Universiteit Leiden

ICT in Business and the Public Sector

The impact of data analytics on the effectiveness of SMEs strategic decision-making processes

Name: Mark Rodenburg
Student-no: s2276925

Date: 13/08/2020

1st supervisor: Prof. Dr. Stefan Pickl
2nd supervisor: Prof. Dr. Wolfgang Bein

MASTER'S THESIS

Leiden Institute of Advanced Computer Science (LIACS)
Leiden University
Niels Bohrweg 1
2333 CA Leiden
The Netherlands

Foreword

This thesis is my final work to complete the master ICT in Business: “*The impact of data analytics on the effectiveness of SMEs strategic decision-making processes*”. A result of seven months extensive research in the midst of the coronavirus outbreak. I started this project because I experienced, as a part-time consultant at a small IT-service provider, how SMEs struggle to find meaningful ways to implement data analytical solutions. And because of my interest in strategic decision-making, I decided to investigate how those concepts could interrelate.

I would like to thank my supervisor Prof. Stefan Pickl for the support and freedom he gave me while doing research. Unfortunately, I could not visit Munich because of the corona outbreak. However, I still enjoyed collaborating together. Furthermore, I would like to thank Prof. Wolfgang Bein for being my second supervisor.

I really enjoyed the two-year programme at Leiden University and want to thank all employees of the LIACS department.

Mark Rodenburg

Abstract

Small and medium-sized enterprises (SMEs) have for a long time been challenged by large competitors around the globe. However, this pressure increases now even more because large companies have so far been more able to utilize data analytics for the identification of emerging opportunities and for the improvement of their internal infrastructures. Literature suggests that SMEs tend to lag behind, because they have less assets available to invest in the required human resources, R&D and infrastructure (Michalkova & Bianchini, 2019). Besides that, there is also a lack of scientific literature available tailored to their limited amount of resources to spend. This study has attempted to fill this literature gap, by investigating how SMEs could take advantage of data analytics technology in order to improve the effectiveness of their strategic decision-making processes.

The research question was as follows: *"How can data analytics improve the effectiveness of SMEs strategic board decision-making processes?"*.

The Grounded Theory has been applied in order to develop two new theories based on the researcher's interpretations of existing literature in order to:

1. Explain how data analytics could be applied by SMEs despite their limited amount of resources to spend.
2. Explain how data analytics could increase the effectiveness of SMEs strategic decision-making processes.

The theories have been able to confirm that SMEs should be able to take advantage of data analytics solutions for their strategic decision-making processes (i.e. portfolio and value-chain performance analysis), because the emerging challenges can be avoided by SMEs by applying a data analytics strategy (Awwad, Kulkarni, Bapna, & Marathe, 2018), by focusing on simplicity and by outsourcing (uncertain) solution development (Collinson & Jay, 2012). Because off-the-shelf Business Intelligence software and (SaaS) cloud solutions allow SMEs to analyse a wide range of data sources (e.g. web logs, telephony logs, transaction logs, sensor data, social media, e-mail databases), and are offered for free or apply a subscription based model, it will be possible to avoid the risk of facing a low return on investment.

These data analytics tools should be able to increase the effectiveness of SMEs strategic decision-making processes because they are able to visualize trends in patterns and behaviour in large datasets, which is expected to eliminate decision-makers biases by opposing them during SWOT-analysis. Additionally, they will be able to provide SMEs with (real-time) feedback on their company's business strategy by visualizing business performance, which will allow them to learn how their decisions are affecting business performance and take corrective actions in response to changing circumstances.

Table of Contents

1. Introduction

Motivation

Data analytics promises to help companies with decision-making by providing them answers based on large amounts of historical data (Jeble, Patil, & Kumari, 2018). The technology can be used to visualize patterns, correlations and associations in data, and would be helpful in various fields. New insights provided by data analytics tools will, for instance, be able to help improve a company's sales performance by identifying customer needs and market segments. Data analytics would also help companies with inventory planning, demand forecasting, capacity planning and procurement planning. Industry leading enterprises, such as Google, Walmart, Amazon and Netflix, are already mastering data analytics and apply the technology as a tool to perform simulations, make predictions or just to describe data.

Research suggests that organizations who strongly agreed that the use of data analytics differentiates them within their industry, are twice as likely to be a top performer (LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011). They approach business operations differently: they apply data analytics for future strategy- and day-to-day decision-making. Organizations applying data analytics tend to be more able to turn challenges into opportunities and tend to make rigorous decisions twice as often in comparison to low performers. In addition to that, is data analytics also increasing the 'clock speed' of organisations (Galbraith, 2014), as it is being used to gather real-time feedback on business performance. These developments have been instructive for other firms to implement data analytics tools within their organisations in order to keep up with those competitors.

However, firms struggle to adopt data analytics in strategic decision-making processes (Merendino et al., 2018). Becoming a data-driven organisation requires a cultural change, as Big Data analytics shifts the power away from board members who had been making decisions based on their experience and judgements for decades. Furthermore, data analytics newcomers would have to be integrated in the company's established decision-making process (Galbraith, 2014). This and a lack of cognitive capabilities among board members to be able to analyse the results of Big Data analytics form barriers for adoption (Merendino et al., 2018). Ensuring senior management involvement and creating a data-driven culture may be the most significant challenges for organisations, however, firms also face other challenges such as the design of an effective data architecture and technology architecture and a scarcity among elite data scientists (Henke et al, 2016). Furthermore, organisations may also be challenged by a lack of access to meaningful data, a lack of interoperability, cybersecurity, privacy regulations and complex ecosystems.

Problem statement

Large organizations are likely to have sufficient means to overcome the obstacles and thus take advantage of the opportunities offered by data analytics. However, a recent mapping of the usage of data analytics shows how on average only 10% of the small business (10-49 persons employed) in Europe use Big Data analytics compared to an average of 33% of the large companies (250 persons employed or more) (Michalkova & Bianchini, 2019). This gap is even bigger in the Netherlands (18% vs 53%) and Belgium (17% vs 55%). Medium-sized businesses are behind as well; only 19% have adopted it. According to the Organisation of Economic Co-Operation and Development, SMEs lag behind because of a lack of investment in complementary assets, such as human resources, R&D, organisational changes and process innovation (Michalkova & Bianchini, 2019). This backlog with regard to the use of data analytics is concerning, as SMEs form the backbone of any country's economy, due to their flexibility and innovativeness (Bokdam, 2019). Especially if a company's success in the future is going to be determined by its access to big data sources, SMEs could have an additional disadvantage in comparison with larger enterprises.

The predominantly larger companies who are able to take advantage of data analytics solutions could obtain three competitive advantages (Davenport & Dyché, 2013); (1) reduced costs, (2) lower time of computing tasks and (3) data-driven product and service offerings. UPS for instance, started to track more than 46.000 vehicles to be able to analyze and reconfigure driver's pickups and drop-offs in real-time and by doing so, they have been able to save more than 8.4 billion gallons of fuel, 85 million of miles for daily routes, one mile daily per driver and more than 30 million dollar. And Macy's applied data analytics for price optimization, because it had been a time-consuming activity for a shop with 73 million product offerings. LinkedIn's 'People You May Know', 'Groups You May Like', 'You May be Interested In' and 'Who Viewed' sections has already been able to bring millions of customers to LinkedIn. Streaming platform Netflix has been using big data to help create proprietary content and Verizon sells mobile location data to measure the effectiveness of outdoor media campaigns. And those who have been using Google's search engine could notice how they are served by ads which are targeted to their personal interest.

The introduction of big data analytics is also increasing the wedge between SMEs and large enterprises, because it will become harder for SMEs to catch up with large competitors, when they have a (potentially disruptive) tool to target their resources on process improvement, tackling key business challenges and new business development (Sen, Ozturk, & Vayvay, 2016). Big data analytics will allow companies to continuously generate new knowledge and increase the strength of a company's competitive advantage because the knowledge gathered will add up over time. Meanwhile, 99.8% of all enterprises are SMEs, they employ 66.9% of the people among all enterprises and are responsible for 57.8% of the value of the European economy. Therefore it is very important that they are able to take advantage of big data analytics, due to their share in the value of the economy. If SMEs would be able to identify new opportunities by using data analytics, it would strengthen the European economy and generate a larger GDP per capita (Sen, Ozturk, & Vayvay, 2016). However, if SMEs would fail to take advantage of data analytics, they could lose customers over competitors who may be located in the US or Asia (e.g. Amazon), which could lower their contribution to the European economy and increase unemployment rate.

Large enterprises who have invested large amounts of resources in data analytics do not necessarily have to be a threat to SMEs, because SMEs are much more flexible than they are (Ann & Perks, 2011) and should be able to catch up by reinventing themselves. Smaller businesses are more flexible than large enterprises because they deal with fewer levels of management, a lower degree of formalization, deal with less resistance to change and are easier to control because top-management is more visible and closer to the point of delivery. These conditions allow them to respond faster to environmental changes. In addition to that, SMEs do also have to deal less with legacy infrastructure in comparison with large enterprises (Sen, Ozturk, & Vayvay, 2016), which makes it easier for them to develop new businesses in order to take advantage of emerging opportunities. SMEs should be more able to focus on new opportunities, while large enterprises are more likely to make incremental improvements within their existing infrastructure.

However, if SMEs considered data analytics as a disruptive technology, they would probably be willing to invest in the required complementary assets, despite their limited amount of resources available, as long as it’s profitable and affordable. Research indicates that there are a couple of additional factors contributing to the poor adoption rate of data analytics among SMEs in particular (Coleman et al., 2016). First of all, SMEs representatives have an extremely low understanding of big data analytics. On the one hand they don’t know whether they have appropriate data and on the other hand they don't know if investing in data analytics is going to bring the benefits claimed by data science enthusiasts. SMEs general management functions tend to be poorly covered because the majority of the staff are domain specialists, which results in a reduced awareness of new opportunities and trends. Furthermore, a lack of exemplary use cases corresponding to SMEs interests and financial barriers causes them so far to be reluctant to investigate data analytics tools.

Nonetheless, data analytics is being reported as a technology which will reshape day-to-day operations and strategic decision-making processes (LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011), and it is about time for those SMEs who lag behind to have a look at the technology as it has matured over the years. However, considering SMEs tend to have less access to financial resources, which makes them cautious about investments outside their business scope (Coleman et al., 2016), and the currently low rate of adoption, they need more guidance to be able to take advantage of this emerging technology. And especially because SMEs have a limited amount of resources to spend, it is important for them to find cost-effective ways to implement data analytics, to be able to maximize returns on investment and to justify their investments. In order to maximize the cost-effectiveness of the introduction of data analytics, it has to make an impact by transforming businesses. This means that data analytics should make a change on strategic level for more leverage. Therefore, SMEs should consider applying data analytics to improve the quality of their strategic decision-making processes, because board members have been making decisions based on gut-feeling and experience for decades (Merendino et al., 2018), which is susceptible to decision-makers biases and reported to be less effective. However, SMEs' limited amount of knowledge with regards to data analytics, a lack of available resources to spend on new technology and a lack of exemplary use cases has so far restrained adoption.

Research objective

The purpose of this research is to gain a deeper understanding of how SMEs could take advantage of data analytics to improve the effectiveness of their strategic decision-making processes. SMEs should be aware of this to avoid customers leaving for competitors who will become more able to determine customer’s needs and improvement potential within their company’s internal infrastructure. Investigating how SMEs could apply data analytics despite having a limited amount of resources to spend, should clarify how they could get started as well and close the gap between them and larger enterprises. And by focussing on the application of data analytics to increase the effectiveness of strategic decision-making processes in particular, SMEs will obtain an understanding of how data could help to transform their businesses, increase operating margins and/or company's growth.

Academic contribution

Current literature discusses the potential of data analytics in a general manner. However, SMEs tend to have, compared to larger enterprises, less resources to spend on R&D, human resources, organizational changes and process innovation. This is why SMEs should have access to customized literature tailored to their needs. Unfortunately, this has not been available for this specific research topic and this might explain why SMEs have so far been lagging behind. Therefore, explorative research has to be conducted to investigate how SMEs could take advantage of data analytics by applying the technology to support their strategic board decision-making processes. Here we may delve into the research question.

Research question

The formulation of the main research question is than as follows:

"How can data analytics improve the effectiveness of SMEs strategic board decision-making processes?"

Three hypotheses have emerged and are provided with a preliminary answer by developing new theory validated based on a wide range of existing literature, namely:

- **Hypothesis 1** - SMEs are able to make use of data analytics to provide information required for strategic decision-making processes (i.e. portfolio and value chain performance analysis) despite having a limited amount of resources to spend.
- **Hypothesis 2** - All SMEs should be able to gather (new) insights by using tools for data analytics, regardless of their organisational context.
- **Hypothesis 3** - Tools for data analytics will allow SMEs to improve the effectiveness of strategic decision-making processes by opposing decision-makers biases while performing portfolio- and value chain analysis.

Terms of definitions

This research brings three main topics together: data analytics, strategic board decision-making and how data analytics could improve the effectiveness of strategic board decision-making processes. Just to make sense of the research problem, these three terms will be defined briefly:

- **Effectiveness -** “the ability to be successful and produce the intended results” (Cambridge Academic Content Dictionary, n.d.)
- **Data analytics -** “the process of examining data sets in order to draw conclusions about the information they contain, increasingly with the aid of specialized systems and software“ (Rouse & Stedman, 2019).
- **Strategic board decision-making -** “non-routine, organization-wide resource allocation decisions that affect the long-term performance of an organization” (Nahum & Carmeli, 2019)

Both topics data analytics and strategic board decision-making (processes) have been extensively reviewed and discussed in chapter 2: Literature Review.

Subject Relevance

The importance of data analytics

Data analytics has the potential to become disruptive according to McKinsey (Henke et al, 2016). Industries assets are underutilized, industries deal with supply/demand mismatches, data is siloed and fragmented, decision-making is subject to human biases and the speed of decision-making is limited by human constraints (Figure 1). These six challenges could potentially be solved by data analytics. Besides that, data analytics could be applied to generate large amounts of personalized data, improve accuracy of predictions and can deliver insights for the improvement of a core business model.

McKinsey estimated data analytics to deliver a potential increase of net margin by 60% for US retail, 25% decrease in operating cost and an estimated 0.5% productivity growth for the European Public Sector (Henke et al, 2016). For manufactures, data analytics could lead to a 50% decrease in product development cost and a 30% increase in gross margin. Global research conducted by McKinsey in 2017, six years later, revealed how 70% percent of 503 executives report that data analytics have caused at least a moderate change within their industries competitive landscape (Gottlieb & Rifai, 2017). 50% of the respondents mention that they are facing new entrants launching new-data focused innovative solutions, undermining traditional business models. 36% of the respondents have witnessed firms improving their core businesses by applying data analytics. Another 36% of the respondents have seen companies extracting new insights from data. And 27% percent have seen traditional businesses introducing new data driven products and services.

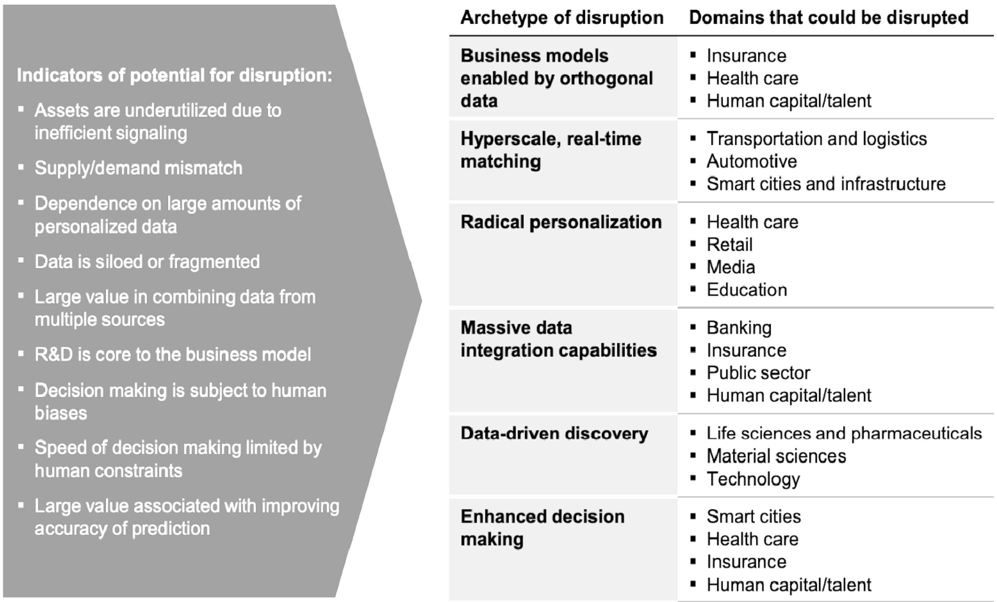


Figure 1: Indicators of potential for disruption (Henke et al., 2016)

Yet there is still a lot of value to be obtained, even at small scale (Hoerl & Redman, 2019). Three examples of small data opportunities: (1) reduce non-value adding work (e.g. dealing with errors), (2) reduce waiting time and (3) streamline handoffs. These issues can be identified with only small amounts of data sets, analysed by a small team of analysts. Fixing these types of issues are tightly focussed and accessible for any company. The profit of small data analytics projects with the duration up to several months, are ranging from $10,000 to $ 250,000 annually. Companies can expect to complete 20 projects with a team of 40 data scientists per year, delivering a large cumulative benefit.

Another case study carried out at the AGGORA group, a mid-sized company specialized in catering equipment solutions (Raj, Wong, & Beaumont, 2016), reported that they had been able to save $ 29,262 within their first six months of using descriptive data analytics software. Their Business Intelligence (BI) solution provided the following benefits: (1) cost savings ($ 4,202) by using data analytics for self-service reporting, (2) increased productivity, efficiency and revenue through new reports ($18,060) and (3) less time spend on report generation ($7,000). This case study also describes how data analytics software helps them to understand, evaluate and compare their current performance against their company's objectives (KPIs). And any user familiar with Excel Pivot tables and graphs would be able to use and benefit from data analytics software, as long as the data retrieval and analysis process is properly designed.

Relevance in strategic decision-making processes

To be able to understand the relevance of data analytics in strategic decision-making processes, Porter's value chain needs to be introduced. Porter's value chain (Figure 2) is a classic model used for strategic planning. The model shows how each company consists of primary and secondary activities in order to provide value for the customer. Value is the amount of money a buyer is willing to pay for the product or service delivered (McGee, 2014). Each activity in the value chain involves cost for 'purchased inputs' such as human resources, services and other goods. And every activity delivers information needed for the organization to be able to understand what's going on in their business.

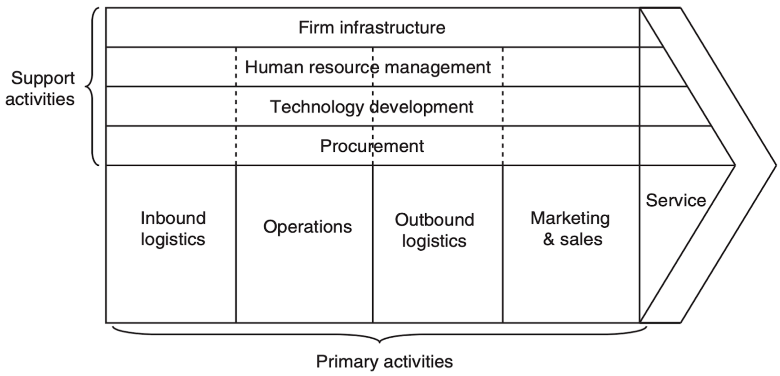


Figure 2: The value chain of Porter (McGee, 2014)

Analysing the data provided by each activity in the value chain, will help to find the most cost-effective ways to lower costs or add value to the offered product or services. This is important, because the value of a product or service offered minus the costs of each activity determines the operating margin for the companies involved.

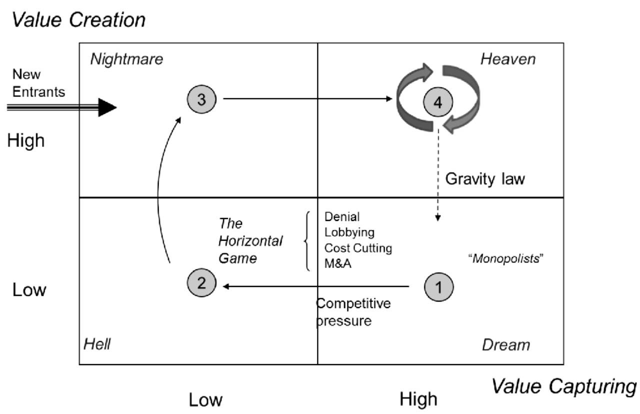


Figure 3: The Value Creation-Value Capturing framework

(Verdin & Tackx, 2015)

Adding value to a product or service or reducing the costs involved is not only about increasing operating margins. Firms need to improve continuously to remain competitive as well. This is because the value of a product or service offered by a firm decreases over time (Figure 3). New entrants and competitors are continuously challenging existing value propositions as they attempt to provide more benefits to the customers. In case competitors are able to provide more benefits to the customer for a given price, firms will suffer from a lower willingness to pay for their offerings, diminishing returns and customers leaving for competitors. This is how companies such as Kodak, BlockBusters, Compaq, Polaroid, BlackBerry and Nokia have seen their customers leaving.

Research scope

This research will be focused on data analytics and its impact on the effectiveness of strategic board decision-making processes. By investigating data analytics and strategic decision-making processes, it will be possible to see how these concepts could interrelate and how data analytics could be used by SMEs to increase the effectiveness of their decision-making processes. The research findings will then be used to answer the main research question, defined earlier: "*How can data analytics improve the effectiveness of SMEs strategic board decision-making processes?*".

Preliminary literature review forms a starting point for this research, in order to acquire an understanding of the main concepts: data analytics and strategic decision-making. And a conceptual model will be developed to visualize how data analytics and strategic decision-making processes are assumed to be interrelated, based on a set of hypotheses. These hypotheses will then be tested by developing new theories based on a collection of phrases in existing literature and reasoning, according to the Grounded Theory research approach. Whether the hypotheses are confirmed or invalidated will depend on the extent to which literature can be used to support the developed theory. The theories developed will then be used to answer the central research question.

2. Literature review

Data analytics and strategic decision-making are broad concepts with various theories and different views involved. This chapter critically reviews existing literature related to both topics and has been used to specify the research problem into greater detail.

Data analytics

Unlike what you would expect from a relatively new concept, there is already a fairly large agreement when it comes to the definition of data analytics. There are various, similar, definitions of data analytics described in today's literature. These are two definitions explicitly defining 'data analytics':

- "Data Analytics (DA) is the process of examining data sets in order to draw conclusions about the information they contain, increasingly with the aid of specialized systems and software" (Rouse & Stedman, 2019).
- "Data and analytics is the management of data for all uses (operational and analytical) and the analysis of data to drive business processes and improve business outcomes through more effective decision-making and enhanced customer experiences" (Gartner, n.d.).

In some cases, data analytics is being called 'data analysis' (Cambridge University Press, n.d.) or 'business (data) analytics' (Liu & Shi, 2015). Although the Cambridge Dictionary defines the process of examining information, especially by a computer as data analysis instead of data analytics, the use of data analytics would be more appropriate during this research because 'data analytics' indicates analysis that is performed with computational power and 'data analysis' could have been done manually as well. Furthermore, the definitions of 'business data analytics' or 'business analytics' refer specifically to the usage of business performance data to resolve business issues (Liu & Shi, 2015). However, organisations do not exist in a vacuum and could apply data analytics technology as well to analyse external data. Because of this and the explorative nature of this project, the broader definition of data analytics described by Rouse & Stedman is more appropriate for this research to avoid loss of information that may occur if a too narrow definition is applied.

And when talking about data analytics, it is commonly described as ‘big’ data analytics. So, when do we speak about ‘big data analytics’ instead of ‘data analytics’? Whether big data is evolved depends on the volume, variety, velocity and veracity of the data (Sun & Huo, 2019). The reason why data analytics and big data analytics need to be distinguished, is because Big Data cannot be analysed by typical ICT tools (such as spreadsheets) due to its volume and complexity. This is in contrast to ‘small’ data, which comes in a volume and format that makes it accessible, informative and actionable (Rouse & Wigmore, 2014). Big data analytics may include the analysis of information widely available on the internet (social media, blogs, news), information from multiple environments, information from multiple languages and slang words (Sun & Huo, 2019). Big Data is pre-eminently too complicated for humans to get meaning out of it without the help from a computer. And considering the amounts of data generated inside a digital workplace of a SME, it might not relatively be that ‘big’ and unstructured compared to the large amounts and variety of data generated on the internet (Domo Inc., 2019). But as we know, humans have very limited data processing power compared to a computer. The amount of data generated by today’s digital services is quickly too much for humans to analyse with simple tools such as spreadsheets. This might explain why data analytics is often called big data analytics in today’s literature.

Another term commonly used as an umbrella term to refer to data analytics, the tools, the architectures and processes involved is Business Intelligence (Pratt & Fruhlinger, 2019) and often abbreviated as BI. The definitions of data analytics and Business Intelligence tend to have a similar meaning although Business Intelligence seems to be more commonly used for descriptive decision support systems, which are less advanced compared to data analytics solutions because these tools tend to include predictive capabilities as well.

Other terms occasionally used to describe data analytics tools are deep analytics, advanced analytics and implicit analytics (Cao, 2017), and tend to emphasise some aspects such as a focus on in-depth understanding of data, a focus on actionable insight or both. Researchers attempted to distinguish the terms Big Data, Big Data Analytics, Business Intelligence and Data Analytics (Dedić & Stanier, 2017) and found that Business Intelligence tools are more focused on the analysis of structured data and depends on the existence of a data warehouse, while Big Data analytics tools do not necessarily require the usage of a data warehouse, because not everything has to be stored in a structured manner in order to be able to obtain meaning out of it. Figure 4 visualizes the interrelationships of each of these concepts. In addition to that, should one be aware that there is not a general accepted framework available and each of these concepts overlap each other, therefore, there is still confusion, even among domain experts about which terminology to use in scientific literature. However, about the potential data analytics has to add value to knowledge discovery processes is much more consensus in scientific literature.

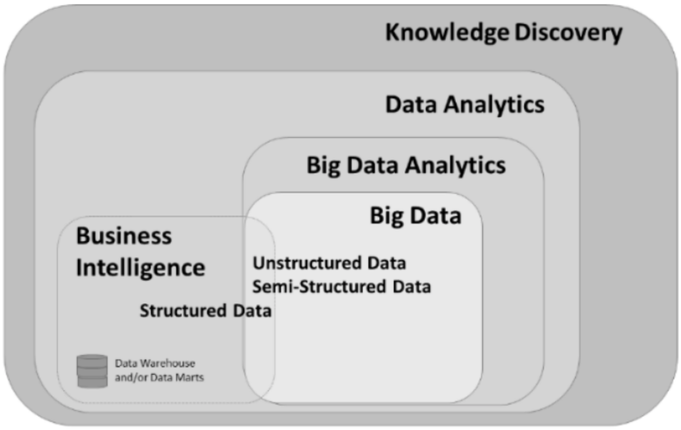


Figure 4: Data analytics and its relationships with other domains
(Dedić & Stanier, 2017)

Data analytics and its relationships with other domains

Data analytics tools can be used to generate knowledge, which is important for organisations to be able to identify, capture, store, analyze, share and develop new solutions (Mbassegue, Escandon-Quintanilla, & Gardoni, 2018). This knowledge management process goes hand-in-hand with (Big) data analytics. Algorithms help to contextualize large amounts of data, which is necessary to be able to transform the data into information. To be able to transform information into knowledge, it has to be applied in decision-making processes. The outcomes of decisions made based on the information collected, results in usable, structured and valuable knowledge for an organization. And because information systems can be used to simulate (i.e. predict) the outcome of decisions based on data analytics technology; new knowledge can be discovered a lot quicker than before. This knowledge will help to determine the optimal allocation of scarce resources and by doing so creating value.

The data analytics approaches can be categorized as follows (Ajah & Nweke, 2019):

- **Descriptive analytics**: simple statistics are applied to describe what is contained in a data set or database. This includes measures of central tendency (median, mean, mode), graphs, charts, frequency distributions, probability distributions and sampling methods. Descriptive analytics can be used to identify opportunities based on knowledge from historical patterns.

- **Predictive analytics** is about applying advanced statistical information software or operations research methods to identify predictive variables and building predictive models. This can be used to predict opportunities in which the firm can take advantage by improving their products and services.
- **Diagnostic analytics** analyses historical patterns to explain why certain events occur. This is often applied for health monitoring and prognosis, fault detection and maintenance.
- **Prescriptive analytics** is about applying mathematical techniques (linear programming) to determine and prescribe the best use of allocated resources. Linear programming is combined with decision theory to optimize constrained budget allocation and to maximize profit.

Processing methods

Raw data has to be collected, cleaned, standardized, consolidated and organized to be able to process the data with a data analytics tool (Michalkova & Bianchini, 2019). The data pre-processing will transform unstructured data into structured data, which is required to be able to apply a statistical analysis on it. Every phase will add value to the data set because the data set becomes more useful to gather insights from. The Organisation for Economic Co-operation and Development (OECD), an inter-governmental economic organisation with 36 member countries, illustrates this process as follows (Figure 5).

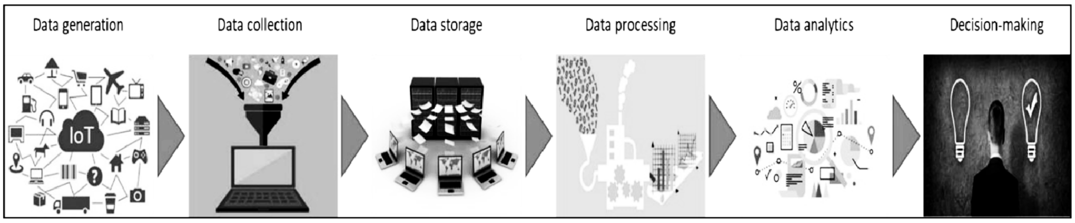


Figure 5: The value chain of data (Michalkova & Bianchini, 2019)

This process model is largely equal to a model described by three researchers in a paper published by HAL (Mbassegue, Escandon-Quintanilla, & Gardoni, 2018). Both models illustrate well how the data analytics process starts with data from various sources and how pre-processed data is used by data analytics to provide insights in order to support decision-making. The difference in Figure 6 is that it defines the outcome of each step more specifically and it shows how the concepts data mining, data transformation, interpretation and evaluation are parts of a data analytics process.

- Data mining is an activity aimed to look for statistical relationships in a dataset (Rouse & Hughes, 2019). These statistical relationships play an essential role because they will tell how much a given column probabilistically affects an outcome.
- The process of changing the format of data for the use of it by another application is called Data Transformation (Hurwitz, Nugent, Halper, & Kaufman, n.d.). This includes data mapping, in which applications are being told how they should process the data.
- Interpretation is the process of trying to understand what the data means (Boyd & Crawford, 2012) and evaluation is about judging or calculating the importance, value, quality or amount of something (Cambridge University Press, n.d.).

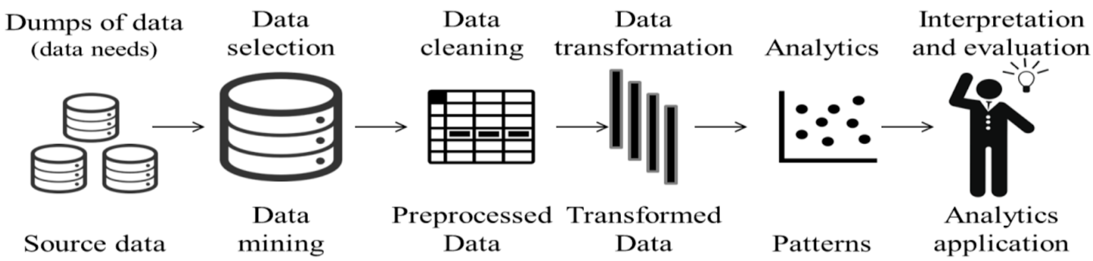


Figure 6: A data analytics process model (Mbassegue, Escandon-Quintanilla, & Gardoni, 2018)

The problem with sequential, waterfall-alike approaches such as the one described above, is that they require a lot of upfront planning, cause struggles with changing requirements during the project, might fail in a late stage of the project and do not deliver any value until the end of the project (Kisielnicki & Misiak, 2017). These limitations make a sequential type of process more appropriate for structured and predictable environments instead of uncertain and unpredictable environments. And taking the highly explorative nature of data analytics projects into consideration, a more iterative way of working would be appropriate.

The industry has also published various models, including two commonly used standards: Daimler Chrysler (then Daimler-Benz), SPSS (then ISL) and NCR's Cross Industry Standard Process for Data Mining (Shafique & Qaiser, 2014) and IBM's Analytic Solutions Unified Method for Data Mining (IBM, 2016). When looking at the differences between the two models discussed and the two standards proposed by the industry, the prevalent difference is the that both the CRISP-DM (Figure 7) and ASUM-DM (Figure 8) model are addressing the need for a cyclic or iterative approach for continuous improvement. The CRISP model as well puts emphasis on the need to align business and data understanding (Shafique & Qaiser, 2014), to make sure the selected data and data quality is appropriate to meet the criteria to solve business problems. This requires some form of theory in order to describe and understand the relationship between the selected data and business objectives. Furthermore, the CRISP model shows how the modelling and data preparation phases occur jointly instead of sequentially. This will result in multiple models, trying to solve the same problem, which will be evaluated to determine the extent of which the objectives are met.

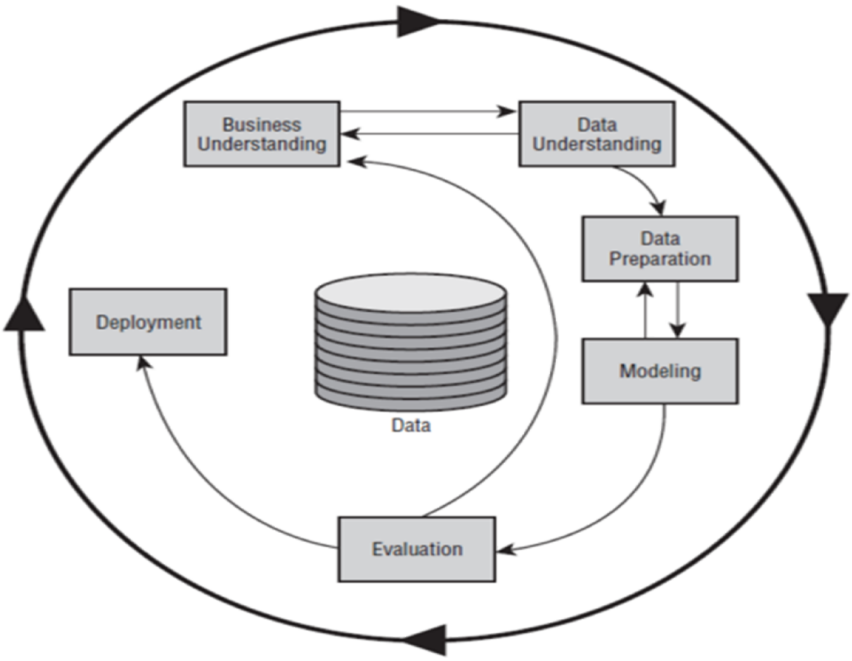


Figure 7: Cross Industry Standard Process for Data Mining (Shafique & Qaiser, 2014)

Researchers have observed that the CRISP methodology lead to a better understanding of the business requirements compared to traditional Agile approaches (Kanban and SCRUM), during an experiment among 85 master students (Saltz, Shamshurin, & Crowston, 2017). However, a drawback of emphasising the need for business and data understanding was that it delayed the analytical modelling phase. The challenges involved with analytical modelling appear much later into the project and may lead to difficulties in meeting the deadline. Besides that, the CRISP model does not pay attention to the visualisation of the data, which plays an essential role for successful end-user adoption.

The ASUM approach (Figure 8), is according to IBM, developed based on real-world experiences combined with industry practises. The approach iterates through five sequential phases, namely: Analyze, Design, Configure & Build, Deploy and Operate & Optimization. Each phase consists of the following activities (IBM, 2016):

- Analyze - defining which features and non-functional attributes the solution needs to accomplish and reach an agreement with all parties involved.
- Design - definition of all solution components and their dependencies, resources and prepare a development environment.
- Configure & Build - configuring, building and integrating components incrementally. The solution has to be tested and validated based on the V-model.
- Deploy - writing a maintenance plan, migration to the production environment and introducing the solution to the business user audience
- Operate & Optimize - solution is running in production and requires periodic maintenance and checkpoints to preserve its health

These iterations are supported by processes which assist with maintenance, monitoring and management of the project.

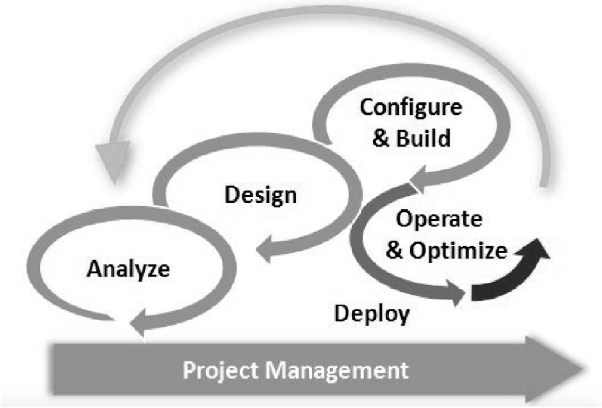


Figure 8: Analytics Solutions Unified Method (IBM, 2016)

Taking each phase of the ASUM model into consideration, it appears to be not as agile as the model itself suggests. It prescribes to start by analysing both functional and non-functional attributes of the solution as a whole, to define all components, their dependencies and to apply prototyping sprints to define requirements, before something has been built. And after that, a maintenance plan will be written, the solution will be migrated to the production environment and the deployment will then be communicated to the business audience. Hence, if the business audience is not satisfied, one would have already fully developed the solution, defined each requirement, every component, each dependency and written a maintenance plan. Moreover, every phase is iterative, which might result in long lead times because every change has to be coordinated to avoid inconsistencies due to the interdependency of each phase. The ASUM model emphasises on the importance of requirement engineering but lacks an early prototyping activity for feasibility analysis.

Furthermore, the ASUM-model seems, despite its name (Analytic Solutions Unified Method for Data Mining), to be a very generic project management model and is not addressing the challenges involved in a data analytics project. This might be because IBM is offering off-the-shell solutions. The results of implementing an off-the-shell solution is probably more predictable than the development of a data analytics solution from scratch and might not require early prototyping because they already know what they can and cannot do with their product due to their previous experiences.

The CRISP model seems to be a leaner, more appropriate, process model for data analytics projects with a greater degree of uncertainty. The CRISP model starts immediately with assessing both business and data to determine in an early stage whether the data can be used to meet the business objectives. Besides that, it’s a methodology tailored to both data analytics and agile principles. However, a visualisation phase is missing and therefore might not receive enough attention during the deployment phase to address the barriers of the business audience when embedding data analytics modelling in business processes.

Data sources

The data used for data analytics solutions can come from a variety of both internal and external sources such as embedded sensor data, phone conversations, documents, video uploads or feeds, social media, web logs and customer surveys. Besides that, the type of data varies across industries: a health care provider deals with very different data compared to a logistics centre or retailer. However, each type of source data can be categorized as ‘structured’, ‘semi-structured’ or ‘unstructured’ data:

- Structured data is identifiable data because it is organized in a structure (Ramesh Kumar & Monica, 2013) such as electronic spreadsheets and databases. Structured data is stored in rows and columns and can be understood by computers and also efficiently be processed by humans.
- Semi-structured data is less structured data such as web posts, blogs, forums, wiki pages, tweets and instant messages. Semi-structured data is data which is still conforming to some set of rules and has tags or other markers to separate fields and enforce hierarchies within the data (Ramesh Kumar & Monica, 2013).
- Unstructured data is data at which the main content does not have a defined structure such as web-pages, Word documents, PowerPoint, e-mails, PDF-files, Audio files, Videos or books (Ramesh Kumar & Monica, 2013).

Figure 9 shows examples of unstructured, semi-structured and structured data sources and how structured data comes with less variety, volume and velocity compared to unstructured data. It also shows how core transactional data is more structured, compared to data stored in internal systems (Varela Rozados & Tjahjono, 2014). External data sources such as Twitter feeds, weather data, blogs, news and Facebook data are more likely to be unstructured as well. Furthermore, it shows that velocity and volume are correlated: quicker data generation leads to more volume.

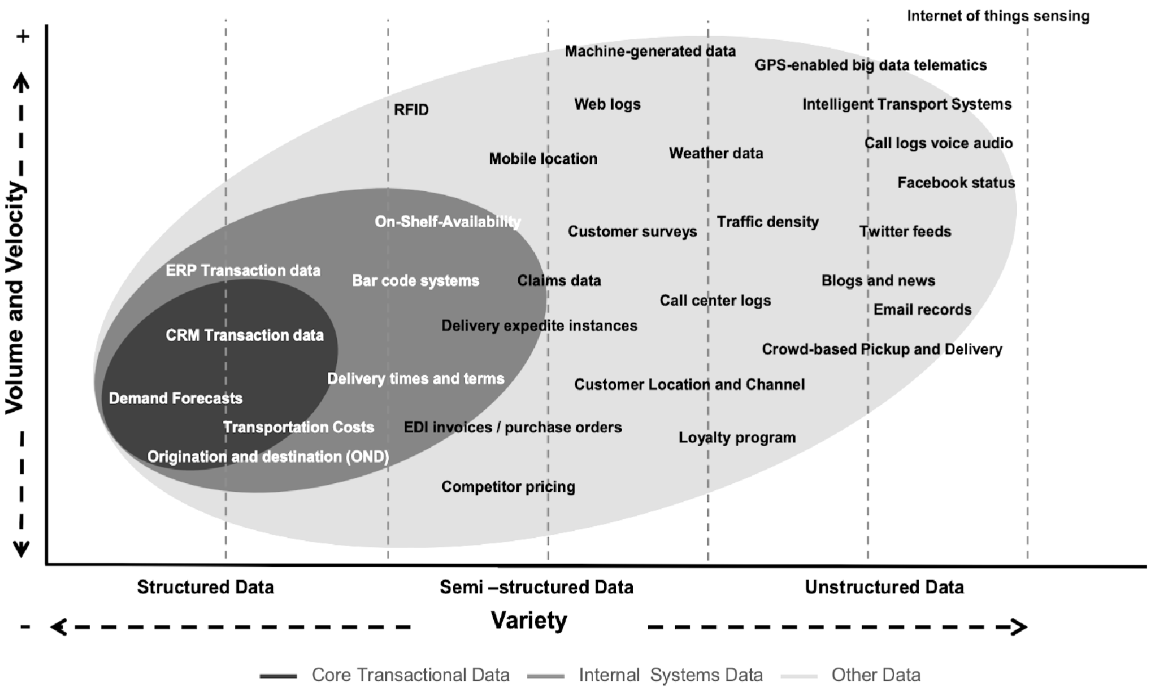


Figure 9: Characteristics of data sources (Varela Rozados & Tjahjono, 2014)

The data sources in a company are typically used to address one main concern, namely: Supply Chain Management. Supply Chain Management is about managing the core activities of a company: procurement, operations, transportation and marketing (Tjahjono, Esplugues, Ares, & Peleaz, 2017). Figure 10 shows how each of the data sources are related to one or more of these domains.

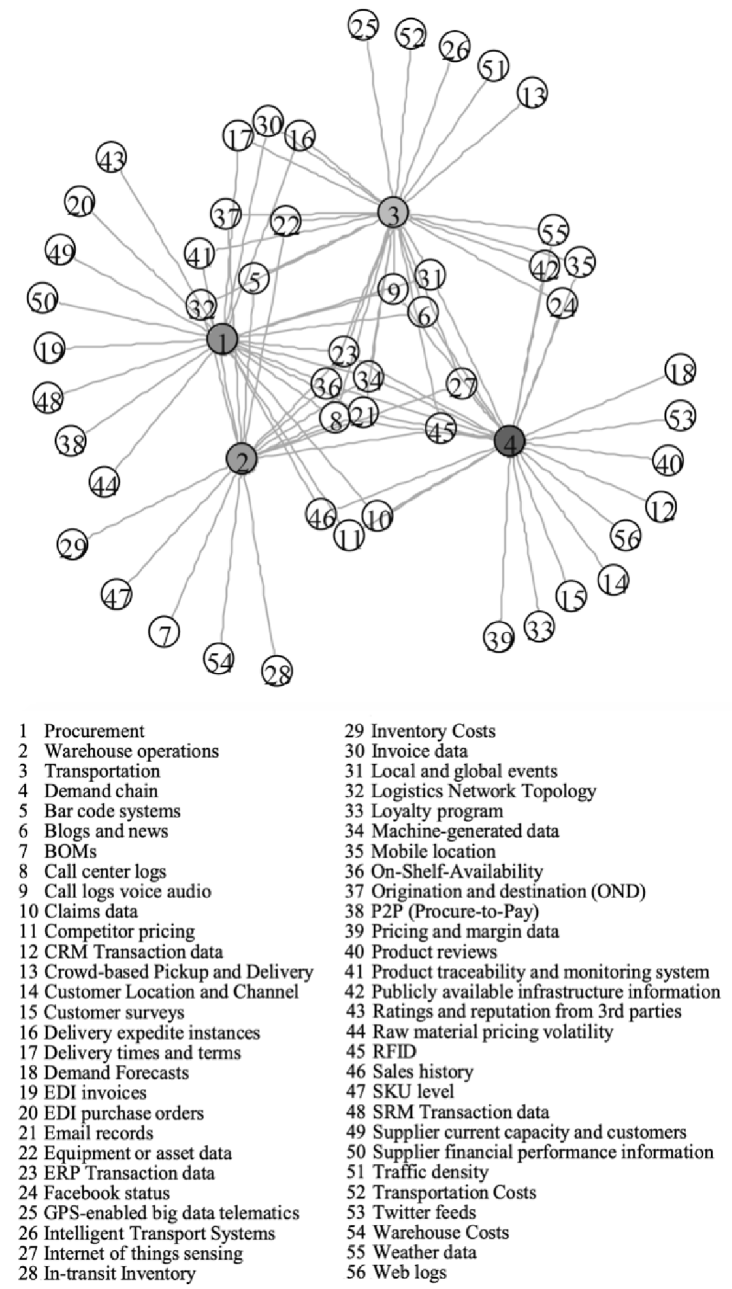


Figure 10: Data sources in Supply Chain Management activities (Varela Rozados & Tjahjono, 2014)

The data used in data analytics applications should meet certain quality requirements to be able to provide meaningful insights. However, the problem with quality is that it's a broad concept and even experts do not universally agree on what the key data quality dimensions are (Askham et al., 2013). Quality metrics can be related to the intrinsic, contextual or representational quality and accessibility (Laranjeiro, Nur Soydemir, & Bernardino, 2015) and people tend to have different views on software as their view is highly depending on the context they are in. Nonetheless ten representatives, including prominent employees from

Microsoft, Lloyds Banking Group, Aston Martin and Aviva UK Health, collaborated in 2013 with DAMA UK to describe the key features of data that can be measured or assessed against defined standards in order to determine the quality of data (Askham et al., 2013). The key quality characteristics identified are: completeness, consistency, uniqueness, validity and accuracy (Figure 11).

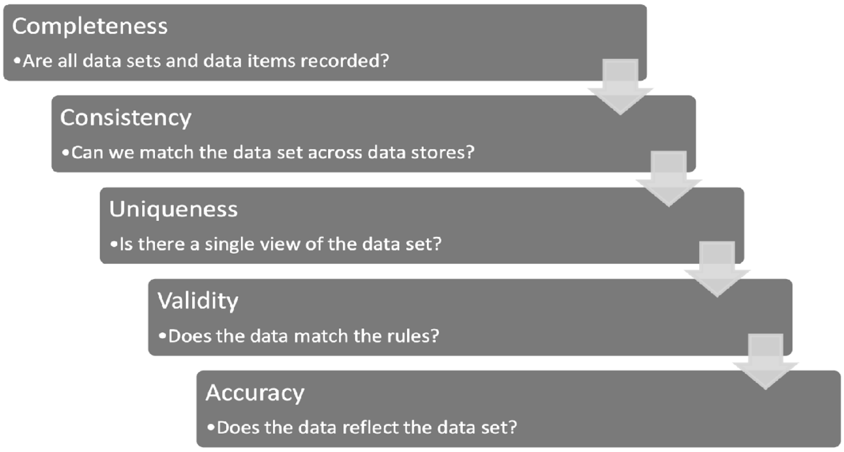


Figure 11: Six primary dimensions for data quality assessment (Askham et al., 2013)

Selecting these primary dimensions means that there will be less attention to other characteristics, such as usability or objectivity, understandability, confidentiality, interoperability, volume, timing and added-value of the data. However, the use of these six quality metrics tend to be more practical for the assessment of the quality of source data compared to other quality assessment methods because it is more specifically related to the source data itself, instead of a data analytics solution as a whole.

Data pre-processing

Data preprocessing is in most cases a mandatory step to be able to use the data for analytical modelling. Raw data in online databases is not likely to be prepared for the use in analytical modelling, which will result in errors, in meaningless models or the data cannot be considered as accurate knowledge (García, Luengo, & Herrera, 2015). This could be because the raw data contains missing values, or because the data needs to be merged with other data to place measurements in context, the data may not been scaled properly (e.g. requiring conversion to nominal, ordinal, interval or ratio scale), requires smoothening, elimination of outliers or random errors or the quantity of data may has to be reduced to be able to make a analytical model work properly. And analyzing raw data that has not been screened can lead to misleading results. Therefore, it is recommended to perform data pre-processing, unless the quality of the source data meets the required quality upfront.

One traditional approach of data preprocessing is ETL processing. ETL processing extracts, transforms, cleans and loads On-Line Transaction data (OLTP), which is stored in day-to-day operational systems, into an On-Line Analytical Processing database (OLAP) (Singh & Sood, 2013). The OLAP database contains historical OLTP data (Figure 12), is non-volatile and does not have to be as large as the source data warehouse as historical trends can often be summarised. OLAP databases tend also to contain multi-dimensional tables, which makes them quickly explorable with data analytics tools compared to relational datatables, because they do not have to be aggregated in order to respond to complex queries.

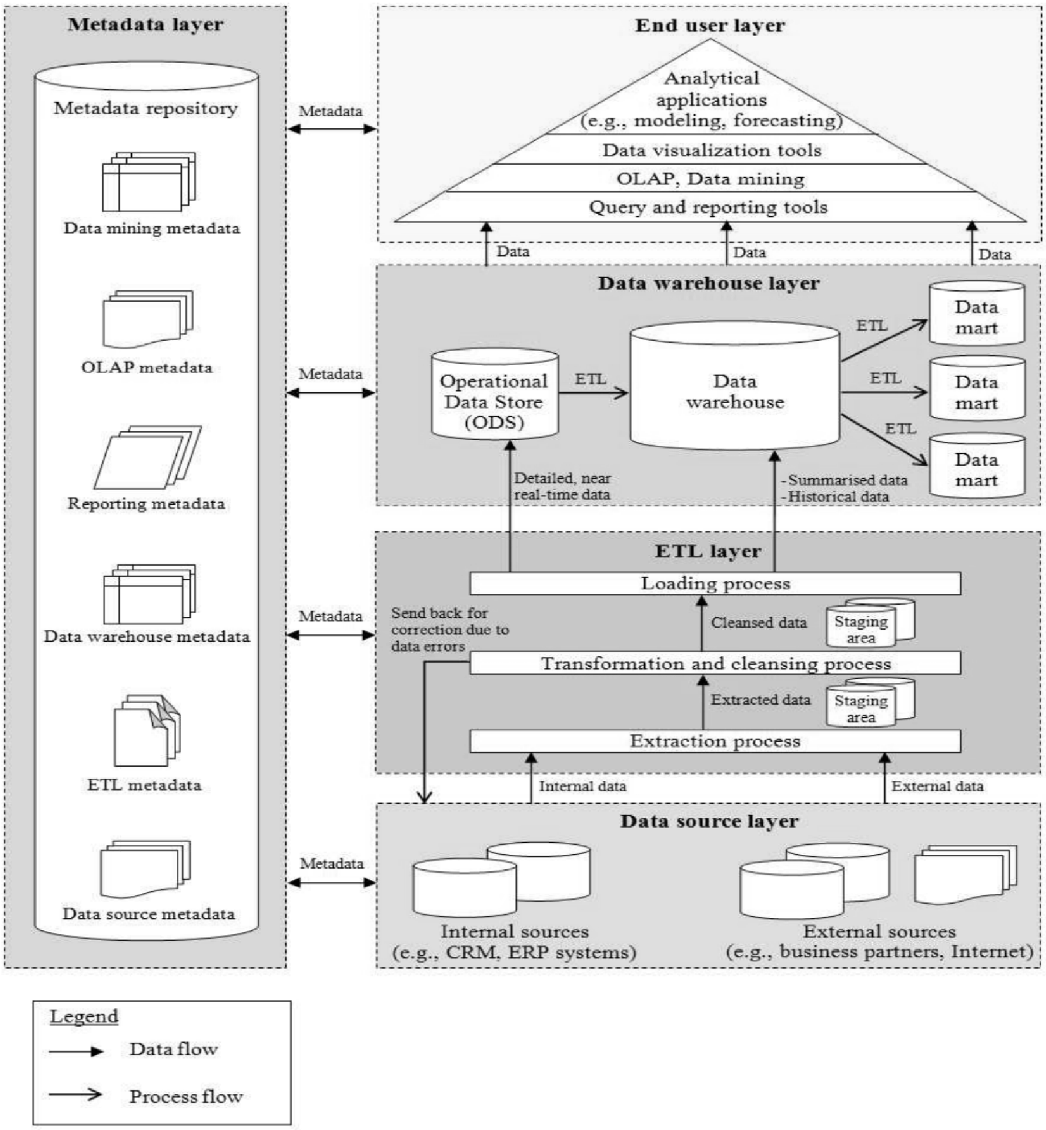


Figure 12: data pre-processing (Lih Ong, Hwa Siew & Fan Wong, 2011)

The biggest advantage of ETL processing is that it delivers pre-structured datasets, allowing quick and stable data analysis. However, ETL processing comes as well with some limitations (Marín-Ortega, Dmitriyev, Abilov, & Gómez, 2014), especially for big data analytics. ETL is not suitable for high volumes of complex data because the volume of data has grown much quicker than the performance of ETL processing. Due to this reason, ETL processing forms increasingly a bottleneck for real-time decision-making. Furthermore, ETL processing is used to transform data before loading it into a data warehouse, however, during a data transformation process, a vast amount of data will be thrown away and the output required depends on the type of algorithm to be applied on the data set. By transforming the data during a data pre-processing stage, one limits their capabilities to respond to changing business needs because the data warehouse may not be ready to work with new analytical modelling algorithms. And if the data cannot be acquired in a timely fashion, the data will be less relevant. Another issue is related to the multidimensional OLAP datatables produced by ETL-processes (Cuzzocrea, Bellatreche, & Song, 2013). Big data sets tend to contain a very large number of dimensions due to their unstructuredness compared to traditional, relational, databases. The increased size of the OLAP datatables results in poor end-user performance, especially during the aggregation and query phases and the complexity of large OLAP datatables makes them difficult to explore. Furthermore, the quality of OLAP datatables will easily become poor due to the highly unstructured nature of big data sets.

An ELT (extract, load and transform) pre-processing strategy can in some degree eliminate these drawbacks. By transforming datasets on demand, one can apply numerous algorithms and respond quicker to emerging business needs (Marín-Ortega, Dmitriyev, Abilov, & Gómez, 2014), because the source data will not yet have been altered or optimized for another algorithm. Furthermore, an algorithm may only need some specific data stored in a large dataset, and therefore can the use of ELT processing make the data transformation itself less time consuming, as it does not transform the entire amount of transaction data.
The use of ELT data preprocessing strategy will make a data analytics platform more flexible because data transformation can be applied multiple times on the same raw data set. However, it will require more storage, because both the original data and transformed data is being kept. Besides that, the training of a (new) model will also be slower because the data has to be transformed first, which is already done if an ETL strategy has been applied.

At the end of the day is data pre-processing going to require a lot of processing power, especially if terabytes or petabytes of source data need to be analysed (Dittrich & Quiane-Ruiz, 2012). Therefore, a new framework called MapReduce is introduced to deal with this concern, which can be used to distribute the load among large computing clusters. However, this requires a lot of resources (e.g. hardware and energy) to facilitate. With MapReduce, will source data initially be assigned to a computing node, to perform some computation in order to make intermediate results which consists of a key/value pair (Khezr & Navimipour, 2017). The next, intermediate, stage is called the shuffle stage, as key/values are being moved between computing nodes and sorted. This key/value pair will then be used by the reduce function, in which the key/value pairs will be combined to form smaller sets of values. The applications of MapReduce vary, it can be used for social network analysis, message mining, intelligent transport systems and healthcare scientific applications. MapReduce is applied by large enterprises such as Google, Amazon, Facebook and IBM. For smaller enterprises, the use of a MapReduce framework may be too advanced, considering the need to set up and maintain computing clusters and after reviewing online tutorials and documentation. However, even if MapReduce is out of reach, for any reason, an SME could still decide to distribute the load of data pre-processing, by splitting data analytics tasks among various systems or use cloud services with these capabilities.

Analytical Modelling - Machine Learning Techniques

Analytical modelling is driven by data science: a concept unifying machine learning, statistics, data analysis and their related methods in order to understand and analyse actual phenomena with data (Alzubi, Nayyar, & Kumar, 2018). Machine Learning provides computer-systems the ability to automatically learn and improve from experience (Henke et al., 2016) without human intervention or assistance. So how does machine learning actually work? A machine is able to learn by applying statistical algorithms to analyse data (data mining). The statistical algorithms will help a computer to discover relationships and patterns in large amounts of data (Hurwitz & Kirsch, 2018). The discovered information will then be used to solve business problems by describing and visualizing data, forecasting or to make inferences based on data. The type statistical algorithm applied depends on the type of problem to be solved. DHL Logistics and IBM collaborated and published a report on the implications and use cases of artificial intelligence for the logistics industry and five problem areas (Figure 13) for machine learning technology to solve: classification, regression, clustering, dimensionality reduction or reinforcement related problems.

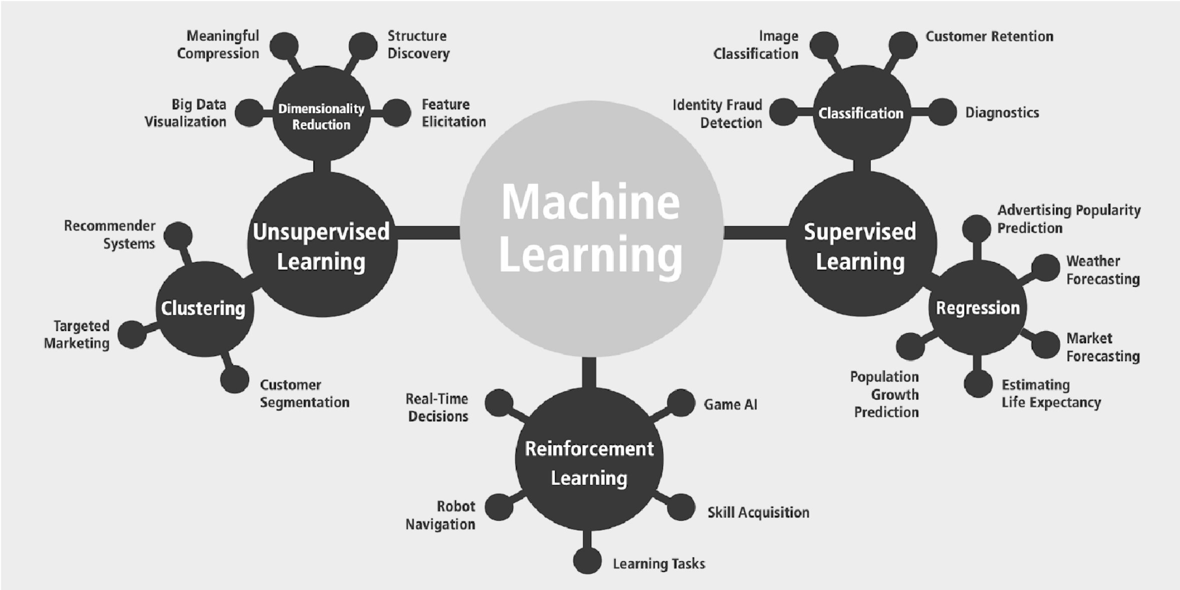


Figure 13: An overview of machine learning problems
(Gesing, Peterson, & Michelsen, 2018)

Classification problems are problems with a fixed number of output such as: Yes or No and True or False. The problem may be a binary or multi-class classification problem (Alzubi, Nayyar, & Kumar, 2018). And regression related problems require an algorithm which is able to provide a continuous and numeric output, and solve problems such as: ‘how much’ or ‘how many’. Both classification and regression related problems can be solved using a supervised learning algorithm (MathWorks, n.d.). A supervised learning algorithm is able to develop (train) a model based on a known set of input and output data. However, unsupervised learning algorithms are used to draw inferences from data sets which consists of input data without labelled responses (MathWorks, n.d.). Clustering and dimension reduction are common use cases for unsupervised learning algorithms. Clustering algorithms try to learn structures within the data and attempt to make clusters based on the similarity in the structure of data (Alzubi, Nayyar, & Kumar, 2018). The algorithm, after being trained, is able to put new unseen data in one of the clusters. Clustering can for instance be applied for customer segmentation, recommendation systems, targeted marketing and object recognition (MathWorks, n.d.).

Reinforcement algorithms are used when a machine's decision has to be made based on previous decision-making outcomes (Alzubi, Nayyar, & Kumar, 2018). The machine learns behaviour by using trial and error and interaction with its continuously changing environment. Machine agents are programmed using the concept of rewards and penalties without specifying how a task has to be accomplished. Reinforcement algorithms are often applied in robotics (Microsoft, 2019), where the set of sensor readings in time is a data point, and the algorithm must choose the robot's next action.

A problem area unmentioned in Figure 13 is anomaly detection, which has been described in a paper published by National Conference on Computational Intelligence (Alzubi, Nayyar, & Kumar, 2018). However, considering abnormality detection is another classification problem: normal or abnormal, it should not be handled as a different problem area. Furthermore, recommender systems are in some papers considered to be a separated problem area. This might be because recommendation systems can work based on both supervised and unsupervised type of machine learning algorithms (Portugal, Alencar, & Cowan, 2018).

Every machine learning algorithm has its own speciality or bias, a specific problem can be solved by several algorithms, and not every algorithm is appropriate for a problem to be solved. Guidelines provided are only meant to be a rule of the thumb (Microsoft, 2019). One algorithm should be compared with another and may even have to be combined to be able to develop the most accurate model. An overview of algorithms per program area is provided by Table 1.

Algorithm	Description	Use
Logistic regression	An algorithm that estimates probability of dichotomized outcome from multiple covariates using logistic function.	Classification
Decision tree	A flow chart–like algorithm that divides data into branches by considering information gain. The final branches represent output of the algorithm (class or value).	Classification/regression
(simple) Neural network	An algorithm inspired by human brain architecture. Layers consisting of nodes are connected to one another with edges weighted as per training results.	Classification/regression
K nearest neighbor	A simple algorithm that classifies observations by comparing k examples that exist in the nearest locations (=examples with the most similar features).	Classification/regression
Support vector machine	Support vector machine draws a boundary line that maximizes margins from each class. New observations are classified using this line.	Classification/regression
K means	A clustering method that makes k clusters in which each observation belongs to the cluster that has its mean in the nearest locations from the observation.	Clustering
Hierarchical clustering	A type of cluster analysis that builds a dendrogram with a hierarchy of clusters. Pairs of clusters are merged to form clusters as they move up the hierarchy (agglomerative approach).	Clustering
Principal component analysis	An algorithm that converts high dimensional data into lower dimensional data with keeping important information as much as possible by orthogonal transformation	Dimensionality reduction

Table 1: An overview of mainstream machine learning algorithms (Kagiyama, Shrestha, Farjo, & Sengupta, 2019)

Model evaluation

A supervised machine learning model is validated by using an out-of-scope sample of test data to see how well the model is able to make accurate predictions. This is being done by splitting the input data in a training and test dataset. The true responses and predicted responses in the test set are compared and by doing so the model is able to determine its accuracy. A model is a good model if it has a high determination coefficient (in percentage), without underfitting or overfitting. Overfitting can occur if a model is too complex and because of that starts to model too many details and includes random noise (Kagiyama, Shrestha, Farjo, & Sengupta, 2019). Models based on large data sets (which are used for AI) are likely to become too complex and deal with issues related to overfitting (Figure 14). This will lead to large variations in predicted values and a misleading low bias. An overfitted model does not capture real associations in data and cannot work well for new data. In contrast, an under fitted model captures associations in data too loosely. This results in high fluctuations between the predicted values and actual values.

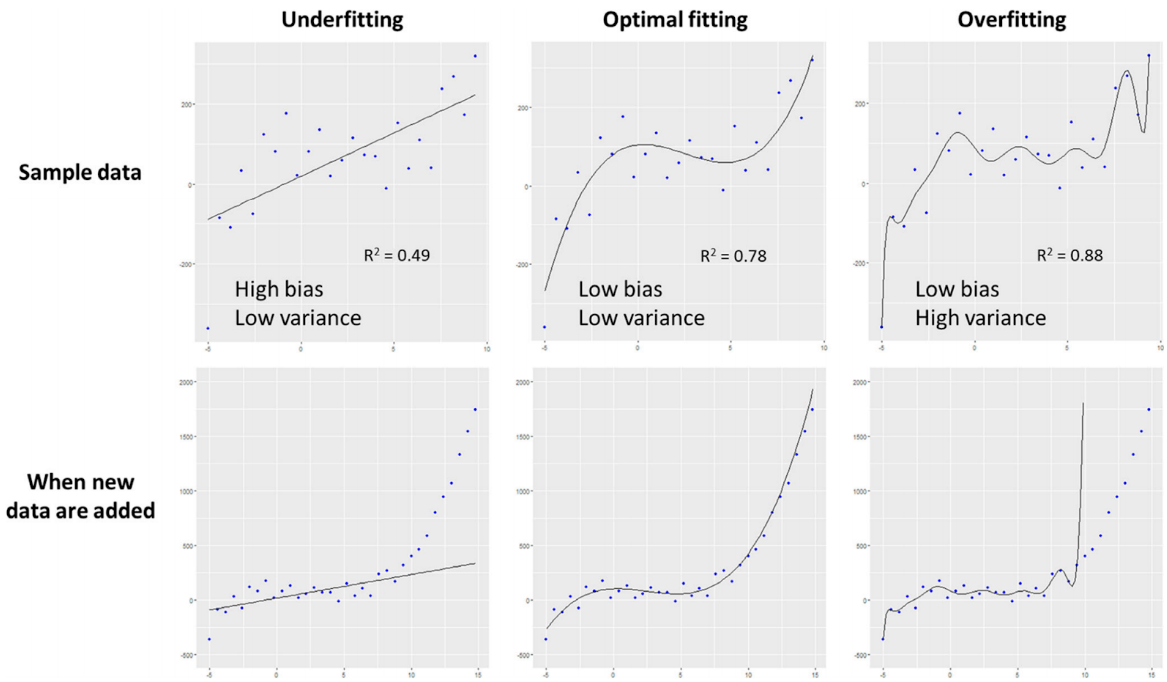


Figure 14: Underfitting, optimal fitting, and overfitting (Kagiyama, Shrestha, Farjo, & Sengupta, 2019)

The accuracy of classification algorithms is determined by comparing the amount of true positives, false positives, true negatives and false negatives in a confusion matrix (MathWorks, n.d.).

The evaluation of a supervised model is an integral part of the process and there are well-defined evaluation measures. However, unsupervised model evaluation is not as well developed (Palacio-Nino & Berzal, 2019). In contrast to supervised algorithms, unsupervised algorithms such as clustering cannot be validated by external data because there is no external information available. This is why internal validation has to be used. This can be done by determining the optimal number of clusters and by assessing how closely the elements of the same cluster are to each other, or by quantifying the extent of separation between clusters.

Practical scenarios for the application and evaluation of unsupervised clustering algorithms are hard to find. A possible explanation is because real world data can rarely be divided into obvious clusters. And even if there are clusters detected, the computer is unable to tell why the data points are similar because a clustered algorithm works with unlabelled data. However, one use case could be to help determine in how many segments a dataset has to be split for another algorithm such as an association rule algorithm in order to increase the effectiveness of, for instance, a recommendation system. The performance of the clustering algorithm could then be assessed by evaluating whether the association rule algorithm has become more accurate if it is aware of the existence of clusters of similar customers.

Dimensionality reduction and reinforcement algorithms can be evaluated by assessing the speed and accuracy under various stress levels (Zubova & Kurasova, 2018) (van Wesel & Goodloe, 2017). In addition to that, are reinforcement models able to quantify actions with a bad outcome (regrets) (van Wesel & Goodloe, 2017).

Data visualization

Data visualisation is about helping people to understand and gain knowledge on the basis of very large, dynamic, complex and often conflicting data sets in a timely manner to support them in their decision-making processes (Kohlhammer, Keim, Pohl, Santucci, & Andrienko, 2011). The cohesion of data, automated data analysis, analytical modelling, data visualisation, user interaction and parameter refinement should provide knowledge to decision-makers (Figure 15), which could be used by them to provide feedback for new ideas. Data visualisation plays an significant role for knowledge discovery because people are able to process visualisations 60,000 times faster than text and 90 percent of information sent to our brain is visual (National Science and Technology Council, 2017). The visualisation for large amounts of data enables decision-makers to make sense of the data, discover patterns and to comprehend the presented information (Sadiku, Shadare, Musa, & Akujuobi, 2016). Commonly used forms of data visualisation are line graphs, pie charts, scatter plots and bar charts. However, optimizing the visualisation of large amounts of data for people is a challenging process, as people differ in their ability to make use of data visualizations and are affected by time constraints. This makes the visualisation of large amounts of data a time-consuming process, as it is difficult to determine which visualisation to be used. Another challenge for data visualisation is speed, as it requires a high amount of processing power to decompose and transform raw data into a meaningful message. The visualisation of large amounts of complex data for end-users is therefore a science by itself (Kohlhammer, Keim, Pohl, Santucci, & Andrienko, 2011), which is called 'Visual Analytics'.

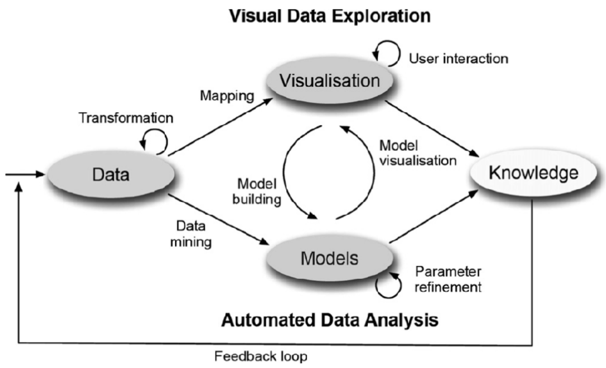


Figure 15: The data visualisation process (Kohlhammer, Keim, Pohl, Santucci, & Andrienko, 2011)

Visual analytics

Visual analytics is about the concerns related to the design of interactive visual interfaces for data analytics tools in order to help facilitate analytical reasoning (Keim et al., 2008). The (limited) cognitive abilities, perceptions and decision making theory of humans should be supported with interactive interfaces to smoothen the interaction with the machines performing the data management, data visualisation, data mining, data compression, data filtering, statistical analysis by developing (Figure 16). Interactive visual analytic interfaces are able, if designed properly, to help people to gather an understanding of large amounts of dynamic, ambiguous or even conflicting data, discover the unexpected, communicate assessments in a timely, defensible and understandable manner. This is being done by highlighting important features and by showing different views to data in a human perception optimized manner (Järvinen, Puolamäki, Siltanen, & Ylikerälä, 2009). Humans are for instance able to understand the meaning of sensory symbols without learning, which allows them to process the information fast. Other factors determining the human perceptual processing rate are contrast, shape, motion, visual attention, distance and size-effects, pattern recognition and visual interaction between the user and data.

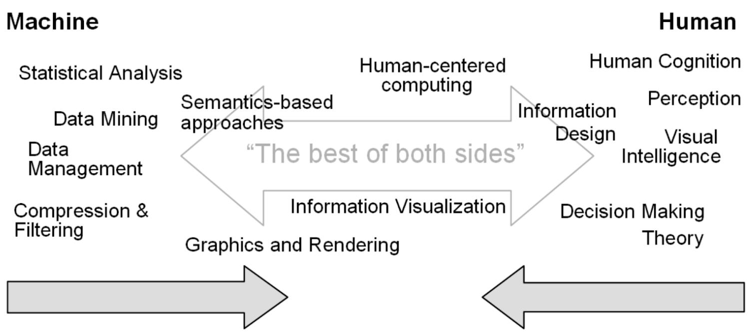


Figure 16: Visual analytics integrates various scientific disciplines to optimize human, data & machine interaction (Keim et al., 2008)

The interactivity between a user and a data analytics dashboard plays a significant role to be able to make them discover new knowledge, as it allows people to test their emerging hypotheses (Sacha et al., 2014). The data will be able to provide direct feedback whether their perception is right or wrong, which could lead to a moment of enlightenment, in which a user realizes that there is a relation between two or more properties or in which the user starts to realize which property is more or less important (Figure 17). However, when visual analytics is introduced to gather knowledge by testing hypotheses, cognitive biases may still occur (Wall, Blaha, Franklin, & Endert, 2017). Users may interact more with a subset of data than others or a small range of values. This may occur because people rely more on specific information instead of abstract signals with a lack of detail. Furthermore, people may become biased because they will only focus on the information presented to them. This may occur when they filter a subset of data without taking the missing information into account. Moreover, people may reject the information presented in an interactive dashboard once there is a uncertainty to be taken into account or continue to believe in information even after it has been discredited. However, these concerns could be addressed by educating users about the biases they may face while using interactive visual analytics interfaces.

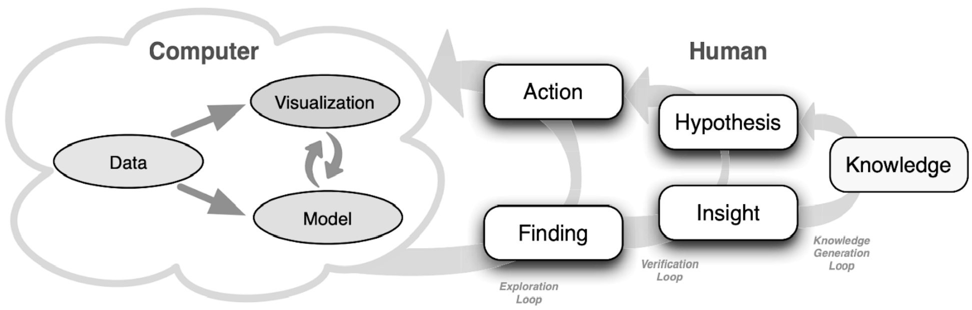


Figure 17: application of visual analytics for new knowledge discovery (Sacha et al., 2014)

Considering that the use of interactive visual analytics dashboards helps people to acquire new knowledge out of historical data, it might also help strategic decision-makers to learn how to make (more) effective decisions. However, in order to determine whether and how visual analytics dashboards could help strategic decision-making processes, it is needed to understand what strategic decision-making processes are and what information they need.

Strategic decision-making

Strategic board decision-making is about making “non-routine, organization-wide resource allocation decisions that affect the long-term performance of an organization” (Nahum & Carmeli, 2019). Strategic decision-making may occur either sequentially, non-sequential, rational or non-rational. A classic view on strategic decision-making was developed by Dewey in 1910 (Ahmed, Bwisa, Otieno, & Karanja, 2014). This model assumes that decision-makers, go rational and sequentially through seven activities: assess the situation, gather facts, assess the unknowns, identify alternatives, establish decision criteria, weigh alternatives, select the best alternative and review the decision. This approach is practical, widely known and understandable for managers. However, it denies the presence of gut feeling and intuition and furthermore it does not address the iterative nature required to develop clarity and to formulate a feasible course of action.

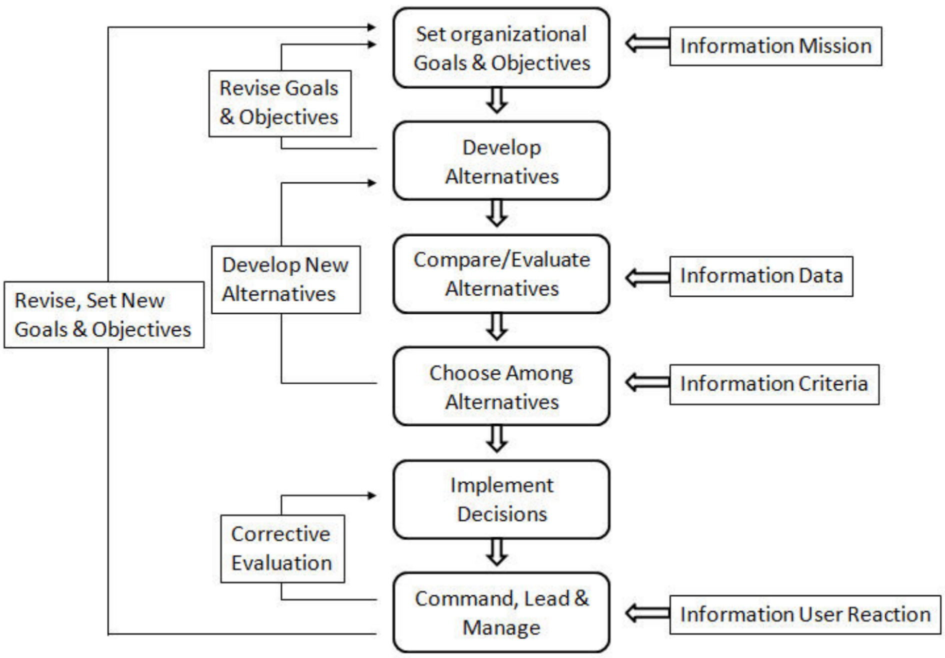


Figure 18: The Military model (Ahmed, Bwisa, Otieno, & Karanja, 2014)

Another rational view on strategic decision-making originates from the Military model (Figure 18) and belongs to the U.S. army (Ahmed, Bwisa, Otieno, & Karanja, 2014). The Military model assumes that the organisation has provided a clear mission and objectives and that these are the driving force in decision-making processes. Furthermore, the Military model has a more iterative nature. The model suggests to revise the organization's mission and objectives and develop new alternatives based on the alternatives available to choose from.

The problem with the military model is that it is unrealistic to expect from SMEs that they are following such a rational approach because a process like the military or classical decision-making model is very static and time consuming. A much leaner and more practical strategic decision-making model is the Cynefin framework. The Cynefin framework shows how the right decision-making approach depends on the type of context a decision-maker is in (Figure 19). It shows how a complex, complicated, chaotic or simple context requires a different decision-making approach (Puik & Ceglarek, 2015). In some cases, it’s unclear to which context a strategic dilemma belongs: these situations are situations of disorder.

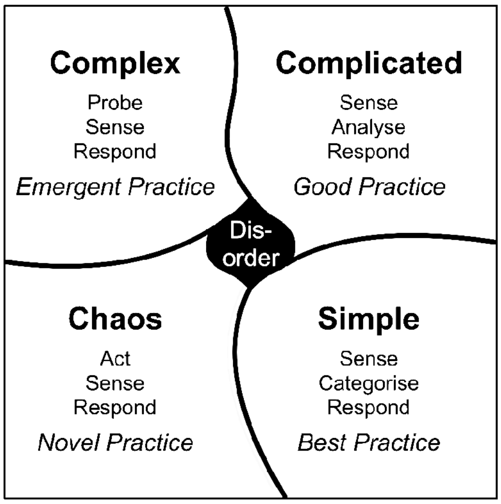


Figure 19: The Cynefin Management Science framework (Puik & Ceglarek, 2015)

Simple situations are clear, predictable, self-evident and linear. In a simple context, one can rely on previously determined solutions (categorise) and there are known best-practises (Puik & Ceglarek, 2015). In contrast to simple management dilemmas, complicated decisions require expertise and are not self-evident. Complicated situations require a decision-making process with sensing, analysing and responding involved. The goal in complicated situations is to achieve ‘good practise’, because there is no best practise available. Even more uncertainty is involved with complex situations, in those cases the cause and effect are yet unclear and require experimenting to determine which actions provide successful outcomes, and which actions do not (Puik & Ceglarek, 2015). Cause and effect become only clear in hindsight and experimenting may lead to unpredictable outcomes and novel approaches of doing things. The fourth type of situation, chaotic situations, are situations without a clear relation between cause and effect, requiring an unconventional and new approach to restore order. In those cases, there is no data to fall back on.

Strategic planning

Strategic decision-makers deal with strategic planning related concerns and address the following four main concerns, according to the '4-Step approach to strategic planning' (developed by Price Waterhouse):

1. Where are we now?
2. Where are our future goals?
3. How are we going to get there?
4. How will we know when we got there? (Schmidt & Laycock, n.d.)

Another theory covering the main concerns of strategic planning is called the 'ABCs of Strategic Planning' (Beheshti, Mahdiraji, & Zavadskas, 2016). This theory describes (Figure 20), just like the 4-Step approach to strategic planning, how an organisation should be concerned of where it is now and where it wants to be, before deciding on how to get there.

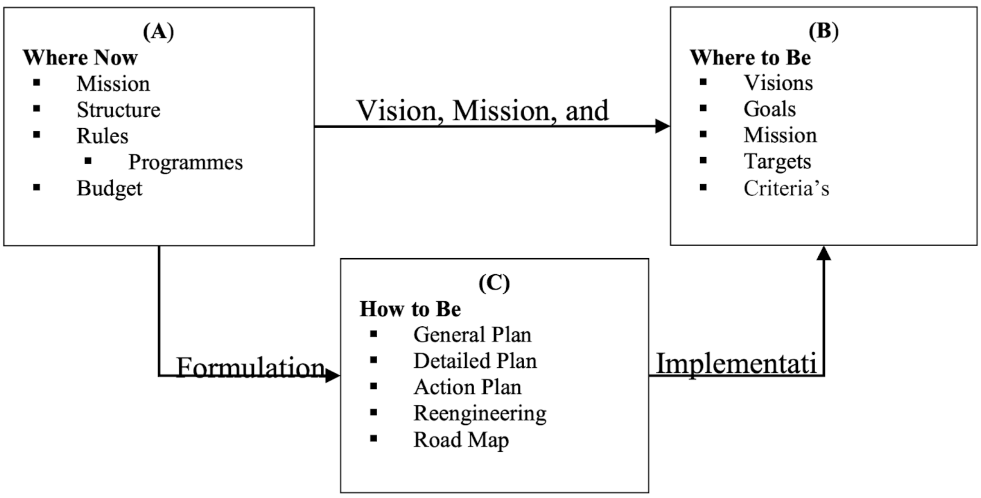


Figure 20: the ABCs of Strategic Planning (Beheshti, Mahdiraji, & Zavadskas, 2016)

The mission and vision of a company should be the company's motivator to develop a long-term strategic planning. Strategic goals are set based on the company's view on how the future looks like and the role the company wants to play in this field. Strategic formulation is necessary for a company since many actors are satisfied with a status quo due to the fact that they are afraid of the consequences of change (Beheshti, Mahdiraji, & Zavadskas, 2016). Strategic management should be an ongoing process of questioning the current status of planned initiatives, responding to changes in the organisation's environment, responding to new requirements for learning and continuous adjustments to the plan.

Figure 21: Interdependency of business actions, activities and business strategy (Purwono & Rahbini, 2013)

Figure 21 provides an overview showing how strategy formulation can be applied to transform a mission and vision, into objectives, a strategy, new policies, programmes, budget allocation and procedures of doing things in businesses. These business activities are in large companies usually divided under top-management, middle-management and lower-management. However, these roles are more likely to be consolidated in smaller companies. Furthermore, Figure 17 and Figure 18 do not include the core values (the identity and ethics) of an organisation. This is another factor influencing the objectives and strategy of organisations (National Treasury - Republic of South Africa, 2010). Besides that, the models above do not pay attention to the performance monitoring of the strategy developed. Performance monitoring would allow companies to adjust their strategy in a timely fashion in case of budget overrun, lack of progress or to respond to its ever-changing environment. This could be done by periodic reporting and by setting milestones.

Strategic positioning

Organisations should continuously be identifying opportunities and challenges that will shape the industry in the upcoming years and determine which type of internal strengths and weaknesses the organisation has, and how their resources can be used to take advantage of opportunities ahead. The Strength-Weakness Opportunity and Threats (SWOT) analysis model is commonly applied by strategic decision-makers to provide structure to this business process. SWOT-analysis attempts to answer the following questions (United Nations, n.d.):

- What do we want to protect that we have or are good at doing?
- What do we want to improve that we have or are not good at doing?
- What do we want to take advantage of to help our organisation?
- What do we want to defend against to help our organisation?

The SWOT-analysis can be used as a starting point for the collection of input required for strategy formulation. Additional management tools can be helpful to tell companies what to look for so that organisations are able to properly answer these four questions. One example of such a management tool, providing structure to discover (external) Opportunities and Threats, is PESTEL. The name PESTEL abbreviates the six concerns a company faces in its competitive area: Politics, Economic, Sociocultural, Technological, Environmental and Legal (Grünig & Kühn, 2015). A company could apply even more models (such as Porter's Five Forces, TOGAF, Hofstede's Cultural Dimensions) to assess each of these problem areas individually. However, PESTEL is different because it's able to describe strategic key concerns in a generic, panoramic way, so that the concerns described by PESTEL are applicable to any organisation. And as mentioned earlier, a business strategy is not based on what is happening today, instead it should be a vision of what the future will look like. This means that the information used in a PESTEL analysis, and the Opportunities and Threats identified, should be future oriented. This vision of the future can then be used to identify critical capabilities required to add value and sustain as a company.

The input gathered during the PESTEL analysis can be used for the next part of collecting input for a company's SWOT analysis: portfolio analysis (Grünig & Kühn, 2015). In a portfolio analysis is real market growth and company's market share (Boston Consulting Group portfolio analysis) or market attractiveness and company's competitive strength (McKinsey's approach) for each product or service being mapped in a matrix. Each product or service in a portfolio analysis could either be a question mark (high potential, low market share), a star (high potential and large market share), a cash cow (high market share, low potential) or a ‘dog’ (low potential and low market share). The results of a portfolio analysis could be used to unveil strengths; products with high market share and high potential, but also weaknesses; a lack of cash flow or a lack of cash flow and future potential.

A portfolio analysis could also be used for trend analysis if a matrix is used. The direction in which a product or service is heading, would be able to indicate opportunities and threats. A growing market could be a basis for a product or service to decide to (dramatically) increase allocated resources and marketing and take advantage of an opportunity (Figure 22). However, a weak growth market would be a threat and requires a company to either preserve its market share and invest defensively or to minimize investment if it’s not paying off due to the strength of competitors or lack of demand.

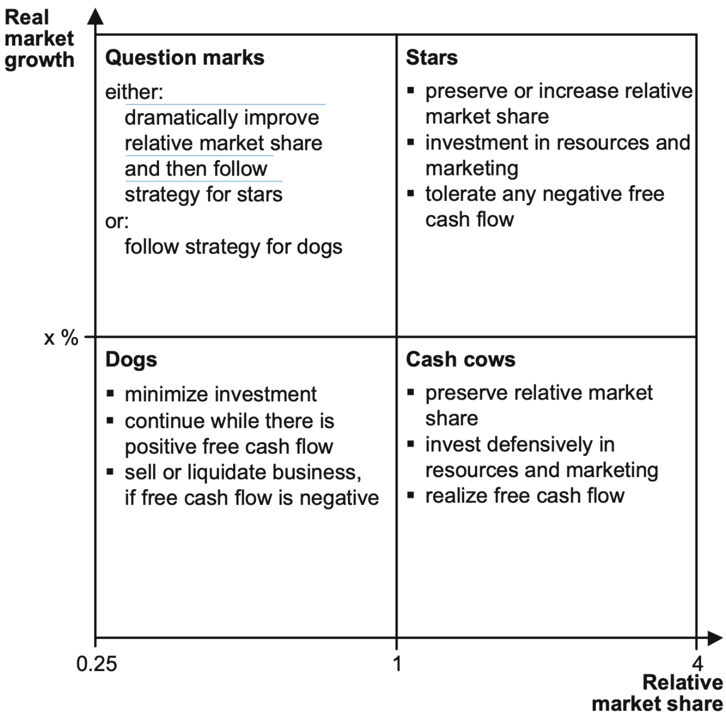


Figure 22: Boston Consulting Group’s portfolio analysis approach (Grünig & Kühn, 2015)

While attempting to draw a conclusion about a question mark, one should take the market life cycle into account. If a product is a question mark just after its introduced by the company or in a new market it’s acceptable, however if a product is still a question mark after a longer amount of time or a declining market, one might need to decide to follow the strategy for dogs (Grünig & Kühn, 2015). Furthermore, some products or services may be considered to be ‘in between’ between two or more classifications and require decision-makers to make a choice: invest, harvest or divest. And besides that, the McKinsey and Boston Consulting Group portfolio analysis are both only two-dimensional models, ignoring service quality, reputation, financial stability, personal quality and marketing effectiveness in their approach for portfolio analysis (National Minority AIDS Council, 2015). And in some cases, an organisation might have legitimate reasons to not liquidate a business and continue with it despite having a low market share and market growth, such as a niche market with high margins or to cover the fixed costs of (underutilized) assets.

However, changes made within a firm's product and service portfolio must be aligned with a company’s internal infrastructure, because it will determine whether a company will be able to develop or change a particular product or service. Strategic decision-makers should therefore not only identify new external opportunities and threats, but also analyse internal strengths and weaknesses in strategic decision-making processes.

Value chain analysis

A value chain analysis can be applied to identify strengths and weaknesses within an organization's internal structure. The Value Chain of Porter (Figure 2), which had been discussed earlier, illustrated how the set of primary and secondary activities performed by a company as a whole form a value chain and how the cost and value of all these activities combined determine a company's operating margins. The value of the company's value chain depends on the scarcity, relevance, durability, transferability, replicability, property rights, bargaining power, the embeddedness of products and services and its ability to address customers’ needs (Grant & Jordan, 2015). The value of a primary activity can be quantified by determining the customers willingness to pay (Popescu & Dascălu, 2011). And the more value a company is able to deliver compared to the resources spent, the more competitive it becomes, which is necessary for a company to defend itself against new entrants (Verdin & Tackx, 2015). Therefore, decision-makers should be deciding upon how firm’s strengths are going to be deployed in strategic important activities to add value to this firm's value chain (Popescu & Dascălu, 2011). Strategic decision-makers should as well identify key weaknesses such as non-value adding activities and decide upon whether these activities should be outsourced or whether some business infrastructure should be changed. The goal is to put a company in favor of other companies by outperforming them by providing an equal product for a lower price or by making them more desirable to make more sales or by achieving superior margins and earn a higher profit.

However, a firm has to perform market research and analyse feedback to be aware of how they could add value to the customer by responding to their needs (Figure 23). Furthermore, a quality management system can be applied to systematically control and improve the performance of business processes to make sure customer needs are continuously met (Popescu & Dascălu, 2011). The determination of appropriate performance levels of business activities is also a strategic decision, which depends on the methods and resources applied and the company's desired competitive advantages. And in order to improve the performance of business activities, a strategic decision-maker could either decide to make small improvements in existing processes or to change the firm's business infrastructure radically to strategically align their organisation to address customers' needs. This could for instance be aimed at adding value by increasing stakeholder satisfaction (e.g. customers, employees, external environment, employers), reducing risk and non-conformities, or by improving the efficiency of a company's supply chain.

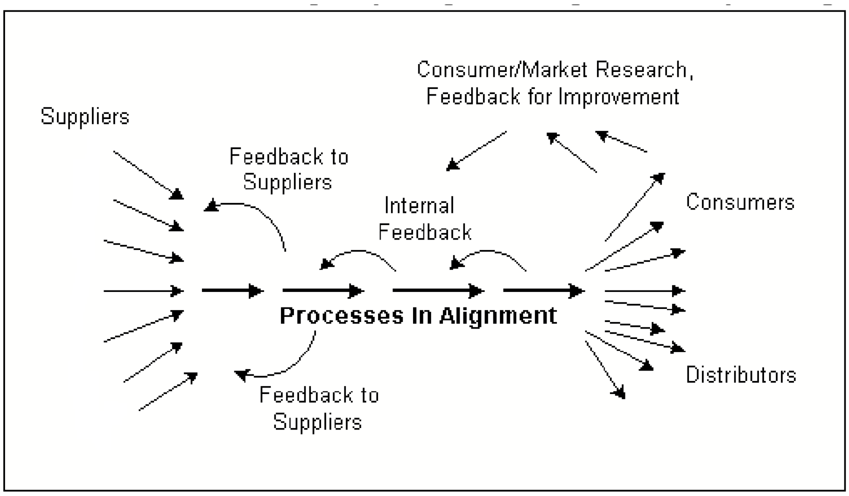


Figure 23: Customers' needs and value chain performance alignment (Popescu & Dascălu, 2011)

The assessment of quality characteristics and performance indicators of a value chain by strategic decision-makers allows them to identify internal strengths and weaknesses, which could be used as a basis to take corrective actions in order to optimize the alignment of value chain performance and business strategy. A decision-maker could for instance influence the value delivered by a value chain by deciding upon the use of (IT) systems, tools or methodologies such as a PDCA-cycle (ISO, 2015), or new manufacturing approach.

Decision-makers may also decide upon new measurements to redesign processes, with regards to the quality of internal or external sources of input, the input itself, process activities, the outputs and the desired quality of subsequent processes (Figure 24).

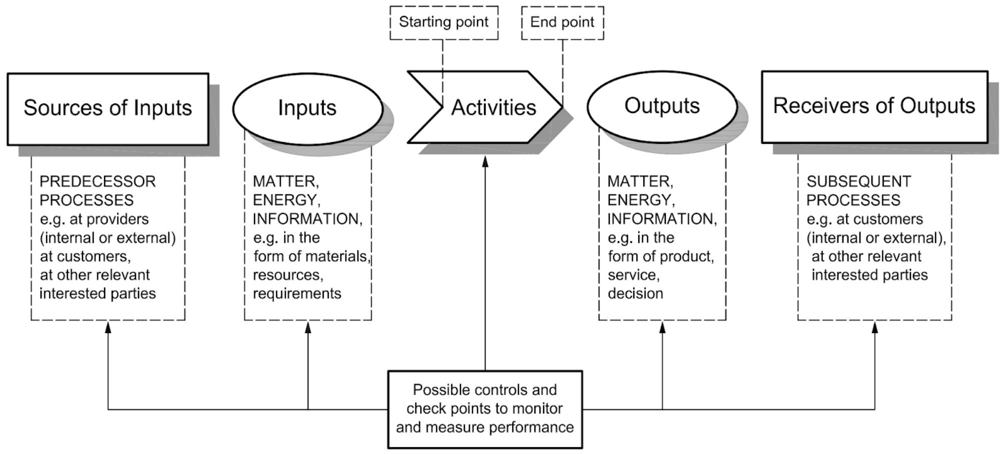


Figure 24: Assessing process quality within a company's internal value chain (ISO, 2015)

Conceptual mapping

Strategic decision-making is based on a company's vision of how their future organisational context will look like and the role the company wants to play within this field. The company's vision can be used as a reference point to determine which opportunities are ahead of them and to determine which internal strengths are required to utilize those. However, a company needs to collect external data to be able to develop a vision of the future of their industry. Without data analytics companies would only be able to use secondary information (e.g. experts, trade fairs, analyst reports and news) to get informed and to obtain awareness of what's happening within a certain industry. Unfortunately, these information sources listed are unreliable, because the information presented may originate from 'hear-say' or poor quality research. These sources are also less able to present information tailored to the company's specific needs, are likely to provide incomplete information, and do not deliver in a timely fashion. These issues may be avoided by analysing published data sources, by using data analytics to measure what's happening. Therefore, the analysis of published data sources is expected to allow organisations to increase the quality of external input for the development of their company's vision which will help them to perform more effective portfolio analysis. The use of data analytics tools should also help to gather external input much quicker (or real-time) because these tools are able to analyse primary data sources.

And in order to be able to determine what's happening in a company, a company needs to collect data from within as well. The data collected could then be used for the identification of internal strength and weaknesses within the company's internal value chain. This is expected, just like in the case of portfolio analysis, to increase the effectiveness of value chain analysis by eliminating biases and by providing (new) insights quickly.

From conceptual mapping to hypotheses

Figure 25 shows the conceptual map developed based on the findings from the literature review, and illustrates how data analytics could affect SMEs strategic decision-making processes. The figure suggests that data generated by internal business activities and published data sources, could be analysed with data analytics tools, in order to provide input for portfolio and value chain performance analysis. This will indirectly affect the SWOT-analysis performed by strategic decision-makers, because the strengths, weaknesses, opportunities and threats identified during the portfolio and value chain performance analysis are now evidence-based. And because this will make strategic decision-makers data-driven, they are expected to set more effective strategic objectives.

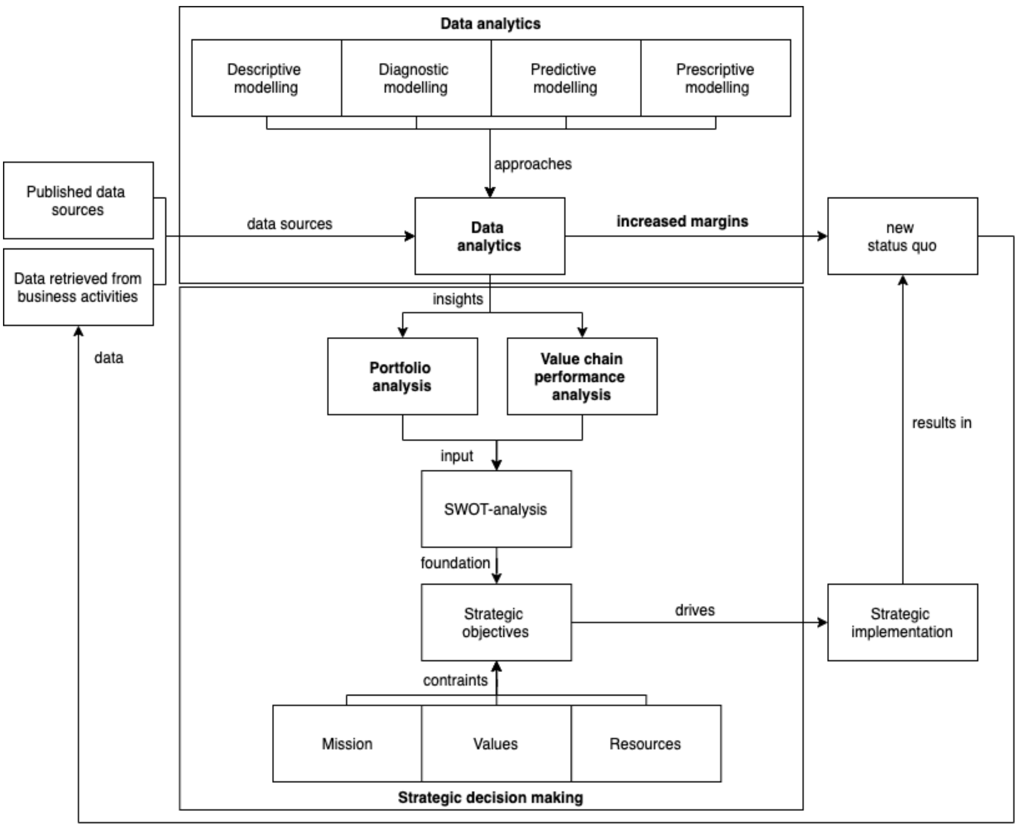


Figure 25: conceptual modeling

This conceptual model is developed based on the following assumptions:

- **Hypothesis 1** - SMEs are able to make use of data analytics to provide information required for strategic decision-making processes (i.e. portfolio and value chain performance analysis), despite having a limited amount of resources to spend.

The first assumption is that SMEs are not required to develop data analytical solutions themselves in order to take advantage of this technology by using SaaS cloud solutions and off-the-shelf software. This will avoid the need to invest in the infrastructure and knowledge required to develop new analytical solutions in small and medium-sized businesses.

- **Hypothesis 2** - All SMEs should be able to gather (new) insights by using tools for data analytics, regardless of their organisational context.

Small and medium-sized businesses may think that they do not have enough or not the right data or infrastructure to be able to take advantage of data analytical solutions to support strategic decision-making processes. Others might think they are not in the right industry to take advantage of data. However, they might underestimate the value of the data generated by their business activities and may be able to make use of published data sources.

- **Hypothesis 3** - Tools for data analytics will allow SMEs to improve the effectiveness of strategic decision-making processes by opposing decision-makers biases while performing portfolio- and value chain analysis.

Tools for data analytics are expected to improve the effectiveness of decision-making processes by opposing decision-makers biases. However, additional research is required to explain why this hypothesis would be true for strategic decision-making processes.

3. Research approach

A research approach has been developed using guidelines from the book “How to do your research project” (Thomas, 2017) in order to provide the right structure to this research and is tailored to the objectives of this project. The approach is discussed in the next paragraphs.

Research gap

The preliminary literature review (chapter 2) formed a starting point for this research project and helped to gain a broad understanding of the information required for strategic decision-making processes and how data analytics can be used to support these business processes. So far, existing literature has insufficiently been able to explain how data analytics would be able to help SMEs to improve the effectiveness of strategic board decision-making processes (Figure 26). There are some potential explanations for this phenomenon:

1. Companies using data analytics for strategic decision-making may not want to share information about their practices as it gives them a competitive advantage.
2. Smaller businesses are less likely to make use of formal strategic decision-making procedures, causing them less likely to be described.
3. SMEs do not have the required resources to understand and take advantage of data analytics for strategic decision-making processes.
4. SMEs might have had insufficient interest in data analytics and/or strategic decision-making processes so far.

However, since a business strategy forms the foundation of a company's success, the potential impact of data analytics cannot be underestimated, as data analytics might be able to help a company to increase their ability to be successful and produce their desired outcomes (e.g. increase operating margins, grow, achieve their company’s objectives). It is therefore important to develop new literature tailored to the needs of SMEs, to help them overcome existing barriers for adoption and take advantage of data analytics. To fill this literature gap, will this research try to answer the following research question: “*How can data analytics improve the effectiveness of SMEs strategic board decision-making processes?*”.

Research design

The Grounded Theory research methodology will be used to analyse existing literature and to develop a theory which explains how data analytics could improve the effectiveness of SMEs strategic decision-making processes. The Grounded Theory is a common approach for the development of a theory by using empirical data (Creswell & Poth, 2018), however, in this case will the Grounded Theory research methodology be applied on existing literature instead because of the confidentiality of strategic decision-making processes, which would cause organisations likely to be unwilling to cooperate and supply unreliable information to an outsider. Another reason to develop a theory based on existing literature is because SMEs tend to be lagging behind when it comes to the use of data analytics, so it would likely be less fruitful to conduct empirical research among them. The concepts data analytics, small and medium-sized businesses and strategic decision-making are also individually well-described in a wide range of literature. This raised the opportunity to combine and develop a new theory based on existing literature by investigating how these concepts could be interrelated. Moreover, the Grounded Theory is an interpretive approach, which takes the researcher's ability to interpret, give meaning, and exclude irrelevant data into account, which will be necessary to develop a theory that applies to SMEs. The research design based on the Grounded Theory research method is visualised in Figure 27 underneath.

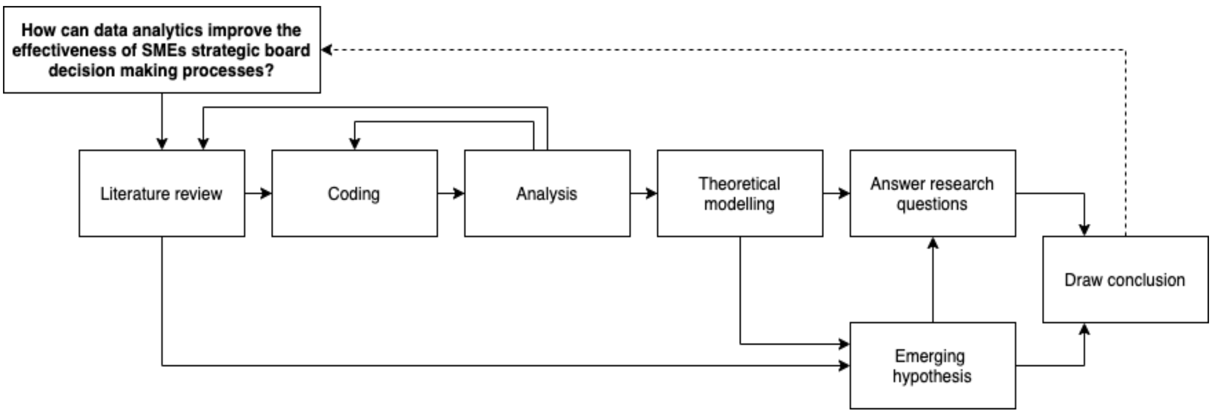


Figure 27: developing a new theory based on existing literature

The Grounded Theory will be applied to answer the following two sub-questions in particular: "How can SMEs use data analytics to provide information required for their strategic decision-making processes, considering their limited amount of resources to spend?" and "How can the use of tools for data analytics improve the effectiveness of strategic decision-making processes?” to be able to answer the main research question. The theories developed will answer these research questions by using existing literature, reasoning and emerging hypotheses. An overview of the research approaches applied per sub-question is provided in the table underneath (Table 2).

Table 2: research approach - an overview

Objective	**Research question**	**Approach**
To gain an understanding of what data analytics is, what it could do and how it works.	How can data analytics help to provide information for strategic decision-making processes?	(preliminary) Literature review
To gain an understanding of what strategic decision-making processes are and what information they need.	What information is needed to support strategic decision-making processes?	(preliminary) Literature review
To determine how SMEs could make use of data analytics for strategic decision-making processes, while having a limited amount of resources to spend.	How can SMEs use data analytics to provide information required for strategic decision-making, despite their limited amount of resources to spend?	Literature review, Grounded Theory
To gain an understanding of how data analytics tools could improve the effectiveness of strategic decision-making processes.	How can the use of tools for data analytics improve the effectiveness of strategic decision-making processes?	Literature review, Grounded Theory

Preliminary literature review has been conducted to gain an understanding of both the concepts of data analytics and strategic decision-making processes, and resulted in three hypotheses, which are going to be tested to determine whether or not data analytics could be applied by SMEs to improve the effectiveness of their strategic decision-making processes. An overview of the three hypotheses and approaches to validate them are shown in Table 3.

Table 3: hypothesis validation

Hypothesis	**Validation**
1. SMEs are able to make use of data analytics to provide information required for strategic decision-making processes (i.e. portfolio and value chain performance analysis), despite having a limited amount of resources to spend.	Examples of data analytics tools addressing the needs of SMEs strategic decision-makers and success stories (if available).
2. All SMEs should be able to gather (new) insights by using tools for data analytics, regardless of their organisational context.	Examples of generic, independent or published data analytics tools which could be applied regardless of SMEs existing infrastructure.
3. Tools for data analytics will allow SMEs to improve the effectiveness of strategic decision-making processes by opposing decision-makers biases while performing portfolio- and value chain analysis.	Examples of how the use of data analytics tools could keep decision-makers away from ineffective decisions and identify (undiscovered) improvement potential while performing portfolio or value chain analysis and success stories (if available).

Literature review

The literature review for this research project is conducted according to a systematic approach. Inclusion and exclusion criteria have been set up for the assessment of literature to make sure that only credible and reliable information is included during this research.

Literature satisfying the following requirements has been included during this research:

- Peer-reviewed sources or having multiple authors.
- Sources from governmental organisations, authorities, NGOs, leading enterprises, academic journals, textbooks, articles from organisations for standardization, charities and post publication reviewed papers.
- Literature published within the last ten years or is still current because the relevance of the literature has not significantly degraded over time and is still applied as a reference point (such as classic management theories, mathematics, statistics, proven facts and superseded theories in science).
- Literature specifying how the data has been collected and applies proper referencing.

The following exclusion criteria has been set:

- Literature which has been published in a language other than english,
- The literature does not provide any additional insights.

The literature has been derived from online databases such as Researchgate, OECD iLibrary, Gartner, HAL-Inria, Semanticscholar, Sciencedirect, arXiv, Elsevier, Springer, the UN Multimedia Library and websites such as Harvard Business Review, IBM, Microsoft, NASA, MathWorks, Techtarget and university repositories. The literature was found by using Google and Google Scholar's search engine. Papers are assessed by scanning the title, date, author(s), abstract, introduction, tables, figures, and conclusion first. Besides that, the reference list is assessed to determine if it contains credible referencing and to discover more relevant literature. Relevant chapters are being read entirely by the researcher. Only papers providing new information, contributing to answering the research questions and papers contributing to a gain understanding of each concept will be included in this research.

Data analysis

The Grounded Theory will, as explained earlier, be used as an approach to generate a theory by analysing scientific literature. This process consists of three stages, namely; Open coding, Axial coding and Selective coding (Figure 28*)* (Mitchell & Noble, 2016).

1. **Open coding**

The first stage, Open coding, consists of reading literature, comparing one paper with another, and coding (c.q. labelling) exemplary quotations (Williams & Moser, 2019). The researcher will determine which quotations will be coded. Some reasons to code a quote from literature would be because it represents a recurring subject, an important idea, explains further, or provides examples, to support the theory developed to answer the research question. Line-by-line coding could as well assist the researcher to methodological analyse literature (Williams & Moser, 2019), however, this approach will not be applied for this research because it would require an excessive amount of time to label every phrase in (scientific) literature while adding a very small amount of value for the purpose of this research. The coding of exemplary quotations will break down the literature into conceptual components and will help to compare, understand and reflect on the data collected.

2. **Axial coding**

In the second stage, called Axial coding, will the researcher be searching to understand and define the relationships between codes by dividing them into sub-categories (Williams & Moser, 2019). Overlapping codes are merged and provide additional evidence for the existence of each sub-category. The coding and axial coding will occur simultaneously, this means that each finding will constantly be compared with earlier findings while collecting data from literature. This so-called 'constant comparison' helps the researcher to learn during the research (Lewis-Pierre, Kovacich, & Amankwaa, 2017) and will also help to verify the validity of earlier findings. The focus area of the literature review will constantly change by the questions raised during the constant comparison.

3. **Selective coding**

In the Selective coding stage are sub-categories from the Axial coding stage selected to be merged into themes in order to achieve an even higher level of abstraction (Williams & Moser, 2019). These themes will then be used to formulate a cohesive and meaningful story, by aligning themes to a main theme. By interrelating each theme, a theoretical story will be developed, explaining how a complex of variables are interrelated.

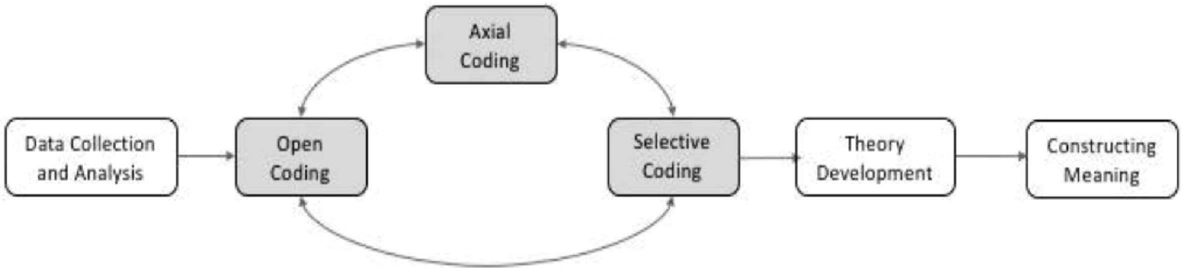


Figure 28: The Grounded Theory approach
(Williams & Moser, 2019)

The researcher is aware of the limitations of the Grounded Theory. First of all: labelling and constant comparison may be time consuming (El Hussein, Hirst, Salyers, & Osuji, 2014), however, in this particular topic there is not much literature available, so there will not be an excessive amount of coding and categorising needed. And another limitation of the model to be developed is that it will only represent the researcher's interpretation of the literature available. A different researcher might develop an entirely different model despite having the same information sources. This means that the model developed will not be able to represent a single truth (El Hussein, Hirst, Salyers, & Osuji, 2014). The quality of the theory developed using Grounded Theory will rely heavily on the researcher's ability to interpret the literature gathered, the quality of the literature and the researcher's ability to interrelate a diverse set of concepts and his ability to develop a theoretical model. These threats to validity will, as mentioned earlier, be avoided by constantly comparing each finding and by 'grounding' interpretations in scientific literature as much as possible.

4. Results

The Grounded Theory has been applied to develop two new theories based on existing literature (Appendix) to answer the research questions “How can SMEs use data analytics to provide information required for strategic decision-making considering their limited amount of resources to spend?” and “How can the use of tools for data analytics improve the effectiveness of strategic decision-making processes?”.

How can SMEs use data analytics to provide information required for strategic decision-making, despite their limited amount of resources to spend?

Data analytics offers an opportunity for SMEs in particular because they tend to have less fixed costs from existing infrastructure. This makes them more able to change direction and respond to new emerging trends in the data they have collected. But as described earlier, the uncertain ROI, complexity and capabilities required to benefit from data analytical solutions make SMEs reluctant to invest. This makes, based on the literature review in chapter 2, sense if it’s about the development data analytical solutions. The development of data analytical solutions is costly and requires a lot of effort (Polkowski, Khajuria, & Rohadia, 2017). However, these costs can be avoided by outsourcing solution development (Collinson & Jay, 2012) by focussing on off-the-shelf BI products and cloud solutions designed to be used by anyone. Some examples of (industry leading) off-the-shelf BI products are Microsoft Power BI and Tableau (Richardson, Sallam, Schlegel, Kronz, & Sun, 2020). Furthermore, the market is offering a wide range of web analytical solutions (Zheng & Peltsverger, 2015), Search Analytics Tools (Ferreira, 2019), Social Management Tools (Ideya, 2018) and Email Marketing Tools (Schaeffer & Olson, 2014) and do not require data science expertise. These tools can be implemented by small and medium sized businesses despite their limited amount of resources to spend. Some of the data analytical tools can be used for free (such as Power BI and Google Analytics), some offer a free trial and most analytical tools offer premium (paid) features on the basis of a monthly subscription.

The reason why off-the-shelf data analytics solutions are affordable for SMEs, despite their limited amount of resources to spend, is because the required development costs are shared among a global customer base and because they do not have to be maintained by SMEs themselves (Papachristodoulou, Koutsaki & Kirkos, 2017). Developing a data analytics software solution in-house would require SMEs to invest in a (continuous cycle) of requirement engineering, designing, coding and testing. Off-the-shelf solutions are also frequently updated by manufacturers (e.g. Microsoft), are more likely to be supported by third party vendors, are designed by a large team of experts and can be installed immediately. And SMEs will not have to spend much resources on handling new feature requests and bugs reporting because they are already being submitted by a large community. Furthermore, off-the-shelf solutions are also often well-described in publicly accessible resources such as books, user group websites, classes and trainings, which will allow SMEs to spend less time on research and development. However, the use of off-the-shelf solutions comes along with some limitations in comparison with customized software systems. SMEs will have limited, if any, power to influence the features or changes implemented by a (commercial) software supplier and SMEs may not be able to modify off-the-shelf solutions by themselves (DragonPoint, Inc., 2013). An inability to change the software to fit SMEs business processes may force a need to work with insufficient or excessive features, an inappropriate workflow or an abnormal vocabulary. And in some cases can change requests only handled by a supplier's high-cost specialists. Off-the-shelf solutions may as well come along with (high) maintenance or licensing fees. Furthermore, the off-the-shelf solution might not be compatible with other software applications (Cohn, 2014) and will not offer a company a competitive advantage. However, for SMEs, the pros are outweighing the cons if it is about the use of off-the-shelf solutions because it makes data analytics accessible to them, despite their limited amount of resources to spend.

SMEs are smaller than large enterprises and have less infrastructure to collect data from (Schaeffer & Olson, 2014). Yet there is still likely to be a lot of data available to analyse, such as transactional (e.g. purchasing, sales, finance) data, customer service logs, social media, web logs (Marr, n.d.), emails, telephony logs (Schaeffer & Olson, 2014) and sensor data (Awwad, Kulkarni, Bapna, & Marathe, 2018) available to gather insights from. A data analytical tool such as Microsoft Power BI would be able to load such types of data as long as its stored in a common form (Iseminger & Sharkey, 2020). Some examples of supported internal data sources are SQL, PostgreSQL, DB2, Access, Oracle, SAP HANA and MySQL databases. In addition to that are also text, CSV, Excel, XML, PDF, JSON files and data published on (public) websites supported. However, there is also an option to load data directly from online services such as SharePoint Online, Dynamics 365, Microsoft Exchange 365, Facebook, Salesforce and Google Analytics. This feature allows SMEs to load and analyse data from cloud services with minimum effort. Some data source integrations (e.g. SurveyMonkey, GitHub, Zendesk) are still in the stage of a beta version. Hence, SMEs do have a broad range of alternatives to integrate their data sources with off-the-shelf BI software, but they remain restricted to the functionality provided as well, because they are unlikely to have the ability to influence a commercial software party's roadmap due to their size, so they are restricted by the data source integrations offered. Furthermore, some of SMEs legacy applications may not be ready for data analytics integration (Adams, n.d.), which could cause uncertain results while performing data analysis. It may also be hard to get access to external data because organisations might not be willing to share it due to its confidentiality, or it might be hidden behind a paywall because of the value the data is representing (Mbassegue, Escandon-Quintanilla, & Gardoni, 2018).

(SaaS) Cloud solutions can also play a significant role for SMEs to avoid emerging challenges while trying to implement data analytical solutions. Cloud solutions can help SMEs to keep away from dealing with legal restrictions, prohibiting data transfers across borders and/or to local storage, this is in particularly the case with privacy sensitive data (Michalkova & Bianchini, 2019). An SME may face difficulties while trying to access certain data, causing analytics processes to be time-consuming (Awwad, Kulkarni, Bapna, & Marathe, 2018). Besides that, it would not be feasible for SMEs to collect all online data because of the volume and costs to store all this data in a data warehouse (Adams, n.d.). Not only the volume of the data to be collected is a challenge, the availability of the data should also be real-time for ideal results, which would require SMEs to invest in a streamlined ETL process. Cloud solutions can also help SMEs to avoid the significant upfront cost to set-up and maintain the infrastructure required to store and analyse a large amount of multidimensional data (Adams, n.d.) because of their 'pay-as-you-go' service model. This lowers the risk of a low ROI, which had been holding SMEs back to invest in new technology (Michalkova & Bianchini, 2019) (Adams, n.d.).

Another challenge to overcome is related to setting up the processes and software solutions required to gain insights from the data being collected (Mbassegue, Escandon- Quintanilla, & Gardoni, 2018) (Adams, n.d.). The implementation of data analytical capabilities should cause a transformation in business processes to be able to take advantage of it (Adams, n.d.). This requires a motivating, supportive, data friendly culture (Deloitte, 2015), in which stakeholders collectively embrace the use of data to support operations and business processes. A simple internal (data analytics) infrastructure would help as well. Simplicity is important to make sure SMEs are able to identify emerging trends (Collinson & Jay, 2012) despite the limited amount of resources they have invested in R&D, and to make sure that they will be able to respond to them. In a simple organisation is less experience, domain knowledge and ability to interpret data required to take advantage of data analytics solutions. Hiring new staff with experience, domain knowledge, analytical skills, and an ability to interpret the usability of data is hard, because it is a combination of skills which is difficult to find (Awwad, Kulkarni, Bapna, & Marathe, 2018). A simple infrastructure avoids the need to spend a large amount of resources to deal with complexity, although SMEs are usually less affected by this compared to larger competitors due to their size.

In addition to the use of off-the-shelf solutions and a focus on simplicity, could a data analytics strategy also play a key role for small and medium-sized businesses, due to a couple of reasons. First of all, a data analytics strategy is recommended by various literature as a tool to prioritize and align the development and implementation of (big) data analytics solutions with the company's business objectives (Awwad, Kulkarni, Bapna, & Marathe, 2018). By addressing the company's 'Where are we now?' and 'Where do we want to be?' questions in a data analytics strategy, and by investigating what kind of information data analytics could provide to the business (e.g. descriptive, diagnostic, predictive or prescriptive models), it will be possible to assess the added value for the business upfront, set a budget, and by doing so, further reduce the uncertainty regarding ROI. Secondly, a data analytics strategy could be used to address the concern 'How are we going to get there?', in order to identify the most cost-effective solutions, while trying to avoid the previously described pitfalls by focussing on simple off-the-shelf solutions. However, there are different views on what else should be included in a BI strategy. Tableau, for instance, suggests to include topics such as: Roles and Responsibilities, Enterprise architecture, Use cases and Content Governance, Education Role Mapping, Users, Community and Upgrade planning (Tableau, 2020). In practise, it will be hard to make a blueprint for each of these topics upfront, especially for SMEs who are unfamiliar with the use of data analytics, considering the amount of challenges SMEs may face while implementing data analytics into their organisations. Besides that, most off-the-shelf solutions are delivered with their own documentation. And in a smaller business are employees likely to know each other better, compared to employees in a large enterprise, so there will be less of a need to formally describe each design decision. Therefore, an SMEs data analytics strategy should instead focus on identifying the most cost-effective solutions to address their needs while avoiding common challenges by comparing off-the-shelf solutions.

Applying data analytics to provide information for strategic decision-making processes

The theory developed based on the analysis of existing literature, has resulted in the following theoretical model (Figure 29) which explains how data analytics could be used to provide new information for strategic decision-making processes while considering SMEs' limited amount of resources to spend. The model proposes a set of critical success factors to mitigate the challenges SMEs are expected to face when they try to use data analytics to support strategic decision-making processes and a set different analytical tools which are accessible to businesses of any size. The interrelations show in particular why SMEs need to develop a data analytics strategy, focused on outsourcing solution development, and how simplicity is required to make the intended organisational changes.

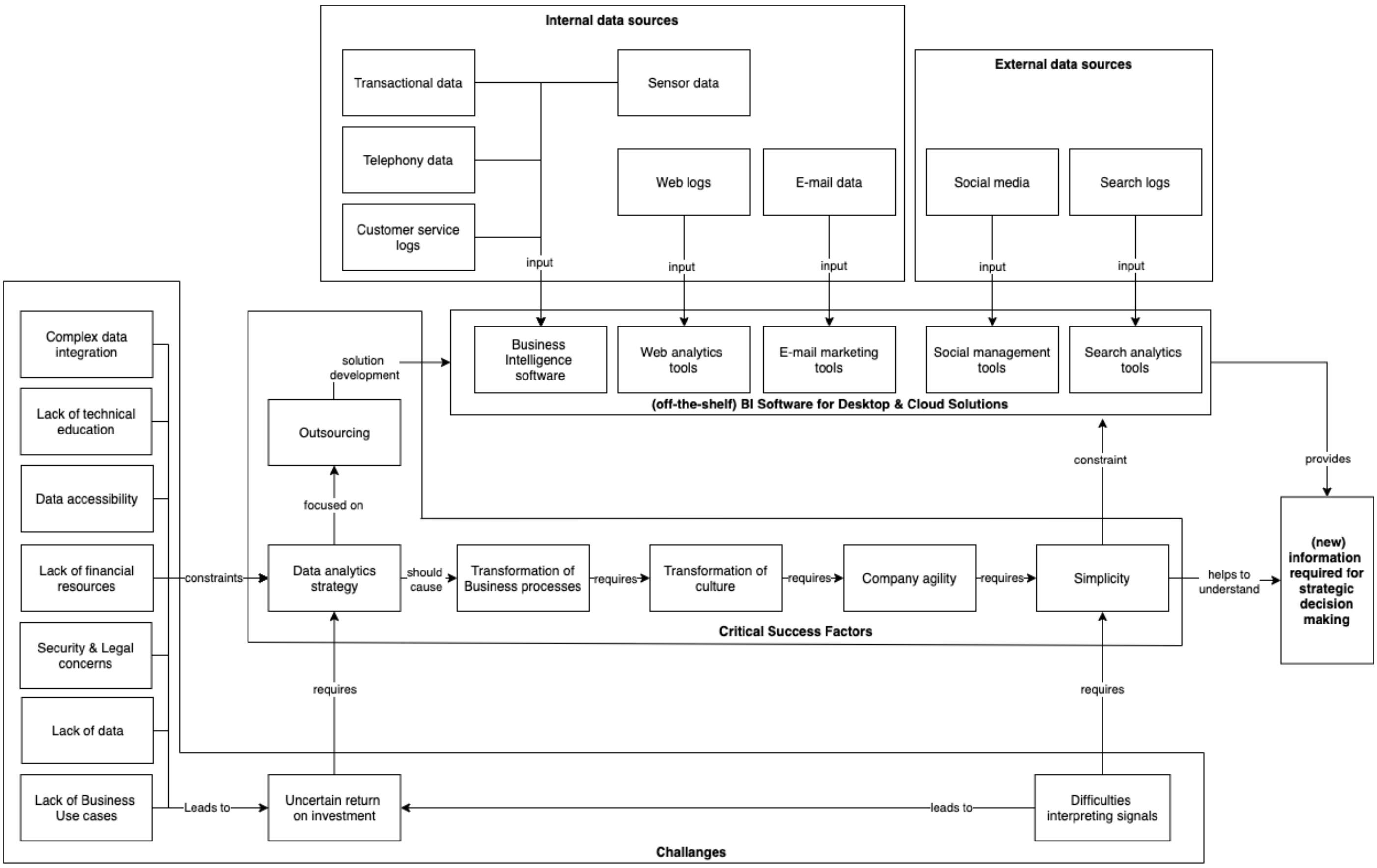


Figure 29: Applying data analytics to provide information for strategic decision-making processes

By responding to the need for simplicity, strategic decision-makers will have less difficulties trying to interpret signals from data analysis tools. And simplicity will also minimize the need for (excessive) training and education, which makes data analytics tools easier to implement and more likely to be adopted by decision-makers. Furthermore, changing a complex data analytics tool would also require more effort, which would make it less able to respond to emerging needs. Therefore, this criteria is identified as a constraint which SMEs should apply while assessing off-the-shelf desktop BI Software and Cloud solutions.

Conclusion

By applying the Grounded Theory methodology in order to generate a theory based on existing literature, it has become clearer how SMEs could make use of data analytics tools despite having a limited amount of resources to spend. The theory developed proposes the use of off-the-shelf and cloud solutions, which could be applied by SMEs, as they will avoid the need for SMEs to develop and maintain data analytics tools by themselves. Various tools (e.g. Microsoft Power BI, Tableau, Google Analytics) are available for free on or on a subscription basis, which makes them accessible to SMEs despite them having a lack of financial resources and minimizes the uncertainty surrounding the return on investment. The use of off-the-shelf solutions is recommended because they will avoid the need to invest large amounts of resources on technical education and staff, especially when they are designed as a generic tool which can be used by end-users. However, because SMEs have a limited amount of resources to spend, it is also recommended by the theory to align data analytics tools with business objectives, to be able to justify investments and to identify the most cost-effective data analytics tool. And in order to make the organisational change, the theory developed recommends to search for simple data analytics tools to make sure decision-makers are able to interpret the presented signals, and because literature suggests that simplicity is a critical success factor for company agility and business transformation.

How can the use of tools for data analytics improve the effectiveness of strategic decision-making processes?

The application of the Grounded Theory on existing literature lead to the generation of a theory, which is largely based on reasoning and assumptions with regards to the inter-relationships between data analytics and strategic decision-making processes, due to a lack of available scientific literature addressing this concern. The theory developed proposes the idea that data analytics could improve the effectiveness of strategic decision-making processes because it allows companies to use data to provide evidence-based input for their SWOT-analysis, which could lead to more effective strategic objectives. This assumption emerged because of the existence of various literature, unrelated to strategic planning or

SWOT-analysis, which had been reporting on how data analytics would allow organisations to become data-driven (Michalkova & Bianchini, 2019), and by doing so improve the quality of decisions made by eliminating biases (Deloitte, 2015). Analytical tools are able to help decision-makers because they allow them to analyse large amounts of data instead of gut-feeling (Deloitte, 2015), in order to assess the performance of a company's internal value chain and to identify strengths and weaknesses. Furthermore, data analytics tools could also be applied to perform popularity assessments for the identification of external opportunities and threats so that companies can optimize their product- and services portfolio.

Data Analysis and Data Analytics tools

A search analytics tool such as Google Trends could for instance be used to measure the popularity of a search term, and by doing so, be used to identify rising or declining trends (Ferreira, 2019). And an early identification and response to an emerging trend could result in a faster growing company and higher operating margins, as it allows companies to respond quicker to changing demand. A search analytics tool could also be applied to determine the company's popularity in comparison with nearby competitors, in order to determine its market share. If a company is underperforming, this could be an indicator for a company to review the competitiveness of its product- and service portfolio. Furthermore, the use of external data for popularity assessment could widen a decision-makers horizon. If a company would only use internal business data, it would not be able to identify emerging trends outside of their product and service portfolio. However, internal sales data could be used as a complementary data source for portfolio analysis, in order to determine how a company performs in comparison with market developments. Another example of an external data source that could be used to assess the popularity of a company's product and service portfolio is social media such as Facebook, YouTube, TikTok, LinkedIn, Twitter or Instagram. Social Media Management tools allow companies to perform sentiment analysis, by analysing trends, subscriber rates, followers, likes, shares, mentions and comments (Confetto & Siano, 2018). The analysis of trends in feedback provided by customers on social media could help companies to understand customer needs and allow companies to respond to them (Salesforce, n.d.). A search analytics tool, could only be used for the identification of new opportunities, or assess a firm's popularity, however, it would not be able to capture 'context', because trends in search behaviour does say very little about the company. The analysis of social media, especially comments, could help strategic decision-makers to gain a better understanding of why a company's popularity is rising or declining.

The identification of internal strengths and weaknesses, will allow decision-makers to improve the effectiveness of strategic decision-making processes even further because it will become clearer which events or departments are decreasing a company's value chain performance (weaknesses), and it will also help for the identification and allocation of excess capacity (strengths). Value chain inefficiencies could for instance be discovered by quantifying the amount of after sales requests per customer and by measuring the amount of return deliveries per product by analysing transactional data. The analysis of internal value chain performance data could as well help to identify bottlenecks by measuring the average queue lengths in each department, the average waiting time for customers on hold, the average delivery time per region and the average required processing time of a service request in each department. These value chain inefficiencies are not only slowing the value chain down, they also indicate a suboptimal quality of service, which lowers customer satisfaction rate and ultimately will lessen customer's willingness to pay, retention rate, word of mouth marketing and company's growth. The elimination of the value chain inefficiencies should therefore be of interest for strategic decision-makers.

Data analytics is not only able to identify value chain inefficiencies, it could also help determine the most impactful ways to eliminate them, in order to increase the throughput of a value chain and improve a company's quality of services. This could be done by inter-relating value chain performance with events (such as downtime, weather, modifications, human error, defects, delays) and the impact of strategic decisions (e.g. changes in capacity allocation, value proposition, quality of service, staff composition, offered trainings, policies, location, used equipment, level of automation, inventory levels, use of materials). For instance, a value chain which is, on a yearly basis, significantly affected by an excessive amount of downtime. The impact of this downtime could be an indicator for strategic decision-makers to invest in new machinery. And after the implementation of new machinery, one could measure whether the performance has improved. If it did, this might be an incentive to invest in more machinery. Without data analytics, the impact of the downtime might have been underestimated and overlooked by decision-makers. Another example could be the use of data analytics tools to analyse the impact of hiring additional staff. BI tools should be able to analyse whether queue lengths have declined and whether the order processing rate has increased. If it did not, this could be an indicator for strategic makers to invest next time in something else, such as training. Data analytics will allow strategic decision-makers to learn and continuously improve a firm's performance, while keeping them away from ineffective decisions, and therefore become increasingly more competitive.

Data analytics tools are able to eliminate decision-maker's biases because they are opposed by the visualization of behaviours and patterns, stored in large sets of (historical) data, and because they allow business performance to be measured and monitored (real-time) by using KPIs. The use of KPIs would help strategic decision-makers to set priorities for strategic decision-making processes, because they will be able to report (real-time) on the company's key-concerns (Microsoft, 2019), typically by the means of a BI dashboard. The use of data analytics tools to keep track of KPIs could also be useful for strategic decision-makers to measure to what extent the company's strategic objectives are met. And because data analytical tools can process data a lot faster than humans, they will also help organisations by delivering quick, visualised, feedback on the performance of a company's strategy. This will allow strategic decision-makers to retrieve knowledge from earlier decisions and change strategy much quicker if it is required.

The elimination of the use of gut-feeling in strategic decision-making processes will, based on these possibilities, assumingly lead to an increased effectiveness of strategic decision-making processes. However, in some cases is the use of gut-feeling still more appropriate compared to the use of reasoning (Okoli & Watt, 2018), such as in a crisis situation. Maintaining an open mind is difficult, and the use of complex calculations to assess each available alternative, will slow down decision-making processes, which is inappropriate in fast-paced and dynamic crisis environments. In addition to that may intuitive decision-making as well be more appropriate for creative minds who are able to think 'out-of-the-box', such as Steve Jobs, which attributed his success to intuition (Umoh, 2018), as entirely new ideas or approaches, cannot be harvested from historical data. Furthermore, the need to collect data may slow down a company's innovation speed, which would lower a firm's potential to become disruptive (Ringel, Taylor, & Zablit, 2015). Even Albert Einstein criticized the society's excessive focus on deliberation at the expense of intuition (Okoli & Watt, 2018). However, assuming SMEs strategic decision-makers are not trying to deal with a crisis situation, such as a near bankruptcy, they should have sufficient time to weigh investment alternatives. And despite strategic decision-makers may not be able to discover new ideas in historical datasets, they could use data analytics tools in order to incrementally improve their ideas, to discover customer needs and to optimize resource allocation. In addition to that, SMEs could also decide to store an excessive amount of data into their data warehouse (ELT pre-processing), just to be able to respond quickly to the emerging needs of strategic decision-makers when they arise.

Embedding data analytics to improve the effectiveness of SMEs strategic decision-making processes

The theory developed based on these suggestions (Figure 30) has been applied to redefine the conceptual model developed during the Literature Review in chapter 2 and attempts to explain how data analytics should be able improve the effectiveness of strategic decision-making processes. The model summarizes how (off-the-shelf) BI software and cloud solutions are able to visualize trends in patterns and behaviour to provide information for popularity assessments and for the detection of value chain inefficiencies. These popularity assessments and detected value chain inefficiencies could then be used to support portfolio and value chain analysis. The outcomes of the portfolio and value chain analysis will then be used as input for strategic decision-makers to perform SWOT-analysis to develop a strategy and to set strategic objectives. The trends in behaviour and patterns visualized by data analytical tools will allow decision-makers to become more effective because they will allow them to make data-driven decisions, by eliminating biases. And the use of data analytics will also allow them to measure, monitor and set priorities based on KPIs as mentioned earlier. The data will also be able to help them to make predictions of the future because they will be able to see the trend in which the organisation is heading, and use earlier data to predict the impact of some earlier decision or event on the businesses performance. Ultimately, is the increased effectiveness of a strategic decision-making process likely to result in higher margins or growth, because they become more able to set value adding and efficiency improving strategic objectives in a cost-effective manner.

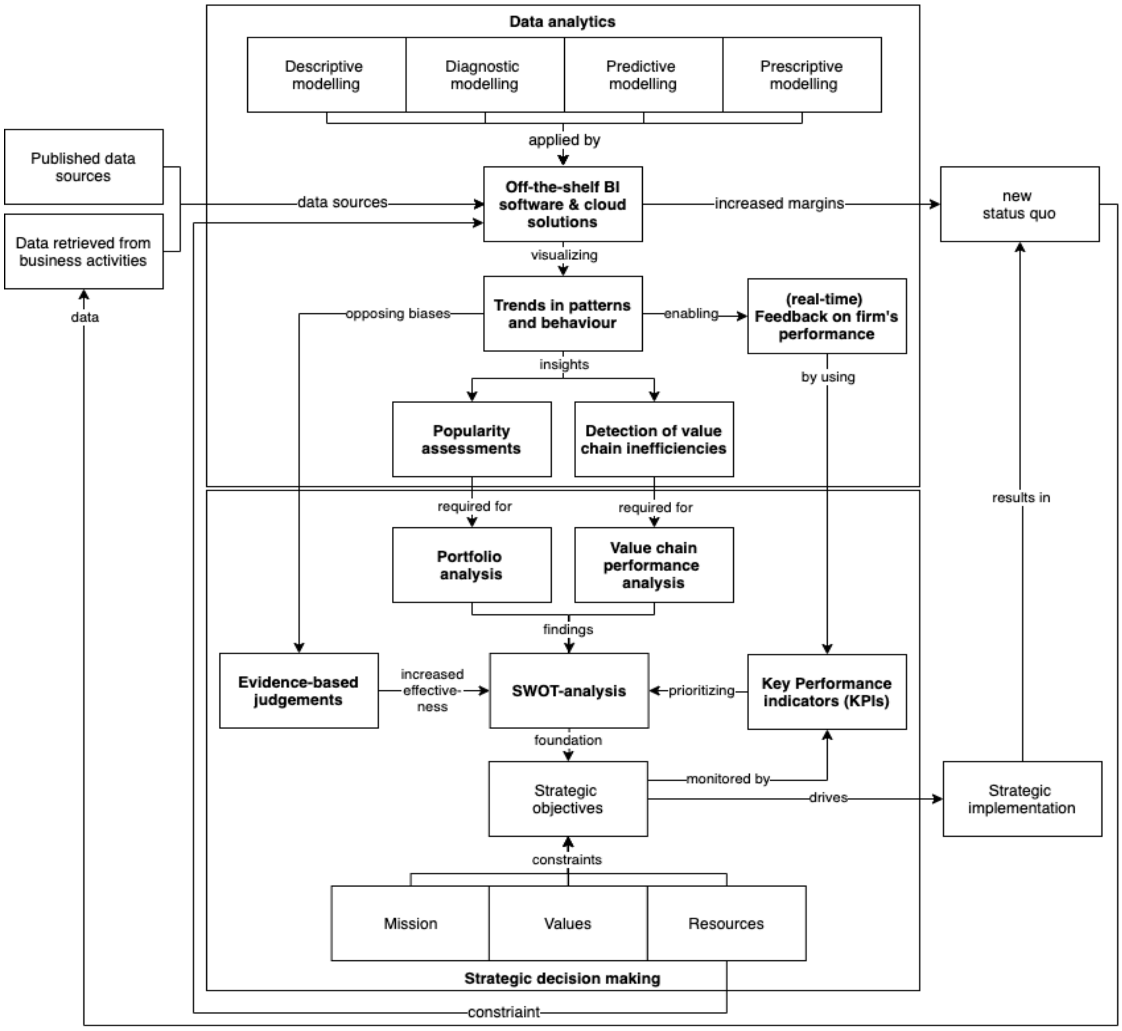


Figure 30: Embedding data analytics to improve the effectiveness of SMEs strategic decision-making processes

In addition to this theoretical model has an artificial success story been developed in order to deliver a specific example in which the use of data analytics could be used to increase the effectiveness of a strategic decision-maker. In this example, data analytics helps Agriculture B.V. by opposing nursery owner Ron’s false belief his Ilex is still in demand just like they have been in the early years (1997-2004) when he started his company. By visualizing the downward trend of Ilex in search popularity and the increased search popularity of Boxwood, he realized, based on a single graph, how his declining sales performance is, likely caused by a decline in demand for Ilex and an emerging preference for Boxwood. As Ron started to realize that the has been ignoring signs of his poor business performance, he decided to ask his local IT service provider to implement a BI dashboard as well, as he wanted to avoid this from happening again. The new BI-dashboard allowed him to identify and improve a suboptimal quality of service and helped him to see how his salesmen performed.

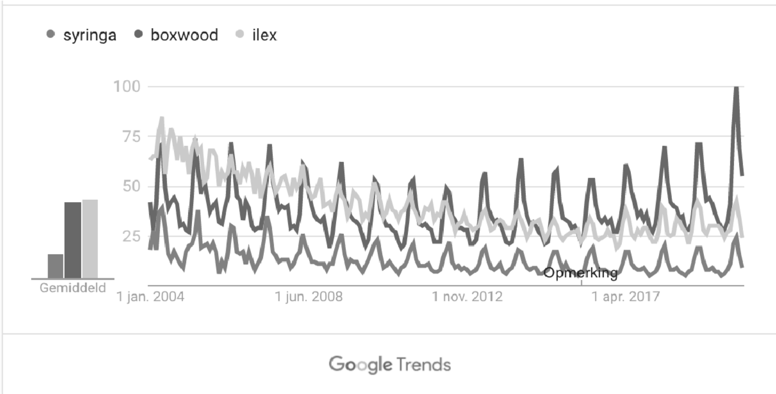
Success Story Example:

Falling sales figures: introducing data analytics to deal with a lack of demand for Ilex at Agriculture B.V.

Agriculture B.V. is a small nursery (24 FTE) founded in 1997 and specialized in growing and supplying Ilexes to various customers within Europe. The Dutch company located in Boskoop has grown rapidly in the early years of existence because Ilex had been very high in demand. Unfortunately, Agriculture B.V. is experiencing a steady decline in demand over the past ten years. The nursery struggles to find new customers and has not been able to achieve a positive operating result over the past year because they were forced to sell their plants at a loss because the remaining customers of Agriculture B.V. have also become less willing to pay for Ilexes. Company owner Ron is looking for new opportunities to make his company competitive again, by either making Ilexes more attractive to customers, lowering his operating cost or to add new species to his assortment. Ron believes that Ilex is still in demand, and assumes that new entrants are the cause of lowered prices. In order to respond to these circumstances, he had already invested last year in new machinery as his company spends a significant amount of money on labour. Unfortunately, this has not changed the situation he is in, as he still needs as many employees as before. Ron also hired a new salesman three years ago, which struggles to generate additional revenue. However, Ron needs to make sure he can pay his bills, therefore he needs to minimize the chance of making additional losses. Ron wondered if data could help him increase his firm's performance.

A quick search on Google Trends

Ron decided to look for data to find out what is really happening. However, he is not a data scientist so he initially searched for an online tool and came up with Google Trends. He decided to compare various types of hedges to see how well Ilex is performing in comparison with other species (https://trends.google.nl):



Ron figured out that the interest for Boxwood plants has skyrocketed while the demand for Ilex has declined steadily. In May 2020, people searched 229% more often for Boxwood than Ilex. However, Ron wondered how he should interpret this data, and found some research, suggesting that it can be incorporated in sales forecasting in order to identify social and economical trends (Boone, Ganeshan, & Hicks, 2015). By identifying customer consumption patterns with Google Trends, it will be possible for decision-makers to respond to them. Another research even indicates that Google Trends was able to predict stock returns (2008 - 2013). A high search volume could predict positive returns in the first one or two weeks. However, they also tend to come with subsequent negative returns (Bijl, Kringhaug, Molnár, & Sandvik, 2016). Therefore, it is important for Ron to distinguish long-term and short-term customer behavior. Google Trends may also not be useful to determine the popularity of a company itself, because a quick search on Google Trends indicates that the search traffic for small and medium-sized enterprises does not generate enough data to show a graph. However, the use of Google Trends to identify new long-term trends, such as an increase in demand for a specific plant species, seems to be helpful for a nursery such as Agriculture B.V., especially because Ron considers to change his product portfolio, due to the decline in demand for Ilex.

Direct management information

Ron also realized that he has been ignoring signs of his company's declining sales performance. Obviously, he noticed some difficulties finding new customers, however, he kept believing in his product as he remembers some highly profitable years in the past. In an attempt to avoid this from happening again, he wants to become more data-driven, and asked his IT-service provider to develop an interactive BI-dashboard for himself. By using a BI-dashboard, he would be able to oversee his business performance (Ferrari & Russo, 2016) and take (corrective) decisions for sustainable long term growth much earlier, which could have avoided him making a loss. Soon after, a local IT-service provider configured a Power BI dashboard upon Ron's request, which allowed him to explore customer spendings, average queue length, amount of aftersales requests, revenue per salesmen, inquiry processing time, established phone calls, delivery times, operating margins, revenue per employee and profit margins per product with an interactive dashboard. These parameters helped Ron to identify various inefficiencies, such as an underperforming salesman, slow response-times on new inquiries and an increase in aftersales requests over the years and declining margins.

New knowledge discovery

While Ron initially only planned to invest in new machinery to lower his operating costs, he started to realize he has to improve his quality of service as well, in order to keep customers satisfied. However, due to the declining operating margins, he also wants to measure how many man-hours are needed for his orders. By using the data of his time registration system, he would be able to measure how many people he has working on his nursery, and see how much new machinery could lower this. By visualising how new machinery impacts business performance, Ron would become more able to assess the effectiveness of his decisions, which could help him for future investments.

Planning for growth

The introduction of data analytics at Agriculture B.V. allowed Ron to determine in which new direction his company should be heading. In order to deal with the long-term downward trend of the popularity of Ilex, Ron decided to start growing 50.000 Boxwood after having spoken to some of Agriculture B.V.'s largest customers. Four out of five of his customers reported a rise in popularity of Boxwood, including the small ones, which could be grown in one year. One of Agriculture B.V.'s long-time customers, would even agree upon the purchase of 25.000 Boxwood for the next year, because they had a hard time finding suppliers of Boxwood.

Moreover, Ron used his new interactive BI-dashboard to identify weaknesses in his business organisation, such as an underperforming salesman, slow response-times and a large amount of aftersales requests. Based on what he saw, he decided to send his salesman to a training program and he also decided to promote one of his employees to become a quality manager, which will be responsible for lowering response-times and amount of aftersales requests.

The use of data analytics provided Ron with various new insights, which he previously would not had without data analytics. Without data telling Ron how his business is performing, he would not have paid attention to his salesman, his customer response-times and amount of after-sales requests, because he was focused on new machinery. This machinery would not have fixed the issues related to the quality of service his firm is offering. This made Ron realize how the visualisation of trends in patterns and behavior opposed his own biases and allowed him to make more effective corrective strategic decisions.

Conclusion

The theory developed suggests that embedding data analytics in strategic decision-making processes could be applied to be able to perform evidence-based SWOT-analysis, which is expected to result in more effective strategic objectives. The visualisations of trends in patterns and behavior stored in historical data should be able to eliminate decision-makers biases by opposing them, and allow them to learn from the impact of their decisions, which can help them to set more effective strategic objectives. However, the theory developed can not be proven yet as there is limited amount of scientific literature and success stories available to support this claim. The visualisation of trends in patterns and behavior is likely to be able to successfully oppose decision-makers biases because humans can process visualisations 60,000 times faster than text and 90 percent of information sent to our brain is visual (National Science and Technology Council, 2017). And the ability of strategic decision-makers to process and generate knowledge out of historical data will increase even more when those visualisations are embedded in interactive dashboards as they help facilitate analytical reasoning (Keim et al., 2008). Especially when strategic decision-makers are able to understand the presented data easily, they are expected to take it into account as it requires much less effort to process.

Performance optimization

Strategic decision-makers should be able to develop (new) knowledge as data analytics tools allow them to test their emerging hypotheses by exploring historical data. However, this new knowledge may not necessarily result in more effective decision-making, because they could misinterpret the data (Mbassegue, Escandon-Quintanilla, & Gardoni, 2018). Some literature indicates that the use of data analytics will result in higher operating margins, as firms utilizing data analytics are twice as likely to be a top performer as they approach business operations differently (LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011). Another case study reported upon a SME achieving a cost reduction of $ 29,262 within six months of using descriptive data analytics software by using data analytics for self-service reporting, increased productivity, efficiency and revenue through new reports and less time spend on report generation (Raj, Wong, & Beaumont, 2016). If these benefits would also apply to the performance of strategic decision-makers, it could have a great impact on business performance due to the leverage they have. Another success story highlights the importance of direct management information regarding the company’s KPIs (Microsoft, 2019) as it provides decision-makers with feedback on the company’s performance.

Ultimately, it is based on the current literature considered to be likely that the use of data will be able to increase the effectiveness of strategic decision-making processes, as the visualisation of trends and patterns and behavior will allow decision-makers to perform evidence-based SWOT-analysis, learn from (real-time) feedback on firm's performance and take corrective actions if needed in order to respond (quickly) to changing circumstances.

5. Discussion

The purpose of this research was to gain a deeper understanding of how SMEs could take advantage of data analytics to improve the effectiveness of their strategic decision-making processes and by doing so protect them from customers leaving for more skilled (global) competitors. SMEs required special attention because preliminary research indicated that only 10% of the small business (10-49 persons employed) and 19% of the medium-sized businesses in Europe have been using (Big) Data analytics compared to an average of 33% of the large companies. Literature suggested that this would be caused by their limited amount of resources to spend on R&D, human resources, organisational changes and process innovation. However, the preliminary literature review also revealed a lack of literature tailored to fit the needs of SMEs, and might explain why they have not been able to take advantage of data analytics solutions for their strategic decision-making processes.

Two new theories have been proposed (Chapter 3), in an attempt to explain how SMEs could take advantage of data analytics to provide information for strategic decision-making processes and to explain how data analytics could improve the effectiveness of their strategic decision-making processes while considering their limited amount of resources to spend. The first model (Figure 29) suggests SMEs to make use of a data analytics strategy and simple (i.e. end user-friendly) off-the-shelf solutions, to avoid or mitigate common data analytics challenges and to avoid the risk of a low ROI as off-the-shelf BI solutions and SaaS cloud solutions tend to be offered for free or on a subscription basis. This model has been grounded in a wide range of literature to validate the researcher's assumptions. In addition to this theory, is there a second theory developed (Figure 30) to describe how data analytics could be integrated to improve the effectiveness of strategic decision-making processes and to address strategic decision-maker's needs. The theory developed suggests that the main reason why the effectiveness of strategic decision-making processes could be increased is because data analytics solutions are able to visualise patterns trends and behaviour, which will help to eliminate biases by contradicting them and because they will allow decision-makers to learn from (real-time) feedback on the performance of the company's strategy.

Furthermore, based on the literature findings and the theory developed, the first hypothesis can be confirmed; SMEs should be able to make use of data analytics solutions to support strategic decision-making processes despite their tight budget.

Off-the-shelf and SaaS cloud solutions allow SMEs to implement data analytics solutions without the need to develop such tools by themselves, are designed to be used by end-users and are accessible for free or on a subscription basis. Unfortunately, as very few (scientific) success stories have been published so far, it is hard to prove this specific hypothesis. One exemplary case study was found on Microsoft's own website, describing how Power BI could be used to provide direct management information and retrieve feedback on business performance, by using KPIs. Other published case studies that were found, tend to include too much customized infrastructure, which makes those success stories insufficiently generalisable. And other case reporting was considered irrelevant, because they did not specifically prove how off-the-shelf solutions could provide information for strategic-decision making processes. Despite this lack of available success stories, it seems to be reasonable to confirm this hypothesis because literature indicates a wide variety of available tools, which could be applied by strategic decision-makers and do not require large investments.

The second hypothesis, which is broader; stating that all SMEs should be able to gather (new) insights by using tools for data analytics, regardless of their organisational context, is valid, considering the opportunity to make use of standalone SaaS solutions such as Google Trends, social media analytics tools, published data sources and the availability of transactional- and e-mail databases in any small and medium-sized business. And generic data analytics tools are able to deal with common data formats such as CSV, XML and SQL. The literature surrounding generic off-the-shelf data analytics tools indicates that the use of data analytics tools does not necessarily require data science expertise. Furthermore, SMEs could apply a data analytics strategy focussed on simplicity to ensure non-data scientists are able to preprocess and interpret data appropriately. Research also suggests that simplicity will help organisations to make the necessary organisational changes. Altogether, there is sufficient evidence to confirm the hypothesis is valid.

The third hypothesis, which states that the implementation of data analytics tools will lead to more effective strategic decision-making processes by opposing decision-makers bias, cannot be confirmed based on current literature. Additional empirical research will be needed to collect data which could help to determine whether it is correct to state that the visualisation of data will successfully eliminate SMEs strategic decision-makers biases, and will result in an increased effectiveness of strategic decision-making processes. However, based on existing literature, it seems likely to be true that the visualised trends in patterns and behavior will increase the effectiveness of strategic decision-making processes.

The visualisation of patterns and behavior in historical data is assumed to be able to contradict decision-makers biases, as it requires a much less effort for strategic decision-makers to process, in comparison with text and speech. And because it will be easier for decision-makers to take the visualised evidence into account, they are more likely to use it and know when their gut-feeling is wrong during decision-making processes. Moreover, various research suggests that the use of visual interactive dashboards, will lead to knowledge discovery and one study reported how firms utilizing data analytics are twice as likely a top performer. Taking these findings into consideration, it seems to be reasonable to assume this hypothesis is true, however, more evidence is needed to be able to prove this.

The theories developed should be able to contribute to the adoption of data analytics for strategic decision-making processes by SMEs because they propose some practical approaches to take advantage of this emerging technology despite their limited amount of resources to spend. And despite the researcher had not been able to use interviews to collect new data, because strategic decision-making processes are confidential to outsiders, the theories do add value compared to what has so far been described by existing literature, because the information surrounding the key concepts have not explicitly been interrelated in a theoretical model until now. *The proposed theories have been able to suggest how data analytics should be able to increase operating margins of organisations into greater detail compared to what has so far been described by existing literature*.

However, the validity of the new theories proposed are constrained because they are developed based on the researcher's interpretation of a wide range of existing literature by applying the Grounded Theory research approach. Which is somewhat unusual, because the Grounded Theory research methodology was designed to develop a theory from empirical data. The results show that the approach of coding phrases and categorising them can be used to develop a theory by using literature. However, a lack of feedback from respondents does make the development of a theory harder and threatens its validity.

6. Conclusion

This research was conducted in a search to understand how small- and medium sized businesses should be able to take advantage of the data analytics tools, despite their limited amount of resources to spend.

The formulation of the central research question was as follows:

“How can data analytics improve the effectiveness of SMEs strategic board decision-making processes?“

An extensive literature review was combined with the Grounded Theory research method in order to develop a theory and to provide an answer to this research question. The conclusion based on these findings is presented underneath.

Research findings

By applying the Grounded Theory on existing literature surrounding SMEs, data analytics and strategic decision-making processes, it has become clearer how data analytics could improve the effectiveness of strategic decision-making processes: the visualised data enables decision-makers to perform evidence-based SWOT-analysis’s (Chapter 3). This finding has emerged because data analytics tools can be used by decision-makers for popularity assessments and to identify inefficiencies, which is needed for the portfolio and value chain performance analysis performed by them. The visualisation of trends in patterns and behaviour will help decision-makers to interpret large sets of data quickly and could cause them to make less biased decisions by opposing their gut feeling. Research suggests that a human brain processes visualisations 60,000 times faster than text and 90 percent of information sent to our brain is visual (National Science and Technology Council, 2017). This will make strategic decision-makers therefore likely to take the data into account as it requires much less effort to process and understand when their gut-feeling is wrong.

The visualisation of the trends in patterns and behaviour could also help decision-makers to monitor key business performance indicators (Microsoft, 2019). This would be beneficial because data analytical tools are able to provide (real-time) feedback on the company’s business strategy, which will allow strategic decision-makers to change direction much quicker, if required. And because a firm’s business performance is visualized it will also be possible for strategic decision-makers to see how their decisions impacted business performance. This will help them to learn, as they will see which actions are more or less effective in certain situations. Another advantage SMEs strategic-decision makers would obtain by visualizing large amounts of historical data is that they will know in which direction trends in patterns and behaviour are moving (Marr, n.d.). This input could be used by them to develop a vision of the future, on which their strategy is usually based (Beheshti, Mahdiraji, & Zavadskas, 2016) (Purwono & Rahbini, 2013). Without data analytics, strategic decision-makers would have to rely on secondary sources such as (manual) departmental reports, experts, trade fairs, analyst reports and news to get informed and to obtain some awareness of what's happening within a certain industry, which tend to be less reliable.

SMEs should be able to take advantage of the opportunities offered by data analytics tools by developing a data analytics strategy (Awwad, Kulkarni, Bapna, & Marathe, 2018), by focussing on simplicity and by using off-the-shelf software. The use of a data analytics strategy is recommended as it could help SMEs to assess the added value of data analytics tools for their business upfront by aligning the use of data analytics tools with business objectives. This will allow them to justify investments, set budgets, and by doing so reduce the uncertainty regarding ROI. And by addressing the concern 'How are we going to get there?' in a data analytics strategy as well, SMEs could look out for the easy-to-use and cost-effective off-the-shelf solutions. Simple data analytics tools are proposed as a means to avoid strategic decision-makers having difficulties while trying to interpret signals and will lower barriers for adoption such as a need for (excessive) training and education.

The usage of off-the-shelf solutions is recommended instead of customized software to avoid the challenges which SMEs are likely to face while trying to develop a data analytics tool by themselves. Preliminary literature review (Chapter 2) indicated that data analytics could require a lot of expertise to set up the required infrastructure, retrieve data, perform data pre-processing, configure machine learning algorithms, assess analytical models, visualise data, make the organisational changes and to deal with privacy concerns. The development of a customized data analytics tool from scratch would therefore be costly, time-consuming and uncertain due to the complexity. By using off-the-shelf BI software and SaaS Cloud solutions, designed to be used by anyone, it will be able for SMEs to avoid these challenges (Collinson & Jay, 2012), as these tools will take care of the complexity for the end-user with (advanced) built-in features. A couple examples of off-the-shelf solutions are tools such as Microsoft Power BI (Richardson, Sallam, Schlegel, Kronz, & Sun, 2020) and Tableau (Tableau, 2013) for the analysis of internal transaction data. In addition to that are various (SaaS) cloud solutions available such as Google Trends (Ferreira, 2019), social media management tools (Ideya, 2018) and email marketing tools (Schaeffer & Olson, 2014). These data analytics solutions tend to offer trials, some are available for free or on the basis of a monthly subscription and are designed to be implemented and used by non-data scientists (Richardson, Sallam, Schlegel, Kronz, & Sun, 2020).

Literature suggests that off-the-shelf solutions are more affordable for SMEs because they tend to have a global customer base to retrieve earnings from (DragonPoint, Inc., 2013). The use of off-the-shelf solutions also eliminates the need for SMEs to maintain the software by themselves. Furthermore, do off-the-shelf data analytic software allow companies to make use of publicly available information sources to educate themselves in a cost-effective manner as well. However, because off-the-shelf solutions are typically developed to serve the mass, the workflow and vocabulaire used by the software may be a poor fit for SMEs' existing business processes. Despite these drawbacks, the pros of the use of off-the-shelf data analytics tools outweigh the cons because it makes data analytics accessible to SMEs, despite their limited amount of resources to spend.

Implications

Small and medium-sized enterprises that have not yet explored data analytics should take into account that they are likely to be less effective while making strategic decisions in comparison with their competitors. Fortunately, they should be able to bridge this gap, considering the wide range of available off-the-shelf solutions, which are able to provide data for portfolio and value chain performance analysis. However, the introduction of data analytics into their decision-making processes should initiate a change towards a more data driven organisation in order to be able to take advantage of it. Those who apply data analytics for strategic decision-making may also benefit from having a competitive advantage compared to large enterprises because they are more able to respond to identified opportunities because they have less resources invested in existing infrastructure. The introduction of data analytical solutions will provide organisations (real-time) feedback on their strategy and therefore allowing them, and force others, to make quicker decisions.

Limitations

The Grounded Theory research approach, that has been applied, is developed to generate theories based on empirical data. And although the researcher initially planned to interview SMEs to investigate how they make use of data analytical tools; it became clear that finding enough organisations willing to share information about their strategic decision-making process was not feasible due to the confidentiality of this topic. This has led the researcher to use existing literature instead of interviews. And despite the credible input provided by scientific literature, it might have been more fruitful to share success stories among SMEs. But then again, the confidentiality of the topic would prevent those ahead from sharing their knowledge because they could disclose a competitive advantage by doing so. Furthermore, the theories developed are only representing the researcher's interpretation of existing literature and should therefore not be considered as a single version of the truth.

Future research

This research attempted to bridge the gap between theory and SMEs practises when it comes to (big) data analytics. The theories developed should help them take advantage of data analytics by improving the effectiveness of strategic decision-making processes. However, the theories developed (Figure 29 & Figure 30) have not been tested to determine their validity. Testing both theories in practise should be considered because the lack of available scientific literature about SMEs strategic decision-making processes and a lack of case studies related to the use of data analytics (Coleman et al., 2016), made both theories heavily relying on the researcher's reasoning and interpretations. The lack of available case studies is a concern which could be addressed by researchers, which could help small businesses to get started with data analytics. Another question which remained open was how strategic decision-makers will respond to visualisations of trends in patterns and behaviour, which oppose their gut-feeling. Decision-makers could for instance decide to reject the information presented to them, and data analytics may even introduce new biases if decision-makers would only interact with a subset of data (Wall, Blaha, Franklin, & Endert, 2017) or when they focus too much on the information presented to them.

Furthermore, additional research could help to determine how common it is for data to be misunderstood by SMEs strategic decision-makers and how severe this can impact business performance.

Appendix: Evaluation - Grounded Theory

The Grounded Theory has been applied to analyse existing literature and to generate two new theories in order to answer the research questions. This appendix shows how exemplary quotes have been labelled in order to form a basis for the developed theories.

Part 1: How can SMEs use data analytics to provide information required for strategic decision-making, despite their limited amount of resources to spend?

Category	Subcategory	Code	Exemplary quotations	Literature
Technologies, products and services	**Business Intelligence software**	**#Personal_analytics_tool**	"Power BI Desktop can be used as a stand-alone, free personal analysis tool. Installation of Power BI Desktop is required when power users are authoring complex data mashups involving on-premises data sources."	(Richardson, Sallam, Schlegel, Kronz, & Sun, 2020)
		#Analytics_by_anyone	"Tableau Software helps people see and understand data. Placed into the coveted "Leader" quadrant by Gartner in 2012, Tableau helps anyone quickly and easily analyze, visualize and share information. More than 10,000 customers across most industries get rapid results with Tableau in the office and on-the-go."	(Tableau, 2013)
	Web Analytics tools	**#Website_data_analytics**	"The fundamental basis of web analytics is collection and analysis of website usage data. Today, web analytics is used in many industries	(Zheng & Peltsverger, 2015)

			for different purposes, including traffic monitoring, e-commerce optimization, marketing/advertising, web development, information architecture, website performance improvement, web-based campaigns/programs, etc.”	
			“Systems, or large parts of systems, may already be in use in a small business. For example, a small business with a web presence can track online visitors to its site via a simple technology such as cookies or more sophisticated visitor analytic tools. Small business owners can track who is visiting the site, where they are coming from, how long they stay, and which pages they browse. This information can be used in targeted marketing campaigns.”	(Schaeffer & Olson, 2014)
	Search Analytics Tools	**#Search_analytics**	“Google Trends is a trends search feature that shows the popularity of a search term in Google. You can view whether a trend is on the rise or declining. You can also find demographic insights, related topics, and related queries to help you better understand the Google trends.”	(Ferreira, 2019)
	Social Management tools	**#Social_media_management_tools**	“Organizations employ SMM technologies to tap into the vast ocean of social media data to	(Ideya, 2018)

			reveal mentions of their brand, topic of interest, companies, and products to gain real time actionable insights and respond appropriately. SMM tools can be defined as software applications, which enable companies to gather, categorize, analyze, monitor, and possibly engage in online conversations about companies, brands, products, competitors, industry and other topics across different social media platforms.“	
	E-mail marketing tools	**#E-mail_ marketing_ tools**	“Cenicola (2013) identified several technologies that small businesses may have in place via which data can be captured – customer relationship management (CRM) systems, websites and visitor/lead analytic tools, email marketing tools, blogs, or electronic payments. Data could also be captured via web-based Voice of Internet Protocol (VoIP) telephone systems (Totka, 2013). “	(Schaeffer & Olson, 2014)
	Publicly available cloud solutions	**#Online_BI _Systems**	“BI suppliers have designed and developed applications and tools to meet real small businesses needs. There are BI systems that are available online. These systems are affordable, easy and they belong to the category of cloud systems. Such solutions are suitable for SMEs, as they do not incur	(Papachristodoulou, Koutsaki & Kirkos, 2017)

			additional installation and maintenance cost. Tools and IT system applications are not considered a privilege of large companies, as the services offered are designed for the needs and requirements of SMEs, which can be just as competitive and successful. "	
Data sources	**Transactional data**	**#Sales_data**	"Useful data sources include traditional in-house data (like sales data and customer service logs), social media, browser logs, text analytics, and large, public data sets (such as census data)."	(Marr, n.d.)
	Customer service logs	**#Customer service_logs**		
	Social media	**#Social_media**		
	Web logs	**#browser_logging**		
	(unstructured) Text	**#Text_analytics**		
	Public datasets	**#Public_data_sets**		
	Sensor data	**#RFID_sensors**	"According to a statistical observation about the growth of data from different sources, the data collected from Transactions, Social Media, Sensors and RFID scans or POS (Point of Sale) were recorded to be 88%, 43%,43%, 42% and 41% respectively, ranking them amongst the top 8 sources for data generation and analytics."	(Awwad, Kulkarni, Bapna, & Marathe, 2018)
	E-mail data	**#E-mail_marketing_dat**	"Cenicola (2013) identified several	(Schaeffer & Olson, 2014)

		a	technologies that small businesses may have in place via which data can be captured – customer relationship management (CRM) systems, websites and visitor/lead analytic tools, email marketing tools, blogs, or electronic payments. Data could also be captured via web-based Voice of Internet Protocol (VoIP) telephone systems (Totka, 2013). “	
	Telephony data	**#Telephony_logs**		
Challenges	**Lack of Financial Resources**	**#Increased_operating_cost**	“The increase in operating costs to store and analyze scalable multidimensional data through traditional databases is often a barrier to companies working with a reduced financial bankroll (Kalan & Unalir, 2016). “	(Adams, n.d.)
		#Financial_constraints	“Despite decreasing costs of data storage, processing and analytics, as well as improving conditions to access to finance for SMEs in most OECD countries, lack of financial resources can represent a key constraint for SMEs to make the necessary investments in building data infrastructure or entering into a service contract to implement data analytics solutions. “	(Michalkova & Bianchini, 2019)
	Lack of Technical Education	**#Difficulties_implementation**	**“**SMEs face difficulty implementing complex software solutions (Coleman, et al., 2016) and must conduct a cost-benefit analysis for investing in potentially	(Adams, n.d.)

			uncertain results from a data-naive team. “	
	Difficulties interpreting signals	**#Interpreting_weak_signals**	“...data is useful when it is being interpreted by someone who has sufficient knowledge of the domain to be able to understand and read the weak signals that can indicate a new trend.”	(Mbassegue, Escandon-Quintanilla, & Gardoni, 2018)
	Lack of Business Use Cases	**#Lack_of_business_cases**	“Not only is there a challenge for companies to see internal use cases for big data analytics but there is a shortage of external success stories for SMEs to model their business implementation after (Iqbal, et al., 2018). “	(Adams, n.d.)
			“...experts complain that “we have a lot of data and we are not doing anything with it.”	(Veeramachaneni, 2016)
	Return on Investment	**#Lack_of_focus_on_added_value**	“...machine learning experts often didn’t build their work around the final objective—deriving business value.”	(Veeramachaneni, 2016)
		#Need_for_justified_investments	“Modern IT systems as well as data analysis systems are a huge expense, often unjustified in the case of SMEs. Profits that SMEs would achieve using such extended methods would be incomparably small in relation to the incurred costs of purchasing, implementing and maintaining Big Data systems. “	(Polkowski, Khajuria, & Rohadia, 2017)
	Complex	**#Delay_by_**	“Outdated data sharing	(Adams, n.d.)

	Data Integration	**legacy_systems**	capabilities coupled with legacy applications not suited for cloud-based integration lead to data latency and the inability to derive confident analysis from data”	
	Security & Legal Concerns	**#Data_security_risks**	“Without the resources of larger companies, SMEs are faced with balancing the risk of data breaches or attacks with the cost of storing data on premises (Iqbal, et al., 2018)”	(Adams, n.d.)
	Data accessibility	**#Data_accessibility**	“Social media companies have understood the value of data and, thus making it difficult for smaller companies to access social media information. Patents are also usually not easily accessible. “	(Mbassegue, Escandon-Quintanilla, & Gardoni, 2018)
		#Transfer_restrictions	“...restriction of data transfer across borders and/or to local storage requirement, in particular for sensible data (e.g. personal information of citizens). “	(Michalkova & Bianchini, 2019)
	Lack of data	**#Lack_of_big_data**	“Big data is inherently high variety; but the data generated and used by small businesses may be more limited in scope. The limited variety of data inhibits the predictive capabilities and discovering the unexpected information that big data projects tout as beneficial. “	(Schaeffer & Olson, 2014)
Critical success	**Adoption of Cloud Computing**	**#Pay_as_you_go #**	“Cloud-based services such as Hadoop or AWS EMR provide SMEs with	(Adams, n.d.)

factors			elastic and parallel resource availability with pay-as-you-go service models, reducing the high cost of entry into data science (Kalan & Unalir, 2016). “	
	Transformation of Business Processes	**#Real_time_decision_making**	“Generating a systematic way to process data allows for businesses to transform data into tangible information (Sen, Ozturk, & Vayvay, 2016) acting as a catalyst for real-time analysis and decisions”	(Adams, n.d.)
	Data analytics strategy	**#Data_analytics_strategy**	“According to a study by Accenture (2014), companies with a disciplined strategy of utilizing Big Data Analytics have had bigger returns for their respective investments in Big Data Analytics...” “Prioritizing the development of a Big Data analytics strategy will help your organization overcome these Supply Chain challenges: By utilizing Big Data and Big Data Analytics, a Supply Chain should function with goals of Improving in areas such as prediction of customer needs, assessment of Supply Chain, efficiency of the overall Supply Chain, reaction time, Risk assessment (ComputerWorld, 2018).”	(Awwad, Kulkarni, Bapna, & Marathe, 2018)
	Transformation of Culture	**#management_commitment**	“... a data-friendly culture is equally as important as the technical	(Adams, n.d.)

			foundation for data driven systems. This transformation of culture can be measured against, “how supportive and motivating is the company leadership and established culture, towards the effective use of data for running the operations and business processes” (Coleman, et al., 2016, pg 8).“	
			“Symptoms of failure are for example too little stakeholder engagement, unclear (analytics) strategy, vaguely defined goals, misalignment of priorities and ambitions across business units, hidden agendas, lack of leadership commitment and too much focus on gathering data, building models and technology solutions. “ “By developing the analytical tool together, you create ownership and understanding.”	(Deloitte, 2015)
	Company Agility	**#Continuous_integration**	“When companies are able to take data science as an agile process, one that will evolve as a company grows in maturity and experience, they are able to realize benefits through continuous integration, delivery and value (Swanson, 2017).”	(Adams, n.d.)
		#Change_readiness	“Conducting a change readiness assessment can help you understand what possible risks and opportunities exist within	(Deloitte, 2015)

			your organization related to the change.”	
		#Dynamic_information_architecture	“To be able to profit from it, processing these data requires systematic collection, and a consistent information infrastructure. The purpose is to give a dynamic character to the data and to highlight the evolutionary aspect.”	(Mbassegue, Escandon-Quintanilla, & Gardoni, 2018)
	Simplicity	**#Need_for_simplicity**	“Complex organizations find it difficult to act on new information and move decisively in the right direction. Overly complex organizations become very internally focused, so they are less able to identify relevant external change and respond quickly to it. “	(Collinson & Jay, 2012)
	Outsourcing	**#Benefit_from_professionals**	“...the company can profit from professionals who might be able to gear them in new directions.”	(Mbassegue, Escandon-Quintanilla, & Gardoni, 2018)
			“Outsourcing may be beneficial for economic reasons, but also to obtain competence and robustness. A service provider can deliver specialized services that may be difficult to provide inhouse.”	(Collinson & Jay, 2012)

By evaluating and interrelating phrases from existing literature, it became clear that there are various off-the-shelf BI tools and cloud solutions available which should allow SMEs to analyse internal and published data sources (Papachristodoulou, Koutsaki & Kirkos, 2017), because they are developed as a generic tool for non-data scientists. Some examples of these tools are business intelligence (desktop) software (Richardson, Sallam, Schlegel, Kronz, & Sun, 2020), web analytics tools (Zheng & Peltsverger, 2015), search analytics tools (Ferreira, 2019), email marketing tools (Schaeffer & Olson, 2014) and social media management tools (Ideya, 2018). These tools should allow SMEs to analyse internal transactional data (Marr, n.d.), customer service logs, web logs, (unstructured) texts, social media data and public datasets, sensor data (Awwad, Kulkarni, Bapna, & Marathe, 2018), email data (Schaeffer & Olson, 2014) and telephony logs. However, SMEs have so far been reluctant to adopt data analytics tools, which could be explained by various challenges SMEs face while considering to adopt these tools. SMEs are more likely to have a lack of access to financial resources (Michalkova & Bianchini, 2019), tend to have less access to technical knowledge (Adams, n.d.), are less able to identify data analytics use cases (Schaeffer & Olson, 2014), find it harder to deal with privacy and security concerns (Adams, n.d.) and are likely to face difficulties while trying to interpret signals (Mbassegue, Escandon-Quintanilla, & Gardoni, 2018). SMEs have also less data to work with because of their size (Schaeffer & Olson, 2014) and may have to deal with difficulties while trying to access (external) data sources (Michalkova & Bianchini, 2019). These challenges cause an investment in data analytics for SMEs to remain unjustified (Polkowski, Khajuria, & Rohadia, 2017), which might explain why only a limited amount of small businesses have adopted this technology. However, SMEs should be able to avoid or mitigate these challenges and minimize the risk of facing a low return on investment, by applying a data analytics strategy (Awwad, Kulkarni, Bapna, & Marathe, 2018) focused on outsourcing (uncertain) solution development (Papachristodoulou, Koutsaki & Kirkos, 2017) and simplicity (Collinson & Jay, 2012) in order for SMEs to be able to implement the solutions and to make the required organisational changes, despite having a limited amount of resources to invest.

Part 2: How can the use of tools for data analytics improve the effectiveness of strategic decision-making processes?

Category	Subcategory	Code	Exemplary quotations	Literature
Change in procedures	**Data driven decision-making**	**#Data_driven_decision_making**	“The impact of data analytics and data-driven decision-making on enterprise performance mostly happens through five channels (OECD, 2013[7]): enhancing research and development (data-driven	(Michalkova & Bianchini, 2019)

			R&D); developing new goods and services by using data either as a product or as a major input (data products and data-intensive products); optimising production or delivery processes (data-driven processes); improving marketing through targeted advertisement (data-driven marketing); developing new organisational and management approaches or significantly improving existing practices (data-driven organisation). “	
	Unbiased judgements	**#Evidence_bases_decision_making** **#Elimination_of_bias**	“Where once gut feeling and proven experience were the most important factors in business decision-making, today we have the ability to use data and advanced analytics to identify and understand complex business processes and interdependencies, predict what will happen next or analyze the needs of clients in great detail. The value of business analytics lies in enabling evidence-based decision-making and avoiding biased judgements. “	(Deloitte, 2015)
	KPIs	**#KPI**	“By defining business critical questions that should be prioritized, drive a focused work effort that is aligned from the CEO to the individual contributors performing the analysis. “	(Adams, n.d.)

Augmented insights	**Spot trends in behaviour and patterns**	**#Spotting_behaviour_and_patterns**	“Spotting and monitoring behaviors and patterns allows us to take a stab at predicting where things are heading, how demand for our products or services will change over time, and what will prompt that change. “	(Marr, n.d.)
Improved Identification Opportunities & Threats	**Popularity assessments**	**#Popularity_assessments**	“Google Trends can offer insights on the popularity of a brand or product, and social media analysis can illustrate popularity (i.e. how often a company is mentioned) and show what customers are saying. “	(Marr, n.d.)
Improved identification Strengths & weaknesses	**Value chain optimisation**	**#Improve_customer_interactions**	“Companies can also better interact and engage with customers by analysing customer feedback in order to improve a product or service. “	(Marr, n.d.)
		#Identify_ineffieciencies	“With any business process that generates data (for example, machinery on a production line, sensors on delivery vehicles, customer ordering systems), you can use that data to make improvements and generate efficiencies.”	
		#Optimize_inventory_management	“Retail companies are able to optimise their stock keeping based on predictions generated from social media data, web search trends and weather forecasts. This allows stores to stock up on the most popular items, ensuring they don’t miss out on sales and	

			reducing the amount of unwanted stock lying around.”	
		#Supply_chain_route_optimalisation	“Supply chain or delivery route optimisation is another business process that is benefitting heavily from big data analytics. Here, GPS and sensors are used to track goods or delivery vehicles and optimise routes by integrating live traffic data”	
Effectiveness	**Higher margins**	**#High_ROI**	“Recent research has found that investing in data and analytics capabilities has high returns, on average: firms can use these capabilities to achieve productivity gains of 6 to 8 percent, which translates into returns roughly doubling their investment within a decade. “	(Henke et al, 2016)
		#Outperforming_peers	“We found that organisations that strongly agreed that *the use of business information and analytics differentiates them within their industry* were twice as likely to be top performance as lower performance. Top performance approach business operations differently than their peers do. Specifically, they put analytics to use in the widest possible range of decisions, large and small. They were twice as likely to use analytics to	(LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011)

			guide *future strategies*, and twice as likely to *use insights* to guide *day-to-day operations*." "The correlation between performance and analytics-driven management has important implications to organizations, whether they are seeking growth, efficiency or competitive differentiation."	
	Increased speed of decision making	**#Increased_clock_speed**	"Historically, computer designers have been increasing the clock speed at which the computer operates. Similarly, organization designers need to increase their organizations' clock speed. Units such as advertising, customer management, new product development, and supply chain management have to synchronize around increasing clock speeds. The ultimate target is the making of decisions in real time."	(Galbraith, 2014)
		#Real_time_customer_analytics	"As a strategic priority, real-time customer analytics is through the roof," says Geoffrey Moore, business consultant and author of Zone to Win: Organizing to Compete in an Age of Disruption. "The early adopters bought in entirely, built monster capabilities, and are getting results." Leading	(Harvard Business Review, 2018)

			digital companies have spent a decade fine tuning algorithms to offer relevant product recommendations that increase the likelihood of effectively upselling and cross-selling products at every touch point. “	

Exemplary quotes from existing literature have also been coded and interrelated in an attempt to explain how data analytics should be able to improve the effectiveness of strategic decision-making processes. The introduction of data analytics tools is expected to increase the effectiveness of strategic decision-making processes, because they are able to provide evidence based input (Michalkova & Bianchini, 2019) for the popularity assessments and value chain performance analysis performed by strategic decision-makers. The tools will allow decision-makers to perform data-driven (Deloitte, 2015), strengths, weaknesses, opportunities and threat analysis, because they are able to visualize historical patterns in trends and behavior. This is expected to increase the effectiveness of strategic decision-makers, because the visualisation of data will help decision-makers to eliminate biases and result in a higher ROI (Henke et al, 2016) (LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011). Furthermore, are data analytics tools able to provide real-time feedback on business performance (Adams, n.d.), which allows decision-makers to take early (corrective) actions in response. In addition to that, are visualised trends in patterns and behavior able to help strategic decision-makers predict in which direction they are heading (Marr, n.d.), which would allow them to respond quicker to emerging opportunities and threats (Galbraith, 2014).

References

Adams, A. (n.d.). Critical Success Factors for SMEs in Big Data Analytics. Retrieved from http://allisongadams.com/papers/Critical%20Success%20Factors%20for%20SMEs%20in%20Big%20Data%20Analytics.pdf

Ahmed, A., Bwisa, H., Otieno, R., & Karanja, K. (2014). *Strategic Decision Making: Process, Models, and Theories (ISSN 2157-6068)*. Macrothink Institute. https://doi.org/10.5296/bms.v5i1.5267

Ajah, I. A., & Nweke, H. F. (2019). *Big Data and Business Analytics: Trends, Platforms, Success Factors and Applications* (Vol. 3, No. 32). Big Data Cogn. Comput. https://doi.org/10.3390/bdcc3020032

Alzubi, J., Nayyar, A., & Kumar, A. (2018). *Machine Learning from Theory to Algorithms: An Overview*. IOP Publishing. https://doi.org/10.1088/1742-6596/1142/1/012012

Askham, N., Cook, D., Doyle, M., Fereday, H., Gibson, M., Landbeck, U., & Schwarzenbach, J. (2013). *The six primary dimensions for data quality assessment*. DAMA UK. Retrieved from https://www.whitepapers.em360tech.com/wp-content/files_mf/1407250286DAMAUKDQDimensionsWhitePaperR37.pdf

Awwad, M., Kulkarni, P., Bapna, R., & Marathe, A. (2018). *Big Data Analytics in Supply Chain: A Literature Review*. IEOM Society International. Retrieved from https://www.researchgate.net/profile/Mohamed_Awwad3/publication/327979282_Big_Data_Analytics_in_Supply_Chain_A_Literature_Review/links/5be1fc824585150b2ba30566/Big-Data-Analytics-in-Supply-Chain-A-Literature-Review.pdf

Beheshti, M., Amoozad Mahdiraji, H., & Zavadskas, E. K. (2016). *STRATEGY PORTFOLIO OPTIMISATION: A COPRAS GMODM HYBRID APPROACH* (Vol. 15, No 3C). Transformations in Business & Economics. Retrieved from https://www.researchgate.net/publication/314231540_Strategy_Portfolio_Optimization_A_COPRAS_G-MODM_Hybrid_Approach

Bianchini, M., & Michalkova, V. (2019a). *Data Analytics in SMEs: Trends and Policies* (No. 15). Committee on Industry, Innovation and Entrepreneurship and the Working Party on SMEs and Entrepreneurship of the OECD. https://doi.org/10.1787/f493861e-en

Bijl, L., Kringhaug, G., Molnár, P., & Sandvik, E. (2016). *Google searches and stock returns* (Vol. 45). International Review of Financial Analysis. https://doi.org/10.1016/j.irfa.2016.03.015

Bokdam, A. (2019). Why SMEs lag behind in big data. Retrieved from https://pure.tue.nl/ws/portalfiles/portal/122437614/Thesis_Abel_Bokdam_Definitief.pdf

Boone, T., Ganeshan, R., & Hicks, R. L. (2015). *Incorporating Google Trends Data Into Sales Forecasting*. Foresight. Retrieved from https://static1.squarespace.com/static/5b9e942a8f5130f854dbef81/t/5be9a6208985835f898e3a0f/1542039073203/incorporating-google-trends-data-into-sales-forecasting.pdf

Boyd, D., & Crawford, K. (2012). *Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon* (Vol 15. No. 5). Information, Communication, & Society. https://doi.org/10.1080/1369118X.2012.678878

Cao, L. (2017). *Data Science: A Comprehensive Overview* (Vol. 50, No. 3). ACM Computing Surveys. https://doi.org/10.1145/3076253

Cambridge Academic Content Dictionary. (n.d.). effectiveness. Retrieved from https://dictionary.cambridge.org/dictionary/english/effectiveness

Cambridge University Press. (n.d.). Data analysis. Retrieved from https://dictionary.cambridge.org/dictionary/english/data-analysis

Cambridge University Press. (n.d.). evaluation. Retrieved from https://dictionary.cambridge.org/dictionary/english/evaluation

Coleman, S., Göb, R., Manco, G., Pievatolo, A., Tort-Martorell, X., & Seabra Reis, M. (2016). *HOW CAN SMES BENEFIT FROM BIG DATA? CHALLENGES AND A PATH FORWARD* (Vol 32. No. 6). Quality and reliability engineering international. https://doi.org/10.1002/qre.2008

Cohn, C. (2014). Build vs. Buy: How to Know When You Should Build Custom Software Over Canned Solutions. Forbes. Retrieved from https://www.forbes.com/sites/chuckcohn/2014/09/15/build-vs-buy-how-to-know-when-you-should-build-custom-software-over-canned-solutions/

Confetto, M. G., & Siano, A. (2018). *Social Media Content: A Management Framework* (Vol. 13, No. 6). International Journal of Business and Management. https://doi.org/10.5539/ijbm.v13n6p84

Collinson, S., & Jay, M. (2012). *From Complexity to Simplicity*. Retrieved from https://www.macmillanihe.com/resources/sample-chapters/9781137006219_sample.pdf

Cuzzocrea, A., Bellatreche, L., & Song, I. (2013). *Data Warehousing and OLAP over Big Data: Current Challenges and Future Research Directions*. International Conference on Information and Knowledge Management. https://doi.org/10.1145/2513190.2517828

Davenport, T. H., & Dyché, J. (2013). *Big Data in Big Companies*. International Institute for Analytics. Retrieved from https://docs.media.bitpipe.com/io_10x/io_102267/item_725049/Big-Data-in-Big-Companies.pdf

De Aguiar Monteiro, L., Rodrigues, R., Cavalcanti De Lima, A., & Escobar, F. (2019). Survey on Trends in Big Data: Data Management, Integration and Cloud Computing Environment. Retrieved from https://www.researchgate.net/publication/332240604_Survey_on_Trends_in_Big_Data_Data_Management_Integration_and_Cloud_Computing_Environment

Dedić, N., & Stanier, C. (2017). *Towards Differentiating Business Intelligence, Big Data, Data Analytics and Knowledge Discovery*. International Conference on Enterprise Resource Planning Systems. Retrieved from https://link.springer.com/chapter/10.1007/978-3-319-58801-8_10

Deloitte. (2015). From gut feeling to data-driven decision-making. Retrieved from https://www2.deloitte.com/content/dam/Deloitte/nl/Documents/deloitte-analytics/deloitte-nl-data-analytics-todays-organizational-challenge.pdf

Dittrich, J., & Quiane-Ruiz, J. (2012). *Efficient Big Data Processing in Hadoop MapReduce* (Vol. 5, No. 12). Proceedings of the VLDB Endowment. Retrieved from http://vldb.org/pvldb/vol5/p2014_jensdittrich_vldb2012.pdf

Domo Inc. (2019). Domo Releases Seventh Annual “Data Never Sleeps” Infographic. Retrieved from https://www.bloomberg.com/press-releases/2019-07-09/domo-releases-seventh-annual-data-never-sleeps-infographic

DragonPoint, Inc. (2013). Custom Software vs. Off the Shelf. Retrieved from http://www.dragonpoint.com/wp-content/uploads/2016/06/guide-to-choosing-the-right-software.pdf

El Hussein, M., Hirst, S., Salyers, V., & Osuji, J. (2014). *Using Grounded Theory as a Method of Inquiry: Advantages and Disadvantages* (Vol. 19). The Qualitative Report. Retrieved from https://www.researchgate.net/publication/268812156_Using_Grounded_Theory_as_a_Method_of_Inquiry_Advantages_and_Disadvantages

Ferreira, N. M. (2019). How to Use Google Trends: 10 Mind-Blowing Tricks for Entrepreneurs. Retrieved from https://www.oberlo.com/blog/google-trends

Galbraith, J. R. (2014). *ORGANIZATION DESIGN CHALLENGES RESULTING FROM BIG DATA*. Aarhus, Denmark: Organizational Design Community. https://doi.org/10.7146/jod.8856

García, S., Luengo, J., & Herrera, F. (2015). *Intelligent Systems Reference Library* (Vol. 72). https://doi.org/10.1007/978-3-319-10247-4

Gartner. (n.d.). Gartner Glossary. Retrieved from https://www.gartner.com/en/information-technology/glossary/data-and-analytics

Gottlieb, J., & Rifai, K. (2017). Fueling growth through data monetization. Retrieved from https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/fueling-growth-through-data-monetization

Grant, R. M., & Jordan, J. (2015). *Foundations of Strategy* (2nd ed.). Retrieved from https://books.google.nl/books?hl=nl&lr=&id=uCFiBwAAQBAJ&oi=fnd&pg=PR9&dq=Grant+Jordan+2015+&ots=IIja3KPlNi&sig=RDyrWRA_F8i8WX-_2tQX28DzCrw#v=onepage&q=Grant%20Jordan%202015&f=false

Grünig, R., & Kühn, R. (2015). *The Strategy Planning Process*. Heidelberg, Germany: Springer. https://doi.org/10.1007/978-3-662-45649-1

Harvard Business Review. (2018). The Key to Unlocking Customer Insights & Driving the Customer Experience. Retrieved from https://www.sas.com/content/dam/SAS/en_us/doc/whitepaper2/real-time-analytics-109676.pdf

Henke, N., Bughin, J., Chui, M., Manyika, J., Saleh, T., Wiseman, B., & Sethupathy, G. (2016). THE AGE OF ANALYTICS: COMPETING IN A DATA-DRIVEN WORLD. Retrieved from https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Analytics/Our%20Insights/The%20age%20of%20analytics%20Competing%20in%20a%20data%20driven%20world/MGI-The-Age-of-Analytics-Full-report.ashx

Hoerl, R. W., & Redman, T. C. (2019). Most Analytics Projects Don’t Require Much Data. Retrieved from https://hbr.org/2019/10/most-analytics-projects-dont-require-much-data

Hurwitz, J., & Kirsch, D. (2018). *Machine Learning For Dummies*. Hoboken, US: Wiley & Sons. https://doi.org/10.1007/978-3-662-45649-1

Hurwitz, J., Nugent, A., Halper, F., & Kaufman, M. (n.d.). The Role of Traditional ETL in Big Data. Retrieved July 22, 2020, from https://www.dummies.com/programming/big-data/data-science/the-role-of-traditional-etl-in-big-data/

IBM. (2016). Analytics Solutions Unified Method. Retrieved from ftp://ftp.software.ibm.com/software/data/sw-library/services/ASUM.pdf

Ideya. (2018). Social Media Monitoring Tools and Services Report Public Excerpts 2018. Retrieved from http://ideya.eu.com/images/SMMTools%20Excerpts%202018%20Final.pdf

Iseminger, D., & Sharkey, K. (2020). Data sources in Power BI Desktop. Retrieved from https://docs.microsoft.com/en-us/power-bi/connect-data/desktop-data-sources

ISO. (2015). *Quality management systems — Requirements* (ISO/FDIS 9001:2015). International Organization for Standardization. Retrieved from https://groupe.afnor.org/produits-editions/bivi/FDIS%20ISO%209001E.pdf

Järvinen, P., Puolamäki, K., Siltanen, P., & Ylikerälä, M. (2009). *Visual Analytics* (117). VTT. Retrieved from http://users.ics.aalto.fi/kaip/online-papers/978-951-38-7178-9.pdf

Jeble, S., Kumari, S., & Patil, Y. (2018). *Role of Big Data in Decision Making* (Vol. 11, No. 1). Operations and Supply Chain Management: An International Journal. https://doi.org/10.31387/oscm0300198

Kagiyama, N., Shrestha, S., Farjo, P. D., & Sengupta, P. P. (2019). *Artificial Intelligence: Practical Primer for Clinical Research in Cardiovascular Disease* (Vol 8, Issue 17). Journal of the American Heart Association. https://doi.org/10.1161/JAHA.119.012788

Keim, D., Andrienko, G., Fekete, J., Görg, C., Kohlhammer, J., & Melançon, G. (2008). *Visual Analytics: Definition, Process, and Challenges*. Springer. https://doi.org/10.1007/978-3-540-70956-5_7

Khezr, S. N., & Navimipour, N. J. (2017). *MapReduce and Its Applications, Challenges, and Architecture: a Comprehensive Review and Directions for Future Research* (Vol. 15, No. 3). Journal of Grid Computing. https://doi.org/10.1007/s10723-017-9408-0

Kisielnicki, J., & Misiak, A. M. (2017). *EFFECTIVENESS OF AGILE COMPARED TO WATERFALL IMPLEMENTATION METHODS IN IT PROJECTS: ANALYSIS BASED ON BUSINESS INTELLIGENCE PROJECTS* (Vol. 9, Issue 1.). https://doi.org/10.1515/fman-2017-0021

Kohlhammer, J., Keim, D., Pohl, M., Santucci, G., & Andrienko, G. (2011). *Solving Problems with Visual Analytics* (Vol. 7). The European Future Technologies Conference and Exhibition. https://doi.org/10.1016/j.procs.2011.12.035

Laranjeiro, N., Nur Soydemir, S., & Bernardino, J. (2015). *A Survey on Data Quality: Classifying Poor Data*. 21st Pacific Rim International Symposium on Dependable Computing. https://doi.org/10.1109/PRDC.2015.41

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). *Big Data, Analytics and the Path From Insights to Value* (Vol. 52, No. 2). https://tarjomefa.com/wp-content/uploads/2017/08/7446-English-TarjomeFa.pdf. Retrieved from https://tarjomefa.com/wp-content/uploads/2017/08/7446-English-TarjomeFa.pdf

Leard. (2012). Principles of research ethics. Retrieved from http://dissertation.laerd.com/principles-of-research-ethics.php

Lewis-Pierre, L. J., Kovacich, J., & Amankwaa, L. (2017). *The Application of Grounded Theory: An Example from Nursing Workforce Research* (Vol. 22, No. 5). The Qualitative Report. Retrieved from https://nsuworks.nova.edu/tqr/vol22/iss5/7

Ong, I. L., Siew, P. H., & Wong, S. F. (2011). *A Five-Layered Business Intelligence Architecture* (2011). IBIMA Publishing. https://doi.org/10.5171/2011.695619

Liu, K., & Shi, J. (2015). *A Systematic Approach for Business Data Analytics with a Real Case Study* (Vol. 2, No. 4). International Journal of Business Analytics. Retrieved from https://pdfs.semanticscholar.org/5607/e98d75285c78c398c33a461811824a085f6d.pdf

Marín-Ortega, P. M., Dmitriyev, V., Abilov, M., & Gómez, J. M. (2014). *ELTA: New Approach in Designing Business Intelligence Solutions in Era of Big Data* (Vol. 16). Elsevier. https://doi.org/10.1016/j.protcy.2014.10.015

Marr, B. (n.d.). How Can Small Businesses Use Big Data? Here Are 6 Practical Examples. Retrieved from https://www.bernardmarr.com/default.asp?contentID=1442

MathWorks. (n.d.). Machine Learning in MATLAB. Retrieved from https://www.mathworks.com/help/stats/machine-learning-in-matlab.html

MathWorks. (n.d.). Plotconfusion. Retrieved from MathWorks. (n.d.). Plotconfusion. Retrieved from https://www.mathworks.com/help/stats/confusionmat.html

Mbassegue, P., Escandon-Quintanilla, M., & Gardoni, M. (2018). Knowledge Management and Big Data: Opportunities and Challenges for Small and Medium Enterprises (SME). HAL. Retrieved from https://hal.inria.fr/hal-01699732/document

McGee, J. (2014). *value chain* (3rd ed., Vol. 12). Retrieved from https://www.researchgate.net/publication/280246631_value_chain

Merendino, A., Dibb, S., Meadows, M., Quinn, L., Wilson, D., Simkin, L., & Canhoto, A. (2018). *Big data, big decisions: The impact of big data on board level decisionmaking* (Vol. 93). Journal of Business Research. https://doi.org/10.1016/j.jbusres.2018.08.029

Microsoft. (2019). How digitizing paperwork and client information has increased customer confidence and satisfaction and enabled strong growth. Retrieved from https://customers.microsoft.com/en-us/story/756790-multideur-dynamics365-office365-powerbi-netherlands

Microsoft. (2019). Machine Learning Algorithm Cheat Sheet for Azure Machine Learning designer. Retrieved from https://docs.microsoft.com/en-us/azure/machine-learning/algorithm-cheat-sheet

Mitchell, G., & Noble, H. (2016). *What is grounded theory?* (Vol. 19, No. 2). Evid Based Nurs. https://doi.org/10.1136/eb-2016-102306

Muller, P., Julius, J., Herr, D., Koch, L., Peycheva, V., & McKiernan, S. (2017). *ANNUAL REPORT ON EUROPEAN SMEs 2016/2017*. European Union. https://doi.org/10.2873/742338

Nahum, N., & Carmeli, A. (2019). *Leadership style in a board of directors: implications of involvement in the strategic decision‐making process*. Springer Science+Business Media. https://doi.org/10.1007/s10997-019-09455-3

National Minority AIDS Council. (2015). *Strategic planning*. Retrieved from http://www.nmac.org/wp-content/uploads/2015/04/Strategic-Planning.pdf

National Science and Technology Council. (2017). *FRONTIERS OF VISUALIZATION II: DATA WRANGLING WORKSHOP SUMMARY*. United States of America. Retrieved from https://www.whitehouse.gov/wp-content/uploads/2018/07/NITRD-FOVII-Workshop-Report-20180711.pdf

National Treasury of South Africa. (2010). *Framework for Strategic Plans and Annual Performance Plans*. Retrieved from http://www.treasury.gov.za/publications/guidelines/SP%20APP%20Framework.pdf

Okoli, J., & Watt, J. (2018). *Crisis decision-making: the overlap between intuitive and analytical strategies* (Vol. 56, No. 5). Management Decision. https://doi.org/10.1108/MD-04-2017-0333

Palacio-Nino, J., & Berzal, F. (2019). Evaluation Metrics for Unsupervised Learning Algorithms. Retrieved from https://arxiv.org/pdf/1905.05667.pdf

Papachristodoulou, E., Koutsaki, M. & Kirkos, E. (2017) *Business intelligence and SMEs: Bridging the gap*. Journal of Intelligence Studies in Business. (Vol. 7, No. 1). https://doi.org/10.37380/jisib.v7i1.216

Patel, P. M., & Deshpande, V. A. (2017). *Application Of Plan-Do-Check-Act Cycle For Quality And Productivity Improvement - A Review* (Vol. 5, No. 1). International Journal for Research in Applied Science & Engineering Technology. Retrieved from https://www.researchgate.net/profile/Vivek_Deshpande4/publication/318743952_Application_Of_Plan-Do-Check-Act_Cycle_For_Quality_And_Productivity_Improvement-A_Review/links/597ae3b00f7e9b0469e78636/Application-Of-Plan-Do-Check-Act-Cycle-For-Quality-And-Productivity-Improvement-A-Review.pdf

Polkowski, Z., Khajuria, R., & Rohadia, S. (2017). *BIG DATA IMPLEMENTATION IN SMALL AND MEDIUM ENTERPRISES IN INDIA AND POLAND* (Vol. 16). Scientific Bulletin – Economic Sciences. Retrieved from https://www.researchgate.net/publication/324965898_BIG_DATA_IMPLEMENTATION_IN_SMALL_AND_MEDIUM_ENTERPRISES_IN_INDIA_AND_POLAND

Popescu, M., & Dascălu, A. (2011). *VALUE CHAIN ANALYSIS IN QUALITY MANAGEMENT CONTEXT* (Vol. 4 (53), No. 2). Bulletin of the Transilvania University of Braşov. Retrieved from http://webbut.unitbv.ro/BU2011/Series%20VII/BULETIN%20VII/20%20popescu%20dascalu%20BUT%202011%202.pdf

Pratt, M. K., & Fruhlinger, J. (2019). What is business intelligence? Transforming data into business insights. Retrieved from https://www.cio.com/article/2439504/business-intelligence-definition-and-solutions.html

Puik, E., & Ceglarek, D. (2015). *The quality of a design will not exceed the knowledge of its designer; an analysis based on Axiomatic Information and the Cynefin Framework* (Vol. 34). 9th International Conference on Axiomatic Design. https://doi.org/10.1016/j.procir.2015.07.040

Purwono, B. S. A., Rahbini, R., & Suyanta. (2013). *Biogas Digester as an Alternative Energy Strategy in the Marginal Villages in Indonesia* (32). Energy Procedia. https://doi.org/10.1016/j.egypro.2013.05.018

Raj, R., Ha Sylvia Wong, S., & Beaumont, A. J. (2016). *Business Intelligence Solution for an SME: A Case Study* (Vol. 3). Proceedings of the 8th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2016). https://doi.org/10.5220/0006049500410050

Monica, N., & Ramesh Kumar, K. (2013). *Survey on Big Data by Coordinating Mapreduce to Integrate Variety of Data* (Vol. 2, No. 11). International Journal of Science and Research. Retrieved from https://www.researchgate.net/publication/266557067_Survey_on_Big_Data_by_Coordinating_Mapreduce_to_Integrate_Variety_of_Data

Richardson, J., Sallam, R., Schlegel, K., Kronz, A., & Sun, J. (2020). 2020 Gartner Magic Quadrant for Analytics and Business Intelligence Platforms. Retrieved from https://info.microsoft.com/ww-landing-2020-gartner-magic-quadrant-for-analytics-and-business-intelligence.html?LCID=EN-US

Ringel, M., Taylor, A., & Zablit, H. (2015). The Rising Need for Innovation Speed. Retrieved from https://www.bcg.com/publications/2015/growth-lean-manufacturing-rising-need-for-innovation-speed.aspx

Rouse, M., & Hughes, A. (2019). data mining. Retrieved from https://searchsqlserver.techtarget.com/definition/data-mining

Rouse, M., & Stedman, C. (2019). data analytics (DA). Retrieved from https://searchdatamanagement.techtarget.com/definition/data-analytics

Rouse, M., & Wigmore, I. (2014). small data. Retrieved from https://whatis.techtarget.com/definition/small-data

Sacha, D., Stoffel, A., Stoffel, F., Chul Kwon, B., Ellis, G., & Keim, D. A. (2014). *Knowledge Generation Model for Visual Analytics* (Vol. 20, No. 12). IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS. https://doi.org/10.1109/TVCG.2014.2346481

Sadiku, M. N. O., Shadare, A. E., Musa, S. M., & Akujuobi, C. M. (2016). *DATA VISUALIZATION* (Vol 2, No. 12). International Journal of Engineering Research And Advanced Technology. Retrieved from https://www.researchgate.net/publication/311597028_DATA_VISUALIZATION

Saltz, J. S., Shamshurin, I., & Crowston, K. (2017). *Comparing Data Science Project Management Methodologies via a Controlled Experiment*. Proceedings of the 50th Hawaii International Conference on System Sciences. Retrieved from http://hdl.handle.net/10125/41273

SAS. (n.d.). Big Data Analytics. Retrieved from https://www.sas.com/en_us/insights/analytics/big-data-analytics.html

Salesforce. (n.d.). Speak to customers with the right message, in the right tone, at the right time. Retrieved from https://www.salesforce.com/products/marketing-cloud/overview/?mc=sfmc

Schaeffer, D. M., & Olson, P. C. (2014). *Big Data Options For Small And Medium Enterprises* (Vol. 18, No. 1.). Review of Business Information Systems. https://doi.org/10.19030/rbis.v18i1.8542

Schmidt, J. C., & Laycock, M. (n.d.). Theories of strategic planning. Retrieved from https://www.healthknowledge.org.uk/public-health-textbook/organisation-management/5d-theory-process-strategy-development/strategic-planning

Sen, D., Ozturk, M., & Vayvay, O. (2016). *An Overview of Big Data for Growth in SMEs* (235). 12th International Strategic Management Conference. https://doi.org/10.1016/j.sbspro.2016.11.011

Shafique, U., & Qaiser, H. (2014). *A Comparative Study of Data Mining Process Models (KDD, CRISP-DM and SEMMA)* (Vol. 12). International journal of innovation and scientific research. Retrieved from https://pdfs.semanticscholar.org/bd13/74cd5f7ee859da6fdf32ea4e646969e90334.pdf

Singh, J., & Sood, A. (2013). *A Comparative Study of OLTP and OLAP Technologies* (Vol. 4, No. 2). International Journal of Computer Science and Technology. Retrieved from https://pdfs.semanticscholar.org/1a32/2adb73a3dfe5fea48b0a05108538a550704d.pdf

Sun, Z., & Huo, Y. (2019). *The Spectrum of Big Data Analytics*. Journal of Computer Information Systems. https://doi.org/10.1080/08874417.2019.1571456

Tableau. (2013). Making Sense of Your Organization's Big Data. Retrieved from https://www.tableau.com/sites/default/files/whitepapers/making-sense-big-data.pdf

Tableau. (2020). Tableau Blueprint Overview. Retrieved from https://help.tableau.com/current/blueprint/en-us/bp_overview.htm

Thomas, G. (2017). *How to do your research project* (3rd ed.). Thousand Oaks, US: SAGE Publications.

Tjahjono, B., Esplugues, C., & Ares, E. (2017). *What does Industry 4.0 mean to Supply Chain?* (Vol. 13). Manufacturing Engineering Society International Conference. https://doi.org/10.1016/j.promfg.2017.09.191

United Nations. (n.d.). *STRATEGIC PLANNING Guide for Managers*. Retrieved from https://hr.un.org/sites/hr.un.org/files/4.5.1.6_Strategic%20Planning%20Guide_0.pdf

Umoh, R. (2018). LEADERSHIP Steve Jobs attributed his success to intuition — here's how to use yours. Retrieved from https://www.cnbc.com/2018/05/23/how-to-trust-your-intuition-like-apple-co-founder-steve-jobs.html

van Wesel, P., & Goodloe, A. E. (2017). *Challenges in the Verification of Reinforcement Learning Algorithms*. National Aeronautics and Space Administration. Retrieved from https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20170007190.pdf

Varela Rozados, I., & Tjahjono, B. (2014). *BIG DATA ANALYTICS IN SUPPLY CHAIN MANAGEMENT: TRENDS AND RELATED RESEARCH*. International Conference on Operations and Supply Chain Management. Retrieved from https://www.researchgate.net/publication/270507070_OSCM-2014

Verdin, P., & Tackx, K. (2015). *Are you creating or capturing value?* (No. 36). Mossavar-Rahmani Center for Business and Government. Retrieved from https://www.hks.harvard.edu/sites/default/files/centers/mrcbg/files/Verdin_final.pdf

Wall, E., Blaha, L. M., Franklin, L., & Endert, A. (2017). *Warning, Bias May Occur: A Proposed Approach to Detecting Cognitive Bias in Interactive Visual Analytics*. IEEE Conference on Visual Analytics Science and Technology. https://doi.org/10.1109/VAST.2017.8585669

Wang, G., Gunasekaran, A., Ngai, E. W. T., & Papadopoulos, T. (2015). *Big Data Analytics in Logistics and Supply Chain Management: Certain Investigations for Research and Applications*. International Journal of Production Economics. https://doi.org/10.1016/j.ijpe.2016.03.014

Wang, Y., Kung, L., & Byrd, T. A. (2016). *Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations*. Technological Forecasting & Social Change. https://doi.org/10.1016/j.techfore.2015.12.019

Williams, M., & Moser, T. (2019). *The Art of Coding and Thematic Exploration in Qualitative Research* (Vol. 15, No. 1). International Management Review. Retrieved from https://pdfs.semanticscholar.org/5dd2/51ddfbb6a563b900e53a9b3476a8c4b2557b.pdf

Zheng, J. G., & Peltsverger, S. (2015). *Web Analytics Overview* (3). Encyclopedia of Information Science and Technology. Retrieved from https://www.researchgate.net/publication/272815693_Web_Analytics_Overview

Zubova, J., Kurasova, O., & Liutvinavičius, M. (2018). *Dimensionality Reduction Methods: The Comparison Of Speed And Accuracy* (Vol. 47, No. 1.). Journal of Information Technology and Control. https://doi.org/10.5755/j01.itc.47.1.18813