ECOLE DES HAUTES ETUDES COMMERCIALES

EHEC ALGER

Dissertation in partial fulfillment of the requirement for a Master's degree in Business Sciences

Major: Distribution & Supply Chain Management

Research subject:

The impact of big data analytics in customer relationship management

Case study : YASSIR SPA

Developed by:

HOCINE Abdennour

Supervised by:

Ms. Naima MESSAOUDI

Marketing Lecturer

8th promotion

June 2021

ECOLE DES HAUTES ETUDES COMMERCIALES

EHEC ALGER

Dissertation in partial fulfillment of the requirement for a Master's degree in Business Sciences

Major: Distribution & Supply Chain Management

Research subject:

The impact of big data analytics in customer relationship management

Case study: YASSIR SPA

Developed by:

HOCINE Abdennour

Supervised by:

Ms. Naima MESSAOUDI

Marketing Lecturer

8th promotion

June 2021

Dedications

To my loving mother,

My dear grandparents, my aunts, and my uncles who have my Eternal Gratitude, May God give you long life, health and happiness.

To my sisters Aicha, Abir, Chaimaa, my dear brothers Abdelouaheb, Allaedinne, Brahim who without them I would not be where I am today.

Also, to my Brother-in-law Madjid and their beautiful children Tasnime, Abdelmalek, Djihane.

To all my friends and anyone who has marked me during this glorious journey.

Abdennour

Acknowledgement

Our praise goes above all to God Almighty, in whom I have faith and who has rewarded me with this present work.

Our thanks go first to all Yassir company members where we have done our internship, Abdelouaheb Hocine, Imad Boutbali, Othman Benallal, Adam Bougedra, Khadidja Yasmine Bourga, Mustapha Debih, Souhila Benbrahim, Arslan Halimi.

I also would like to thank them for the continuing support, the supervision, the availability and the time they dedicated for me, I can never thank them enough for the generous collaboration and information sharing for the advancement of this work.

Besides, I would like to express my deep gratitude to my thesis supervisor Dr. Naima Messaoudi, for the help she has given me in carrying out this research paper.

Furthermore; we warmly thank all the experts from outside the enterprise that accepted to help and answer the given questions even with the lack of time in their respective professions, and for the start-up managers for the very warm welcoming and quick integration in the team also the answers to our interviews' questions.

I thank all of them for the time they have given me but also for the exchanges and discussions that we have had and which allowed me to carry out this research work.

We would like to give an affectionately thanks: Fatma salmi my mother, and to all the family members, friends and all those who supported me, from near or far, throughout the years of study.

Sincere and honest thanks to all of you; because you have always been present. May this work be a testimony of my gratitude and deep respect.

By Abdennour Hocine

Abstract

In today’s connected world, the traditional practices of supply chain management are far from being sufficient, because customer behaviour nowadays is becoming unpredictable, and in order to retain its consumers, the companies realize the importance of using new techniques in order to find sustainable solution as a context for the fierce market.

The explosion of big data has created a new opportunity and made a pressure about companies to start transforming big data to knowledge and to start leveraging on Big Data solutions in order to sustain the competitions by becoming more data-driven in managing their supply chains.

Customer relationship management is a strategy by which the company aims to understand, anticipate and manage the needs of its current and potential customers, with the aim of creating profitable and lasting relationships with customers and creating a competitive advantage.

In this present work, we have tried to better understand how big data analytics techniques is improving customer relationship management's performance and its impact on each level of the customer relationship management.

Key-words: supply chain management, big data, data-driven, customer relationship management, big data analytics.

Résumé

Dans le monde connecté d'aujourd'hui, les pratiques traditionnelles de gestion de la chaîne d'approvisionnement sont loin d'être suffisantes, car le comportement des clients de nos jours devient imprévisible, et afin de fidéliser ses consommateurs, les entreprises réalisent l'importance d'utiliser de nouvelles techniques afin de trouver une solution durable comme un contexte de marché féroce.

L'explosion des mégadonnées a créé une nouvelle opportunité et fait pression sur les entreprises pour qu'elles commencent à transformer les mégadonnées en connaissances et à tirer parti des solutions Big Data afin de soutenir la concurrence en devenant davantage axées sur les données dans la gestion de leurs chaînes d'approvisionnement.

La gestion de la relation client est une stratégie par laquelle l'entreprise vise à comprendre, anticiper et gérer les besoins de ses clients actuels et potentiels, dans le but de créer des relations rentables et durables avec les clients et de créer un avantage concurrentiel.

Dans ce travail, nous avons essayé de mieux comprendre comment les techniques d'analyse de big data améliorent la performance de la gestion de la relation client et son impact à chaque niveau de la gestion de la relation client.

Mots-clés : gestion de la chaîne d'approvisionnement, big data, gestion des données, gestion de la relation client, analyse des big data.

ملخص

في عالم اليوم المتصل، فإن الممارسات التقليدية لإدارة سلسلة التوريد بعيدة كل البعد عن أن تكون كافية، لأن سلوك العملاء اليوم أصبح غير متوقع، ومن أجل الحفاظ على عملائها، تدرك الشركات أهمية استخدام تقنيات جديدة من أجل إيجاد حل مستدام خاصة في ظل المنافسة الشرسة.

لقد أدى الانفجار في البيانات الضخمة إلى خلق فرص جديدة وفرض ضغوطًا على الشركات لبدء تحويل البيانات الضخمة إلى معرفة والبدء في الاستفادة من حلول البيانات الضخمة من أجل الحفاظ على المنافسات من خلال أن تصبح أكثر اعتمادًا على البيانات في إدارة سلاسل التوريد الخاصة بهم.

إدارة علاقات العملاء هي استراتيجية تهدف الشركة من خلالها إلى فهم وتوقع وإدارة احتياجات عملائها الحاليين والمحتملين، بهدف إنشاء علاقات مربحة ودائمة مع العملاء وخلق ميزة تنافسية.

في هذا العمل، حاولنا أن نفهم بشكل أفضل كيف تعمل تقنيات تحليلات البيانات الضخمة على تحسين أداء إدارة علاقات العملاء وتأثيرها على كل مستوى من مستويات إدارة علاقات العملاء.

الكلمات المفتاحية: إدارة سلسلة التوريد، البيانات الضخمة، إدارة علاقات العملاء، تحليلات البيانات الضخمة.

List of figures

List of tables

List of schemas

List of abbreviations

SCM Supply chain management

SC Supply chain

SCA Supply chain analytics

BD Big data

BDA Big data analytics

SQL Structed query language

NOSQL Not only SQL

RDBMS Relational data base management system

DBMS Data base management system

HTTP Hyper-Text Transfer Protocol

IOT Internet of things

CSV file Comma-separated values file

IBM International Business Machines

IDC international data corporation

CRM customer relationship management

SRM Suppliers relationship management

EDI Electronic Data Interchange

ERP Enterprise resource planning

RFID Radio frequency identification data

ML Machine learning

AI Artificiel intelligence

BI Business intelligence

BA Business analytics

BDBDA Big data business analytics

ED Enterprise dynamics

LSCM Logistics & Supply chain management

SCOR supply chain operation Reference

MILP Mixed-integer linear programming

ROI Return on investment

AHP Analytical hierarchy process

DS Data science

R&D Product research and development

PPC Production planning and control

QM Quality management

WMS Warehouse management system

CPFR Collaborative planning, forecasting and replenishment

ITS Intelligent transportation system

ORION On-Road Integrated Optimization and Navigation

GPS Global position system

E-CRM Electronic customer relationship management

S-CRM Social customer relationship management

R&D Research and development

CAGR Compound Annual Growth Rate

TNC Transportation network companies

RHs Ride-hailing services

Summary

General introduction

General introduction

More than ever before, the economic world is experiencing a rapid pace of change characterized by complex globalization, increasing competition, and unprecedented technological advances; almost every business sector has been affected by the abundance of information called big data, the supply chain is no different.

The organizations are faced with streams of data flooding in from various channels at an accelerating rate. This data can hamper an organization's ability to keep up with data inflows and derive valuable insights. Moreover, this has created a new analytics technique that has come up with tools and techniques to make data-driven supply chain decisions. Analyzing and interpreting results in real-time can assist enterprises in making better and faster decisions to satisfy customer requirements. It will also help organizations improve their supply chain design and management by reducing costs and mitigating risks. Recently, various research studies have indicated the benefits of using big data methods in logistics and supply chain management.

This giant and complex sector, both dynamic and competitive, requires quick adaptability to customers' evolving needs and desires. Also, the need for new analytics techniques creates profoundly changing aspects, as it became a significant concern for all supply chains and, more specifically, customer relationship management. The customer's behavior has been changing radically because of the tremendous development of information and communication technologies, and the customer is becoming more and more demanding. His loyalty has become a challenge to hand. The companies have to reconsider their strategies and adopt a customer-centric approach. They have moved from a "product" orientation to a "customer" orientation. Thus, the BDA takes a central position.

We decided to formulate this research to investigate the use of big data analytics techniques in improving customer relationship management performance. As a result, the overall question study to which this thesis will answer:

How is big data analytics impacting Customer relationship management performance?

In order to answer this problem, we have found it helpful to ask the following sub-questions:

- What levels does Big Data Analytics affect customer relationship management?
- How can the use of big data analytics affect value creation in customer relationship management?

- What are the implications of applying big data analytics techniques on customer relationship Management's strategy & operations, and how it relates to the supply chain performance?

To better understand our problematic and answer these questions, it seems necessary to examine the following hypothesis:

- **H1:** Big data analytics can identify known risks and help to predict future risks by spotting patterns and trends throughout the supply chain. Increase accuracy in planning.
- **H2:** Big data analytics will enhance CRM strategies in delivering personalization and customization of services for the customer.
- **H3:** Integrating big data with Customer Relationship management empowers quickly finding out how products or brands are perceived by customers online, identifying the weaknesses, and revising the marketing and sales strategies.

To verify the affirmation or the cancellation of these assumptions, we will adopt the following methodology: In-depth documentary research, through the consultation of several books, articles, specific websites as well as academic works related to our subject of studies, also internal documents of the company as well as quantitative research through customer segmentation.

To accomplish our research, we have structured our work in three chapters:

The first chapter is devoted to the supply chain analytics, which firstly presents a general presentation of big data and big data analytics, then, we have presented the new concept which combines analytics techniques and supply chain, and, finally, the importance of applying big data analytics in supply chain management's sector.

The second chapter is dedicated to electronic customer relationship management, including general notions about customer relationship management, then we have described the integration of big data analytics and customer relationship management.

The last chapter highlights the use of analytics techniques in customer relationship management, more precisely in Yassir SPA, including the company presentation, the survey approach, and the presentation of the results.

Chapter One

Supply chain analytics

Introduction

Nowadays, in the current competitive environment, companies are beginning to realize the importance of the availability of large amounts of data, and started to use it in order to make the right decisions and support their strategies, moreover, supply chain professionals are struggling to deal with big data to reach an integrated, efficient, effective and supply chain.

In this first chapter, we will present and discuss the new concept that combines big data analytics and supply chain management, which is called Supply chain analytics. First, we will highlight the factors that add value to the whole supply chain's network in general. We will then explore how big data analytics impacts the supply chain's network at all levels.

This chapter must start by understanding what Big Data is, and that is what the first part of section one will be about. Then, the second part will present extensive data analytics, Its definition, and its importance in the business. Finally, we will clarify the new notion of supply chain analytics and discuss it.

Keywords: Supply chain, big data, big data analytics, supply chain analytics.

I. Section one: Understanding Big data analytics

To fully understand the meaning, impact, and application of BDA, we first need to understand what Big Data is.

1. Understanding Big data

1.1 Big data evolution

They have been using the term 'Big Data' since the 90s. "***John R. Mashey*** is given the credit of making the 'Big Data' term popular ". [1] Big Data is not wholly new or used only from the last years. People have been trying to use data analytics and diagnostic techniques to support the decision-making process for many years.

The tremendous increase of structured and unstructured data sets made traditional data analysis difficult and thus transformed into 'Big Data' in the last decade. The evolution of Big Data can be classified into 3 phases, where every Phase has its characteristics and capabilities and has contributed to the contemporary meaning of Big Data.

- **Phase I:** Big Data originates from the domain of database management. It mainly depends on the storage, extraction, and optimization of data stored in Relational Database Management Systems (RDBMS). Database management and data warehousing are the two core components of Big Data in the first Phase. It gives a foundation to modern data analysis and techniques such as database queries, online analytical processing, and standard reporting tools.
- **Phase II:** From the early 2000s, usage of the Internet and the Web started offering unique data collections and data analysis opportunities. Yahoo, Amazon, and eBay expanded their online stores and started analyzing customer behavior for personalization. The HTTP-based content on the Web massively increased the semi-structured and unstructured data. Organizations now had to find new approaches and storage solutions to deal with these new data types and analyze them effectively. In later years the growth of social media data aggravated the need for tools, technologies, and analytics techniques to deal with this unstructured data.
- **Phase III**: over the last decade, the large-scale usage of smartphones with different internet-based applications gives the possibility to analyze behavioral data,

[1] https://www.bigdataframework.org/short-history-of-big-data/ (13/04/2021 at 16:00)

also, location-based data (GPS-data). Simultaneously, the rise of sensor-based Internet-enabled devices termed the 'Internet of Things (IoT) is making millions of T.V.s, thermostats, wearables, and even refrigerators generate zettabytes of data every day. "This explosion in the volume of 'Big Data' now started a race to extract valuable insights and information from these new data sources".[2]

Table 1.1: the summary of the three phases in Big Data.

Phase-I	**Phase-II**	**Phase-III**
DBMS-based, structured content	**Web-based, unstructured content**	**Mobile and sensor-based content**
1.RDBMS & data warehousing 2.Extract Transfer Load 3.Online Analytical Processing 4.Dashboards & scorecards 5.Data mining & statistical analysis	1.Information retrieval and extraction 2.Opinion mining 3.Question answering 4.Web analytics and Web intelligence 5.Social media analytics 6.Social network analysis 7.Spatial-temporal analysis	1.Location-aware analysis 2.Person-centered analysis 3.Context-relevant analysis 4.Mobile visualization 5.Human-Computer interaction

Source: BANU (Ayesha) & YAKUB (Mhd): «*Evolution of Big Data and Tools for Big Data*,» Journal of Interdisciplinary Cycle Research, XII(X), October (2020), p 312.

[2] BANU (Ayesha) & YAKUB (Mhd): «*Evolution of Big Data and Tools for Big Data*,» Journal of Interdisciplinary Cycle Research, XII(X), October (2020), p.312.

1.2 The Data Deluge

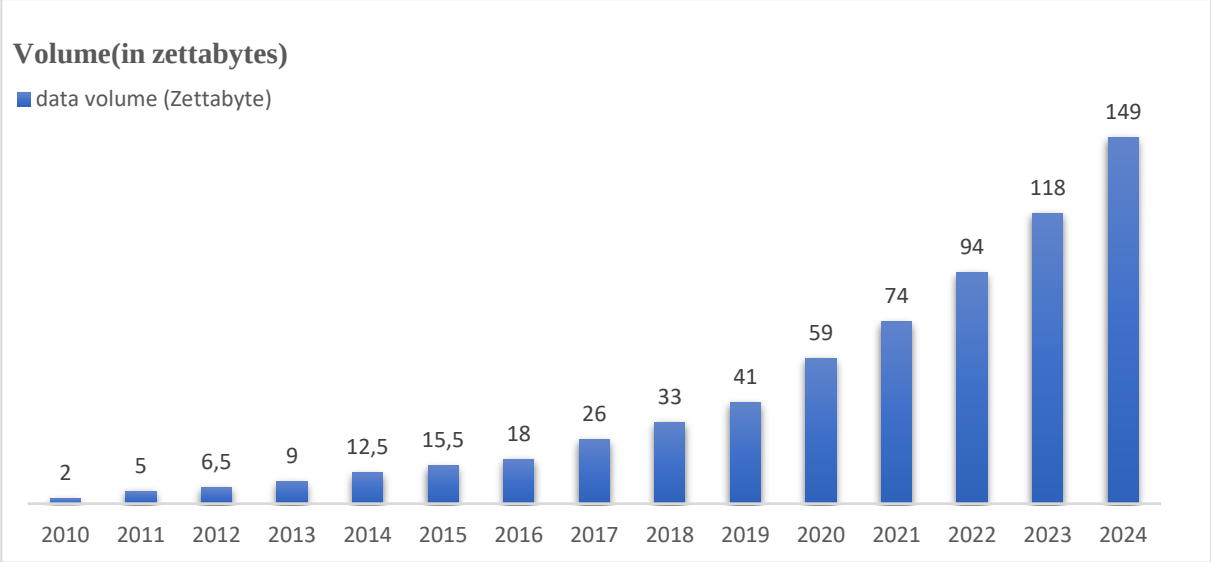
Big Data is a broad term with many related meanings on the technical, computing, data, and usage levels; it is not an entirely new term; its history starts many years before the present buzz around Big Data. Many decades ago, people started using analytic techniques to support their decision-making process. However, in the last two decades, due to tremendous usage of the Internet and social networking, the amount with which Data increased became beyond measures of human comprehension.

Earlier, the Data was quantized in terms of kilobytes, megabytes, and gigabytes. Nevertheless, today, Data is counted in terms of terabytes, petabytes, and Zettabytes. According to tech jury 2020, "the data produced by humans every day is 2.5 quintillion bytes, and 463 exabytes of data will be generated each day by humans as of 2025".[3]

The explosion in the amount of data, called "data deluge" ,"now days is becoming a reality this Data's volume evolution is forcing us to use new units of measurement with which we have yet to familiarize ourselves, During the day when you are reading this, more data will be produced than the amount of information contained in all printed material in the world", [4], "according to IBM, 90% of global data has been created in the last two years "[5], and "the Internet Data Centre (IDC) estimated the growth of data to be of a factor of 300 between 2005 and 2020, expecting to rise from 130 Exabytes to 20,000 Exabytes" [6].The total amount of data created, captured, copied, and consumed in the world is forecast to increase rapidly, "reaching 74 zettabytes in 2021 and reach 149 zettabytes by 2024". [7]

[3] https://techjury.net/blog/how-much-data-is-created-every-day/#gref (13/04/2021 at 18:00)
[4] The amount of information contained in all printed material is estimated to be around 200 Petabytes, while IBM estimated that in 2012, 2.5 Exabytes (1 Exabyte = 1024 Petabytes) of new data was in each day created, and the amount continues to increase.
[5] https://www.digital-business-trends.com// (13/04/2021 at 18:21)
[6] GANTZ (John) & REINSEL (David): *« The digital universe in 2020*: *Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East »*, Internet Data Center (IDC), December 2012, p.3.
[7] https://www.statista.com/statistics/871513/worldwide-data-created// (13/04/2021 at 18:21)

Figure1.1: Volume of data created, captured, copied, and consumed worldwide from 2010 to 2024



Source: https://www.statista.com/statistics/871513/worldwide-data-created// (13/04/2021 at 18:21)

1.3 Big data definition

Being a complex polymorphic object, its definition varies according to the communities that are interested in it as a user or provider of services., Big Data is not a simple concept to define precisely since the very notion of big in terms of volume of data varies from one area to another. A set of technologies does not define it. On the contrary, it defines a category of techniques and technologies. This is an emerging field, and as we seek to learn how to implement this new paradigm and harness the value, the definition is changing.

It can be seen that there are several definitions and used and are accepted depending on the community.

“The term "*Big Data*" refers to the evolution and use of technologies that provide the right user at the right time with the correct information from a mass of data that has been growing exponentially for a long time in our society. The challenge is to deal with rapidly

increasing volumes of data and the difficulty of managing increasingly heterogeneous formats and increasingly complex and interconnected data". [8]

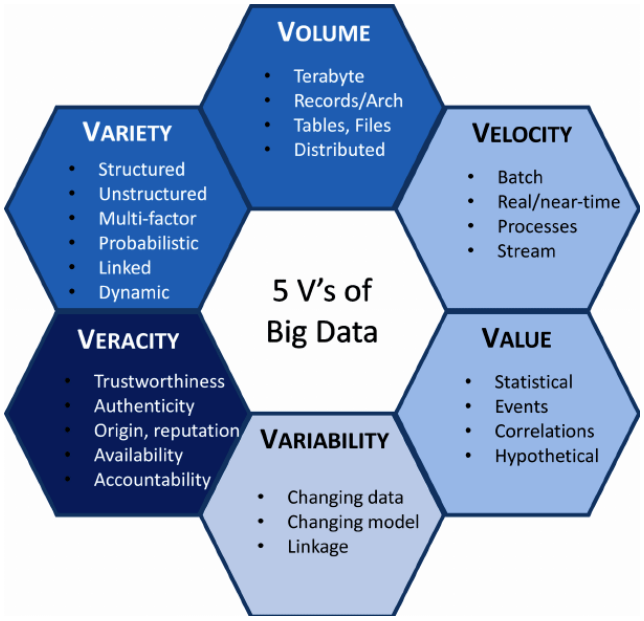
"It refers to the exponential growth of the data volumes, the actual processing of the large volumes of data, or, more generically, to all the computing phases required to extract useful information". [9]

However, to better characterize the concept of big data, and prevent potential confusion about its definition, a set of attributes were identified as defining characteristics. According to their common name initial, this set of attributes is referred to as the V's of Big Data. Considering that the Big Data ecosystem is highly dynamic, the set is expanding to include new V's (i.e., challenges) continuously identified.

1.4 Big data Characteristics

We can say that what characterizes Big Data is the concept of 5 v's these 5 v's are what explain the big data concept The convergence of these five dimensions helps to define Big Data:

Figure 1.2: the 5v's of big data



Source: SERRAO (Carlos), et al.: *« Handbook of Research on Trends and Future Directions in Big Data and Web Intelligence, »* Information Science Reference, global book series (ADMDM), 2015, p.22

[8] RIAHI (Sara): *« Big Data and Big Data Analytics: concepts, types and technologies*,» International Journal of Research and Engineering, (2018), p.524–528.
[9] RIAHI (Azzedine) & RIAHI (Sara)*: «The big data revolution, issues and application,»* IJARCSSE, August 2015, p.167.

1.4.1 Volume: The volume represents the vast amount of data, and it is called a large size, understood as unmanageable data, and beyond the processing capacity of the traditional system, it represents the amount of data generated, stored, and operated. The increase in volume can be explained by increasing the amount of data generated and stored and the mass quantities of data that organizations are trying to use to improve decision-making processes.

the volume of data continues to increase at an unprecedented rate. *(According to **Statista website**, by 2024, there will be 149 zettabytes of data created)*, However, what constitutes genuinely "high" volume varies by industry and even geography and is smaller than the petabytes and zettabytes often referenced. Many companies consider datasets between one terabyte and one petabyte to be Big Data. Still, everyone can agree that whatever is considered "high volume" today will be even higher tomorrow.

1.4.2 Variety: Big Data is generated in multiple varieties; variety refers to the number of data types managed by an information system. This multiplication leads to a complexity of links and link types between these data. "The variety also relates to the possible uses associated with raw data. Technological advances allow organizations to generate various types, and we cannot talk about the variety without defining the significant data types (structured, semi-structured, and unstructured data)". [10]

1.4.2.1 Big data types:

1.4.2.1.1 Structured Data owns a dedicated data model. It also has a well-defined structure, follows a consistent order, and is designed to be easily accessible and used by a person or a computer. Structured Data is usually stored in well-defined columns and also Databases. **Example**: ID-Codes in Database (DBMS)

1.4.2.1.2 Semi-Structured Data is another form of Structured Data. It inherits a few properties of Structured Data, but the central part of this kind of data fails to have a definite structure, and also, it does not obey the formal structure of data models such as an RDBMS[11]. **Example:** Comma Separated Values (CSV) File.

1.4.2.1.3 Unstructured Data "is altogether a different type of which neither has a structure nor obeys to follow the formal structural rules of data models. It does not even have a consistent format, and it found to be varying all the time., However, rarely, it may have information related to Data and time".[12]. **Example**: Videos, Images, Audios

---

[10] BANU (Ayesha) & YAKUB (Mhd): «*Evolution of Big Data and Tools for Big Data*,» Journal of Interdisciplinary Cycle Research, XII(X), October (2020), p.312.

[11] https://www.edureka.co/blog/big-data-characteristics/ (27/06/2021 at 16/51)

[12] IDEM

1.4.3 Velocity: The speed and directions from which data come into the enterprise are increasing due to interconnection and advances in network technology, so "it is coming in faster than we can make sense of it. Furthermore, the faster the data come in and the more varied the sources, the harder it is to derive value (meaning) from the data". [13] , "velocity refers to the speed at which data are generated and processed, and it Could be defined as the speed of data traveling from one side to another or moves around and the speed of processing it with the high rate of receiving data and information".[14]

High-velocity Data is generated at such a pace that it requires distinct (distributed) processing techniques. An example of data that is generated with high velocity would be Twitter messages or Facebook posts.

1.4.4 Veracity: "It refers to the level of reliability associated with certain types of data; the trustworthiness of Data impacts their value"[15], therefore, one of the newly identified challenges when dealing with Big Data is veracity, which generically synthesizes the Data's correctness, quality, and accuracy. "The concerns related to the veracity apply both to the input, and the result harvested when mining it. Thus, the quest for high data quality is an essential Big Data requirement and challenge. However, even the best data cleansing methods cannot remove the inherent unpredictability of some data, like the weather, the economy, or a customer's buying decisions, which makes veracity an essential V of big data."[16]

Data veracity focuses on data quality and accuracy and defines how data can be trusted when an important decision needs to be made regarding the collected data. "Data is categorized in good, bad, or undefined data, which could be because of data inconsistency, incompleteness, ambiguity, latency, deception, and approximations". [17]

1.4.5 Value: Data value refers to data usefulness in decision making. Big data delivers more prosperous and more subtle insights by nature than can ever be achieved with standard data collection and analysis, so collecting, storing, and analyzing Big Data is useless unless it produces value. Therefore, this aspect goes alongside and determines any previous or future challenge of Big Data. It can be safely stated that "Value" is the ultimate goal of Big Data.

---

[13] https://bigdatapath.wordpress.com/2017/12/28/what-is-big-data-more-than-volume-velocity-and-variety/

[14] HUSSIEN (Abdou) et al.*: « How Many Old and New Big Data V's Characteristics, Processing Technology, and Applications (BD1) »*, IJAIEM, September 2020, p.19.

[15] VACCARO (Carlo): *« Big Data in Official Statistics and Official Statistics*,» Ph.D. Thesis in Computer Science, University of Camerino, 2014, p.56.

[16] VENKATESH (G), ARUNESH (K*): « An Overview of Big Data Characteristics and Enhancement of Big Data Application, »*, International Journal of Computer Science & Technology, Jun 2019, p.16-17.

[17] MILLS (Steve) et al.: «*Demystifying big data: a practical guide to transforming the business of government,»* Tech America Foundation, Washington, 2012, p.10.

Being the driven factor of this technological revolution, value is the importance of the data, and this is a very significant feature in B.D. The possible value of B.D. is substantial. However, without timely access, its value cannot be exploited.

2. Big data analytics

The concept of big data has been around for years; most organizations now understand that if they capture all the Data that streams into their businesses, they can apply analytics and get significant value from it.

Big data analytics is one of the best techniques, which can help organizations to overcome their problem. The new benefits that big data analytics brings to the table it, however, are speed and efficiency because it provides a tool for extracting valuable patterns and information in a large volume of data. Whereas a few years ago, a business would have gathered information, run analytics, and unearthed information that could be used for future decisions, today, that business can identify insights for immediate decisions. The ability to work faster – and stay agile – gives organizations a competitive edge they did not have before.

2.1 Big data analytics definition

Big data analytics "examines large amounts of data to uncover hidden patterns, correlations, and other insights. According to SearchBusinessAnalytics, BDA is the often-complex process of examining big data to uncover information; it is a form of advanced analytics, which involve complex applications with elements such as predictive models, statistical algorithms, and what-if analysis powered by analytics systems that can help organizations make informed business decisions".[18]

In simple word, we can say that BDA "is the use of advanced analytic techniques against enormous, diverse big data sets that include structured, semi-structured, and unstructured data, from different sources, and in different sizes from terabytes to zettabytes". [19]

Combining big data and analytics makes the different tools that help decision-makers get valuable, meaningful insights and turn information into business intelligence. After all, big data without analytics are just lots of data because Data itself has no value. On the other hand, analytics without big Data is simply mathematical and statistical tools and applications.

[18] https://searchbusinessanalytics.techtarget.com/definition/big-data-analytics (17/04/2021 at 04:02)
[19] https://www.ibm.com/analytics/hadoop/big-data-analytics (17/04/2021 at 04:04)

2.2 Big data analytics types

2.2.1 Descriptive analytics

It looks at Data and analyses past events for insights as to how to approach the future. In addition, it is looking at the reasons behind past failure and success. In this process, the following questions are answered: What has happened, what is happening.

"Descriptive analytics is a commonly used form of data analysis whereby historical Data is collected, organized, and then presented in an easily understood way. Descriptive analytics is focused only on what has already happened in a business and, unlike other methods of analysis, it is not used to draw inferences or predictions from its findings. Instead, descriptive analytics is a foundational starting point used to inform or prepare data for further analysis down the line".[20]

Generally, descriptive analytics uses simple maths and statistical tools, such as arithmetic, averages, and percent changes, rather than the complex calculations necessary for predictive and prescriptive analytics. "Visual tools such as line graphs and pie and bar charts are used to present findings, and online analytical processing (OLAP) systems are used and supported by reporting technology (e.g., RFID, GPS, and transaction bar-code) and real-time information identify new opportunities and problems. It is used to collect, describe, and analyze the raw data of past events". [21]

This process can find possible business-related opportunities because it analyses and describes past events and makes it something interpretable and understandable by a human.

2.2.2 Predictive analytics

While descriptive analytics focuses on historical Data, predictive analytics, as its name implies, is focused on predicting and understanding what could happen in the future.

"Predictive analytics techniques are used to answer what will happen in the future or likely to happen by examining past data trends using statistical, simulation, algorithms, and machine-learning techniques to identify the likelihood of future outcomes based on historical Data. These techniques seek to discover the causes of events and phenomena and predict the

[20] https://studyonline.unsw.edu.au/blog/descriptive-predictive-prescriptive-analytics (18/042021 at 20:25)
[21] SADEGHI DARVAZEH (Saeid), RAEESI VANNI (Iman), & MUSOLU (Farzaneh Mansouri): *«Big Data Analytics and Its Applications in Supply Chain Management,»* March 25th, 2020, p.2.

future accurately or fill in the data or information that already does not exist. However, statistical techniques cannot be used to predict the future with 100% accuracy". [22].

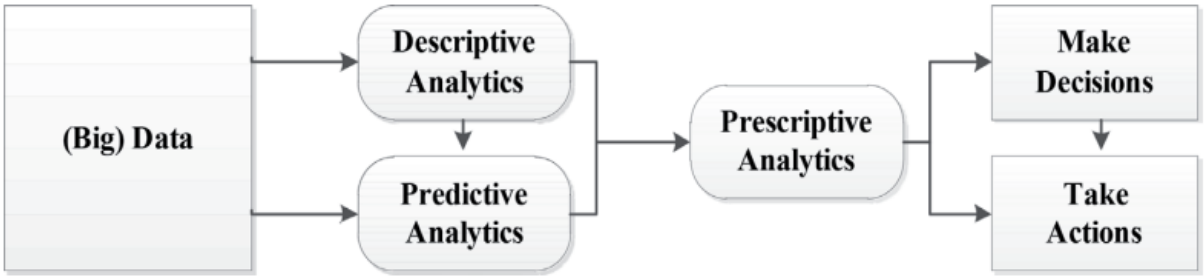
"Predictive analytics predict purchasing patterns, customer behavior, and purchase patterns to identify and predict sales activities' future trends. These techniques are also used to predict customer demands, inventory records and operations, so organizations can feel more confident that they are making the best possible business decision. Some of the most common predictive analytics applications include fraud detection, risk, operations, and marketing".[23]

2.2.3 Perspective analytics

Predictive analytics tells you what could happen. It deals with the question of what should be happening. Guides alternative decisions based on predictive and descriptive analytics using simulation, mathematical optimization, or multicriteria decision-making techniques. This methodology is the third, final, and most advanced stage in the business analysis process and the one that calls businesses to action, helping executives, managers, and operational employees make the best possible decisions based on the data available to them.

The application of prescriptive analytics is relatively complex in practice, and most companies are still unable to apply it in their daily activities of a business. Furthermore, the correct application of prescriptive analytics techniques can lead to optimal and efficient decision-making.

Figure 1.3: Using descriptive, predictive, and prescriptive analytics to make decisions and take actions



Source: SADEGHI DARVAZEH (Saeid), RAEESI VANNI (Iman), & MUSOLU (Farzaneh Mansouri): *«Big Data Analytics and Its Applications in Supply Chain Management,»* March 25th, 2020, p.3.

[22] https://www.sas.com/en_us/insights/analytics/big-data-analytics.html (13/04/2021 at 14:21)
[23] SADEGHI DARVAZEH (Saeid), RAEESI VANNI (Iman), & MUSOLU (Farzaneh Mansouri): *«Big Data Analytics and Its Applications in Supply Chain Management,»* March 25th, 2020, p.3.

3. Big data analytics techniques in Supply chain management

3.1 Statistics analysis

The statistical analysis consists of two types of analysis: descriptive and inferential. In descriptive statistics, past data are used to describe or summarize the feature of a phenomenon; it uses either graphs or tables or numerical calculations.

Inferential statistics are used to deduce the properties of phenomena and predict their behavior based on a sample of past data. Both quantitative and qualitative methods can be used simultaneously to take advantage of both the methods and the right decisions. "Statistical analysis is used when faced with uncertainty, such as distribution, inventory, and risk analysis. Statistical multivariate techniques are also used for supply chain monitoring to effectively manage the flow of materials and minimize the risk of an unintended situation". [24]

"Given the volume, variety, veracity, and velocity of big data, the supply chain needs robust and easy techniques for analysis. Unfortunately, traditional statistical methods are no longer responsive because the massive data leads to noise accumulation, heterogeneity, etc. Therefore, proposing and applying effective statistical methods are very important, and significant attention has been paid to this issue recently".[25]

3.2 Simulation

"Manufacturers need simulation tools to optimize the product development process, increase creativity, speed the time-to-market product, reduce production costs, and create innovation. Simulation provides many proven benefits for each stage of the product design and manufacturing process, such as producing more innovative products with greater efficiency for the customer and creating a better experience". [26].

Modeling and simulation techniques should be used to develop extensive data, such as simulation-driven product design. In today's competitive environment, the use of simulators to produce innovative products is considered a challenge. "Because manufacturers have to continually drive their operational efficiencies, meet the cost, require the time-to-market

---

[24] PUIGJANER (Luis), et al.: *« Supply chain monitoring: a statistical approach,»* European Symposium on Computer-Aided Process Engineering, Barcelona, January 1st, 2005, p.3.
[25] DARYANTO (Yosef), TIWARI (Sunil): *«Big data analytics in supply chain management between 2010 and 2016: Insights to industries »*, computers and Industrial Engineering. November 2017, p.8.
[26] KAMBATLA (Karthik) et al.: *« Trends in big data analytics*,» Journal of Parallel and Distributed Computing, July 2014, p.14.

product, and predict the customer preferences, Modeling and simulation help developers run the "what-if" analysis under different system configurations and complexity". [27]

"A simulation model to analyze the vast data collected from an intelligent manufacturing system's surrounding and shop floor environment has been developed. This model improved the decision-making in this production system." [28]. *For example*, as a predictive tool, simulation can help manufacturers predict the need for machines and additional equipment based on customer order forecast and learning from other historical data such as cycle time, throughput, and delivery performance. "Supply chain simulation can be used, for example, for predicting the service, testing the inventory policy, analyzing the production capacity, determining the asset utilization", [29] . Moreover, it is validating the optimization result. SCA provides new methods for the simulation problem with a large amount of data. Nowadays, several simulation software allows an evaluation of a system's performance before its creation. Enterprise Dynamics (E.D.) is one of the most powerful and most used software that researchers and practitioners use to simulate SCM issues.

3.3 Optimization

"The use of optimization techniques as part of SCA helps improve the accuracy of demand forecasting and supply chain planning. In addition, it is a powerful tool for supply chain data analytics".[30]

Optimization techniques by extracting the insights and knowledge of the enormous data generated by complex systems that include multiple factors and constraints such as capacity and route can analyze multiple objectives such as demand fulfillment and cost reduction.

Finally, "using supply chain optimization techniques and multiuser collaboration, performance tracker, and scenario management enables organizations to achieve their different goals. The use of optimization techniques supports supply chain planning and increases planning accuracy but presents a large-scale optimization challenge". [31], Several "signal processing and statistical learning techniques have been used for analytic optimization,

---

[27] RANJAN (Rajiv): *« Modeling and simulation in performance optimization of big data processing frameworks,»* IEEE Cloud Computing. 2014, p.16

[28] SHAO (Guodong), SHIN (Seung Jun) & SANJAY (Jain)*: «Data analytics using simulation for smart manufacturing,»* In Proceedings of the Winter Simulation Conference. IEEE; December 7th, 2014, p.19.

[29] LLamasoft: *« Supply chain simulation: why its time has come,»* White Paper, 2016, p.13.

[30] SARAVANAN (Balaraj): *« Optimization model for improving supply chain visibility*.» Infosys Labs Briefings. 2013, p.15.

[31] GANG (Gunasekaran) et al.: *« Big data analytics in logistics and supply chain management: Certain investigations for research and applications*.» International Journal of Production Economics. 2016, p.104.

principal component analysis, dictionary learning, compressive sampling, and subspace clustering".[32]

Optimization techniques are suitable for data analysis in LSCM. It helps analyze highly complex dynamic systems with substantial data volumes, multiple constraints, and factors and can gain insights that allow decision-makers to make appropriate decisions,

In addition, optimization helps analyze supply chain performance measures, such as cost reduction and demand fulfillment, among others. Another benefit associated with optimization is its flexibility because it can uncover new data connections, turn them into insights.

3.4 Machine Learning

In artificial intelligence, well-established methods such as neural networks, support vector machines, and statistical machine learning are classified as machine learning. Machine learning provides algorithms for computers to discover knowledge and make decisions by first learning from the given data. Machine learning methods have to be improved in big data analytics, both for the supervised and unsupervised learning schemes. "Deep machine learning, parallel support vector machines, fast learning" [33], "distributed machine learning" [34],Ontology learning and models like MapReduce are all machine learning techniques for dealing with big data problems.

3.5 Others

In addition to the above four mainstream and big primary data analytics techniques, other techniques such as social network analysis, clustering algorithm analysis, data envelopment analysis, data mining, and visualization analysis are helpful for big data analytics.

---

[32] SLOVAKIA (K), GIANNAKIS (G) & MATEO (G): *« Modeling and optimization for big data analytics:(statistical) learning tools for our era of data deluge*.» IEEE Signal Processing Magazine. 2014, p.19.
[33] HUANG (Guang-Bin) et al.: « *Efficient and rapid machine learning algorithms for big data and dynamic varying systems,»* IEEE Trans. Syst. Man Cybern. Syst. 47(10): 2625–2626, 2017, p.2.
[34] KUMAR (Abhimanu) et al.: « *Petuum: A new platform for distributed machine learning on big data*,» IEEE Trans. Big Data 1(2): 49–67, March 2015, p.5.

Table 1.2: Major Functionalities and Current Trends of the Major Big Data Analytics Techniques

Methods	Functionalities in big data analytics
Statistics	• Determine correlations and data patterns, and identify the data relationship (e.g., by regression) in a quick manner • Show whether a sample can denote the population, which helps reduce data requirements and computational time. • Serve as an essential technique to support other techniques such as data mining.
Machine learning	• By intelligence, learn, evolve, and capture behaviors of the systems understudies. • Can capture complex relationships in the systems but require substantial training time and memory • Flexible and able to support the processing of different data types by image processing, pattern recognition, text recognition, etc • Extract useful information from data by employing statistical and machine learning models • Clustering analysis, segmentation analysis, dynamic data mining with multi-source massive datasets
Optimization	• Find the optimal solution of quantitative analytical models • Require extensions to deal with real-time processing, parallel processing, etc. in large-scale optimization

Source: CHOI (Tsan-Ming), STEIN (W. Wallace) & WANG (Yulan)*:« Big Data Analytics in Operations Management,»* Production and Operations Management Society, Hong Kong, December 2017, p 1871.

II. Section tow: Supply chain analytics

1. Evolution of supply chain analytics

In the past, supply chain analytics was primarily limited to statistical analysis and quantifiable performance indicators for demand planning and forecasting. Data was stored in spreadsheets that came from different participants in the supply chain,

By the 1990s, companies adopted Electronic Data Interchange (EDI) and Enterprise Resource Planning (ERP) systems to connect and exchange information among the supply chain partners. These systems provided more accessible access to Data for analysis and assisting businesses in their designing, planning, and forecasting.

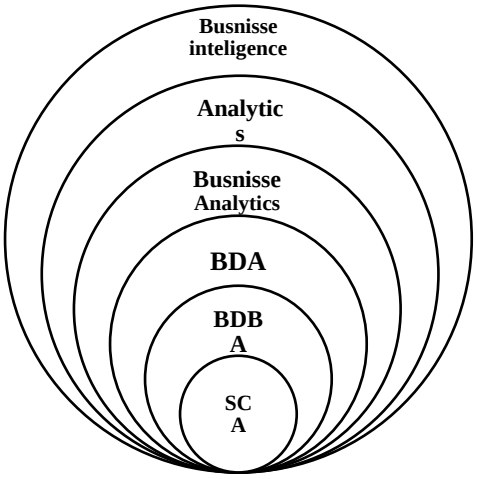
In the 2000s, businesses began turning to business intelligence and predictive analytics software solutions. These solutions helped companies gain a more in-depth knowledge of how their supply chain networks were performing, make better decisions, and optimize their networks.

Today's challenge is how companies can best use the vast amounts of data generated in their supply chain networks. As recently as 2017, "a typical supply chain accessed 50 times more data than just five years earlier".[35].

However, "less than a quarter of this Data was being analyzed. Further, while approximately 20 percent of all supply chain data is structured and easily analyzed, 80 percent of supply chain data is unstructured or dark."[36]. Furthermore, "IDC estimates that by the end of 2020, 50 percent of all business software will incorporate some cognitive computing functions."[37] . Today's organizations are looking for ways to analyze this dark Data best.

[35] ELLIS (Simon), SANTAGATE (John)*: « The path to a thinking supply chain,*» IDC Technology Spotlight, Aug 2018, p.3
[36] Watson customer engagement: *« The A.I. journey: Artificial intelligence and the supply chain*,» IBM, November 2018, p.2.
[37] Mc GOVERN (Matt)*: « Creating a thinking supply chain for the cognitive era,»* Watson Customer Engagement, Mar 2017, p.5.

Figure1.4: Analytics evolution including terminology delimitation



Source: GORMAN (M. F) & KLIMBERG (R. K): *« Benchmarking Academic Programs in Business Analytics. Interfaces »*, INFORMS, June 2014, p.329-341.

2. Supply data analytics definition

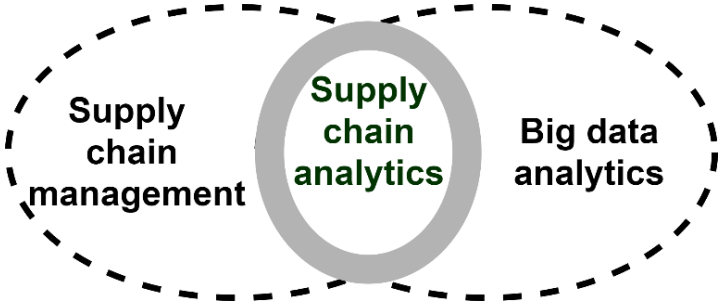
The supply chain is the network of businesses and firms involves in providing goods and services to the last customer, including all the processes. It is not only the management of physical flows but also the management of financial and information flows, and typically, these Supply chain networks generate a considerable amount of data. "Supply chain analytics aims to make sense of all this Data, uncovering patterns and generating insights. In simple words, Supply chain analytics (SCA) means using BDA techniques in order to extract hidden valuable knowledge from the supply chain."[38].

Analytics represents the ability to make data-driven decisions based on a summary of relevant, trusted data, often using visualization in the form of graphs, charts, and other means.

Supply chain analytics is also the foundation for applying cognitive technologies and advanced methods, such as machine learning (ML) and artificial intelligence (A.I.), to the supply chain process.

[38] WANG (G), et al.*: « Big data analytics in logistics and supply chain management: Certain investigations for research and applications,»* International Journal of Production Economics, 2016, p.5.

Figure 1.5: Supply chain analytics



Source : Made by us

3. SCM Big Data sources

The traditional sources of extensive data in the supply chains according to volume and speed about the variables are shown in **Figure 1.6.** The three central ovals of the diagram show primary transaction data, data from the internal supply chain, and various external sources of data. Oval (other data) on the broadest horizon along the black shaft, showing the various natures of data sources outside the supply chains. In addition, the volume and speed of the external data are higher on average than the primary transaction and internal system data.

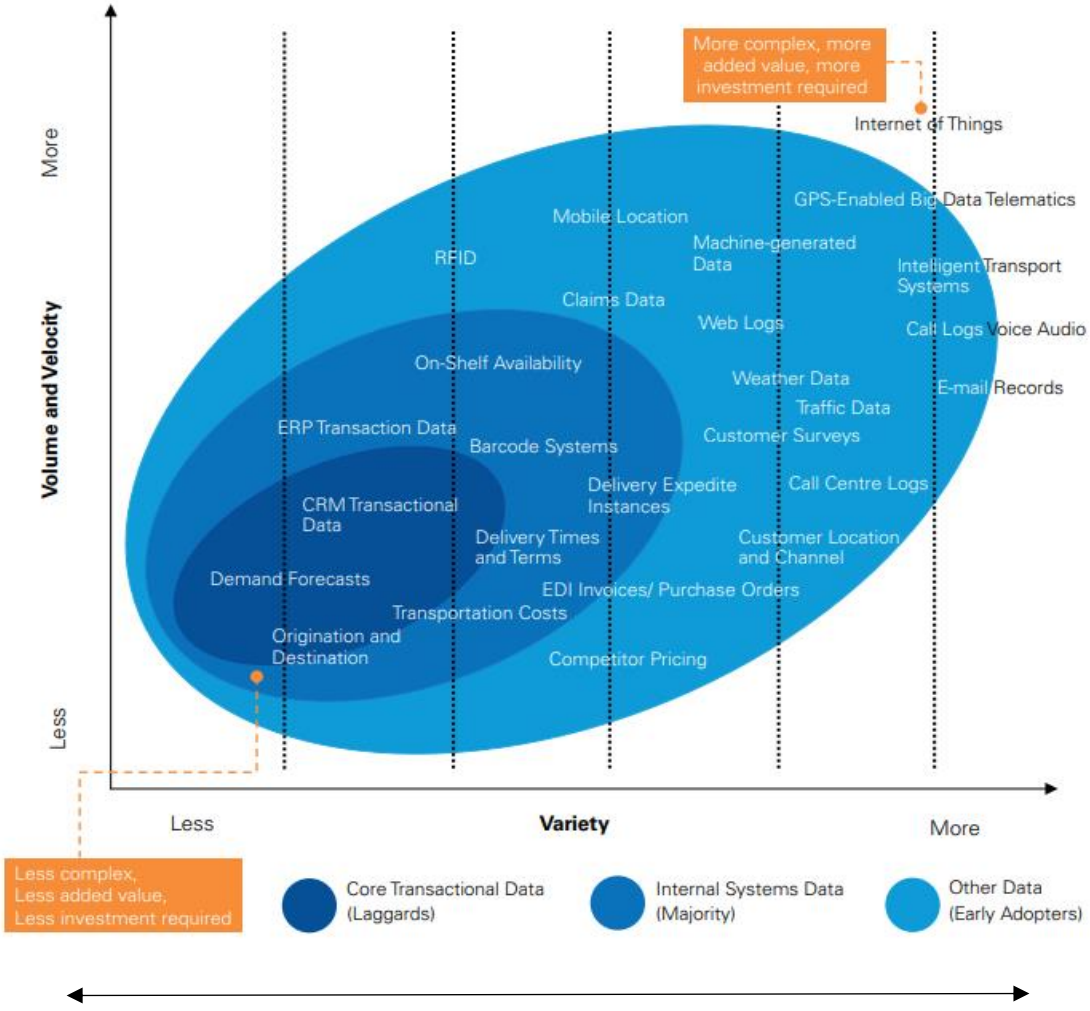
Complete identification of data sources used in work matters and successful implementing guidelines achieved from the systematic review presented a list of 52 primary sources for extensive data through the supply chain. Each of the sources has been reported in one or more of SCM on four handles, with an instance level of 0 (not shown in this handle) to 4 (kernel for processes on this handle). Similarly, each data source is classified according to its volume and speed reported on a scale of 0 to 4. The variety is described in a classification of 3 levels: structured, semi-structured, or unstructured. Although these three subcategories are statistically dependent on the data source scores, in order to facilitate the analysis of interest areas discussed later, they are reported separately,

Figure 1.6 shows the average volume and velocity versus the variety of the data sources in a model such as $\mathbf{E(Y \mid X) = f(X, \beta)}$ with **Y=0.5(Volume +Velocity)** and **X=Variety.**

Each of the three shaded areas contains data sources between primary transaction data, internal or other system data, respectively. The limit of three zones has a much larger horizon when moving on different formats (horizontally) than on the other two (vertical). If the model E above is a linear regression, all β vector parameters are strictly positive. In practice, this fact relates to a positive correlation between significant volumes and the information rate in unstructured formats. Many practitioners and universities support this proposal, and although

there is no decisive quantitative analysis, it is considered a basic rule that 80% of helpful business information is unstructured. Validation of this trend in SCM has clear consequences in the Data Management approach to BDA. Although transaction data in relational databases of different system [39].

Figure 1.6: Supply chain significant data sources (Classification)



Source: ROWE (S), & POURNADER (M)*: « Supply Chain Big Data: how big data is shaping supply chains of tomorrow,»* Series Part 1, KPMG, March 2017, p 8.

[39] ROWE (S), & POURNADER (M)*: « Supply Chain Big Data: how big data is shaping supply chains of tomorrow,»* Series Part 1, KPMG, March 2017, p.8-9.

4. Application of BDA in SCM areas

BDA play a critical role at all operational, tactical, and strategic levels of the supply chain; for example, in the strategic level, SCA is used for product design, network design, and sourcing; in the tactical and operational levels, SCA can also be used for procurement, demand planning, logistics, and inventory.

4.1 BDA in supplier relationship management

Businesses in the supply chain sector need accurate supplier data to develop effective supplier relationship management strategies because establishing a close relationship with key suppliers and enhancing collaboration with them are essential factors in discovering and creating new value and reducing the risk of failure.

Strategic resources and supplier relationship management (SRM) are the success factors of organizations, focusing on relationship management and collaboration by establishing discipline in strategic planning and managing all interactions with organizations' suppliers to reduce the risk of failure and maximize the value of these interactions.

"Using BDA techniques can provide accurate information on organizational spending patterns that help manage supplier relationships. "[40]. For example, big data can provide accurate information on the return on investment (ROI) of any investment and in-depth analysis of potential suppliers. "In a study, fuzzy synthetic evaluation and analytical hierarchy process (AHP) was used to supplier evaluation and selection, given the high capacity of big data processing as one of the evaluated factors has been used, and Organizations will have diverse parameters for gauging which supplier matches their needs and demands". [41].

4.2 BDA in procurement management

Procurement consists of a series of action mechanisms and contracting, and it deals with the relationships at the upstream supply chain. "Logistic organizations, given the high volume of widely dispersed data generated across different operations, systems, and geographic regions, need advanced systems to manage these enormous data, as well as skilled professionals who can analyze these data and extract valuable insights and knowledge into them in order to

[40] JIN (Y), J.I. (S)*: « Partner choice of supply chain based on 3d printing and big data »,* Information Technology Journal, 2013, p.6822.
[41] WANG (G), GUNNASEKARAN (A) & NGAI (E, W): *« Distribution network design with big data: Model and analysis,»* Annals of Operations Research. 2018, p.539-551.

apply them in their planning and decisions". [42], BDA in procurement lies in supply management, spend analysis, and risk prediction. Research on supplier selection and relationship development inherits a conventional process where several selection criteria are set up in the first place; it can be used to manage supplier's performance and supply chain risk.

The application of BDA techniques enables these criteria to be able to expand to a considerable multifaceted scale, which reflects the complexity of supply management the supplier selection often accompanies with the effort of reducing cost through allocating order quantity between key suppliers, improving purchasing process, predicting uncertainties and estimating pricing risk. Moreover, "BDA also expedites risk management through an established warning system of abnormal events, detecting supplier's cartel and corruption. Procurement professionals will also be able to carry out intelligent and predictive supplier selection activities because the application of BDA techniques such as Machine learning and its intelligent algorithms can help these businesses collect and analyze suppliers' data and improve transparency in procurement activities create more opportunities for long-lasting relationships". [43]. By applying BDA techniques such as Machine learning and its intelligent algorithms, it can help organizations collect and analyze suppliers' data, improve transparency in procurement activities, and create more opportunities for long-lasting relationships.

4.3 BDA in manufacturing

"Big data and analytics can already help improve manufacturing. For example, energy-intensive production runs can be scheduled to take advantage of fluctuating electricity prices. BDA in manufacturing is to streamline the manufacturing process by identifying core determinants that influencing S.C. performance as a whole"[44]. Moreover, she was then taking actions to improve them in four aspects continuously: product research and development (R&D), production planning and control (PPC), quality management (Q.M.), diagnosis, and maintenance.

---

[42] SADEGHI DARVAZEH (Saeid), RAEESI VANNI (Iman) & MUSOLU (Farzaneh Mansouri): «*Big Data Analytics and Its Applications in Supply Chain Management,*» March 25th, 2020, p.2.

[43] VAN NGUYEN (T): « *Exploring applications of Big Data Analytics in Supply Chain Management*,» Ph.D. Thesis in Ph.D. Philosophy, University of Greenwich, London, September 2019, p.20-22.

[44] ALICE (K) et al.: « *Big data and the supply chain: The big supply chain analytics landscape,*» McKinsey & Company Operations Extranet, 1(of 2), 2016, p.3-4.

Product R&D heavily relies on extracting various forms of customer feedback. Through BDA, the result of these feedback analyses can lead to enhanced innovation capability and ongoing monitoring products.

"Applying BDA in diagnosing machine fault and planning for maintenance is another area that enhances the S.C. performance. It is expected that the system would automatically detect the failure and be capable of taking action without human intervention. The fault refers not only to machines but also workers' abnormal behaviors".[45].

4.4 BDA in warehousing

Warehousing is one of the most critical processes in the logistics chain because the efficient handling and storing of products and materials are a vital role that directly impacts the supply chain performance and the traditional way of managing the warehouses with legacy or ERP systems is now outdated. Nowadays, many factors are affecting the way warehouses are managed, such as customer expectation and globalization, etc. This led to the development in Warehouse management (particularly inventory management) by modern identification systems after the successful introduction of RFID. Within this group, the most significant data clusters are related to an automated sensing capability, mainly as the Internet of Things and extended sensors, connectivity, and intelligence to material handling and packaging systems applications evolved.

Big data analytics is one of the best techniques to help organizations manage their warehouses and distribution centers. With this technique, warehouse managers can understand how customers' behavior changes and their expectations from manufacturers and supply chain managers. Also, the warehouse managers can have a detailed insight into the process of loading, carrying, unloading, and delivering. As a result, they can plan the routes and schedule the deliveries better to increase safety and decrease petrol expenses.

"Data mining-based storage assignment methods have been used to process data on orders, products, and customers constantly and automatically saved in warehouse management

[45] GUO (S.Y.) et al.: *«A Big-Data based platform of workers behavior: Observations from the field,»* Accident Analysis & Prevention, Elsevier Ltd, 2016, p.299-309.

systems (WMS) and Enterprise Resource Planning (ERP) systems. In addition, mining of data collected through RFID".[46].

BDA applications in this supply chain function have focused "firstly on material handling and layout zoning to maximize space utilization, minimize the distance travelled to fulfill orders, and consequently, minimize storage and material handling costs and risk of hazard events. One example is "chaotic" storage approaches that enable the efficient use of warehouse space and minimize travel distances for personnel".[47].

Second, on inventory availability, a fundamental element of customer service is inventory availability. Therefore, inventory control has significant impacts across the whole SC process directly dictates warehouse operations. Its primary task is to determine the correct stock level that balances the inventory holding cost and the cost of lost sales. From this viewpoint, BDA has emerged as a critical tool to support inventory managers. Indeed, "the adoption of BDA stimulates collaborative planning, forecasting, and replenishment (CPFR) implementation, which enables the real-time access, amalgamation, and extraction of massive data from heterogeneous inventory points including procurement, production, distribution, and point of sale in order to produce fast, accurate, and reliable inventory replenishment predictions".[48].

4.5 BDA in Transportation

Regarding transportation management, the ultimate objective of BDA adoption is to develop an Intelligent Transportation System (ITS) that allows real-time traffic operation control and proactive safety management. "Traffic operation efficiency can be improved by accurately and timely predicting short-term traffic flow demand or using intelligent routing to avoid traffic congestion, while safety is mainly studied through crash analysis".[49].

Transportation Truck companies already make use of analytics to improve their operations. For example, they use fuel consumption analytics to improve driving efficiency;

---

[46] CHIANG (D), LIN (C.-P). And CHEN (M.-C*): «The adaptive approach for storage assignment by mining data of warehouse management system for distribution centres,»* Enterprise Information Systems, 2011, p 219–234.

[47] ALICE (K) et al.: *« Big data and the supply chain: The big supply chain analytics landscape,»* McKinsey & Company Operations Extranet, 1(of 2), 2016, p.4-5.

[48] PRAJOGO (D), & OLHAGER (J): *« Supply chain integration and performance: The effects of long-term relationships, information technology and sharing, and logistics integration*,» International Journal of Production Economics, 2012, 135(1), p.514–522.

[49] WANG (Y): *« Robust causal dependence mining in big data network and its application to traffic flow predictions*,» Transportation Research Part C: Emerging Technologies, 2014, p.292–307.

and they use GPS technologies to reduce waiting times by allocating warehouse bays in real-time. In addition, courier companies have started real-time routing of deliveries to customers based on their truck's geo-location and traffic data. UPS, "for example, has spent ten years developing its On-Road Integrated Optimization and Navigation system (Orion) to optimize the 55,000 routers in the network. The company's CEO David Abney says the new system will save the company $300 million to $400 million a year".[50].

Big analytics will also enable logistics providers to deliver parcels with fewer delivery attempts by allowing them to mine their data to predict when a particular customer is more likely to be at home. In addition, companies can cut costs and carbon emissions on a more strategic basis by selecting suitable transport modes. Analytics also can enable the organization to understand when goods need to be shipped rapidly by truck or when there is time for slower barge or train delivery.

To logistics and transportation planners, the primary concern on inventory management is defining appropriate logistics plans to maintain the quality and safety of the product during the in-transit process. From this viewpoint, BDA provides unprecedented opportunities for tracking, assessing, and monitoring the product conditions following the in-transit context, including temperature, vibration, moisture, light exposure, and humidity level.

4.6 BDA in demand planning

Many supply chain executives are keen to improve demand forecasting and production planning with big data. Accurate demand forecast has always been a significant puzzle on SCM; BDA can determine trace consumer loyalty, demand signal, and optimal price data. However, one of the challenges the organizations face is applying advanced hardware and software and algorithm architecture, which can drive creativity and innovation in business services.

Demand sensing combined with BDA techniques has enabled companies to incorporate detailed short-term and real-time demand data into their forecasts. For instance, some studies have demonstrated how online reviews and social media data can determine the predictors of product sales in both e-commerce and retail businesses. Using sentiment analysis and text clustering, these studies were able to translate unstructured comments to describe and predict customer behavior, where BDA allows identifying new market trends and determining root

[50] http://www.wsj.com/articles/at-ups-the-algorithm-is-the-driver-1424136536 (30/04/2021 at 15:30)

causes of issues, failures, and defects. Data analytics can predict customers' preferences and needs by examining customer behavior, driving creativity and innovation in business services.[51]

"Another research stream has made efforts to analyze the readership and helpfulness of online comments for both vendors and consumers." [52]. Search traffic information, loyalty cards, and mobile network data have also proven helpful in sensing demand through customer segmentation. Detecting customer behavior and predicting market volatility have underscored the opportunity to sense and react near real-time to changes in the demand.

In one study, "a model was presented to predict the electric vehicle charging demand that used weather data and historical real-world traffic data. This model enables operators to plan the generation profiles and operation by determining the charging demand.[53] Another study presents a model for predicting demand for air passenger demand, which uses big data to estimate air passenger demand. The results of this study show a 5.3% prediction error".[54].

After forecasting and sensing demand, BDA can be applied to shape demand, such as price management, marketing, and advertisement, managing online reviews, long-term marketing campaigns and planning policies, improving branding, improving customer experience, and product life cycle management.

---

[51] SADEGHI DARVAZEH (Saeid), RAEESI VANNI (Iman), & MUSOLU (Farzaneh Mansouri): «*Big Data Analytics and Its Applications in Supply Chain Management*,» March 25th, 2020, p.2.

[52] SALEHAN (M), KIM (D.J.): *«Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics*,» Decision Support Systems, ElsevierB.V.,2016, p.30–40.

[53] KIM (S): *« Forecasting short-term air passenger demand using big data from search engine queries*.» Automation in Construction. 2016, p.98-108.

[54] FAN (J), HAN (F), LIU (H): *« Challenges of big data analysis*.» National Science Review, 2014, p.293-314

Table 1.3: Some examples of practical applications of BDA in SCM

SCM lever	Functional Problems	Type of data	BDA proposed solution	BDA techniques
Marketing	Sentiment analysis of demand new trends	Blogs and news feeds, ratings and reputation from third parties, web lop; loyalty programs, call centers records, customer surveys.	1. Create lexicons from training datasets that identify key terms that relate to the demand of a product. 2. Integrate all data sources that relate to a product into a unified text corpus. 3. Use supervised learning algorithms to predict sentiment scores of the corpus' term-document matrix based on training datasets_	Natural language processing Text mining with R to package: (Corpus, term-document matrix) Logistic regression, random forests, CART, Naive Bayes, k-NN;
Procurement	Informing supplier negotiations	SRM Transaction data, Supplier current capacity & top customers, supplier financial performance information	1. Capture performance requirements for procurement contracts (quality measures). 2. Require or publicly capture data regarding previous supplier transactions with other third parties in similar characteristics (delivery locations, lead times).	Suitable supervised learning algorithms, expert systems modeling
Warehouse Operations	Warranty Analytics	Internet Of sensing, user demographics, historical asset usage	1. Aggregate multiple sensing sources in real-time with reports on monitored assets together with user demographics. 2. Aggregate patterns in user and usage clusters in order to generate multidimensional segmentation.	T-distributed stochastic neighbor embedding (t-SNE)
Transportation	Real-time optimization	Traffic density, weather conditions, transport systems constraints, intelligent transport systems, GPS-enabled Big Data telematics	1. In order to address time variability for deliveries in predefined networks, model the delivery network and update it with the current position of delivery units. 2. New requirements for delivery are entered into the system. Taking into account all network availability factors, from each delivery unit, a spatial regression predicts the time/cost of serving a delivery to another point of the network_	Spatial regression modeling

Source: VARELA (I. R), & TIAHIONO (B): *« Big data analytics in supply chain management: trends and related research*.» 6th International Conference on Operations and Supply Chain Management, 1(1), Bali, 2014, p.12.

5 Analytics on SCM

Big data create different supply chain capabilities that provide networks with greater data accuracy, insights, and clarity and create a greater e-contextual intelligence shared across the supply chains. BDA is a powerful tool for solving supply chain issues and driving supply chains ahead. The use of intelligent, embedded analytics benefits every part of a supply chain at all levels. To succeed in competitive markets, the supply chain decision-makers must continually seek ways to effectively integrate and manage significant data sources to gain more value and competitive advantage. The practical and appropriate use of significant data sources and techniques resulted in enormous improvements in processes of the supply chain:

- **Planning and scheduling:** Planning and scheduling are some of the essential parts of any supply chain. If an organization Gets this wrong, it is going to lose much money. Supply chain analytics provide better insights into inventory levels across different locations. It also gains high-quality decision support through near real-time information, which can be vital if something goes wrong.

- **Improved responsiveness:** Supply chain analytics help to determine what stock is needed. Organizations will predict what is likely to happen to determine what items to buy, what items to discard, and what items they will need more of—armed with this information. As a result, it can be much more responsive to trends and events.

- **Improved demand planning:** With supply chain analytics, an organization can examine past trends to predict the stock and what amount the customers will require in the future. They can see what items are selling well and what items did not sell well- This increases customer satisfaction and can inform product development.

- **Order optimization:** Supply chain analytics lets organizations optimize both the items they are ordering and the overall ordering process- they can increase the number of on-time orders. By Decreasing the cost of acquiring items and making sure they get what they order. Order optimization brings upstream benefits as well. For example, production is more efficient when assembly lines are not waiting for parts.

- **Real-time supply chain execution:** organizations can see their orders and supply chain in real-time and be alerted as soon as an issue arises. They do not have to worry about where the orders are, what is happening with them, or whether they need to make changes. Taking this one step further, prescriptive-based analytics would not just see what happened; it shows the options they have to do something about it right then — for example, what other suppliers are ready to source from immediately.

- **Inventory management:** With supply chain analytics, organizations can plan, forecast. Furthermore, fully optimize their inventory, so they do not waste space or waste money with stock that may or may not be working the way it should. They can look across networks to consider consumption rates, inventory levels, and other aspects of their supply chain. As a result, they can create the right balance of inventory to keep the supply chain stocked adequately while releasing cash for other activities.

- **Replenishment planning:** supply chain analytics enable the supply chain profession to keep a close eye on inventory levels, noting where more stock Will be needed and on what schedule. Further, they can see across their network of trading partners to discover who has the stock that you need and how you can avoid overstocking when it is not required.

- **Improved partner performance:** Supply Chain analytics give much greater visibility into the trading partners' activities and performance. Organizations can see quickly which ones are delivering the most profitable business for them. They can determine which suppliers are suitable to work with and are reliable when an order is placed. Armed with this information, they are well-positioned for contract and price negotiations.

- **Improved cash management:** By helping to focus on elements like delivering perfect orders and controlling inventory levels, supply chain analytics helps with cash management. An organization can deliver more orders on time and in full. They can reduce the errors and exceptions that drain profitability and damage customer

relationships. They shorten the order-to-cash and cash-to-cash cycles and grow revenue by reducing waste.[55]

- **They build agile or responsive supply chains** by predicting and gaining a better understanding of the market trends and customer expectations and preferences. BDA can facilitate the real-time monitoring of the supply chain and manage data that enhance the speed, quality, accuracy, and flexibility of supply chain decisions. Utilize a wide range of data from news, social media, weather data (SNEW), and events, as well as direct data inputs from multiple static and dynamic data points, provide the capability to predict and proactively plan all supply chain activities.
- **They build reliable and intelligent supply chains** by applying the Internet of Things (IoT), machine learning, and deep learning techniques in each supply chain activity. For instance, IoT can provide real-time telemetry data by the real-time monitoring of the supply chain to reveal the details of production processes. Machine learning algorithms that are trained to the data can accurately predict imminent machine failures. Deep learning techniques can also be used to predict customers' demand and their preferences and expectations accurately.

- **It is supporting the creation of sustainability in SCM**. BDA undoubtedly will enhance social, environmental, and financial performance measures. For example, detailed planning for timely delivery of the product can be done by analyzing the real-time traffic data provided by the GPS that reduces the production of carbon emissions and the cost of fuel consumption.

- **They enable global supply chains to adopt a preventive** rather than a reactive measure to supply chain risks (e.g., supply failures due to natural hazards or fabricated, contextual, and operational disruptions). BDA techniques can help supply chain managers predict future external events and adopt a proactive against them in a more complex global supply chain. [56]

[55] WILEY (J), SONS (L): «*Supply chain analytics for dummies*,» Open-Text, special edition, The atrium southern gate Chichester, 2017, p.14-16.

[56] SADEGHI DARVAZEH (Saeid), RAEESI VANNI (Iman), & MUSOLU (Farzaneh Mansouri): «*Big Data Analytics and Its Applications in Supply Chain Management*,» March 25th, 2020, p 2.

Conclusion

Through this chapter, we have presented the concepts of big data analytics and Supply chain analytics by first addressing the explosion of data in the last years, which led to the development in the tools, techniques, and methods of analyzing it. We have also shed light on the most important sources of data generated by the supply chain network and the importance and benefits of applying big data analytics techniques in supply chain management at all levels.

We conclude from the above that in the current marketplace, the companies cannot withstand big data analytics because big data analytics is the most crucial factor in building strong supply chains and contributes to the development of new mechanisms to confront risks even before they occur. Hence, Big data analytics capability (BDA) is one of the best techniques, which can help organizations to overcome their problem by providing a tool for extracting valuable patterns and information on a large volume of data generated by the supply chain networks

Chapter two:

Electronic customer relationship management

Introduction

The increasing development of the Internet has led to new communication channels with customers (electronic communication channels) through mobile applications and social networking.

In the current competitive environment, the companies are starting to realize the importance of building a solid relationship and understanding customer behavior with customers, and supporting their customer relationship management at all levels. This is what is known nowadays by electronic customer relationship management (e-CRM).

In this chapter, we are going to present and discuss electronic customer relationship management (e-CRM). We will try to present its definition, elements, components, and its applications. Then, we will explore how the extensive data analytics technique used in extracting insight and value from the considerable amount of data generated by the e-CRM and how it can be used to improve customer relationship management at all levels.

Keywords: electronic communication channels, electronic customer relationship management (e-CRM), customer relationship management.

I. Section one: Electronic customer relationship management

1. Customer relationship management an overview

1.1 Evolution and growth of customer relationship management

The evolution describes the stages of the development of CRM. From the 1990s until today, it has grown from a tactical marketing tool to a strategic element in all marketing decisions. The growth of the Internet also has increased the adoption rate of CRM in many industries. To provide a historical perspective.

1.1.1 First Generation (Functional CRM): The collection of activities that later took on the umbrella acronym CRM developed initially as two independent product offerings:

- **Salesforce automation (SFA):** These products addressed presales functions such as maintaining prospect and customer data, telemarketing, and generating leads.
- **Customer service and Support (CSS):** activities, such as help desks, contact and call centers, and field service support.

Although fragmented and poorly integrated with the back office, early SFA/CSS applications delivered the promise of sales and service improvements, though their combined market niche remained small.

1.1.2 Second Generation (Customer Facing Front-End Approach): Innovations in CRM during the 1990s matched those of ERP, including the integration of different independent subsystems into one package, CRM technology was expected to fill the gaps left by ERP functionality and address the business needs of the company's customer-facing front end. For the most part, this goal was not achieved during the 1990s, leading to increased disillusionment with CRM technology and implementations. Industry observers began talking about the demise of CRM. Even as the Internet fuelled new expectations, it became clear that revenue increases through technology were challenging to implement, realize, and measure without a more strategic understanding of the process.

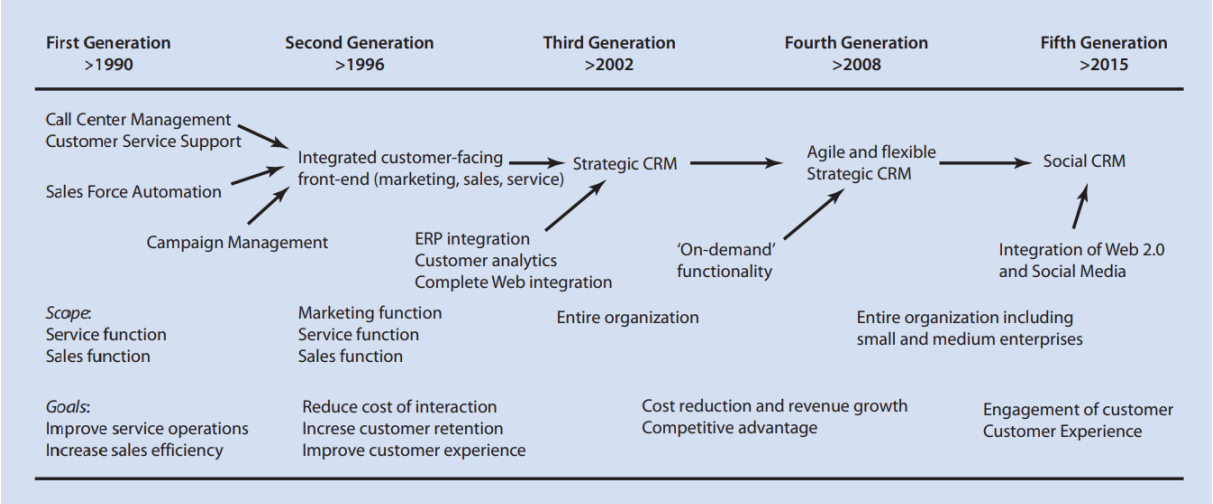
1.1.3 Third Generation (Strategic Approach): By the end of 2002, the CRM market had started to pick up, and the gap between customers' perceived value and value realized was closing. The integration of Internet technology helped to boost CRM. Companies learned from experience and their failure to implement prior versions of CRM and recognized the eventual goal of CRM to grow revenue, not just control costs.

1.1.4 Fourth Generation (Agile and Flexible Strategic CRM):

At the end of the first decade of the twenty-first century, we face the start of the fourth generation of CRM. Strategic CRM is widely accepted and established as an essential element of the marketing strategy. An ever-increasing number of small and medium-sized companies adopt this management tool and its related technologies to drive their businesses Agility, flexibility, and low fixed costs are critical—the emergence of social media and increased self-service, as well as the growing prevalence of web-based services. CRM technology on a pay-per-use basis can provide on-demand functionality.

1.1.5 Fifth Generation (Social CRM): The development of the new technological advances and the unprecedented reach of social media gave rise to the fifth generation of CRM, social CRM, which is characterized by the engagement of the customer through the integration of the web 2.0 and social media and by the use of data-driven insights to optimize the overall customer experience. Companies encourage active customer participation online while they use software applications to track real-time social data. Additionally, the combination of data across different social media platforms allows companies to determine the customer value based on profitability and based on their online behavior in terms of referrals, knowledge dispersion, and influencing other social media community members.[57]

Figure 2.1: Timeline of the CRM evolution



Source: KUMAR (V), & REINATRZ (W): « *Customer relationship management, concept, strategy, and tools*, » Third edit, Springer Texts in Business and Economics,2018, p.12.

[57] KUMAR (V), & REINATRZ (W): « *Customer relationship management, concept, strategy, and tools*, » Third edit, Springer Texts in Business and Economics,2018, p.12.

1.2 The definition of CRM

Customer Relationship Management (CRM) could be defined as the infrastructure that enables the delineation of an increase in customer value and the correct means to motivate valuable customers to remain loyal - indeed, to buy again. [58]. CRM is the core business strategy that integrates internal processes and functions and external networks to create and deliver value to targeted customers at a profit. It is grounded on high quality customer-related data and enabled by information technology.

In summary, CRM is a technology-enabled approach to the management of the customer interface. Most CRM initiatives expect to have an impact on the costs to serve and revenue streams from customers. The use of technology also changes the customer's experience of transacting and communicating with a supplier. For that reason, the customer's perspective on CRM is an essential consideration in this book. CRM influences customer experience, and that is of fundamental strategic significance.[59]

1.3 CRM Types

Due to the CRM functions, it can be divided into three basic types Operational CRM, Analytical CRM, Collaborative CRM. [60].

- **Operational CRM**, often referred to as front-office CRM, covers most areas of customer-company contact. CRM applications collect, process, and store data about customers so that later this data can be used in analytical CRM [61], it is the component that helps to improve the efficiency of day-to-day customer operations. Therefore, this is concerned with automating processes involving communication and interaction with customers and integrates front, back, and mobile offices. This CRM component is the initial producer of data and includes typical corporate functions of sales automation, enterprise marketing automation, order management, and customer service or Support.[62]

---

[58] DYCHE (Jill): « *The CRM Handbook: A Business Guide to Customer Relationship Management*,» Addison-Wesley, Information Technology Serie, Boston (Mass.), 2002, p.30.
[59] BUTTLE (F), & MAKLAN (S): « *Customer Relationship Management, Concept and Technologies*,» Third Edition, Routledge, 2015, p.3.
[60] JELINEK (D): « *The Role of the CRM System in the Development of the Customer Capital of An Enterprise Processes of Capital Supply in Production Enterprises*,» Serie Monographs No 1, 2006, p.103-109.
[61] JELINEK (D), CHLUSKI (A): «*Possibilities of using CRM systems in health care institutions*,» Information technologies in public administration and healthcare., SGH Publishing House, Warsaw, 2010, p.42-43.
[62] PEPPERS (D), RODREGS (M): *« Managing Customer Relationships; A Strategic Framework*,» 2nd Edition. U.S., 2010

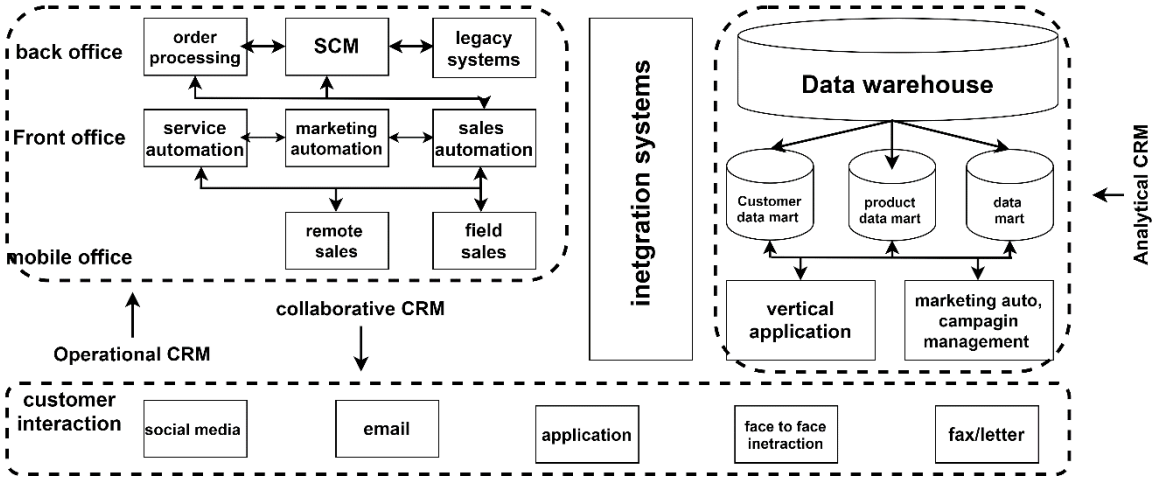
- **Collaborative CRM** can be seen as a communication center connecting companies and their customers, suppliers, and business partners. Indeed, it allows customers, staff, salespeople, and partners to access, distribute and share data. Personalized publishing, email, communities, conferences, and web-enabled relationship interaction centers are examples of collaborative services. These services make teamwork more straightforward and more productive, enabling companies to improve processes and improve customer satisfaction.[63]

- **Analytical CRM** is also known as a back-office CRM. It involves understanding the customer's activities that occurred in the front office.[64]

Analytical CRM mainly focuses on analyzing the data collected and stored to create more meaningful and profitable customer interactions. By analyzing customer data, companies can understand behaviors, identify buying patterns and trends and discover causal relationships. The insights obtained from the data help to model and predict future customer satisfaction and behavior more accurately and may constitute a quantified foundation for strategic decision making,

Analytical CRM solutions allow the management of the relationship with customers effectively, and to achieve this purpose, the data is processed, interpreted, and reported using several tools.[65]

Figure2.2: The model of CRM system that includes all the types



Source: JELINEK (D) "*The Evolution of Customer Relationship Management System."* Proceedings of the 19th International Conference on Computers, 2015, p.30.

[63] GREENBERG (P): «*CRM at the Speed of Light: Capturing and Keeping Customers in Internet Real Time.* », McGraw-Hill (Osborne), USA, 2001, 1st Edition.
[64] DYCHE (Jill): « *The CRM Handbook: A Business Guide to Customer Relationship Management*,» Addison-Wesley, Information Technology Serie, Boston (Mass.), 2002, p.34.
[65] GREENBERG (P): *« CRM at the Speed of Light: Essential Customer Strategies for the 21st Century*», McGraw-Hill Osborne Media; Third Edition, January 1, 2004

2. Understanding electronic customer relationship management

2.1 E-CRM Definition

Advancements in technology, computer, and telecommunication have supported the development of internet technology. The rapid advancement of technology in all sectors has also encouraged the progressive development of the CRM concept, leading to the emergence of the electronic Customer Relationship Management (e-CRM) concept.

E-CRM concept refers to the set of marketing activities, tools, and techniques that make it possible to build and improve the relationship between the company and its customers and develop relationships with existing and potential customers via the Internet [66], it can be described as the use of technology to support customer relationship management strategy and is seen as the integration of customer relationship management with e-business applications.[67]

The development of marketing through online media is increasing in which e-CRM helps manage the relationship with customers and provides several benefits for the companies like increasing the companies' profit, creating satisfactory services by using integrated information and displaying consistency, procedure, and process of handling problems. [68]

2.2 Differences between CRM & E-CRM

The differences between CRM and E-CRM are subtle but essential; they concern the underlying technology and its interfaces with users and other systems. For example, E-CRM provides the ability to take care of customers via the Web, or customers being able to take care of themselves online. Many E-CRM systems provide the customer with a self-service browser-based window to place orders, check order status, review purchase history, request additional information about products, send emails and engage in a mass of other activities. The customer is no longer limited to contacting an organization during regular business hours, and the organization does not have to provide a live contact at the other end for customer inquiries and

---

[66] METTARGARUNAGUL (M), PENSIRI (P): « The use of e-CRM to enhance customer relationship: *the case of Toyota Mahanakorn*,» Master thesis in information technology management, Malardalen University (Sweden), 2011, p.14.

[67] KENNEDY (A): « *Electronic Customer Relationship Management (e-CRM): Opportunities and Challenges in a Digital World*.» Irish Marketing Review, 2006, 18(1). P.63.

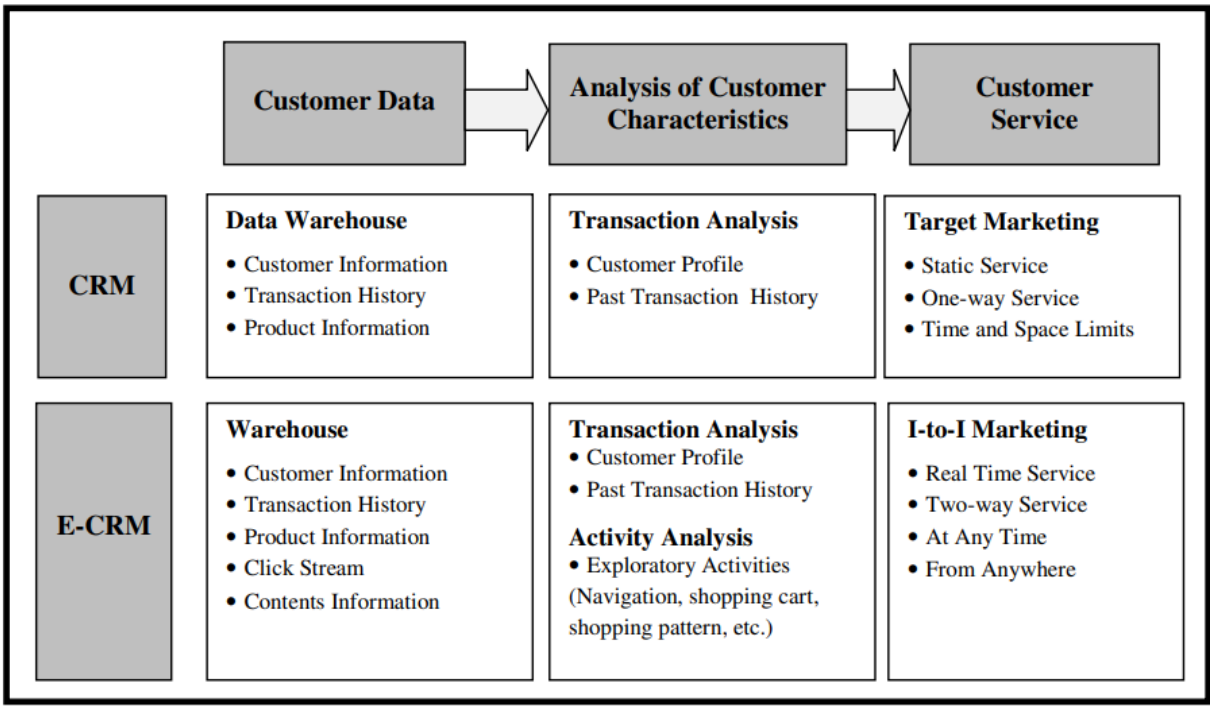
[68] KIM-SOON (N), ZULKIFLI (M. F): « *The impact of e-CRM Business performance of small company*.» Journal of Engineering and Technology 3, p.145

requests. In effect, within an E-CRM environment, customers do most of the work for themselves and not for the businesses. [69]

CRM applications are designed around products and job functions. In E-CRM, the applications are designed with the customer in mind. They give the customer the 'total experience' on the Web. Web-enabled CRM is usually designed around one department or business unit, not the entire enterprise. In E-CRM, on the other hand, applications are designed for the entire enterprise, including all customers, suppliers, and partners. Traditional CRM has limitations in supporting outside multi-channel customer interactions that combine telephone, the Internet, email, fax, chat, and so on. Unlike traditional CRM, current E-CRM solutions support marketing, sales, and service.[70]

Figure 2.3 below summarises the main differences between the two technologies CRM & E-CRM.

Figure 2.3: The Differences between CRM and E-CRM



Source: PAN (S. L), LEE (J. N): *« Using e-CRM for a unified view of the customer*,» Communications of the ACM, 46(4), 2004, p.95–96.

[69] CHANDRA (S), STRICKLAND (T. J)*: « Technological Differences Between CRM & E-CRM.» Issues in Information Systems, 5*(2),2004, 408–413.
[70] PAN (S. L), LEE (J. N): *« Using e-CRM for a unified view of the customer*,» Communications of the ACM, 46(4), 2004, p.95–96.

2.3 E-CRM components

E-CRM is the "care component" of customer relationship management for companies in e-commerce and a critical element for innovation and improvement of existing products. This can improve customer satisfaction by creating more value for customers and differentiating the company from competitors, with higher quality products, lower prices, quicker responses, better performance, and convenient services. E-CRM comprises four main components: *Engage, Purchase, Fulfill, and Support.*

2.3.1 Engage: E-CRM enables companies to attract more customers for online shopping engagement. Products or services are displayed more significantly on popular online search engines: Google, Yahoo, Facebook, Twitter, Instagram, or commerce sites: Bay, Amazon, and Alibaba. Customers feel more satisfied because they can get more information about the products.

2.3.2 Order: E-CRM enables companies to effectively improve the order process with enterprise resource planning applications, from inventory and processing to shipping. Customers feel more convenient and accurate to place orders because they can change, modify, add or cancel orders anytime, as long they are not shipped. They can track shipment status over the Internet by the tracking number. This helps to improve customer satisfaction and decrease customer frustration caused by complex order systems or a lack of communication. Customers can select different payment options, such as credit cards or electronic invoices. Hence, it costs less to obtain valuable customers.

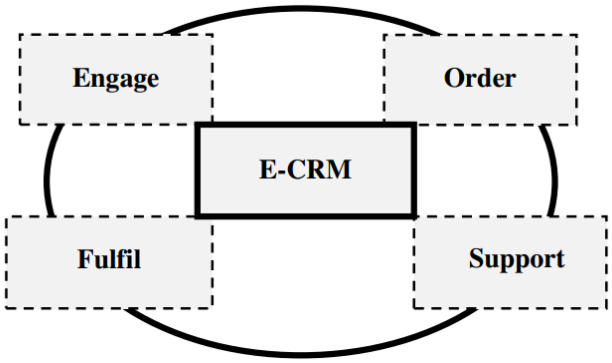
2.3.3 Fulfill: E-CRM enables companies to fulfill responsibilities to customers by assuring that they will receive the proper order at the right time. E-CRM allows companies to track and compensate for loss or damage of products caused by their own- or third-party processing and packaging.

2.3.4 Support: E-CRM enables companies to effectively notify customers, through the database of customer information, about updates or upgrades of products or services. The "follow" function on web browsing enhances the link between companies and customers. This helps to increase customer satisfaction by solving customer problems or responding quickly to their inquiries. Customers can get direct Support through real-time chatting, voice chats, forums, and other online browser features provided by companies. [71]

---

[71] JULTA (D), CRAIG (J), & BODORIK: *« Enabling and Measuring Electronic Customer Relationship Management Readiness*,» in the Proceedings of the 34th Annual Hawaii International Conference on System Sciences, Big Island (Hawaii), Organizational Systems and Technologies Track, Hawaii, 2001, p.15.

Figure 2.4: E-CRM components



Source: made by us

2.4 E-CRM applications

The following is a set of e-CRM capabilities that collectively comprise a total e-CRM solution:

2.4.1 Customer analytic software predicts, measures, and interprets customer behaviors, allowing companies to understand the effectiveness of e-CRM efforts across both inbound and outbound channels. Analytics provides a set of e-CRM metrics that describe the state of the customers across the enterprise. Customer analytic software, and its reports, can be used by virtually anyone across the organization, from marketers and salespeople to corner-office executives. Ideally, analytics should be delivered as a packaged but extensible application that uses a set of industry-specific data models and reports. This packaging will help companies implement solutions quickly and limit risk. Most importantly, customer analytics should integrate with customer communications software to enable companies to transform customer findings into ROI-producing initiatives.

2.4.2 Data mining software builds predictive models to identify customers most likely to perform a particular behavior such as purchase an upgrade or churn from the company. Modeling must be tightly integrated with campaign management software to keep pace with multiple campaigns running daily or weekly. For example, a marketer might select a propensity model from an existing library to dynamically score individuals within an already defined customer segment. He or she might then select individuals with scores above a certain threshold to include in an up-sell campaign. This segment-selection process improves response rates and campaign effectiveness and lowers campaign costs by reducing the size of the original target segment.

2.4.3 Campaign management software leverages the data warehouse to plan and execute multiple, highly targeted campaigns over time, using triggers that respond to timed events and customer behavior. For example, a retail campaign may present a high-profit customer with a gift on her birthday or send an email promoting various specials if the customer has been silent for three months. Campaign management software tests various offers against control groups, capture promotion history for each customer and prospect, and produces output for virtually any online or offline customer touchpoint channel, such as personalized e-messaging, call centers, branch offices, direct-mail

2.4.4 Business simulation, used in conjunction with campaign management software, optimizes offers, messaging, and channel delivery before executing campaigns and compares planned costs and ROI projections with actual results.

2.4.5 A personalized e-messaging system should deliver either text or HTML pages, scale to support very high volumes, have an automated mechanism to answer responses, and drive recipients to Web pages through trackable URLs embedded within messages. For example, a company might include three trackable URLs within a particular email and monitor success at driving an individual customer or prospect to one, two, or all three sites, and in which order.

2.4.6 A real-time decision engine coordinates and synchronizes communications across disparate customer touchpoint systems. It contains business intelligence to determine and communicate the most appropriate message, offer, and channel delivery in real-time and support two-way dialogue with the customer. Its decision powers assess a customer's activity at the touchpoints (i.e., Web, email, or customer care) in conjunction with that customer's profile stored in a data warehouse. This information responds to an inbound message, selecting from any number of staged offers or messages. For example, the decision engine might inform a Web personalization product of the best offer to present to a Web visitor in real-time. Business rules determine responses. For example, a rule might require a sales representative to wait one day before contacting that individual if a visitor searches a Web site for its mutual funds. [72]

[72] FRAWLEY (A): *« Evolving to e-CRM: How to optimize interactive relationships between you and your customers,»* An exchange, Inc white paper (First Serie), 2003, p.11-12.

2.5 Framework of E-CRM

Companies need to take firm initiatives on the E-CRM frontier to optimize the value of interactive relationships and coordinate marketing initiatives across customer channels. So, E-CRM must address customer optimization along three dimensions:

- **Acquisition:** By increasing the number of customers
- **Expansion:** By increasing the probability
- **Retention:** By Increasing the amount of time that customer stays

Today, although CRM is on top of a firm's priorities, there are many unclear points about its definition and role. In order to fully understand these ambiguities, we should view CRM from a strategic and systematic approach and recognize the components of customer value. A conceptual framework with a strategic approach for E-CRM covers every part's definition and functions and includes adapting the model to different firms, which helps to use E-CRM as a managerial, strategic approach. This framework is similar to the porter value chain in shape and in having value-adding activities. In the porter value chain, two types of activities:

- **Primary Activities** (including inbound logistics, Operations, Outbound Logistics, Marketing and sales, and after-sales service)
- **Supportive or Secondary Activities** (including firm infrastructures, human resource management, R&D, and procurement) run through a value chain that at last result in athletic goal achievements and gaining of competitive advantage.

Customizing the framework for a customer-centric approach, we mention that the only thing that determines the final profitability is the customer and its satisfaction and loyalty. So, all activities that help this would be value-adding and vital to the firm's success and position in the market. However, these CRM-related activities could not be practical without supportive activities helping primary ones run parallel to them. In this framework, similar to the porter value chain, we have two types of activities, Primary and Supportive ones, with the difference being customer-centric and CRM related. As the company runs its primary activities of Customer Acquisition, Retention, and Expansion, it needs supportive activities such as channel integration and information management. Information management activities deal with customer databases, including three types of information: customer information, customer information, and customer information. An essential point in the framework is the initial phase which is Strategy Formation that should be considered before any CRM implementation

2.5.1 Strategy formation

Before speaking about CRM technological issues, the organization should consider CRM in the context of their overall business strategies. This task needs a two-dimensional emphasis:

- On business strategy
- On customer strategy of the firm

The more these two dimensions are interrelated with each other, the more successful CRM strategy would be. Through developing and reviewing business strategies, organizations identify their fundamental capabilities and the way they could transfer them as valuable to their consumers.

In customer strategies, customer types and their categories would be identified. In this categorization, customers may be classified as contradictory to their expectations and preferences. They also could be put to a category that is not correct from their point of view. E-CRM provides the ability to have a one-to-one relationship between firms and customers, satisfy their expectations and solve this problem.

2.5.2 Multiple channel integration activity

One of the keys and strategic components for moving toward E-CRM is to develop a consolidated and relationship-based strategy for existing channels (old ones) and web-based technologies (new and modern ones) that are designed to enhance the relationships with customers (not only reducing costs). So, it can be mentioned that E-CRM is multiple channels that get consolidated. If the goal of the channel strategy is only cost reduction, undesirable effects will appear. It should be considered that a scattered channel strategy leads to customers' dissatisfaction, which eventually abolishes the whole business. So, channel strategy should:

- Provide integration for all channels to get integrated to use one customer base answer this question: what kind of interaction, customers prefer to have with the firm.
- Consider which channel is appropriate for what kind of customers and for what type of transactions. [73]

[73] IDEM

2.5.3 Information management activity

This process includes customer data collection from all touchpoints and omission of redundancies, and using them to create an up-to-date and complete customer profile to enhance the quality of interaction between firms and customers. In this process, some tools are handy and should be used appropriately and in incoordination. These are listed as below:

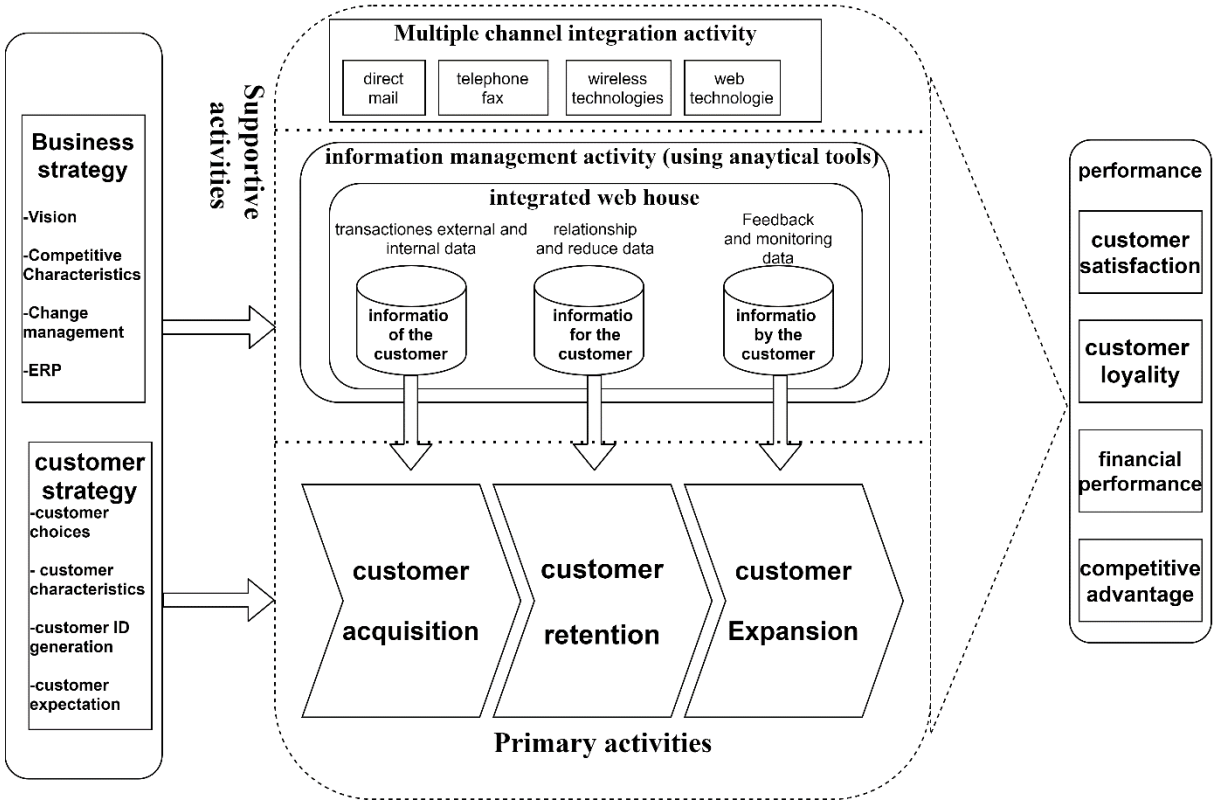
2.5.3.1 Customer database: Customer's information would be saved and maintained in an integrated and consolidated Web House (a web-based Data Warehouse) that could be accessed from all parts of a firm. Before entering the web house, a transformation process would be done on information to save the integration of W.H. and to prevent the entrance of redundant information.

2.5.3.2 Analytical tools: To use web house efficiently, several types of analytical applications like data mining, Web Mining, and OLAP would be implemented on these W.H.'s. So, meaningful patterns and relations could be extracted from customer's interactions and behaviors.

2.5.3.3 Customer acquisition: From a strategic perspective, all the primary activities for attracting customers, like advertising strategies and marketing efforts, are for new customer acquisition and creating an image of the firm brand name. If these activities are run through the Web, firms should do different tasks to recognize customer interests. Site visitor tracking and clickstream databases are some of the activities that help recognize customer interests and offer better products or services to them. Companies try to convert Web site visitors and browsers into (hopefully repeat) buyers during customer acquisition. Visualization tools can play a crucial role in product evaluation and selection at this stage. It is very straightforward to say that companies must acquire customers to do business. Even firms with high retention rates lose customers and must continuously acquire new customers to stay in business. Because customer acquisition is a costly affair and the relationship developed during the acquisition phase strongly influences retention and add-on selling, it must be carefully managed and evaluated.

2.5.3.4 Customer retention: Even by ignoring the fact that the cost of new customer acquisition is S times of retention of an old one, yet retention and expansion are more critical due to lack of sufficient information about new customers. The goal of customer retention is to leverage customer acquisition investments. Because customer acquisition usually is more expensive than retention, it makes sense to find ways and mechanisms to extend the relationship between firm and customer. Of course, the focus must be set mainly on the most profitable customers.

Figure 2.5: Framework of e-CRM



Source: KARIMI (F): «*A Conceptual Framework for Electronic Customer Relationship Management (e-CRM):" A Strategic Approach*, » 2020, p.15.

2.5.3.5 Customer expansion: CRM Businesses invest their money on those customers, which will bring the most value for the firm. Customer expansion with a focus on retention of the most value-adding customers' needs precise strategies to do. In the expansion phase, core customers (the most valuable ones for the firms) actively participate in the two-way interactions with the CRM and expand the CRM's customer base by word-of-mouth marketing. Feedback or suggestion from these core customers (by-the-customer information) may prove crucial for the CRM to introduce new products, improve business processes, and satisfy customer needs.

2.5.3.6 Performance measurement and evaluation activity: To evaluate the performance, some criteria would be determined through the e-CRM system design stage. These criteria depend on different goals of serving customers and should be aligned with overall operational business criteria. Here the focus is on two elements of performance: customer satisfaction and operational performance. In traditional models, customers' satisfaction is examined through surveys and complaint systems. Today, web technologies enable e-CRM to get feedback from customers immediately after the interaction, and in minimum time, even feedback could be a

part of the interaction process. It is hoped that customer satisfaction led to customer loyalty which is another criterion for measurement of performance.[74]

2.6 The benefits of e-CRM

The well-planned implementation of an ECRM system aims to create a long-term relationship with customers with minimum cost, Reduce the customer defection rate, Increase the profitability from low-profit customers, and focuses on high-value customers.[75], it produces a winning situation for customers and companies alike. Improvements in the overall customer experience led to greater customer satisfaction, which positively affects the company's profitability.

- **E-CRM increases customer loyalty** because information stored in this software helps a company to look actual cost of winning and retaining the customers. By using this information, a company can use its time and resources for the most profitable customers. In this way, a company can find the best customers by managing them as a premium group. This shows that it is not advisable to treat all customers in the same way. [76]

- **E-CRM gives more effective marketing** because this information is used to predict what kind of product a customer likes to buy and the timing. It allows to make the campaign targeted and to track it more effectively. This customer data is used to analyze it more effectively, like which marketing campaign is the best and effective and its impact on sales and profitability.[77]

- **E-CRM improves customer service** and support because it helps to receive, update and fulfill orders remotely, and this finest tool is used to complete this service in the best practical way.

- **E-CRM is an efficient and cost reduction tool** that integrates all customer's data into a single database. It permits marketing teams, sales forces, and all departments to exchange

---

[74] FJERMESTAD (J), ROMANO (N): «Electronic Commerce Customer Relationship Management,» International Journal of Electronic Commerce, 2006, p.83-85.
[75] Kennedy (A): *« Electronic Customer Relationship Management (e-CRM): Opportunities and challenges in a Digital World,»* Irish Marketing Review, No.1&2,2006, p.60.
[76] KOTLER (P) et al.: *« Principles of Marketing,»* 4th European Edition, Pearson Education Limited, England, 2005, p.6.
[77] GREENBERG (P): *« CRM at the Speed of Light: Capturing and Keeping Customers in the Internet Real Time,»* 2001, p.17.

information and achieve the corporation's common objectives by using the available statistics.[78]

- **To Enhance Customer Interactions and Relationships** By gathering online customer information. It is already in an arranged format and can be pulled to analyze without data entry, compared to traditional channels. This data streamlining technique enhances information quality with less time. Organizations can pick more information via online channels, which leads better analytical decisions to overview customer behavior; thus, targeted and customized relationships are established. By using CRM, both parties, customers, and sellers receive benefits because customers receive those products or services more closely related to their desires while organizations receive high-value and low-risk customers. Designing a correctly and implemented CRM can remove many administrative demands in the organization and give better information to the customers at a low cost. Cost-effective marketing is achieved by well-defined segmentation, which increases profit. In association with other technologies used at the back end like customer databases, data mining and warehousing, value-adding, and personalized products and services create a competitive advantage.[79]

- **Managing Customer Touch Points** When customers deal with an organization, they move between traditional and online channels. E-CRM supports these multi-channel touchpoints with the company, and there should be consistency in customer experience. When multiple interaction points are offered to customers, they will not be ready to repeat the processes if it is not integrated with other departments. Customers should have real-time information via all the channels when they switch to alternative channels. The customer can interact with the company through the sales department, reseller, or over the Internet. This has created multi-channel management for successful CRM strategy in organizations. The most significant advantage of E-CRM is to link all the operations in a business which affects the customer experience. Technology allows companies to have customers' feedback from touchpoints or social media. The technology makes it possible for agents, managers, partners, and other

[78] LLOYD (G) (2001): *« Electronic Customer Relationship Management: Benefits and Pre-Implementation Considerations,»* America's conference on information systems, 2001, p.842-843.
[79] AB HAMID (N.R.): *« E-CRM: are we there yet? »*, Journal of American Academy of Business, 2005, p.60.

users to instantly maintain a single view of the customer and gain organizational information. This ability of a single view of customers has improved customer service.[80]

- **Personalization and E-Loyalty** With E-CRM, it has become possible for companies to tailor the overall customer experience at an individual level. This tailoring is based on customer data, dynamic personalization, which include information of contents presented and the products offered and supported by advertising from other organizations. The direct-to-customer channel is a crucial enabler to handle automated systems that make it possible to offer highly relevant content because of the volume of data collected (e.g., links clicked in emails, products viewed but did not purchase online, etc.). E-CRM improves Internet customer's loyalty level and enhances customer E- loyalty.[81]

- **Source of Competitive Advantage** When well-designed and correct E-CRM is implemented, it increases the digital loyalty cycle, which becomes a lasting competitive advantage. When a firm uses E-CRM technology and redesigns its business processes to acquire customers and retain them, it strengthens customers' purchase decisions, including pricing, quality of the product, marketing, sales, and customer service. It creates a more digital loyalty cycle.[82]

[80] CROSBY (L), JOHNSON (S): *«CRM and management*,» Marketing Management, Jan.–Feb 2002, p.10.

[81] AB HAMID (N.R.): *« E-CRM: are we there yet? »*, Journal of American Academy of Business, 2005, p.65.

[82] KENNEDY (A), KELLEHER (C) & QUIGLEY (M): *«CRM best practice*,» getting it right the first time, ESBI, Proceedings of the Irish Academy of Management (IAM) Conference, GMIT, Sept 2005, 7–9.

II. Section two: analytical customer relationship management

1. Understanding analytical customer relationship management

Analytical customer relationship management is designed to deeply analyze the customer's information and data and disclose essential intension of the customers' behavior on which the organization can do capitalization. A need to find new sources of revenue is one of the fundamental requirements for business growth in every given aspect of the business. This perspective enhances reaching the current customers and prospects to encourage them to do more business with the organization. Analytically, from the customer's perspective, there are confident choices in products, services, delivery, and packaging that the effort of an organization to sell more products may result in an onslaught of the information, which may lead to a piece of confusing information. As far as the business perspective is concerned&-, where the information is exchanged between the business and the prospective and current customers, the customer support centers, direct mail, telesales, direct sales, e-commerce, and the business website pages accessed by the customers, hence might be challenging factors to figure out which methods deliver the best results. [83]

The acquisition of customers is at the fore of any organization in designing CRM campaigns. Through the current years, customer acquisition techniques have evolved spurred by technological advancements and customer heterogeneity. They are basing on data improvement collection capabilities. Storage data capabilities, the ability to analyze the collected data, and customer profiles have paved the way to increase intricate and focused acquisition techniques. In attaining competitive acquisition techniques, there is a need for the organization to implement mass-level strategies to foster market organizational development.[84]

The collaborative CRM integrates customer interaction and communication channels. For example, through the phone, email, website, and faxes. The main agenda of this part is to refer the customers systematically and consistently. Primarily it maximizes profitability, revenue, and customer satisfaction. Certain technologies support the business organizations, for example, gaining customer attention, storage and analysis of customer satisfaction, and internal process information. The critical and essential part of the type of CRM above is software

---

[83] TEKLITZ, (F), McCarthy (R.L.): *«Customer Relationship Management: The Ultimate Guide to the Efficient use of CRM*,» 1st Edition, Wiesbaden, Germany,2001, p.64.

[84] KUMAR (V), PETERSEN (A) & PETERSEN (J.A.): *«Statistical Methods in Customer Relationships Management*,» 1 S.T. Edition, Hoboken New Jersey, 2012

installation with customer-oriented strategies. Basing on technology, it is one of the powerful tools to support the CRM initiatives, and with no alliance, the tool of less benefit to the organization's business or the company. The type of CRM particularly supports the processes performed at the office level that affect the customer's activities and enhance the customer relationship. The CRM type addresses maintenance, planning, marketing, advertising, finance, billing, invoicing, and information technology. Enhances various departments in the business to share helpful information, improving the quality of customer service delivered to the clients. Implementation can be done through essential software or in more complex installation, depending on the requirements of a company. The CRM type shows how the current business situation undergoes a vast cultural shift, transforming from mass marketing to the product-siloed thinking of the industrial age to the customer-based culture of the information age, analyzing 15 primary goals to build relationships with individual customers become more valuable to the enterprise. The current market situation has enhanced the organizations to being customer-centric without being controversial where the employees are perceived to matter compared to the customers. The enterprise might face a hard time delivering the services to the customers desired consumers of the customers products. [85]

The competitive market situation in the current business development opportunities forces the companies to develop new marketing techniques that can enable them to survive through the competitive marketing storms. The highlighted above factors force the business organizations to develop innovative marketing activities to capture customer needs and improve customer satisfaction and retention for long-term benefits for both parties, hence improving marketing decision support. Organization managers are faced with a rapidly changing competitive marketing environment., hence being forced to become more competitive through better and critical decision making. Hence the managers need to implement the collaborative CRM in their take in the business organizations managed. Angela Hausman elaborates better in the journal writing of Doing business in a Digital World.[86]

[85] PEPPERS (D), RODREGS (M): *« Managing Customer Relationships; A Strategic Framework*,» 2nd Edition. U.S., 2010

[86] HAUSMAN (A): *« Doing Business in A Digital World. Journal of Business & Industrial Marketing»,* Emerald Group Publishing Limited. Bingley, United Kingdom, 2005

2. Integration of big data analytics and customer relationship management

Advances in analytic techniques, information technology, technology cost reduction, and attempts to capture and retain customers boost the growing use of Big Data. Large organizations have used that technology to raise awareness that may lead to a competitive edge. Big Data represents a new way to manage and know customers. It is a decision-making process based on data and new types of data-enriched products and services that have been given new perceptions from customers, products, and markets. [87]

The relationship between relationship marketing and Big Data may allow companies to relate with clients more effectively, creating opportunities for automation of relationship processes and managing helpful business information. In this context, CRM and other technologies may be used in several business segments and cause more interactivity among people with possibilities of information sharing to create and define customer relation strategies. [88]

The integration of BDA and CRM is what we will describe here, the backend support needed for analytical CRM. Specifically, we first outline a generic architecture and then focus on the two key components: data warehousing and data analytics.

2.1 Data Warehouse

Building a good data warehouse is a crucial stepping stone in getting started with analytical CRM. Data sources for the warehouse are often the operational systems (websites, mobile app) for the external data, sales growth, customer profiles, etc. For the internal data, they are providing the highest level of data. Data sources are designed for operational use and decision support, and the data reflect this fact.

Through direct communication and touch with customers via communication channels, customers generate a massive amount of data, these data are often from different sources and different types, and because of that, the data company store are not always structured. It is often unstructured (no-SQL) unusable like this. It stores in a no-SQL database, Here the data engineers team converts that data and transforms it into usable structured data (rows and

[87] PRICEWATERHOUSE (PWC): *« Capitalizing on the promise of Big Data*,» 2013, p.2.
[88] DAVENPORT (T. H), DYCHE (J): «*Big Data in Big Companies*.» SAS Institute Inc, July 13, 2016, p.19-20.

columns) and puts it in the CRM data warehouse where it is available to the data scientists to use it.

It is essential to clean warehouse data since critical CRM decisions will be based on it. The three classes of data extraction tools commonly used are - data migration which allows simple data transformation, data scrubbing, which uses domain-specific knowledge to scrub data. Moreover, data auditing will discover rules and relationships by scanning data and detects outliers. Loading the warehouse includes some other processing tasks, such as checking integrity constraints, sorting, summarizing, building indexes, etc.

2.2 Data analytics

The next generation of analytic CRM requires companies to span the analytical spectrum and focus more on looking forward. The 'what has happened, and the 'why has it happened.

Combined with data explosion, time-to-market pressures force many organizations to struggle to stay competitive in 'less time. With the need to be more proactive, organizations focus their analytical efforts to determine what will happen, what they can do to make it happen, and ultimately automate the entire process. Big data analytics is now viewed today as an analytical necessity. The primary focus of data analytics is to discover knowledge previously unknown, predict future events, and automate the analysis of enormous data sets.

Once the data has been cleaned up in the data warehousing step (the data integration), various data analytics techniques (Machine learning, data mining, etc.) can be applied to extract models from it. Several data analytics techniques have been developed, and the one to be applied depends on the specific purpose at hand. Once a model has been developed, it can be used for two kinds of purposes:

- **The first** is to understand the present behavior of the customers and extract value and knowledge from it. A type used for this purpose is called *descriptive analytics*. by the application of its techniques (descriptive statistics, etc.), a company can extracts insights from it and understand the customers by answering the question of "what has happened? and why?" and segment them into homogeneous clusters and try to understand the nature of their growth in the last periods to have a clear vision about its market.

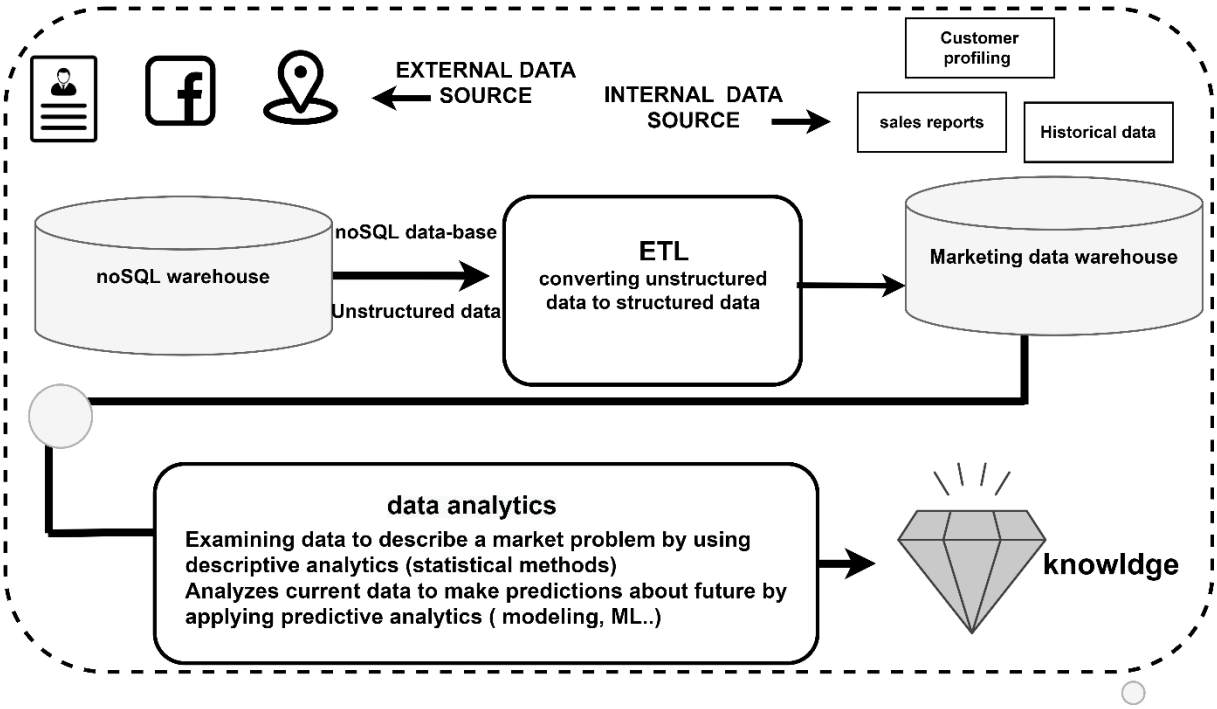
- **The second** is to use the type to make predictions behavior of the customers. A model used for this purpose is called *predictive analytics.*

By applying predictive analytics techniques (data mining, machine learning, deep learning, etc.), the company can predict customer lifetime value, behavior, profitability, etc., to develop a tight strategy that achieves its primary and secondary goals.

The descriptive model, extracted from past behavior, is a starting point for a predictive model. Such an approach is quite successful, as it is based on the assumption that past behavior is a good predictor of future behavior ñ with appropriate adjustments. [89]

Based on the knowledge provided by the team of data scientists, the customer relationship management team applies perspective analytics by studying the results and comparing them with their knowledge of the market, results, and previous experiences to develop the best plan to meet the upcoming and current challenges and take the right decisions.

Schema 2.1: The integration of BDA and CRM



Source: made by us

[89] SRIVASTAVA (J) et al.: *« A case study for analytical customer relationship management »*, PAKDD, Berlin, 2002, p.22-24.

3. The Application of Big data analytics in Customer Relationship Management

3.1 Acquisition of New Customers

The acquisition of new customers is the first step in the life cycle of customer relationship management, is the indicator of business development, and is the key to growing businesses. [90]. Data mining can help companies study the respondents who visit the corporate website, call toll-free, or fill out the application form. Furthermore, companies can make their marketing activities focus on these people. [91]. It can help to complete the division of the potential customer base to achieve the effect of these activities, which are the cost of accessing new customers, effective management, and the improvement of the return.

3.2 Customer Segmentation

A large customer base is divided into several smaller groups based on specific attributes. Customers in the same group are highly similar in one or a few property values, but there is a big difference between them in the different groups. For example, using technologies in data analytics (data mining), such as decision trees, cluster analysis methods, companies can divide into different user groups by the customer's psychology, habits, or the frequency of purchase, in order to achieve the targeted customer services and the development of targeted products and to improve customer satisfaction and lifetime value of the customers.

3.3 Customer Retention and Loss Warning

Customer retention is a process, which maintains established customer relationships to make the customers repeat to purchase products or services. Studies have shown that the cost of a new customer is five times the cost of keeping an old customer and improving customer retention by 5%, the profit will be increased by 25% [92]. Therefore, customer retention is the key to business success. Enterprises can analyze the customer transaction history in database, demographic information, and other data associated, and do some targeted research on the loss of customers, then create a loss of client model, and develop appropriate plans and programs, to implement different strategies for different customers using neural network technology.

---

[90] SUN (Ning): *« The issues of customer relationship management based on data mining*.» Technol Ind 11(6), 2011, p.59–61.
[91] XIA (W), WANG (X) & WANG (Q-S): *« The data mining in the enterprise customer relationship management*.» Technol Manage 6, 2007, p.196–198.
[92] ZHANG (Z): *« The study of SME customer relationship management based on data mining*.» Econ Forum 492(7), 2011, p.202–204.

3.4 Cross-Selling

Cross-selling is the marketing of new products or services sold to an existing client. It is not only an effective means to increase profits by expanding sales to existing customers but also an important strategy to enhance the corporate image, foster customer loyalty, and protect sustainable enterprise development. [93]. Companies can find the characteristics of the buyers, find the critical factors of the buying behavior, and establish a forecasting model to predict and analyze the customer's future buying behavior by using correlation analysis sequence analysis and so on "2/8 rule" is the classical theory of customer relationship management, that is 80% of the profits from the 20% of customers. Therefore, companies need to know which customers are profitable to assign more resources to these customers. Companies can establish the Life Time Value model to predict the profitability of these customers within a predetermined length of time by using the classification analysis techniques and so on.

3.5 Customer Profitability

Analysis "2/8 rule" is the classical theory of customer relationship management, that is 80% of the profits from the 20% of customers. Therefore, companies need to know which customers are profitable to assign more resources to these customers. Companies can establish the Life Time Value model to predict the profitability of these customers within a predetermined length of time by using the classification analysis techniques and so on.

3.6 Analysis of Customer Credit

The customer's credit rating is significant in customer relationship management. Because, when the credit risk and fraud happened, enterprises will face loss of market share, which causes loss of market, customer, competitiveness, and reputation. Through data analysis, companies can use classification analysis, cluster analysis, sequential pattern analysis, and other techniques to predict the causes and the possibility of customer fraud. Establishing a fraud rule base can be a timely warning of various fraud and reduce the loss of business.[94]

---

[93] ZHONG (R-W), WANG (H-P): *«The customer relationship management system based on data mining*.» Shaoguan University, 31(3), 2010, p.20-21.

[94] ZHAO (H-Y), ZHAO (S-H): *«Applications of Data Mining in the Enterprise Customer Relationship Management.»* Scientific paper, International Conference on Information Technology and Management, Beijing, 2013, p.57-59.

4. The benefits of analytical CRM

Business realizes that their most valuable assets are relationships with customers and all stakeholders. Building personal and social relationships becomes a vital marketing area—the importance of relationships as market-based assets that contribute to customers' value. With the amount of data increase, some business organizations use advanced, powerful computers with colossal storage to process big data analytics and increase their performance resulting in tremendous cost savings. [95][6]. Businesses manage structured and unstructured data sources such as social marketing, retail databases, recorded customer activity, logistics, and enterprise data to establish a quality level of CRM strategies by recognizing essential data and its advantage. Some CRM strategies can happen through big data and big data analytics. Since big data can provide a pattern of customers' information, businesses can predict and assume the needs of their customers nowadays.

Extensive data had helped shaped many industries and changed the way businesses operated nowadays. Big companies benefited from this shift, especially companies such as technology giants such as Amazon and Google, and would continue to serve these giants from the sheer volume of data they generated. Data Velocity showed how marketers could access real-time data, such as real-time analytics of internet sites and social media interactions. CRM with the significant data influence, a new paradigm had been created to allow accessibility and availability of information, resulting in greater take-up by big or small businesses alike. Big data offers pervasive knowledge acquisition in CRM activities. Big data will support long-term relationships through understanding customers' life cycle and behavior in a more comprehensive perspective. Customers voluntarily generate a considerable amount of data daily by detailing their interest and preference about products or services to the public through various channels.

CRM with prominent data features becomes agile regarding strategy like push notification through the smartphone to the potential target audiences. Web / Apps users who make a comment, liking page, or comes back visiting Web or Apps are potential customers are targeted for pushed notification. CRM aims to quantify sales transactions, promotion, product awareness while building long-term relationships and loyalty. Therefore, big data analytics can develop a comprehensive view of customers to enhance service fitting with customer attention, engagement, participation, and personalization. The study introduces several fundamental

[95] WATSON (J)*: «Big Data Analytics: Concepts, Technologies, and Applications*.» Communications of the Association for Information Systems, 2014, p.1261.

marketing concepts with big data closely related to customer-based CRM strategies by engaging the customer life cycle.

CRM with big data brings a promise of significant transformation that can affect the organization in delivering CRM strategies. Businesses cannot simplify marketing strategies, only focusing on a short-term relationship with customers without any path in attracting, retaining, and extending for a long-term relationship. In addition, the organization can also create better customer personas by using the profile data as the backbone of creating accurate personifications for the customers. Also, the organization will have data on the customers' needs and preferences and used this data to provide better content for the audience where the content is relevant and valuable to them. All these data can also provide valuable information for the management team to improve marketing budget management by ensuring the operational business process stayed on budget with the help of data and to be more focused and targeted.[96]

There were many benefits for using big data in CRM, such as accurate and update in profiling of target customers, predicting trend on customer reaction toward marketing messages and product offerings, personalize the message that creates emotional attachment and product and the following were just some of the benefits:

4.1 Customer profiling

Whenever a business acquires a new customer through marketing activities, the customer will determine each activity's value. When the customer perceives the value is positive, they will be happy and satisfied. Otherwise, they can consider finding another business even from other competitors that may fulfill their requirements. Therefore, customer profiling for every customer becomes essential for a business to make sure that the whole CRM's life cycle (sales, marketing, and customer service) offers personalized and customized services to experience differently according to their needs and interest.

Big data can help in delivering customer profiling since it includes business activities monitoring. Big data analytics can track purchase histories and their online conversations about their product or services. Businesses will gain a more comprehensive view of customers' expectations and better understand potential customers' interests. The impact of big data analytics is significant, especially on the distribution of marketing channels between service providers and customers, whereas suppliers or service providers are engaging with customers

[96] ANSARI (M): *«Customer Relationship Management and Big Data Enabled: Personalization & Customization of Services*.» Applied Computing and Informatics, 2018, p.6-8.

directly, threatening the sustainability of intermediaries marketing agencies. Customer profiling can gain invaluable insight from big data analytic and create a competitive advantage. All of these organizations derive business value from leveraging personalization. Some of the examples for customer's profiling of services:

Amazon.com developed a system of product recommendation based on their analysis of customers' previous purchases data [97].Target, the supermarket, can develop a predictive model in tracking the purchase made by the pregnant women [98]. UPS, the package delivery company, created an application to redesign their drivers' daily routes to achieve fleet optimization. Customers' profiling is possible through big data analytics because the organizations have access to more accurate data as big data can discover the value of the hidden data connections and patterns. In addition, it can improve business decisions because it provides as much knowledge as possible. [99]

CRM team generates customers' knowledge profiling to enhance businesses and understand precisely target audience, personalize message for each potential customer, and tailoring message fits with customers' interest and preferences. CRM with big data analytic can develop a comprehensive knowledge of customers for decision making. Roll Royce applies big data analytics in the aircraft engine-manufacturing sector and uses the result to predict when and where the breakdown of aircraft engines might occur by installing sensors to collect data. [100].

Hence, they sell engines and packages of both engines and monitoring services that generate profit by charging customers based on usage, repairs, and replacements. This service currently accounted for more than 70% of their annual revenue in their aircraft engine division by leveraging big data to create a competitive advantage.[101]

---

[97] DALEN (C) & DALHLBOM (F): *«Big Data in the Telecom Industry: A Study of How Big Data Affects Innovativeness and Market Dynamics.»* Master Thesis in Marketing Spring, Stockholm School of Economics, Sweden, (2014), p.7.
[98] KIM, YUN (B): *«Big Data: A Temporary Thing? »*, Sungkyunkwan University, Republic of Korea, 2015
[99] FAN (S), LAU(R-Y) & ZHAO (J.L.): *«Demystifying big data analytics for business intelligence through the lens of marketing mix*.» Big Data Research, 2(1), 2015, 28-32.
[100] WANG (H), et al.: *«big data and industrial internet of things for the maritime industry in north western Norway*,» In TENCON 2015-2015 IEEE Region 10 Conference, 2015, p.1-5.
[101] Rolls Royce Shifts In Higher Gear With Big Data (datafloq.com) (13/06/2021 at 23/17)

4.2 Personalization

Personalization is essential for the organizations to focus on the user experience to boost digital marketing efforts by making sure the targeted audience seeing advertisements, social media posts, contests, or creating events that have emotional relation to the audience. These efforts will help to map the customer journey easier to tract since it is on digital marketing. These trends will be able to provide accurate and real-time mapping of customers' choices and also locations. Hence the organization can use this data to create a more personalized leading nurturing process by making the loyal customers feel important, and their relationship with the company is meaningful. In comparison, customization enables the business to sort out specific CRM activities to the targeted audiences more precisely. Business strategies should transform customer strategies and systems to customer engagement and the ability for customization. The one is more focused on the conversation that is going on between organization and customer and the collaborative models that cutting-edge companies are carrying out for customer engagement. Customization strategies will improve customer services, and excellent customer support will increase loyalty, revenue, brand recognition, and business opportunity.

Big data in CRM should target the right market segment for the right product or services to improve market sharing, avoid access production, control budget, efficiency, and effectiveness in the business process. Furthermore, the long-term strategy focuses on producing marketing materials for target customers rather than reaching random customers. Customization is essential in marketing because marketing should be flexible in offering their product or service. Flexibility in offering products shows empowerment for a customer. Businesses can improve their sense of understanding of their customer's behaviors. The marketing team can customize the campaigning agenda that is expected to fit with potential customers' patterns. Customers' profiling is also possible because big data with geolocation analytic promotes quick and appropriate marketing strategies. It can quicker than a competitor in marketing products, appropriate in determining the price by understanding purchasing power of potential buyers from data of financial abilities. Appropriate in displaying product means the supplier has the best possible place to display the product easily reached by customers, the data come from geolocation. [102]

[102]ANSARI (M): «*Customer Relationship Management and Big Data Enabled: Personalization & Customization of Services*.» Applied Computing and Informatics, 2018, p.8-9.

Netflix detects traffic details of customers' views to spot problems in the area and add systems that can help the future demand. They are also able to get more vision of their customer's desire. [103]

4.3 Value creation

Big data brings new opportunities for discovering values since it shows customers' behavior or anything related to society, leading to more precise analysis. CRM strategy ventures potential customers, meeting expectation, understanding their needs, and delivering value. While, generating value can come from big data analytic. As described, big data grow an extraordinary rate huge amount of data sets containing a variety of data types and come up with patterns, correlations variables, latest market trends, customers' choices and preferences, and other valuable business information. [104]

These big data analytics provides effective marketing, new revenue streams, customer relation services, improved value chain efficiency, sustainable competitive advantages over rival organizations, and many other benefits. [105]CRM initiative needs to extract value from various data sources gathered from customers to manage or analyze customers' preferences. Big data generates value to understand people's needs and preferences by having access to more accurate information as big data can discover hidden connections and patterns in customers' behavior through multi-sources and multi-channels. Customers produce data voluntarily through various means, including clickstream activities while visiting the website for e-commerce transactions. [106]

More customers generate and share data in the public sphere means an organization can extract more value. Customers produce data, and business organizations capture those data for their marketing plan. In addition, customers make conversation on social media about products that are helping any audiences, either customers or producers. Customers benefit from the conversations and reviews because they understand product knowledge better. At the same time, the producer understands better about products' expectations. Big data has involved customers in delivering practical CRM activities where marketing teams fine-tune the ideas into

---

[103] Big Data basic concepts and benefits explained - TechRepublic (13/06/2021 at 23:27)
[104] BOLLIER (D) & FIRESTONE (C.M.): *« The promise and peril of big data*.» Washington, DC: Aspen Institute, Communications, and Society Program, 2010, p.1.
[105] ANSARI (M): *« Online Learning: Trends, Issues, and Challenges in Big Data Era*.» Journal of E-Learning and Knowledge Society. 2016, p.124.
[106] W.U. (J H), WANG (S C): *« What drives mobile commerce? An empirical evaluation of the revised technology acceptance model*». Information & management, 2005, p.726.

an executable marketing program. Social media analytics is the analysis of structured and unstructured data from social media channels.

Value is added when the customer believes the information and can then make the appropriate decisions in their marketing strategy. Decision-making is vital in marketing strategies. Decision-making can improve because extensive data analysis provides comprehensive information from multi-channel interactions. For instance, a marketing manager needs to understand customers' behaviors, expectations, and trends to deliver a favorable and impression marketing campaign to potential customers. Decisions on the appropriate marketing contents in specific area can be derived from big data reports and then analyzing these data to discover trends and key issues. Marketers who can identify problems and propose solutions are able to gain a greater understanding of customers' desires. Therefore, a marketing campaign needs to extract the value from various data sources of Weblog files, social media, enterprise content, transaction, and the team must ensure its authenticity and reliable, valid information derived from big data analytic. In short, CRM should support big data analytics. In order to mitigate the gap between expectation and delivery. However, for big data to add value in marketing, marketers need to access and understand the comprehensive customer path from awareness to conversion to close the gap between online and offline customer behavior and experience: all data sources, either structured or unstructured, form customer profiling. Customer profiling can come from geolocation data analytic of customers through a smartphone or any smart mobile devices, face recognitions, browsing behavior and online activities, e-commerce histories, social networks, and so on. [107]. Then, they are analyzed to develop the pattern, trend, or behavior of the customer. For instance, marketing staff who deal with the customer directly are supplied a comprehensive and user-friendly report of potential customers in the speed of data analytics (velocity). The report shows pattern, behavior, and likelihood of customer so that the marketing staff can aggressively and convincingly the value of product or service that fit with his or her need. Fitting customers' needs can come from the personalization of marketing or treating every customer based on their interest and need. Engaging customers in providing feedback will make a product or service superior to a competitor. Velocity is the feature of big data analytics. Therefore, a quick and appropriate response will make improve customer satisfaction which in turn to customer loyalty. Big data offers simplicity for CRM staff since they are equipped with customer patterns in advance to know how to handle every

[107] W.U. (J H), WANG (S C) TAI (W.C.): *« Mobile shopping site selection: the consumers' viewpoint*.» In System Sciences. Proceedings of the 37th Annual Hawaii International Conference, 2004, p.8.

customer. CRM department can offer extra facilities to the potential customer based on their interest and behavior.

Big data facilitates multi-channel interaction. For instance, automotive manufacture releases a new sedan's car and manufactures concerns about the customer's and competitors' responses. The manufacture can monitor all social networks and media to evaluate customers' reactions to improve the product. In the end, big data provides cost-efficient for overall CRM activities because they can pinpoint of effective CRM strategy for the customer.

Conclusion

In this chapter, we have highlighted two points. The first one is electronic customer relationship management. The rapid advancement of technology in all sectors has also spurred the progressive development of the CRM concept and created new communication channels that bring Electronic Customer Relationship Management (e-CRM) to the surface; we have presented its definition and components that can help and improve the customer relationship management.

The second point we have discussed is the analytical customer relationship management and the integration of big data analytics; we have presented big data analytics techniques for building customer relationship management and the benefits that companies can get from it.

We have concluded that e-CRM displays product catalogs and promotion media yet also used for dialogs, discussions, and consultations with customers online. Hence, the customers can be invited to be directly involved in the design, development, marketing, and selling of a product. Moreover, by integrating big data analytics and CRM, companies can extract value and insights from the amount of data generated by the Customers and understand customer behavior to provide a customized, personalized product and service.

Chapter three

case study

Introduction

This last chapter constitutes the culmination of research work. It translates our understanding of the theoretical notions developed in previous chapters. This case study aims to verify the extent to which the big data analytics applied in all the customer relationship management levels impact the value creation and the changes that touched customer relationship management with this many revolutions in big data the world is living. Furthermore, this work was done to take the example of an experienced company using analytics techniques in CRM.

This chapter will start by presenting and discussing the global ride-hailing market evolution, and the Algerian one too, then we are going to present the host organization, and we will try to concentrate more on their CRM strategies and how they use the analytics techniques on support decision-making, and this is the content of the first section.

The second section will start by presenting the research methodology, the tools, and the techniques we have used. Then, we will present the results and interpret them, and finally, recommend some strategies and actions.

Keywords: CRM, decision-making, analytics techniques, big data analytics, ride-hailing, value creation.

I. Section one: Presentation of the enterprise 'YASSIR SPA'

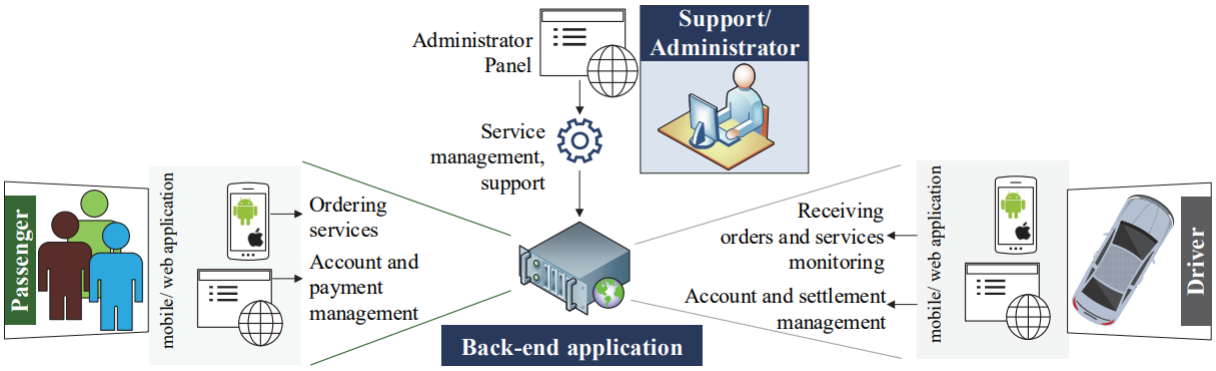
This section is consisted of 3 main points, at first, we are going to start by highlighting the "Ride-hailing" industry in the world and Algeria by discussing the market growth evolution. The second point will present the host company "Yassir S.P.A.," its activities, services, and the organizational chart, and then we will discuss Yassir's CRM system and strategies.

1. An overview on Ride-Hailing service

1.1 Ride-hailing definition

"Ride-hailing is a transportation service that allows passengers to request a ride in real-time via smartphone applications that link passengers to nearby drivers with a transportation network company (T.N.C.) such as Uber or Lyft." [108]These services were first brought to market in 2009 through Uber and in 2012 through Lyft, which provided an app connecting drivers and riders. Most apps have various improvements on traditional taxicabs, such as electronic dispatch, a way for drivers and passengers to rate each other, and dynamic pricing functions to encourage drivers to meet the demand for rides during peak periods. Many terms are used to describe this emerging transportation mode, including ride-sourcing, ride-hailing, e-hailing, and app-based on-demand service.

figure 3.1: Ride-hailing's application mechanism



Source: HUNAITI (Ziad): The Ride-Hailing Mobile Application for Personalized Travelling. Modern Applied Science, October 2018, p.244-245.

The ride-hailing companies aid the driver-partners and the riders to be connected through a smartphone application. The mechanism is illustrated in **Figure 1.3**. The Rider can

[108] RAYLE (L) et al.: « Just a better taxi? A survey-based comparison of taxis, transit, and ride-sourcing services in San Francisco ». Transp policy 45, San Francisco, 2016, p.169-171.

use the application using the passenger's account to request a ride. In response, when a neighboring driver-partner accepts the request, the applications let the Rider know a projection of arrival time to the passenger's location. The application also updates the Rider when the driver-partner is almost near to the location. In some applications, once a driver accepts the request, the passenger can evaluate the waiting time and decide whether to accept the ride. In addition, both the passenger and the driver can use Global Position System (G.P.S.) navigation during the trip, and the application guides the driver to shortcuts and less-congested roadways.

Unlike the traditional taxi, as the ride-hailing cars do not possess any distinctive visual appearance, the applications also provide few valuable information to the Rider, such as the type of the vehicle, license plate number, approved information regarding the driver's identity to facilitate the Rider to identify the driver-partner at pickup location. In addition, the Rider has the freedom to express a preference about the route course to the destination. The destination can be specified through the application before or during the journey. Finally, the journey is terminated at the destination point, followed by an automatic calculation of the fare. [109]

1.2 Ride-hailing market evolution

The Ride-hailing service marketplace has been developing in the past decade and overgrowing, with new and innovative services and choices. As a result, there are exciting opportunities and unforeseen challenges for cities as they ensure mobility, equity, access, and reliability across their systems. By far, the most prominent new entry in this marketplace comes under the heading of "ride-hailing services," also known as "transportation network companies" (T.N.C.s), with Uber and Lyft as the most extensive examples in the U.S. Yassir in Algeria.[110]

In recent years, passenger preferences for utilizing ride-hailing services across the developing countries, such as India, North African, Vietnam, etc., have increased, along with increasing demand in the developed regions, such as the United States, China, and Europe. As a result, the companies enhance the options and expand their operations to be provided in mobile applications, to retain their respective market shares in a highly competitive market.

The major drivers of this market include the growing need for personal mobility in the wake of rising urbanization and a fall in car ownership. Also, growing Internet and smartphone penetration is leading to the high growth of the ride-hailing market. On the other hand, the

[109] IDEM
[110] NACTO: Ride-hailing services*: «opportunities & Challenges for cities,»* New York, June 2016, p.4-5.

factors which restrain the market are resistance from traditional transport services and complex transport policies of different countries.

The global ride-hailing market has declined from $60.5 billion in 2019 and to $42.25 billion in 2020 at a compound annual growth rate (CAGR) of -13.88%. The decline is mainly due to the COVID-19 outbreak, extending lockdown across countries and the measures to contain it [111]. The market is expected to grow to $56.87 billion in 2021 at a compound annual growth rate (CAGR) of 34.6%. The growth is mainly due to the companies resuming their operations and adapting to the new normal while recovering from the COVID-19 impact, which led to restrictive containment measures involving social distancing, remote working, and commercial activities' closure operational challenges. The market is expected to reach $108.15 billion in 2025 at a CAGR of 17%. [112]

The ride-hailing market consists of revenue generated by ride-hailing used for passenger transportation service in which the passengers are connected with the vehicle drivers through a smartphone-based app. The industry comprises establishments primarily engaged in providing passenger transportation by automobile, not operated over regular routes and on regular schedules.

1.3 Ride-hailing in Algeria

Algeria can be considered as an ideal market for the Ride-hailing industry because of many aspects; we tried to present some of them in the three following points:

- The first one is the weakness of the transport offer; In the big cities of the country, the transport offer is weak in terms of quality and quantity. It is not uncommon to wait several tens of minutes or even hours to see the bus or train in Algeria. Moreover, there are no fixed schedules for public transport. In Algiers, only the eastern and southern suburbs of the capital are served by train. At the same time, only one metro line is functional, as is the tramway.
- It is not easy to get around the city on a coach. Even getting around by car is a hassle. However, traffic jams on one side and parking difficulty on the other discourage the most reckless.
- The taxi-Cabs in Algeria does not have the best image. They do not systematically accept to take the client to his/her destination if it is outside the most frequented roads. If

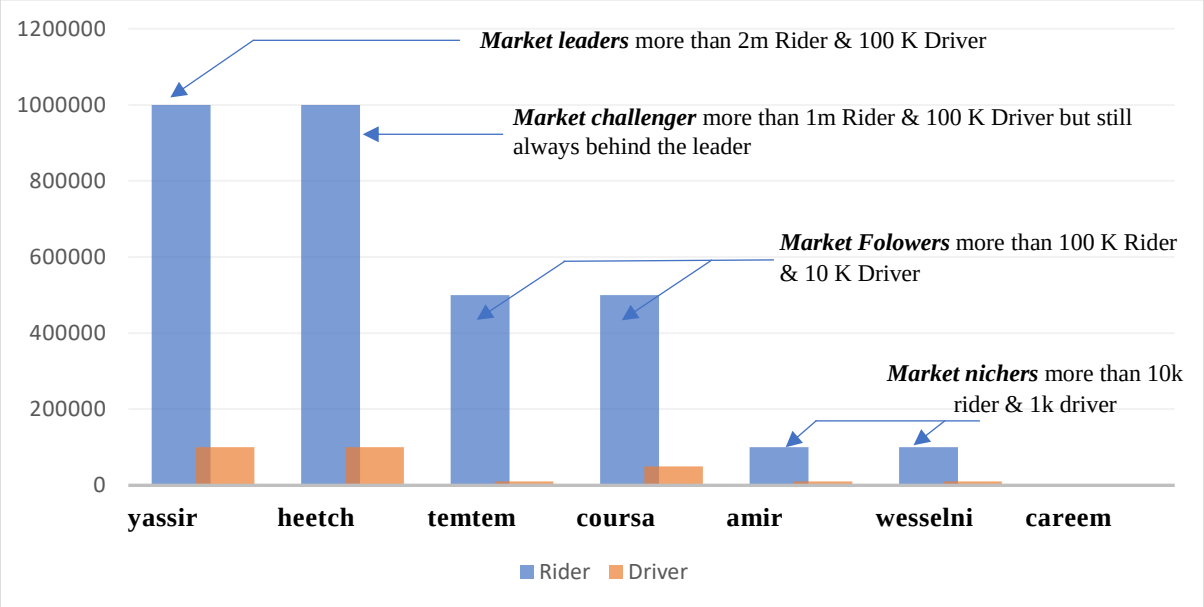
[111] https://www.researchandmarkets.com/reports/5314994/ (30/05/2021 at 23:31)
[112] https://www.researchandmarkets.com/reports/5027885/ (30/05/2021 at 23:31)

by chance they accept, it will be for a price much higher than the one indicated by their meter.

On the other hand, cabs are not hostile to applications, unlike many other countries (Morocco, France ...). On the contrary, many of them work with ride-hailing companies to avoid empty rides and optimize their income. These aspects and adding to them the globalization and the growth of internet users have created the demand.

Since the launch of Yassir and then Temtem at the end of 2017, the Ride-hailing market in Algeria has experienced a solid dynamic. Indeed, nearly a dozen players have launched in the meantime, and competition between them is raging; some of them are an Algerian product (Course & Amir), which have a piece of good knowledge about the Algerian market and the Algerian customer, So, they can supply the best service. The most significant ride-hailing market leaders recently acquire others (Careem by Uber or Heetch). It will be difficult for the Algerian R-H applications to resist them because they already have experience, which can help them avoid making inevitable mistakes and proliferate. Advanced features and almost no bugs, their costs are relatively low, and they can centralize functions like the technical functions.

Figure 3.2: the number of installations of Ride-hailing's app in Algeria



Source: made by us (Adopted from paly-store and app-store, considering Careem there is no data available)

2. Yassir's Customer relationship management system

2.1 The host organization

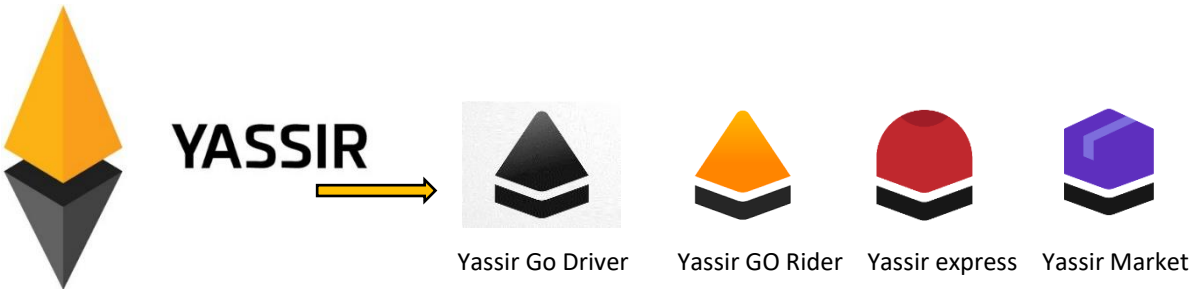
2.1.1 Yassir S.P.A.

Yassir S.P.A., better known by its brand name Y.A. Technologies, is a start-up company with headquarters in Palo Alto, Silicon Valley, U.S.A., and the second one in Algiers, Algeria. This one comes to change the daily life of service providers present in Algeria and Africa through innovative digital/mobile solutions and create a bridge between Silicon Valley and the Algerian ecosystem. It was founded in 2017 by two Ph.D. students El Mehdi Yettou and Noureddine, both alumni of the Polytechnic School of Algiers, with extensive research experience at the Canadian University of Sherbrooke and the American University Stanford, respectively. A model of snowballing start-ups and have an economic and social impact in the Algerian environment.

Figure3.3: Yassir's slogan

.

Figure 3.4: The evolution of Yassir's logo 2017-2021

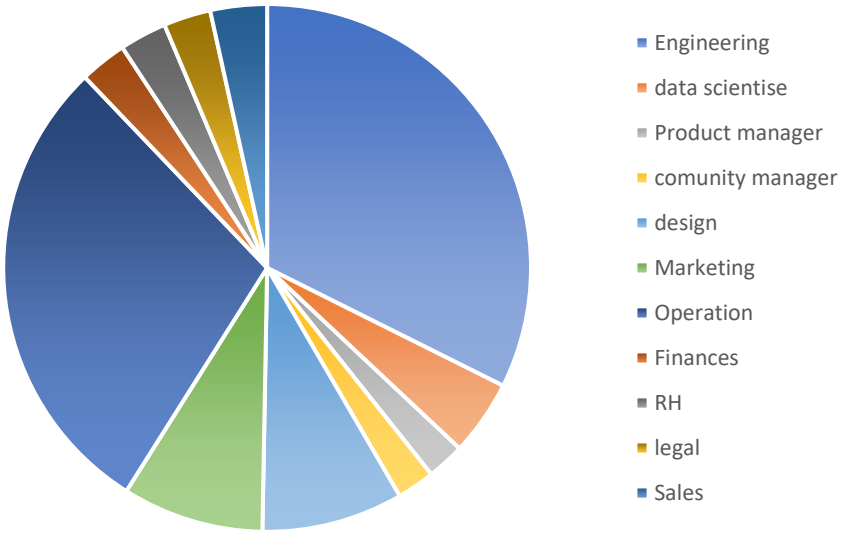


Source: https://yassir.com/ (04/06/2021 at 23:30)

2.1.2 YASSIR Team:

The company has more than 180 employees in Algeria, 30 in another four countries with an average age of 28 years, 70% of the management team left a large multinational company to join Yassir, the team is composed as follows:

Figure 3.5: The breakdown of Yassir's employees by title



Source: internal document

2.1.3 The services of Yassir

2.1.3.1 Yassir G.O.:

The two founders chose the transportation sector to launch the first and the main product of their company "Yassir Go," which is a ride-hailing service that brings a new vision of transportation, offers customers a private driver in one click, and allows customers to move from one point to another with ease and comfort. This service is provided through a mobile application, a web platform for customers, and a mobile application. Using this service is very simple; log in using the phone number, then select the pickup location, the destination, and the type of car the client wants from a range of available car types (Classic, Comfort, Premium, etc.) and order your driver in one click. Yassir Go was first launched in Algiers in September 2017, and since that day, the product has continued to evolve with new features and services. Yassir Go service is now available in 14 wilayas in Algeria and another 11 cities in 4 countries.

The app has recorded more than 10,000,000 downloads since it was created, more than 15,000,000 Km of trips, more than 2 million Rider and 40,000 drivers on the national territory.

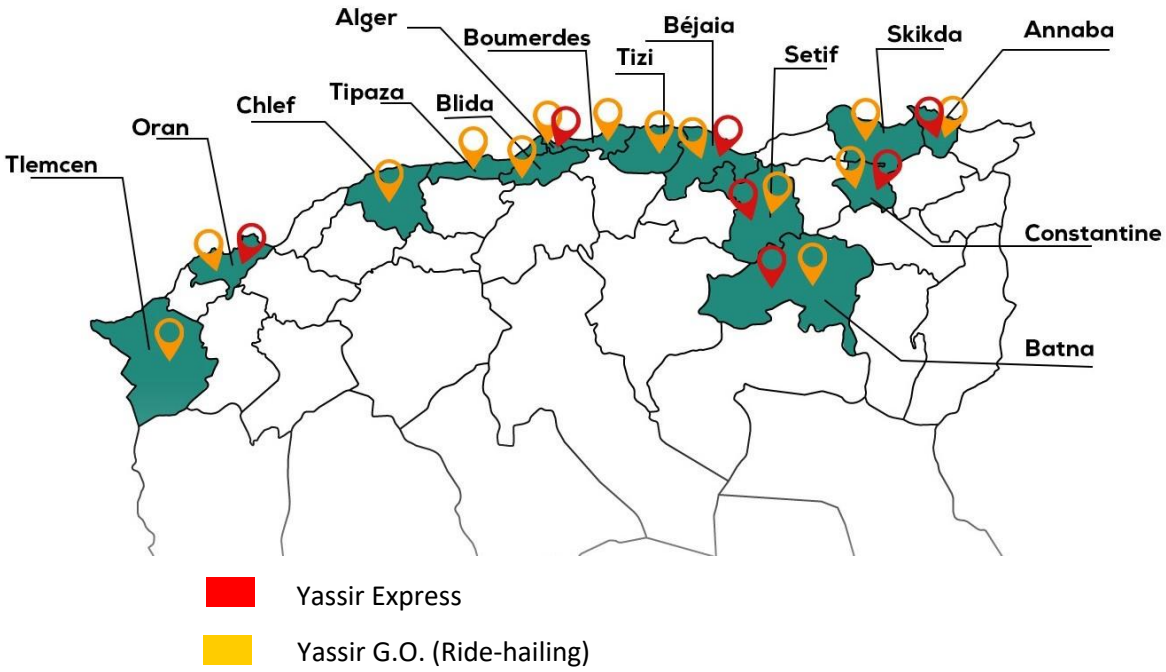
In addition, Yassir Go has established itself in the two neighboring countries, Tunisia and Morocco, in September 2019, France, Canada in 2020. [113]

2.1.3.2 YASSIR Express:

If the adventure began with the service Yassir Go, the creators did not stop there. On November 16, 2019, YASSIR launched the YASSIR Food app, which delivers meals from different restaurants. On April 13, 2020, Yassir Food evolved and became Yassir Express to adapt to the market needs, especially in the containment period due to the coronavirus pandemic. This service is a delivery service via a mobile application that delivers daily products where the customer wants them seven days a week. In addition, an array of stores with different types of products is available on the application. [114]

- **Available store categories:** Restaurant, butcher, convenience store, cosmetics, parapharmacy, wholesale, bakery, fruits and vegetables, and pet shop.
- **Delivery categories:** On foot, By bicycle, By motorcycle, By car

Figure 3.6: geographical distribution of Yassir apps



Source: made by us (adopted from an internal document)

[113] https://yassir.com/client/ (02/05/2021 at 14:18)
[114] https://express.yassir.io/ (05/06/2021 at 02:46)

2.1.3.3 Yassir Business:

YASSIR Business is an application intended for all companies providing them with a solution for their employees' travel at lower costs and with less effort. This solution intended for professionals aims to help all companies optimize their organizational and financial resources without any budgetary commitment. [115]Yassir Business provides many advantages for companies:

- **Dematerialize transportation:** No need to have a fleet of cars to manage employee travel.
- **Better cash management:** A way to eliminate cash for employee travel with the Business offer.
- **Analysis and monitoring:** Ensure better monitoring and analysis of costs related to employee transportation.
- **Security and total control:** Controlled transportation management. All trips are subject to company approval.

2.1.3.4 Yassir Market:

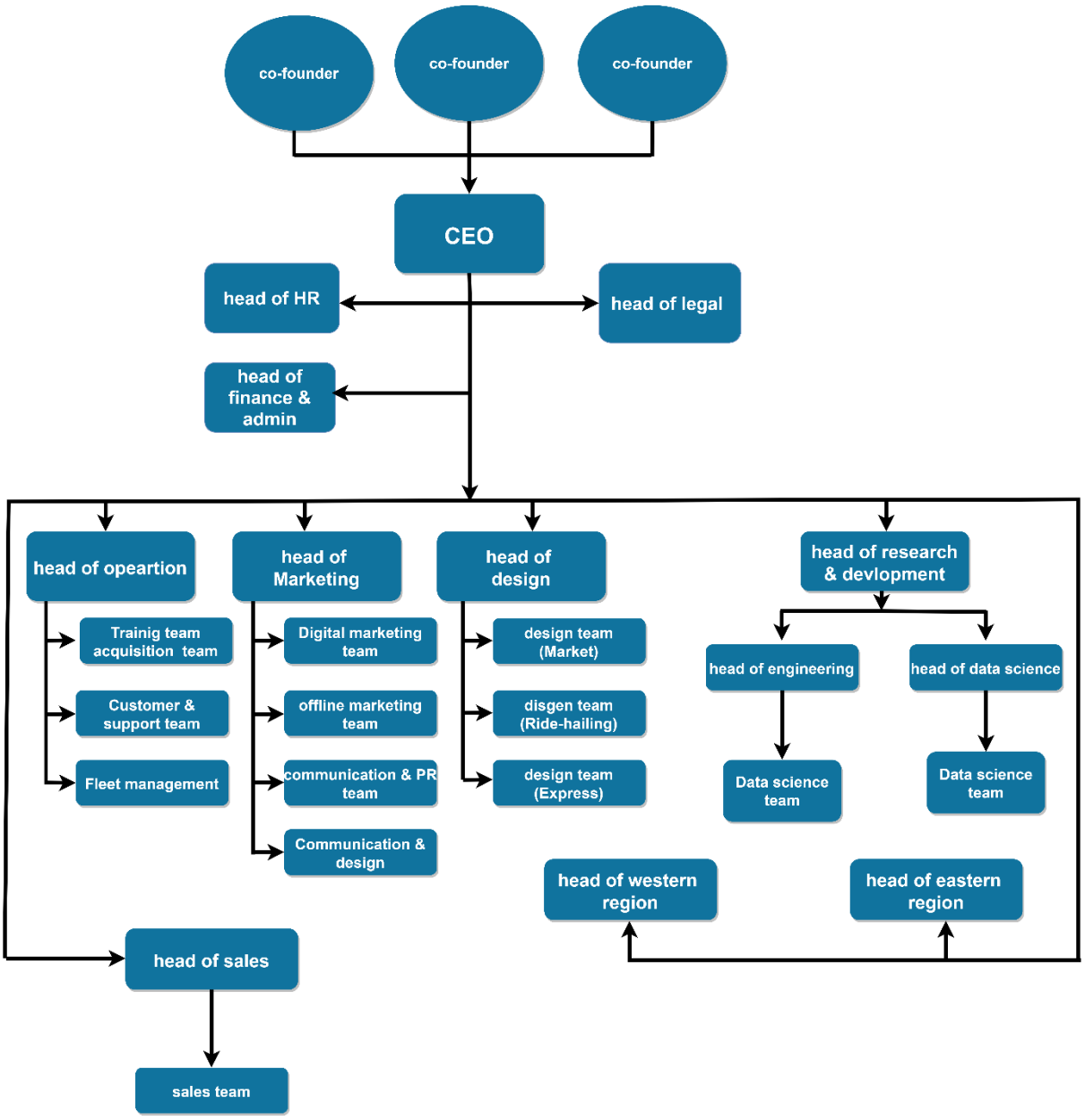
Recently, Yassir has launched a new e-market under Yassir Market, a new "e-commerce" platform where e-suppliers and e-buyers are brought together to contract sales, purchases, and services online. The customer can pay on delivery or by card Dhahabia, C.I.B., visa, and Mastercard and provide the country's delivery.[116]

2.1.3.5 Teleconsultation:

Facing the Covid-19 outbreak and in collaboration with a community of volunteer doctors. Yassir puts at its customer's disposal specialists who will answer any questions or doubts online. In addition, to relieve overcrowding in hospitals and physicians' offices that represent a favorable environment for the spread of the virus, organize meetings between the doctors who wish to volunteer to answer online the needs, doubts, or questions of fellow citizens and the patients (zoom, google-meeting). [117]

[115] https://yassir.com/business/ (05/06/2021 at 02:46)
[116] https://market.yassir.com/ (05/06/2021 at 03:10)
[117] https://yassir.com/en/telemedicine/ (05/06/2021 at 03:10)

2.1.4 organizational chart



Source : Internal document (19/05/2021)

2.2 Yassir customer relationship management system

2.2.1 The integration of B.D.A. and CRM in Yassir

Through direct communication and touch with customers via an electronic communication channel, customers generate a vast amount of data. These data are collected and stored (with customer permission) through Yassir's data sources: "the mobile application and the website,." the company stores the customer's data such as the number of visits per day, times of visits, number of trips...etc., these data are stored in a NoSQL data warehouse (oracle, mongo dB), where it is in the form of unstructured data (no SQL) unusable like this. So here, the data engineers team converts that data and transforms it into usable structured data, and puts it in a cloud platform (amazon or google) where it is available to the data science team to use it.

2.2.1.1 Collaborative CRM's

After the data is placed in the cloud platform, the team of data scientists tries to extract value and knowledge from it through descriptive analytics (descriptive statistics, simulation, optimization, etc.). To extract value, ideas, and insights from it and try to understand customers' problems and answer the question of "what has happened? and why?" and segment them into homogeneous closets and try to understand the nature of their growth in the last periods.

2.2.1.2 Analytical CRM's

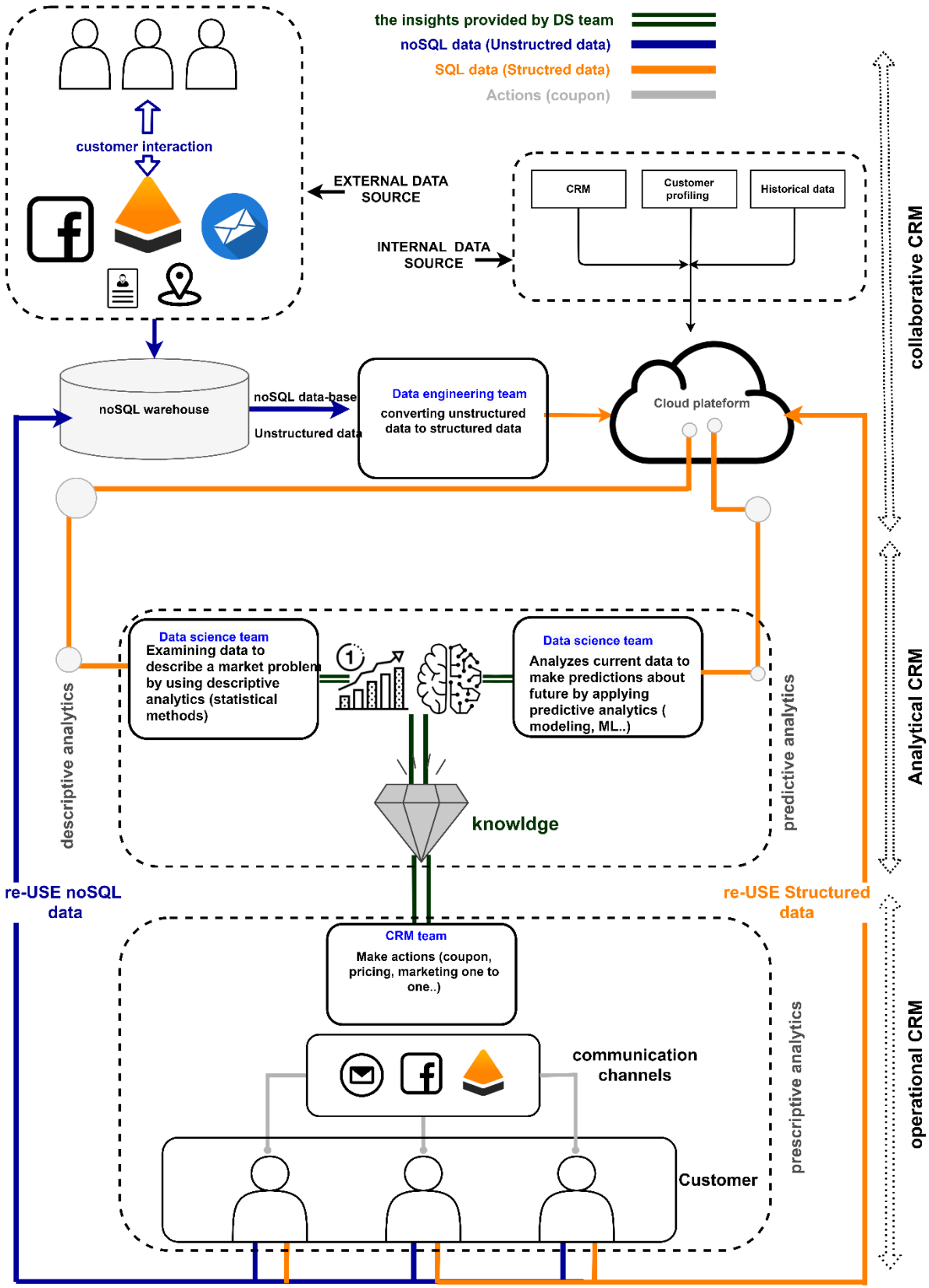
By applying big data analytics' second type, the data science team try to predict customer lifetime Value, its behavior, profitability, etc. by the application of predictive analytics techniques (data mining, machine learning, deep learning, etc.) to develop a tight strategy that achieves the company's primary and secondary goals.

2.2.1.3 Operational CRM's

Relying on the knowledge provided by the team of data scientists, the customer relationship management team applies perspective analytics by studying the results and comparing them with their knowledge of the market, results, and previous experiences to develop the best plan to meet the upcoming and current challenges and make the right actions.

The actions (coupon) in Yassir are generally launched through electronic communication channels (app, website). Yassir's CRM team does not apply the action to all the customer segments, but to a certain percentage of them (Applying only about 90% of the segment), to compare the two groups to evaluate the strategy's efficacy. The data resulting from this process, such as the customer's response and growth rate, etc., are re-collected and stored in the NoSQL warehouse (the first one) to be used again.

Schema 3.1: the integration of B.D.A. & CRM in Yassir



Source: made by us

2.2.2 The CRM strategies of Yassir

The marketing department in Yassir plays an important role; it directly impacts its performance and is considered the 3rd important department in the company structure after the I.T. & the strategic one.

Customer relationship management in Yassir is a crucial factor because it is in direct touch with the customers. Therefore, t aims to build a strong relationship with them, derived by the data science team (data-driven), which they supply the knowledge and the information through market studies based on analytics technics that help decision-makers make the right decision (Data-driven).

Yassir follows a plan that inducts four possible actions. Every one of them is oriented & applied in different customer segmentation and aims to different objectives.

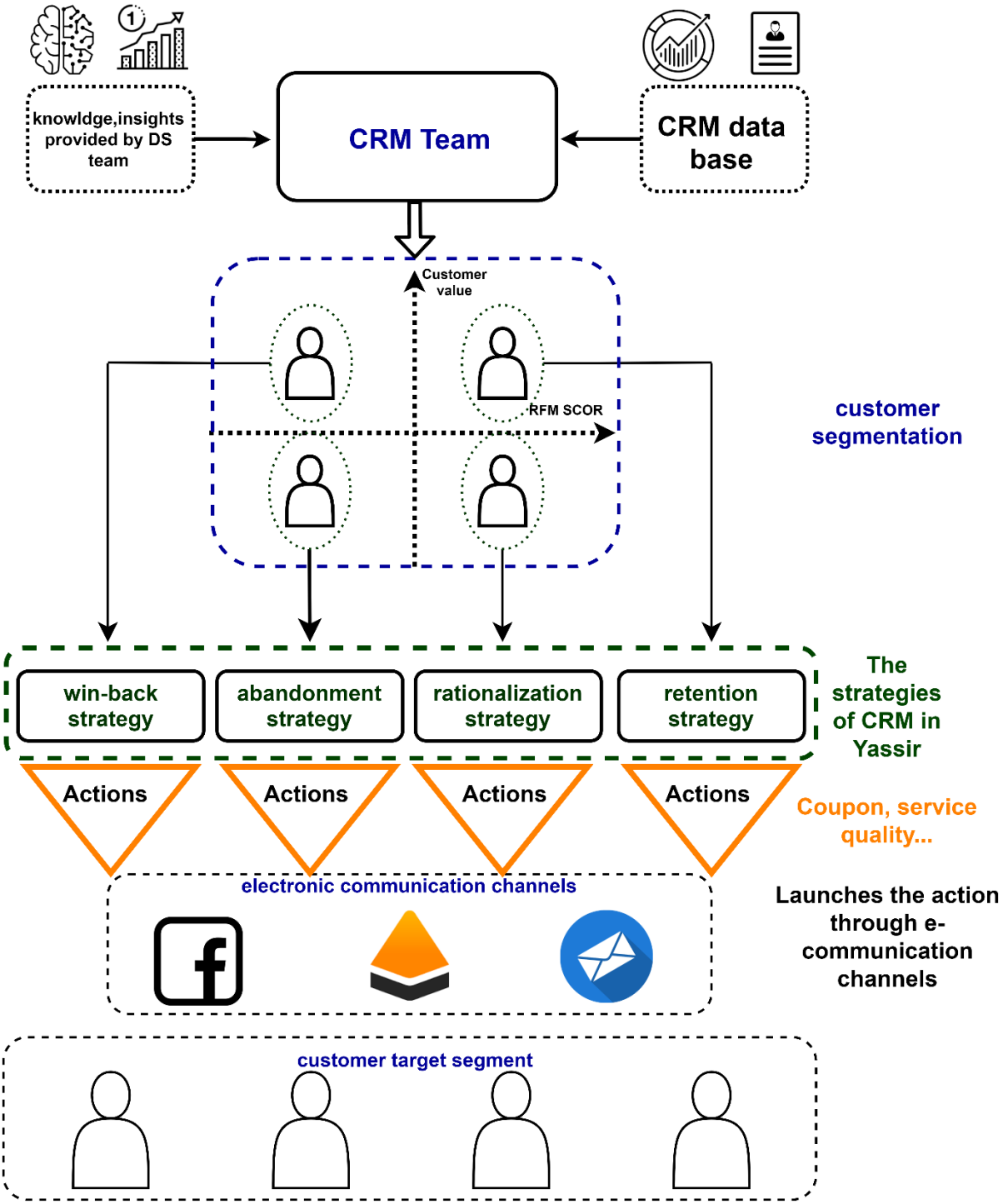
- The first one is ***"win-back strategy,"*** which is an action that aims to regain an existing customer who has churned (In Yassir, we define the churned customer by a rider who did not have to make any finished trip for three months or more). This strategy seeks an answer to the question of "Why our customer churned?" because if Yassir can see the reasons why they can provide him the best possible offer if he is a target yet or abounds him if he is not a potential anymore (for example if a rider bought a car or he will not be a potential anymore).

- The second one is "***the customer retention strategy***," it has known that retain an existing customer is cheaper than winning a new one, So, the primary purpose is to keep the existing customer by rewarding customers who purchase regularly. For example, offering a birthday gift, when a rider requests a trip on his birthday, Yassir offers him a free trip or a coupon that's what will create a solid personal relationship between Yassir & their customers and make them feel that they are unique and the company are carried on their them.

- The goal is not just to retain the customer but also to push them to be more active, and that is what ***"The rationalization strategy"*** will be about, so this strategy seeks to improve the profitability of regular but not very lucrative customers. This rationalization involves reducing the costs associated with these customers.

- The last one is about abounding the customers with a negative profit that's it ***"the abandonment strategy***," So Yassir's revenue from them is less than the costs they spend to keep them ***(R.T. – CT < 0)***, abounding these customers is a strategical step because keeping them would only bring more costs.

Schema 3.2: the strategies of CRM in Yassir



Source: made by us

II. Section two: Customer segmentation using K-means clustering

For the practical writing part, scientific methods to confirm or invalidate our hypotheses are also required. To answer the main problem of the research, we chose to develop a quantitative study and employ a methodological approach adapted to the nature of this research.

In order to achieve these objectives, we will present the research methodology used in achieving our investigation, the results obtained, and some recommendations.

1. Research methodology

In this point, we will describe and present the objective, methods, techniques, and tools that we have used in this case study to collect the information necessary to verify the research hypotheses.

1.1 Research objective

Our study focuses on the contribution of big data analytics and its role in decision-making in customer relationship management. It aims to show how companies can use big data analytics techniques to improve the customer experience by providing the best personalized possible offer and improving the customer relationship management's strategies.

To find an answer to this problem, we developed the following sub-questions:

- What levels does big Data Analytics affect customer relationship management?
- How are the use of big data analytics affects value creation in customer relationship management?
- What are the implications of applying big data analytics techniques on customer relationship Management's strategy & operations, and how it relates to the supply chain performance?

The methodology used in the study is a simulation of the CRM analytics in Yassir and tries to contribute to building a customer relationship strategy. Given the typology of the questions posed, we have resorted to two significant data analytics types, ***"Descriptive analytics"*** & ***"Perspective analytics."***

Firstly, through the descriptive analytics, we have made an R.F.M. analysis on a sample of a Yassir customer database by using a simple machine learning algorithm called K-means, intending to divide them into homogeneous segments, which would help us to understand the behavior of each one of them.

Secondly, through perspective analytics, we have analyzed every segment, but every one of them into the appropriate strategy, assigned specific actions to each one and made some recommendations regarding the best offers for each segment.

Then, we selected one of the segments and made a FRAT analysis on it to segment it into new sous-segments to understand its purchase behavior nature. The results of this segmentation analysis have helped us to build a customer relationship management strategy that is based on upselling and cross-selling to each one of them.

1.2 Research techniques & methods

Here we will present the methods and the techniques that we have used in our case study.

1.2.1 R.F.M. Segmentation

1.2.1.1 What is R.F.M. segmentation?

"R.F.M. analysis is a common data-driven customer behavior segmentation technique to understanding customer buying behavior. As the name implies, it involves calculating and examining three KPIs - recency, frequency, and amount - that summarizes the corresponding dimensions of the customer relationship with the organization. R.F.M. analysis can be used to identify good customers with the highest scores in the relevant KPIs, who generally tend to be good prospects for additional purchases.

It can also identify other purchase patterns and respective customer types that are of interest, such as infrequent large consumers or customers with small but frequent purchases who may also have sales prospects, depending on the market and the specific product being promoted".[118]

R.F.M. dimensions are generally defined as follows:

- **Recency:** This attribute was when customers purchased from the brand the number of days since the customer's last purchase. It has long been known that customers who have recently purchased are more likely to purchase again than customers who have not purchased in a long time. Time, in this case, can be days, months, years, or whatever is helpful in the industry. Indeed, the type of recency depends mainly on the type of items purchased and the industry. "Recency is often considered the most important. However,

[118] TSIPTSIS, (K) & CHORIANOPOULOS, (A): *« Data Mining Techniques in CRM: Inside Customer*,» edition Wiley, Chichester, 2009, p.335.

previous studies have shown that however, according to previous studies, R.F.M. values are generally firm-specific and are based on the based on the nature of the products".[119] *For example, if the customer's last purchase was 42 days ago, their Recency input is "42."*

- **Frequency:** Frequency is the number and rate of purchase transactions. For example, if one customer buys once every 12 months and another customer buys eight times during the same period, it can be said that even though the amount of revenue or profit is the same, the customer who buys more often is more likely to continue to buy and be a loyal customer than the one who buys only once for the same or even more valuable. As a general rule, higher loyalty and customer value are generally obtained from customers who buy more frequently than those who buy less frequently over similar periods. *For example, if a customer purchases three times in one month, their Frequency input is "3 in one month."*

- **Monetary:** This simply means that the value is expressed in revenue, profit, or derived calculations. According to another definition, the monetary value indicator is defined as the average value of transactions (average amount per purchase transaction). Since the total value tends to be correlated with transaction frequency, the rationale for this alternative definition is to capture a different and complementary aspect of purchasing behavior. The amount attributes are generally less predictive than the other two attributes in an R.F.M. model. *For example, If the customer spends a total amount of 400 dinars on the purchase in one month from our company, then their Monetary input is "400 in one month."*

1.2.1.2 The application of the R.F.M. model

- **R.F.M. to improve Customer Lifetime Value: "**How much a customer spends with the company over their lifetime is based on a number of factors."[120]. R.F.M. can help in many aspects: reducing churn, up-selling, and cross-selling to segments most likely to respond, increasing loyalty and referrals, selling higher-priced items, etc.

- **R.F.M. for new product launches:** Promoting new products to loyal customers is a great way to get initial traction and feedback. Contact champions and loyal customers before you

[119] JO-TING WEI et al. ii: *« A review of the application of R.F.M. model*,» African Journal of Business Management, Vol. 4(19), 2010, p. 4199-4206, p.4200.
[120] R.F.M. Analysis For Successful Customer Segmentation - Putler (19/06/2021 at 03:21)

even create a product. They can provide great information on what to build and how to promote it. This group of people will also be happy to refer the product to their circles of influence.

- **R.F.M. to increase user loyalty and engagement:** If there is a loyalty program, the potential loyalist is the first segment to target. It is essential to ensure that their initial experience with the product and service is pleasant and memorable. Follow up with some timely promotions, and they are very likely to buy again. Sending educational content to these customers will also increase their engagement with the brand.

- **R.F.M. to reduce customer churn:** At-risk customers, this is the segment to pay special attention to. It is sending personalized emails or calls to reconnect with these customers. It is even possible to offer repeat purchases at a discount or conduct surveys to address their concerns before losing them to competitors/alternatives.

- **R.F.M. for remarketing/retargeting campaigns:** "Remarketing is a smart technique that delivers ads/promotions to people who have visited the brand's site at least once, but are now on another site. They will see the ads on the other sites they visit - this improves click-through rates and overall effectiveness. A simple way to use remarketing with R.F.M. can be to export a segment of customers - especially recent or promising customers - to Facebook audiences or another campaign management solution. Then show promotions to that group of people."[121]

- **RFM to better understand the business:** Most small businesses do not fully understand their customers. They may not know the demographics of their customers. Collecting and understanding this information can also be time-consuming and costly. R.F.M. analysis becomes a quick method to understand customer behavior. Furthermore, because it is based on transaction history, its quality is much higher. Looking at different R.F.M. segments can reveal information about the business. Asking questions about how segments compare to each other can open up enormous opportunities for growth.

[121] IDEM

1.2.1.3 Limitations of the R.F.M. segmentation:

- The R.F.M. segmentation does not provide a complete understanding of customer behavior because it does not combine with other essential customer attributes such as product preferences or the type of purchase, so the company needs to have a complete view of customer purchase behavior and use all available information to guide its business decisions;
- Sending too many campaigns to a particular segment can annoy customers;
- The activity must be recurring. Without frequency, the model is meaningless, as in the automotive market;
- Neglecting low-value customers: many low-value customers can be just as valuable as few high-value customers

1.2.2 K-means clustering

1.2.2.1 what is K-means clustering

K-means clustering is an unsupervised machine learning algorithm, which is "used when you have unlabelled data (i.e., data without defined categories or groups). The goal of this algorithm is to find groups in the data, with the number of groups represented by the variable K. The algorithm works iteratively to assign each data point to one of K groups based on the features that are provided. Data points are clustered based on feature similarity."[122].

The main objective of K-means clustering is to group similar data points into "clusters."[123]. Here, 'K' means the number of clusters, which is predefined.

Here K defines the number of pre-defined clusters that need to be created in the process, as if K=2, there will be two clusters, and for K=3, there will be three clusters, and so on.

1.2.2.2 How K-means algorithm works

K-means allows us to cluster the data into different groups and a convenient way to discover the categories of groups in the unlabelled dataset on its own without the need for any training. It is a centroid-based algorithm, where each cluster is associated with a centroid. The

[122] Introduction to K-means Clustering (oracle.com) (19/06/2021 at 14:00)
[123] A cluster is a group of similar entities. It refers to a collection of data points aggregated together because of certain similarities

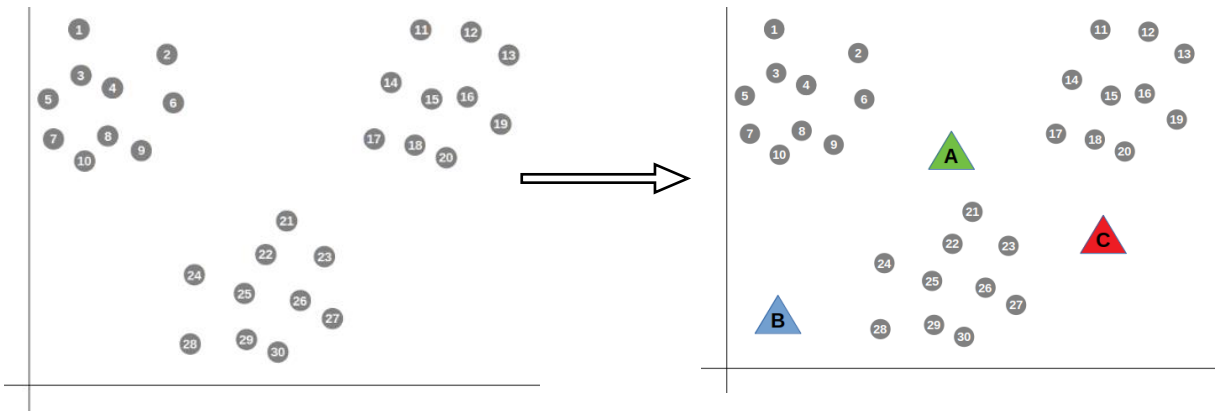
main aim of this algorithm is to minimize the sum of distances between the data point and their corresponding clusters.

"The algorithm takes the unlabelled dataset as input, divides the dataset into k-number of clusters, and repeats the process until it does not find the best clusters. The value of k should be predetermined in this algorithm. The algorithm starts s by placing k different averages (i.e. means) whose values are either initialized randomly or set to real data points on the plane. Let's start with k=3, as the data "seems" to be falling into three groups. For explanation-related purposes, lets randomly initialize the values (i.e., positions) of the averages, then, the algorithm goes through the data points one-by-one, measuring the distance between each point and the three centroids (A, B and C). The algorithm then groups the data point with the closest centroid (i.e., closest in distance)".[124]

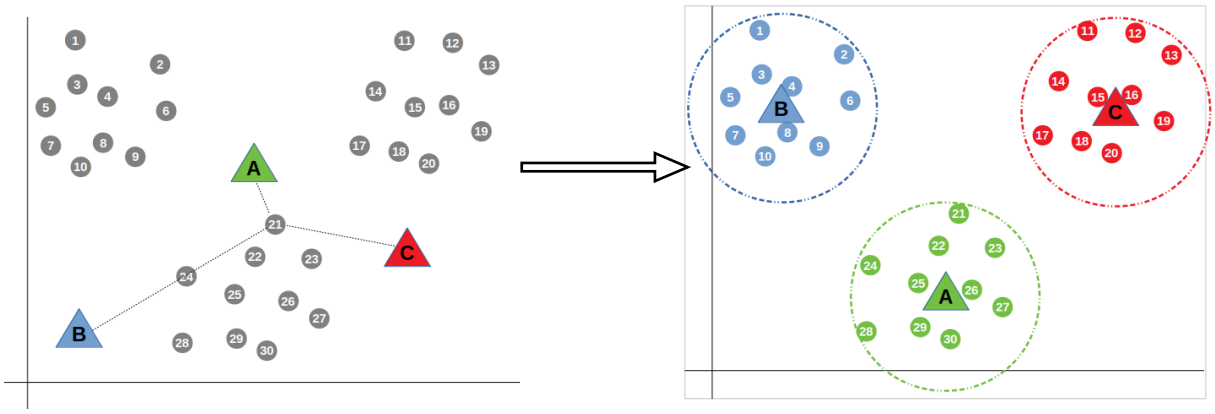
For instance, data point number 21 will belong to group A in the green color, merely because it is closer in the distance to centroid A.

It re-calculates the means (the values of the centroids). It is the sum of the points belonging to that centroid divided by the number of points in the group. It keeps doing the above until no centroid changes its value in re-calculation. Each centroid has centered itself in the middle of its cluster, surrounded by its circular decision boundary.

Figure3.7: K-means clustering works



[124] How Does k-Means Clustering in Machine Learning Work? | by Anas Al-Masri | Towards Data Science (19/06/2021 at 15:37)



Source: How Does k-Means Clustering in Machine Learning Work? | by Anas Al-Masri | Towards Data Science (19/06/2021 at 15 :37)

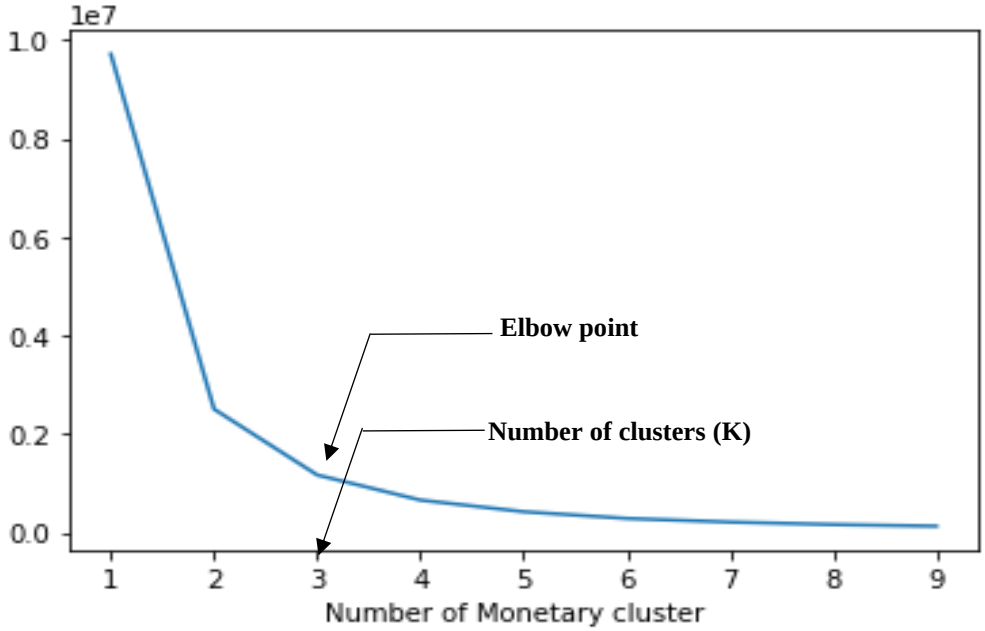
1.2.2.3 Choosing the number of clusters K

In general, there is no method for determining the exact value of K, but an accurate estimate can be obtained using the following techniques.

"One of the metrics that is commonly used to compare results across different values of K is the mean distance between data points and their cluster centroid. Since increasing the number of clusters will always reduce the distance to data points, increasing K will always decrease this metric, to the extreme of reaching zero when K is the same as the number of data points. Thus, this metric cannot be used as the sole target. Instead, mean distance to the centroid as a function of K is plotted and the "elbow point," where the rate of decrease sharply shifts, can be used to roughly determine K."[125]

[125] Introduction to K-means Clustering (oracle.com) (19/06/2021 at 15:37)

Figure3.8: Number of clusters (elbow point)



Source: made by us

1.3 Research tools

In order to achieve your research objectives and answer the hypothesis, we have used some tools, and in the following points, we will present them:

1.3.1 Python

In order to segment customers, we have used Python for data description, analysis, data visualization, and the application of the k-means algorithm on our data set.

"Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built-in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development and use as a scripting or glue language to connect existing components. Python's simple, easy-to-learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse.

The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms and can be freely distributed".[126]

[126] https://www.python.org/doc/essays/blurb/ (19/06/2021 at 16:42)

When it comes to data science, Python is a potent tool that is open source and flexible. It is known for having massive libraries for data manipulation and is extremely easy to learn and use by any data analyst or anyone. The packages and libraries we have used:

- ***Numpy:*** is the fundamental package for scientific computing with Python. This library provides multiple functions to create or save arrays and manipulate vectors, matrices and polynomials.
- ***Pandas*** is a Python library for data manipulation and analysis. It offers in particular data structures and operations for manipulating numerical tables and time series.
- ***Matplotlib*** is a Python library for plotting and visualizing data in graphical form. It can be combined with the NumPy and SciPy scientific computing libraries. Matplotlib is freely distributed.

1.3.2 Google Colaboratory

"Colaboratory or "*Colab*" for short is a product of Google Research. With Colab, any user can write and execute arbitrary Python code in a browser. Colab is particularly suitable for machine learning, data analysis and education. More technically, Colab is a hosted Jupyter notebook service that requires no setup and provides free access to compute resources including GPUs".[127]

1.3.3 Cloud- platform

We have used a cloud platform to select the data we need for our study, and the cloud platform is a relational data warehouse (SQL data warehouse) designed to help users manipulate, Store, query, and select the data intending to help organizations to turn their data sets to insights.

2. R.F.M. Customer segmentation using K-means

The R.F.M. segmentation process goes through 5 steps. We are going to present them in the following points, but firstly,

- Calculate Recency Monetary and Frequency
- Choosing the number of clusters (K)
- Apply k-means algorithm to calculate clusters
- Calculate R.F.M. score

[127] https://research.google.com/colaboratory/intl/de/faq.html (19/06/2021 at 16:52)

- Rider’s segmentation by their R.F.M. score

2.1 Data Exploration

Nevertheless, firstly, we must have a clear understanding of our customer's data to present and describe the columns and the rows in the following point.

2.1.1 General information about the dataset

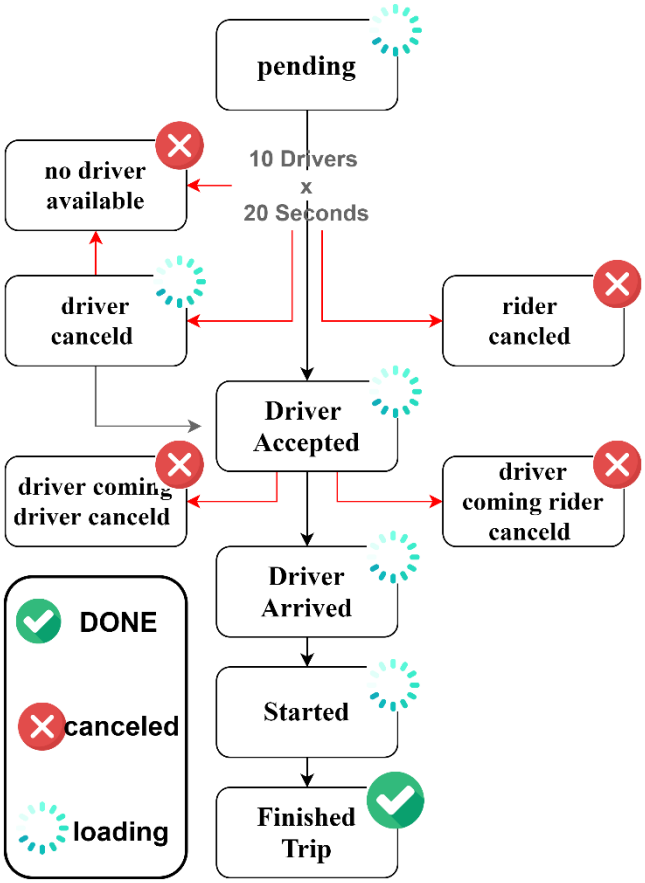
Our research concentrates on customer segmentation. To segment these customers, we need data; that is why Yassir and its data team gave us access to the internship database, which is consisted of 25 columns and more than 200,000 rows, so for our case study, we have chosen to select only the customers whom they made at least one finished trip (trip's statute) from Algiers (pickup wilaya) to anywhere in the last 11 months:

- **trip_id:** The Identification of the trip (series of letters and numbers)
- **request_date:** The date of the trip’s request
- **pickup_address:** The pickup address
- **pickup_daira:** The pickup daira
- **pickup_commune:** The pickup commune
- **dropoff_address:** The drop-off address
- **dropoff_wilaya:** The drop-of wilaya
- **dropoff_daira:** The drop-off daira
- **dropoff_commune:** The drop-off commune
- **accept_date:** The date when the Rider accepted the trip
- **start_date:** The date when the trip started
- **rider:** The identification of the Rider (series of letters and numbers)
- **driver:** The identification of the driver (series of letters and numbers)
- **estimated_distance:** The distance of the trip
- **estimated_eta:** The time the trip took to complete
- **estimated_cost:** The price of the trip with discount
- **original cost:** The price of the trip
- **coupon:** The coupon code used
- **discount:** Discount rate
- **status:** Trip’s status
- **purchase_type:** The type of purchase (premium, classic, comfort.)
- **Finish_date:** The date of finishing the trip

- **rider_rating:** Rider's satisfaction about the trip and the driver (from 1 to 5)
- **rider_comment:** Rider's satisfaction about the trip and the driver (notice or comment)
- **driver rating:** Driver's satisfaction about the trip and the driver (from 1 to 5)
- **driver_comment:** Driver's satisfaction about the trip and the Rider (notice or comment)

For to trip's status, there are ten statuses in Yassir we are going to present them:

Figure3.9: Yassir's trips status



Source: made by us

- **PENDING:** The initial state of a trip at the time of a driver's request.
- **NO-DRIVER-AVAILABLE:** No driver in the vicinity or drivers are not interested in this trip (in fact, the application sends the request to 10 available drivers for 20 seconds each, if no driver accepts, the trip will be recorded as N.D.A.).
- **DRIVER-CANCELED:** The request is refused by a driver, so it is redirected to another driver.
- **RIDER-CANCELED:** Cancellation of the trip by the customer before it is accepted.
- **ACCEPTED:** The trip is accepted by a driver.

- **DRIVER-COMING-CANCELED:** Cancellation of the trip by the driver after it has been accepted.
- **DRIVER-COMING-RIDER-CANCELED:** The customer cancels the trip after a driver has accepted it.
- **DRIVER-ARRIVED:** Driver has arrived at the pickup point to pick up the customer;
- **STARTED:** The trip is started.
- **FINISHED-TRIP:** The trip is finished, and the customer is dropped off at the drop-off point.

2.1.2 Dataset description

Descriptive statistics can give us great insight into the shape of each feature. Often, we can create more summaries than we have time to review. The describe function on the Pandas framework lists eight statistical properties of each attribute.

Figure3.10: dataset description

```
data.describe()
```

	estimated_distance	estimated_eta	estimated_cost	original_cost	discount	rider_rating	driver_rating
count	1.260300e+05	126030.00000	126030.000000	123678.000000	123678.000000	105880.000000	126020.000000
mean	1.511944e+04	1435.70457	623.729118	637.155484	0.014134	4.515518	4.633384
std	2.400503e+04	1161.59326	662.163776	673.366151	0.074370	0.964977	0.887479
min	0.000000e+00	0.00000	0.000000	200.000000	0.000000	1.000000	1.000000
25%	5.534000e+03	806.00000	350.000000	360.000000	0.000000	4.000000	5.000000
50%	1.041800e+04	1240.00000	500.000000	515.000000	0.000000	5.000000	5.000000
75%	1.872575e+04	1815.00000	740.000000	753.000000	0.000000	5.000000	5.000000
max	1.918691e+06	88385.00000	57960.000000	57960.000000	1.000000	5.000000	5.000000

Source: made by us using google colab (20/06/2021 at 00:19)

Firstly, from the following tables, we can see that there are more than 1 million rows, but that does not represent the number of riders; it represents the number of trips in a period of "*11 month*". [128]

We notice that the mean of rider_rating and driver_rating is almost 4,5/5, and 75% of them have given 5/5 that is what gives us a good impression about our customer's and supplier's satisfaction, by looking to the max and the min (0; 57000) of the estimated_cost we can notice

[128] We cannot specify the period's dates for reasons related to the company's credibility with riders.

that there is a big gap between them especially if we took in consideration that 75% of it is less than 750 DA, that is what indicates that our riders have a different purchase behavior.

2.2 Calculating recency, monetary, frequency

The first step is calculating Recency, Monetary & Frequency for every customer:

- **Monetary =** The sum of all estimated cost for every customer
- **Frequency** = The number of times the customer has made a trip
- **Recency** = The number of days between the last trip and "today." [129]

Figure3.11: Monetary, Frequency, Recency

	rider	recency	monetary	frequency
0	59341fff2c2c0d00267b7324	257	730.0	2
1	5936c1edeb5452002683c505	647	2530.0	6
2	5936d1e27eb09900264b037b	320	1170.0	4
3	593828c4fb9bff00269e2f8f	723	460.0	1
4	59386269cc1d0d0026e1c8b8	84	904.0	2
...	...	...	...	...
145535	5e5ab39cb57d320007088eca	0	701.0	1
145536	5e5ab44cba26ab00086890a8	0	300.0	1
145537	5e5ab4b8448d580009d1194c	0	619.0	1
145538	5e5ab56b1c3de20007b8133a	0	854.0	1
145539	5e5ab915ca3f940008c25d62	0	754.0	1

	recency	monetary	frequency
count	145540.000000	145540.000000	145540.000000
mean	217.198798	1969.172333	3.344716
std	183.554144	3471.960691	6.323553
min	0.000000	0.000000	1.000000
25%	59.000000	533.000000	1.000000
50%	174.000000	951.000000	1.000000
75%	338.000000	1950.000000	3.000000
max	783.000000	132352.000000	249.000000

Source: made by us (using google Colab)

If we look at these tables, we can see that we have more than 140,000 riders, and we notice that the mean of the monetary frequency is almost 2000 DA, three times. Its maximum is 130000 DA, 249 times by order. This gap is justified by Pareto's theory (20% of riders representing 80% of the company's revenue), and what has made it sure is that 75% of our Rider's monetary is less or equal than 2000 DA, their frequency is less or equal than three times, and 25% of the recency is less or equal than two months. However, at the same time, we can see that the mean of recency is almost seven months, which gives us a wrong impression about whether our riders are satisfied or not.

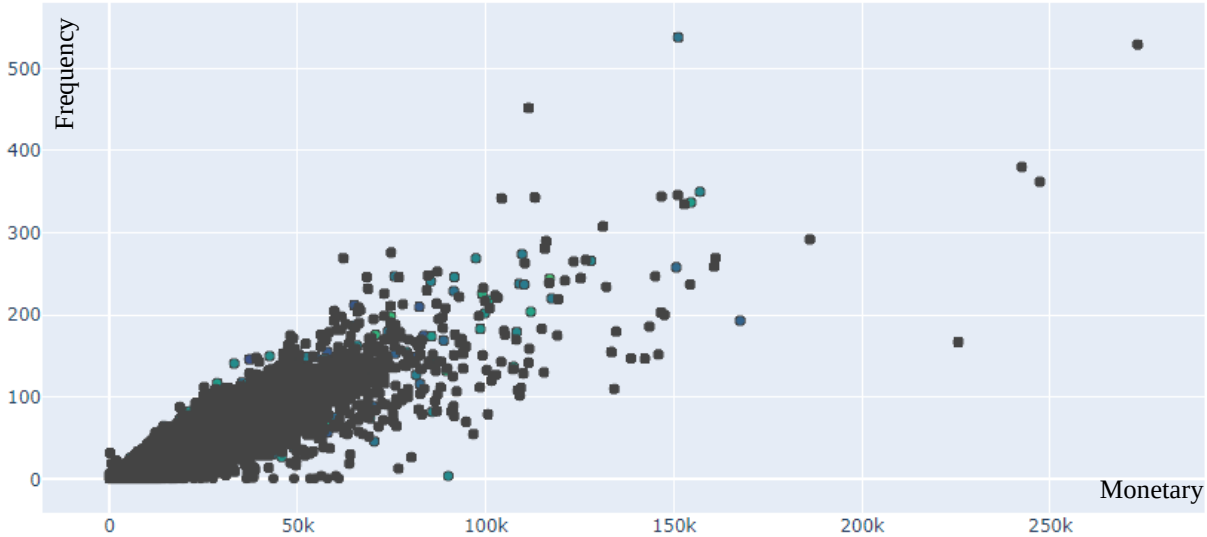
2.2.1 Correlation between variables

The scatters plot shows the relationships between the two variables (frequency and monetary scores). It is common to see visible linear groupings of points on the frequency scale, as frequency often represents a relatively small because frequency often represents a relatively

[129] What we mean by today is the day we made this segmentation

small range of discrete values. For example, if the maximum number of trips is 150, there are only 150 possible frequency values, whereas there could be hundreds of possible recency values and thousands of monetary values. The interpretation of the recency axis depends on the type of recency measure (days or months).

Figure3.12: Scatter plot (Monetary and frequency)



Source: made by us

Table3.1: The correlation coefficient

r^{130}	Monetary	Frequency	Recency
Monetary	**1**	**0,97**	**-0,25**
Frequency	**0,97**	**1**	**-0,22**
Recency	**-0,25**	**-0,22**	**1**

Source: made by us

We notice that there is a strong direct relationship between monetary and frequency. It almost perfect **(r=0,97)**. As the frequency increases, the revenue also increases. So, if a rider makes many trips, his money will also increase, which is clear to see in **Figure3.12.**

If we look at **Table3.1**, we can see an inverse relationship between recency & frequency, monetary. As the monetary or frequency increases, the recency decreases. Customers who recently made a trip generated more revenue than those who made it in the distant past.

[130] The correlation coefficient [1; -1]

However, at the same time, we can see many-valued riders have a high frequency and monetary, but their last trip has been before three months.

The correlation coefficient is so (0,25; 0,2) weak, so it is common to see a customer whose recency is 0, but in the same he has a weak frequency and monetary score, these riders are new, and the reason why their monetary score is low is that Yassir offers a discount of 50% for the first two trips.

2.3 Choosing the number of clusters K

The step before applying the means algorithm is to determine the number of clusters. As we explained before, there is no method for determining the exact value of K, but an accurate estimate can be obtained using the Elbow Method, which helps us determine the optimal cluster number for optimal inertia. It is commonly used to compare and calculate results across different values of K, which is the mean distance between data points and their cluster centroid, mean distance to the centroid as a function of K is plotted, and the "elbow point," where the rate of decrease sharply shifts. By using this graph, we can determine the optimal number of clusters K. Inertia graph are as follows:

Figure3.13: the optimal number of clusters "Recency."

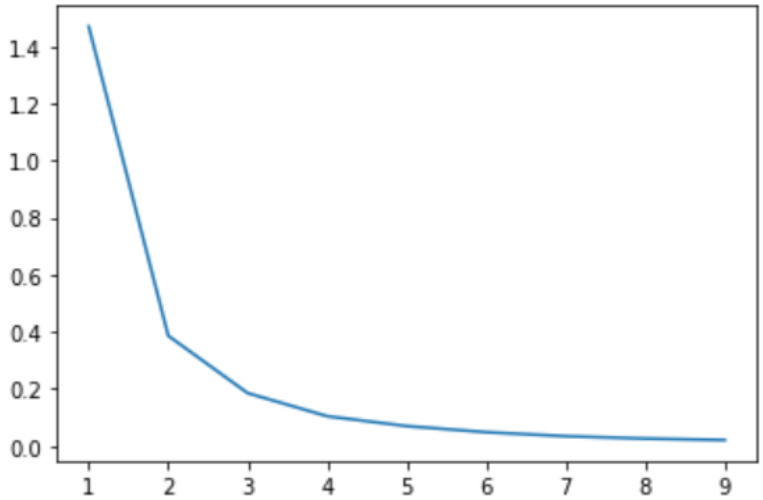


Figure3.14: the optimal number of clusters "Monetary."

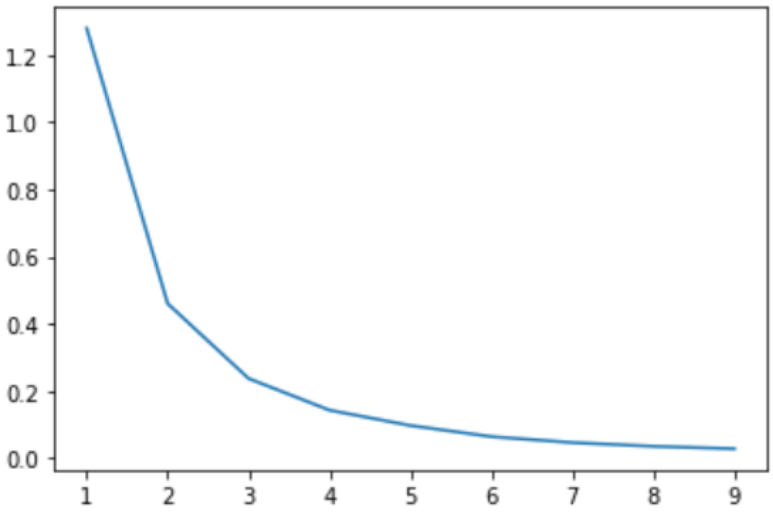
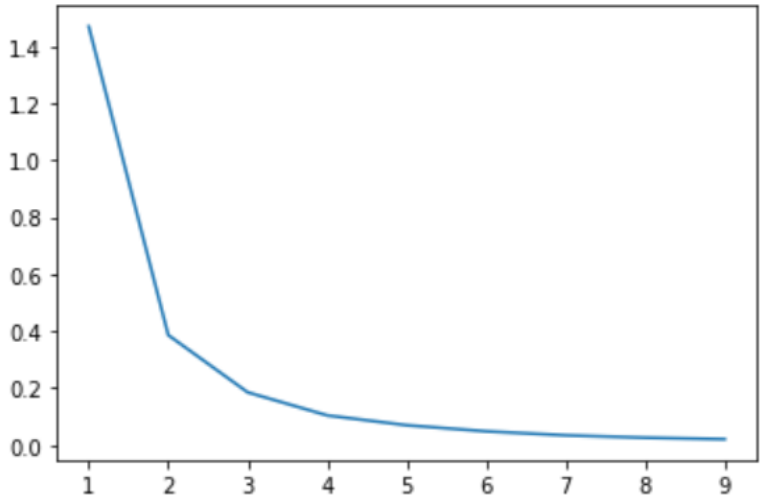


Figure3.15: The optimal number cluster “Frequency."



Source: made by us

As we can see from these graphs, the elbow point for the three variables, recency, frequency, monetary, is three. So, we can say that the optimal number of clusters K is 3, but in our case, it is better to take 4 clusters to help our business, and our riders’ segmentation be more precise, we will choose four bins for each axis of R.F.M., and that is what will help us to divide our riders to homogenous segmentation.

2.4 R.F.M. Score

2.4.1 Calculate the clusters

After choosing the number of groups, we can now apply the K-mean algorithm, which will score between 1 and 4 for the three variables:

- 1 for the low values

- 2 for the medium values
- 3 for the high values
- 4 for the very high values

As for frequency and monetary, a high value means a rider who made many trips and spent much money, and that is a valued customer. However, this will be a problem in the case of recency because the algorithm will give a score of 4 for the Rider who has a considerable recency number, so it will give a 4 to a rider who did not make a trip before five months and gives 1 to a rider who has made a trip before one or two days, so the problem here is the order the Rider who has a low recency value must have a high score. The Rider who has a high recency value must have a low score. Here is the following table, which presents the clusters:

Figure3.16: The clusters of variables

	rider	recency	monetary	frequency	ClusterRecency	ClusterMonetary	ClusterFrequency
0	59341fff2c2c0d00267b7324	248	1810.0	5	3	1	1
1	593ded1662f0700026f9f31d	193	1266.0	2	3	1	1
2	5942a6d00679690026743c80	204	562.0	1	3	1	1
3	5943c7400679690026743cd4	155	1797.0	2	3	1	1
4	5951ccf8daade900269fc029	137	1134.0	3	3	1	1
...	...	...	...	...	...	...	...
247032	5d2d788aead25600072c6755	3	59288.0	131	4	4	4
247033	5d4c7d9a62574b0009db1043	0	166767.0	226	4	4	4
247034	5d6aa73876c3350008d7f220	20	93508.0	155	4	4	4
247035	5d72b9d23c9c630007f20f79	31	88107.0	144	4	4	4
247036	5d9c3ba3e3af0c00085eebd5	2	92641.0	167	4	4	4

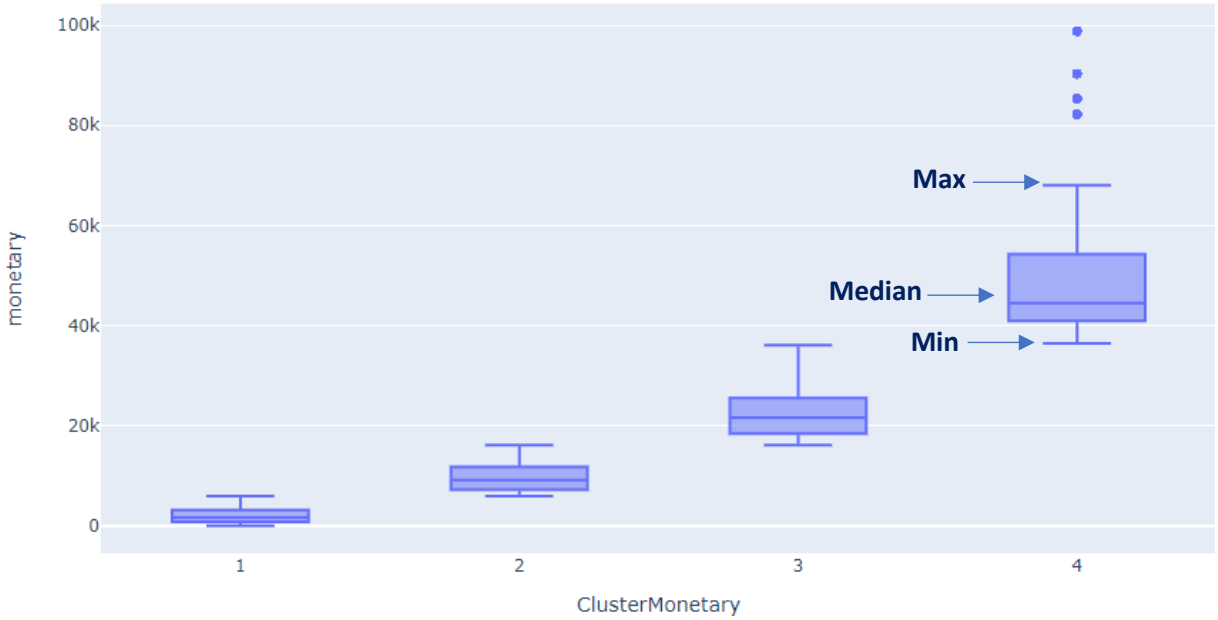
Source: made by us

- **Monetary cluster**: By looking at these figures, we can see the gap between the clusters:
  - The number of riders with one monetary cluster is almost "56 K rider."[131]Furthermore, they represent 40% of all the riders. The max of cluster 1 is almost 6000 Da, and the min is 0 Da. This cluster is representing the riders who have the lowest monetary value.
  - For cluster 2, their 42k Rider represents 30% of all the riders, the min is the max of the first cluster, 6000 Da, and the max is 16000 Da; these riders have a medium value.
  - The monetary cluster 3 represents the riders who have a high value. The number of these riders is 30k (23% of all the riders), the min is 16000 da, and the max 37000 Da.

[131] 1k = 1000

- Cluster 4 is the very high riders. Their min is 37000 Da, and their max is 60000 Da; by a 10 K rider, they represent 7% of them.
- Four riders have generated a monetary of more than 90000 da they do not belong to any cluster; so, we have considered them an extra high.

Figure3.17: Monetary cluster description

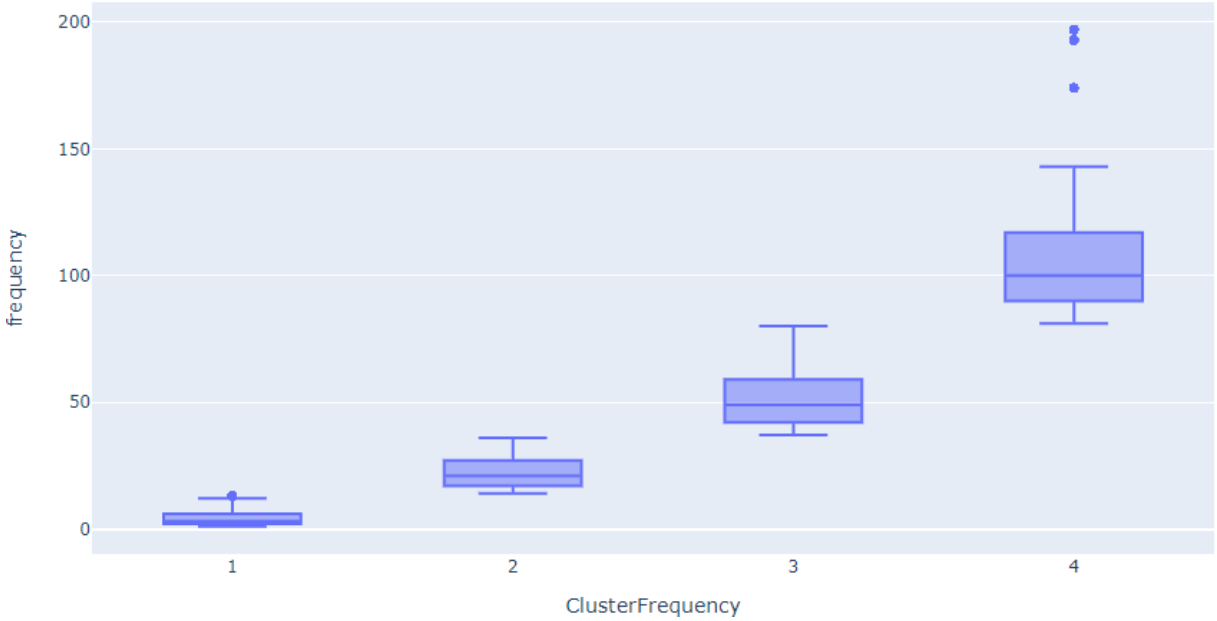


Source: made by us

➢ **Frequency Cluster:**

- The number of riders with one frequency cluster is almost 70 K rider, and they represent 50% of all the riders. The max of cluster 1 is almost 13 times, and the min is one time; this cluster represents the riders who have the lowest frequency value.
- Cluster 2, their 42k Rider, represent 30% of all the riders, the min is 13 times, and the max is 36 times; these riders have a medium frequency value.
- The frequency cluster 3 represents the riders who have a high value. The number of these riders is 21K (15%), the min is 36, and the max 60 times.
- Cluster 4 is the very high riders. Their min is 60 times, and their max is 143 times; by 7 K rider, they represent 5% of them.
- If we look at the chart, we can see three points; these points are three riders who have made more than 170 trips, they have an extra high value.

Figure3.18: Frequency cluster description

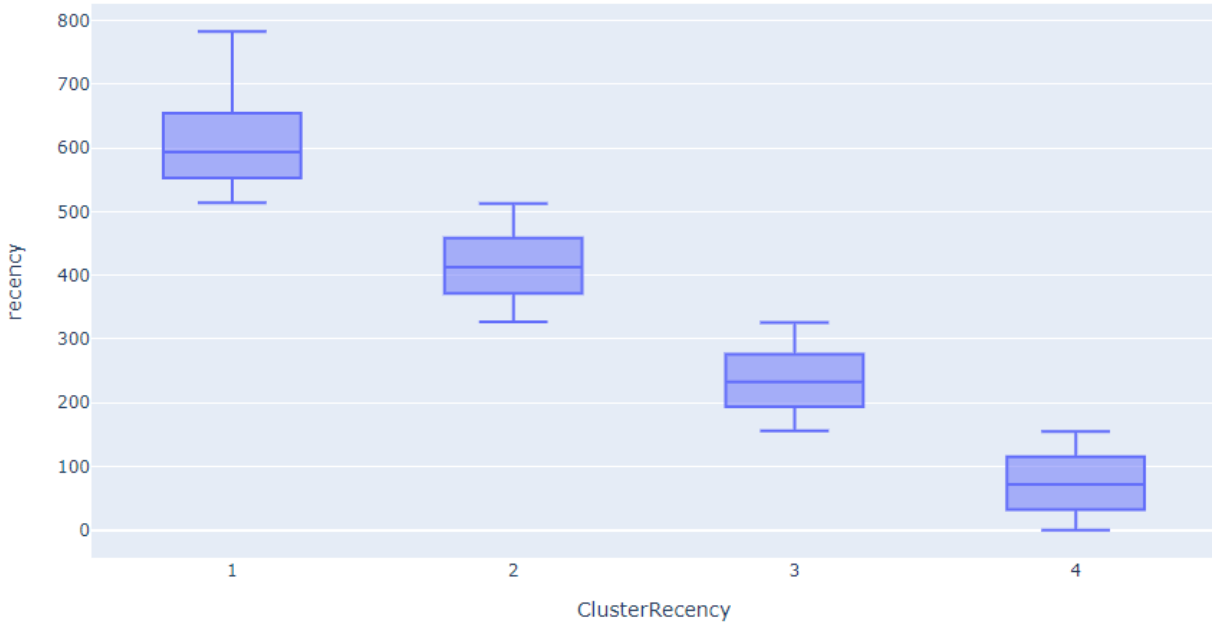


Source: made by us

➢ **Recency cluster:**

- 35% of riders have a recency cluster score of 1, The min of cluster 1 is almost eight months, and the max is one year; this cluster represents the riders who have churned.
- For cluster 2, it represents 45% of all the riders, the min is eight months, and the max is five months; these riders have a medium recency value.
- The recency cluster 3 represents the riders who have a low recency value (12% of all the riders), the min is five months, and the max is three months.
- Cluster 4 is the very high riders; their min is three months, and their max is 0; they represent 8% of them.

Figure3.19: Recency cluster



Source: made by us

2.4.2 Calculate R.F.M. score

Now, after calculating the clusters, we are going to calculate the R.F.M. score using this method:

RFM Score *= (Monetary cluster * 1000) + (Frequency cluster * 100) + (Recency cluster * 10)*

The order is important because we take the monetary value as a primary priority, frequency, and recency value.

Figure3.20: R.F.M. Score

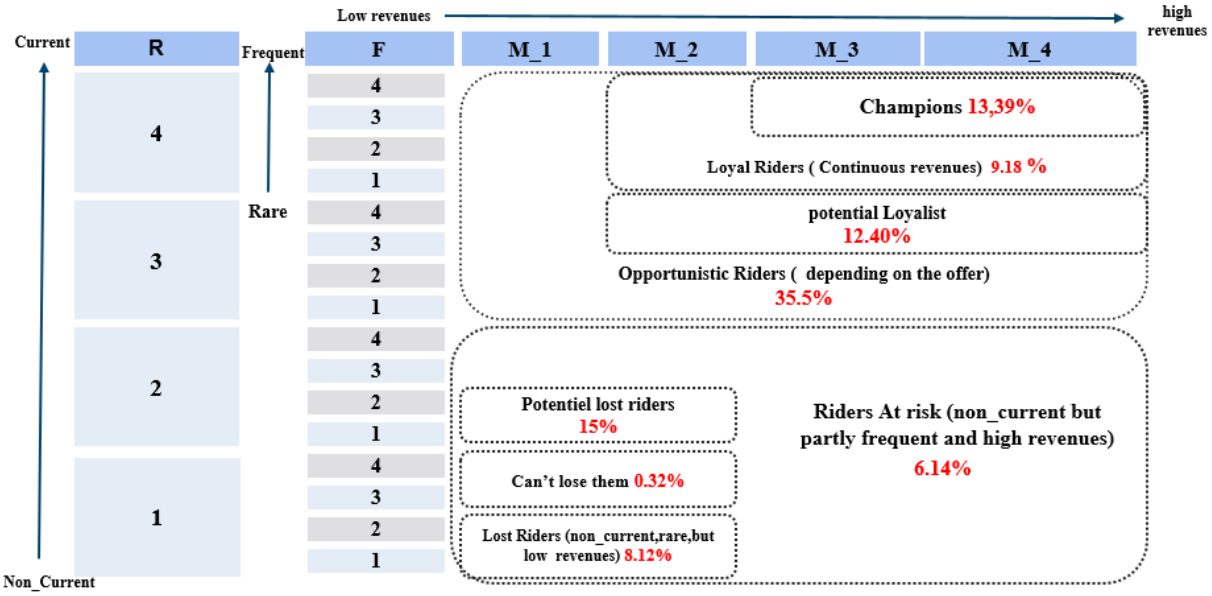
	rider	recency	monetary	frequency	ClusterRecency	ClusterMonetary	ClusterFrequency	RFMScore
0	59341fff2c2c0d00267b7324	248	1810.0	5	3	1	1	113
1	593ded1662f0700026f9f31d	193	1266.0	2	3	1	1	113
2	5942a6d00679690026743c80	204	562.0	1	3	1	1	113
3	5943c7400679690026743cd4	155	1797.0	2	3	1	1	113
4	5951ccf8daade900269fc029	137	1134.0	3	3	1	1	113
...	...	...	...	...	...	...	...	...
247031	5cfc0e26c6938b4259cf0d22	0	65759.0	153	4	4	4	444
247032	5d2d788aead25600072c6755	3	59288.0	131	4	4	4	444
247034	5d6aa73876c3350008d7f220	20	93508.0	155	4	4	4	444
247035	5d72b9d23c9c630007f20f79	31	88107.0	144	4	4	4	444
247036	5d9c3ba3e3af0c00085eebd5	2	92641.0	167	4	4	4	444

Source: made by us

2.5 Customer segmentation

Our research aims to segment the riders into a homogenous segment to provide them the best possible offer. By analyzing and interpreting the R.F.M. score, we divided the customers into eight segments. The following figure explains the eight segments:

Fugure3.21: rider' segmentation



Source: made by us

Table3.2 : segments' description

Segment	Description	RFM SCORE	(%)
Champions	These are the best riders. They have recently made a trip, spent a lot of money, made many trips, are the best riders, and created positive word-of-mouth for around their circle of acquaintances.	444 443 433 434	13,39
Loyal riders	These riders are loyal, even if they did not make many trips recently, but we can notice that they generated a massive amount of profitability every time they made a trip, which explains the high value they have.	442 432 422 414	9,18
Potential loyalist	These riders have a high monetary value, and they have made many trips. They have a high-frequency score; even if they do not take a trip regularly, they still so profitable	343 344	12,4

Opportunistic Riders	These riders are depending on the offer. We call them "The opportunistic riders" they do not make many trips because they make trips only when there is a coupon, we can see that these riders have a different monetary value, and that can be explained by the nature of the coupon sometimes there 100% coupon (free trip), so the monetary value will be low and sometimes the coupon is 40% and that' what explains the high monetary value.	311 312 313 314 341 342 343 344	35,5
Riders at Risks	These riders have a good monetary value and a suitable frequency score too, but they have not made a trip for a long time, and that is what makes them at risk because even if their last trip was before a long time ago, they still a good segment, so, it will be a significant loss if we lose with.	244 243 244 233 243 214	6,14
Cannot lose them	These riders may not have a high monetary value or have not been so active recently, but they have a high-frequency score. They have made many trips before a long ago, so we cannot lose them.	141 142 131 132	0,32
Potential Lost riders	These riders seem like they have "churned rider."[132], they have not made a trip for a long time, their frequency score is so low, and their monetary value low too.	211 212	15
Lost riders	These riders have not been active for a long time and, they have the lowest monetary and frequency value, and they have the highest recency period; these riders have "churned"; they are not a potential rider anymore.	111 112 121 122	8,12

Source: made by us

[132] A churned rider is a rider who stopped using the company's service & product.

3. The CRM strategy for every segment

The aim of our research is not only about Rider's segmentation. It is about putting every segment in the right CRM strategy and recommending the best action for them to improve their profitability. To put every segment in the best customer relationship management's strategy, we must first calculate every segment's lifetime value, then calculate the profile of one of them and comparing them with the R.F.M. score in order to choose the right strategy for every segment of them and recommend the best possible actions.

3.1 Segment lifetime value

The segment lifetime value is the total revenue generated by one segment during its lifetime. We had calculated every segment's lifetime value, and then we calculated every segment's profit. We are going to describe all the steps of calculating in the following points:

Segment Lifetime value *= [Average segment's value * average lifetime (months)* gross margin (%)]*

- ***The Average segment's value:*** it is the sum of the monetary values of all the riders in the segment:

$$\sum_{first\ Rider\ in\ the\ segment}^{last\ Rider\ in\ the\ same\ segment} \frac{MONETARY}{THE\ NUMBER\ OF\ RIDERS}$$

- ***The average lifetime:*** It represents the life of the customer is using the company's service. It refers to the expected period during which the Rider will use the company's product. It calculated by this formula:

$$1/_{Churn\ rate}$$

- ***The churn rate:*** also known as the rate of attrition or customer churn, measures the percentage of riders who stopped using the service for a reason or another.

- **Gross margin:** is the Yassir's net revenue from the driver revenue it is constant 0,2%

In the following table, we are going to present the segment's lifetime value, the average cost of keeping Rider, and finally, the net profit:

Table3.3: the segment's lifetime value

Segment	Average segment value (DA)	Churn rate[133] (%)	Average lifetime (month)	Gross margin (%)	Segment's lifetime value	The average cost of keeping Rider (DA)	Segment profit (DA)
Champion	90000	2%	50	2%	90,000	7000	87,000
Loyal riders	75000	3%	33	2%	49500	5000	44500
Potential loyalist	65000	10%	10	2%	13,000	3000	10000
Opportunistic riders	40000	50%	2	2%	1600	2000	-400
Riders at risk	35000	35%	3	2%	2100	1000	100
Cannot lose them	25000	40%	2,5	2%	1250	1000	250
Potential Lost	10000	50%	2	2%	400	1000	-600
Lost riders	7500	50%	2	2%	300	1000	-700

Source: made by us

[133] Internal document

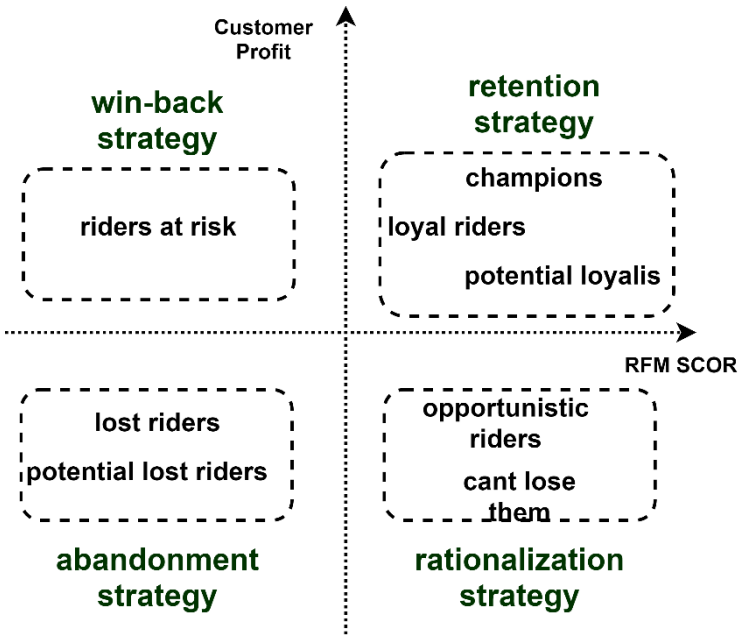
As we can see from the last table:

- The “champions” have the most useful lifetime, and they also have a positive profit.
- The “loyal riders” have the 2nd highest lifetime value, and they have a positive profit.
- The “potential loyalist” riders have a good lifetime value and a positive profit.
- The “opportunistic riders” have a high churn value because they make trips only when there is a coupon code. They have a low lifetime value and a negative profit.
- The “riders at risk” have a not very low lifetime value, and they have a positive profit, even if it is not high.
- The “can’t lose them” riders have a low lifetime value and a positive profit.
- The “potential lost” riders have a low lifetime value and a negative profit too.
- The “lost riders” have a meager lifetime value and a negative profit too.

3.2 Recommend the best CRM strategy for every segment

Based on the main results we get from R.F.M. and the segment's lifetime value and profit, we have recommended the best strategy and the best actions for every segment. The following figure explains how we used the R.F.M. score and segment's profit in order to choose the best strategy:

figure3.22: CRM strategy based on R.F.M. score and segment profit



Source: made by us

3.2.1 Retention strategy:

For the three segments, "champions," "loyal riders," "potential loyalist," The retention strategy is the best strategy for them, these riders are the most profitable, and it has known that retain an existing customer is cheaper than winning a new one, So, the primary purpose is to keep them always related to our service by:

- *First, for the champions' segment:*
- **Offer rewards**: Offer a Comfort Ride (more conformable car than the classic) or Premium Ride (luxurious car) without having to pay extra fees. These customers will feel more special and this will trigger more orders from these customers in the in the future (upselling)
- **Promote new services:** since champions are frequent riders, your new services will always be on their radar. Inform them regularly with attractive newsletters (cross-selling).
- *Second, for the loyal' segment:*
- **Send Push Notifications**: Offer them to try other transportation options (time, comfort, space, or premium cars)
- **Present Bonuses:** Offering bonuses to loyal customers initiates future purchases. For example, offering a free ride for a limited period.
- **Create a personal relationship with them:** by offering a birthday gift, when a rider requests a trip on his birthday, Yassir offers him (her) a free trip or a coupon that's what will create a solid personal relationship between the company & their riders and make them feel that they are unique for the company.
- *Third, for the potential loyalist*
- **Loyalty program:** introduce them to a loyalty program or perhaps an elite club membership so that the company will not lose them.
- **Hold contests:** keep those riders in the loop by holding contests with attractive offers to keep the company permanently in touch with them.

3.2.2Win-back strategy:

The riders at risk are good Riders. Their profitability is good. However, their R.F.M. score is low so, the "win-back strategy," which is an action that aims to regain an existing customer who they have not been so active recently, this strategy seeks an answer to the

question of "Why our riders are not active? ", because if the company can find the reasons why they can provide him the best possible offer, so what we recommend is:

- First, keep these riders in touch by emailing and social networks to determine why they are not as active as they were (pricing, quality service...).
- Then, the company must provide them a personalized, customized offer, an enhanced offer, so they cannot refuse it.

3.2.3rationalization strategy

The opportunistic riders are very active. They generally use the company's service only when there is a coupon. It is why their profit is negative. The "can't lose them" riders also are a good rider, but they can be better, they are active but not always we can see that their recency are so high, but at the same time their frequency and monetary are high, they still always a potential but what we think is these riders are using a competitor service, their profit is negative because their last trip was before a long time. These two segments are a good segment, but they need a re-orientation:

- *First, for the cannot lose them' segment:*
- **Offer them a comeback gift:** by offering them a 50% discount for the first two trips.
- **Examine the service:** most customers may leave because they are not satisfied with the service. Check out the reviews on various sites, or check the driver's rating to understand the real reason.
- *First, for the opportunistic riders' segment:*
- **Create a new coupon strategy:** offer a coupon related to the distance, if the Rider has a trip more than 20 km, he will get a discount of 50%, or if he makes many trips only, he will get a coupon.

3.2.4Abandonment strategy

The difference between the lost riders and the potentially lost riders is the recency score, the lost riders did not make a trip this last year, but the potential losses have been made. We can notice that they have a negative profit and a low R.F.M. score. The company must understand the reason why these riders have churned. It will be helpful to improve their service for the current riders and to avoid the same problem in the future by:

- *First, for the potentially lost riders:*

- **One last promotion:** send a promotional email saying, for example, "We miss you. Here is your gift. Get 25% off your entire order.

➢ *Second, for the lost riders:*

- **Ignore them completely:** This is the last thing the company goes for. Most likely, this type of customer will not come back. Abounding these customers is a strategic step because keeping them would only bring more costs.

4. FRAT Segmentation

The R.F.M. segmentation does not provide a complete understanding of customer behavior because it does not combine with other essential customer attributes such as product preferences or the type of purchase. Hence, the company needs to have a complete view of customer purchase, and this is the reason for using the evolutionary version of R.F.M., The FRAT Segmentation, which refers to the frequency recency amount (Monetary) and including the type of purchase.

We have chosen the champion segment to analyze it using the FRAT segmentation method because they are the best riders. They are the most valuable segment. They have a high monetary, frequency, and lowest recency to understand their purchase behavior and recommend the best possible offer based on upselling and cross-selling strategies.

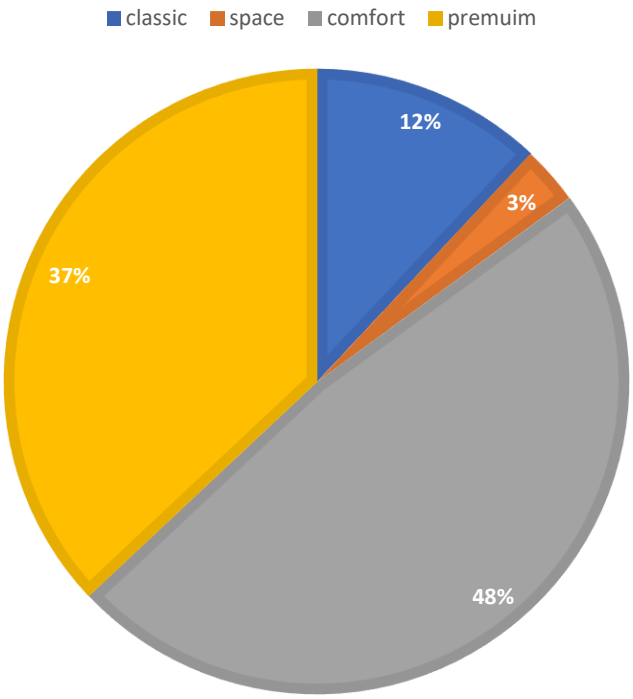
Firstly, we must have a clear understanding of the purchase types that Yassir offers:

- **Classic:** affordable and fast travel (simple cars, a reasonable price).
- **Comfort:** modern and comfortable vehicle (medium cars, a high pricing strategy).
- **Space:** large vehicle with a spacious trunk.
- **Premium:** premium trips are the highest service the company offer; it offers a high luxury car and a high service quality (a high pricing strategy).[134]

In the following figures, we will describe the use of these purchase types by the champions' segment, and we will present the use of coupon too:

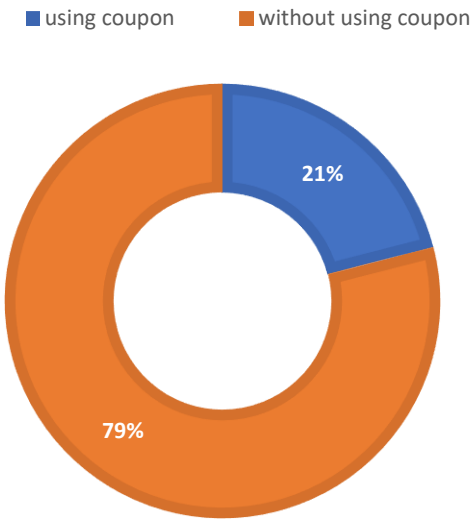
[134] https://yassir.com// (24/06/2021 at 19:18)

Figure3.23: all the trips types that the riders of the champion's segment have made it



Source: made by us

Figure3.24: the use of the coupon by champion's segment redries



Source: made by us

As we can see from the first figure, the riders of the champion's segment use several trip's types:

We can notice from the first figure that 85% of these riders have used the high types. 37% of them have used the premium type, which the most expensive type, the other 48% used the comfort type, which is an expensive service. These riders do not care about how much they will pay. They only want a high-quality service (a luxury car, an excellent driver, a quick service), and what proves this idea is there only 12% of the segment whom they used the classic service, we can also see that only 3% of them use the space trip and that is logical because the use of this types is so rare.

From the second figure, we have relied upon that these riders generally do not use coupons in their trips (only 21% of them used a coupon) because they have some fears about the quality service after using a coupon.

Based on these results and the R.F.M. score; we have made a FRAT segmentation to these riders by including a new score (the type of trip):

- 3 for the premium trip
- 2 for the comfort trip
- 1 for the classic trip

For the space rip, we did not consider it because the percentage is meager, and there are no riders who use this trip regularly, and they only use it in rare cases. We are going to present the segment in the following table:

Table3.4: FRAT segmentation

Segment	**FRAT score**	**Description**
Premium champion	4443	These riders are the best Rider they make many trips, and they use only the premium trips
Comfort champion	4342	These riders have generated a considerable amount of profit because they made many trips, and generally, they use the comfort service

Classic champion	4331	These rides make many trips, but they only use the classic trips and use the coupon

Source: made by us

4.1 Recommendations

As we have explained before, these riders are looking for a high-quality service. They do not care about how much they have to pay. At this point, we have just taken the premium champions because they are the most valuable riders for the company. So, our recommendation actions are based on two strategies "The cross-selling" by satisfying additional, complementary needs that are unfulfilled by the original item and "The up-selling" by providing them a high luxury service to provide additional value to these riders, instead of limiting them to already-encountered products.

- **Up-selling:**

Create a new trip type: The company can provide a new trip type that will be more luxurious than the premium, an extra luxury car, the best possible driver, and high-quality service.

Give them a gold membership: they will feel so special. The company need to do that because that will create a strong personnel relationship by giving them the freedom to choose any driver they want

- **Cross-selling**

Offer new complementary service: provide a complimentary service such as waiting, so the riders can take a trip, and the driver will be waiting for them, or book a driver for one day, for example.

Conclusion

This chapter has tried to concretize the different theoretical notions of our research topic in the previous chapters.

Based on the customer segmentation and its results and the analysis of the different segments, we observed how big data analytics could be applied in customer relationship management. Thanks to the analysis of the results obtained, we were able to detect the usage of the B.D.A. to improve customer relationship management so that we can subsequently confirm our hypotheses.

Through this study, we can conclude that the B.D.A. has become indispensable for the companies to ensure better management, improve their performance, and finally make the right decisions at the right time.

General conclusion

General conclusion

This research study aimed to demonstrate the application and contribution of big data analytics in customer relationship management.

Throughout this thesis, we have reviewed the theoretical approaches by drawing on a vast literature concerning big data, big data analytics, supply chain analytics, and the notions related to electronic and analytical customer relationship management, information within companies, and the information system.

Our project comprises three studies; the first was a literature search (secondary data mining) where we presented the fundamental concepts of big data analytics and studies of supply chain and analytics its applications.

The second and third studies were about variables that drove us to use the observation method during the internship to pick the information we need after collecting the data from an internal source that allowed us to make a simulation and analyze the results using a quantitative methodology. At the end of our research study, we have tried to demonstrate the significant roles of the big data analytics function and its degree of influence on the customer relationship management sector within a company.

In addition, the results obtained and interpreted in the last part of our thesis allowed us to confirm the positive impact of BDA on customer relationship management, to understand and measure the contribution of the latter and thus confirm our hypotheses.

The first hypothesis states that big data analytics can identify known risks and help to predict future risks by spotting patterns and trends throughout the supply chain. the observation of the importance of the BDA in support decision-making in Yassir allowed us to confirm this hypothesis

The second hypothesis, which states that big data analytics will enhance CRM strategies in delivering personalization and customization of services for the customer, is confirmed because we have built and recommended a personalized, customized strategy based on big data analytics techniques.

Thus, we have been able to confirm the third hypothesis. It assumes that integrating big data with customer relationship management facilitate to find out how a product or brand is perceived by customers online and identify the weaknesses and revise the marketing and sales

strategies because using BDA techniques can extract a tremendous amount of information and insights about the customer, the product or the market.

At this stage of analysis, we can say that BDA offers strategic and operational advantages that allow the company to position itself well in managing customer relationships to satisfy its customers more and thus increase their loyalty.

The accomplishment of this modest work was not without constraints. Indeed, we encountered some difficulties on the technical side. It is justified by the absence of practical training in data processing and management software packages, an essential asset to the SCM. The detailed study of these different parameters requires a multidisciplinary culture, which in some cases exceeds the framework of a modest dissertation and finding an internship because, in Algeria, there are only a few companies that use big data analytics in their supply chain or customer relationship management.

Nevertheless, our advantage is that our internship included the necessary aspects to understand the data science and the analysis of the processes within the company, and we hope to have contributed to the reflection on the question of the importance of the BDA techniques in the supply chain management in general and the customer relationship management in particular. We wish that other researchers come to deepen this topic within the Algerian companies.

To conclude, we hope to have drawn the leaders' attention within the companies on the importance of the treated subject, a subject which remains until now little neglected by the Algerian companies. In order to conduct in-depth research on this subject, we propose some topics that require a multidisciplinary culture and incredibly technical skills, such as artificial intelligence and machine learning in SCM in general and in CRM, mainly through the study of the supply chain analytics on a broad aspect of business management.

Bibliography

Books

- BUTTLE (F), & MAKLAN (S): « *Customer Relationship Management, Concept and Technologies*», Third Edition, Routledge, 2015, p.3.
- DYCHE (Jill): « *The CRM Handbook: A Business Guide to Customer Relationship Management* », Addison-Wesley, Information Technology Serie, Boston (Mass.), 2002, p.30.
- GREENBERG (P): *« CRM at the Speed of Light: Essential Customer Strategies for the 21st Century*», McGraw-Hill Osborne Media; Third Edition, January 1, 2004
- GREENBERG (P): «*CRM at the Speed of Light: Capturing and Keeping Customers in Internet Real Time.* », McGraw-Hill (Osborne), USA, 2001, 1st edition.
- JELONEK (D), CHLUSKI (A): «*Possibilities of using CRM systems in health care institutions*», Information technologies in public administration and healthcare., SGH Publishing House, Warsaw, 2010, p.42-43.
- JELONEK (D): « *The Role of the CRM System in the Development of the Customer Capital of An Enterprise Processes of Capital Supply in Production Enterprises*», Serie Monographs No 1, 2006, p.103-109.
- KOTLER (P) et al: *« Principles of Marketing»*, 4th European Edition, Pearson Education Limited, England, 2005, p.6.
- KUMAR (V), & REINATRZ (W): « *Customer relationship management, concept, strategy, and tools*», Third edit, Springer Texts in Business and Economics,2018, p.12.
- KUMAR (V), PETERSEN (A) & PETERSEN (J.A): *«Statistical Methods in Customer Relationships Management»*, 1 ST Edition, Hoboken New Jersey, 2012
- PEPPERS (D), RODREGS (M): *« Managing Customer Relationships; A Strategic Framework»*, 2nd Edition. U.S, 2010
- PEPPERS (D), RODREGS (M): *« Managing Customer Relationships; A Strategic Framework»*, 2nd Edition. U.S, 2010
- SERRAO (Carlos), et al: *« Handbook of Research on Trends and Future Directions in Big Data and Web Intelligence»*, Information Science Reference, global book series (ADMDM), 2015.
- TSIPTSIS, (K) & CHORIANOPOULOS, (A): *« Data Mining Techniques in CRM: Inside Customer*, » edition Wiley, Chichester, 2009, p.335.
- WILEY (J), SONS (L): «*Supply chain analytics for dummies*», Open-Text, special edition, The atrium southern gate Chichester, 2017, p.14-16.

Articles

- AB HAMID (N.R): *« E-CRM: are we there yet? »*, Journal of American Academy of Business, 2005, p.60.
- AB HAMID (N.R): *« E-CRM: are we there yet? »*, Journal of American Academy of Business, 2005, p.65.
- ALICKE (K) et al: *« Big data and the supply chain: The big supply chain analytics landscape »*, McKinsey & Company Operations Extranet, 1(of 2), 2016, p.3-4.
- ALICKE (K) et al: *« Big data and the supply chain: The big supply chain analytics landscape »*, McKinsey & Company Operations Extranet, 1(of 2), 2016, p.4-5.
- ANSHARI (M): *« Online Learning: Trends, Issues, and Challenges in Big Data Era»*. Journal of E-Learning and Knowledge Society. 2016, p.124.
- ANSHARI (M): *«Customer Relationship Management and Big Data Enabled: Personalization & Customization of Services»*. Applied Computing and Informatics, 2018, p.8-9.

- ANSHARI (M): *«Customer Relationship Management and Big Data Enabled: Personalization & Customization of Services»*. Applied Computing and Informatics, 2018, p.6-8.
- BANU (Ayesha) & YAKUB (Mhd): *«Evolution of Big Data and Tools for Big Data»*, Journal of Interdisciplinary Cycle Research, XII(X), October (2020), p.312.
- BANU (Ayesha) & YAKUB (Mhd): *«Evolution of Big Data and Tools for Big Data»*, Journal of Interdisciplinary Cycle Research, XII(X), October (2020), p.312.
- BANU (Ayesha) & YAKUB (Mhd): «Evolution of Big Data and Tools for Big Data», Journal of Interdisciplinary Cycle Research, XII(X), October (2020), p 312.
- BOLLIER (D) & FIRESTONE (C.M): *« The promise and peril of big data*». Washington, DC: Aspen Institute, Communications and Society Program, 2010, p.1.
- CHANDRA (S), STRICKLAND (T. J)*: « Technological Differences Between CRM & E-CRM». Issues in Information Systems*, *5*(2),2004, 408–413.
- CHIANG (D), LIN (C.-P). and CHEN (M.-C*): «The adaptive approach for storage assignment by mining data of warehouse management system for distribution centres »*, Enterprise Information Systems, 2011, p 219–234.
- CHOI (Tsan-Ming), STEIN (W. Wallace) & WANG (Yulan): « *Big Data Analytics in Operations Management »,* Production and Operations Management Society, Hong Kong, December 2017
- CROSBY (L), JOHNSON (S): *«CRM and management»*, Marketing Management, Jan.–Feb 2002, p.10.
- DARYANTO (Yosef), TIWARI (Sunil): *«Big data analytics in supply chain management between 2010 and 2016: Insights to industries »*, computers and Industrial Engineering. November 2017, p.8.
- DAVENPORT (T. H), DYCHE (J): *«Big Data in Big Companies»*. SAS Institute Inc, July 13, 2016, p.19-20.
- ELLIS (Simon), SANTAGATE (John)*: « The path to a thinking supply chain»,* IDC Technology Spotlight, Aug 2018, p.3
- FAN (J), HAN (F), LIU (H): *« Challenges of big data analysis»*. National Science Review, 2014
- FAN (S), LAU(R-Y) & ZHAO (J.L): *«Demystifying big data analytics for business intelligence through the lens of marketing mix»*. Big Data Research, 2(1), 2015, 28-32.
- FJERMESTAD (J), ROMANO (N): «Electronic Commerce Customer Relationship Management», International Journal of Electronic Commerce, 2006, p.83-85.
- FRAWLEY (A): *« Evolving to e-CRM: How to optimize interactive relationships between you and your customers»*, An exchange, Inc white paper (First Serie), 2003, p.11-12.
- GANG (Gunasekaran) et al: *« Big data analytics in logistics and supply chain management: Certain investigations for research and applications»*. International Journal of Production Economics. 2016, p.104.
- GANTZ (John) & REINSEL (David): *« The digital universe in 2020*: *Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East »*, Internet Data Center (IDC), December 2012, p.3.
- GORMAN (M. F) & KLIMBERG (R. K): *« Benchmarking Academic Programs in Business Analytics. Interfaces »*, INFORMS, June 2014, p.329-341.
- GUO (S.Y) et al: *«A Big-Data based platform of workers behavior: Observations from the field»*, Accident Analysis & Prevention, Elsevier Ltd, 2016, p.299-309.
- HAO (H-Y), ZHAO (S-H): *«Applications of Data Mining in the Enterprise Customer Relationship Management»*. Scientific paper, International Conference on Information Technology and Management, Beijing, 2013, p.57-59.
- HAUSMAN (A): *« Doing Business in A Digital World. Journal of Business & Industrial Marketing»*, Emerald Group Publishing Limited. Bingley, United Kingdom, 2005
- HUANG (Guang-Bin) et al: « *Efficient and rapid machine learning algorithms for big data and dynamic varying systems »,* IEEE Trans. Syst. Man Cybern. Syst. 47(10): 2625–2626, 2017, p.2.

- HUNAITI (Ziad): The Ride-Hailing Mobile Application for Personalized Travelling. Modern Applied Science, October 2018, p.244-245.
- HUSSIEN (Abdou) et al*: « How Many Old and New Big Data V's Characteristics, Processing Technology, and Applications (BD1) »,* IJAIEM, September 2020, p.19.
- JIN (Y), JI (S)*: « Partner choice of supply chain based on 3d printing and big data »,* Information Technology Journal, 2013, p.6822.
- JO-TING WEI et al. ii: *« A review of the application of R.F.M. model*, » African Journal of Business Management, Vol. 4(19), 2010, p. 4199-4206, p.4200.
- JULTA (D), CRAIG (J), & BODORIK: *« Enabling and Measuring Electronic Customer Relationship Management Readiness*», in the Proceedings of the 34th Annual Hawaii International Conference on System Sciences, Big Island (Hawaii), Organizational Systems and Technologies Track, Hawaii, 2001, p.15.
- KAMBATLA (Karthik) et al: *« Trends in big data analytics »,* Journal of Parallel and Distributed Computing, July 2014, p.14.
- KENNEDY (A), KELLEHER (C) & QUIGLEY (M): *«CRM best practice»,* getting it right first time, ESBI, Proceedings of the Irish Academy of Management (IAM) Conference, GMIT, Sept 2005, 7–9.
- Kennedy (A): *« Electronic Customer Relationship Management (e-CRM): Opportunities and challenges in a Digital World»,* Irish Marketing Review, No.1&2 ,2006, p.60.
- KIM (S): *« Forecasting short-term air passenger demand using big data from search engine queries»*. Automation in Construction. 2016.
- KIM, YUN (B): *«Big Data: A Temporary Thing? »*, Sungkyunkwan University, Republic of Korea, 2015
- KIM-SOON (N), ZULKILI (M. F): « *The impact of e-CRM Business performance of small company*». Journal of Engineering and Technology 3, p.145
- KUMAR (Abhimanu) et al: « *Petuum: A new platform for distributed machine learning on big data »,* IEEE Trans. Big Data 1(2): 49–67, March 2015, p.5.
- LLamasoft: *« Supply chain simulation: why its time has come»,* white Paper, 2016, p.13.
- LLOYD (G) (2001): *« Electronic Customer Relationship Management: Benefits and Pre-Implementation Considerations»,* America's conference on information systems, 2001, p.842-843.
- Mc GOVERN (Matt)*: « Creating a thinking supply chain for the cognitive era »,* Watson Customer Engagement, Mar 2017, p.5.
- MILLS (Steve) et al: *«Demystifying big data: a practical guide to transforming the business of government»,* Tech America Foundation, Washington, 2012, p.10.
- NACTO: Ride hailing services*: «opportunities & Challenges for cities »,* New York, June 2016, p.4-5.
- PAN (S. L), LEE (J. N): *« Using e-CRM for a unified view of the customer »,* Communications of the ACM, 46(4), 2004, p.95–96.
- PRAJOGO (D), & OLHAGER (J): *« Supply chain integration and performance: The effects of long-term relationships, information technology and sharing, and logistics integration»*, International Journal of Production Economics, 2012, 135(1).
- PRICEWATERHOUSE (PWC): *« Capitalizing on the promise of Big Data»,* 2013, p.2.
- PUIGJANER (Luis), et al: *« Supply chain monitoring: a statistical approach»,* European Symposium on Computer Aided Process Engineering, Barcelona, 1 Jan 2005, p.3.
- RANJAN (Rajiv): *« Modeling and simulation in performance optimization of big data processing frameworks»,* IEEE Cloud Computing. 2014, p.16
- RAYLE (L) et al: « Just a better taxi? A survey-based comparison of taxis, transit, and ride-sourcing services in San Francisco ». Transp policy 45, San Francisco, 2016, p.169-17
- RIAHI (Azzedine) & RIAHI (Sara)*: «The big data revolution, issues and application»,* IJARCSSE, August 2015, p.167.
- RIAHI (Sara): *« Big Data and Big Data Analytics: concepts, types and technologies»*, International Journal of Research and Engineering, (2018), p.524–528.

- ROWE (S), & POURNADER (M)*: « Supply Chain Big Data: how big data is shaping supply chains of tomorrow »,* Series Part 1, KPMG, March 2017, p.8-9.
- ROWE (S), & POURNADER (M)*: « Supply Chain Big Data: how big data is shaping supply chains of tomorrow »,* Series Part 1, KPMG, March 2017.
- SADEGHI DARVAZEH (Saeid), RAEESI VANANI (Iman), & MUSOLU (Farzaneh Mansouri): *«Big Data* Analytics *and Its Applications in Supply Chain Management»,* March 25th 2020
- SALEHAN (M), KIM (D.J): *«Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics»,* Decision Support Systems, ElsevierB.V, 2016
- SARAVANAN (Balaraj): *« Optimization model for improving supply chain visibility»*. Infosys Labs Briefings. 2013, p.15.
- SHAO (Guodong), SHIN (Seung Jun) & SANJAY (Jain)*: «Data analytics using simulation for smart manufacturing»,* In: Proceedings of the Winter Simulation Conference. IEEE; 7 Dec 2014, p.19.
- SLAVAKIS (K), GIANNAKIS (G) & MATEO (G): *« Modeling and optimization for big data analytics:(statistical) learning tools for our era of data deluge»*. IEEE Signal Processing Magazine. 2014, p.19-20
- SRIVASTAVA (J) et al: *« A case study for analytical customer relationship management»*. PAKDD, Berlin, 2002, p.22-24.
- SUN (Ning): *« The issues of customer relationship management based on data mining»*. Technol Ind 11(6), 2011, p.59–61
- TEKLITZ, (F), McCARTHY (R.L): *«Customer Relationship Management: The Ultimate Guide to the Efficient use of CRM »,*1st Edition, Wiesbaden, Germany,2001, p.64.
- VARELA (I. R), & TIAHIONO (B): *« Big data analytics in supply chain management: trends and related research»*. 6th International Conference on Operations and Supply Chain Management, 1(1), Bali, 2014.
- VENKATESH (G), ARUNESH (K): « *An Overview of Big Data Characteristics and Enhancement of Big Data Application,* », International Journal of Computer Science & Technology, Jun 2019, p.16-17.
- WANG (G), et al*: « Big data analytics in logistics and supply chain management: Certain investigations for research and applications»,* International Journal of Production Economics, 2016, p.5.
- WANG (G), GUNNASEKARAN (A) & NGAI (E, W): *« Distribution network design with big data: Model and analysis »,* Annals of Operations Research. 2018, p.539-551.
- WANG (H), et al: *«big data and industrial internet of things for the maritime industry in north western Norway»,* In TENCON 2015-2015 IEEE Region 10 Conference, 2015, p.1-5.
- WANG (Y): *« Robust causal dependence miningin big data network and its application to traffic flow predictions»,* Transportation Research Part C: Emerging Technologies, 2014.
- WATSON (J)*: «Big Data Analytics: Concepts, Technologies, and Applications»*. Communications of the Association for Information Systems, 2014, p.1261.
- Watson customer engagement: *« The AI journey: Artificial intelligence and the supply chain»,* IBM, November 2018, p.2.
- WU (J H), WANG (S C) TAI (W.C): *« Mobile shopping site selection: the consumers' viewpoint»*. In System Sciences. Proceedings of the 37th Annual Hawaii International Conference, 2004, p.8.
- WU (J H), WANG (S C): *« What drives mobile commerce? An empirical evaluation of the revised technology acceptance model»*. Information & management, 2005, p.726.
- XIA (W), WANG (X) & WANG (Q-S): *« The data mining in the enterprise customer relationship management».* Technol Manage 6, 2007, p.196–198.
- ZHANG (Z): *« The study of SME customer relationship management based on data mining»*. Econ Forum 492(7), 2011, p.202–204.

- ZHONG (R-W), WANG (H-P): *«The customer relationship management system based on data mining»*. Shaoguan University, 31(3), 2010, p.20-21.

university papers

- DALEN (C) & DALHLBOM (F): *«Big Data in the Telecom Industry: A Study of How Big Data Affects Innovativeness and Market Dynamics»*. Master Thesis in Marketing Spring, Stockholm School of Economics, Sweden, (2014), p.7.
- METTARGARUNAGUL (M), PENSIRI (P): « The use of e-CRM to enhance customer relationship: *the case of Toyota Mahanakorn* », Master thesis in information technology management, Malardalen university (Sweden), 2011, p.14.
- VACCARI (Carlo): *« Big Data in Official Statistics and Official Statistics»*, PhD Thesis in Computer Science, University of Camerino, 2014.
- VAN NGUYEN (T): *« Exploring applications of Big Data Analytics in Supply Chain Management »*, PhD Thesis in PhD Philosophy, university of Greenwich, London, September 2019.

Webography

- Big Data basic concepts and benefits explained - TechRepublic
- How Does k-Means Clustering in Machine Learning Work? | by Anas Al-Masri | Towards Data Science
- http://www.wsj.com/articles/at-ups-the-algorithm-is-the-driver-1424136536
- https://express.yassir.io/
- https://market.yassir.com/
- https://research.google.com/colaboratory/intl/de/faq.html
- https://searchbusinessanalytics.techtarget.com/definition/big-data-analytics
- https://studyonline.unsw.edu.au/blog/descriptive-predictive-prescriptive-analytics
- https://techjury.net/blog/how-much-data-is-created-every-day/#gref
- https://www.bigdataframework.org/short-history-of-big-data/
- https://www.digital-business-trends.com//
- https://www.edureka.co/blog/big-data-characteristics
- https://www.ibm.com/analytics/hadoop/big-data-analytics
- https://www.python.org/doc/essays/blurb/
- https://www.researchandmarkets.com/reports/5027885/
- https://www.researchandmarkets.com/reports/5314994/
- https://www.sas.com/en_us/insights/analytics/big-data-analytics.html
- https://www.statista.com/statistics/871513/worldwide-data-created//
- https://yassir.com/
- https://yassir.com/business/
- https://yassir.com/client/
- https://yassir.com/en/telemedicine/
- Introduction to K-means Clustering (oracle.com)
- R.F.M. Analysis For Successful Customer Segmentation - Putler
  Rolls Royce Shifts In Higher Gear With Big Data (datafloq.com)

Table of contents