK

Our Data Economy estimation does not include the user benefits and social impacts of data-driven innovation such as changes in quality of life (health, safety, recreation, air quality). Although these benefits may be evaluated in economic (monetary) terms, they are not economic impacts as defined above as they do not induce an increase in the business activities and a consequent growth in GDP.

The Value of Data

The value creation process based on data rests on the elaboration of information and knowledge (OECD 2016), although the boundaries between data, information, and knowledge are sometimes fuzzy. The huge volume of data is a global phenomenon which is sometimes view with suspicion by citizens, consumers and businesses because data flows are seen as an intrusion of the privacy.

Nevertheless, it is now commonly agreed that data analysis can provide benefits to both businesses and consumers. Moreover, the introduction of GDPR (General Data Protection Regulation) in May 2018 helped in managing the usage of information, giving rules to data users as well as providing control over personal data to data owners.

We should remind that the economic theory holds that information encourages competition between businesses for the benefit of consumers. Data do not provide value and benefits as such; data need to be collected, stored, aggregated, combined and analysed in order to be appropriately used for decision making processes. **To create value, data need to be processed (OECD, 2016):**

- **Extracting information from structured and unstructured data**: data analytics techniques are today able to analyse both structured and unstructured data. We should remind here that most data stored by businesses are unstructured. Technologies such as optical character recognition, natural language processing, face recognition algorithms and machine learning algorithms are empowering the use of all data.
- **Real-time monitoring and tracking**: analysis of data in real time is often mentioned as one of the most powerful factors since it supports organisations to make real-time decisions, which, in a fast changing world, is a well-known competitive advantage.

- **Inference and prediction**: until now, prediction was based exclusively on prior information and data series. Data analytics can now enable the creation of information even without prior information. Such information can be created through patterns and correlations of data. Personal information, for example, can be deduced from anonymous or non-personal data. Businesses and organisations demand real time insights rather than historical and periodical information, and for advanced specialised data analytic services. Algorithms allow machine and statistical learning based on non-specific data; businesses can learn and predict a lot about their customers even if they do not have specific data and time series about the issue they are interested in. Machine learning has, as an example, applications in health care where data collected on patients are recorded by imaging, or it supports production processes to increase the quality of production

The diffusion of technology supporting production and analysis of data induces organisations and businesses to base their decisions on data much more than they were used to do. As pointed out by OECD in its recent report, the process to take decisions is also changing. Decision makers do not necessarily need to understand the phenomenon before they act on it. A store can change the product placement based on data analysis without the need to know the reason why such a change should improve the sales. There is therefore a decision automation process: "first comes the analytical factor, then the action, and last, if at all, the understanding" (OECD, 2015).

The impacts of such a new approach to decision making and to the use of data in all the enterprises and organisations' functions are many and varied, so that we believe, such impacts will be object of studies and analysis in the upcoming years. It is, at this point, difficult to classify them and to suggest a taxonomy of such impacts.

Such impacts have been observed through some empirical studies and case analysis. The most relevant ways the benefits appear are the following:

- **Creating more information, knowledge and transparency**: technology is making data more accessible and exploitable to all kind of stakeholders, including SMEs. This increases transparency and decisions are made on a rational process.
- **Improving performance**: having access to a wide information and to a high number of data is changing the way of making decisions. An increasing number of organisations are going to become data-driven organisations, which means that they make decisions based on empirical results. As an example, retailers can adjust prices and promotions, more precisely than they were used to and in real time. This may improve competitiveness. McKinsey underlines that the health sector is achieving a lot of benefits from the new making decisions process: studies on clinical data allow to identify and understand the sources of variability in treatment, to identify the best treatment protocols and to create guidelines for the optimisation of treatment decisions. This does not only increase the effectiveness of treatments, but it also produces saves.
- **Improving customisation of actions for better decisions**: data technology is definitely improving the segmentation of customers and the analysis of their preferences in real time. This allow companies to supply products and services targeted to specific groups of individuals who have specific needs and preferences. Such a segmentation is also useful when supplying public services. Such a segmentation helps define the price precisely and offering exactly what is needed which

means a better quality and also companies avoid offering products and services the consumers are not willing to pay.

- **Innovating products and services as well as business models**: the more information and understanding businesses have about their customers, the better they can serve them. It is important to say that although consumers may fear their privacy is injured, this can also provide them unexpected surplus: real time price comparison services do not only provide better transparency but also allow buying the best product at the most convenient price (for example when buying online airline tickets or when booking hotels). Companies can in fact produce and create new products and services to better satisfy their customers' needs. This is true also for the public sector and specifically for the health care system where preventing care programs can be created.
- **Ecosystem effects**: there are some areas in which there are great opportunities deriving from the use and the exchange of data, and that will be also driving examples for the near future, such as parts provenance and the origin track of food and materials in manufacturing, but also tracking the conditions at which materials and goods are shipped, know your customers and digital identity (for the financial and the public sector), tracking of medical devices and appliances as well as managing data sources of medical information in healthcare.

These effects are reflected in an increase in revenues due to higher market share from the increase in competitiveness or due to a reduction in costs. All these effects are included in the forward indirect impacts; these impacts are delivered on the user industry, and because of the above reasons, these are the impacts we consider new on the overall economic system.

Methodology Approach

Measuring the data economy, broadly speaking, depends on:

- the macroeconomic context
- the availability and diffusion of tools that help companies in their data elaboration and usage;
- the industry and country maturity;
- the integration processes the companies are implementing.

Therefore, the data economy model is based on a set of assumptions on all these factors, including choices about proxy indicators where actual data is missing.

The data economy model is a highly sophisticated model articulated by country and industry which has successfully delivered the current and forecast estimates of economic impacts for the last cycles of the EDM Monitoring Tool measurement. The model is sufficiently flexible that it was possible at the start of 2020 to run a simplified version to provide a rough post-COVID data economy estimate for 2020 and forecast to 2025 for the Baseline scenario. For the next round of the study, we will revise and update the structure and key assumptions of the model.

The main steps are similar to the other indicator models and the following:

- annual round of desk research and data collection.

- Revision and update of the assumptions driving the model and the measurement of each category of impacts.
- Measurement of each type of impacts as follows:
    - Direct impacts: they correspond to the value of the data market (indicator 4)
    - Backward indirect impacts: they correspond to the increase of revenues by data supplier companies and are based on indicator 3 – data supplier companies' revenues.
    - Forward indirect impacts: this is the most difficult type of impact since is based on the estimates of the economic benefits by industry generated by the adoption of data-driven innovation, through the calculation of multipliers.
    - Induced impacts measure the secondary effect of the other categories of impacts together on the overall economy and are calculated through the use of specific indicators and the estimate of appropriate multipliers.
- The impacts measured are then aggregated and their value and growth trend will be cross-checked again for coherence with other indicators. This is the value of the data economy and will be calculated in the first measurement for 2019, 2020 and 2021 (estimate). (for the EU27, separately and in total, the rest of the EEA, the UK and Switzerland).
- Separately, the study team will provide estimates of GDP value for each of the country covered and will calculate the impact of the aggregated data economy on GDP (for the EU27, separately and in total, the rest of the EEA, the UK and Switzerland).

The forecasting to 2030 will be carried out as follows:

For the forecasting scenarios the study team will:

- Review the qualitative assumptions developed for each scenario for the year 2025 (only Baseline) and 2030 (3 scenarios).
- For direct and backward indirect impacts, the forecast will be calculated separately by the data market model and the data supplier companies' revenues model and the results will be included in the data economy model (through a round of cross-check and validation of coherence, robustness and quality of all the results).
- For forward indirect impacts and induced impacts, we will derive assumptions in order to calculate the forecast multipliers under the 3 alternative scenarios to the year 2030.
- The value of all impacts will be aggregated calculating the value of the data economy.
- Estimates of GDP value in 2030 for the countries measured will be generated for the 3 alternative scenarios.
- Finally, we will calculate the impact of the data economy on GDP.
- The results will be again cross-checked for mistakes or lack of coherence by different members of the study team.

It should be noticed that the IDC study team includes 4 different analysts in charge of the indicators models. They are the same analysts who have developed and calculated these models in the past years. This will help the quality control since they will collaborate in the reciprocal cross-check and validation of their models.

Data sources

Each year the study team will carry out ad-hoc desk research to update the data economy model assumptions, particularly the value of multipliers, leveraging other similar studies about the economic impacts of data innovation and other emerging technologies (sources such as McKinsey, Accenture, Deloitte, Everis). The results of the data collection for this study will be used, specifically the measurement of business impacts of data innovation. IDC's annually and bi-annually updated research will be used, including (but not limited to) the following studies:

- IDC Worldwide Semiannual IT Spending Guide Industry and Company Size
- IDC Worldwide Public Cloud Services Spending Guide
- IDC Worldwide Big Data and Analytics Spending Guide
- IDC Worldwide Internet of Things Spending Guide
- IDC Worldwide Artificial Intelligence Spending Guide

IDC spending guides estimate the demand side spending in technologies by industry, country and use case for a period of 5 years (previous year, current year and 3 years forecast). The spending guides are high quality data products, cross-checked for coherence both by geography and by industry. These spending guides will feed into the estimate of the level of take-up and value of data spending by industry as well as into the historical series.

The data on business impacts sourced from the ad-hoc field survey will also be leveraged to improve the estimate of indirect impacts. In order to measure the impact of the diffusion and use of data services and products, we will estimate each component (as defined in the above paragraph) of the impacts separately.

Estimate of Forward Indirect Impacts

As highlighted by OECD, 2013, McKinsey 2011 the impacts provided by the exploitation of data over the economic system include:

- optimising production and delivery processes (data-driven processes)
- optimising marketing by providing targeted advertisement
- enhancing research and development and developing new products and services
- improved decisions making, launch of innovations, creation of new businesses,
- innovating business models
- creating transparency and diffusion of information

While impacts from the data supply-side are immediate and measurable, the impacts on the demand-side are more difficult to catch, especially in the early stage of an emerging industry.

The estimate of the value of the data economy will be based on estimates of the multipliers of the data products and services on the whole economy which depend on (but not only):

- The multiplier effect of data products and services on innovation in the whole economy;
- The multiplier effect of increased revenues by users.

IDC will cluster the industries which may be affected by a high, medium, and low multiplier effect in order to estimate the overall effect on the EU economy. Finance, retail, manufacturing, energy for example are industries where the impact of an intensive use of data is likely to be high. Direct and indirect impacts and the possible multiplier effects are not going to occur within a year, but they may require at least a couple of years.

IDC will develop a detailed model based on other models calculating the economic impacts of IT pervasive innovations. Impacts on economy will clearly depend on the diffusion rate, which in turn depends also on the general economic conditions of next years.

Measuring Data Professionals Skills Gap

Definition and Scope

The **Data Professionals Skills Gap** indicator captures the potential gap between demand and supply of data professionals in Europe.

Monitoring the skills gap is of paramount importance since the lack of skills may become a barrier to the development of the data industry and the rapid adoption of data-driven innovation. It is based on a model balancing the main sources of data skills (from the education system and re-training and other carriers) with the estimated demand (by all data companies).

For the data skills gap the data is provided as always for the 5 largest EU countries and the rest of EU 27 in an aggregated way, mainly because of the difficulty to measure data skills job vacancies for each individual Member State.

Indicator 6 – Description				
N.	**Name**	**Description**	**Type and Time**	**Segmentation**
6	Data Professionals Skills Gap	Gap between demand for and supply of data technical and data business professionals (not segmented)	Absolute number and % on total demand, 2019–20–21 Forecast to 2025 (Baseline scenario) Forecast to 2030, 3 scenarios	By Geography: 5 EU MS: DE, ES, FR, IT, PL Rest of EU27 Total EU 27 UK Switzerland EEA (NO, LI, IS)

Methodology Approach

The measurement of this indicator is based on a model combining the separate estimates and forecasts for the demand for data technical and business professionals and the supply of corresponding data skills by the inflow from the education system and upskilling and reskilling of the existing workforce (Figure 8.1). This includes balancing the main sources of data skills (from the education system and re-training to the provision from other careers) with the estimated demand for data skills (by all data companies).

More specifically, we use the following definitions:

- **The supply of data professionals** is equal to the data skills supply stock (the sum of employed data professionals and the unemployed ones).
- **The demand for data professionals** is the sum of existing and open positions for data technical professionals and data business professionals, that is the number of currently employed data professionals (indicator 1 in this study) plus the unfilled vacancies.

- **The indicator measures the difference between total demand and supply**; if demand is higher than supply there is a data skills gap (excess demand). If supply is higher than demand, there is over supply and potentially unemployment.

Data Sources

As for the other indicators, the study team will carry out annually ad-hoc desk research on data skills supply and demand dynamics. The main sources which will be considered are (but not limited to):

- ILOSTAT (International Labour Organization) Statistics and Databases (2020)
- EUROSTAT Educational enrolment statistics (Last update: 2021).
- IDC's Technology Employment Impact Guide – updated on a semi-annual basis with forecast of employment across 40 technology job roles, including seven data management and analytics roles (Last update: June 2021)
- Cedefop – Skills-OVATE data for vacancy estimations (Most recent data collected between July 2018 and September 2020)
- Cedefop – Skills Index and Skills forecast (Last update 2020)

Measuring Demand

The total demand for data technical and business professionals is calculated for the years 2019, 2020 and 2021 in the first cycle and will move forward one year for each measurement cycle. For the current year of the indicator (in the first measurement it will be 2021) we have added to the number of data professionals sourced from Indicator 1 an estimate of existing unfilled positions (vacancies). The labour market is a dynamic environment characterised by inflows and outflows of human resources, and at any given moment present there are companies looking to hire as well as unemployed looking for a job. Our model includes estimates of these inflows and outflows due to retirements, sickness, deaths, graduations, career changes between companies, industries and job roles, and people entering or exiting the market for training or education activities.

To estimate the current vacancies, we have carried out additional data collection on job search portals such as LinkedIn, Indeed and others to calculate the level of demand for data skills jobs, defined on the basis of the desks research and analysis calculated for Indicator 1 on data professionals. IDC's ongoing research on the demand for advanced ICT and data skills has been leveraged to support the forecasts. In addition, the survey has provided data about companies' difficulty in filling specific data professional positions. This has helped to model the demand forecast and to understand the level of the potential data skills gap.

The forecasted demand for data professionals to 2025 (Baseline scenario) and 2030 under the 3 scenarios calculated by Indicator 1 is considered as the total potential demand (as it incorporates future potential vacancies).

Measuring Supply

Supply has been estimated by aggregating the number of graduates in the relevant disciplines corresponding to the data skills identified in Indicator 1 and the level of inflows from other careers or upskilling. The model considers the inflows and outflows in the data skills market such as retirements, and unemployment.

Since we have changed the definition of data professionals compared to previous years, in this first phase of the study we have also updated the type of data skills to be monitored and the type of fields of study providing them. To do so we have leveraged desk research but also expert interviews.

The relationship between skills demand and supply and the resulting skills gap or over-supply is illustrated in the figure below.

The Data Skills Demand-Supply Balance Model

Source: European Data Market Monitoring Tool, IDC 2021

Annex II – The Ad-Hoc Survey

The EDM Monitoring tool measurements over the past few years leveraged successfully a combination of statistical sources (Eurostat, Ilostat) and annually updated data from IDC surveys. However, we believe that for the new cycle of measurement it is necessary to carry out an ad-hoc survey specifically tailored on the EDM indicators for the following reasons:

- To monitor the reactions of the stakeholders and create a new data baseline after the disruption of the COVID-19 pandemic and the following recession in 2020, which has created a breach in the growth trends of the data market, as documented in the post-COVID scenarios presented in the last reports of the past study[31]. By 2021 we expect to be into the first phase of the rebound recovery: it will be necessary to measure how and to what extent this recovery will evolve.
- Another factor of disruption is the actual implementation of Brexit, the UK exit from the EU, from January 2021. We have already covered this topic extensively in the past years of the EDM Monitoring tool, but a check point through a survey is recommended to make sure that our data for both the 27EU and the UK remain sound and valid.
- To collect data on the new countries which have been added to the Monitoring tool, namely Switzerland, Norway and Iceland.
- To collect data on the agriculture sector, which has been added to the list of industries measured by the study.

Even more important, the survey was needed to collect new evidence on the updated definitions of some of the key indicators:

- **Measuring the value of data** (data monetisation). The definition of the data market in the new study is expanded to include an estimate of the value of data. The survey investigates data monetisation to contribute to this estimate.
- **Measuring data users**. The survey analyses more in depth the take-up of data innovation through the investigation of data use cases and their adoption. This is used to carry out a more articulated identification of data users.
- **Measuring data suppliers**. The survey identifies data supplier companies through questions on their core business without excluding any sector from this analysis.
- **Measuring data skills**. The survey provides fresh evidence on organisations' needs for data skills.

Measuring Business Impacts

The survey includes questions on the business benefits of big data, leveraging the big data benchmarking methodology developed in the DataBench project led by IDC. This is used to inform and strengthen the direct and indirect impacts estimates used for the data economy model.

Survey Sample

The survey was conducted by IDC's European Survey and Analysis Team, and it is based on the following suggested criteria. The size of the sample is mandated by the study requirements and the

[31] "Moving towards the European data-agile economy", May 2020, D2.8 Final report on policy conclusions, Update of the European Data Market Study (SMART 2016/0063)

cost constraints. Our experience guarantees that the quality and reliability of the survey provides sufficiently robust data to feed into the indicator's models.

1st Survey Features	Description
Sample size	1,190 completed interviews
Countries	EU: Czech Republic, France, Germany, Italy, Poland, Spain, Sweden Non-EU: Switzerland, Norway Iceland, UK
Methodology	CATI/telephone 20-minute. Languages: English, French, German, Italian, Spanish, Czech, Polish.
Target respondent	decision makers and influencers for data solutions within the company, manager level+. Mix of IT and non-IT (e.g. analysts, analytics, BI roles etc.)
Target company	companies/organisations with 1+ employees, all verticals. Companies with "1 employee" should not exceed 5% of sample.
Screening questions	Companies must use data to make business decisions. Minimum n=200 must also be "data suppliers". This means these data suppliers must sell data to other companies.
Quotas	Hard quotas by country. Hard quota: minimum 200 respondents must be data suppliers (based on question on core business activity) Soft quotas (+/-10% flexibility) for industry verticals Soft quotas (+/10% flexibility) for size-bands in each country
Timescale	Roughly 12 weeks

This is not what was agreed on in the Inception Report. Here is the agreed list of survey features:

1st Survey Features	Description
Sample size	1,190 completed interviews
Countries	EU: Czech Republic, France, Germany, Greece, Italy, the Netherlands, Poland, Romania, Spain, Sweden Non-EU: Switzerland, Iceland, UK
Methodology	CATI/telephone 20-minute. Languages: English, French, German, Greek, Italian, Spanish, Czech, Polish, Romanian
Target respondent	decision makers and influencers for data solutions within the company, manager level+. Mix of IT and non-IT (e.g. analysts, analytics, BI roles etc.)
Target company	companies/organisations with 1+ employees, all verticals. Companies with "1 employee" should not exceed 5% of sample.
Screening questions	Companies must use data to make business decisions. Minimum n=200 must also be "data suppliers". This means these data suppliers must sell data to other companies.
Quotas	Hard quotas by country. Hard quota: minimum 200 respondents must be data suppliers (based on question on core business activity) Soft quotas (+/-10% flexibility) for industry verticals Soft quotas (+/10% flexibility) for size-bands in each country
Timescale	Roughly 12 weeks

The selection of the countries surveyed was based on the following criteria:

- Geographical balance (representing all main geographical areas in the EU)
- Country size (mix of large, medium, and small Member States)
- Coverage of the new countries to be included in the tool (Switzerland, Norway, Iceland)

The survey quotas for countries, industry, and company size bands used in the survey to establish a credible baseline are specified in the table below. The quotas for the countries are considered hard quotas – i.e., they must be complied with in the survey. The quotas for each industry and company size band are soft quotas – i.e., they are targets for the survey but absolute compliance with the quota is not a requirement.

Compared to the original plan, we have reduced the number of interviews under 10 employees (companies who are less likely to use data-driven innovation meaningfully, and less likely to be data suppliers) and increased the number of large companies (with more than 250 employees) more likely to lead in data-driven innovation. Also, we have kept the UK in the sample but considerably reduced its sample size and made room other three non-EU countries apprehended in this study.

Country - EU	Sample	Country – Non-EU	Sample
Czech Republic	70	Switzerland	60
France	160	Iceland	40
Germany	160	UK	60
Greece	60		
Italy	120		
The Netherlands	90		
Poland	100		
Romania	90		
Spain	110		
Sweden	70		
Total	**1,030**	**Total**	**160**
Overall Total	**1,190**		

Industry	Sample	Industry	Sample
Agriculture	75	Mining & manufacturing	122
Construction	77	Professional services	119
Education	86	Public administration	119
Financial services	106	Retail/ wholesale	122
Healthcare service providers	80	Transportation and storage	85
Information and communications, Telecom, Media	122	Utilities	77
Overall Total	**1190**		

Company Size Band	Sample
1-9	148
10-49	238
50–249	238
250-499	238
500+	328
Overall total	**1190**

Extrapolation of Results to Non-Surveyed Countries

The results of the survey are presented for the 10 surveyed countries. We use a consolidated methodology to extrapolate results to each of the other countries measured in the Monitoring tool, based on data modeling. The extrapolation is based on matching the surveyed countries with those not-surveyed, identifying those most similar. The data modeling is based on mapping the orthogonally closest equivalent country in the survey sample using a 20-dimension vector across industries, company size, and education levels. This involves measuring the share of the total country taken by each of the 12 NACE II industries, five company size bands, and share of population educated to degree level, and matching these shares to the un-surveyed country. The surveyed country with the closest match when measured across all these dimensions is deemed the nearest equivalent country. This methodology has been used successfully in the first round of the EDM Monitoring tool and more recently in the ATI (advanced technologies for industry) study led by IDC with other partners on behalf of DG GROW, currently ongoing, which carried out a similar survey on the take-up of technologies.

Questionnaire Structure

The final questionnaire is below of this chapter. The structure of the questionnaire is articulated as indicated in the following table.

The structure of the questionnaire is based on the following sections:

Sections	Number of questions	What the data is used for
Screener questions (Country/Industry/Size/ Use of data /respondent)	7	Achieve sample quotas and verify the role of the respondent
Section 1 Drivers and use of data	1 + 8	Profiling the use of data including take-up of use cases (both for data suppliers and data users)
Section 2 Data Monetisation	6	Collecting data on data sales, purchases and data sharing
Section 3 Data suppliers	3	Profiling data suppliers, their offering portfolio
Section 4 Business impacts of data-driven innovation	4	Measuring impacts on business outcomes based on KPIs
Section 5 Data professionals	8	Measuring data professionals
Total	30 without screening	

Sections	Draft questions	Type of answer	What the data is used for
Screener questions			
S.1	In which country do you primarily work?	1. Czech Republic 2. France 3. Germany 4. Iceland 5. Italy 6. Norway 7. Poland 8. Spain 9. Sweden 10. Switzerland 11. United Kingdom	Achieve survey quotas by country.
S.2	Which of the following industry classifications best represents the principal business activity of your company/organisation? In the case of commercial businesses, by principal business activity we mean the activity generating the highest share of company revenue.	Categories for analysis/quotas: 1) Agriculture 2) Construction 3) Education 4) Financial services (banking, insurance, other financial services) 5) Healthcare service providers 6) Information and communications (telco, media) 7) Mining & Manufacturing 8) Professional services 9) Public administration 10) Retail/ wholesale 11) Transportation and storage 12) Utilities Banking and credit institutions (retail and corporate banking, global transactions, card/retail payments) Insurance (life, property and casualty, reinsurance, insurance agents and brokers) Other financial services such as brokers, asset managers, and stock exchanges Professional services (consulting, legal and tax, advertising, engineering, architecture, staffing services, software and IT services, real estate, etc.) Telecommunications (fixed or mobile voice, data, or internet services only; *excluding manufacturing of telecommunications equipment*) Media (broadcast, printing, and publishing and other media) Entertainment and other personal and consumer services (arts, entertainment, cultural activities, gambling, libraries, museums, etc.) (OTHER) Passenger transport (water, air, rail, and other land transport)	Achieve survey quotas by sector.

Sections	Draft questions	Type of answer	What the data is used for
		Freight transport/logistics (logistics and supporting services and postal services) Utilities (gas, electricity, water, sanitary services, etc.) Manufacturing (manufacturing of all products; excluding coke and refined petroleum products) Oil and gas (extraction of crude petroleum and natural gas, manufacturing of refined petroleum products, and service activities incidental to oil and gas extraction) (OTHER) Agriculture All other Resource industries (forestry and logging, fishing, mining and quarrying; *excluding oil and gas* and agriculture) Construction (construction of buildings, construction of industrial facilities, construction of transportation infrastructure, construction of utility projects, etc.) Retail trade (stores, e-commerce, hotels and similar accommodation, restaurants, etc.) Wholesale trade (distribution, import and export) Government (local/regional/national) Healthcare facilities or services (hospitals, outpatient services, physician/GP offices, medical testing and diagnostic services, home and community care services) Education (higher education, primary, secondary schools, pre-primary, online open course, language schools, driving schools, other educational support, etc.) Other [TERMINATE]	
S3	Does your company use data to make business decisions?	1. Yes 2. No [TERMINATE]	
S.4	Which of the following titles most closely describe your current title within your organisation?	• CEO, managing director, owner (if company under 100 employees) • VP of IT, CIO/CTO, head of IT • IT director/IT Manager • Head of information management • Head of analytics / insight • BI Architect • BI/Business Analyst • Data warehousing specialist • Database architect • CDO (chief data officer, chief digital offi • Digital Director / Head of Digital • Data Engineers • Data scientist • Data architect • Senior data engineer/senior developer • COO/head of operations • Other [TERMINATE]	Select appropriate respondents

Sections	Draft questions	Type of answer	What the data is used for
S.5	Which of the following best describes your role regarding the use and management of data that is used to support business/organisational decisions in your country?	1. I am the primary decision maker. 2. I am part of a team that makes the decisions. 3. I influence the decisions and I am knowledgeable about the topic. 4. I do not have any involvement. [TERMINATE]	Select appropriate respondents
S.6	How many people are employed by your organisation at all locations across your country?	1. 1 employee 2. 2 to 9 employees 3. 10 to 49 employees 4. 50 to 99 employees 5. 100 to 249 employees 6. 250 to 499 employees 7. 500 to 999 employees 8. 1,000 to 2,499 employees 9. 2,500 to 4,999 employees 10. 5,000 to 9,999 employees 11. 10,000 or more employees 12. Don't know [TERMINATE] **Analysis/quota groups:** 1 to 9 10 to 49 50 to 249 250 to 499 500+	Achieve survey quotas by size band. Smaller quota of companies under 1-9
S.7	Does your organisation provide any data or analytics products, services, tools or technologies to other organisations?	• *Yes [minimum quota of 200]* • *No, but planning to* • *No, and not planning to*	[minimum 200 respondents = yes]
1.0.	(as if q7=yes/plan) Which of the following data/analytics products/services or tools does your company provide (or plans to provide, if S7=plan) to other organisations? Choose all that apply	• data-based vertical solutions/ applications combining mobile, cloud and social technologies. • data-based products and services for online advertising and marketing • solutions and tools based on data analytics and/or elaborating real real-time streaming of social data across a range of social networks. • software and consulting services leveraging Big Data tools and access to the data held by the company itself. • premium datasets and/or access to complex and diverse data sources • data analytics, data mining tools and technologies (including for example computational linguistic, semantic search, natural language processing, artificial intelligence for data and text mining solutions) • business & IT services dedicated to the ICT tools used to analyze and process data.	

Sections	Draft questions	Type of answer	What the data is used for
		• cloud computing services, consumer and business IT products, services, and solutions delivered and consumed in real time over the Internet. • solutions supporting the flow and processing of information (interconnecting hardware systems, software and computers and other devices) • connectivity Infrastructure • data-based products and services to end-users in specific vertical markets high performance computing (HPC)	
Section 1	**Drivers and use of data**		
1.1	Does your company use Big Data and Analytics?	• Yes, using • Yes, Piloting or implementing • Evaluating or considering for future use • Not using and no plans to do so	Profiling user companies
1.2	What does your company mainly use data analytics for?	*Rank in order of priority maximum 3* • Better understand customer behaviour and expectations • Business and financial planning • Product / services improvement and innovation • Understanding market dynamics and our competitors • Optimise business operations • Risk management and regulatory compliance	Profiling user companies Rotate options
1.3	Is your company engaged in the following data-driven use cases?	Yes – Planning – No <See Use Case list> • Horizontal for all • Vertical specific	Measure take-up of use cases
1.4	Which of the following tools does your company use for analysis of data?	Mark all that apply • Structured reporting tools • Data visualisation tools • Multi-dimensional analysis tools • Advanced Analytics tools • AI tools (eg machine learning) • Business Intelligence applications. • Data warehouse platform (cloud or on-premise) • Big data platform, NoSQL database	Monitor the level of sophistication in the use of data Rotate options
1.5	What type of data do you collect and exploit in your company?	Mark all that apply • Machine-activity-based data (IoT sourced data, sensors, alarms) • Customer-activity-based data (transactions, page views, social media) • Location data (people, equipment) • Streaming data (video, audio, security cameras) • Data software.	Rotate options

Sections	Draft questions	Type of answer	What the data is used for
		• Data services. • Personal data (names, status, genomic, facial recognition • Business data (retail, consumer/credit, financial data…)	
1.6	Does your organisation participate in one or more data marketplaces, B2B or industrial data platforms?	• Yes. • No, but planning. • No.	Profile data users
1.7	Do you provide or use open data?	Yes use – yes provide – yes both – no but planning – no	Profile data users
1.8	What are the main barriers to adoption of data analytics and data tools?	Pick only 3 • Regulatory constraints on the use of data (privacy, GDPR) • Lack of business users' understanding of the data • Insufficient skills/ lack of expertise to implement the analytics • Difficult to make the business case/perceive ROI • High cost of tools and data systems • Lack of support from senior management • Siloed, incompatible or unreliable data	Profile data users
Section 2	**Data monetisation**		
2.1	Did you purchase data in the last year?	• Yes. • No, but planning. • No. • Don't know	Estimate data purchases
2.2	If yes to 2.1, how much did you spend for purchasing data last year?	Absolute number	Estimate data purchases
2.2.1	If you are not sure can you indicate a range?	• <10K€ • 10K€–20K€ • 20K€-50K€ • 50K€-99K€ • 100 to 499K€ • Over 500K	Estimate data purchases
2.3	Did you sell data in the last year?	• Yes , our internal data • Yes, internal data combined with external data • No, but planning. • No	Estimate data selling
2.4	If yes to Q2.3: how much revenues did you make from data selling last year?	• Absolute number	Estimate data selling
2.4.1	If you are not sure can you indicate a range?	• <10K€ • 10K€–20K€ • 20K€-50K€ • 50K€-99K€ • 100 to 499K€ • Over 500K	Estimate data selling

Sections	Draft questions	Type of answer	What the data is used for
2.5	If yes to Q2.1 AND/OR Q.2.3 which of the following types of data did you buy or sell?	Mark or that apply Specify buy or sell • Data software and services • Machine-activity-based data (sensors, alarms) • Customer-activity-based data (transactions, page views, social media) • Location data (people, equipment) • Streaming data (video, audio, security cameras) • Personal data (names, status, genomic, facial recognition) • Retail data, consumer/credit data, financial data	Type of data monetised
2.6	Do you make your internal data available for data sharing with third parties?	• Yes to our end customers (B2C) • Yes, with our customer and/or supplier in a value chain (B2B) • Yes, in a B2B platform with multiple partners • Yes with government/public actors (B2G) • No, but planning • No,	Estimate diffusion of data sharing
Section 3	**Data suppliers**		
3.1	How would you define your role in the data market?	***Rank by importance, max 3*** • Vertically integrated data supplier • Information and data broker • Data marketplace • Industrial data platform • Technical data tools and solution provider • Provider of data analytics • Provider of mobile and/or external apps for data. • Provider of ICT enablers, cross infrastructures, cloud computing.	Profiling data suppliers
3.2	What is your main target market ?	• Consumers + list of industries	Profiling data suppliers
3.3	What were your total revenues from data-related business last year ?	Absolute number + ranges	Calculating data companies revenues
3.3.1	If you are not sure can you indicate a range?	• <10K€ • 10K€–20K€ • 20K€-50K€ • 50K€-99K€ • 100 to 499K€ Over 500K	
Section 4	**Business impacts of data-driven innovation**	**Split questions**	
4.1	Did you achieve the following business benefits from data-driven innovation last year?	*Mark only those that apply* • Yes, Cost reduction. • Yes, Revenue growth. • Yes, Profit growth. • None of the above	Identify business impacts Rotate order

Sections	Draft questions	Type of answer	What the data is used for
4.2	For each item selected in 4.1: What was the improvement?	Profit = % improvement Revenue = % improvement Cost reduction = % reduction	Measure business impacts Rotate order
4.2.1	If no answer to 4.2, for each item selected in 4.1 : If you are not sure can you provide a range?	If selected cost reduction: • Decrease by less than 10% • Decrease between 5% and 10% • Decrease between 10% and 20% • Decrease between 20% and 30% • Decrease between 30% and 50% • Decrease more than 50% If selected revenue or profit growth: • Increase by less than 10% • Increase between 5% and 10% • Increase between 10% and 20% • Increase between 20% and 30% • Increase between 30% and 50% • Increase more than 50%	Measure business impacts Rotate order
4.3	Did you achieve any of the following business benefits from data-driven innovation last year?	• Time efficiency • Product/service quality improvements • Customer satisfaction improvement • Business model innovation • None of the above	Identify business impacts
4.4	If Yes to 4.3, for each item: How do you rate the improvement achieved thanks to data-driven innovation?	For each item selected • Very high improvement • High • Medium improvement • Low • Very low	Estimate business benefits for the data economy
Section 6	**Data professionals**		
6.1	How large is the IT department in your organisation in your country?	Absolute number As a % of total organisation	Measure data professionals
6.2	Data technical professionals are data engineers, data analysts, data administrators. How many are there in your company?	Absolute number If don't know: Can you estimate their share on total employees? %	Measure data professionals
6.3	Data business professionals are data scientists and business data analysts. How many are there in your company?	Absolute number If don't know: Can you estimate their share on total employees? %	Measure data professionals
6.4	Which line of business are the biggest consumer of data?	Rank top 3 • Finance/Accounting • Customer Service	Understand where data consumers are

Sections	Draft questions	Type of answer	What the data is used for
	(i.e. where managers and decision makers spend more than 10% of their time using data)	• Human Resources • IT • Operations, incl. procurement, production and supply chain • Sales and Marketing • Security, Risk and Legal • Product/service development, incl. R&D	
6.5	In the past 12 months, have you been hiring for data professional roles?	Data technical professional • Yes • No • Don't know Data Business Professional • Yes • No • Don't know	Understand sourcing of data skills
6.6	How difficult did you find it to hire for these roles? If yes to 6.5	• No difficulties • Some difficulty • Very difficult • Unable to hire	Understand sourcing of data skills
6.7	(excluding answer no difficulties in 6.6) What was the impact of the difficulty to hire data professionals?	• Marginal, we managed without • We had to use external consultants and/ or internal reskilling • Medium, projects got postponed • Large, delays and partially missed business goals • Very large, we missed or cancelled business goals	Understand impact of difficulty in hiring data skills
6.8	In the past 12 months, how did you source data skills you needed?	Select all that apply • Hire experienced data professionals from other companies • Engage/contract with external consultants • Hire graduates and train them up • Upskill internal staff • Hire and upskill professionals from other companies/ industries	Understand supply sources