

**A Work Project, presented as part of the requirements for the Award of a Master in
Management Degree from NOVA – School of Business and Economics**

FlixBus: Toward the Development of Predictive Analytics

Noé Paúl de Jesús López Márquez - #27459

A Project carried out on the Master in Management Program, under the supervision of:

Leid Zejnilovic

LISBON

JUNE 2018

Abstract

Data is considered today by many the new oil of the century. As the amounts of data generated increase at fast pace, so does the challenges faced by companies. With new tools and technology firms try to keep up with the ever changing nature of data in order to gain an advantage within the competitive landscape. Nevertheless there is still a long road to completely understand all data related topics. Marketing is one of the main activities that has evolved and acquired a strong data focus. This paper examines how companies can better understand data and generate more effective strategies in order to incorporate analytics into the marketing mix. The study provides a better understanding of the main data definitions and develops a case study about FlixBus a global mobility leader, which offers a framework to assess big data capabilities, such as the main sources of data, the different analytical methods and its applications and finally the main challenges that might be encountered by companies.

Keywords: Big Data, Big Data Analytics, Big Data Methods, Big Data Challenges, Descriptive Analytics, Predictive Analytics, Prescriptive Analytics, Big Data Analytics Capability.

Agenda

1. Introduction.....	4
1.1 Conducting the Research Question.....	5
2. Big Data.....	6
2.1 Defining Big Data.....	6
2.2 Big Data Analytics Methods.....	7
2.3 Big Data Challenges.....	8
3. Methodology.....	9
3.1 Data Collected.....	10
4. FlixBus Case Study.....	11
4.1 The Origins of FlixBus.....	12
4.2 Current State of Analytics at FlixBus.....	13
4.3 Big Data Challenges Face by FlixBus.....	15
4.3.1 Data Challenges.....	16
4.3.2 Process Challenges.....	17
4.3.3 Management Challenges.....	18
4.4 Creating a Data Strategy.....	20
4.5 Findings.....	20
5. A Marketing Mix Framework for Big Data Management.....	21
6. Discussion.....	23
7. Limitations and Direction for Further Research.....	24

1.- Introduction

In recent years the amount of data generated by business has increased enormously, every day around 2.5 quintillion bytes of data are produced in the world (1 exabyte equals 1 quintillion bytes or 1 exabyte equals 1 billion gigabytes) and 90% of these data is unstructured (Dobre and Xhafa, 2014). By 2020 it is expected that more than 40 Zettabytes (40 trillion gigabytes) of data will have been generated in the world (Gantz and Reinsel 2012). With data being generated from any-where, any-time, and any-device, current technologies are being outpaced quickly by the volumes of big data, making established data processing tools inadequate, for example data warehouses and databases (Sivarajah, 2016). Nevertheless, companies have switched its efforts in order to become data-driven at different level of businesses, big data has become one of the top critical trends for business success and competitive advantage (Savitz, 2012b), categorized by many as today's Digital Oil (Yi, Liu, Liu, & Jin, 2014) and the new raw material of the 21st century (Berners-Lee & Shadbolt, 2011)

Thanks to new BD analysing technologies like NoSQL Databases, BigQuery and Hadoop better insights can be attained in order to improve business strategies and decision-making processes in critical sectors and activities (Yi et al., 2014). Nevertheless, BD is still on an embryonic stage where its economic benefits are still uncertain, however there has been a strong investment made by companies in order to try to capture more value and leverage on big data analytics capabilities (BDAC), this investments were around 2.1 US trillion in 2013 (Lunden, 2013). A recent study by accenture (Columbus, 2014) highlights that “87% of enterprises believe that BDA will redefine the competitive landscape of their industries within 3 years”.

In Marketing BD allows to integrate heterogeneous information from different sources and provides a holistic view in order to generate more accurate marketing intelligence and improve marketing decision-making (Fan, 2014).

1.1 Conducting the Research Question

Regardless of where BD is generated from and shared to, with the reality of BD comes the challenge of analysing it in a way that converts into big value. For fast growing companies and startups data management is quickly becoming a trending practice, although this process is perceived as a requirement to become more efficient, enhance strategic potential, spot new revenue streams and gain competitive advantage over rivals, companies must acknowledge the inherent risks, advantages and challenges that comes with BD before making great investments of resources on BD technology (Sivarajah, 2016). New strategies to approach specific marketing activities and its data needs must be created evaluated in order for companies to benefit the most from BDA.

Therefore, this study will focus on what big data is, its main challenges and the different types of big data analytics. Further, this study seeks to answer the following research question:

How can FlixBus overcome big data challenges and move toward predictive analytics in order to improve marketing activities?

To answer this question a case study about FlixBus, a global mobility company that is currently on the transition from descriptive to predictive analytics was conducted. Additionally this paper expand big data and marketing literature by developing a framework for companies to asses big data implementation, this framework incorporates big data analytics, data sources and its challenges into the marketing mix.

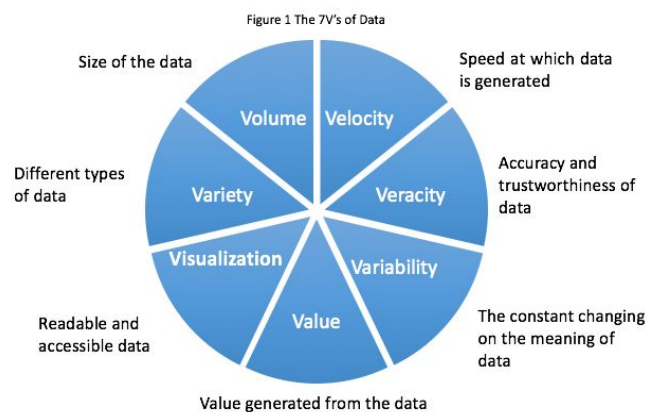
2. Big Data

The following section aims to provide a definition of big data and big data analytics methods as well as summarizing the main big data challenges faced by companies across all industries

2.1 Defining Big Data

Big data as a research discipline and as a concept it still evolving and not yet completely established, therefore a fully realized definition and classification of the concept is yet to be determined.

An online survey conducted by Harris Interactive and SAP of 154 global executives from small and midsize companies aiming to benefit from big data revealed how the term and its characteristics are understood differently across industries and individuals. The most common framework known in literature to define big data focus on 3 V's: volume, variety and velocity (Chen, Chiang, & Storey, 2012; Kwon, Lee, & Shin, 2014), however this model has been modified and expanded by several authors into 4, 6 and even 7 V's (Sivarajah, 2016) (fig. 1).



A simple definition of big data based on its characteristics and adapted from Gartner, Inc. by incorporating the 7 V's framework is: Big data is data with high volume, velocity, variety, variability, veracity, visualization and value that transforms into information assets that demand

cost-effective, innovative forms of information processing for enhanced insights and better decision making.

In order to understand better the term big data its main characteristics must be addressed. The 7 V's can be breakdown on a simple way in, volume which refers to the amount of data generated, with new trends such as IoT or hyper connected consumers, a high volume of data is generated every hour, with 90 percent of all data we have today was generated just in the past couple of years. Variety tell us about multiple data formats that can be structured, semi-structured or unstructured. Velocity emphasizes the high-rate of data inflow generated every second. The veracity dimension focus on the increasingly complexity of data structures anonymities, imprecision or inconsistency in large data-sets. Variability refer to the ever changing meaning of data. The presentation of data on an understandable and digestible way is defined as visualization. Finally, data must generate knowledge or value for end-users.

2.2 Big Data Analytics Methods

Big Data by itself does not provide a lot of value when it is unprocessed, therefore businesses need to create efficient processes and methods to extract value out of it. BDA methods can be classified on three main groups, descriptive analytics, predictive analytics and prescriptive analytics. (fig. 2)

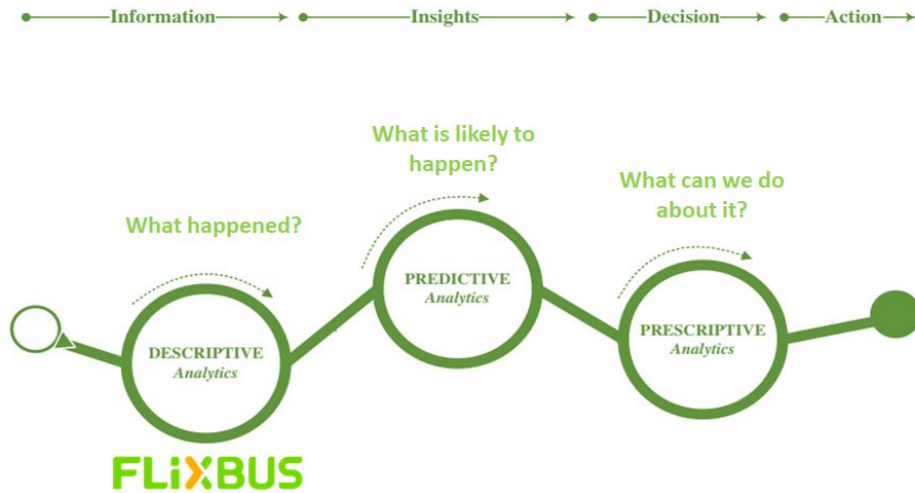


Figure 2 Types of Big Data Analytics and FlixBus Current State

Descriptive Analytics are considered the simplest form of BDA methods, its main goal is to summarize and describe knowledge patterns based on historical data by using simple statistical methods to understand the root cause of specific events, mainly answering the question, what has happened? On the other hand, predictive analytics aims to answer the question, what could happen? Based on forecasting and statistical models to determine future possibilities, normally using supervised, unsupervised or semi-supervised learning models. Finally, the most useful method of analytics is prescriptive analytics which by using optimization and simulation algorithms tries to advice on possible outcomes and tries to determine cause-effect relationships to answer the question, what should companies do?

2.3 Big Data Challenges

According to Akerkar (2014) and Zicari (2014) BD challenges can be broadly categorized into three main groups, based on the data specificities and its life-cycle, the data process and the management challenges (fig 3). Data challenges relate to the specific characteristics of data,

which can be summarized on the 7 V's. Process challenges cover the different techniques needed for data processing, such as collection, interpretation, modelling and results. Finally, Management challenges are related to privacy, security, ethical and governance aspects of data.

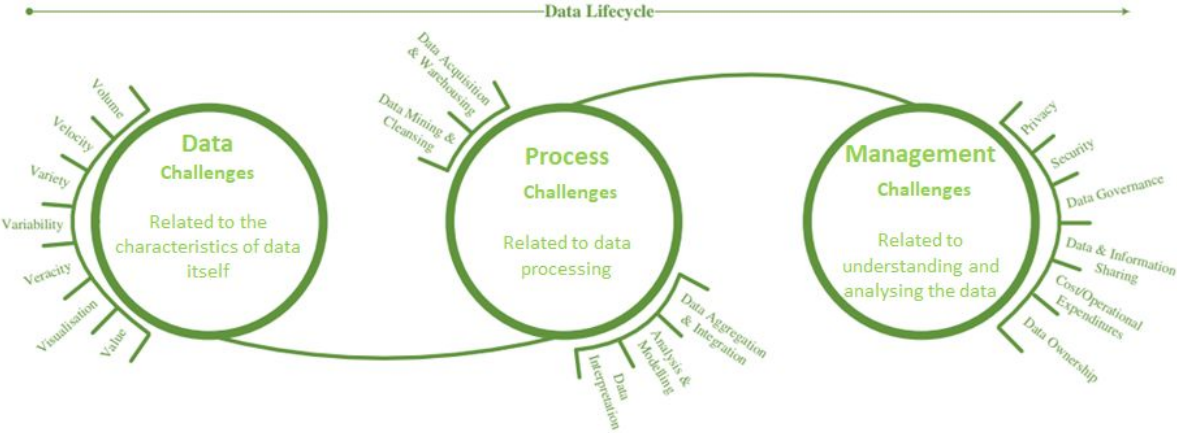


Figure 3 Big Data Challenge Taxonomy

3. Methodology

In order to answer the proposed research question, a qualitative research approach was selected. Since this work focus on the transition from prescriptive to predictive analytics, a case study is the most appropriate methodology to demonstrate how companies can better be prepared to use new big data analytics methodologies and its implications and benefits on the marketing mix. According to Baskarada (2014) a case study research differ from traditional survey based studies in a sense that case studies focus on capturing direct evidence rather than perceptions and attitudes toward specific events and behaviours. Furthermore, a difference from statistical sampling is that case studies utilize theoretical sampling, which does not undertake representative capture of all possible variations but tries to gain a more robust understanding of cases to facilitate development theories.

This case study involves an in-depth analysis of an individual unit, in this case FlixBus and follows the six interdependent stages proposed by Yin (2009) to build a case study research.

It starts with the planning stage where the research question and reason for doing the case must be addressed, outlined on section one. Afterwards on section two the design of the research and theoretical concepts are defined. The third step focus on gaining skills as an investigator and training for the specific case study. Finally, the last two steps focus on the collection of multiple sources of evidence and the analysis of such sources. In the context of case studies data generally consists of examining, tabulating, categorizing or a combination of evidence to draw empirical conclusions (Baskarada, 2014). On a case study research, the investigator evaluates and decodes the meaning of the data, in terms of what it may imply, this process is overtaken during and after the evidence is gathered, all elements are then coded with a classical content analysis approach. Lastly, a framework to identify the key data, methods, applications and challenges needed to improve the marketing mix was developed to help organizations to detect key resources for successful move toward more advanced analytics methods.

3.1 Data Collected

On a case study research, relevant data might be gathered through archival records, interviews, documents, direct observations and physical artefacts. However, it should be held in mind that all these elements may contain biases and thus not always have an accurate reflection of what reality is. The main source of data for this paper was collected through interviews with relevant stakeholders, such as data users and decision makers. Besides, direct observations concerned with the future objectives of the company and upcoming data strategies were gathered during monthly meetings, held to provide a big picture of these strategies across the whole company.

Finally, internal data from FlixBus such as statistics, documents, surveys, analytical tools and web information was used to gain a more holistic view of data utilization within the company.

Primary Data	
Interviews	Michael Amberger - Mobile Marketing Manager
	Marco Scabioli - Social Media Manager
	Tim Hiebenthal - Marketing Intelligence Manager
	Luis Berumen Martinez - Jr. Marketing Intelligence Manager
	Fely Monaco - Paid Search Manager
	Amalia Trivizadaki - CRM Manager
	Sepanta Sharafuddin - Campaign Manager
	Valentina Gravotta - Senior International Marketing
Observations	Data Privacy Conference - FlixBus HQ Munich 05.05.2018
	2018 FlixBus Challenges - FlixBus HQ Munich 04.02.2018
	Weekly Meetings within the Company
Secondary Data	
FlixBus Internal Documents	
FlixBus Statistics and Reports	
FlixBus Employee Surveys	
Web Information	
FlixBus Internal Analytical Tools	

4. FlixBus Case Study

The following case study is structured as follows: An introduction and brief history of FlixBus a global leader in the mobility industry which aim to adopt predictive analytics and use it to improve overall marketing activities. Afterwards the current state of analytics at FlixBus and the main big data challenges that the company faces will be addressed. Lastly, an evaluation of FlixBus data strategy will be performed.

“Analytics competitors understand that most business functions—even those, like marketing, that have historically depended on art rather than science—can be improved with sophisticated quantitative techniques. These organizations don’t gain advantage from one killer app, but rather from multiple applications supporting many parts of the business” - Davenport.

By 2018 FlixBus is Europe's biggest long-distance with 1,200 destinations and 200,000 daily connections, offering green mobility for over 70 million passengers in 26 countries across

Europe and with a coming expansion into the U.S. west coast, becoming its 27th market. Besides its core business of buses, FlixBus expanded its services into the rail industry with its first 2 trains on early 2018. With a unique combination of tech-start-up, e-commerce platform and transformation company, FlixBus aims to provide smart green mobility for everyone to experience the world. With the brand FlixBus the company started revolutionizing the bus travel industry and quickly became Europe's largest intercity bus network.

4.1 The Origins of FlixBus

FlixBus was founded in Germany in 2011 and started operations in 2013 after the liberalization of the German bus market in that same year. After consolidating its market leadership in Germany by merging with Meinfernbus FlixBus started its expansion in 2014 to Austria, Netherlands and Switzerland, since that the network has grown further and today includes countries like France, Italy and Scandinavian region, today being present in most European countries (exhibit 1). After five years of its founding FlixBus is market leader in Europe with 90% market share.

The fast growth of FlixBus was possible to its business model which is radically different from that of traditional bus operators, by using an Umbrella model, where instead of owning buses, FlixBus provides a business platform and takes over sales, marketing, network planning, customer service, among other activities while the bus partner companies overtake the operations. (FlixBus Internal Data, 2018).

As many young successful companies, such as Spotify, Netflix and Uber, FlixBus is a great representation of how data has become one of the most powerful assets in the current digital driven economy. In a surprisingly short time FlixBus became market leader and dominating

mobility firm in Europe without operating its own buses. Thanks to its business model and intelligent platform, the company was able to successfully create and sustain competitive advantage hard to imitate by competition. When you use data as a strategic asset you can adapt your business model and marketing activities accordingly, for FlixBus knowing customer travel preferences allow them to effectively adapt and model supply and demand of its network accordingly.

Nevertheless, due to the fast growth and expansion, the company has been struggling to keep up with the great amounts of data generated by customers, network, buses and other activities making it harder to utilize predictive analytics models and relying mainly on simpler methods. Therefore, creating more robust analytics models and make better use of the data collected has become one of the main challenges faced by the company and one of the top priorities to keep ahead of competitors. In its search of improving big data capabilities and keep its competitive advantage and fast growth FlixBus aims to leave behind its current model of descriptive analytics and implement predictive analytics specially into its online marketing strategy.

4.2 Current State of Analytics at FlixBus

In the past decades the volume of data available for firms have increased exponentially, gathering attention from companies on all industries and impacting the way decision making is done. Data-driven marketing activities have become more important for business in recent years, leading to better customer segmentation and acquisition, supply and demand forecast, helping in product development and so on.

At FlixBus, marketing and sales represent one of the most important value creation entities due to the nature of its business model and therefore huge investments are done in order to keep

developing new marketing strategies and remain ahead of competition, especially in online marketing which has strategic importance for the company and is one of the main sources of data (Exhibit 2).

The online marketing department counts with 6 different channels, SEO, SEA, Mobile, Social, Mail and Affiliate, which together with the website are the main sources of data acquisition for FlixBus. Each channel acquires and uses data for different purposes, but the main focus overall is customer acquisition and customer retention. SEA is one of the most relevant channels, by tracking search behaviour from users in different search engines like Google or Bing, they have been able to successfully help network planning by identifying search patterns. Regarding SEO although they team has been struggling in order to find data that creates value, they have been able to identify potential relevant content for users and transform it into business opportunities. Mobile marketing team is in charge mainly of all the website visits via mobile and of the management of the FlixBus App, the app is one of the most important resources for data creation and value added, this is due to the fact that customer acquire by app tend to have a better life time value and it allows for critical data acquisition that can later be used for better targeting such as location. Social media team is a big asset for the company for two reasons, first, the main target groups which are younger users and students spend most of their time scrolling down in their preferred social media and secondly there is an enormous potential for data acquisition and application in social media platforms, like geo targeting, sentiment analysis, customer profiling, potential target analysis and so on. Affiliate is in charge of finding possible partners such as websites, blogs, coupon sites and the like, and although there is great potential for collecting data of successful sites and optimize the process partner selection, there is still not much done on this

aspect. Finally, mail is able to collect user's data such as location and it is one of the most advanced channel on the use of data and analytics by creating dynamic network selection based on the last departure point and cohort based targeting, although there is still a huge potential for improvement, the team is already taking the first steps into moving toward predictive analytics.

However most of the marketing activities are undertaken with a descriptive approach, which represents the simplest type of big data analytics, where their main focus lays on summarization and description of patterns based on historical data, such as transactions, last departure point and type of device used. Nevertheless, this approach tends to be inaccurate and undermines the potential of most marketing activities, inquiring higher marketing costs for the company, specially during price promotions, where most of the campaigns tend to have a negative value. This also represents a threat against competitors in new markets, where having robust prediction models could mean for FlixBus to overtake competitors within less time and with lower resource investments in the different markets it operates and keep their fast pace growth without compromising its current market leadership and service provided.

FlixBus next step in order to sustain its aggressive expansion and market leadership is to leverage on the power of analytics and move from descriptive to predictive analytics.

4.3 Big Data Challenges Faced by FlixBus

“A companywide embrace of analytics impels changes in culture, processes, behaviour, and skills for many employees.” - Davenport

It was not long ago when FlixBus embraced for the first the need of a marketing intelligence entity, in order to meet its data needs. The Marketing Intelligence team was created the 15th of January 2015 with only one member Philipp J., the next year the team growth incorporating two

new members and as of today it consists of 5 full time positions and 2 interns. The main activities undertaken by the team are split on reporting infrastructures for the different department, such as data sets, building reports and dashboards, as well as providing consulting advice for the different teams regarding tool selections, analysis and campaign performances. Furthermore, the team is also in charge of providing input for marketing and the company management, such a revenue planning and forecasting or cohort creations.

There are three big issues that have not been addressed yet by FlixBus, the first one is related to data integration, the lack of a common database and data with the same structure is an impediment for some department to cooperate with each other or share data more easily. The second issue is lack of third party data integration which would allow the company to gain more insights of potential customers when the data was crossed with their current database. Finally the absence of a Chief Data Officer that is able to implement a coherent data strategy and aligns it with the overall company goals.

With a constraint on resources and human capital, the integration and analysis of the heterogeneous data sources has proven to be a difficult task, with a network operating in 26 countries there is an increasing need to build up robust IT infrastructure and to find the optimal approach given the constraints of labour force. In order to succeed on big data adoption and create big data analytics capabilities FlixBus faces the three types of challenges of big data: data characteristics, processing and management.

As previously seen in figure 3 each of this challenges can be broken down in smaller steps and each of them have a different impact on the company road to achieve predictive analytics.

4.3.1 Data Challenges

There are seven aspects about data that FlixBus must address which are volume, velocity, variety, variability, veracity, visualization and value. These 7 v's can be categorized into high priority and low priority from a company perspective, according to current capabilities and goals. For FlixBus it is important to address first the issues with high volumes of data as its network grows and the company keeps expanding into new markets so does the quantity of data acquired, therefore it is necessary to come up with novel approaches to manage big data problems. Velocity is another aspect that is considered extremely relevant due to the high rate of data inflow such as daily bookings and new customers. FlixBus also need to improve its current visualization strategy since it is the main component that allow key stakeholders to digest data and support its decision making process and at this current stage on time they still lack robust visualization tools, relying solely on simple graph techniques such as excel and PowerPoint. Finally comes value as the company struggles to extract knowledge and create value from this vast amount of structured and unstructured data.

Regarding variability, veracity and variety, these create less conflict for the company. Since there are not many different types of data, making variety less of a challenge, since all the data acquired is from a primary source and there is almost no third party data utilized also veracity tend to be more stable and finally variability still does not represent a big problem for the company on the current stage of analytics.

4.3.2 Process Challenges

For FlixBus process challenges can be categorized in 4 different groups. Firstly, data acquisition and warehousing, due to the lack of efficient algorithms that allow data process in order to clean and reduce data before storing. This leads to vast amounts of data that then have to be cleaned

and mined mostly manually which tend to be a long and non-automated process consuming most of the time of the marketing intelligence team. The main challenge regarding processes lays on data analysis and modelling which tend to be undertaken in excel sheets, as shown by Davenport spreadsheets and databases generated by users tend to lead to multiple versions of key indicators within an organization. Additionally, researchers have shown that approximately between 20% to 40% of spreadsheets contain mistakes, hence, the more spreadsheets generated by companies, the more opportunities for potential errors to arise. Added to this the use of multiple tools, lack of shared data among departments and decentralized data analysis creates different definitions and sources of truth, as each department utilizes different KPIs, analysis methods and tools, leading to several inconsistencies within the company data. Finally, data interpretation is an immense challenge, due to the vast amount of data processed and its different sources, it has clearly become an issue within the company to accurately digest and understand data.

4.3.3 Management Challenges

Typical big data management challenges relate to privacy, governance, security, ownership, cost and information sharing (Sivarajah, 2016). For FlixBus the main challenges centre on security, data and information sharing and governance, on the other hand the company has been able to manage well privacy and ownership, therefore these two aspects does not represent a big problem, while cost and operational expenditure lays in the middle.

Regarding security, the company needs to ensure a robust infrastructure addressing two main issues, internal security and external security. The first one must enable that different staff and employees has access only to relevant data for their department to avoid possible leaks from the inside and possible data mishandling. Furthermore, to ensure security of data against external

entities such as competitors must be overcome, especially due to the fact that the company has been the target of several phishing attacks through employee emails coming from competitors.

Data and information sharing have also proven to be some of the biggest weaknesses of the company. The lack of data and knowledge harmonisation and connection not only with business partners but also across different departments, is barely non-existent, with each department accountable for their own data and with little idea of what data other departments are collecting or using. Furthermore, this becomes even more challenging on offices in different countries, since most of the data management and marketing activities are centralized in the company headquarter in Munich, making it hard seamless information sharing and exchange across offices and departments.

As the company grows so the amounts of data collected, to ensure the quality of the data mined and analysed, data governance is key priority, for FlixBus selecting which data is warehoused, analysed, accessed and visualized is one of the biggest challenges, and the company is trying to approach with caution in order to not incur in unnecessary expenditure and possible data scandals. Also the current status of data governance makes data mining processes slow and does not warranty that all different data users comply with data quality standards.

At this point of time, cost does not represent a big challenge for the company, but it is something that should be looked carefully, with big data come big costs due to storage needs and data analysis processes and tools, hence cost minimization should be borne in mind.

Finally, ownership and privacy are of low impact for the company in the sense that they have been performing an excellent job, this is due to law enforcement regulations in Germany which

lead to the creation of data legal departments within the company which constantly make sure to comply with the highest standards in these two aspects.

4.4 Creating a Data Strategy

In order to overcome the major big data challenges and move toward predictive analytics the company finds itself in a difficult position that urges to identify the key data to collect and store and to better understand how it can improve its current data analysis and applications in order to generate value for the company. Furthermore, FlixBus should adapt a business process perspective when it comes to data rather than the current silo perspective. This would allow to achieve optimal value creation and reduce duplicated work by facilitating cross functional communication among departments and different entities, enhancing overall business processes. With a business process perspective focused on data all departments must work together to have a better understanding of the bigger pictures. Some of the disadvantages with the current approach is sub optimization, due to each department optimize data according to its own specific goals, this also leads to communication gaps and loss of information and key data, which finally results on losing sight of the bigger picture and a less effective strategy, the company could benefit by adding a Chief Data Officer responsible of develop information and analytics strategies aligned with business goals.

Furthermore, the company is on the process of implementing a new enterprise system with the hope to be able to leverage on their data and make a better use of it without the need of making a big expenditure, although some departments are still unsure of weather it will create more benefits or not, but top managers are confident this new approach will improve analysis effectiveness and will take them to the next level of analytics.

4.5 Findings

After the interpretation of all the data collected and its linkage to the relevant literature, the following key findings were determined. Regarding of the different data challenges the company currently faces, FlixBus is putting data on the core of most of its marketing activities and although the current strategy is not the optimal, the company acknowledges this and is working on the improvement of it. Related to each specific data challenges on the way to predictive analytics FlixBus has certain advantages and disadvantages depending on the type of challenge. Regarding data, the company counts with rich sources of data and data from which valuable output can be generated, nevertheless the weakest points of the company relates to the processes of data, specifically with the current approach for mining and modelling which still are on an early stage, nevertheless the company attempts to fix this by hiring new people to cope with the demanding needs of data and plan to implements a new enterprise system to aid on this process. Finally, management challenges are not at its best, however the company organizational culture and managerial approach are great assets that rather than be an obstacle to adopt predictive analytics are big agents of change and constantly seek to become more data oriented. Therefore, FlixBus plans would develop around concentrate its efforts on the improvement of data processes while taking advantage of its data and managerial elements.

5. A Marketing Mix Framework for Big Data Management

The case study has shown that on the road to predictive analytics companies face several challenges related to data characteristics, processes and management. The following framework was developed based from the findings on the case study and theory and it demonstrates the key elements that lead to more successful and smoother transition toward predictive analytics, while

incorporating these elements into the marketing mix (fig. 4). The framework classifies within the 5P model three critical features that must be considered for a successful data strategy, which are key data needed, analytical methods and applications and the main challenges that can be encountered. “Data” refers to the critical data that is generated by each of the marketing Ps and that is useful for the company according to its specific marketing goals. “Method & Application” is divided in two parts. First a split between descriptive and predictive techniques, the reason for this is that although the company goal is to achieve predictive analytics, descriptive analytics are still relevant and generate value in specific cases and therefore cannot be left aside. Secondly depending on the method there are different use cases that can be achieved with the data gathered. The last step “Challenges” deals with the main issues faced on each step of the marketing mix, divided on the 3 main big data challenges; data, processes and management challenges (Sivarajah, 2016).

	People	Product	Promotion	Price	Place	
Data	<ul style="list-style-type: none"> Demographics Social Networks Mail information Search Engine Behavior Payment Methods Usage Behavior/Patterns Survey Data 	<ul style="list-style-type: none"> Network Supply/Demand Load factor App Rating Downloads Bus Travel Data Trip Rating Survey Data 	<ul style="list-style-type: none"> Promotional Data (e.g. Sold Non-Promotion & Promotional Tickets, Campaigns Uplift, Booking Increase) Survey Data 	<ul style="list-style-type: none"> Transactional Data Survey Data 	<ul style="list-style-type: none"> Location-Based social networks Mobile GPS location Last departure/destination points Survey Data 	
Method & Application	Descriptive	<ul style="list-style-type: none"> Customer Segmentation Sentiment Analysis Customer Profiling 	<ul style="list-style-type: none"> Product Reputation 	<ul style="list-style-type: none"> Promotion Analysis 	<ul style="list-style-type: none"> Pricing Strategy Analysis Competitor Analysis 	<ul style="list-style-type: none"> Location Based Targeting
	Predictive	<ul style="list-style-type: none"> Potential customer targeting Cohort forecasting Increase targeting effectiveness CLV forecasting 	<ul style="list-style-type: none"> Network Planning Supply/Demand Forecasting Route Optimization 	<ul style="list-style-type: none"> Recommendation Systems Increase Promotion Value Reduce Customer Cannibalization Relevance of Promotion (context & Precision) 	<ul style="list-style-type: none"> Dynamic Pricing Forecast Willingness to Pay Optimal Price Prediction Budget Allocation 	<ul style="list-style-type: none"> Location Forecasting Destination Recommendation Personalization based on Location Increase offer relevance Timely Targeting
Challenges	Data	<ul style="list-style-type: none"> Volume Variety Veracity Variability 	<ul style="list-style-type: none"> Volume Variety Velocity Veracity 	<ul style="list-style-type: none"> Velocity Value Visualization 	<ul style="list-style-type: none"> Value Visualization 	<ul style="list-style-type: none"> Velocity Variety Visualization Veracity
	Process	<ul style="list-style-type: none"> Mining & Cleansing Analysis & Modeling Interpretation Aggregation & Integration 	<ul style="list-style-type: none"> Analysis & Modeling Aggregation & Integration 	<ul style="list-style-type: none"> Mining & Cleansing Analysis & Modeling Interpretation Aggregation & Integration 	<ul style="list-style-type: none"> Acquisition & Warehousing Mining & Cleansing Analysis & Modeling Interpretation 	<ul style="list-style-type: none"> Acquisition & Warehousing Mining & Cleansing Analysis & Modeling Aggregation & Integration
	Management	<ul style="list-style-type: none"> Privacy Security Ownership Information Sharing 	<ul style="list-style-type: none"> Privacy Security Information Sharing Cost 	<ul style="list-style-type: none"> Information Sharing Cost Governance 	<ul style="list-style-type: none"> Information Sharing Governance Security 	<ul style="list-style-type: none"> Information Sharing Governance Security Privacy

Figure 4 Marketing Mix Framework for Big Data Assessment

This model helped FlixBus to identify not only its current standpoint but also the next steps to follow in order to achieve predictive analytics and use them on a way that generate value by selecting only the most critical data needed and being aware of the main challenges to overcome. Furthermore, this framework challenges organizations to take a more holistic big data approach by addressing different data related topics and incorporate them within the marketing mix. Not only the model aims to identify data needs, its potential applications and the challenge that might be encountered but also provides a better understanding for all the different stakeholders about which data is collected and how it is utilizes in order to meet marketing objectives, allowing for accountability and a better understanding of where the company stands and which are the next steps to take.

6. Discussion

This paper aimed to explain how different big data characteristics and challenges can be incorporated into the marketing mix to have more successful data strategies and become more data-oriented, allowing companies to achieve a smoother transition from descriptive to predictive analytics. Thus, the case study serves as a mean to identify key data challenges, data applications and critical data sources faced by organizations on the endeavour of achieving better analytical models. The findings acknowledge that in order to develop more sophisticated data strategies and move toward more precise analysis methods such as predictive analytics, a traditional approach is not enough. The findings of FlixBus case study showcased that organizations must approach big data with a more holistic view, thinking beyond data as an isolated mean to achieve competitive advantage. The framework resulted from an exhaustive literature review on big data challenges, characteristics and its use on the marketing mix and was adapted from Fan (2015)

providing organizations the critical elements to consider when considering a big data strategy or willingness to switch from a simpler descriptive analytical method to a more robust one. The framework can be used to fulfil a specific task within the marketing department or as a more complex tool to switch entirely the data strategy of the organization.

6.1 Limitations and Directions for Further Research

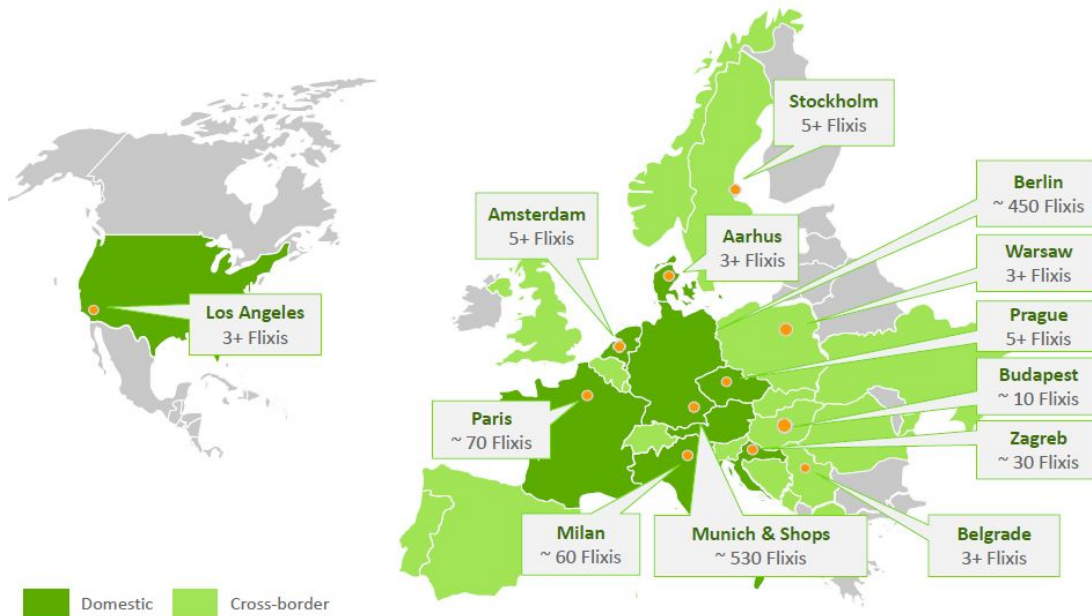
Regardless of the marketing mix framework for big data management, the scope of the research is still limited, hence, researchers, readers and future academics should be aware of these limitations and interpret this paper accordingly. The first limitation is the nature of the framework and its focus on marketing, therefore further research and adaptation for broader departments may be analysed. The second limitation is specific to the characteristics of FlixBus regarding organizational, technology and human capital at the moment of this analysis and the creation of the framework, which might be applied for companies on similar situations but the need of adjustments may exist for companies on early or more advanced stages. Another limitation is the few information about best practices in prescriptive analytics in both theory and practice which makes it hard be contemplated within the framework on the specific case of FlixBus. Finally, regardless of the key findings of the study and support from the literature reviewed, the final analysis and synthesis was based on the researcher interpretation, which might contain error and biases. Therefore, further research should consider the following points. Analyse how to choose the right data sources to achieve specific goals. Study how to select the right data mining techniques and analysis methods. Investigate how to balance and calculate big data return on investment. Examine how to keep this framework refined and actualized as big data technology is continuously evolving.

References

- Akerkar, R. (2014). *Big data computing*. Florida, USA: CRC Press, Taylor & Francis Group.
- Baskarada, S. (2013). *Qualitative Case Study Guidelines*. Victoria: Joint and Operations Analysis Division DSTO.
- Berners-Lee, T., & Shadbolt, N. (2011). There's gold to be mined from all our data. *The Times*, London 1:1–2. Online Available at: <http://www.thetimes.co.uk/tto/opinion/columnists/article3272618.ece> (Accessed on 21st February 2018).
- Chen, G., Chen, K., Jiang, D., Ooi, B. C., Shi, L., Vo, H. T., & Wu, S. (2012b). E3: an elastic execution engine for scalable data processing. *Journal of Information Processing*, 20(1): 65–76.
- Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 36(4): 1165–1188.
- Davenport, T. H., & Dyché, J. (2013). *Big data in big companies*. International Institute for Analytics. Available Online at http://www.demonish.com/cracker/1431316877_1217a9641e/bigdata-bigcompanies-106461.pdf (Accessed 5th March 2018).
- Davenport, T. H., & Harris, J. G. (2007). *Competing on analytics: The new science of winning*. Harvard Business Press.
- Fan, S., Lau, R. Y.K., Zhao, J. L. (2015). Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix. *Big Data Research*, 2: 28-32.
- Gandomi, A., Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35: 137-144.
- Gantz, J., & Reinsel, D. (2012). *The Digital Universe in 2020: Big data, bigger digital shadows, and biggest growth in the Far East*. IDC – EMC Corporation. Online Available at <http://www.emc.com/collateral/analyst-reports/idc-the-digital-universe-in2020.pdf> (Accessed 16th February 2018).
- Hold, A. How FlixBus uses AI to make bus travel sexy. Online Available at <http://alexanderhold.com/how-flixbus-use-ai-to-make-bus-travel-sexy/> (Accessed on 1st May 2018)
- Savitz, E. (2012a). Gartner: Top 10 strategic technology trends for 2013. Online Available at <http://www.forbes.com/sites/ericsavitz/2012/10/23/gartner-top-10-strategic-technology-trends-for-2013/> (Accessed on 10th March 2018).
- Savitz, E. (2012b). Gartner: 10 critical tech trends for the next five years. Online Available at <http://www.forbes.com/sites/ericsavitz/2012/10/22/gartner-10-critical-tech-trends-for-the-next-five-years/> (Accessed on 10th March 2018).
- Sivarajah, U., Kamal, M. M., Irani, Z., Weerakkody, V. (2016) Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 70: 263-286.
- Yi, X., Liu, F., Liu, J., & Jin, H. (2014). Building a network highway for big data: architecture and challenges. *IEEE Network*, 28(4): 5–13.
- Zicari, R. V. (2014). *Big Data: Challenges and Opportunities*. (2014) In R. (Ed.), *Big data computing* (pp. 103–128). Florida, USA: CRC Press, Taylor & Francis Group.
- (2012, June 26). Small and midsize companies look to make big gains with “big data,” according to recent poll conducted on behalf of SAP. Retrieved from <http://global.sap.com/corporate-en/news.epx?PressID=19188>

Appendix

Exhibit 1



Countries where FlixBus operates and where offices are located

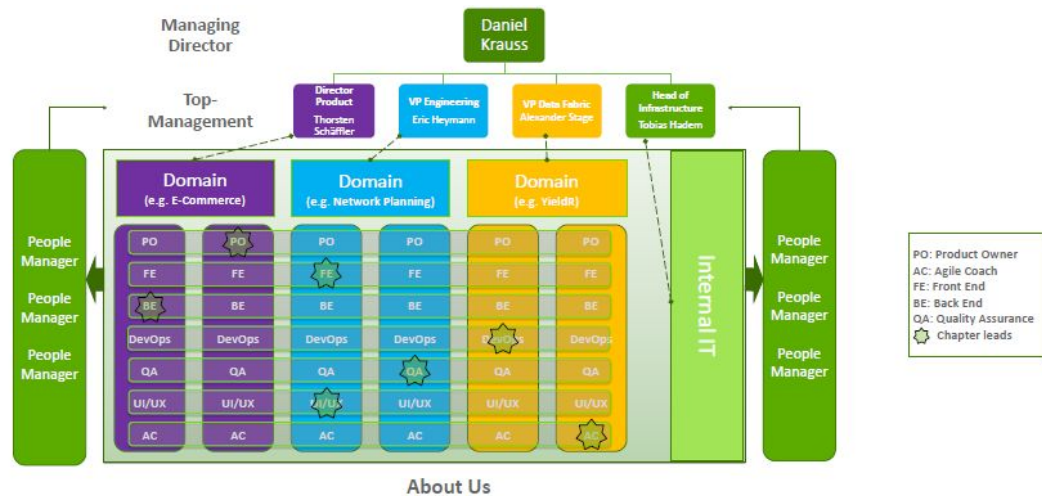
List of Countries:

U.S.A., Portugal, Spain, France, Netherlands, Luxembourg, Belgium, Austria, Italy, Germany, Denmark, U.K. Poland, Czech Republic, Hungary, Sweden, Norway, Switzerland, Croatia, Bosnia, Slovakia, Romania, Slovenia, Serbia, Ukraine, Montenegro, Macedonia.

Exhibit 2



Overall Organizational Structure



Top Management – Board

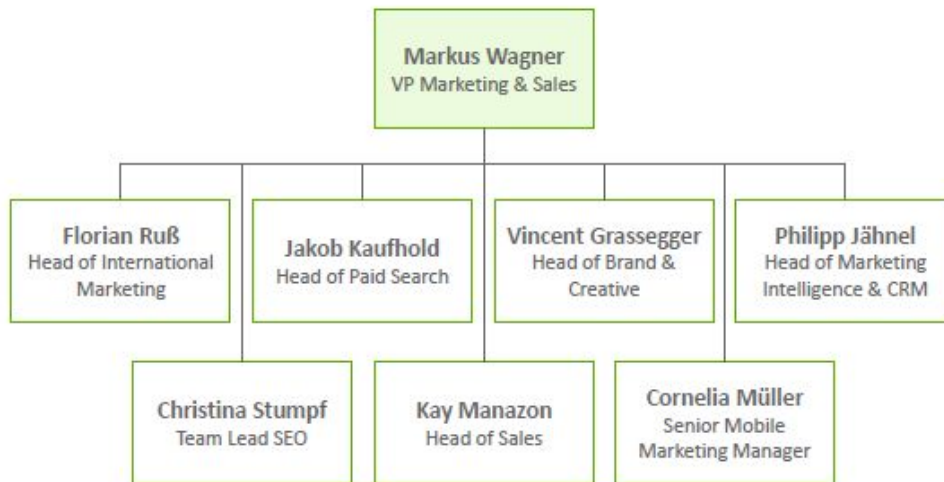
- CIO holds the ultimate accountability for FlixTech
- VP Data Fabric embeds data deeply into the DNA of FlixBus
- VP Engineering is accountable for delivering software products of extraordinary quality

- **Director Product** drives innovation specifically towards the end customer
- **Head of Infrastructure** and Operations is accountable for the operations of all production systems

Leadership Roles

- People Managers: Mastery
 - Disciplinary responsibilities
 - Hiring and Controlling
- Agile Coaches: Autonomy
- Product Owner: Purpose
- Chapter Representative: Governance

FlixTech Functions and Organizational Structure



Marketing Organizational Structure