UNIVERSITA' DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI "M. FANNO"

CORSO DI LAUREA MAGISTRALE / SPECIALISTICA IN

BUSINESS ADMINISTRATION - MANAGEMENT

TESI DI LAUREA

"ARE SME AWARE OF DATA-CULTURE? THE CENTER ITALY SME CASE"

RELATORE:

CH.MO PROF. ENRICO RETTORE

LAUREANDO/A: DANIELE STROZZIERI

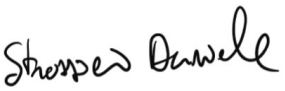
MATRICOLA
N. 1189715

ANNO ACCADEMICO 2018 – 2019

Il candidato dichiara che il presente lavoro è originale e non è già stato sottoposto, in tutto o in parte, per il conseguimento di un titolo accademico in altre Università italiane o straniere.
Il candidato dichiara altresì che tutti i materiali utilizzati durante la preparazione dell'elaborato sono stati indicati nel testo e nella sezione "Riferimenti bibliografici" e che le eventuali citazioni testuali sono individuabili attraverso l'esplicito richiamo alla pubblicazione originale.

The candidate declares that the present work is original and has not already been submitted, totally or in part, for the purposes of attaining an academic degree in other Italian or foreign universities. The candidate also declares that all the materials used during the preparation of the thesis have been explicitly indicated in the text and in the section "Bibliographical references" and that any textual citations can be identified through an explicit reference to the original publication.

Firma dello studente



Summary

Introduction

The term 'Big Data' refers to the collection of all that data and our ability to use it to our advantage across a wide range of areas, including business. Data is no more an invention, nowadays can cover everything from spreadsheets to photos, videos, sound recordings, written text, and sensor data.

Part of the reason for this explosion in data is the IoT (Internet of Things), that refers to devices that collect and transmit data via internet, and covers everything from your smartphone, smartwatch, even TV and refrigerators. Connected devices can not only connect to the internet, but they can also connect and share information with each other. Already, there are more connected devices than people in the world, it is predicted that by 2025, 41,6 billion devices will be capturing data on how we move, work, and operate the machine on which we depend (WEF, The state of the connected world , 2020). Soon, it is not unreasonable to imagine your refrigerator knowing when your milk is out of date and automatically telling your smartphone to order more in the next online shop. Despite this trend was growing year by year, Covid-19 ha gradually transformed the role of IoT in just few months by tracking and monitoring the disease around the world. Despite the rapidly growing presence of IoT devices, network, and analytical systems, many of these technologies are still in the early stage of development. In the next year, with 5G connection and more and more data stored, I will not be surprised to see applications such as real-time traffic coordination for autonomous vehicles or Natural language processing (NPL) within smart speakers, reducing costs and improving the user experience.

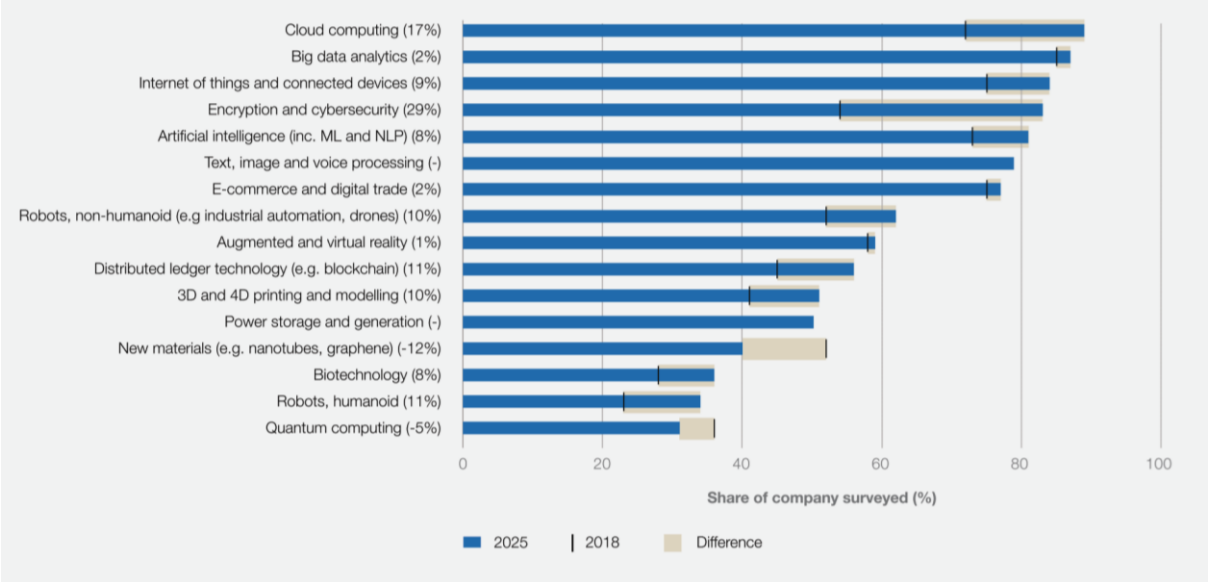


Figure 1 – Technologies likely to be adopted by 2025 (WEF, Future of Jobs Survey, 2020)

Figure 1 present the technologies likely to be adopted by 2025 by companies according to Future of Jobs Survey, conducted by the World Economic Forum. Cloud computing, Big Data Analytics (BDA) and e-commerce remain high priorities. Not only, but there was also a significant rise in the number of companies expecting to adopt encryption, non-humanoid robots, and artificial intelligence. The adoption of these new technologies will bring with them new job roles and skill sets.

Figure 2 - The top 20 roles increasing demand across industries (WEF, Future of Jobs Survey, 2020)

To govern a world contaminated by technology, there is the need of capable people to rule it. Trends show that, people that can work with data will be requested in every company, across every industry: data analyst, Big data, AI and Machine Learning specialist, digital marketing and strategy specialist, business developers, IoT and software specialists (Top 10) and many more (figure 2).

Alongside job roles demand, people should adapt to a different mindset too, indeed, among top 15 skillsets for 2025 (WEF, Future of Jobs Survey, 2020), having an analytical thinking and an innovation orientation will be fundamental, so as, active learning, complex problem solving,

critical thinking and creativity. We must start reshaping our job perception being aware of all that mechanical jobs that can be done by a machine will be replaced with jobs that require emotions, creativity, values, something that machines cannot afford (for now).

B. Top 15 skills for 2025

1	Analytical thinking and innovation	9	Resilience, stress tolerance and flexibility
2	Active learning and learning strategies	10	Reasoning, problem-solving and ideation
3	Complex problem-solving	11	Emotional intelligence
4	Critical thinking and analysis	12	Troubleshooting and user experience
5	Creativity, originality and initiative	13	Service orientation
6	Leadership and social influence	14	Systems analysis and evaluation
7	Technology use, monitoring and control	15	Persuasion and negotiation
8	Technology design and programming		

Figure 3 - Top 15 skills for 2025 (WEF, Future of Jobs Survey, 2020)

How is Italy performing in term of digital skill? According to the Digital Economy and Society Index (European Commission, 2020), Italy ranks 25th out of 28 EU member states in 2020 edition of DESI report.

	Italy			EU
	DESI 2018 value	DESI 2019 value	DESI 2020 value	DESI 2020 value
4a1 Electronic information sharing % enterprises	**37%** 2017	**37%** 2017	**35%** 2019	**34%** 2019
4a2 Social media % enterprises	**17%** 2017	**17%** 2017	**22%** 2019	**25%** 2019
4a3 Big data % enterprises	**9%** 2016	**7%** 2018	**7%** 2018	**12%** 2018
4a4 Cloud % enterprises	**NA** 2017	**15%** 2018	**15%** 2018	**18%** 2018
4b1 SMEs selling online % SMEs	**8%** 2017	**10%** 2018	**10%** 2019	**18%** 2019
4b2 e-Commerce turnover % SME turnover	**6%** 2017	**8%** 2018	**8%** 2019	**11%** 2019
4b3 Selling online cross-border % SMEs	**6%** 2017	**6%** 2017	**6%** 2019	**8%** 2019

Figure 4 - DESI 2020, integration of technology (European Commission, 2020)

Despite Italy has a good rank in terms of 5G preparedness, there is a significant gap in terms of human capital, indeed, compared to EU average, Italian people records very low levels of basic and advanced digital skills. The number of ICT specialist and ICT graduates recorded in Italy are also below compared to EU average. The digital skills gap can be explained by a low use of digital services, with 74% of Italians are regular internet users, among technology adaptation by companies, results are below EU average: 7% of enterprises use big data versus 12% of European mean, 15% cloud solutions versus EU mean of 18%. A deep gap is present in SME that sell online, 10% in 2019 compared 18% of EU mean. Despite this delay, something seems

be moving with “Italia 2025”, a government plan five-year plan that puts digitisation and innovation at the centre of a ‘process for the structural and radical transformation of the country’. The strategy started by digitalizing the public sector with new app and web services for citizen and with the promise to help in developing new ideas. Lastly, with recovery found plan, Italian government should incentive even more the digital transformation, at that point will be companies’ responsibilities to challenge the digital revolution.

The purpose of this master thesis is to increase companies’ awareness about Analytics solution, a trend that is increasingly present nowadays. In the following chapter it will be explained the definition and the characteristics of a digital company, to then discover how to create value through data, how to assess their initial value, what are the possible business models, implementation, and the main challenges. The next step will be analysing the current state of centre Italy small and medium enterprises, are they aware of data-driven culture? With a special focus on Teramo province SME. Lastly, there is a review of three different Big Data & Analytics solutions that can be implemented in every context (provided that a company has enough data) thanks to business intelligence, social media listening and process mining tools. A sort of BDA guide to start exploiting your data potential.

A big thanks to Professor Enrico Rettore for helping me during the writing of the thesis, his lead was precious along this final project. A big thanks to Professor Diego Campagnolo and Martina Gianecchini for their precious advices and references.

The digital company

An overview

The digital company model is observed when there is an appropriate combination of two factors: the orientation towards automation and the orientation towards innovation (Kenneth C. Lauron, 2010). This model assumes that all, or at least an important part of the internal management processes, is organized around the knowledgeable use of Information Technologies (IT), as well as services, products and customer relations are governed by effective leveraging offered by these technologies.
With the definition of industry 4.0, in which computers and automation come together in an entirely new way, enable the digital company (or smart factory) idea, in which cyberphysical systems monitor the physical process of the factory and make decentralized decisions. A company to be considered 4.0, must have the following features (Marr, 2017):

1) Interoperability
2) Information transparency
3) Technical assistance
4) Decentralized decision making

In digital companies, information stream is available anywhere in the organization, the company's critical assets are governed by digital tools; in its inter-organizational processes, it has set up a dedicated IT system to manage a significant part of the relationships that employees and management maintain, daily, with external actors (such as suppliers and customers).

The definition of digital company would be incomplete or meaningless without presenting the real actor of digitalization process: the IT system. The company information system can be defined as a set of interconnected elements that collect, catalogue, search, process, store and distribute data, transforming it into useful information to support the decision-making and control activities of a company (Kenneth C. Lauron, 2010).
It is accounted for:

a) Collecting data (input)
b) Data elaboration (transformation)
c) Report creation (output)

Data is the heart of an information system, but it is an original representation and does not interpret the phenomenon, while the final information is a data that has been subjected to a transformation process that has made it meaningful for the agents and the decision-making process.
As years passed by, data becomes the key to competitive advantage, the more a firm can leverage data, apply analytics, and implement new technologies, the more will succeed. Data and the ability to transform it into business value will become increasingly important in every sector. According to the International institute for Analytics, businesses that work with data will see US $430 billion in productivity benefits over competitors who are not using data by 2020 (Press, 2016).

As technologies will improve, the more future works sophisticated will be. The downside is that these technological revolutions might not create as many jobs as they eliminate. We will need more programmers, statisticians, engineers, data analyst and IT personnel to create and manage these sophisticated computers but not every factory line worker can easily shift gears and becoming a data analyst.
Boston Consulting Group forecasted that by 2025, as much as a quarter of jobs currently available will be replaced by either smart software or robots (Wakefield, 2015) while a University of Oxford study suggested that up to 35% of existing jobs in the UK could be at risk of automation within the next 20 years.

Although the digital company seems the ultimate business solution, it has its cons: often the theory is difficult to work in practice, such as the governance of inter-organizational relations is not always possible since, all the external counterparts of the company are not digitally developed, and many technologies haven't been standardize yet, usually making the company act as an early tester.

Correlation between IT investments and performances

"You can see the computer age everywhere but in productivity statistics" (Robert S. Solow 1987). Despite heavy investments and improvements in IT technologies, this is hardly reflected in productivity statistics. The benefits of IT are often considered as something intangible; from the organizational point of view it impacts simultaneously on several variables and it is difficult to separate the effects from those of other (non-IT) resources. For these reasons, their evaluation

requires not only efficiency measures but also effectiveness parameters. Benefits in terms of productivity are manifested when the technology has reached a level of diffusion close to 50%, and only from that moment, it would be possible to really measure the proceeds from these technologies (Fichman, 2004). IT integrations usually have an incremental flow, it cannot be evaluated in the short term, but only in the long term. This is also why they take an element of inimitability by competitors.

One of the first studies, dated early 1990s, to find a correlation between IT investments and productivity, identify two main positive factors:

- *Capital deepening*: when workers are supported by more capital, there is a greater investment in input.
- *Multifactor Productivity*: it is a phenomenon that occurs when technical advances in the production process or the quality of the output make it possible to increase the level of the output level without involving additional investments in inputs.

Another research conducted by Weill (Weill, 1992), who had the merit of defining a variable, called the *conversion effect* that relate IT investments and value creation. The higher the conversion effect is, the higher are the benefits deriving by IT investments; it assumes higher values when the following elements are high:

- The level of user satisfaction
- The degree of organizational turbulence
- The level of commitment of the top manager towards the project
- The company's experience with IT

It follows that two companies, given the same level of IT investment, and several other conditions, the one that performed best was the one with the greatest *conversion efficiency*.

Starting from the mid-90s we are witnessing a sort of miracle in terms of IT-performance linked to an increase in productivity in the ICT sector itself and IT technologies used in the manufacturing industry. This would be due to an acceleration of the *Moore's law* for which the power of the processors started to double no more after 18 months but after only one year. In particular, the sectors that benefited most from these technologies are those that have also been able to develop certain innovations at the level of business processes and company organization.

In fact, the competitive advantage would derive, starting from the development of IT resources through a process that also implies the development of *complementary organizational resources*, corporate *culture* and so on. One of the first studies on the subject was provided by Franke (1987) in the banking and insurance sector (ATM and PC), in his conclusions, he believed that the low returns on investments in IT were a consequence of the *learning curves* associated with the introduction of these new technologies and predict a trend change in the medium-term future.

However, not all that glitters is gold idiom, the risk of investing in IT is often greater than investing in other types of capital: IT pass through a very rapid process of obsolescence which involves a higher impact on fixed costs. Besides, a simultaneous introduction of changes at the IT level and at organizational level may be a risk: IT resources can produce value for the business only when an IT culture is absorbed by the organization, becoming a routine element of the value of an enterprise.

Lastly, Francalaci and Morabito (Morabito, 2008) showed how the complexity of IT as a business tool has a greater impact on small and medium-sized companies, where economies of scale and scope are more difficult to achieve, computer literacy is often not high, and management is rarely IT-oriented.

How to create Value from Big Data

The recent literature about Big Data & Analytics (BDA) has been focused on how it can be used to increase technical organizational capabilities rather than study its impact on organizational value (Varun Grover, 2018).

BDA is the application of statistical, processing and analytics techniques to big data for advancing business (Nicola, Ferreira, & and Ferreira, 2014). In the latest years BDA is becoming important to address unique costumers' requirements and solutions that ranges from the operations, marketing (with CRM), HR and Administration departments, deeply influencing the decision making.

Big Data is a new powerful source for a company, on one hand, for gaining a competitive advantage and for realizing a strategic business value, on the other hand, it must be followed by an organizational capital assets and human talent.

Companies reported that they spend more than 10 percent of their IT budget on data alone (Fleurence, et al., 2014), seeing BDA as a strategic asset to guide their decision making and improve their business processes and outcomes (Gopalkrishnan, Steier, Lewis, & and Guszcza, 2012). By analysing streams of structured and unstructured data can be answered questions that businesses have not even considered before.

Despite, there have been conducted much research about Big Data, few shed lights on how investment in BDA can yield tangible business value and how to assess a value to it (also called "data monetization" and "data valuation") (Varun Grover, 2018).

Before to translate BDA into value, companies must understand its key elements, necessarily to trigger a strategic planning, including anticipated returns, potential impact on competitiveness, current operational ecosystems, as well as opportunity cost and risk for the investment (ISACA, 2014).

We can start pointing out that data yields more value when they are integrated with data from other sources; moreover, data represent only the input of the whole value creation. To transform data in powerful knowledge and insight for the company itself and its customers, it is necessary combine them with their actual use: it is necessary to have a high-quality data, an appropriate information system, analytics tools, and mostly, talented human capital.

It is all about data strategy.

"Without a smart plan of action to use the data to produce business insights the data itself becomes a white elephant – expensive and useless" (Marr, 2017). As every business activity, it must have a well design data strategy at the base that guides companies to achieve the prefixed goals. This means defining the business-critical questions that companies need to answer, to collect and analyse those data to provide right solution. A strategy roadmap can then harness existing and new data assets. BDA initiatives without clear business goals and strategies will fail. Usually, IT teams use to focus on short-medium term (such as data storage, ownership, and integrity) rather than the business' long-term strategic goals and how data can help in reaching those goals.

To succeed, a company has to have pointed in mind what is the path that brings a data strategy plan to exploit its value and which key elements it has to touch to trigger innovation and a competitive advantage. A BDA strategy that depicts the objectives and approach to BDA is the glue that holds disparate BDA initiatives together. Lastly, a good data strategy is not determined by what data is readily or potentially available, it is about what your business wants to achieve and how data can help you get there**.**

Assessing value to Data

Success of big data strategies requires more than just the data asset, the techniques to collect and manage big data, and knowledge and implementation experience of analytics methods and tools. It also requires an understanding of the mediating process and mechanisms so that BDA can serve as a resource to harness strategic business value and keep firms competitive. To start this long path, it is of vital importance to address a value to data.

We can try to assess a value to BDA using the resourced *based view framework* and through the *Functional versus Symbolic value.*

Resource-based view applied to data.

The VRIN framework, based on a *resource-based view*, is a test to assess if the company has the right resources and capabilities for sustainable competitive advantage, where VRIN stand for *Valuable, Rare, Inimitable and Nonsubstitutable*. While the first two tests determine

whether a resource or a capability can support a competitive advantage, the last two determine whether the competitive advantage can be sustained (Arthur Thompson, 2015):

1. *Valuable*: are the company able, through its BDA structure, to generate valuable insights to exploit new business opportunities and/or mitigate competition threat? If the answer is yes, then BDA can truly become a valuable resource for business; the more a company can generate data and transform them in useful insight, the more can exploit its business process improvement, product and service innovation, customer experience and the overall internal organization; BDA can analyse both internal and external sources of data, the mix can generate a highly competitive advantage; for example, I remember, during the *organization development and behaviour* class, the case of an entrepreneur, of an important local company, working in the cooling system field, had run out of stock in Netherland since they didn't forecast that there during that summer, temperatures had been one of the highest recorded; imagine the company had a proper BDA linked not only to internal processes but even to external sources such has countries' temperature, probably they would have had adequate inventory capacity to the new condition, in this way to dramatically increase their annual sales;

   Therefore, companies can use insights about customers and markets to leverage customer satisfaction and loyalty, lock in customers and suppliers and/or create a niche market; lastly, companies can enrich their business intelligence, develop a dynamic organization structure to better respond to a world that change months by months.
2. *Rare*: are the company's Big Data content, analytics capabilities, or a mix of them rare? Can competitors acquire or possess them? We define BDA rare when competitors struggle to acquire or possess the same or similar capabilities; we can evaluate two key answering to this question: a) by the content of big data itself, b) by analytics capability (VARUN GROVER, 2018); the former point can be summed up thinking to Salvagnini S.p.a, an Italian company that started installing sensors on their machineries long time ago and before their competitors; the company in this case realized a competitive advantage driven by the large amount of data that were stored through the years that made their experience (learning curve) ahead compared to their competitors; the latter point concerns a business context where has been developed analytical knowledge, human talent and experience in managing big data and in using proper tools to exploit relevant insights; maintaining a company data assets competitive and rare is an ongoing process that requires continuous assessment and a continuous update of models and skills;

3. *Inimitable*: how is costly/difficult for a competitor without a BDA capability to develop it? Is it almost impossible to imitate what the company has already achieved? The reasons why BDA can be costly to imitate is typically linked to the IT infrastructure and to analytics algorithms; the former is linked to company strategic plan, even many companies already own an IT infrastructure, many are unlikely to sustain an efficient analytics task and with timeliness, moreover it represent a relevant investment that many companies see as unnecessary compared to manufacturer's tangible assets; regarding analytics algorithms, it is a learning by doing process where the algorithms grow aside with data it receives and the ability of skilful employees themselves; besides, algorithms adapts to the company context, structures and story, something unlike to be transfer elsewhere;
4. *Nonsubstitutable*: does the company organization business strategies and culture support the development of valuable, rare, and inimitable BDA resources? Even if the implementation of BDA is IT-oriented, the BDA strategy must be business oriented, linked to firm's long-term business strategy; here the key value is driven by a perfect match between the BDA technologies and organizational resources and capabilities; all the company organization must move along with the data strategy since the ability to generate powerful insights and new opportunities grow along with IT implementation and company's learning process; in early covid-19 situation, Amazon provided a great example in using its data-driven strategy to forecast three weeks before what was going to happen in Europe, they were able to organize inventories, making stock of essential products and getting rid of superfluous ones; moreover the use of transversal team made possible to reorganize the full internal organization, included recruitment activity, with new online test assessment and application format.

Strategic Business Value: Functional versus Symbolic

In addition to the *resource-based view framework*, the strategic value created by BDA can be assessed in two different dimensions: functional and/or symbolic, where we define:

- *Functional value* referred as the performance improvement directly resulting from adopting BDA, such as market share and financial performance.
- *Symbolic value* referred as the "signalling effect" or positive externalities deriving from investing in BDA; the outcomes are generally quantified in positive brand image and reputation, mitigating environmental pressure.

From a strategic fit perspective, functional value can be seen as the fit between technology and organizational task while symbolic value the fit between technology and organizational environment. It is important to consider that the two types of strategic value may not be mutually exclusive (Varun Grover, 2018).

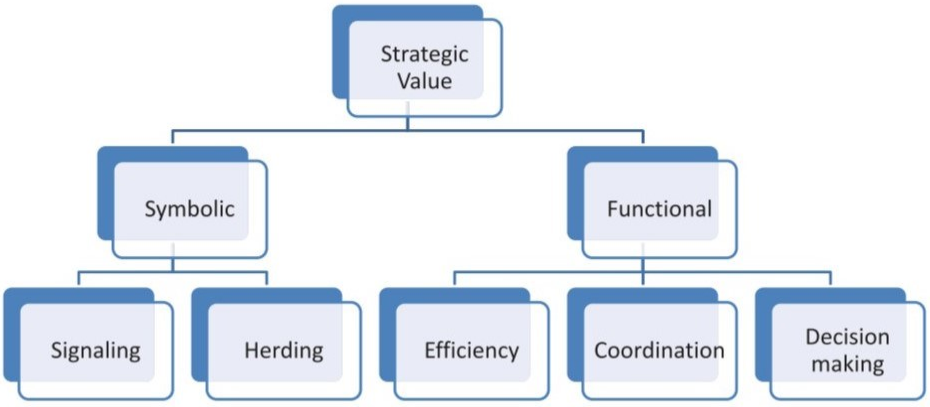


Figure 5 - BDA Strategic Value (Varun Grover, 2018)

We observe functional value when we convert assets to tangible (or intangible) value, this is possible when BDA and complementary resources are synergically linked to create unique capabilities that empower value creation processes. The results are shown through an improvement of efficiency, coordination and decision making, as long as performance improvement, cost saving, innovation and a better understanding and relationship with customers.

While functional values can be shown through an overall increase in company's performances, symbolic values be a positive externality. With a good fit BDA organization and activity, markets receive a positive signalling effect of organization innovation that can be described just looking at the market capitalization rank, where at the first ten positions, 7 out of 10 companies belongs to tech industry (CorporateInformation, 2020). Signalling theory indicates that the communication of private information by an agent reduces information asymmetry between agent and the principal; moreover, even a reputational effect can be disclosed, showing company commitment in innovating in new technologies and tools to face new challenges of the future.

McAfee and Brynjolfsson (McAfee, 2012) conducted a research for assessing direct data value, interviewing executives of 330 public North American companies about their BDA organizational structure and performances. Results showed how the more companies characterized by a data-driven philosophy, the better they performed on objective measures of

financial and operational performance. Moreover, companies in the top third of their respective industry in the use of data-driven decision making were approximately 5 percent more productive and 6 percent more profitable than their competitors.

Strategic roles of BDA

Combining symbolic and functional value of BDA strategic value, it is possible to illustrate four strategic roles addressed to BDA:

- *Image Builder:* combining and high symbolic value with a low functional value BDA can be a great image builder to generate a positive signal to stakeholders.
- *Reactive Defender:* when symbolic and functional value are low, the firm is probably adopting a defensive stance against competitors rather than starting itself the innovation process.
- *Performance Enhancer:* when symbolic value is low and functional value is high, BDA may assume a Performance enhancer role, resulting in a boost in productivity and in internal processes.
- *Strategic Transformer:* when symbolic and functional value are high, companies may enhance both internal and image value in the marketplace, exploiting the full beneficial of a BDA strategy.

Figure 6 - Strategic roles big data analytics

The Path to create value from BDA.

Creating value through BDA is not as simple as it is described since it required a huge investment regarding both IT-infrastructure and talented human capital. According to Grover, Chiang, Liang, and Zhang conceptual framework (Varun Grover, 2018), it is possible to compare strategic changes dimensions by Pettigrew and Whipp (Pettigrew, 1991) with a BDA strategy to create value:

1. *Content in BDA*: in this field, the strategic change comes from the development of *dynamics capabilities,* that is an ongoing capacity of a company to modify its existing resources and capabilities or creates new ones (Arthur Thompson, 2015).
2. *Process in BDA:* change should be made following the *IT-value modelling* which describes how IT investments enable an impact on both processes and variance representation.

Figure 7 describes how value is created by Big Data Analytics, mainly by two fundamental processes: *capability building* and *capability realization.*

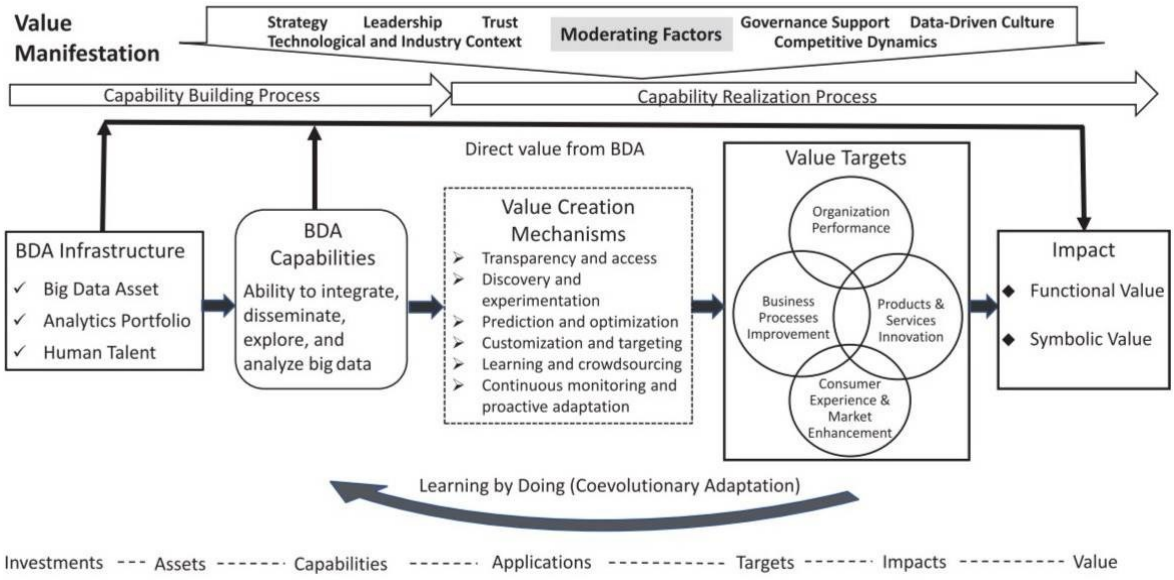


Figure 7 - Value creation by BDA (Varun Grover, 2018)

Capability Building Process

It is well known that capabilities are built by a *learning by doing* path or, in alternative, they are acquired from the outside. In a BDA context, capabilities are meant as the ability to both manage and analyse data to create new insights. The path for building BDA capabilities starts from coordinating the BDA infrastructure, composed by big data asset, analytics portfolio and human talent, and the capability to exploit new insights: asset and knowledge must be mixed.

1° step: establish the infrastructure.

BDA infrastructure is the first investment a company must achieved in order to start a BDA value strategy; it is possible divide the relevant investments in:

Big data assets

A BDA infrastructure typically includes *data sources* and a *platform* that has the aim to collect, integrate, share process, store, and to manage big data; *here*, we try to classify the major IT systems, useful for decision making, according to a hierarchical way or by the management areas (Kenneth C. Lauron, 2010).

The first macro classification provides various types of IT systems arranged along with pyramidal hierarchical view, starting from the base to the top:

- At the base of the pyramid we find the *transaction processing systems* (TPS); these are computer-based information systems (CBIS) that keep track of all routine information in organizations; These systems are often central to companies, as they constitute the most important information base of all subsequent information of the organization.

- The *management information systems* (MIS), it acquires data from the TPS and generate reports for the organization's management through structured pre-period representation about company's operations.

- *Decision support systems* (DSS), are account for supporting non-routine managerial decisions, specifically aimed at supporting corporate decisions; with them, it is possible

for managers to simulate hypotheses to verify or test the validity of management; DSSs use data from TPS and MIS and often add data from external sources, such as commodity prices, market rates, currencies, and so on.

- At the top of the information systems hierarchy we find the *executive information systems* (ESS), reserved for senior management at the highest level of the organization; these systems are at the full service of non-routine strategic decisions; ESS often offers digital dashboard throughout which senior management can have management trends under control by aggregating internal and external information and summarizing it with the most important indicators that must be checked.

The second macro classification presents the types of IT systems according to the most important management areas of the company:

- *Enterprise resource planning* (ERP) systems are dedicated to the integrated management of all company data using a single centralized database on which the functional areas of company management are interfaced.

- *Customer relationship management* (CRM) systems: these are systems dedicated to the management of customer relations, overseeing the commercial area and the marketing activities, pre- and post-sales services; in many cases the CRM feeds important data warehouse systems where all the data on the history of customer relations are stored, from which marketing and commercial management plans are extracted.

- *Supply chain management systems* (SCM): are those information systems dedicated to the management of all the processes of acquiring resources and raw materials and their transformation, including the final process of shipping goods to customers (in the case of tangible goods), but also for procurement and stock management, production planning, optimal warehouse management.

Analytics portfolio

Investing in a strong analytics infrastructure is fundamental to exploit the true potential of data analytics, by prioritizing meaningful data and integrating them to generate meaningful insights. The company must make decision in term of data aggregation, transformation, tool selection and analytics models that suits most to the company's needs. Generally, the output of this process provides ready to use application such as consumer sentiment analysis, financial risk modelling, marketing campaign analysis, cross-selling opportunities, employees, and performance optimization.

Human talent

Big data and Analytics portfolio assets would be useless without a proper human talent infrastructure, becoming a critical element of the BDA value creation process. To design and implement a BDA strategy and exploit its value creation, company's' employees have to possess the right skill and a data driven culture. This represent one of the biggest challenges for firms since talented employee in this field, nowadays, are still rare and their profession is not well design.

The Chief Data Officer

In many organizations, the skills required to select, purchase, install, manage, and maintain these assets often are referred to in the following ways: *Chief information officer*. Usually, big companies may see the benefit in hiring a Chief Data Officer (CDO), a well determinate role inside the company who is responsible for data as a companywide asset, mostly when the magnitude of data becomes difficult to manage. Business opportunities, monetization, data security, privacy is way too complex and fragile to be discussed at the bottom level of the hierarchy. A centralised figure that manages, decide about data strategy is needed to avoid inefficiencies and harmful expenses, therefore companies engaged in serious data projects ought to have a CDO at the top in charge of data management structure.
Gartner statistics showed that by 2019 (STAMFORD, 2016) it is expected that 90% of large organizations will have hired a CDO. Ideally a CDO would have both a technical and business background (Marr, 2017).

For smaller companies, a CDO may not be necessary or possible, in this case a similar function will be carried out by the leadership team itself, perhaps with the help of an external data consultant.

Core qualitative and requirements of good CDO includes:

- *High level vision*: it is responsible for the big picture of the data priorities and strategy for the entire company; it has a high level of understanding of both data strategy and the business needs in line with the company strategy.

- *Implementation:* it is responsible for data strategy implementation at every level of the hierarchy in the company, it ought to be skilful in managing large teams on technical projects and build a business case around technical projects with many variables and uncertainties.

- *Data accuracy, security, and privacy:* it has the full responsibility in collecting and maintaining accurate data, ensuring data security, and devising and implementing data privacy policies; the CDO becomes the ethical conscience of the company, defining and following ethical guidelines about how data are collected and stored.

- *Identifying business opportunities:* the CDO is the person most responsible for identifying business opportunities discovered through data, moreover, CDO should increase revenue or decrease costs based on the information learned through data initiatives.

- *Data-driven culture leader:* CDO must also be the 'charismatic leader' for a data-driven culture through convincing and including everyone in the company about the importance of data, security, and privacy as well as its the business value.

- *Data as commodity:* CDO will also be the visionary responsible for recognizing the value in monetizing an organization's data.

Data Analysts

When we think about BDA professionals, the first names we came up with are data scientists, developers, programmers, analysts, and modelers that can serve significant roles in both

managing and analysing data, but usually they lack strategic view. The most intensive use of people occurs during the input (design of BDA strategy) and output (interpretation of results) stages. Nowadays, the market is asking for a diversification of skill, where Data scientists are hardly defined. Indeed, "data scientist" today can be used to describe anyone, from a data engineer who set up behind the scenes, systems that collect and store data to statisticians who analyse numbers. Therefore, to make this subject clearer, below it is described the five essential data science skills that employees must have to turn data into powerful insights (Marr, 2017):

1. *Business skills:* any data scientist should be able to understand what keeps the business ticking, what cause it to grow, the key business processes and what is the company advantage against competitors; along with business; communications skills are also important to extract the maximum value from data and then to be able to communicate them to all parties involves.

2. *Analytical skills:* it has to have the ability to spot patterns, links and understand the causal effect among data until they produce the desired results, as well as to be able to interpret reports and data visualization.

3. *Computer science:* the ability to create sophisticated machine learning, natural language processing algorithms and be able to use the latest data software solutions.

4. *Statistics and mathematics:* statistician's skills are relevant to defining relevant populations and appropriate sample size to maintain a high data quality appropriate for the company needs.

5. *Creativity:* nowadays it represents the key skill, especially when working with data, the possibility to discover new ways to create values from data; it is the engine that starting from the technical skills use them to produce powerful insights and business opportunities.

The ultimate combination of skills can be hard to find, it may make more sense to combine skills in a more creative way that works for your business, such as match an employee with relevant analytical skills with someone who has great communication skills to communicate them to a wide audience.

2° step: developing BDA capabilities.

Having a good IT infrastructure is only an input, firms need to have strong capabilities of integrating, managing, sharing, and analysing big data to transform them from "raw elements" to value creation (the desire output).
BDA offers to the company a forward-looking view, enabling firms to anticipate and execute future opportunities based on real-time insights discovered from high-volume streaming data sources, current events, and ongoing business processes. Indeed, BDA includes: (a) descriptive analysis that reports on the past; (b) predictive analysis that develops models based on past data for future prediction; and (c) prescriptive analysis that uses models to specify optimal behaviours and actions (Varun Grover, 2018).

There are two main routes to develop data-related competencies. One is boosting your in-house talent and the other is outsourcing the data analysis.

Creating a data culture

"A data culture is about everyone across the business understanding the value of data and how it can help the business succeed" (Marr, 2017). A culture that embraces data and evidence-driven approaches to business decisions, and governance that delineates responsibility and accountability for data, are both drivers for BDA value creation.
A good way to build a strong data culture is to engage key personnel in the data strategy process, both in developing the strategy and its implementation, and let them become advocates, in this way they may spread the culture along all the structure. Communication is fundamental, leaders should spend time engaging people in the data strategy and demonstrate the positive impacts of data through case studies, specific books/articles and mostly with numbers! Company culture and governance are critical to creating the right environment for BDA.

Despite what has been written above, a top-down approach cannot work if there is a cultural shift resistance within the organization. Therefore, it is necessary to introduce the concept of *absorption capacity* and how it drives companies to achieve value creation.

Figure 8 - Absorption capacity as link between IT integration and business performance

Organizational absorption capacity

The concept of *absorption capacity* represents a dynamic concept; it is the ability for a company to identify, assimilate and exploit external knowledge for commercial purposes, triggering a source of innovation resulting from reengineering rather than ex novo (Levinthal, 1990).
Absorption Capacity is highly visible in those organizations where they learn how to effectively use the new knowledge, available for business processes, and use it for implementing change, capable of increasing companies' competitiveness.
The absorption capacity of an organization is not the simple sum of the skills of each employee, but it is the cooperation between several parties involved that makes the absorption of new knowledge and the process of change effective.

The *absorption capacity* can also be defined according to the *resource-based view*: the ability of a company to translate a change defined by a combination of resources into organizational performance. To develop an excellent absorption capacity, it is necessary to continuously improve and invest in the knowledge that a company has acquired on certain resources, especially in the IT field, where the rapid obsolescence of skills and changes in the organizational could compromise the firm's advantage. Therefore, a lack of investment in each area of expertise can compromise the future development of technological skills dedicated to that area.

In the IT field, *absorption capacity* is seen as a dynamic element characterized by four orientations (Kenneth C. Lauron, 2010)

1. *Process orientation:* a close relationship between IT and business processes makes technology becomes a powerful enabler of value creation without the company being process-oriented; the continuous exchange of cross-functional information allows greater speed in planning and control activities.

2. *Training orientation:* acquiring new knowledge related to new IT domains, investing in training, is strategy, in supporting evaluations about new opportunities offered by future technological developments; the main IT trainings are *lecturing, formal training and on-the-job;* in particular, the combined use of formal training and on-the-job is essential for the development of new application software and its adoption at an organizational level.

3. *The orientation to change:* defined as the degree to which a manager and workers can influence change, in contrast with how much organizational inertia they can oppose (Damanpour, 1991); a change-oriented organization promotes an organizational culture that supports the exploration, assimilation and application of new technologies and associated business solutions; these activities include the analysis of the surrounding environment aimed at acquiring new knowledge and a cross-functional dialogue with colleagues and other members of the organization.

4. *Orientation to flexibility:* the ability to adapt or change products and organizational processes in a relatively short time and at low costs; IT represent a strategic driver in promoting those benefits, it favours the adaptation process at the organizational level, throughout training policies, sharing and integration of knowledge at every company level.

To sum up, implementing a cultural shift in an organization is not a quick and easy job. It takes time and dedication to get companywide buy-in and it requires a shift in mindset away from gut-based decisions or the 'this is how we have always done it' mentality. Moreover, thanks to absorption capacity, the degree of rootedness of the effects of IT integration in its business processes, becomes difficult to imitate by competitors due to the complexity of managing all the variables that made up the change process.

Outsourcing data analysis: BDaaS

Since building up an IT infrastructure and the correlated capabilities need lots of time and resources, small and medium enterprises have the tendency to make extensive use of outsourcing. In the latest few years, more and more business started offering cloud-based data services to business customers, such as Amazon with AWS and Microsoft with Azure.
This fast-growing market is commonly known as "big data as a service" (BDaaS), and it has been estimated that the business IT spending in cloud-based services activities will increase from about 35% by 2021, around US $88 billion, $30 billion address to BDaaS (Marr, 2017).
BDaaS can refer to a wide variety of data function, from supplying data to supplying analytical tools to carrying out the actual analysis for the customer and providing insights via reports. This new market is a great solution for SME, they can simply rent the provider's cloud-base storage and analytics services, and benefit from vast data sets that may otherwise not be able to access. Moreover, BDaaS services drastically lower or eliminates upfront infrastructure costs, lowering barriers of entry and removing many hurdles associated with implementing a data strategy.

Organizational implications

How can a company keep a flexible structure when most of the investments for the IT infrastructure are represented by fixed costs and licenses? What is the degree of centralization of decision making about IT infrastructure? How is the information flow coordinated? Who identifies and defines the needs of the various divisions? Do they themselves or should these decisions be controlled by top management? These are just a few questions that management must answer during an IT development strategy.

Regarding the degree of flexibility of a company, it is possible to attribute this capacity to the term *scalability*, that is, the ability of a PC, product, or system, to expand to serve many users without collapsing (Kenneth C. Lauron, 2010). It is managers responsibility to define the threshold of acceptability about response times of the systems and the degree of availability for critical systems, to achieve the prefixed goals and maintain while keeping the desired level of performance. If the investment is oversized, the infrastructure may remain unused and slows down the firm's economic and financial performance. If it is undersized, it can prevent the development of important product/services and opportunities hurting the company competitiveness.
To make reasoned decisions about buying your IT assets or renting them from external suppliers, we can recall the *rent or buy* assessment.

First, it is good to consider the competitive strategy of the company, what objectives it wants to achieve and how long it would take, the current evaluation of the IT assets the company already has and the strategic positioning of competitors.
A fundamental analysis is represented by the *Total Cost of Ownership (TCO):* it can be used to analyse direct and indirect costs to determine the actual costs of the various technological implementations. Are you aware that the TCO of a PC can exceed three times the initial purchase cost? The initial cost of hardware and software represents only about 20% of the TCO (Kenneth C. Lauron, 2010), so managers must pay close attention to administrative costs and hidden costs not disclosed since the beginning.
To reduce TCOs, the company may centralize and standardize hardware and software resources. In a centralized structure, the systems could be managed by a central office capable of solving any problems and exploiting BDA potential. By decreasing the number and models of PCs and software available for staff, companies may also reduce the number of technical personnel needed to support the whole infrastructure.

To develop an IT strategy, management must stimulate the sharing of information through people interactions. It may be done using:

- *task forces*: temporary bodies dedicated to the analysis and solution of a significant business problem.
- *committees*: permanent but discontinuous bodies which are entrusted with the tasks of aligning information between the various company functions.
- *liaison roles*: have the function of junction between one department and another, between one function and another.
- *project/product managers*: the use of organizational roles which are entrusted with purely tasks of transversal coordination action in the organization.

In more complex cases, the global set-up could change and migrate towards a matrix organizational model in which a constant organizational alignment is expected both horizontally and vertically, made through people who hold ad hoc roles and functions.
BDA requires data to be collected and analysed with centralized governance, which ensures that all big data projects within an organization apply the same standards, protocols, methods, and tools. Meanwhile, organizations can also benefit from having a federated model for delivering BDA projects, which can improve the speed of analytics and ensure that gained

insight is available to decision makers. Therefore, an organization needs to establish a governance framework that standardizes the operations of BDA across diverse operational areas while supporting federated project delivery (Marr, 2017).

Capability Realization Process

"The amount of big data should not make it a big deal. It is the ability to actually do something meaningful and valuable with it (Fisher, DeLine, Czerwinski, & and Drucker, 2012)*"*. Nowadays, BDA has become a key element of sustainable business models, that derives from the ability to generate actionable insights and apply them to business practices to accelerate innovation, drive optimization, and improve business performance (Varun Grover, 2018): from discovering new services and product, to improve lead times, to a better understanding of customers and market trends.

Value Creation Mechanisms

Before explaining which are the possible BDA applications (Targets), it is important to make clear what are those mechanisms that creates a linkage between BDA capabilities (stressed above) and value targets that the company wants to reach.

1. *Transparency*: the ability to generate descriptive data to create a dashboard that may provide real-time access to company current performances and market status; not only the top managers would be aware of company strategic position, but even the rest of the hierarchy, improving communication and the internal culture.
2. *Discovery:* usually it is the most emphasized aspect of BDA, the ability to provide deep and pragmatic insights to generate important outcomes for various BDA targets; the ability to discover new solutions for the company itself and its customers, along with new possible markets.
3. *Forecasting:* the ability to generate, on the one hand, probabilistic outcomes for the future and determine "best path" to follow; on the other hand, the ability to warn and adjust for situational abnormalities, data that derivers from IoT applications inside the company.
4. *Customization:* the possibility to generate powerful report to enable a better segmentation of customers and markets; the result is in a major customization of

products and services, as well as targeting different market segments with digitally versioned products.

5. *Learning:* the ability to continue to learn from newly fresh data and to apply it to new solutions; nowadays, machine learning has been applied to many different contexts from online education to automated vehicles, while crowdsourcing is being used for predictions and leveraging innovative talent.

Targets of Value Creation

Value targets are defined as the output of a BDA strategy, giving to data a visible value creation and/or internally improvements. In determining value targets, it is important to involve the views and priorities of stakeholders too, not only what data says. Here, we identify five distinct targets of BDA value creation:

1. improving decision making
2. improving operations
3. product and service innovation
4. customer experience
5. monetization of data.

First, BDA enables companies to collect better market and customer intelligence by improving organizational decision models: analysing streaming data in real-time may improve performances such as providing a just-in-time markets scenarios or inventory status. This can be accomplished by providing broad and consistent access to data across all the organization, supported by powerful and reliable structures to act on the data.
Second, BDA helps companies by improving the effectiveness, efficiency, and productivity of business processes, which leads to better execution and less time spent on process breakdowns: from tracking machine performance to optimizing delivery routes to even recruiting the very best talent, big data can improve internal efficiency and operations for almost any type of business and in many different departments. Sensors integrated in IoT machineries play a fundamental role, not only to track the productivity flow, but mostly to prevent breakdowns, stock outs and loss in efficiency. At this point, a great supporting system is needed to support

the continuous improvement of business processes. There are three types of business process analysis: validation, verification, and performance, where all require a large volume of process and event data (Vera-Baquero, Colomo-Palacios, & and Molloy, 2013).

Third, BDA may create value for product and service innovation by analysing customer clickstreams or purchasing patterns. Product reviews, consumer discussion forums, tweets, Facebook pages data has become increasingly analysed for gaining business insights. These virtual community and a new key figure as influencers, can influence each other. By analysing these online consumers, the company may enable promotional strategies and opportunities to improve a product or even come up with a new one. A company may identify the frequently reported product flaws/issues or consumer desired features of certain products or services, which provide insights for product/service innovation. Some companies nurture crowdsourcing innovation platforms to facilitate this.

Another possibility is to find and engage the highest relevant people in a community (Influencers) for better targeting or to let them test company's products: it is possible to assist to a change in communication strategy, closer to final customers through a person they trust to.

Fourth, BDA can also deliver a better customer experience and more competitive.

services, resulting in higher customer satisfaction and retention. Through social network analytics firms can acquire new customers and identify potential competitors through online comments and link social activities, profiles, and purchase history of existing customers to create a better and more comprehensive understanding of their customers through BDA.

Companies may identify customers specific needs answering to questions such as "what do other customers like me think of this product or service?" or "What's the top features attributed to my products and services?". Answering such questions not only increases the understanding of what customers are looking for but may also identify other communities to link company's product to. For example, thanks to software like Awario or Brandwatch, it is possible for a company to track consumer sentiment analysis and discover new topics linked with our products; during a Marketing project about Gluten free product, it was possible, to identify a linkage between gluten free products with the gym /healthcare world, where protein elements of some gluten free pasta could match athlete's needs, shifting the product preposition from "food for ill people" to "protein food for a healthy body".

Five, BDA also provides the opportunities to build BDA itself in a product offering.

There are mainly two streams of revenue: the former is about to transform data as an asset that increases the overall value of the company, the latter is about the ability to monetize data by selling back to costumers or other interested parties. Since data becomes a core asset, the need for careful data governance becomes even more pressing (Marr, 2017).

Above has been treated the former way to use data as a core asset, while the latter it is increasing its business relevance year after year. If a company generate or gathering data, it may consider offering it in a second market (if any). For companies such as Facebook and Google, BDA represent their main business stream of value: they use to offer free access to their platform and services for free, apparently, customers exchange that free services for a most valuable currency, data! Through data gained by customers, these companies resell them to companies that want use their advertisement platform to target companies own customers; at this point it easy to understand that the companies that use these Ad platforms generate an additional stream of data that will increase Facebook and Google BDA value creation (decision making, improve in operations, etc).
This latest example showed how these targets are interrelated, as many initiatives involve multiple targets.

Why data strategies fail.

Despite the growing popularity of BDA, many organizations have failed to reach their strategic goals after investing substantial resources while a very little number has been able to exploit BDA value creation and to successfully write business cases. Many companies fail to execute their data strategies successfully, sometime because it is no more achievable, or it may be too vague, or ill-defined that nobody knows where to start. A survey conducted by *Bein & Company* about executives at more than 400 companies, reported that only 4 percent of companies were good at analytics; 56 percent did not have an appropriate system to capture the data that they needed or did not collect useful data; and 66 percent lacked the right technology to store and access data (Wegener, 2013). With Covid-19 emergency, more and more companies are going to leverage their IT infrastructure and BDA skills, many of them would succeed in implementing the new strategy and many would not, below I am going to list the main pitfalls a company could encounter and how they could mitigate them.

Communication

Communication, or lack of it, is a big stumbling block. If employees are not involved into the thinking behind a strategy, they may not agree with it or, even worse, not to believe in it. This can result in a decrease in performance or low morale. As well as individuals, lack of

communication between departments can also be an issue. To generate value from BDA, data functions need to be able to communicate successfully with other departments and leadership, and vice versa. Companies should look to build and maintain strong links between whoever is analysing the data, whoever is reporting the insights, and business leads. It is also important that employees feel they have a voice in the strategy implementation. Therefore, organizational links (*liaison roles, cross-functional teams, project/product managers*) to boost communication is vital when your data strategy is running.

Data Quality

Inaccurate, outdated, or inconsistent data with other sources of information lead to wrong decisions, withdrawal of products from the market, financial losses. Data quality problems add a further layer of complexity to the real-time and actionable use of big data. The main common data quality challenges are:

1. *data are often noisy, erroneous, or even missing* misspelled words, incorrect grammars, tone misunderstanding in a social media content, as well as data recorded by mobile and wearable devices, poses significant challenges for computational linguistic and technical analysis.

2. *ensure data trustworthiness*: do you trust in your data? Having truthful data becomes too difficult as they grow exponentially, it is a much bigger challenge than volume, velocity, and variety in BDA; how many TripAdvisor, Amazon, eBay reviews are sincere and original? Well, it is estimated that approximately 20–25 percent of online consumer reviews are fake (Qiao, Zhang, Zhou, & Wang, 2017).

3. *partially or even unlabelled data*: in the real-world data are often unlabelled or only partially labelled. A lack of labelled data causes significant problems when using machine learning algorithms for model building since the less labelled data a company has, the more algorithms will generate false positive/negative. The company need more research on machine learning with partially labelled data, such as semi supervised machine learning.

4. *data imbalance* occurs while training a classification model with samples of data, these do not represent different classes equally.

Most of the errors result from mistakes made during data entry. The incidence of such errors increases as companies move their business to web solutions, allowing customers and suppliers to enter their data on websites that directly update internal database.
For example, during the first lockdown, Italy national social welfare institution (INPS) director, Pasquale Tridico, reported that in response to almost 1 million of 'Bonus €600' requests (aimed to sustain Italian entrepreneurs), around 250.000 citizens failed to write the right bank account number, around 25% of initial requests (Baratta, 2020). *Data cleansing* is a "must be done" process that companies must pass through. Data cleansing consists of activities aimed at detecting and correcting data in a database that are incorrect, incomplete, poorly formatted, or redundant (Kenneth C. Lauron, 2010). Data cleansing not only corrects errors but establishes consistency between data sets that originated in different IT systems and that needs to be done prior to any analytics. Therefore, data hygiene through data cleaning, filtering, and selection to detect and remove noise and abnormality from data automatically becomes essential. The more consolidated, cleaned, accurate, and consistent data a firm has, the more likely it can make better decisions.

Data Integration

The true value in incorporating a BDA initiative comes from integrating and unifying diverse sources of existing and new data, both structured and unstructured, where the variety dimension of big data determines the high complexity of data integration. Many companies fail in exploiting BDA potential due to a disaggregation of sources, maybe they still use different database for each department that makes unable to transform big data in wide data. Independent data sources may provide data collected at different time intervals and with different levels of granularity and aggregation, on the other hand, not all the organization would be aware and able to use data all in once. Could happen that due to a lack of data integration, a company may skip business opportunities or mitigate risks. For example, data coming from marketing department unintegrated with data from operational department: the former could launch a marketing campaign without being aware of the total lead time or machinery capacity, it may transform results in stock outs or in excessive lead time.

In general, there are two possible approaches to face data integration challenges:

The first is the *loose mediator-based integration*, in which data are kept at their original sources. A data management system acquires and combines data relevant to answering users' questions. Although this approach avoids data replication and leverages the storage and maintenance of the source data by its respective owners, it may be necessarily to develop resolution strategies to cope with conflicting conclusions drawn from different data sources.

The second method is the *tight integration approach* that extracts, collects, and then replicates all relevant data from individual sources into one large integrated big data repository (a centralize database). This process involves data extraction, cleansing, transformation, and loading. Then analytics tasks this time will be directly performed on the dedicated data repository. All the pre-processing is done in advance and offline, which can shorten the time needed for analytics and decision making. However, it must cope with issues of up-to-date replication and data consistency.

Data Security

The growing cybersecurity threats imply that organizations should be prepared for data breach, and design and implement measures to detect such breaches in advance to avoid and minimize their negative impacts. Organizations are expected to comply with data security regulations and report to both the regulator and individual consumers affected when a data breach takes place. As data volume and computing infrastructures get large, traditional methods of data security mechanisms are becoming inadequate, as a result, effective security mechanisms are needed at different layers to provide data security. An onion model (Kune, Konugurthi, Agarwal, Chillarige, & and Buyya, 2016) of defence for big data security may be the right choice. The computing infrastructure layer, as an onion will be slit in several layer, in this way, big data setup will be confined to an enterprise or several enterprises. These layers can be described in:

- at the *distributed data level*, companies need to develop a privacy preserving mechanisms, data encryption techniques, data access control mechanisms, and security of data models.
- at the *analytics security level*, companies need to develop secured frameworks that allow them to use analytics securely based on authentication mechanisms.
- at *the user level,* confidentiality, integrity, and user authentication mechanisms must be established to validate users; from a mobile point of view, this layer has become very central, smartphones are a needful accessory, that internalise a pack of sensors that could

truly compromise user's privacy, for these reasons biometric unlock and two factor authentication solutions are essential.

Human Talent

For companies operating outside of the centres of tech gravity filling machine-learning or analytics posts can be a challenge. In the U.S., 10% of the pool of advanced analytics talent sits in greater metropolitan New York City, and 14% is in the San Francisco Bay area (Brahm, 2019). In Europe, just three countries are home to half of all the EU's AI talent: 24% can be found in UK, 14% in Germany and 12% in France (LinkedIn, 2019). The US employ twice as many AI skilled individuals than EU, despite its total labour force being just half the size. Figure 9 describes AI talents distribution in Private and Academia sectors from a sample of six European countries. Academia and research centres dominate, especially in those country where AI intensity is the lowest (Italy and Spain), while the northern country, especially France, AI talents are mainly employed in the private sector. In the UK case, the 35% of AI talents works for global technology companies.

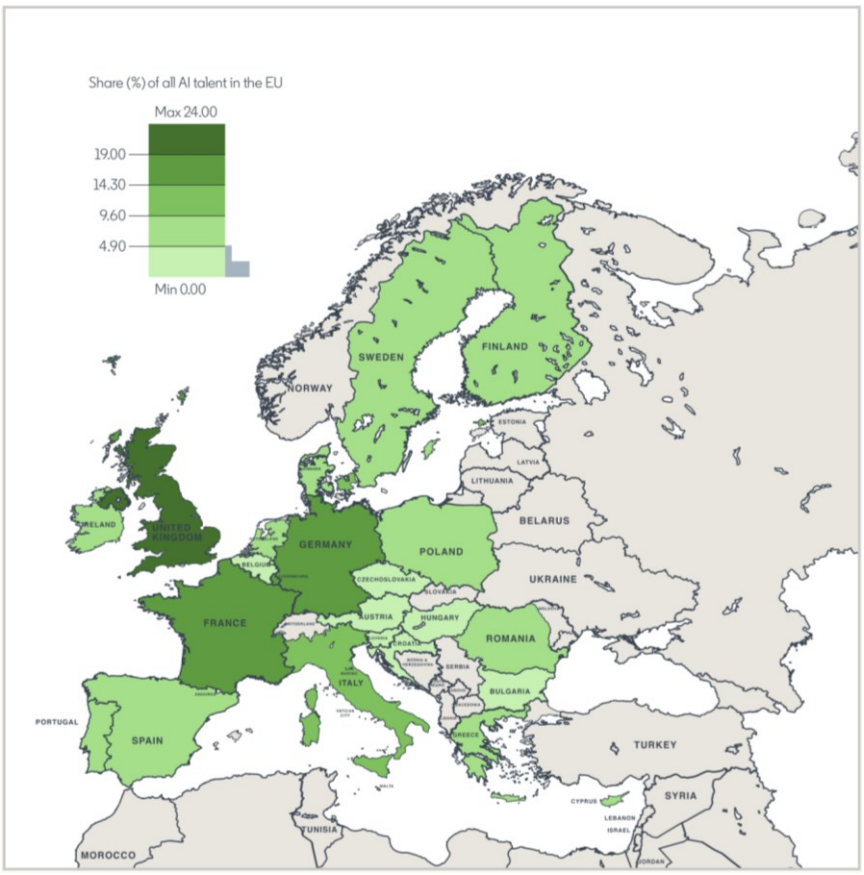


Figure 9 - AI talent distribution in the European Union (LinkedIn, 2019)

A similar scenario is seen in India, where advanced analytics talents are situated respectively in Bangalore and Delhi while in China the most power hubs are situated in Shanghai and Beijing. Since the number of people trained to work with data are in short supply, especially the one involved in machine learning and AI, the demand for big data expertise is growing every day, as more and more companies want to exploit the power of data. BDA is a strategy and operational activity; an organization needs to build a team of people with sufficient BDA skills and talent to capitalize on the promise of big data.

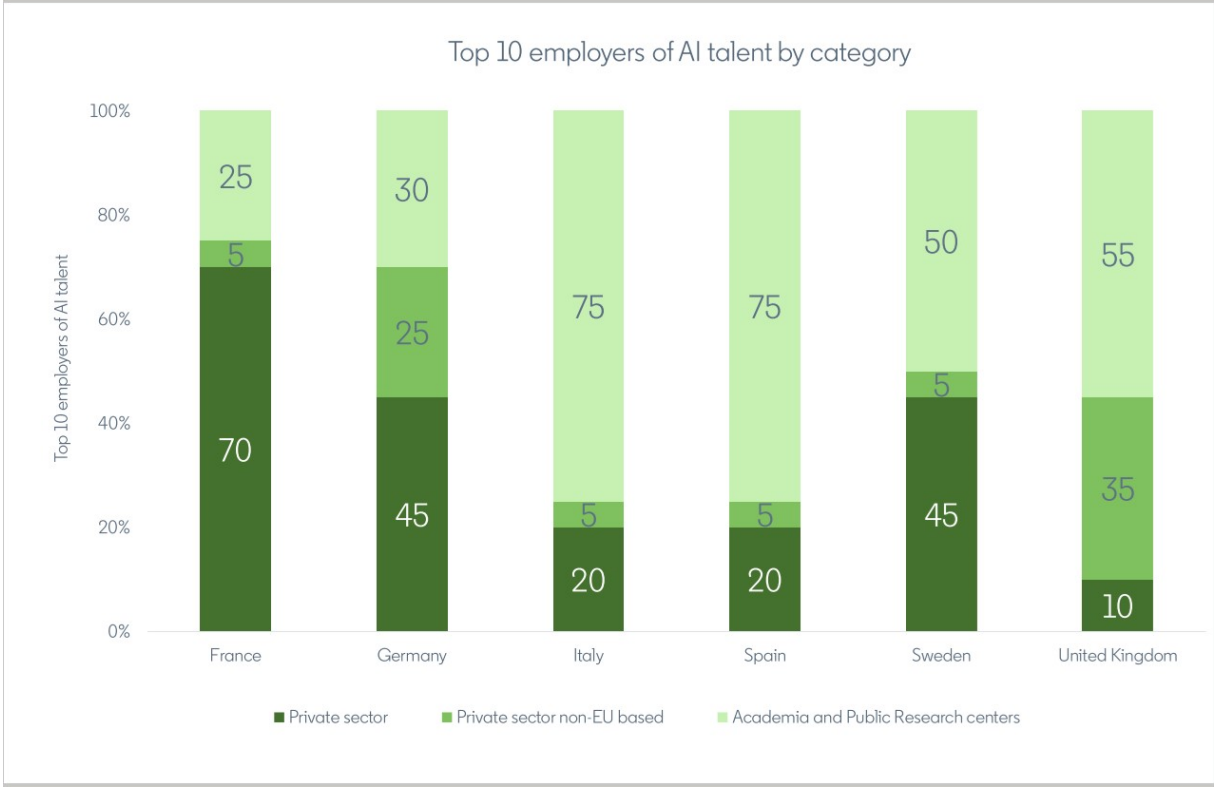


Figure 10 - Top 10 employers of AI talent by category (LinkedIn, 2019)

In 2015, IBM, Cisco, and Oracle together had 26,488 open positions that required big data expertise. In Bain & Company's survey, 56 percent of executives said that their companies lacked the capabilities to develop deep, data driven insights (Wegener, 2013). At Walmart, when a new member joins the analytics team, he or she has to take part in the analytics rotation program, in which he or she will spend time with different departments to understand how BDA is being leveraged across the company.

What actions can an organization outside high developed centres take to acquire BDA talents?

- *Create dedicated outposts within Tech Hubs*: talents are rare and expensive in the hypercompetitive Bay Area, a smart choice is to create multiple centres around the world linked with the main tech hub; these teams will work alongside and grow under

the same company, the aim is to acquire talent from different part of the world where they may be less expensive; moreover it could be a high opportunity to create a talent path where skilled people can experience an international program, enabling the diversity management potentiality; .

- *Invest in Academic Centres:* with companies hiring thousands of data professionals, more and more universities are expanding their course offers with artificial intelligence and data-science programs; cities like Toronto, Montreal, Atlanta, Pittsburgh, and London have become important centres for Google, Facebook, Amazon, and Uber, based on the strengths of their university programs and their linkage with universities themselves.
- *Expand Reach with a Hybrid Approach*: the idea is to mix internal capability with external partners that matches the breadth of advanced analytics expertise they'll need in the future; in this structure, a core, in-house analytics team focuses on developing critical strategic tasks (around 30%), such as data-science team and model development while third-party service firms handle less-critical advanced analytics work like tactical data management and model maintenance (70%).
- *Support Internal Training Opportunities:* the idea is to train in-house talents by offering retraining and continuous learning to help them develop new analytical skills; as BDA is a young field, advanced analytics often lacks experienced manager, by helping existing employees develop these new skills, companies tap into a group that already knows the company, the industry, and how to operate effectively across the organization.

Analytics

LaValle (LaValle, Lesser, Shockley, Hopkins, & and Kruschwitz, 2011) identified three levels of analytics capability of firm: aspirational, experienced, and transformed.

Organizations with *aspirational analytics capability* are new on the field, they focus on efficiency or automation of existing processes and in findings ways to cut costs, their analytics capability are low and are the latest in achieving their analytical goals. They lack some essential IT assets for collecting, understanding, or acting on analytics insights.

Organizations with *experienced analytics capability* have gained some analytics experience, they are looking for developing effective methods to collect, incorporate, and act on analytics for business optimization and often intend to go beyond cost savings.

Organizations with *transformed analytics capability* can be considered veteran, they gained a professional experience with BDA and with organizing people, processes, and analytics tools. Transformed organizations are three times more likely to outperform their competitors s than aspirational organizations, indeed, they mostly focus on driving profitability and making targeted investments in niche analytics.

Value Assessment

A Gartner survey reported that when IT and business leaders were asked about ROI related to BDA investments, a large proportion of companies (43 percent of those planning to invest and 38 percent of those that have already invested) did not know if their ROI would be positive or negative (Heudecker, 2016). Not being aware of IT-Profitability relation can be considered deathly since IT investment are riskier compared to other non-IT investments due to fast obsolescence and the change from an organizational point of view. Moreover, without setting a goal in term in performance will make difficult budgeting process, mostly while setting data strategy.

Are Small and Medium Enterprises aware of data value creation? The centre Italy case

As we can imagine, BDA solutions and data-driven culture is mainly found among Big Tech companies, such as Amazon, Microsoft, Apple, and Facebook. In Italy, BDA solutions are mainly applied to Banking and Insurance industries, where those technologies are used to forecast the rate of succeed of customers to repay a loan or the percentage of accidents in a determined area and under certain conditions. Not by chance, those two major industries, in the latest years, are protagonists of many merging activities, where banks are acquiring insurance products and vice versa.

Regarding the other industries, seems that they have not developed a data-driven culture yet. During a call with a Data strategy specialist, who works in one of the largest technology consulting companies of Italy, she pointed out how much difficult it is dealing with customers in offering data solutions in sectors not related with banking and insurances. Companies feel sceptics about the potentiality of those solutions and not interested at all. Moreover, companies seem to adopt data-solutions not to exploit their potential on the market, instead their main purpose for adopting a BDA solution are cost savings.

Despite those initial clutches in spreading a data-driven culture in Italian Big companies, Covid-19 emergency has been creating an open door in favour of those technologies, where digital transformation has become an important source to survive.

An important element that must be taken in consideration is that, According to Prometeia's latest estimates, in 2017 there were about 5.3 million SMEs that employed over 15 million people and generated a total turnover of 2,000 billion euro (Infodata, 2019). According to Cerved, Italian SMEs have produced a turnover of 886 billion euros and an added value of 212 billion, equal to 12,6% of Italian GDP (Cerved, 2018). SME employs 82% of workforce in Italy (well above the EU average) and represent 92% of active enterprises, sign of a distinctive sign of Italian economy (Infodata, 2019). Their activities are concentrated in the services, construction, and agriculture sectors (72% of SME employees in Italy). Furthermore, for the southern regions, for example, SMEs represent 83% of production, compared to an average national contribution of 57%. New SME trend show that new entrepreneurs are on average younger (38.4% are under the age of 35) and more educated. An important contribution is offered by foreigners, which are equal to 10%, and by women: in 2014, among new entrepreneurs, 28.4% were female (up to reach 31.1% among self-employed workers). Women

entrepreneur, unlike their male colleagues, have opened their own businesses in sectors with a high technological and knowledge value: 48.4% against 39% registered among men.

The territorial differences seem not to be significant exception for two aspects: The South and the Islands have the highest percentage of young people, 41% under 35 against 38.4% on the national average, and the lowest of foreign entrepreneurs, 5% against 10% (ISTAT M. |., 2019).

The state of technological innovation and digital challenges

According to *permanent census of companies 2019* (ISTAT, 2020), in the period between 2016-2018, 77,5% of companies, with more than 10 workers, invested, or utilized, in one of the 11 technologies considered as key component for the digitalization process. Most companies utilize a small range of technologies, prioritizing IT infrastructures (cloud solutions, optic fiber, ERP software). Only at an advance digital degree level, IT investments become more structured and integrated among them: more than 90 thousand over 210 thousand companies studied, utilize 3 of the 11 key technologies.

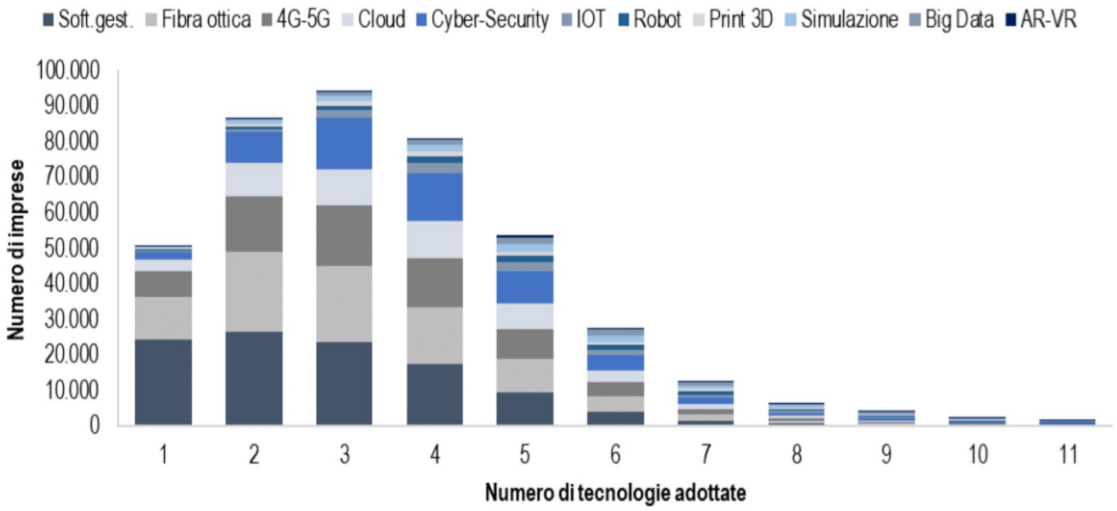


Figure 11 - Digital technologies adopted by SME with more than 10 employees within 2016-2018 (Istat, 2020)

The adoption of IT infrastructures already reaches saturation among less digitized companies compared to the adoption of complex applications with greater impact on business process, which spread slower. Indeed, are 16,6% the companies that adopted at least one complex

technologies among IoT, augmented reality, BDA analysis, advanced automation, and 3D printing; this can be considered as an ongoing transition with a high grow potential.

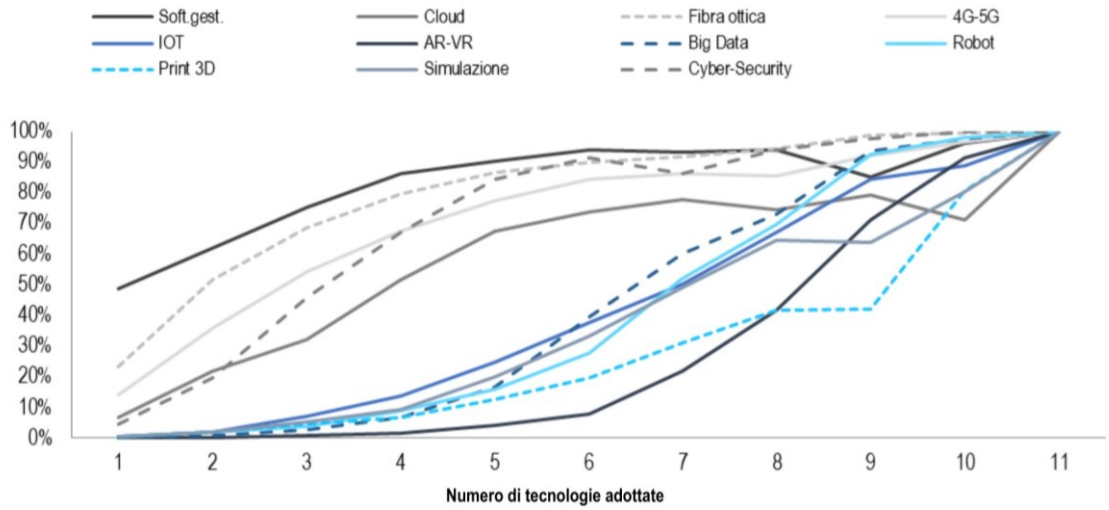


Figure 12 - SME with more than 10 employees that adopted digital technologies within 2016-2018 (Istat, 2020)

Taking in consideration companies' dimension, digital investment gap is present but not as high as it could seem: 73,2% of companies with 10-19 employees against 97,1% of companies with more than 500 employees invested in IT.
Surprisingly, regional differences aren't relevant, indeed, 73,3% of south-Italy companies invested in IT compared to 79,6% of North-Est companies.

The digital lead is mostly represented by service companies with an adoption rate higher than 90%: the telecommunication sector, R&D, IT, auxiliary financial activities, publishing and insurance. If we consider the manufacturing sector, the first position is represented by pharmaceuticals with 94,1% followed by the chemical sector with 86,6%.
For expected IT investments between 2019 and 2021 there will be a trend toward security (+33,5%) and web connection (+13,1%); moreover, for those companies that already have adopted complex technologies (such as IoT and BDA), the trend shows an enlargement of adaptation of those technologies, with a +180,7% for augmented technologies, +117,6% for 3D printing, +111,9% for BDA and 109,9% for robotics.
A digital shifting and IT investments are more present in those companies that are involved in innovation of products and processes: a greater intensity of digitization overlap with innovation.

Companies with at least 10 employees that invest in digital technologies note a positive impact in sharing information and knowledge within the company (52%). Only 32.4% observed greater efficiency of production processes and 13.7% greater ease in acquiring knowledge from the outside while an insignificant percentage of companies perceived negative effects of digitization, at least in its initial phase, on efficiency and productivity.

Some critical aspects emerged observing the personnel digital skills, where the need for training related to the introduction of digital technologies was verified. The use of new management software seemed to be the most critical aspect with 39.6% of all companies with digital investments had to intervene to fill a skills gap in this area. A similar result, with the need of training activity for the introduction of Internet connections (15.9%), IT security (13.1%) and the adoption of cloud technologies (10%), for simulation between interconnected machines (60.5%), robotics (56.2%), 3D printing (51%) and 47% for the analysis of Big Data. Digitization changes staff skills needs and brings out new priorities and potential critical issues.

The macro-skills considered most relevant and the most critical identified by the companies were, at the first position, Security together with the staff's ability to communicate and share information and collaborate using digital technologies.

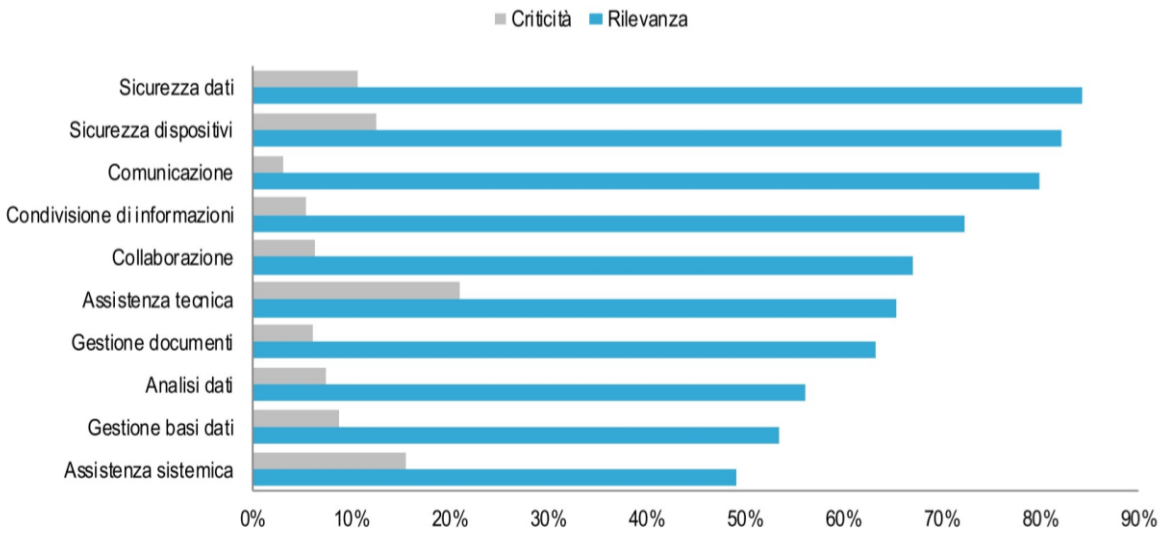


Figure 13 – relevance and criticality of digital skills in 2018 (ISTAT, 2020)

Lastly, companies with 10 employees or more, when asked how to manage any critical skills issues, determined by digital transformation, indicates the following practices: acceleration of

turnover (72.5%), self-training of personnel (57.0%), offer of training to personnel (56.9%), increase in investments in automation (56.3%), use of consultants (48.3%), improvement of the selection processes for new hires (44.3%).

The centre Italy case: the state of digital transformation in the north Abruzzo district

An overview

Abruzzo is the northern region of the south Italy area, with a high influence by the centre Italy regions, especially by Marche district and Rome. With a surface of 10.831,84 km^2, 1.333.939 citizens which are distributed throughout the territory according to a demographic density which is rather low, 123.2 inhabitants per sq. km, 14th place among the Italian regions (Camere, 2013). There are 149,334 registered companies, a number that does not allow the province to go beyond the 14th position in the general ranking for the number of entrepreneurial initiatives. In a context that sees commerce as the predominant sector, the importance of agriculture stands out: 19.1% of the total number of local businesses, a figure slightly higher than the reference area and above the national value, 13%. The incidence of craft activities (22.8%) is quite high, higher than the average value of the southern area (17.8%), and in line with the Italian average (23.2%).

Abruzzo is one of the main manufacturing regions in Europe in terms of industrial intensity (Symbola, 2019), with a production system with a GDP of over 30 billion euros is seventh in Italy for industrial specialization, seventh for the incidence of exports on GDP (8.7 billion of which about 50% linked to the automotive sector), sixth for trade surplus and second in terms of trade value (more than 200 euros are exported per 100 imported), a system that has one of the highest rates of production diversity in Italy (721 categories of activities present in the area out of 800). Among the most prominent specializations, electronic industries are the first in Italy, L'Aquila is becoming the centre of innovation and research, starting from 5G testing and projects toward Smart city's principles, not forgetting Gran Sasso laboratories, centre of innovation in physics field. Automotive sector is another Abruzzo's main industry, with more of 30.000 personnel specialized in cars and bikes components, as well as Pharmaceutical sector, with around 1200 personnel with some big national and international companies linked with local universities. Others important sectors are: Food & Beverage, famous for its Doc wines and local dishes, as well as Design and research centre affiliate to universities.

Like other regions of south Italy, the highest risk is the high percentage of young graduates who emigrated to find employment and from the difficulties of the Region to keep their talents within the regional borders, a risk that could further slow the innovation and digital transformation process.

Survey BDA: Are companies aware of data-driven culture?

Between early December 2020 and January 2021, there have been collected (using an online survey tool), among the central Italy regions (with 77% of replies are from Teramo province), companies' managers opinions about their awareness and vision toward adopting a data-driven culture. Despite the initial goal was to collect around 100 replies from the 300 Small and Medium Enterprises contacted, only 52 replied to the survey (17,3%): all companies contacted were comparable with the final sample, for dimension (with at least 10 employees, with a minimum of IT structure) and sector (mostly manufacturing activities and food & beverage). According to the ISTAT census (2018) about Abruzzo companies, there are 3947 companies between 10 to 249 employees, which it is the range took in consideration in this research: not too small to invest in Big Data & analytics solutions, not too big to be considered as big company (>250 employees). If we only consider Teramo province (figure 14), the total of companies is 23984: 95 % small enterprises (less than 10 employees), 4,4% small to medium (10-49 employees), 0,4% medium enterprises (50-249 employees) and only the 0,004 big companies. Comparing our, relevant and significant, sample of 32 companies with more than 10 employees, from Teramo province, the final sample cover the 2,1% of small to medium enterprises, 7,1% of medium enterprises and the 36% of big enterprises.

Location	Teramo									
Period	2018									
Data type	number of active companies					number of employees of active companies (average annual				
Number of employees	0-9	10-49	50-249	250 +	total	0-9	10-49	50-249	250 +	total
Ateco 2007										
0010: TOTAL	22818	1057	98	11	23984	42691,48	18595,15	9839,38	5096,41	76222,42
B: extraction of minerals from quarries and mines	16	1	1	..	18	45,73	15,62	174,69	..	236,04
C: manufacturing activities	2285	408	53	9	2755	5853,59	7876,33	5148,54	4259,61	23138,07
D: supply of electricity, gas, steam and air conditioning	106	2	..	..	108	58,55	28,11	..	..	86,66
E: water supply, sewerage, waste management and remediation activities	52	19	6	..	77	136,43	339,79	813,38	..	1289,6
F: buildings	2862	139	7	..	3008	5523,14	2392,96	847,93	..	8764,03
G: wholesale and retail trade, repair of motor vehicles and motorcycles	5651	159	10	1	5821	10203,37	2669,78	767,89	533,93	14174,97
H: transport and storage	433	48	4	..	485	1005,58	768,54	407,03	..	2181,15
I: activities of accommodation and catering services	2038	156	3	..	2197	6326,68	2185,25	341,26	..	8853,19
J: information and communication services	442	12	1	..	455	810,9	190,35	57,15	..	1058,4
K: financial and insurance activities	409	8	1	..	418	630,51	142,79	147,05	..	920,35
L: real estate activities	894	5	..	..	899	1036,36	79,86	..	..	1116,22
M: professional, scientific and technical activities	3733	15	..	..	3748	4665,33	206,19	..	..	4871,52
N: rental, travel agencies, business support services	694	30	7	..	731	1202,38	667,34	661,54	..	2531,26
P: education	102	3	..	..	105	155,23	43,81	..	..	199,04
Q: health and social assistance	1273	16	4	1	1294	1811,72	276,69	342,72	302,87	2734
R: artistic, sports, entertainment and fun activities	435	14	..	..	449	816,55	259,12	..	..	1075,67
S: other service activities	1393	22	1	..	1416	2409,43	452,62	130,2	..	2992,25

Figure 14 - ATECO 2007 - Company census of Teramo province (ISTAT, 2018)

The survey was built with the use of close questions, open questions, and Likert scales: the first part of the survey has the aim to filter company by size, region and by role of the interviewed subject since the goal was to collect information from top managers who have been working in

the company for long time, that knows the culture and the vision. The second part of the survey is dedicated to discovering the state of IT infrastructure of the company, if they have got the right personnel to perform Analytics task and in which way, they would like to compensate in case they lack analytical skills. Lastly, there were asked about their strategic positioning according to the BDA matrix.

Question N.1: what is your company size?

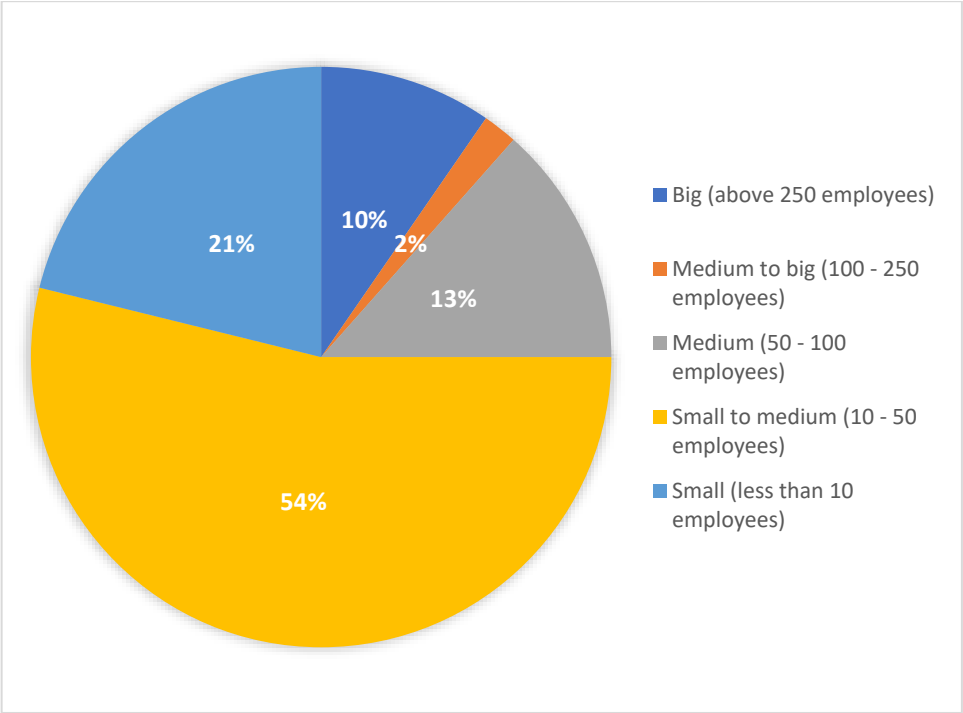


Figure 15 - BDA survey, question N.1

The reference sample is composed for 54% of small to medium enterprises (with 10 to 50 employees), 21% of companies with less than 10 employees. 13% of medium enterprises (50 to 100 employees), 2% by medium to big companies (100 to 250 employees) and 10% by big companies (above 250 employees). In this sense, the survey goal was reached, indeed, 88% of the sample belongs to SME category, a range of company that should be attracted by BDA technology capabilities but struggle to implement them in their company due to lack of infrastructure or personnel skillset.

Question N.2: Your industrial sector?

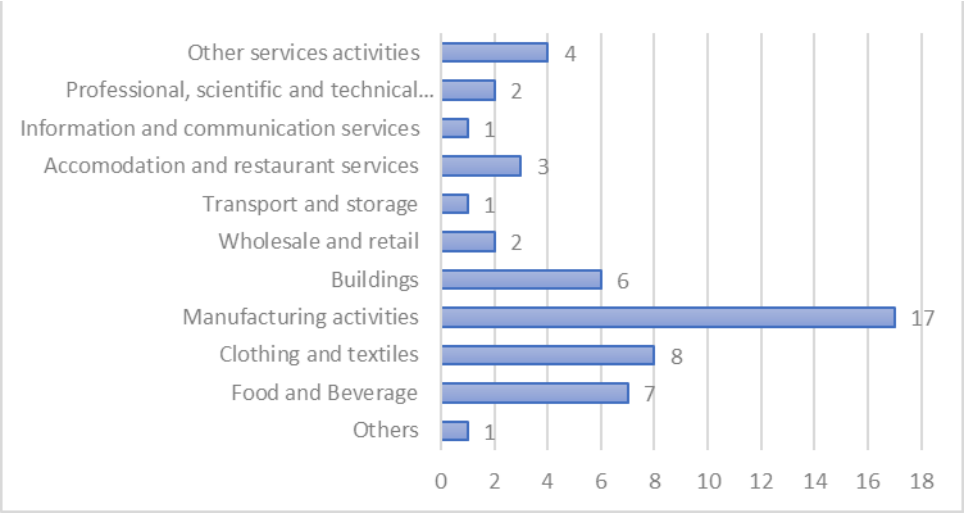


Figure 16 - BDA survey, question N.2 (absolute values)

The industrial sector of our sample is composed by 33% of manufacturing activities, followed by clothing and textiles (15%), food & beverage (13%), and buildings sector (12%). The result is in line with Abruzzo and Marche region SME activities: many product producer, lower service providers.

Question N.3: Province / region in which the company operates?

In the reference sample, the 89% of companies operates in Abruzzo region, 40 over 52 from Teramo province, only 11% from Marche region. In Teramo province, 53% companies belong to small to medium enterprises (10-50 employees), concentrated in manufacturing and building sectors. The Teramo area will be the focus for further insights in the appropriate section.

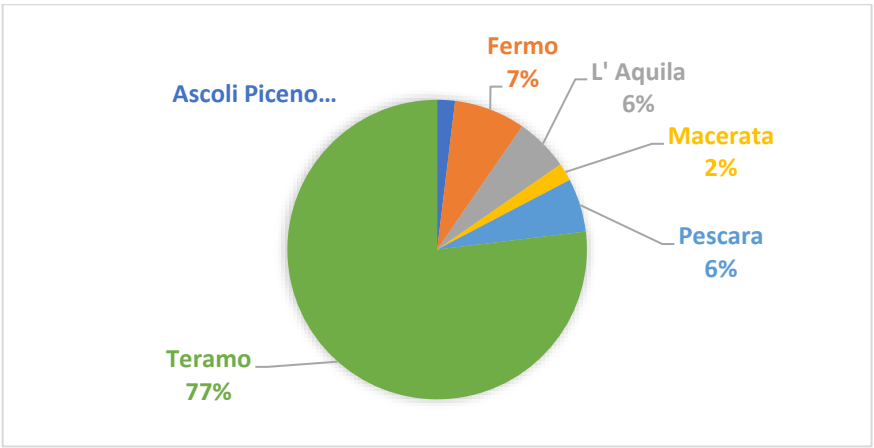


Figure 16 - BDA survey, question N.3

Question N.4 and N5: your role in the company? How long have you been working in the company?

The aim of question N.4 and N.5 was to filter those replies provided by roles different from decision making area. 69% of the sample were entrepreneurs (owners) and company CEO themselves, 21% belongs to the managerial area and 10% were specialist. Regarding question 5, 94% of the sample should be aware of the company vision and culture, being them inside the company for long time or founders of the company.

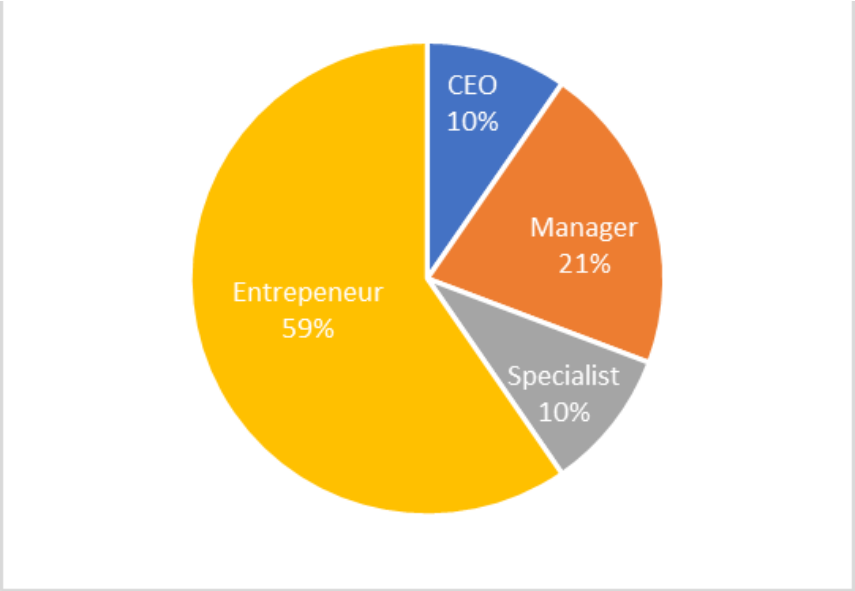


Figure 18 - BDA survey, question N.4

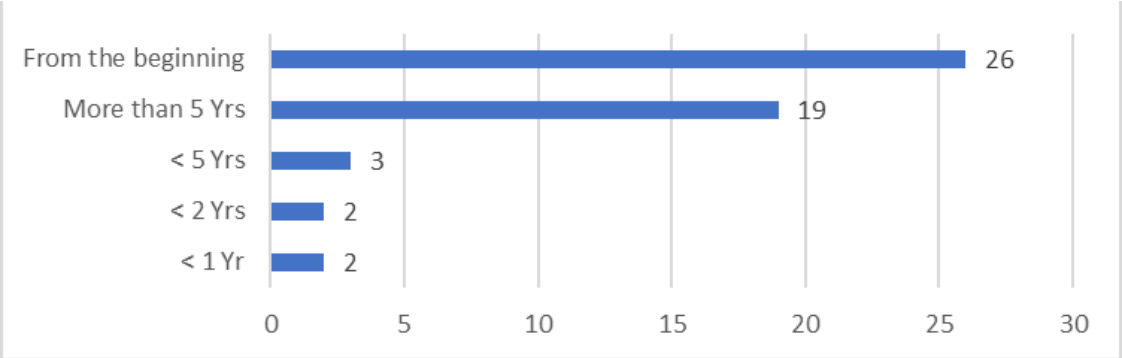


Figure 19 - BDA survey, question N.5 (absolute values)

Question N.6: do you think that moving towards digitization through a data-driven culture, which gives value to data collected, could be the winning card for obtaining a competitive advantage?

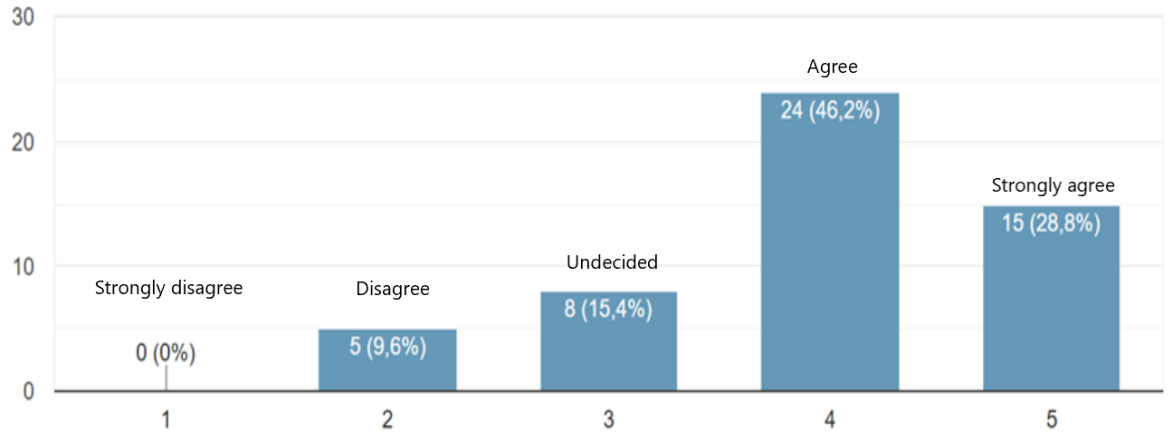


Figure 20 - BDA survey, question N.6

Question N.6 is the turning point of the research. After explaining them what a data-culture is, by inserting its definition and related examples in the initial part of the survey (before data-culture related questions) interviewers were asked about rating the importance of shifting their current mindset in a data-driven one (from a value from 1 to 5) to foster a competitive advantage in the future. 46,2% of the companies gave a mark of 4 out of 5 and 28,8% of 5. The average score is 3,94.

By filtering overall score by company size, the result shows:

- Small companies scored on average 3,81.
- Small to medium companies scored on average 3,78.
- Medium companies scored on average 4,42.
- Medium to big companies scored on average 4.
- The few big companies in the sample scored on average 4,4.

Despite the small sample size, results shows that the larger the company size, the higher their awareness around data-driven culture. Medium companies seem to be attracted by this data trend, nonetheless, as it will be shown in the next questions, they lack tools and personnel to exploit their data potential.

Question N.7: In which area do you think that the study and enhancement of company data can be useful / winning for your company?

Question 7 has the aim to address relevance to BDA field of application, in this specific case, companies were asked to choose two areas of interest that might help their companies to gain a competitive advantage. The available fields of application were:

- Financial /administrative management: I want to be able to obtain an automatic billing and / or accounting system with data updated in real time (Roi, cash flow, recurring payments).

- Marketing and sales: I want to get information about how my competitors are doing and /or monitor the progress of my marketing campaigns in real time and /or look for new opportunities for the launch of a new product and / or service.

- HR (Human Resources): I want to be able to monitor employees' performances and/or identify possible talents; better management and optimization of shifts and workspaces.

- R&D: I want to develop new products/services through analysing data from company machineries, my current products and/or from the market and/or from customer comments on social networks.

- Operations: I would like to predict the flows of demand and adjust my offer accordingly in real time, avoid running out of stocks, manage orders automatically, optimize inventory.

Companies chose to be mostly interested in Marketing and Operations field, right ahead Administrative and financial management. Small companies are mostly interested in marketing and operations (26% and 32% respectively) while among medium enterprises, 35% of them prefer solution that can improve their marketing and sales activities, followed by operations and administrative and financial management (both 22% of medium enterprises preferences). Regarding preferences filtered by industry category, Food & Beverage companies prefer, with the 55%, data-driven solution that can improve their marketing and sales activities, followed

by operations while the other industries do not show a significant difference among proposed solutions.

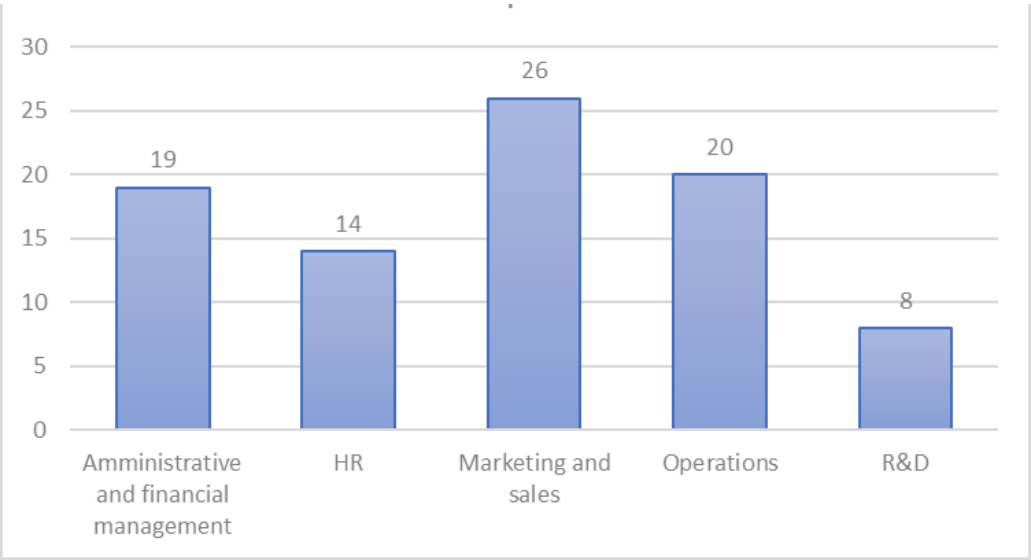


Figure 21 - BDA survey, question 7 (absolute values)

Question N.8: How is the company's IT infrastructure (information Technologies) structured?

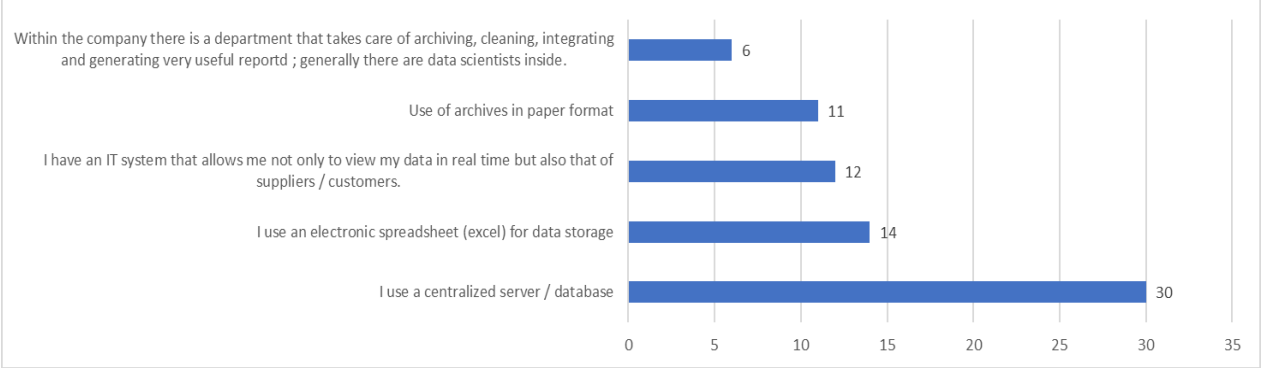


Figure 22 - BDA survey, question 8 (absolute values)

Question N.8 had the aim to investigate IT companies' infrastructure. Most of companies adopted a centralised databased (58%) and the few companies that have adopted a specific department for data analytics are the big ones (only 5 of the total sample). Around 27% of the companies utilize a mixed structure with both central databases and electronic spreadsheet or in paper format. Paper format archaization is mostly prevalent in small enterprises, around 50% of the cases, but more than half of them implemented a centralised database.

Question N.9 & 10: Do you think, in your company, there are people capable of enhancing corporate data? If you feel that you do not have adequate / sufficient human resources, how do you plan to act?

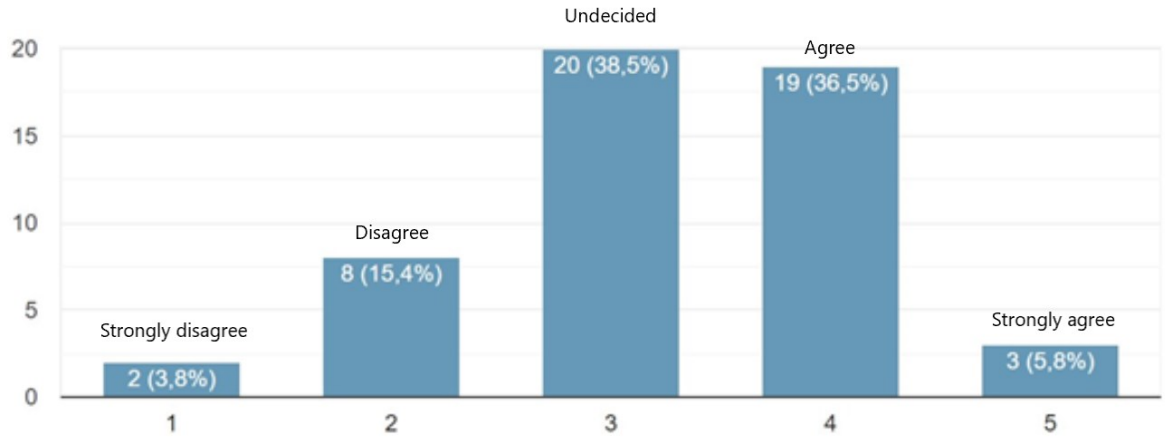


Figure 23 - BDA survey, question 9

Question N.9 investigate the opinions of top management roles about their employee's ability to enhance corporate data to create value creation. Results showed a score, on average (on a scale from 1 to 5) of 3,94, in absolute value the most frequent score were 3 (38,5%) and 4 (36%). There are not high differences among company dimensions: score of 4 for small enterprises and a score of 3,93 for medium enterprises. Apparently top managements are aware that their personnel are not already qualified to adopt a data strategy approach to create a significant value creation.

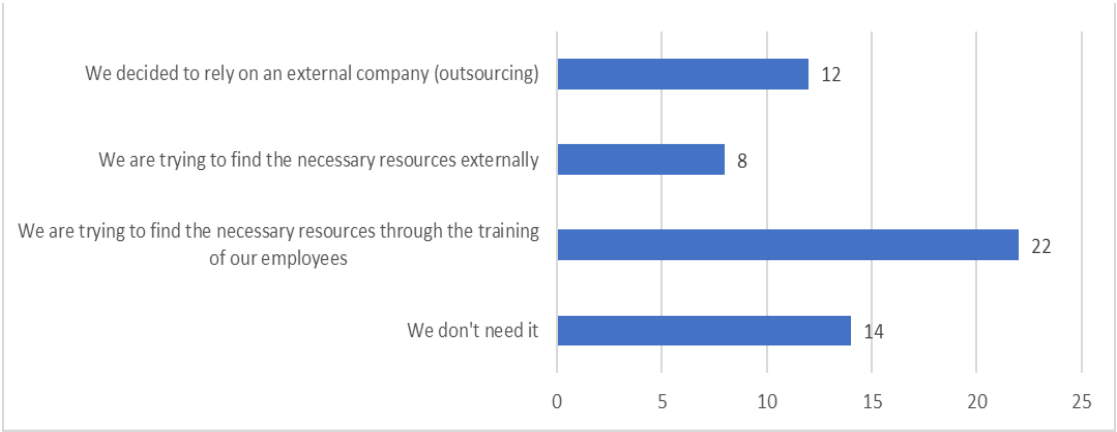


Figure 24 - BDA survey question 10 (absolute values)

Indeed, question 10 results showed that, companies mostly prefer to access to digital and analytical skills by training their employees (42%), followed by those which do not need any or are not searching for digital skills (mainly small enterprises), while it is interesting how, 23% of companies decided to rely on external companies' capabilities.

Question N.11: who has access to company data? - Data Governance

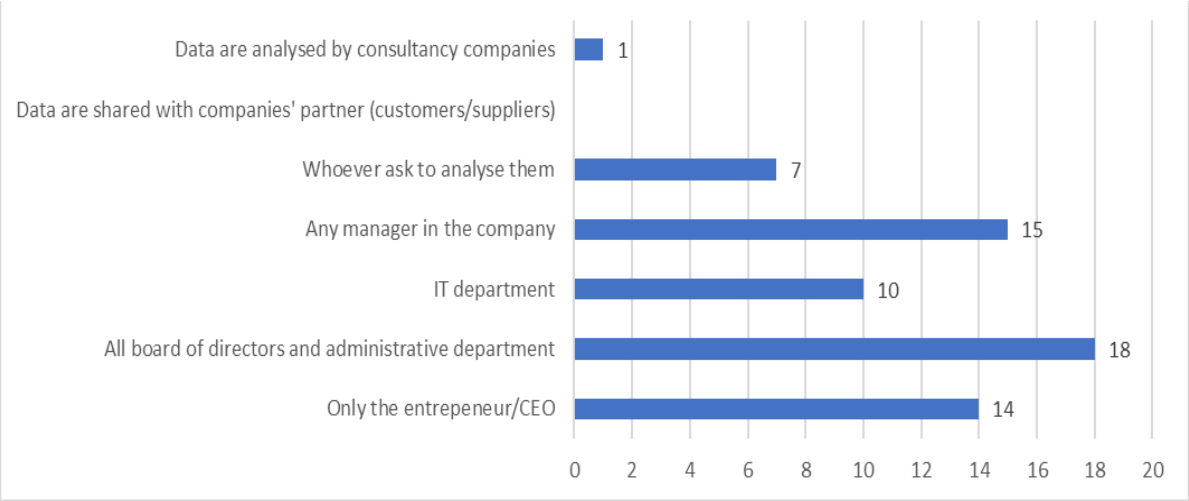


Figure 25 - BDA survey, question 11 (absolute values)

Question N.8 had the scope of investigating communication problems within companies, in the reference sample it is possible to notice that data are shared among top management, IT and administrative department. Contrary to question N.10, where 12 companies replied that they would like to rely on external data consultancy, only one company replied that data are accessible by consultant companies. Moreover, data are not shared with companies' partners, like suppliers, pointing out that there is a lack of connection, in real time, with the whole production chain: it might be impossible to use a pull approach, where the production follows market demand and raw materials are supplied as fast as the company needed it. In the sample case, companies might use a push approach, where the companies must forecast their production according to their own conjectures. It is not surprising since to adopt a pull approach a company needs BDA technologies, both suppliers and the company itself.

Question N.12: how complete is the data you use to make decisions?

Question number 12 had the aim to investigate the level of top management awareness toward their data during the decision-making process. The 62% of top managers stated that they believe to have the key data to make decisions but are not sure to fully exploit them. This result shows

a good base to work with companies' data, sign of a possible improvement of data exploitation by adding to an internal/external analytical figure that could lead to the full exploitation of company's data.

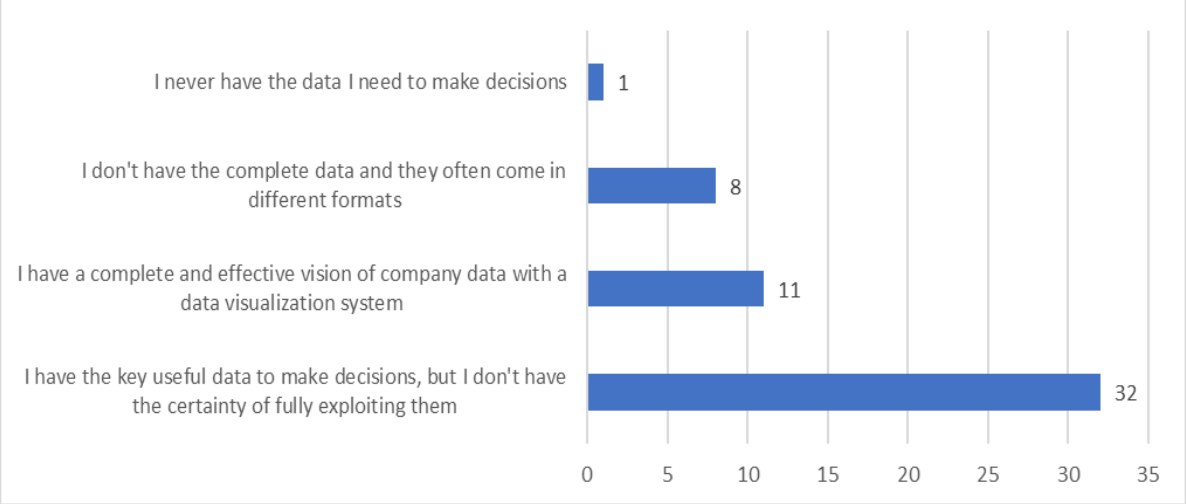


Figure 26 - BDA survey, question 12 (absolute values)

Question N.13: do you use company's data to perform forecasting activities?

When company were asked about forecasting activities, the 40% of them make periodic forecast on the main KPI, the 37% would like to start forecasting activities but they lack technology or skilled personnel to perform this activity, the 13% have a real-time view on the main economic and financial indicators while the 10% have never thought about forecasting. Most of the medium enterprises are interested or are performing forecasting activities, 45% of them regularly forecast the main KPI while 40% of them would like to adopt forecasting technologies but do not have technological or skills to perform them, only the 15% have never thought about forecasting activities. Among small enterprises, only the 27% perform forecasting activities.

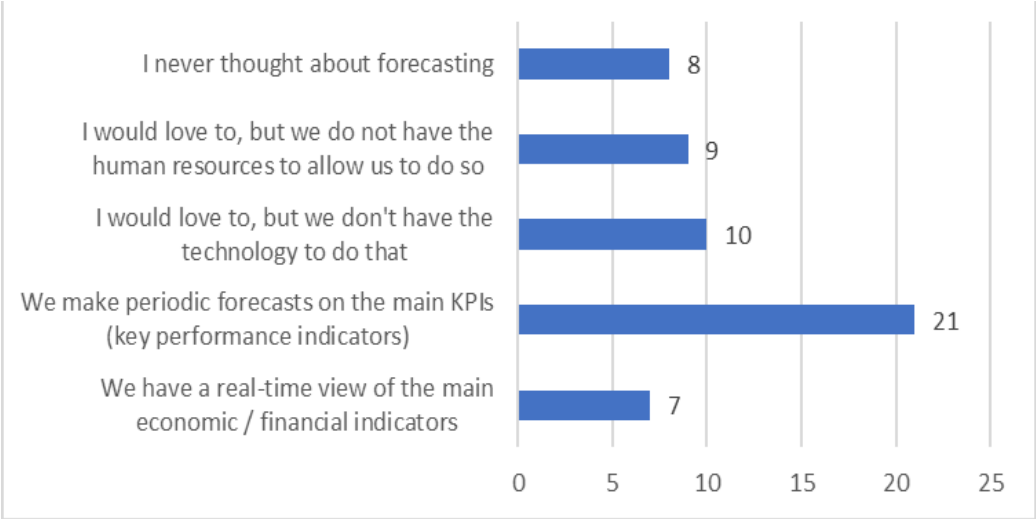


Figure 27 - BDA survey, question 16 (in absolute values)

Question N.14: how is your business data integrated?

One of the most hurdles when a company tries to exploit value creation through data is the lack of integration between data sources, making the entire data analysis and decision-making process confuse and not precise.

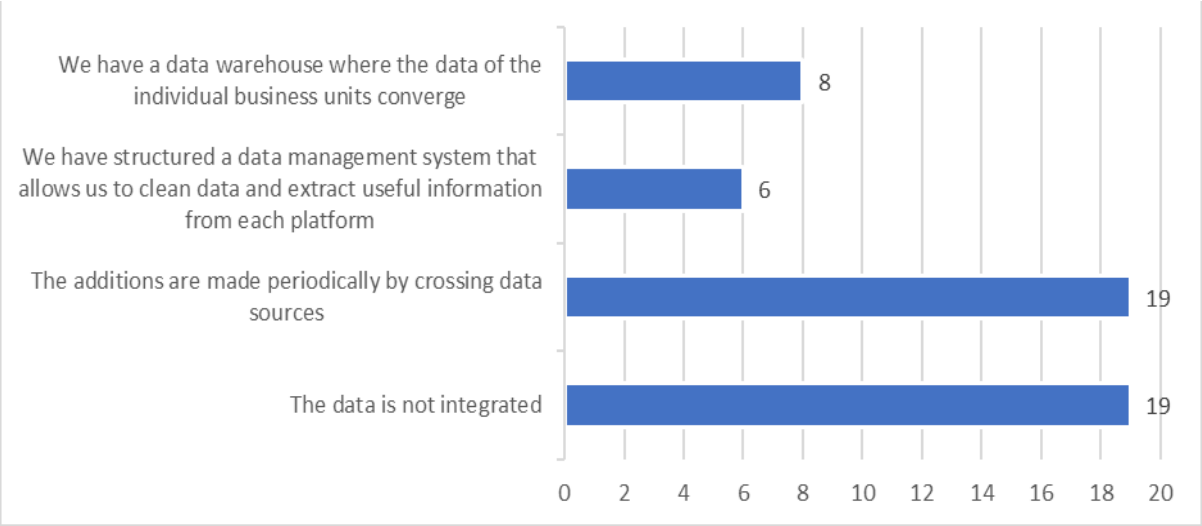


Figure 28 - BDA survey, question 14

In the reference sample, 36.5% of companies integrate their data periodically, on the other hand, there are companies (36,5%) that do not integrate their data at all. The few big companies in the sample integrate completely their data, with some dedicated activity to clean and extract useful information. Regarding medium enterprises, 34% of them do not integrate their data while 43% integrate periodically their sources of data, the remaining 13% collect data of single business units but do not integrate them in a way to have a complete vision of the company.

Question N.15: what technologies are present in your company?

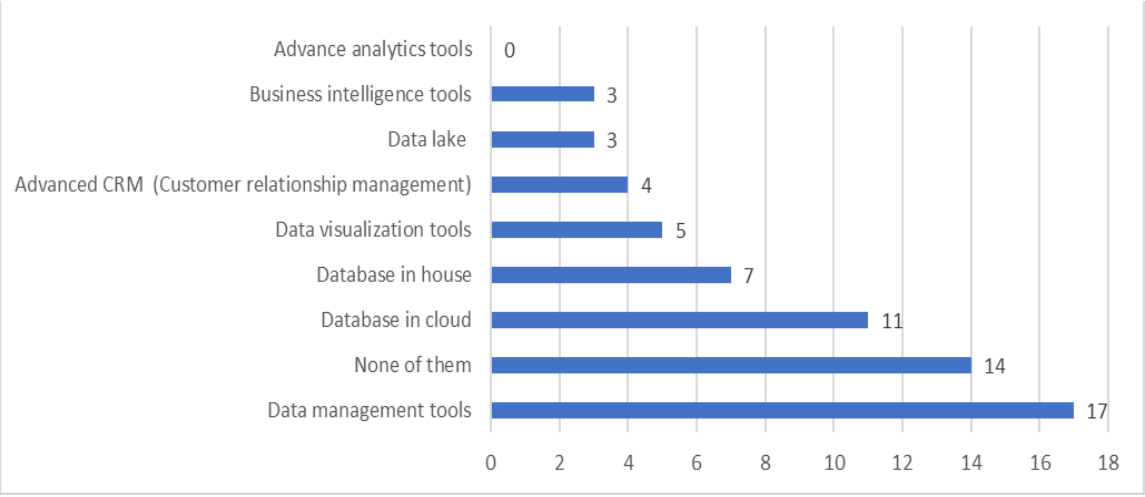


Figure 29 - BDA survey, question 15 (absolute values)

Question N.15 reveals the real state of the reference sample, indeed the 27% do not use any advance BDA technologies tool, the remaining 73%, the 45% use a data management tool and 29% use Cloud based solutions, the 13% use data visualization tools and only the 8% business intelligence tools. Data visualization tool users are only 11%. By filtering data according to companies' dimension, only the 6% of medium enterprises use advanced CRM tools and only 14% use a data visualization tool, while the 31% use data management tool. No companies use advance analytics tools, such as machine learning, artificial intelligence, and scenario analysis technologies, technologies that now are the greater enhancer of the BDA value creation process.

Question N.16: after investing in an IT structure, in the literature it is possible to identify four profiles of a digital company; which of these do you identify with?

Figure 19 shows the self-assessment distribution of SME companies in the BDA matrix. Most of them identify their company role as performance enhancer (40%), 30% as strategic transformer, 15% as image builder and 15% as reactive defender.

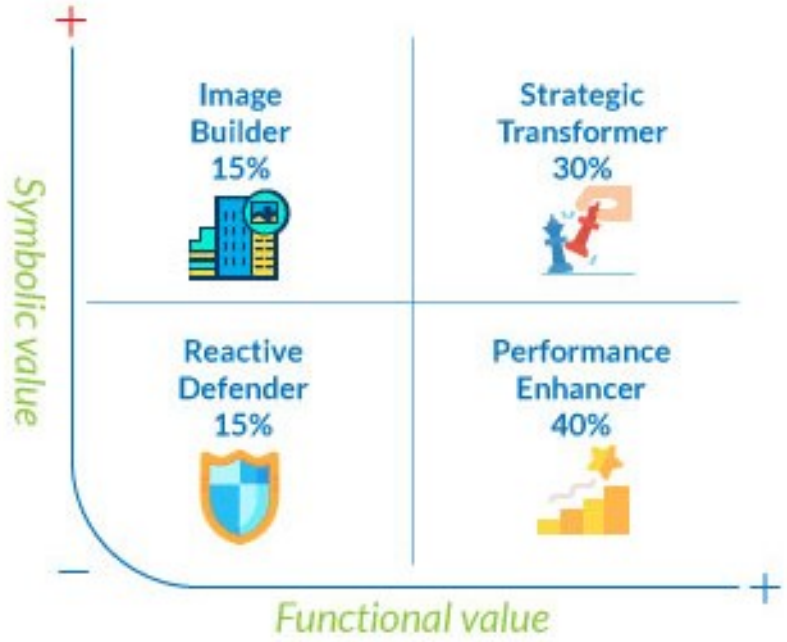


Figure 30 - BDA survey, question 16

Unfortunately for them, becoming a strategic transformer is not as easy as they believe, and without the implementation of analytical tools (none of them have adopted such technology) in their company probably will never obtain that status. It is more credible that most of the company may achieve the role of performance enhancer by improving internal processes thanks to data analytics.

Question N.17: we have reached the end, at this point, do you think you have enough data, and above all of value, to be able to offer an exclusive solution to your customers?

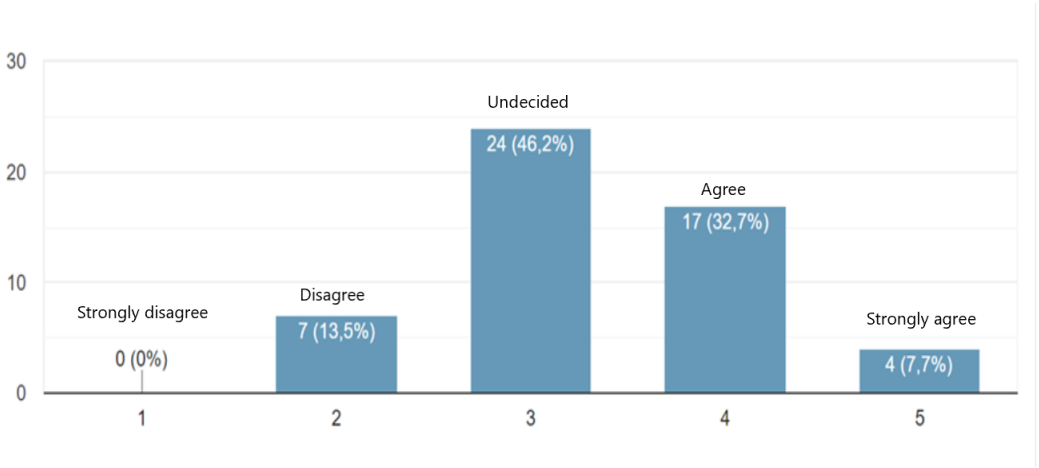


Figure 31 - BDA survey, question 17

The last question had the aim to assess the potentiality by companies to offer an exclusive solution/product to their customers. The overall SME score, on average, is 3,31, with small enterprises that score 3,29 while medium enterprises score 3,32. The most frequent value is 3. Apparently, SME cannot, at the moment, offer a real solution driven by a data-culture, in line with the previous results.

Insights about SME in Teramo province and final considerations

By filtering the sample by province, top management and strictly to significant SME, it is possible to obtain a sample of 32 companies operating in the province of Teramo, an area with a prevalence of SME mainly operating in the tourism and manufacturing industry. The main insights are:

- 56% of companies analysed were classified as small to medium dimension (10 to 50 employees), the 25% small enterprises and the 19% medium enterprises.

- The most prevalent industries in the sample are buildings (19%), clothing and textiles (19%) and metallurgy (13%).

- Companies believe that by adopting a data-driven culture it might lead to a competitive advantage with a score (from 1 to 5) of 3,81.

- SME prefer to obtain BDA solutions that could improve their marketing & sales performances (25%), their operations (27%) and in the administrative and financial field (19%).

- 37% of SME use a centralized database to store their data, 20% a spreadsheet, 20% paper files, 17% have an integrated infrastructure that can exchange their data with partners and 7% have a dedicated department for data management purposes; 28% of them adopted a mixed solution.

- Top managers, on average, scored 3,28 (from 1 to 5) when requested to assess their personnel digital skill to analyse data.

- To access to an analytical skillset, if needed, 25% of the companies prefer to train their personnel while the 17% are searching those roles externally, and the 19% decided to rely on an external company.

- 90% of company do not share company's data, they are only accessible to top management roles.

- 60% of top managers believe they have the right data for decision-making process, but they do not think they are use them at their full potentiality.

- 31% of companies make use of forecasting activities on the main KPI while the 44% cannot access to forecasting activities both for a lack of technology and qualified personnel.

- Most of the companies do not integrate data (44%) while 38% periodically integrate their data sources.

- Only 5% of SME companies have an advanced CRM tool, the 23% a data management tool, 15% use a cloud solution and 23% of them do not use any of BDA listed solutions.

  16% of the companies recognised themselves as image builders, the 41% as performance enhancers, the 13% as reactive defenders and the 31% as strategic transformers.

- 62% of the companies gave a score of 3 (out of 5) to their capacity to provide a unique solution to their customer based to their capacity to extract insights from their data, the overall average is of 3,25.

The bottom line of this research work is that SME are not aware of how to trigger a value creation by the study of their own data. Despite they evaluated the necessity to adopt a data-driven culture as an important factor to foster a competitive advantage in the future, they lack the awareness and the knowledge to act in that direction. During the survey distribution, I had the chance to directly interview few top managers included in the sample, some of them express the need to adopt a data-driven culture as not relevant to renew the product or R&D activities. Data analytics has not to be considered as a stand-alone field, rather it must be considered as a staff function in the organization chart: data analytics generates deep insights to make every activity of the company to perform better. Most of SME lack of technology and personnel adapt to compute a BDA strategy, mostly lacking advance analytical tools fundamental to exploit such strategy. The "reactive defendant" strategy to start thinking and aligning to the companies of the next future should start by restructuring the company organization in order to periodically integrate their data and spread the communication not only among top management roles but even at lower hierarchical levels. External consultant companies, in this unskilful context, might gain a high relevance among SME, BDA awareness and knowledge must be spread among top managerial figures to start together a path that led to digitalize Small & Medium enterprises. Covid-19 emergency is accelerating the timing of innovate company infrastructure and activities; indeed, many government incentives will be given to deal with digital transformation, an occasion that every SME (and not only) must pursue not to disappear in this global market.

Bringing up solutions: bring your data to life.

Most of the time, companies do not realise they are sitting on a treasure; they do not even see it: data. Data has been recorded since the beginning of a company's life when on the balance sheet is inserted the equity values and the relative assets that have been acquired throughout the years, not to speak about orders and customers which whom they interact every day.
Nowadays, it is easy to incur in companies that have already started to track their operations, recording a set of raw data about operations flows but without analysing them in depth. In the meanwhile, during covid-19 pandemic, new category of data is coming up to life at the eyes of analyst: imagine tracking data about how many hours people spent their time video-chatting among them, maybe it would be possible to translate this data source in powerful insight for companies. By identifying, based on daily interactions among employees, who had better be present physically in the office or who may stay at home. This and many ideas can come up studying and analysing data, the main problem is that the major of people are not aware of this potentiality or they struggle in linking data analysis with, for example, product innovation, thinking that those two "categories" travel in two different road, while BDA has to be considered not as a road, but as a light system that helps "drivers" (alias managers/companies) to choose one or more roads consciously, with the possibility to exploit new opportunities and avoid possible threats.

The aim of this chapter is to propose affordable solutions that can be adopted by any company, no matter for their size, with a particular attention to SME, solutions that can bring value to companies' data.
Through the chapter it will be possible to explore sales solutions through the creation of an interactive dashboard created in excel, with forecasting options and easily to update as the company gain more data, an early solution for early adopters' entrepreneurs. The second solution will cover the needs of integrating multiple sources of data via multiple platforms and databases to create a unique and flexible dashboard to easily adapt company's objectives and relevant KPI, moreover, it will be presented a Social listening tool demo, capable to "listen" to the market and intercept new trends and new needs.
The last solution will be relevant for those companies which want to improve their operations, discovering bottlenecks and excessive lead times throughout the value chain process.

Let's get started!

Improve your Sales' awareness through the creation of Dashboards.

How many times does it happens to own a series of lines of text about orders, quantity, prices and at the end they finish unused?
Many times, people think that to create something meaningful with data it necessary to buy expensive programs or that you must be a high skilled figure. The reality is that everyone can manage data in an effective way, at an early level, in which manager of small enterprises can benefit from it. For this "democratic" solution, it is been used Microsoft Excel, a tool that should be present in every PC's desktop.

Theoretical background

In the dashboard are present both statical tools to forecast future revenues and sales indexes useful to *have* a clear vision about "where the company is and where it is going".

Sales Velocity

Sales velocity measures how quickly your business is making money, by determining it, analyses how fast deals are concluded through your Sales pipeline, showing how much revenue a company can expect to make in one day or over a certain period.

Figure 32 – Sales Velocity Index (Bauer, 2020)

Sales velocity can be considered as a litmus paper about the health of a business: if sufficiently high the company can succeed and follow a growing path.

Understanding the rate at which your customers move through your pipeline will help teams pick out which steps in the sales process are working and which ones are causing customers to exit.
To calculate Sales Velocity, a manager needs to track four key factors: number of opportunities in the pipeline, the average deal size, win rate (or conversion rate) and the Pipeline length (or Sales Cycle Length).

1. *Number of opportunities:* it calculates the number of deal opportunity happened in your pipeline in a specific period; here is important to filter bad opportunities from the good one, meaning those that were analysed following the BANT framework (Budget, Authority, Need and Time), in other words, it is important taking in account those opportunities that were worth pursuing; this is crucial, better having a small number of worthy opportunities than high and not worthy, why? Bad deals will provide a low Sales velocity score lead by Win rate.

2. *Average deal size:* it refers to the average dollar (euro) value for a sale, the only factor of sales velocity that directly correlates with companies' profits; as you can imagine, average deal size is a relevant metric for service companies that adopt a subscription pricing model; to calculate average deal size you need to know total revenue from all deals divided by the total number of deals.

3. *Win rate:* it is the percentage of deals that ended in a positive way, where the potential customer become a paying customer; it is a good indicator about the salespeople performances, it is calculated dividing the number of closed won deals and the number of opportunities (interaction).

4. *Average Pipeline length:* it is the average amount of time a potential client takes to become a buying client; it is calculated by dividing the total number of days for all deals over the number of deals; both customers and salesperson hope to close the deal as fast as possible, the lower it is, the better is.

The combination of those four index determinates the Sales Velocity, to have it as big as possible, meaning that the company is making a huge amount of money day by day, we would

like to have a high win rate, that is why we want to take in account the deals that worth, meaning that how salesforce is very capable to sell company's product/services, and the ability.
to close deals as fast as possible, of course, without making pressure on the client.

Forecasting techniques

Why should a company forecast its future? Well, the aim of forecasting is to reduce the range of uncertainty within management judgements must be made. It can be seen as a well-planned vacation; you know exactly where you are going and what could happen along the way, in this way you can react consequentially to provide the best experience possible. Companies that are used to forecast are aware of where they are going and in which way to react according to the market changes. One way to classify forecasting procedures is done by distinguish between long term and short-term forecasting. Long term forecasting is used to set the general course of an organization for the long run, linked to vision and mission, becoming a particular focus of top management. Short term forecasting is used to design immediate strategies used by middle and first-line management to reply to the immediate future (John E. Hanke, 2014).

To perform a forecasting process, a company must pass through those five steps:

1. *Problem formulation and data collection:* the relevant data must be available and correct.

2. *Data manipulation and cleaning:* it is possible to have too much or too little data, some relevant and some not, some may contain errors or blank space, or it could happen that the same data are expressed in different format; it is important to manipulate data to create a standard, easily to update and reliable for the forecasting procedure.

3. *Model building and evaluation:* this step involves fitting the collected data into a forecasting model that can minimize forecasting errors; as a rule of thumbs, it is better to follow Parsimonious approach, between two forecasting models that provides similar results, choose the simplest one.

4. *Model implementation:* it is the implementation of forecasting model; data with recent historical periods are often held back and later used to check the accuracy of the process.

5. *Forecast evaluation:* it involves comparing forecasting values with actual historical value, forecasting errors are analysed; examinations of errors patterns may lead the analyst too modify the forecasting model.

For our purposes, it will be described three simple approaches to forecast a time series: naive, averaging, and smoothing methods.

Naive method is a perfect solution for small datasets, unlike other methods, it does not need a huge quantity of data and its application is easy. Naïve forecasts are one possible solution since they are based solely on the most recent information available. It assumes that recent periods are the best predictor for the future:

Equation 1 – Naïve Method

$$\hat{Y}_{t+1} = Y_t$$

Where $\hat{Y}_{t+1}$ is the forecast made at time t for period t+1. Naïve method is also called the "no change" forecast since what happen today will happen tomorrow. Due to is simplicity, that can be considered his strength, it suffers of random fluctuations that are tracked as faithfully as other fundamental changes. In case data values increase over time, they are said to be nonstationary in level or to have a trend. If we use the standard Naïve equation, forecasts will be consistency low, it is possible to adapt this method to a trend by adding the difference between this period and the last period:

Equation 2 – Naïve Method with Trend

$$\hat{Y}_{t+1=} Y_t + (Y_t - Y_{t-1})$$

It might happen that the time series has a strong seasonal pattern, in this case an appropriate forecasting equation for quarterly data might be:

Equation 3 – Naïve Method with seasonality

$$\hat{Y}_{t+1} = Y_{t-3}$$

Seasonal Naïve equation says that the next quarter will have the same value id did in the corresponding quarter one year ago.

Since Naïve method are very basic and suffers of considering the latest data to forecast the next period, it is usually used as a term of comparison between two models for evaluate the parsimony rule.
A method that can be applied easily by management when forecasts need to be updated daily, weekly, or monthly, and there is a large dataset, averaging or smoothing techniques performed best. These techniques weight average of past observations to smooth short-term fluctuation. The aim here is to use past data to develop a forecasting model for future periods.

Simple average is an appropriate technique when the forces generating the series to be forecasted are stabilized and the environment, in which the series exists is generally unchanging. Simple average uses the mean of all relevant historical observation as the forecast for the next period. A field of application might be the quantity of sales of a product in its maturity stage. Nonetheless, when forecasting many series simultaneously data storage may be an issue since it takes in account all the series.

Equation 4 – Simple average

$$\hat{Y}_{t+1} = \frac{1}{t}\sum_{i=1}^{t} Y_i$$

If a manager has interest in considering only a certain part of the time series, using more recent observation, a Moving averages method might be the right techniques: as a new observation is available, a new mean is computed by adding the newest value and dropping the oldest. *"A moving average of order k is the mean value of k consecutive observations. The most recent moving average value provides a forecast for the next period*" (John E. Hanke, 2014)

Equation 5 – Moving average.

$$M_t = \hat{Y}_{t+1} = \frac{Y_t + Y_{t-1} + \cdots + Y_{t-k+1}}{k}$$

Unfortunately, moving average technique does not perform well with nonstationary series (trends and seasonal components), to overcome this issue it is possible to use Double moving averages. To compute double moving average, once single moving average has been calculated, a second set is computed as a moving average of the first set.

Equation 6 – Double moving average

$$M'_t = \frac{M_t + M_{t-1} + \cdots + M_{t-k+1}}{k}$$

At this point to develop a forecast based on moving averages in "p" periods (Equation 7), it is possible to compute a_t coefficient (Equation 8) by adding to the single moving average the difference between the single and the second moving averages. Then we add an additional adjustment factor b_t (Equation 9) which is like a slop that can change over time.

Equation 7 – Moving averages forecasting equation.

$$\hat{Y}_{t+p=a_t+b_t p}$$

Equation 8 - a_t coefficient

$$a_t = M_t + (M_t - M'_t) = 2M_t - M'_t$$

Equation 9 - b_t adjustment factor

$$b_t = \frac{2}{k-1}(M_t - M'_t)$$

where:

k = the number of periods in the moving average

p = the number of periods ahead to be forecast

Despite of moving averages methods take in account only the most recent observation, there is no possibility to weight them, with simple exponential smoothing techniques, it is possible to provide an exponentially weighted moving average of all previously observed values. The aim is to continually revises an estimate in the light of more recent experiences. The most recent observation receives the largest weight α, with range $0 < \alpha < 1$, while the next most recent observation receives less weight, $\alpha(1-\alpha)$, and so on.

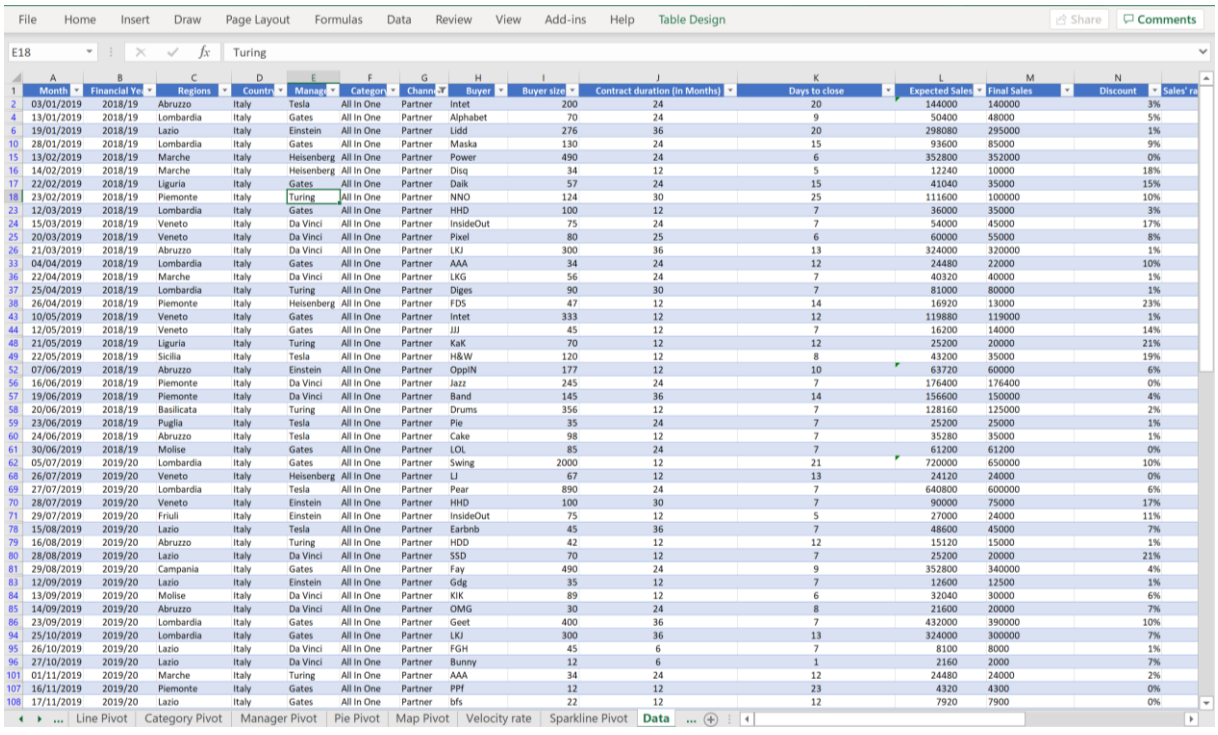
Equation 10 – Exponential smoothing formula

$$\hat{Y}_{t+1} = \alpha Y_t + \alpha(1-\alpha)Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + \cdots$$

The speed at which past observations lose their impact depends on the value of α: when it is close to 1, the new forecast will be quite similar the current observation, while when it is close to 0, the new forecast will be like the old forecast. A trick to identify the most useful α is to perform different tests with differ α coefficient, the one with the lowest MSE (Mean Square Error) will be the most adequate for that situation.

Build a Sales Dashboard in excel

For those companies that are becoming aware of the potential of data, I would like to suggest building your own dataset by recording or extrapolating (from ERP) orders details in an excel spreadsheet, like the one in figure 33. For this experiment, it was imagined a service company selling cloud services (Phone, Collaboration and Mail) through two main channels: online store and Partner (All-in-One). The All-in-One solution is accessible only for those customers that subscribed a yearly subscription directly with the company, involving a direct deal, in which alle the three services are included.



Month	Financial Ye	Regions	Countr	Manag	Categor	Channe	Buyer	Buyer size	Contract duration (in Months)	Days to close	Expected Sales	Final Sales	Discount	Sales' ra
03/01/2019	2018/19	Abruzzo	Italy	Tesla	All In One	Partner	Intet	200	24	20	144000	140000	3%	
13/01/2019	2018/19	Lombardia	Italy	Gates	All In One	Partner	Alphabet	70	24	9	50400	48000	5%	
19/01/2019	2018/19	Lazio	Italy	Einstein	All In One	Partner	Lidd	276	36	20	298080	295000	1%	
28/01/2019	2018/19	Lombardia	Italy	Gates	All In One	Partner	Maska	130	24	15	93600	85000	9%	
13/02/2019	2018/19	Marche	Italy	Heisenberg	All In One	Partner	Power	490	24	6	352800	352000	0%	
14/02/2019	2018/19	Marche	Italy	Heisenberg	All In One	Partner	Disq	34	12	5	12240	10000	18%	
22/02/2019	2018/19	Liguria	Italy	Gates	All In One	Partner	Daik	57	24	15	41040	35000	15%	
23/02/2019	2018/19	Piemonte	Italy	Turing	All In One	Partner	NNO	124	30	25	111600	100000	10%	
12/03/2019	2018/19	Lombardia	Italy	Gates	All In One	Partner	HHD	100	12	7	36000	35000	3%	
15/03/2019	2018/19	Veneto	Italy	Da Vinci	All In One	Partner	InsideOut	75	24	7	54000	45000	17%	
20/03/2019	2018/19	Veneto	Italy	Da Vinci	All In One	Partner	Pixel	80	25	6	60000	55000	8%	
21/03/2019	2018/19	Abruzzo	Italy	Da Vinci	All In One	Partner	LKJ	300	36	13	324000	320000	1%	
04/04/2019	2018/19	Lombardia	Italy	Gates	All In One	Partner	AAA	34	24	12	24480	22000	10%	
22/04/2019	2018/19	Marche	Italy	Da Vinci	All In One	Partner	LKG	56	24	7	40320	40000	1%	
25/04/2019	2018/19	Lombardia	Italy	Turing	All In One	Partner	Diges	90	30	7	81000	80000	1%	
26/04/2019	2018/19	Piemonte	Italy	Heisenberg	All In One	Partner	FDS	47	12	14	16920	13000	23%	
10/05/2019	2018/19	Veneto	Italy	Gates	All In One	Partner	Intet	333	12	12	119880	119000	1%	
12/05/2019	2018/19	Veneto	Italy	Gates	All In One	Partner	JJJ	45	12	7	16200	14000	14%	
21/05/2019	2018/19	Liguria	Italy	Turing	All In One	Partner	KaK	70	12	12	25200	20000	21%	
22/05/2019	2018/19	Sicilia	Italy	Tesla	All In One	Partner	H&W	120	12	8	43200	35000	19%	
07/06/2019	2018/19	Abruzzo	Italy	Einstein	All In One	Partner	OppIN	177	12	10	63720	60000	6%	
16/06/2019	2018/19	Piemonte	Italy	Da Vinci	All In One	Partner	Jazz	245	24	7	176400	176400	0%	
19/06/2019	2018/19	Piemonte	Italy	Da Vinci	All In One	Partner	Band	145	36	14	156600	150000	4%	
20/06/2019	2018/19	Basilicata	Italy	Turing	All In One	Partner	Drums	356	12	7	128160	125000	2%	
23/06/2019	2018/19	Puglia	Italy	Tesla	All In One	Partner	Pie	35	24	7	25200	25000	1%	
24/06/2019	2018/19	Abruzzo	Italy	Tesla	All In One	Partner	Cake	98	12	7	35280	35000	1%	
30/06/2019	2018/19	Molise	Italy	Gates	All In One	Partner	LOL	85	24	7	61200	61200	0%	
05/07/2019	2019/20	Lombardia	Italy	Gates	All In One	Partner	Swing	2000	12	21	720000	650000	10%	
26/07/2019	2019/20	Veneto	Italy	Heisenberg	All In One	Partner	LJ	67	12	13	24120	24000	0%	
27/07/2019	2019/20	Lombardia	Italy	Tesla	All In One	Partner	Pear	890	24	7	640800	600000	6%	
28/07/2019	2019/20	Veneto	Italy	Einstein	All In One	Partner	HHD	100	30	7	90000	75000	17%	
29/07/2019	2019/20	Friuli	Italy	Einstein	All In One	Partner	InsideOut	75	12	5	27000	24000	11%	
15/08/2019	2019/20	Lazio	Italy	Tesla	All In One	Partner	Earbnb	45	36	7	48600	45000	7%	
16/08/2019	2019/20	Abruzzo	Italy	Turing	All In One	Partner	HDD	42	12	12	15120	15000	1%	
28/08/2019	2019/20	Lazio	Italy	Da Vinci	All In One	Partner	SSD	70	12	7	25200	20000	21%	
29/08/2019	2019/20	Campania	Italy	Gates	All In One	Partner	Fay	490	24	9	352800	340000	4%	
12/09/2019	2019/20	Lazio	Italy	Einstein	All In One	Partner	Gdg	35	12	7	12600	12500	1%	
13/09/2019	2019/20	Molise	Italy	Da Vinci	All In One	Partner	KIK	89	12	6	32040	30000	6%	
14/09/2019	2019/20	Abruzzo	Italy	Da Vinci	All In One	Partner	OMG	30	24	8	21600	20000	7%	
23/09/2019	2019/20	Lombardia	Italy	Gates	All In One	Partner	Geet	400	36	7	432000	390000	10%	
25/10/2019	2019/20	Lombardia	Italy	Gates	All In One	Partner	LKJ	300	36	13	324000	300000	7%	
26/10/2019	2019/20	Lazio	Italy	Da Vinci	All In One	Partner	FGH	45	6	7	8100	8000	1%	
27/10/2019	2019/20	Lazio	Italy	Da Vinci	All In One	Partner	Bunny	12	6	1	2160	2000	7%	
01/11/2019	2019/20	Marche	Italy	Turing	All In One	Partner	AAA	34	24	12	24480	24000	2%	
16/11/2019	2019/20	Piemonte	Italy	Gates	All In One	Partner	PPf	12	12	23	4320	4300	0%	
17/11/2019	2019/20	Lazio	Italy	Gates	All In One	Partner	bfs	22	12	12	7920	7900	0%	

Figure 33 – Example of a simple excel Database

In detail, in the database that has been used, there were recorded the following data:

- Date and Financial year of the order
- The region and country of destination of the order

- The manager accounted for the deal
- The product/service sold, and the channel used
- The buyer credentials
- Months of contracts
- The forecasted price and the final price
- The discount applied and salesperson satisfaction

By combining those data has been possible to generate a useful dashboard that has the purpose to make managers aware of his current business state. It is possible to update the database just writing a new line of data right after the last row of the current dataset.

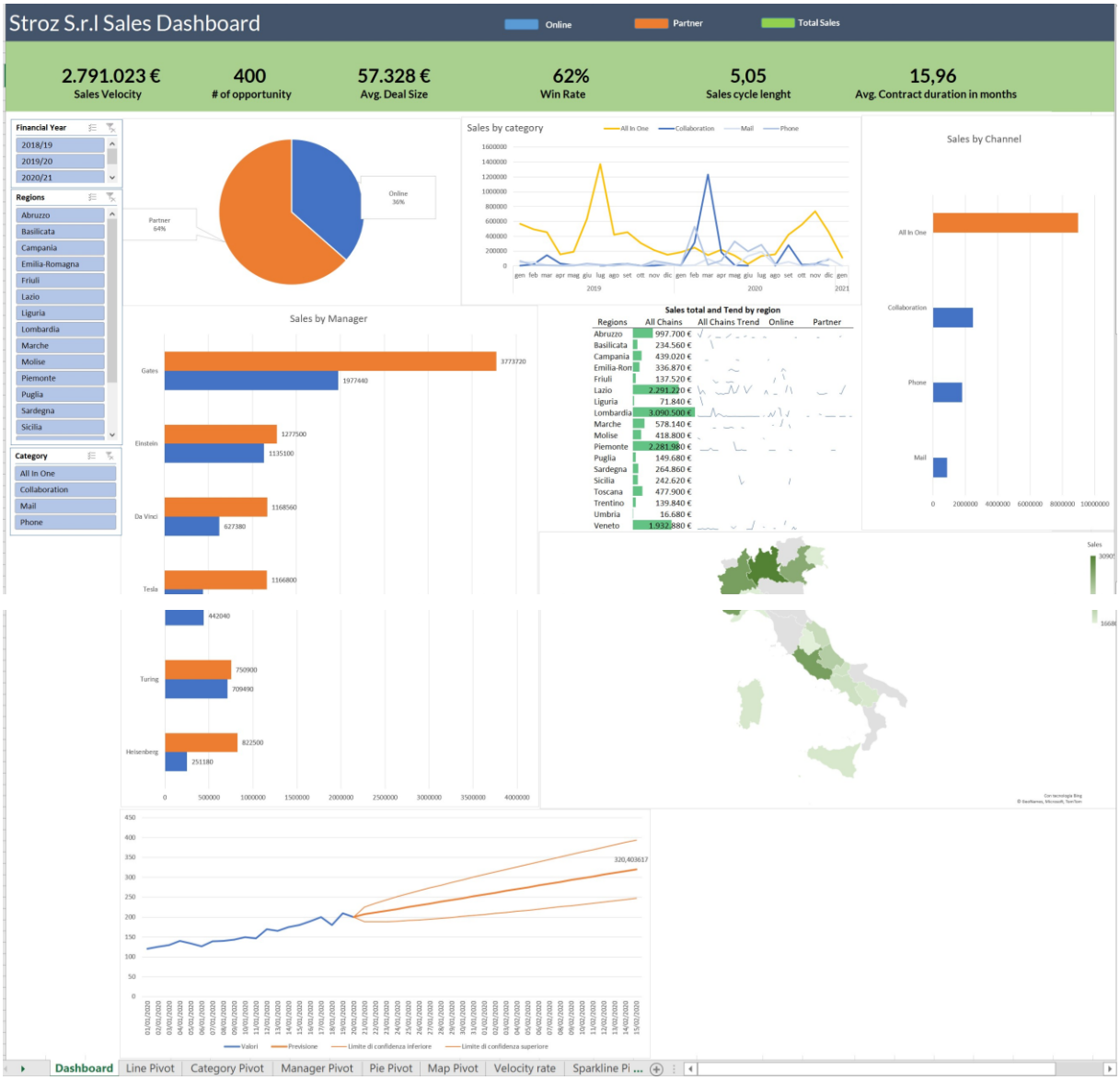


Figure 34 – Sales dashboard

The dashboard is compose by different elements:

- *Pie chart*: it has the aim to show the persentage of sales between the two channels, in this experiment, Partner contracts generated more sales compared to the online contracts;

- *Category Chart*: it has the aim to display the stream of sales of every category among All-in-One, Phone, Mail and Collaboration solution; by observing this chart, a manager can be aware of which product was performing better in a certain period and which was the peak of sales;

- *Manager chart*: it has the aim to diplay which salesperson/sales manager is performing best, in which area and which service is selling best; it is important to be aware of how people are performing inside the company, in this way management can adjust their role according to they strenght points;

- *Sales by region:* by analysing the sales performances in a specific area or in the overall country makes possible to set up ad hoc strategies and entry barriers to deny access to possible new entrants, discont policy and be aware of the company market share;

- *Sales forecasting:* by selecting the "Data" tab, select the "Data Analysis" tool (figure 34), here it is possible to compute some forecasting techniques that were illustrated above, and even more possibilities. In the example it was possible to forecast sales for the second part of january and the first part of february; remember that forecasting activities are important not to incur in stock outs or to be prepared for market fluctations.

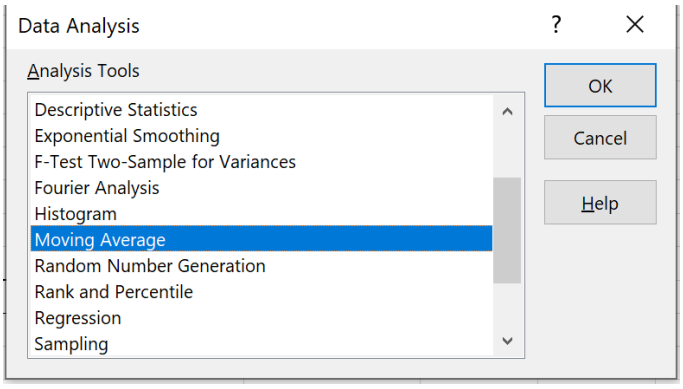


Figure 35 – Data Analysis tab

	A	B	C	D	E
1	Sequenza temporale	Valori	Previsione	Limite di confidenza inferior	Limite di confidenza superior
2	01/01/2020	120			
3	02/01/2020	125			
4	03/01/2020	130			
5	04/01/2020	140			
6	05/01/2020	134			
7	06/01/2020	127			
8	07/01/2020	139			
9	08/01/2020	140			
10	09/01/2020	143			
11	10/01/2020	150			
12	11/01/2020	147			
13	12/01/2020	170			
14	13/01/2020	165			
15	14/01/2020	175			
16	15/01/2020	180			
17	16/01/2020	190			
18	17/01/2020	200			
19	18/01/2020	180			
20	19/01/2020	210			
21	20/01/2020	200	200	200,00	200,00
22	21/01/2020		206,9402	188,34	225,54
23	22/01/2020		211,47874	188,22	234,74
24	23/01/2020		216,01727	188,88	243,16
25	24/01/2020		220,55581	190,02	251,09
26	25/01/2020		225,09435	191,49	258,70
27	26/01/2020		229,63288	193,21	266,05
28	27/01/2020		234,17142	195,13	273,21
29	28/01/2020		238,70996	197,21	280,21
30	29/01/2020		243,24849	199,42	287,08
31	30/01/2020		247,78703	201,74	293,83
32	31/01/2020		252,32557	204,16	300,49
33	01/02/2020		256,8641	206,66	307,06
34	02/02/2020		261,40264	209,24	313,57
35	03/02/2020		265,94118	211,88	320,00
36	04/02/2020		270,47971	214,59	326,37
37	05/02/2020		275,01825	217,34	332,69
38	06/02/2020		279,55679	220,15	338,97
39	07/02/2020		284,09532	223,00	345,19
40	08/02/2020		288,63386	225,89	351,38
41	09/02/2020		293,1724	228,82	357,53
42	10/02/2020		297,71093	231,78	363,64
43	11/02/2020		302,24947	234,78	369,72
44	12/02/2020		306,78801	237,81	375,77
45	13/02/2020		311,32654	240,87	381,79
46	14/02/2020		315,86508	243,95	387,78
47	15/02/2020		320,40362	247,06	393,75

Figure 36 – Example of forecasting tool in excel.

The real power of this dashboard is the possibility to insert sliders. Sliders help in filtering our dataset according to company needs, in the experiment it was possible to filter by financial year, by region and by product. Below it is shown a demonstration of Veneto region.

In Veneto region example, by easily selecting the desirable region from the slider on the left side of the screen, it is possible to observe that: Veneto region theoretically generate € 447.061 per day (Sales velocity), around 16% of the overall company' sales in Italy; the All-in-one solution has been the most sold solution in the region and Gates is the best seller.

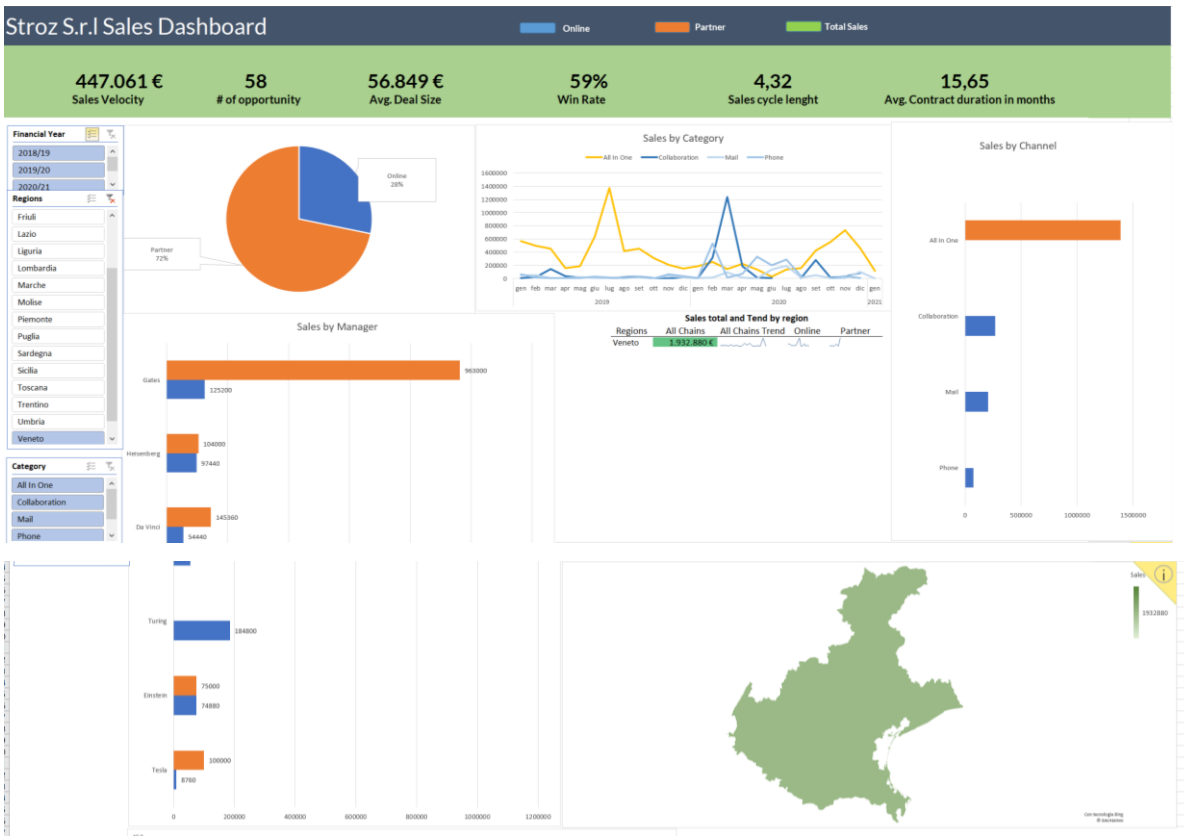


Figure 37 – Veneto region insight

Despite the simplicity and limits of the solution, it is a first step to get used to manage different data and visualize it in an easy and clean dashboard that might help in decision making process. It is important to pointed out that, the difference is made not by data itself but the way they are presented and visualized, easily to share and to comment among the company organization. In the following solution, it is been shown a further integration with tools specialized in business intelligence and data visualization.

The ultimate data-driven solution: become a Strategic Transformer

The master goal of a company, when it invests in a new technology, is to exploit its hidden potentiality and obtain a superior competitive advantage on the market, leading to superior returns. For this reason, a company with the aim to become a strategic transformer must adopt a data-driven culture that carries to internal operation enhancement and it must be recognise as an innovative firm by the market. To accomplish this role, a company must be aware of both its internal and external data sources. In this chapter it will be explained how to integrate and exploit companies' data using Microsoft Power BI, and how to research new opportunities from the market through Social listening tools.

Improve your data awareness with Microsoft Power BI

Microsoft Power BI is a business intelligence tool with the aim to quickly connect your data and easily visualize them in a user-friendly customable dashboard, sharing insights with your team to *"bridge the gap between data and decision making"*.

Figure 38 – Magic Quadrant for Analytics and Business intelligence Platforms – Gartner (2020)

According to Gartner, Power BI is recognized as a leader for the thirteenth consecutive year in the Gartner 2020 Magic Quadrant for analytics and Business Intelligence, an easy and accessible tool. Power BI desktop version is free, while the Pro is available for a low monthly price per user (around € 9), a costly efficient solution that offers to everyone the possibility to experience analytics capabilities and sharing a data-culture mindset in your company. The flexibility of software, united to a wide range of possible integrations, makes possible to create ad hoc dashboards for every need: HR, operations, sales, marketing campaign, accounting insight, R&D. Among the wide range of choice of data source, it is possible to link different sources of data together in few clicks. Examples of ready to use data integration sources are excel, Azure cloud solution, Google Analytics, Facebook, LinkedIn, and Survey Monkeys.

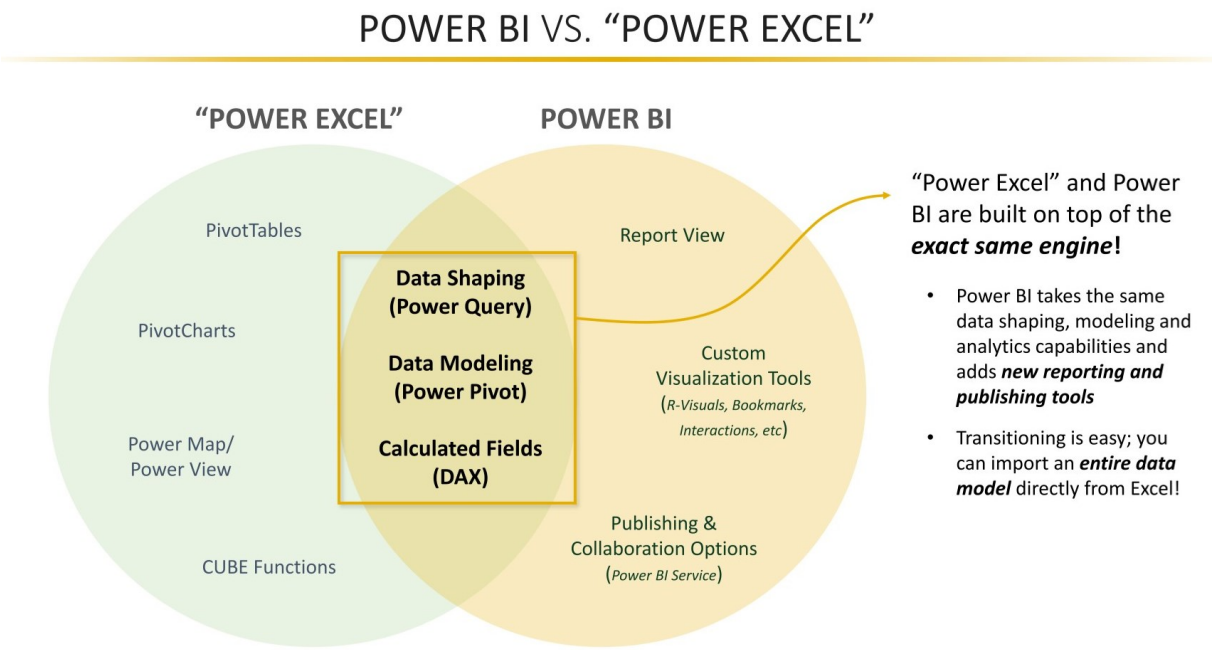


Figure 39 – Power BI vs Excel (Dutton, 2019)

The following Demo show the case of a reseller willing to integrate multiple databases to obtain powerful insights about the current business state, products, and customers.

The datasets uploaded in Power BI are composed by:

- *Customer Lookup*: contains all the information about the customer such as full name, birthday, current age, customer priority, domain, education level, income level, occupation and so on.

- *Calendar Lookup:* it contains all the information about the data of the orders, date name, months, and year.

- *Territory Lookup:* contains information about sales location around the globe, sorted by continent, country, region.

- *Product Lookup, Category and Subcategory* contains every information about products characteristics such as model name, colour, size, name, SKU, price.

- *Sales:* contains all the relevant information about the order, sales, revenue, profit, costs, average price and so on.

- *Returns:* contains all the relevant information about returns, returns rate.

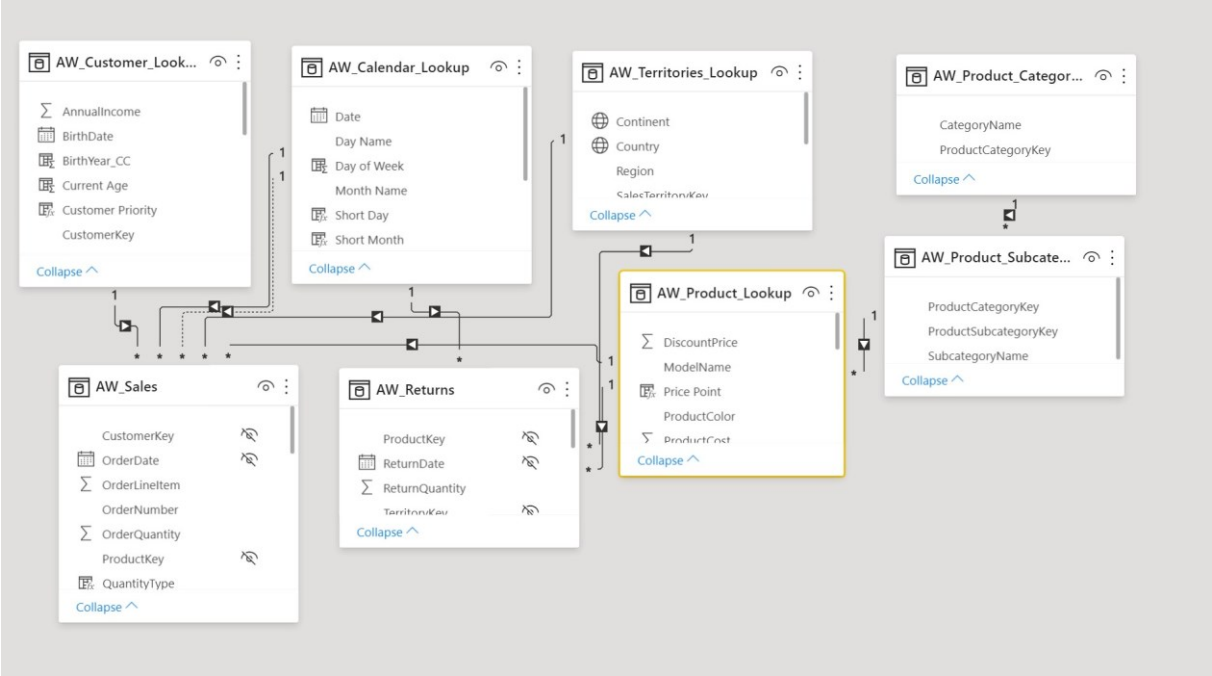


Figure 40 – datasets integrated in Power BI demo

Sales and Returns datasets are called *Data tables* while the others are called *Lookup tables*: the former contains mostly numerical value while the latter mostly descriptive values. Sales and Returns datasets are the central component of the case, indeed, as it is possible to observe in figure 40, both are linked with the other lookups' datasets through a series of foreign and primary keys. For example, *Sales* is linked to the Customer lookup table through the *Customer key* row, present in both the dataset. Once datasets are fully loaded in power BI, the analyst needs to operate a *normalization procedure,* that is the process of organizing the tables and column in a relational database to reduce redundancy and preserve data integrity. To normalize

a database, an analyst should eliminate redundant data to decrease the table size, minimize errors and anomalies and simplify queries, each table should serve a distinct and specific purpose.

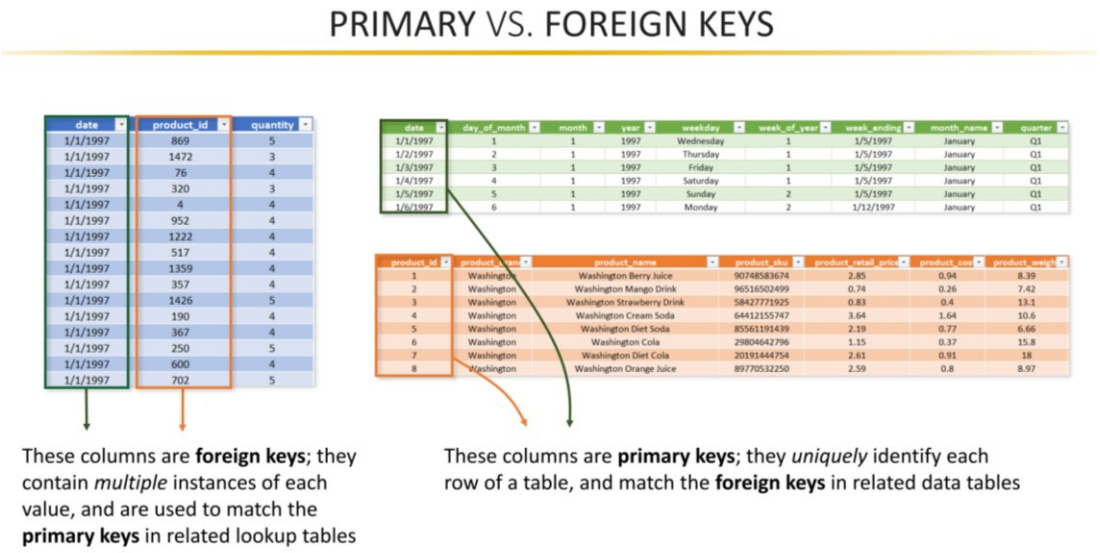


Figure 41 – Primary vs Foreign keys (Dutton, 2019)

The Executive summary dashboard

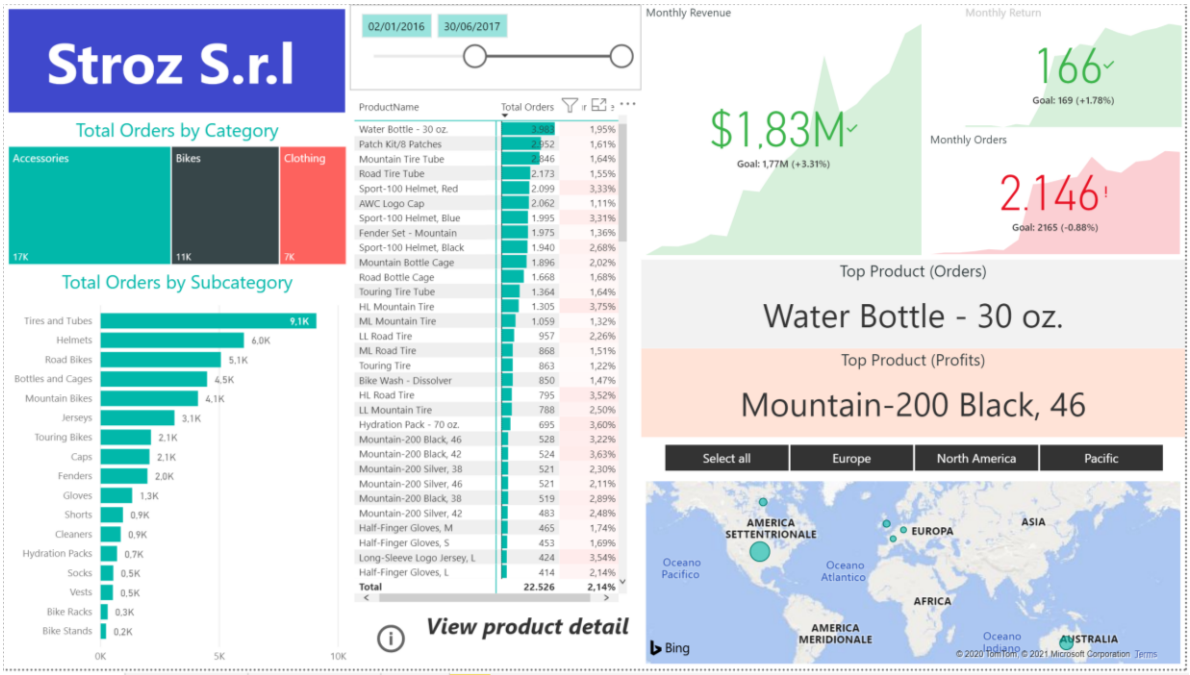


Figure 42 - Power BI Demo: Executive summary dashboard

The executive summary dashboard has the aim to resume the overall orders' situation according to sales' measures. It is a powerful tool since every manager can check the status of company's products and filter them just clicking on every field insert in there. Moreover, in Power BI it is possible to set managerial role, in this case was filtered by continent: a manager that operates in Europe would have the access only to European data.

Starting from the left side of the dashboard (figure 42), it has been placed a Tree map with the aim of resuming the number of orders by category (accessories, bikes, and clothing). By just looking up at it a manager can understand orders proportion by rectangles dimensions. By clicking one category it is possible to filter the whole dashboard according to that specific category. Just below the tree map, it is placed a stacked bar chart that visualise every subcategory by total orders, in this case, tires and tubes are ranked first.

Next to the company label, it is inserted a data slider, by adjusting the slider it is possible to select a determined range of period, useful to visualise company's performance in different period and create benchmarks. Below the data slider it is placed a matrix that ranks total orders according to product name and returns rate, the latter, the higher it is, the worst is, so, it is been used a conditional formatting to highlight in red those products with a high return rate.

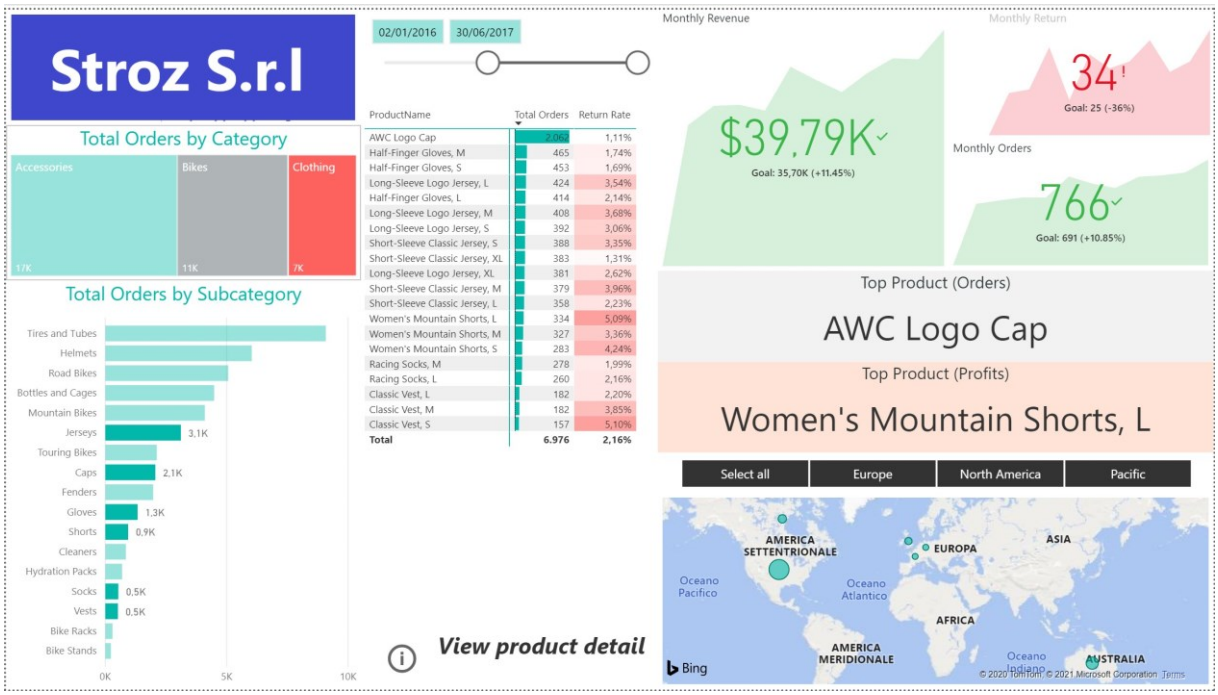


Figure 43 - Power BI executive summary filtered by Clothing

On the right side of the dashboard, it is place mainly KPI tools to assess if the company is achieving its monthly goals in terms of revenues, return and orders: in green colour if the company is succeeding, in red if not. Right below, two Card fields are placed to display top products by orders and profits. Lastly, a Map is inserted in the bottom right corner to sort orders by continent, country, region.

Product detail dashboard

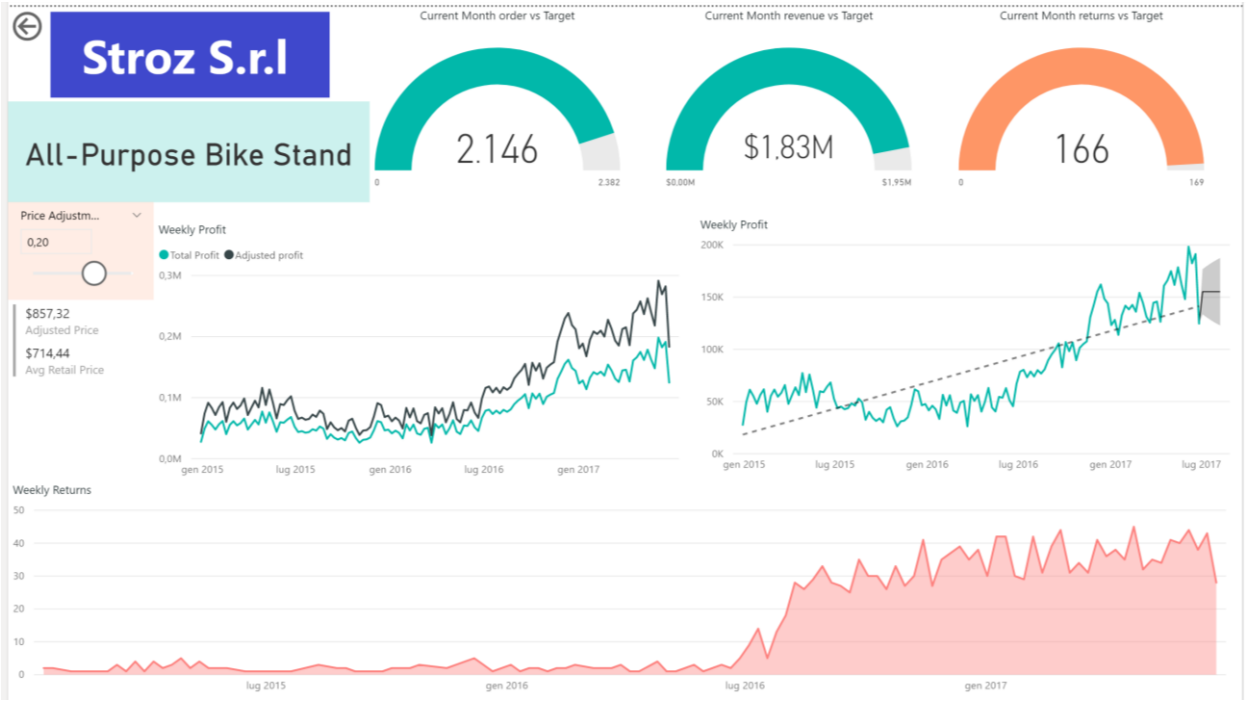


Figure 44 - Product detail dashboard

The second tab created for this demo is the product detail dashboard, it has the aim to obtain a deep understanding about single/group product performances over a certain period.

Starting from the top (figure 44), there were inserted three Gauge data visualisation with the aim to highlight current product situation versus company target. Stroz s.r.l in this case did not still reached its achievements, both in month order and revenues while the month returns for that period was reached.

Right below Gauge tools, have been inserted line charts with the aim to display economic measures. The two weekly profit charts have two differ aim: the former can be filtered by price adjustment filter, creating instantaneous "what-if scenarios", in figure 44 was simulated the case of increasing "All-Purpose Bike stand" price by 20%, indicated with the green line (below the black line), the graph shows that with an increase in price, weekly profit would fall; the latter shows sales trend and forecasting with a confident interval set to 95%. Lastly, the bottom chart shows weekly return, in this case the company had seen an increase in weekly revenues by July 2016. By investigating what happened during 2016 period, Stroz s.r.l expanded bikes category, by introducing MTB line. Not surprisingly, in the executive summary tab it is showed Mountain-200 black as the most profitable product.

Like the previous tab, here it is possible to filter the whole dashboard by products and periods, a useful tool to track products performances and returns.

Customer detail dashboard

The customer detail dashboard has the aim to let the manager visualize customer composition, by gender, income level, occupation, and age (figure 45). Moreover, it is possible to rank customers according to total orders and total revenues.

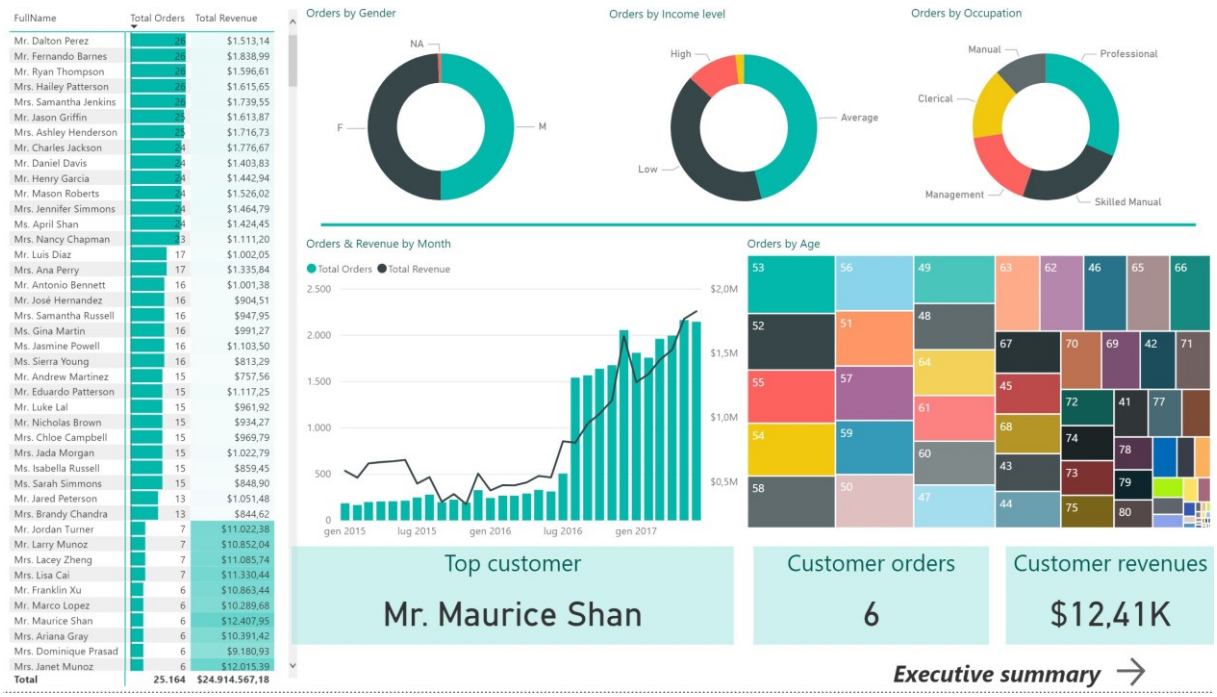


Figure 45 - Customer detail dashboard

Figure 46 - Customer detail dashboard, filtered by client

For example, if we sort customers by total revenue, Mr. Maurice Shan (figure 46) is company's top customer, as we click on him, the dashboard displays all customer details: he is a male of 70 years old, with an average income, he is a professional profile and generates $ 12,41 thousand of revenue, and his top peak of sales was in July 2016.
Though a customer details dashboard a company might improve their communication and discounting practices by visualising the different target groups, creating solution ad hoc for each customer group and forecast new market opportunities just by observing internal data.

Lastly, in Power BI it is possible to employ the AI algorithm to discover new relationships between databases, this can be applied by using the Q&A tool (Figure 47).

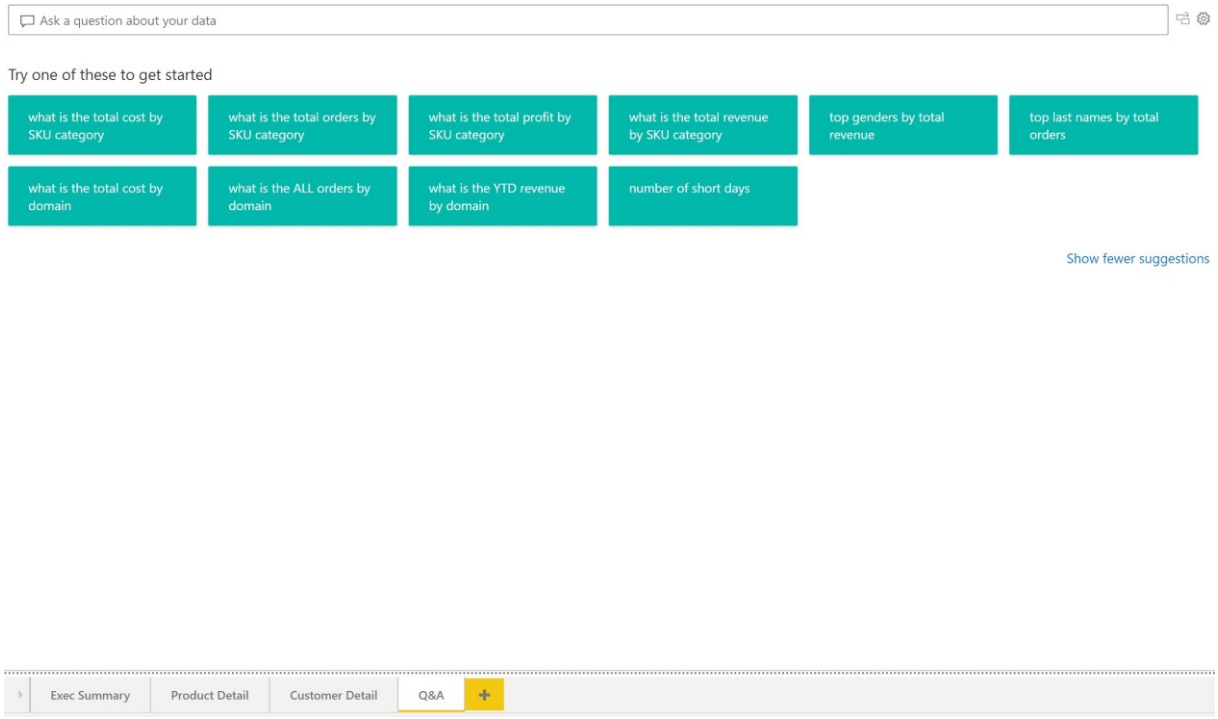


Figure 47 - Power BI Q&A tool

Q&A tab is a powerful tool, by suggesting different databases' interrogations, even if it is possible to manually insert new relationship. By writing in the research bar, Power BI will provide to the managers the data and relationship they asked for. For example, by asking "Profit by bike colour" (figure 48), Power BI displays a bar chart: black bikes are the one that generates most profits over the other colours.

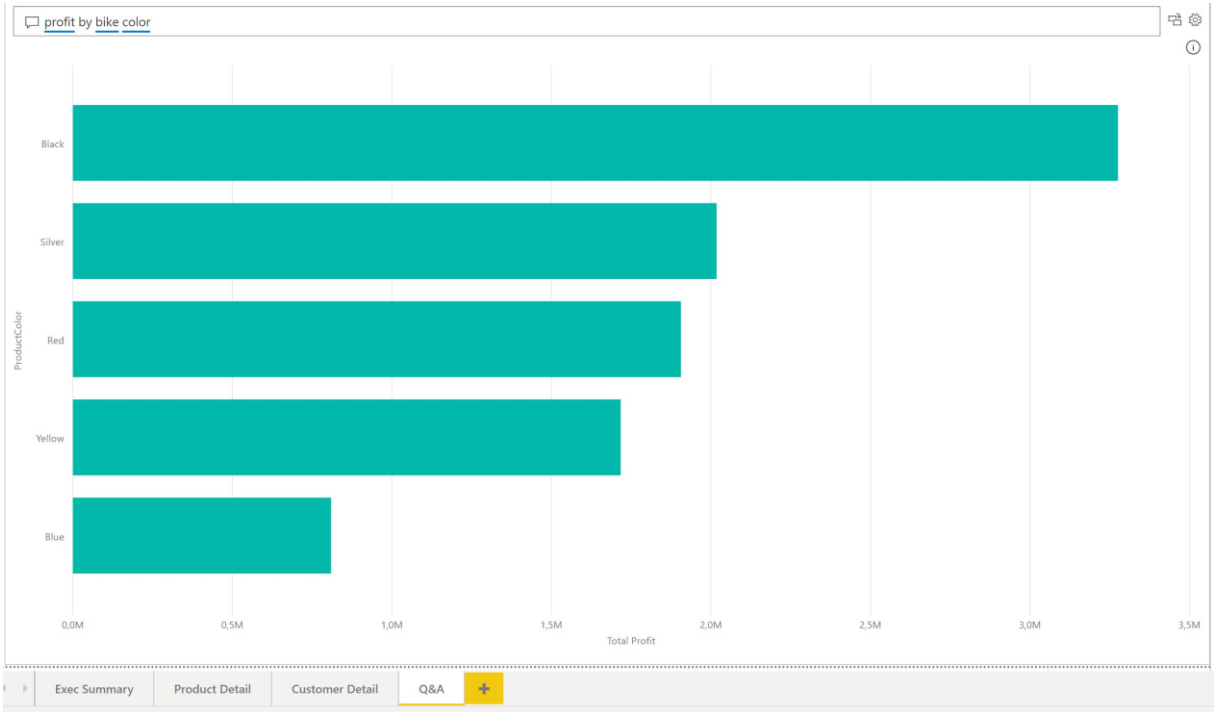


Figure 48 - Power BI Q&A tool example

Key Influencers tool

The latest Power BI tool presented here is the *Key influencer tool*. Its purpose is to visualize concrete data by changing the formatting, bring it to life in a visual form; it helps managers to understand and expose individual factors that drive some sort of an outcome. This tool uses regression models, either a logistic regression for categorical variables or linear regression for numerical variables. For this demo it was been used a dataset based on Kickstarter campaigns.

project_id	project_name	subcategory	category	goal_amount	launched	project_outcome	backers	country	amount_pledged (USD)	Year
1098244109	Ghost Vinyl: Insert Meme Here	Indie Rock	Music	$5.000	11/10/2017 15:12:38	failed	0	US	$0	2017
1143821987	"Live My Life Again"	World Music	Music	$5.000	24/06/2017 22:42:35	failed	0	US	$0	2017
1154095172	Gig and Studio Equipment (Keyboard, laptop)	Jazz	Music	$5.000	24/01/2017 16:50:19	failed	0	US	$0	2017
1176678994	Passion in Perfection Music Visual	Music	Music	$5.000	15/01/2017 14:38:57	failed	0	US	$0	2017
1199494968	Afro Desi	World Music	Music	$5.000	16/08/2017 19:42:13	failed	0	US	$0	2017
1224229585	(Thanksgiving Challenge) Thanksgiving (Thanks4giving)	Hip-Hop	Music	$5.000	10/11/2017 07:06:51	failed	0	US	$0	2017
123989156	Music with meaning	Music	Music	$5.000	11/01/2017 08:19:58	failed	0	US	$0	2017
127603380	The Dark Night and The Morning After- (3 Hip Hop Albums	Music	Music	$5.000	08/09/2017 21:09:17	failed	0	US	$0	2017
1323784832	SnapN'Rock	Music	Music	$5.000	30/05/2017 16:54:11	failed	0	US	$0	2017
1338677766	Electric Kid Dance Party	Music	Music	$5.000	15/05/2017 17:21:42	failed	0	US	$0	2017
1346600121	Anything's Possible EP Promotion & Touring	Music	Music	$5.000	07/04/2017 19:14:11	failed	0	US	$0	2017
1367599041	I.N.C. Music & Talk www.rustywilkens.com	Music	Music	$5.000	17/10/2017 17:37:35	failed	0	US	$0	2017
139243480	Sound Off Onstage Showcase & Tour	Music	Music	$5.000	06/10/2017 18:10:32	failed	0	US	$0	2017
1443647987	AMD1 ILLUMINATED THE OSIRIS CODE	Hip-Hop	Music	$5.000	04/01/2017 08:43:50	failed	0	US	$0	2017
1445480862	Pride, Prejudice & Power Project	Pop	Music	$5.000	02/05/2017 21:54:28	failed	0	US	$0	2017
1454866263	AMD1 The Illuminated One	Hip-Hop	Music	$5.000	11/08/2017 04:11:29	failed	0	US	$0	2017
1506402901	Starving Artist	Hip-Hop	Music	$5.000	04/07/2017 16:45:17	failed	0	US	$0	2017
1506756111	Eusini - Salem Bottom	Indie Rock	Music	$5.000	27/08/2017 02:42:00	failed	0	US	$0	2017
1528451272	My First Album A House Is Not A Home *Rebooted*	Pop	Music	$5.000	25/09/2017 14:49:48	failed	0	US	$0	2017
1545624098	Just trying to make it in the game	Hip-Hop	Music	$5.000	09/09/2017 08:52:04	failed	0	US	$0	2017
15826392	Chico Gonzalez: NEW ALBUM & 3 Music Videos	Hip-Hop	Music	$5.000	13/10/2017 11:47:22	failed	0	US	$0	2017
1583444157	Blue: In My Feelings	R&B	Music	$5.000	11/06/2017 00:30:04	failed	0	US	$0	2017
1596078892	backers for smg	Hip-Hop	Music	$5.000	25/05/2017 01:42:39	failed	0	US	$0	2017
1609904344	Jay Davi for you EDM Fans! Electronic Dance Music Songs	Electronic Music	Music	$5.000	31/03/2017 22:06:37	failed	0	US	$0	2017
1613308771	Bring Jazz Back To Chattanooga	Jazz	Music	$5.000	05/09/2017 19:35:01	failed	0	US	$0	2017
1615650015	Big Black Truck Album Project	Country & Folk	Music	$5.000	31/05/2017 12:09:21	failed	0	US	$0	2017
1652253982	Jacqua Cooper's "Solo Music Album" away from O.K. Drun	Electronic Music	Music	$5.000	06/10/2017 18:15:32	failed	0	US	$0	2017
1684634046	support my first audio and video record.	Faith	Music	$5.000	12/11/2017 21:40:06	failed	0	US	$0	2017
1690045152	Studio Album Financing	Country & Folk	Music	$5.000	03/08/2017 20:47:29	failed	0	US	$0	2017
170674562	Kpop Fan Meet & Greet in the US	World Music	Music	$5.000	24/02/2017 19:41:52	failed	0	US	$0	2017

Figure 49 - Kickstarter dataset (Dutton, 2019)

Kickstarter is the most famous web crowdfunding platform where start-up can raise money to develop and sell new products and solutions. The database is composed by the following data: product ID, project name, category and subcategory, funds goal amount, launch date, project outcome (succeed/failed), number of backers, country of origin, amount pledged and the year.

The demo starts by analysing a categorical field as the project outcome; the goal is to discover and visualise what variable impacted successful projects by indicating to Power BI the possible explanations for the outcome.

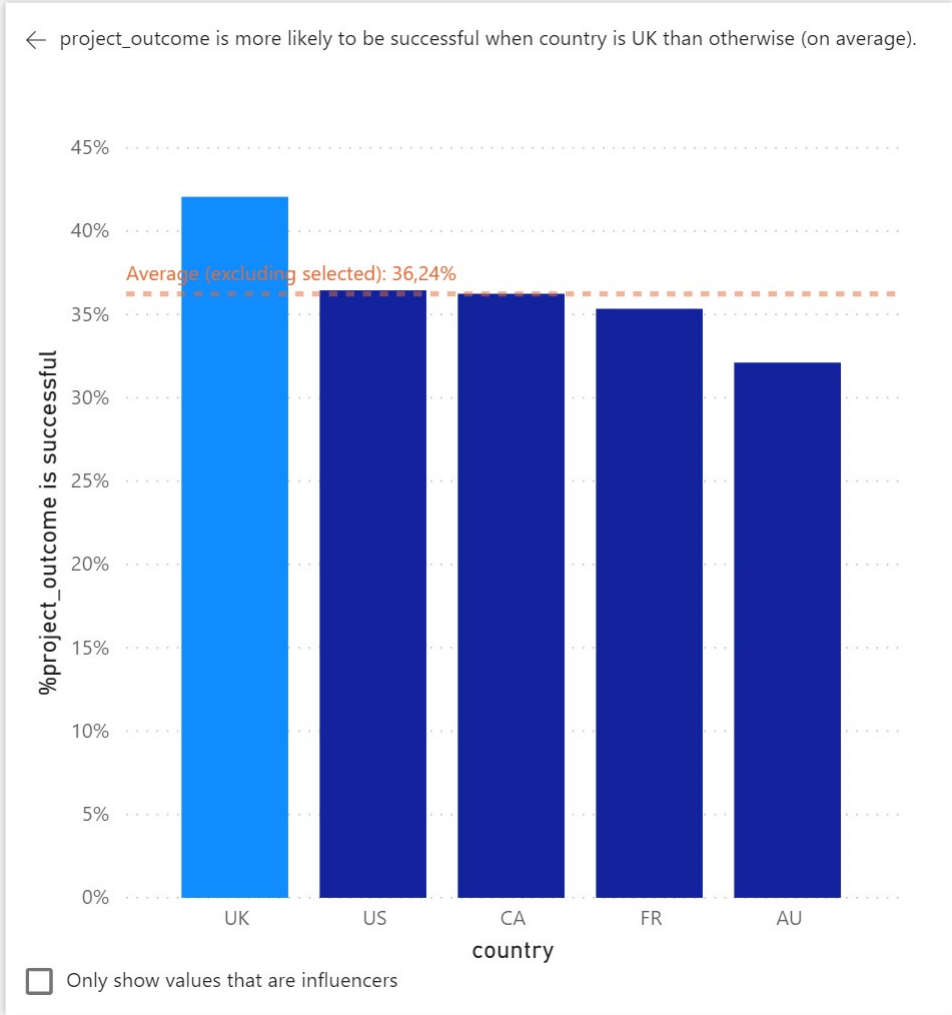


Figure 50 - Kickstarted successful project by country

By dropping in the country variable, Power BI suggest that those projects started in UK, had the likelihood to be successful increases by 1.11x. The variables used by Power BI in the regression model to address to UK the start-up green field leadership were success rate compared to the average and the volume of projects. Since in the dataset is present many other

significant factors that can explain and influence project success, this time we will drop in category, backers, country, and amount pledged, we obtain a totally different result (figure 51).

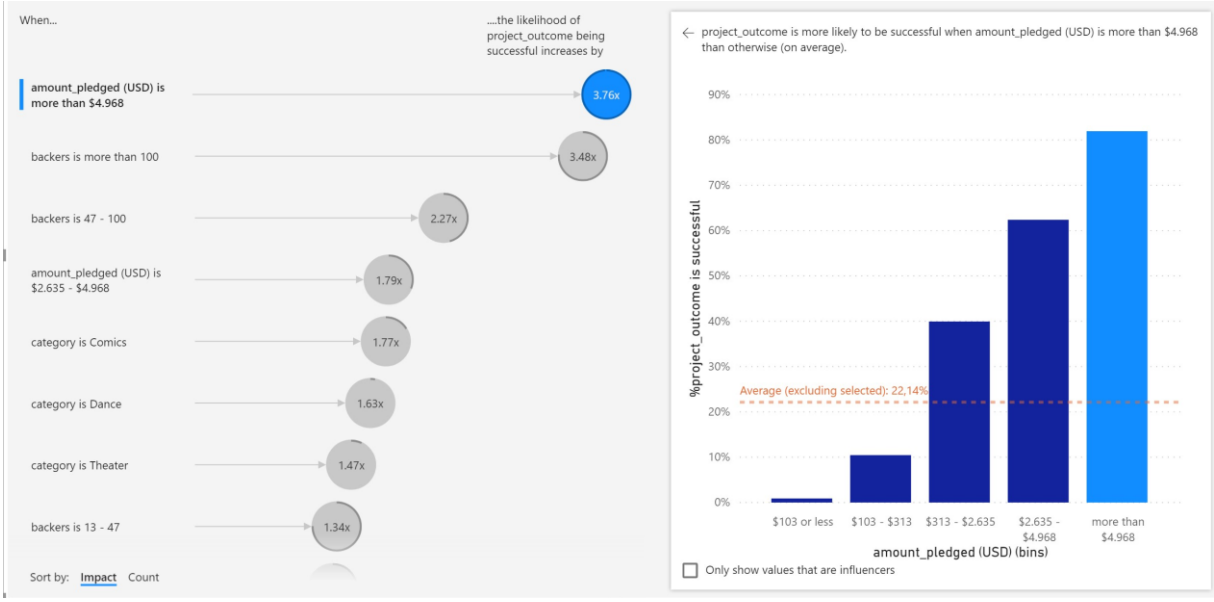


Figure 51 - Kickstarter categorical segmentation

- When the amount pledged is more than $ 4.968 the likelihood a project will being successful increases by 3.76x.
- When backers who believed in the project are in range 47 – 100 the likelihood a project will being successful increases by 3.48x
- When the project belongs to Comics category the likelihood it will be successful increases by 1.77x

With the addiction of more explanatory variables, now the highly chance that a project succeeds being in UK is ranked 12 (with 1.11x).

Moreover, clicking on “Top segment” tab, Power BI compute a clustering analysis, segmenting the population based on project outcomes (figure 52). This is a useful tool for those companies that wants so find a specific target group according to some relevant variables. By clicking on each bubble, Power BI will show the variables that affected the segmentation, for example, start-ups which belongs to segment 1 had 88% of success rate when backers are greater than 100 and category is not gaming. Start-ups which belong to segment 2 succeeded 83,2% of the time when backers are greater than 100 and are all related to game category. On the other hand, segment 6 contains all those start-ups with 46,8% of succeed rate when the amount pledged is

greater than $ 313 and it is less than or equal to 47, and backers are greater than 13 but less or equal to 47.

Figure 52 - Power BI Top segments analysis

The next step is to analyse a numerical variable, the amount pledge, what variable allowed it to increase its value? We will start the analysis by insert the backer's variable in the "explained by" field. The result is shown in figure 53 & 54.

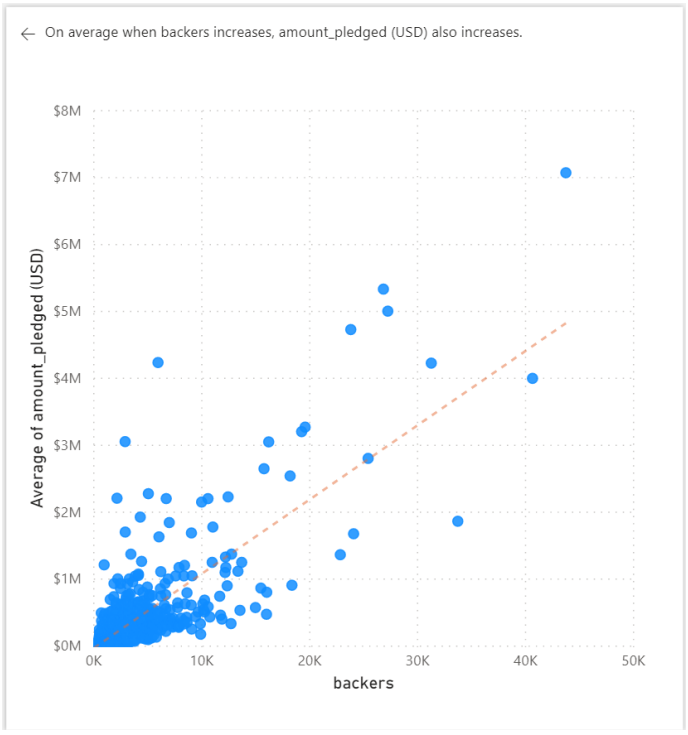


Figure 53 – Key influencers – Linear regression

The result shows, not surprisingly, that when the numbers of backers increase, the average of amount pledge increases by $ 80 thousand. Next, we want to analyse the average amount pledge by category. To compute this analysis, it was been asked to Power BI to analyse the average amount pledged, explain by the amount pledged, expand by “category” variable. The result showed that Design, Tech and Games categories attract more investors.

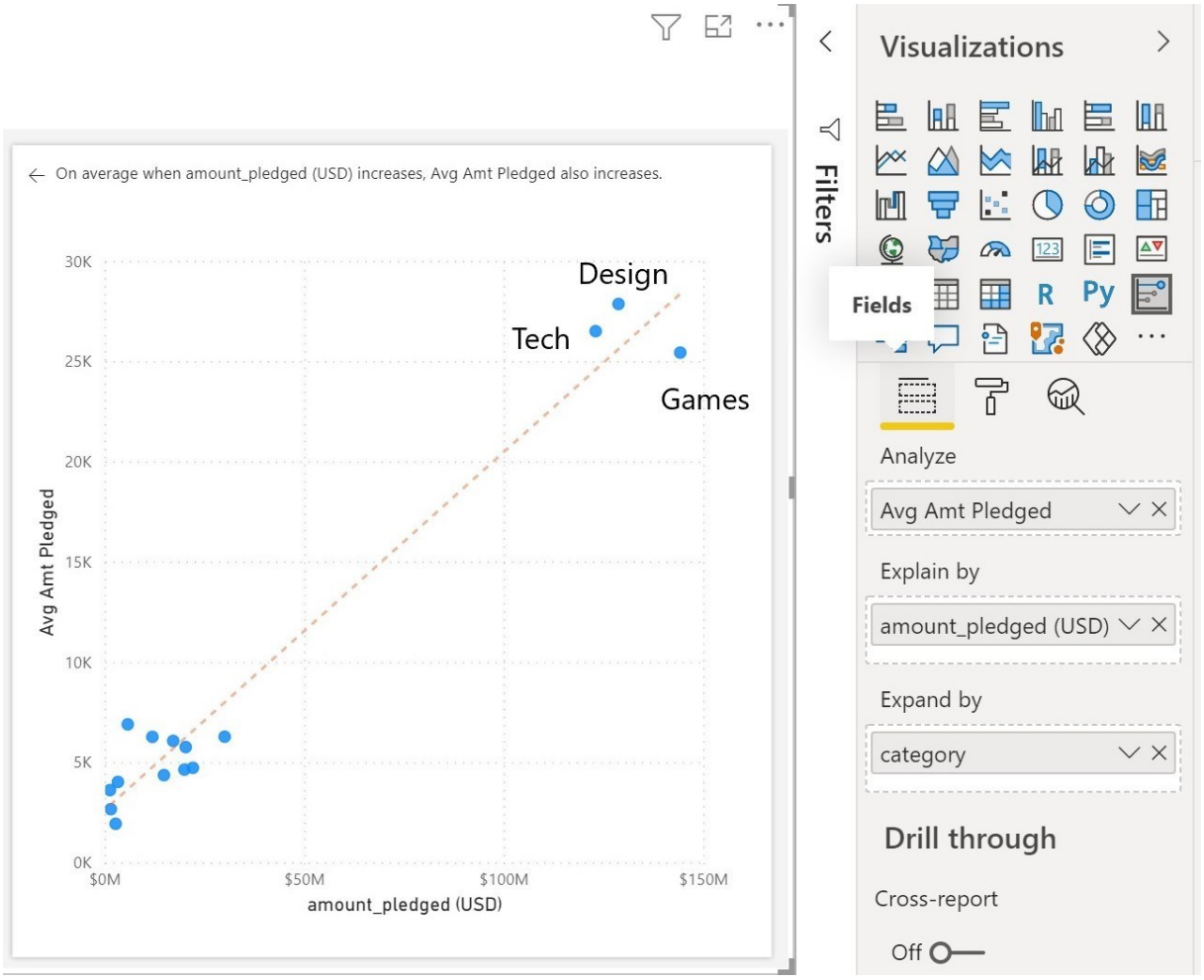


Figure 54 - Amount pledged by category

To conclude this demo, the last Power BI tool presented is the decomposition tree. This usefull tool generate a visual hierarchy based on, in this case, the amount of pledge collected (figure 55). By decomposing the tree structure, it is possible to visualise that the category with the most amount of pledge collected was Games, the best subcategory was Tabletop Games and “The 7th Continent game” was the one that collected the highest pledge, $7.072.757. It is possible to create the composition tree manually, by exploring category per category, or by using Power BI integrated AI that automatically suggest the highest or the lowest score.

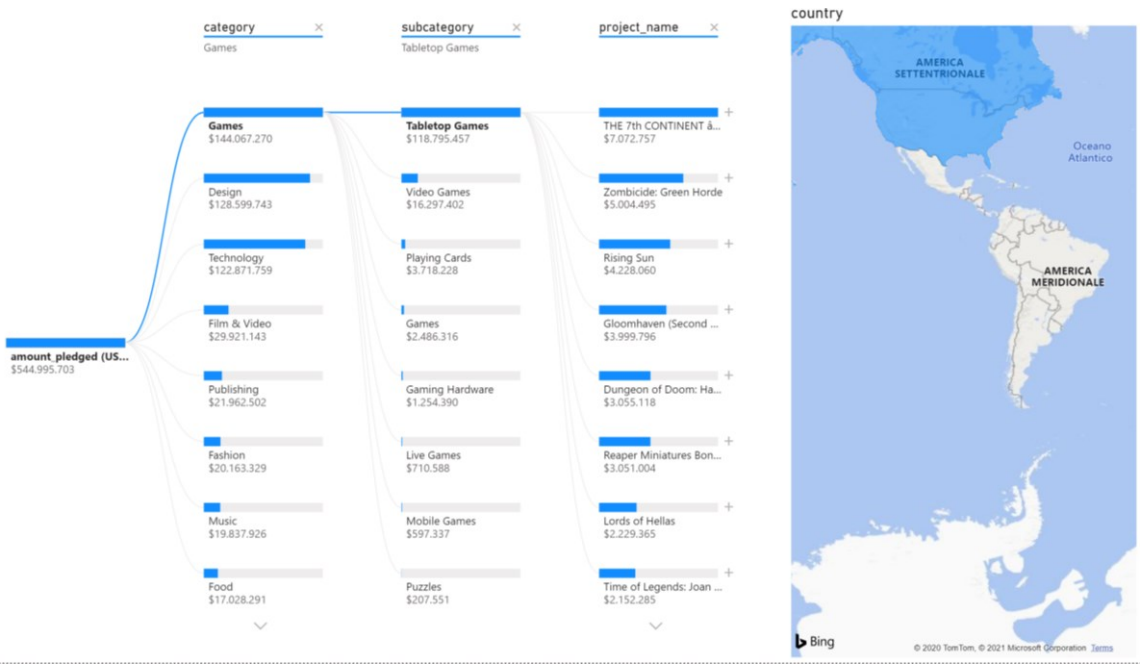


Figure 55 - Power BI decomposition tree

Boost your sales by listening to your clients: Social Listening.

It is true, companies can achieve more just by exploiting their own data and visualising them in a proper and effective way, nonetheless, listening to the market is a fundamental source of success, for this purpose in this paragraph we are going to introduce social listening tools to discover new business opportunities and customer's need.

What is social media listening? It is an activity consisting in monitoring, identifying, and assessing everything that has been said or is being said about a brand, a product, or a person. Social listening analyses unstructured data deriving by social networks (mostly twitter) and let the analyst to visualize those data in an easy and effective way. Those unbranded conversation are the result of multiple interaction created by user, possibly made by the target group of interest. Social listening goal is to intercept the power of word-of-mouth, one of the most promotional tools (Nielsen, 2015), by filtering recommendation from friends/family and online communities.

Why is social listening important?

- To address indirect complaints
- To uncover unique opportunities
- To get unfiltered feedback

Social listening is not only restricted to marketing campaigns and CRM; indeed, it can be used to recruit talented candidates by “listening” LinkedIn posts and tags, analyse market trends, competitors and to anticipate the future demand.

How can a company start social media listening activities? Companies can use ad hoc tools specialised in capturing and interpreting online conversation distinguish among the relevant ones and noisy ones.

In the following demo it was used Awario, a social listening tool that offer a subscription plan of $30 per month per user. Awario can listen to online conversations of multiple platforms as Twitter, Facebook, Reddit, YouTube, Instagram, forum news and web in general.

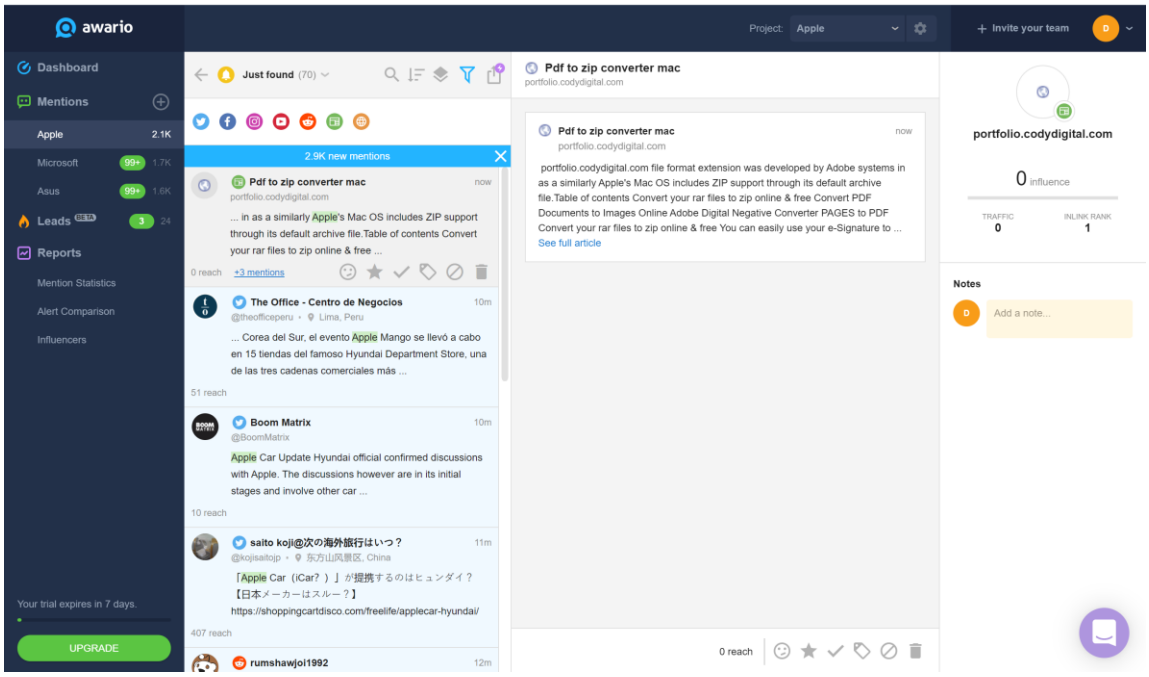


Figure 56 - Awario Dashboard

Through Awario, a company can follow what it is being said about its brand and product by filtering different sources. In this case, the aim was to discover what is being said about Apple products and some of its competitors. From the dashboard is possible to filter conversation, directly reply to tweets and visualize sentiment analysis, find the most relevant influencers, understand what the main languages are used in the discussion, country of origin and many more options. One interesting feature is the topic cloud, a cloud of words that it is possible to explore just by clicking on each world. For example, by investigating Apple’s topic cloud, it is possible to discover other different kinds of topics that we might dig into. Another example, by taking in consideration the University of Padua, the topic cloud displays many mentions about “Law”, “Padova”, “technology”, “family”, “dsea” keywords. Those keywords can be utilised

to dig into new opportunities and find relationship between your current product and what your customer are asking for.

Figure 57 - University of Padua topic cloud

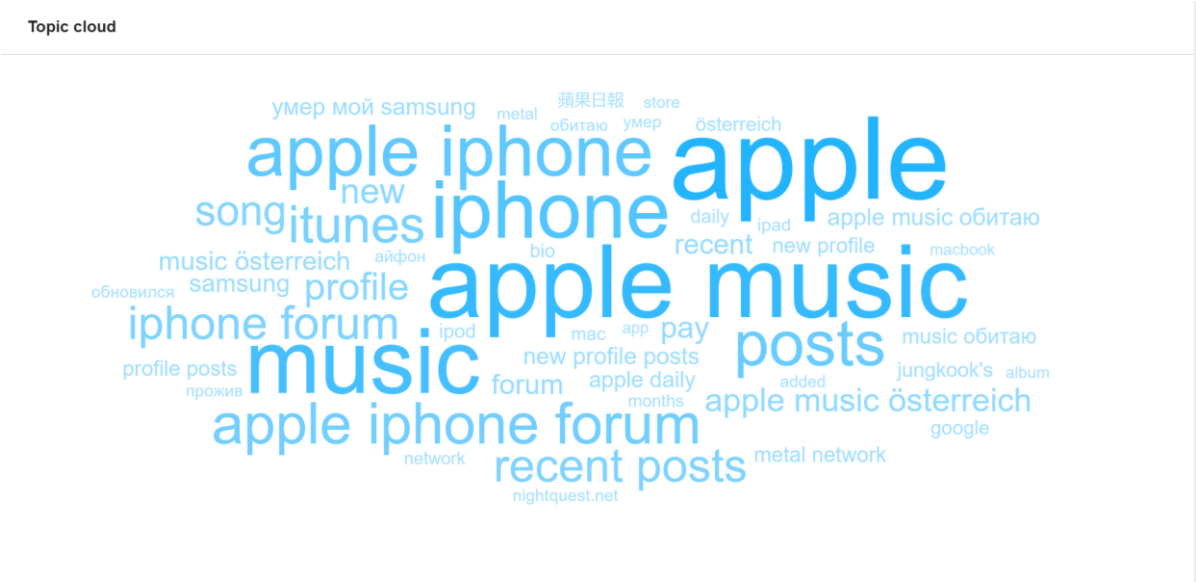


Figure 58 - Apple topic cloud

Process mining: improve your operations through data analysis

Business process management had been for long time few back-bunker issues, two to be precise (Spanyi, 2019):

1- The creation of "current state" processes: most of the time in business process reengineering, organizational are mainly focused on improving "to be" process and less the "as is", that is the description on how a business process is being performed today.

2- The lack of connections between business processes and an organization's enterprise information system.

The former problem can be address to a mindset that is looking forward to the next level and not at the current state, where the company live. Understanding current process is critical to know whether it is worth investing in improvements to avoid performance problems and the degree of variation among the process itself. Surprisingly, most of the company skip this stage, finding shortcuts to it or paying external consultant to analyse the current state processes ("as is"). The latter problem is linked to a lack of connection between enterprise systems and process oriented, not interrogating themselves about the process is being executed from the information system, skipping important information about process' performances day to day.

Process Mining tools have the aim to solve those problems, during the latest years this technology has become more and more popular, indeed, Gartner published a market guide for process mining in 2018 including many use cases. Process Mining software helps to capture information from enterprise transaction system and provide detailed (data-driven) information about how key processes are performing, creating logs that makes visible how computer-mediated work is really happening, from the receiving of the order, delivery and when payment is made. Process analytics create key performance indicators for the process, which enables a company to focus on the priority steps to improve. The main idea is to make visible in an easy way the full process to detect the root causes of variation. Organizations that have been digitalized their process and where there is still some unstructured will exploit the best value from Process Mining.

Process Mining in action: optimizing a purchasing process

To compute Process Mining simulation, it is been used *Disco*, a process mining tool distributed by Fluxicon (www.fluxicon.com/disco/). In the simulation, we are going to play the role of a manager who is accounted for the purchasing process in an organization, the goal is to make the process more efficient. The steps are the following:

- once the request is being submitted it has to be approved by the manager.
- the request is forwarded to the purchasing department where an agent has to choose the best option and place the order with the supplier.
- the supplier sends an invoice that will be paid by the financial department.

Figure 59 – Purchase process (Leoni, 2020)

We suppose that the whole process is being recorded by an Enterprise Resource Planning (ERP) system.

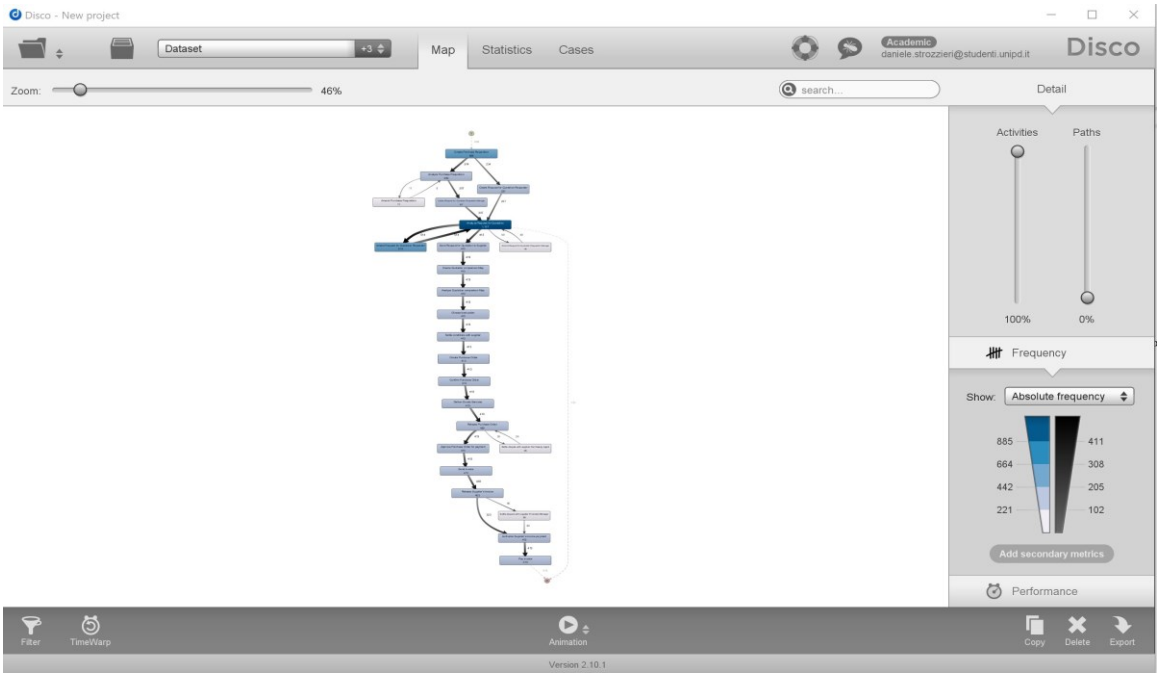


Figure 60 – process breakdown

In figure 60, it is possible to observe, once we loaded our database, that the breakdown of the entire process follows a waterfall scheme; the grey circle at the top (with a triangle) indicates the beginning of the process while the grey circle at the bottom (with a square), indicates the end of the process. Since the beginning, we can split the model in two parts: the bottom part is just a sequence of activities while the one at the top looks more interesting.

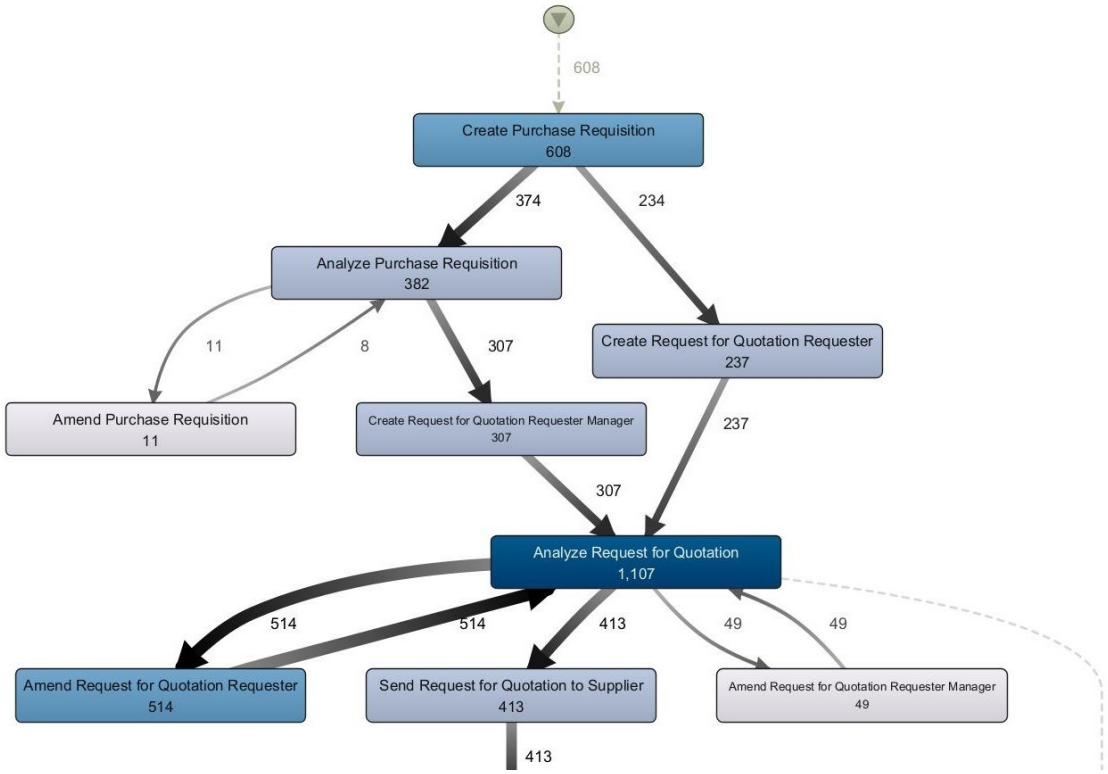


Figure 61 – top side of the process

In the top side we can notice the stream of initial process in a purchase situation, our attention will focalize on blue rectangles, the bluer they are, the higher the frequency request is. Let's focus on "Amend Request for Quotation", with 514 requests (over 608 purchase requisition) it is well highlighted by the software. According to the process owner, this activity is supposed to be used only in exceptional situation, that is, when a change is being made to an existing request. In a normal situation, "Amend Purchase Requisition" should be the one with the high number or requests, sign of efficiency, of normal status, while here is the exception. This is a clear source of inefficiency; indeed, amendments can really slow down the applications. If we want to visualize the main process behaviour, we can easily visualize it moving the "Activities" slider to 0%. Once, we become more aware about our process path, we can animate it through the "Animation" button to check possible bottlenecks in our process.

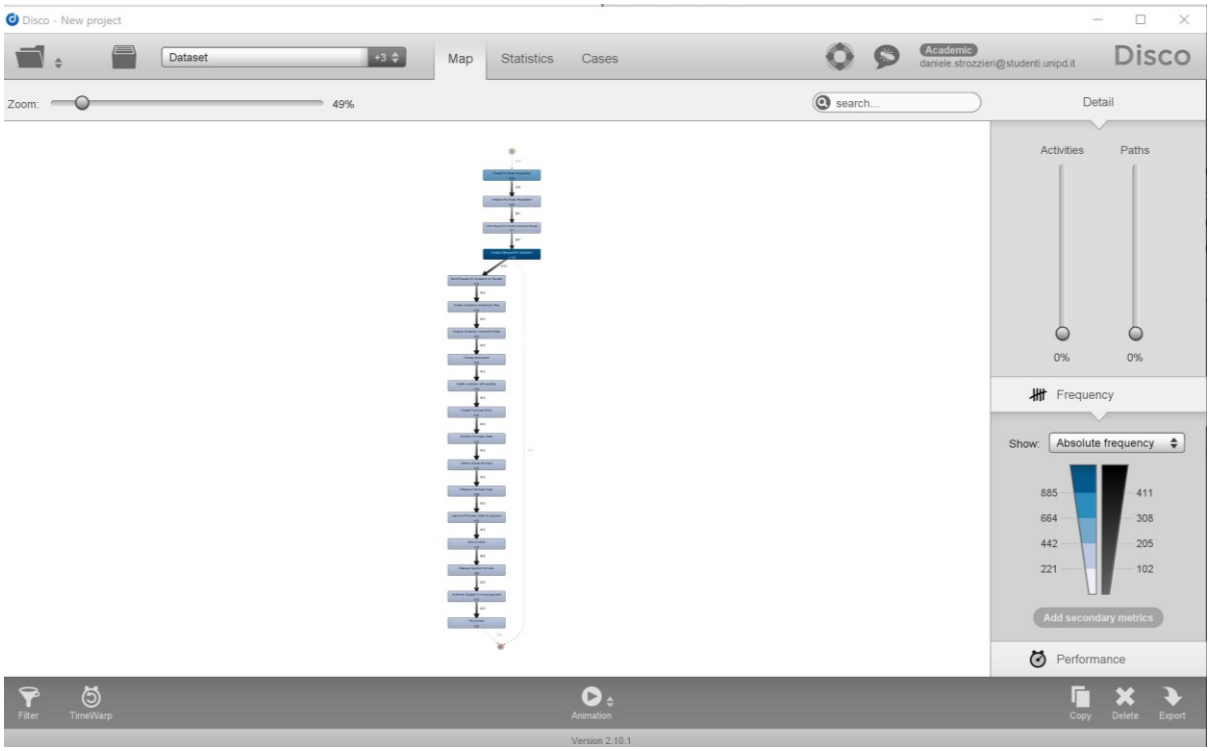


Figure 62 – main process behaviour

Breaking down our dataset, we are able to see the main processes, starting from the top: create purchase requisition, analyse purchase requisition, create request for quotation requester manager, analyse request for quotation, send request for quotation to supplier, create quotation comparison map, analyse quotation comparison map, choose best option, settle conditions with supplier, create purchase order, deliver goods services, release purchase order, approve purchase order for payment, send invoice, release supplier's invoice, authorize supplier's invoice payment, pay invoice.

Next step let's animate the process (Figure 63), at this point a new window will pop up and there it is possible to watch and analyse time by time how our process is performing, which progress is working more, which should be optimized, and witch process is working less. Is it necessary? Is it normal? Can we smooth it?

In this case, we can visualize and bring the discovered bottleneck to the attention of everyone inside the company since there are no number and strange tables, but only a simple video that can be export in few clicks. In this case a possible bottleneck is present between "Amend request for quotation" and "Analyse request for quotation", underlying another time the anomaly of that unnatural path.

After investigating the process map, we should have a clearer awareness of what our process looks like "as is", now, let's investigate the Statistics tab (figure 64).

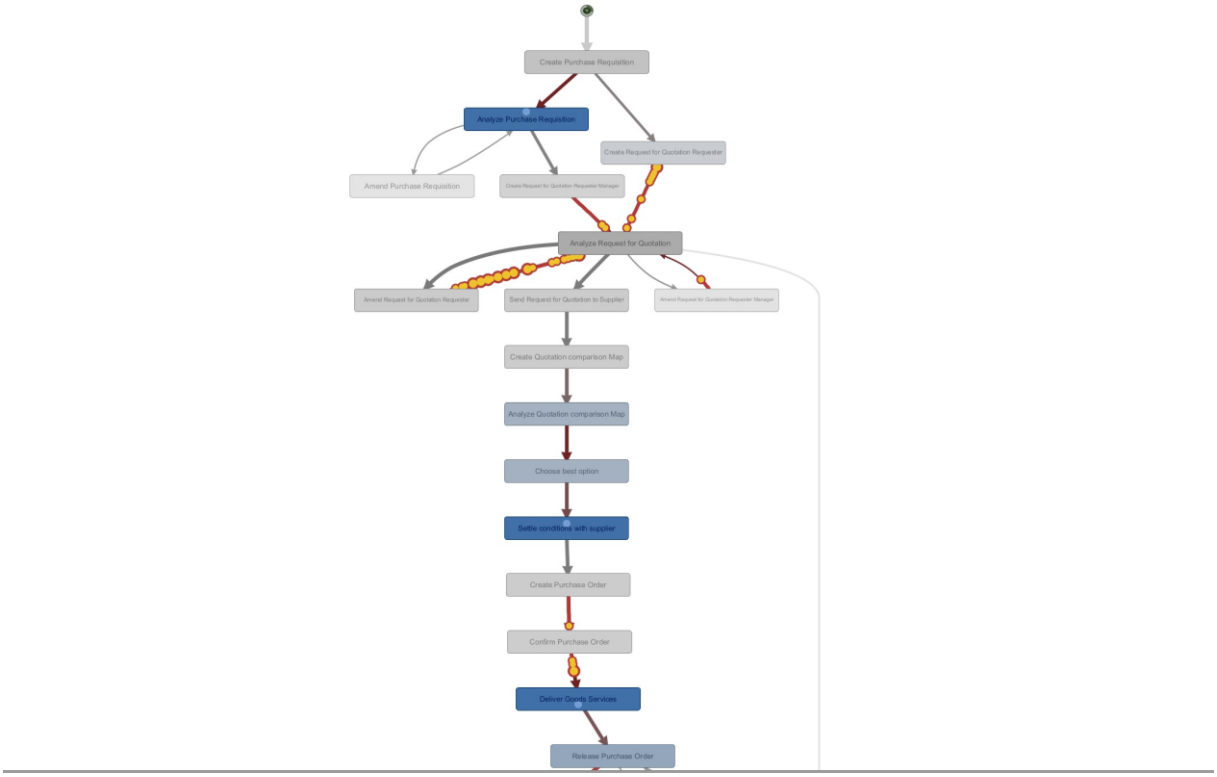


Figure 63 – animated process with bottleneck individuation

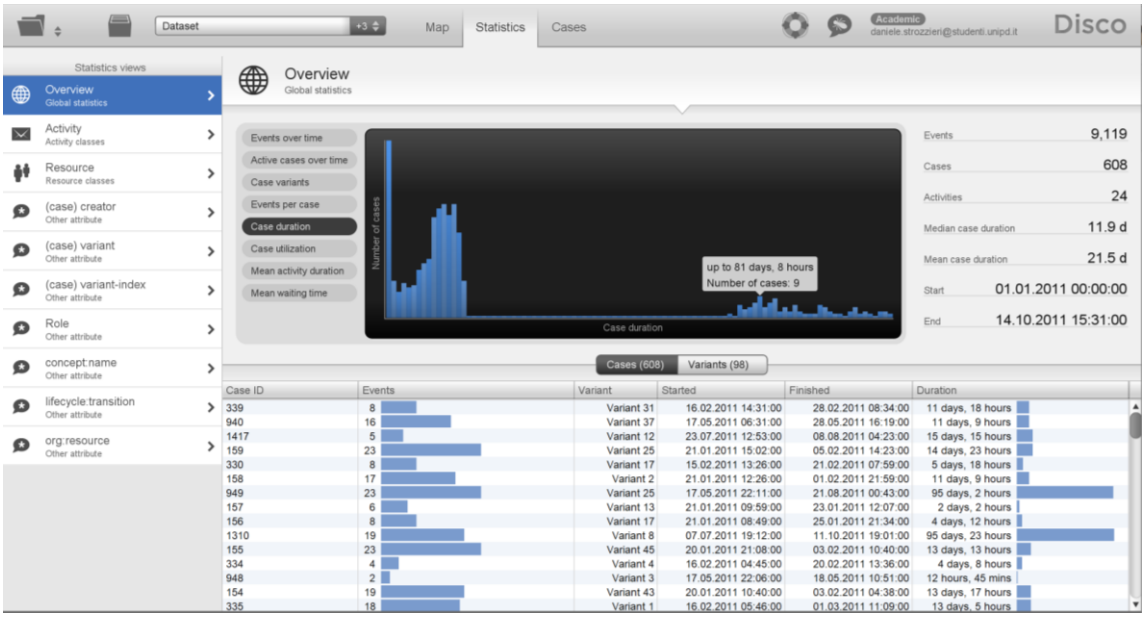


Figure 64 – statistics tab view

Despite it is a small dataset, on the right side we can find some statistics about our dataset: 9119 events (rows in the dataset), 608 events (purchase orders), the number of activities, median case duration and start and end of the dataset period, in this case it last 10 months and 15 days. As manager, we received complaints about the throughput time, to find the possible causes, the manager might analyse the Cases duration tab (below the histogram). In the histogram, moving

the mouse over it, we can clearly watch how most of the cases are completed within 16/17 days in total, on the other hand there are quite few of them that takes more of 80/90 days! It looks a serious problem that takes the throughput time in the process longer and longer. As manager we want to investigate which case is taking so long: let's switch visual by clicking the "Cases" tab (figure 65).

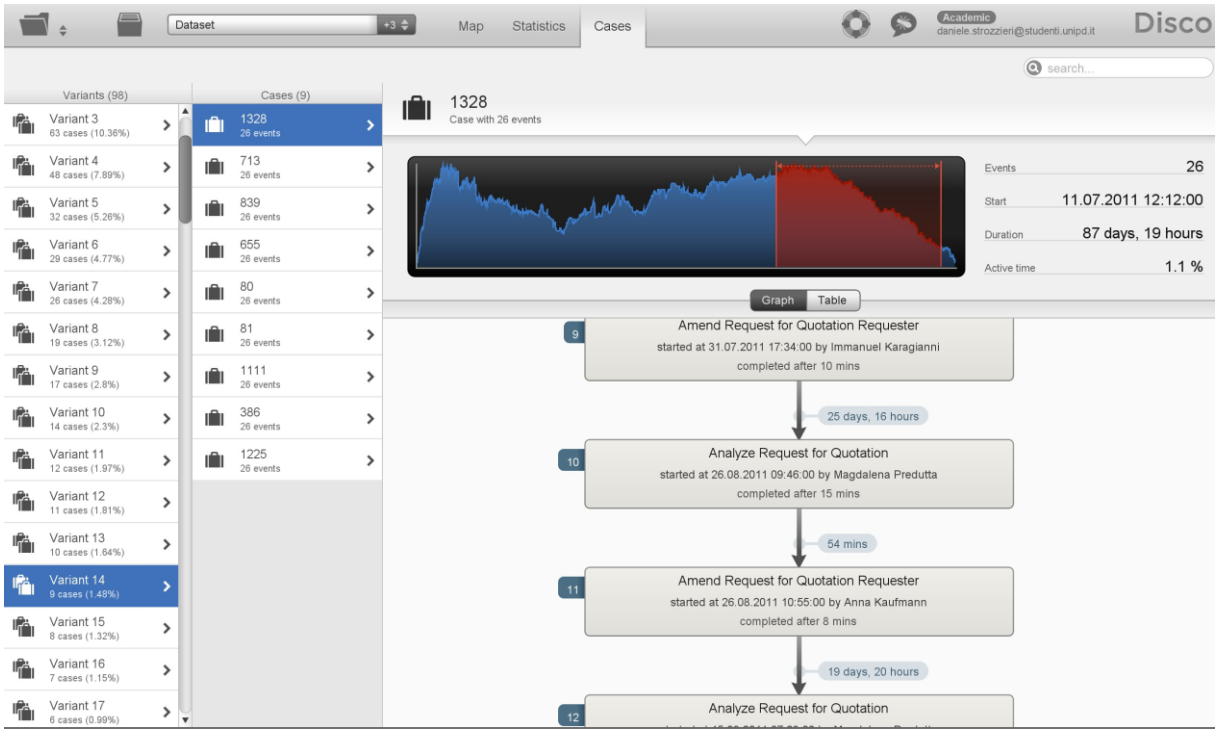


Figure 65 – Cases tab view

In the "Case" tab it is possible to find the history of an individual case; on the right it is possible to explore case's variants. A variant is a sequence of steps, from the beginning to the end. If two cases have in common the same path the software group them in the same variant. Usually by observing the case tab, a manager can get understanding about how the process usually perform. Often the main scenarios cover the 60-80% of the process.
By analysing the statistical tab, we noticed that the "Variant 14" (but also variant 8 and 25) have a high lead time, more than 80 days, almost 96 in variant 8 (that covers the 5% of cases). In this case, analysing Variant 14, through the graph tab, we discover that to pass from "Amend Request for quotation" to Analyse Request" took 25 days and 18 hours! Similar results could be found in the other variants.

Now, let's investigate the "Variant 3", equal to the 10,36% of the cases, it is possible to notice that this time the process is composed by only two cases: "Create purchase requisition" and "Analyse purchase requisition", the process has been stopped after only two steps, it ended

since the request was rejected. This specific event was already highlighted in the Map tab with a dot line at the right of "Analyse Purchase Requisition".

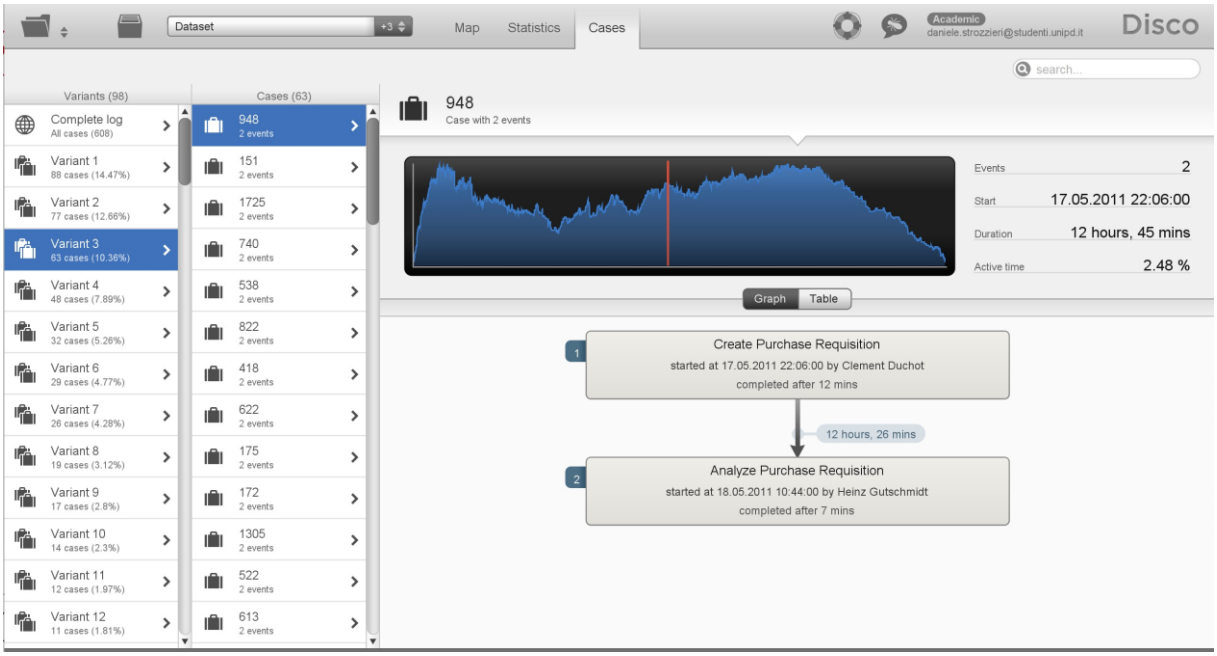


Figure 66 – Variant 3 view

To understand why some cases are taking so long, we will use Performance' filters by adding them through the filter list on the left down corner in Disco (Figure 67). We will filter those process that takes longer than usual cases (21 days): Disco show us that around 15% of all cases in the dataset fall outside of the service level target for this process. By applying the filter is clear that 92 cases have remained, where activity "Amend Request for Quotation Requester" happens 269 cases, around 2,92 times per case, much higher compared to the general case where the value is about 0,84 times per case (514 times out of 608 cases).

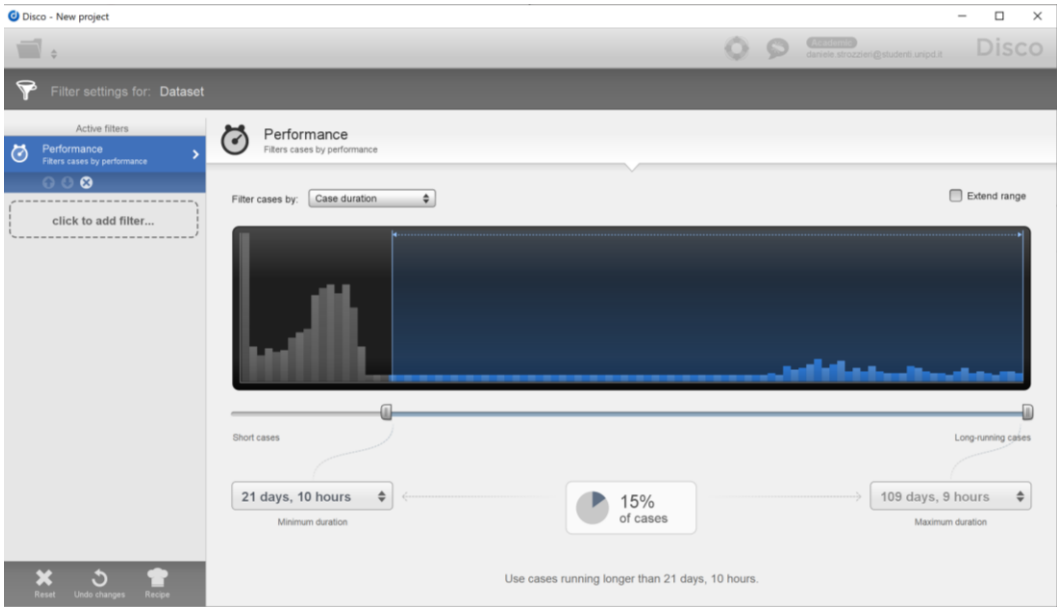


Figure 67 – Disco filters' tab

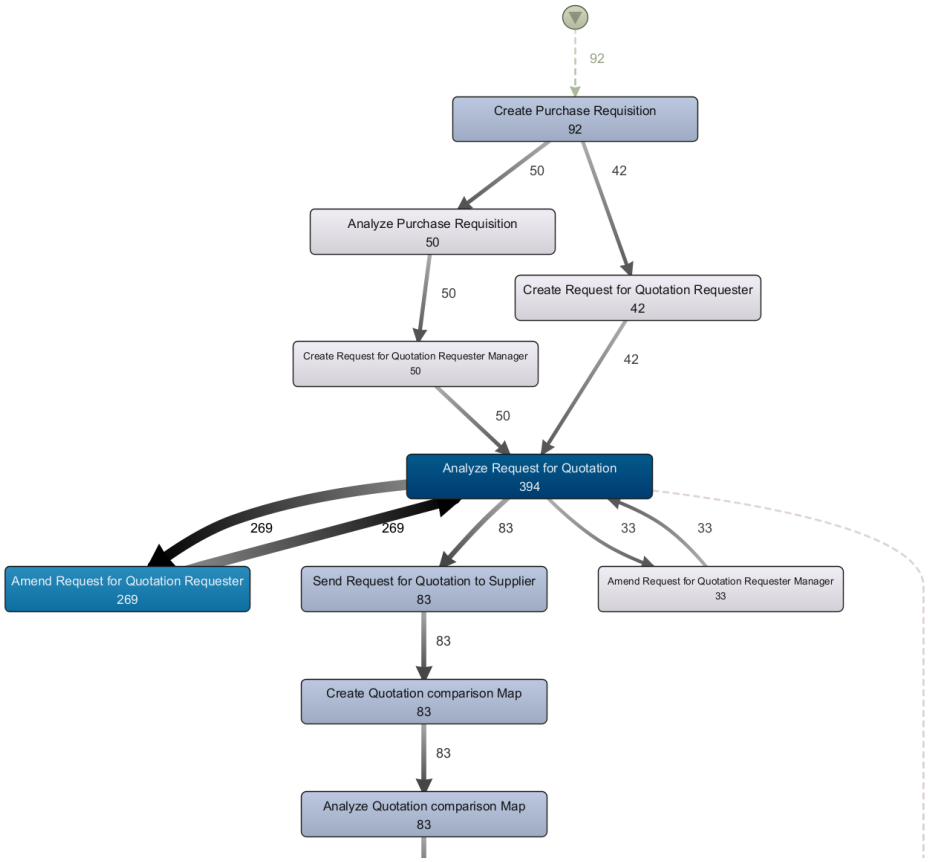


Figure 68 – process breakdown filtered.

Lastly, let's consider the Time Analysis tool. First, it is possible to analyse it by looking at the activity duration and the time between the activities. If we go back to the Map tab and switch from the frequency view to the performances vie (on the right bottom side) and select "Median duration", we can see the average duration of activities and the average time between activities. The activities are usually of small duration (around a couple of hours). On the other hand, many delays are observed between the activities related to the amendment of requests, especially the Amend Request for Quotation, which we have observed to be relatively frequent.

Looking at the bottom part of the process (figure 69), it is possible clearly notice the red rectangle among the yellow one, "Deliver Goods Services" that takes long time to be executed, more than one day. We can say that this timing is quite normal, since this kind of process requires time, it is difficult to improve.

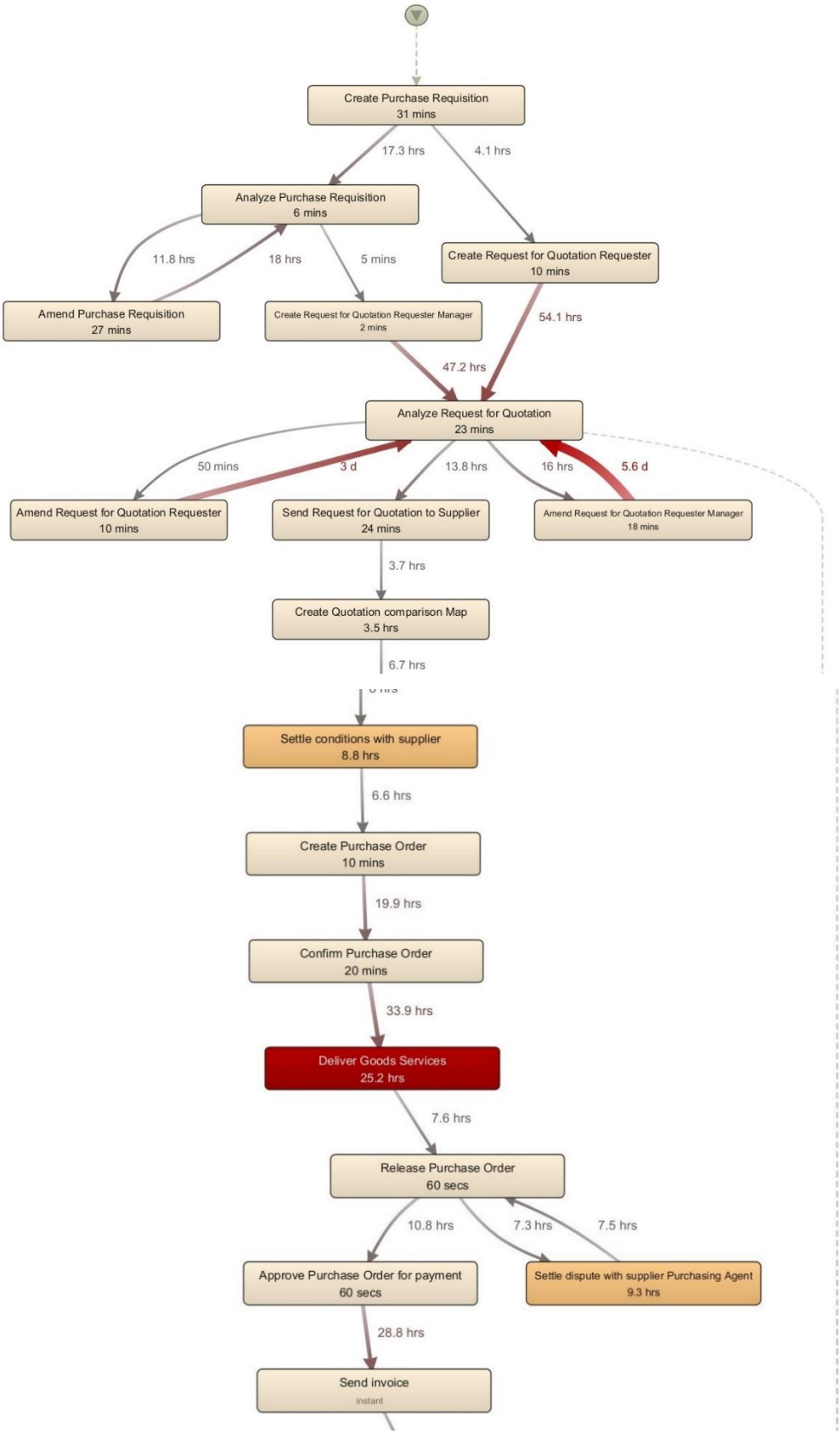


Figure 69 – Disco performance analysis

Final considerations

The inefficiencies of this process are related to frequent amendments that since the beginning of our analysis the process mining tool highlights as problematic. At this point further internal investigation should be carried on and trying to avoid the problem since the beginning. After modifying some internal processes, the manager had better check again through the process mining program if the problems have been erased and/or there is margin of improvements.

Conclusions

Big Data & analysis is a powerful business tool, typical of digital companies, that has the aim of exploit value creation from company' data, both internally and externally. Through data analytics, managers can produce powerful insights that improve decision making, customer relationship, operations and even use them as a real asset to generate revenue. To be implemented, an upstream strategy is required, where the whole organization must adopt a data-driven culture. To enable a BDA strategy, it is necessary to go through two fundamental processes: Capability Building Process and the Capability Realization Process.

In the first step, it is necessary to build an AI infrastructure, while in the second one analytics skills and tools are fundamental to transform raw data in powerful insights. The third step is represented by human resources, through absorption capacity, they able to adapt quickly and effectively to a new change, allowing them to embrace new analytics tools as an opportunity rather than an evil. Nonetheless, as in every change, problems are always behind the corner: communication, quality, data integration, security, analytics, and value assessment are the main hurdles. An external and more complicated hurdle is the recruitment of human capital, rare to find and distributed in few BDA districts, mostly in big urban areas. Despite BDA is a new trend, that is finding opposition by big companies not related to bank and insurance sector, small and medium enterprises should start to embrace this new technology to exploit their full potential. Centre Italy SMEs are still at the "before-early" stage in analytics capabilities; despite they are becoming familiar with the potential of BDA, they do not know the way to go for generating value through data. It has been proved in the previous chapter that implementing BDA tools nowadays are not as difficult as it seems, new tools are affordable and not as difficult to use. Technologies as business intelligence, social media listening, and process mining can truly transform the business of a company, from unsuccessful to successful, improving every single aspect of the company and finding new solutions, ideas, and competences. Maybe, new economical sustains destinated to digital transformation will trigger the data trend, and for that time will be important that top managers, and moreover, external consultants, help companies to pass through the digital process. It is a change that soon or later every company will deal with, the choices are two: abandon or ride the trend. In a globalize world, there is not anymore, the time to look what happen inside the regional borders, companies must break those borders and look in the face the digital change, with the hope to rebuild a better and world.

Bibliography

(2020, November 3). Retrieved from CorporateInformation: https://www.corporateinformation.com/Top-100.aspx?topcase=b

Arthur Thompson, A. J. (2015). *Crafting and Executing Strategy: Concepts and Readings* (20 ed.). McGraw-Hill Education.

Baratta, L. (2020, April 20). Retrieved from Linkiesta: https://www.linkiesta.it/2020/04/sito-inps-600-euro-audizione-tridico-hacker-bufala/

Bauer, E. (2020, February 14). *Marketo Blog*. Retrieved from https://blog.marketo.com/2017/11/sales-velocity-critical-sales-metric-team-probably-isnt-measuring.html

Brahm, C. (2019, October 3). *Solving the Advanced Analytics Talent Problem*. Retrieved from Bain & Company: https://www.bain.com/insights/solving-the-advanced-analytics-talent-problem-mit-sloan/

Camere, G. S. (2013). *Unionecamere.gov.it*. Retrieved from Geo Smart Camere: https://www.unioncamere.gov.it/Atlante_2015/regioni/abruzzo/popolazione.html

Cerved. (2018). *Rapporto Cerved PMI 2018.* Retrieved from https://know.cerved.com/wp-content/uploads/2018/11/PMI-2018-intero.pdf

Damanpour, F. (1991). Organizational Innovation: a meta-analysis of effects of determinants and moderators. *Academy of management journal, 34*(3), 555-590.

Dutton, C. (2019). *Up & Running with Power BI Desktop.*

European Commission. (2020). Digital Economy and Society Index 2020 Italy. *DESI*.

Fichman, R. (2004). Real options and IT platform adoption: Implications for theory and practice. *Information Systems Research*, 132–154.

Fisher, D., DeLine, R., Czerwinski, M., & and Drucker, S. (2012). Interactions with big data analytics. *Interactions, 19*(3), 50–59.

Fleurence, R., Curtis, L., Califf, R., Platt, R., Selby, J., & and Brown, J. (2014). Launching PCORnet, a national patient-centered clinical research network. *Journal of the American Medical Informatics Association, 21*(4), 578–582.

George, A. Z. (2002). Absorptive capacity: A review, reconceptualization, and extension. *Academy of management review, 27*(2), 185-203.

Gopalkrishnan, V., Steier, D., Lewis, H., & and Guszcza, J. (2012). Big data, big business: Bridging the gap. *BigMine '12: Proceedings of the 1st International Workshop on Big Data, Streams and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications*, (pp. 7-11). Beijing, New York: ACM.

Heudecker, J. H. (2016, september 19). *Survey Analysis: Big Data Investments Begin Tapering in 2016*. Retrieved from Gartner: https://www.gartner.com/en/documents/3446724

Infodata. (2019, July 10). *Infodata | Il sole 24 ore*. Retrieved from https://www.infodata.ilsole24ore.com/2019/07/10/40229/?refresh_ce=1

ISACA. (2014). *Generating value from big data analytics*. Retrieved from isaca.org: . http://www.isaca.org/ Knowledge-Center/Research/Documents/Generating-Value-from-Big-Data-Analytics_whp_

ISTAT. (2020). *CENSIMENTO PERMANENTE DELLE IMPRESE 2019.* Retrieved from https://www.istat.it/it/archivio/238337

Istat. (2020, February 7). *Istat*. Retrieved from https://www.istat.it/it/archivio/238337

ISTAT, M. |. (2019, october 18). *Startup innovative: tutti i dati al 30 settembre*. Retrieved from https://www.mise.gov.it/index.php/it/per-i-media/notizie/2040345-startup-innovative-tutti-i-dati-al-30-settembre

John E. Hanke, D. W. (2014). *Business Forecasting.* Pearson Education Limited.

Kenneth C. Lauron, J. P. (2010). *Management dei sistemi informativi.* Pearson - Prentice Hall.

Kune, R., Konugurthi, P., Agarwal, A., Chillarige, R., & and Buyya, R. (2016). The anatomy of big data computing. *Software: Practice and Experience, 46*(1), 79–105.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M., & and Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT Sloan Management Review, 52*(2), 21–32.

Leoni, M. d. (2020). Hands-on Process Mining.

Levinthal, W. C. (1990). Absorptive Capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*(35), 128-152.

LinkedIn. (2019). LinkedIn AI talent in the European labour market 2019.

Marr, B. (2017). *Data Strategy - How to Profit from a world of Big Data, Analytics and the Internet of Things.* Kogan Page.

McAfee, A. a. (2012). Big data: The management revolution. *Harvard Business Review*, 60–68.

Morabito, C. F. (2008). IS integration and business performance: The mediation effect of organizational absorptive capacity in SMEs. *Journal of Information Technology*(23), 297-312.

Nicola, S., Ferreira, E., & and Ferreira, J. (2014). A quantitative model for decomposing and assessing the value for the customer. *Journal of Innovation Management, 2*(1), 104–138.

Nielsen. (2015). *Nielsen's global trust in advertising report.*

Pettigrew, A. a. (1991). *Managing Change for Competitive Success.* Oxford: Basil Blackwell.

Press, G. (2016, January 6). *6 predictions for big data analytics and cognitive computing in 2016*. Retrieved from Bloomberg: https://www.bloomberg.com/professional/blog/6-predictions-for-big-data-analytics-and-cognitive-computing-in-2016/

Qiao, Z., Zhang, X., Zhou, M., & Wang, G. a. (2017). A domain oriented LDA model for mining product defects from online customer reviews. *50th Hawaii International Conference on Systems Sciences*, (pp. 1821–1830). Waikoloa.

Spanyi, T. H. (2019). What Process Mining Is, and why companies should do it. *Harvard business review*.

STAMFORD, C. (2016, January 26). *Gartner Estimates That 90 Percent of Large Organizations Will Have a Chief Data Officer by 2019*. Retrieved from Gartner: https://www.gartner.com/en/newsroom/press-releases/2016-01-26-gartner-estimates-that-90-percent-of-large-organizations-will-have-a-chief-data-officer-by-2019

Symbola. (2019, september 23). Retrieved from Symbola: https://www.symbola.net/approfondimento/23867-2/

Varun Grover, R. H.-P. (2018). Creating Strategic Business Value from Big Data Analytics: A Research Framework. *Journal of Management Information Systems, 35*(2), 388-423.

Vera-Baquero, A., Colomo-Palacios, R., & and Molloy, O. (2013). Business process analytics using a big data approach. *IT Professional, 15*(6), 29–35.

Wakefield, J. (2015, September 13). *Intelligent Machines: The jobs robots will steal first*. Retrieved from bbc.com: https://www.bbc.com/news/technology-33327659

WEF. (2020). Future of Jobs Survey.

WEF. (2020). The state of the connected world .

Wegener, R. (2013, september 17). *The value of Big Data: How analytics differentiates winners*. Retrieved from Bain & Company: https://www.bain.com/insights/the-value-of-big-data

Weill, P. (1992). The Relationship Betweeninvestment In Informationtechnology And Firmperformance: A Study Ofthe Valve Manufacturing Sector. Center for Information Systems ResearchSloan School of ManagementMassachusetts Institute of Technology.