The Importance of Big Data for Understanding Consumer Buying Habits, Increasing Competitive Advantages in Retail Trade

BULIAN, Braian de Souza [1]
ALENCAR, Cícero Aparecido [2]


abstract

In the current market scenario, companies are increasingly needing to stand out. Understanding how customers and the market itself behave is essential for companies to program to better serve them, in order to achieve competitive advantages over their competitors. The purpose of this paper was to present how Big Data can help retailers better understand their customers' specific purchasing and behavioral habits and how it can help them make internal processes more efficient, which enables them to increase their competitive advantage. In addition, a theoretical model for implementing Big Data was presented. The present study evolves according to the position of Taurion (2013). "... Big Data creates value for companies by discovering patterns and relationships between data that were previously lost not only to internal data warehouses, but to the Web itself, tweets, Facebook comments, and even YouTube videos. "It was concluded that the processing and analysis of Big Data, if properly applied, provides improvements in several processes within companies, such as: marketing, logistics, decision making and others. It is worth remembering that companies need to carry out a feasibility study on the application of this technology in order to know if their company has minimum structures required for the application of tools capable of reading and analyzing this large amount of data.

Key words: Big Data, Retail, Competitive Advantages.

1. INTRODUCTION

Although very important, due to the creation of formal jobs, economic movement, capital turnover and others, retail has been suffering with considerably negative results. According to Globo (G1) by Daniel[3] Silveira and Marta Cavallini, since the year 2015, sales in the industry have been declining. With the current unstable political-economic scenario, the tendency is for the market to become increasingly competitive. As a consequence, customers are looking for cheaper products. The proof of this is the increase in the billing and expansion of markets that follow the model "atacarejo", modality of commerce that mixes wholesale of self-service and retail type hypermarket, thus enabling the sale of products at a lower price. According to the magazine[4]ne Exame, a survey presented during the launch of the Brazilian Association of Autoservice Wholesalers (Abaas) points out that this type of store reaches 46.4% of households, surpassing this year (2016) in this criterion the supermarkets, after having already surpassed hypermarkets in 2015.
Consequently, people change their habits in order to adapt, either by looking for the lowest prices or aiming at the best cost-benefit of each product. Based on this scenario, companies are looking for ways to stay in the market, which can be based on simpler solutions such as layoffs, or even more complex as improvements to their processes. All this aiming at reducing costs while satisfying customers, increasing their competitiveness in the market and ensuring the survival of the company, or even its prominence.

The use of Big Data together with specific tools has become an almost necessary option for companies to stand out today. A given step, a purchase made, an access on your social network, a website, a click. Every second, for each action performed, data is generated. They have always existed, but with each passing minute, they are generated in greater quantity and speed. In addition, the data storage and reading capacity has been increasing dramatically, which also makes it possible to turn them into useful information for the market, whatever its field of activity.

Big Data refers to an immense amount of data and therefore, it is necessary to have specific tools capable of reading, so that all are interpreted, generating information that will be useful for the purpose. Companies, by not using this large volume of data, stop interpreting data that, anyway, is generated. Consequently, information about specific habits such as shopping and locomotion is not collected. In this way, companies could improve processes already carried out, such as logistics, marketing, decisions and others, always aiming at the best way to serve their customers, whatever their needs, and thus, obtain competitive advantages in the current market so disputed. In this context, the question is: How can Big Data help retailers understand the specific buying habits of each consumer, and how can this help them?

This paper aims to analyze how Big Data can help retailers to better understand the specific purchasing habits of their customers and thus gain competitive advantage. First it was presented through a bibliographical reference how companies can collect data, find patterns and buying habits of their customers and apply Big Data. Next, it was demonstrated how the interpretation of the data collected helps to increase the efficiency of operational processes already performed by these companies, increasing their competitive advantages. Finally, a model was introduced for the implementation of Big Data in companies, aimed at improving administrative processes based on consumption and market patterns. Thus, it can be affirmed that this work may be relevant for the retail companies and of additional information for the academic area. This is caused by the implantation of Big Data in the companies in this contemporary scenario. In addition, the feasibility of this work is due to the extensive content that can be found on the subject serving as a theoretical basis, and that also present real results and comparisons between companies that use Big Data, and companies that do not use it.

2. THEORETICAL FOUNDATION

This chapter addresses the theoretical basis of the project and provides key information for understanding the concept of Big Data and its role in the contemporary world. In addition, it will show the importance of the application of this technology in companies in order to increase the efficiency in their processes, consequently improving their competitive capacity.

2.1 BIG DATE

Despite its emergence in the early 1990s, Big Data is still considered something new since it is becoming more common only in recent years. Because of this, there is still no exact concept that defines what Big
Data really is. So it is possible to find many definitions on the subject revolving around the same idea.

According to Kenneth Laudon and Jane Laudon (2014), recently most of the data collected by the organizations consisted of transactional data that could easily be organized and related in management programs, however, lately a data explosion occurred due to the enormous web traffic through e-mail and social networking content. Still about the authors, Big Data does not refer to any specific amount of information, but rather data in the range of petabytes and exabytes, from billions to trillions of records, all from diverse sources.

According to Enomura (2014), defining the term "Big Data" is very complicated, and its first definition in a way known today was in 2001, with Doug Laney of the company Gartner Group, the famous Three V's (volume, speed and variety). They are databases whose capacity requires innovative forms of information processing for better perception and decision making.

Mayer-Schonberger and Cukier (2013, p.4) bring the following idea about the term:

There is no strict definition for the term. At first, the idea was that the volume of information had grown so much that the quantity examined no longer fit into the processing memory of computers, so the engineers had to refine the instruments they used for analysis. This is the source of new processing technologies, such as Google's MapReduce and its open-source equivalent, Hadoop, launched by Yahoo. They allow much more data to be managed than before, and the data - this is important - need not be allocated in rows or in the classic tables.

For further development and understanding of the work, below is a table containing other important Big Data concepts cited by leading authors in the field.

Table 1: Authors / Organizations and their respective Big Data concepts

<table>
<thead>
<tr>
<th>Author / Organization</th>
<th>Concept</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cezar Taurion (2013, p.26)</td>
<td>&quot;... unquestionably, we are talking about a very significant volume of data. But in addition to abyssal volumes, [...] there are variety of [...] data and speed. [...] I personally add two other variables that are: veracity of data and [...] value for the business.</td>
</tr>
<tr>
<td>McKinsey Global Institute (2011)</td>
<td>&quot;The intense use of online social networks, from mobile devices to the Internet connection, transactions and digital content and also the increasing use of cloud computing has generated untold amounts of data. The term Big Data refers to this data set whose growth is exponential and whose</td>
</tr>
</tbody>
</table>
Every day, we create 2.5 quintiles of data bytes - so much so that 90% of the data in today's world has been created in the last two years. These data come from everywhere: sensors used to gather information about the weather, messages to partner media sites, digital photos and videos, transaction records of purchases, and GPS signals from cell phones, to name a few. This data is Big Data.

Source: Adapted from Big Data, 2013; Big data: The next frontier for innovation, competition, and productivity, 2011; What is Big Data? 2016; Big Data: Impacts and Benefits, 2013.

Among the concepts, it is always mentioned about volume, variety and speed of data. In some cases, often more recent, there is also the value and veracity of the data. Therefore, it is also important to define and understand what they are and why they impact the Big Data concept.

According to Intel (2012), the Three V's characterize what Big Data is as a whole, but also define the main issues and needs to address, being:

Volume: The massive scale and growth of unstructured data surpass traditional storage and analytical solutions;

Variety: Big Data is collected from new sources that have not been mined in the past. Traditional data management processes can not handle the diversity and variation of Big Data data, which comes in formats as diverse as e-mail, social networks, videos, images, blogs, and data sensors; and

Speed: The data is generated in real time, with the requirements of useful information to be served.

For Taurion (2013), there are still two other variables to consider, which are:

Veracity: because we need to make sure the data makes sense and is authentic; and

Value: because it is absolutely necessary that the organization that implements Big Data projects obtain a return on these investments.

2.1.1 The impact of Big Data on companies

The application of Big Data to companies can have positive and negative impacts across industries. Therefore, a study is necessary so that the company has the necessary information about the system
before implementing it, avoiding to the maximum losses and negative results.

Big Data can impact current and future process models in a variety of ways. In addition to business impact, data aggregation can affect planning, use, assurance and privacy governance and management (ISACA®2013b):

- **Governance** - so that care is taken to define how the Big Data governance will be, and what data will be included or not.
- **Planning** - Consists of collection and organization of results, which help to justify adjustments and/or improvements of processes that can be identified using specific research techniques, develop negotiation programs following an orientation to certain conditions that trigger events, encourage standardization of purchases as a customer compares and researches services and products, uses location-based information along with other data to verify customer data (loyalty to product or company, routes, demands of different products and), helps in just-in-time (JIT) of inventory and inventory based on changes in demand according to specific times, assistance in managing real-time logistics operations.
- **Use** - The use of Big Data requires a good infrastructure. Smaller companies may take much longer to deploy this concept because of these.
- **Warranty** - After developing a good strategy to drive Big Data, the company should be concerned with setting up a guarantee structure, so that Big Data can be controlled and protected. Guarantee and data quality are the biggest concerns for companies.
- **Privacy** - special attention should be given to personal data and information, even when authorized by users, always avoiding exposure and protecting them from people of bad faith. In addition, each government has its own laws, prohibiting or legalizing certain actions, one must be aware of this as well.

### 2.1.2 How Big Data can help companies

At all times people, and even things, are generating data. With the advancement of technology, the volume of such data has increased significantly, in addition to the increased capacity to monitor them. In the same way, the technology provided tools capable of separating, organizing and reading this data, so that they could become important information.

Recent research indicates that the proper use of Big Data can play an economic role by [...]promoting innovation, competitiveness and productivity in all segments (LIMA, Cecília A., CALAZANS, Janaina H. C., 2013). In addition, McKinsey Institute (MANYIKA, James, CHUI, Michael, BROWN, Brad et al) says that using Big Data will become an essential basis for the competitiveness and growth of individual companies. From the point of view of competitiveness and value capture potential, all companies should take Big Data seriously.

ISACA® (2013b, page 07) takes the following view on the opportunities that Big Data can bring:

The Big Data opportunities are significant as well as the challenges. Companies that dominate the emerging Big Data management discipline can gain significant rewards and differentiate themselves from their competitors. In fact, a survey by Erik Brynjolfsson, an economist at the Sloan School of Management at the Massachusetts Institute of Technology, shows that companies using "data-based
decision making" get a five-to-six productivity. Proper use of Big Data goes beyond collecting and analyzing large amounts of data; it also requires understanding how and when to use data when making crucial decisions.


Big Data analysis can positively impact (ISACA®, 2013b) on product or market development, operational efficiency, customer experience and loyalty, as well as market demand forecasts. According to Taurion (2013) the use of Big Data by companies helps them improve several actions, previously unthinkable, such as:

- Optimize cross-selling or cross-selling. According to Taurion (2013), about 30% of Amazon's sales are driven by recommendations "you might also want" or "you might want to."
- Location-based marketing or location-based marketing. By identifying a customer's location, the company can send a specific message to it with some promotion. According to Taurion (2013), 79% of Starbucks customers receiving the personalized message become more likely to visit the store.
- Analysis of customer behavior in the store. About 80% of customer time within a store is spent circulating in the search for products and not effectively interacting or buying. Tracking and analyzing customer behavior in the store allows you to invest in stocks that take advantage of this wasted time.

Tessarolo and Magalhães (2015, p.33), give other examples of how Big Data helps Brazilian companies:

Brazil, telecommunications companies like Rede Globo, Record and SBT use Big Data to measure the popularity of their programs. They use social media services that search information on social networks like Facebook and Twitter if certain subject matter covered in any program was accepted by the target
viewers or not. In this way, producers and directors are able to make better decisions based on this information collected in the media.

The use of Big Data, as already mentioned, has a high capacity to help companies get the most efficiency out of all their processes. With this, the use of available resources and a total reduction of waste of the same, allows companies to have the lowest cost and consequently the highest profit possible.

2.1.3 Data collection and analysis

As previously mentioned, Big Data is about collecting data, organizing it, and analyzing it. Currently, the data is divided into structured and unstructured data. The information stored in the databases is known as structured data, because it is represented in a strict format. For example, each record in a relational database table. (ELMASRI, 2011, p. 416)

For Tessarolo and Magalhães (2015), when dealing with a huge amount of data, they come from different sources and manners. It is a great challenge to mine, clean, organize, correlate, link and transform that data into relevant information. For this to be possible, it is necessary to create a database where all this information is captured through applications, created for the purpose of guiding this data, storing them in an organized and clear way, thus facilitating the search and manipulation these data.

Within the internet network only a quarter of these data are stored in databases. Much information is created daily and less than 10% can be mined and organized, making every day huge amounts of information become junk mail, making it difficult to search and often becoming useless. (TESSAROLO AND MAGALHÃES, 2015. P. 02)

According to The Dama Guide (2012), unstructured data is any documents, files, graphics, images, texts, reports, forms or video or audio recordings that have not been encoded or otherwise structured into rows and columns or records. According to many estimates, about 90% of all stored data is kept out of relational databases. Unstructured data is presented as information stored in context. In fact, there is always some structure in which data provides information, and this structure may even be tabular in its presentation.

Already for Tessarolo and Magalhães (2012, page 03):

Unstructured data make up the network. Of all the data in the world that was generated in recent years only 10% of this data is structured. The remaining 90% are unstructured and gather largely on social networks like Facebook, Twitter, Pinterest, among others.

According to Tessarolo and Magalhães (2012), it is necessary to reorganize and mine these digital data so that all the information created daily is not lost and becomes useful. The solution to this, without a doubt, is Big Data.

2.1.4 Human Resources for Big Data

In spite of all the requirements of Software and Hardware, it is also necessary for people to know how to handle these tools so that companies obtain positive results. According to Taurion (2013), Big Data's
growth scenario points to the emergence of new job opportunities for IT professionals and other industries.

A new position, called a "data scientist" or data scientist is a good example. It usually requires training in Computer Science and Mathematics as well as the analytical skills needed to find the providential needle in the haystack of data collected by the company. "A data scientist is someone who is curious, who analyzes the data to detect trends," said Anjul Bhambhri, IBM vice president of Big Data products recently. (TAURION, 2013. P. 79)

The author summarizes that three basic profiles of professionals engaged in Big Data can be identified. The first is based on data scientists, who are professionals trained in statistics, computer science and / or mathematics capable of analyzing large volumes of data and extracting new business opportunities from them. The second is based on business analysts who can formulate the right questions, perform the analysis of the answers and from them, make strategic decisions of the company. The third and final one consists of technology professionals who will take care of the infrastructure and its technical support to support Big Data.

Still by Taurion (2013) Big Data's big challenge in coming years is to have skilled professionals, since technology is evolving fast and will not be deterrent. The bottleneck then lies in people, not technology. Besides the data scientist, there is also space for other professional activities. Strong demand also comes from developers and system administrators, for example, who specialize in Big Data-driven tools such as Hadoop, a technology designed for data-intensive distributed applications and used by well-known sites such as Facebook, Yahoo and eBay.

### 2.1.5 Risks and care in Big Data

Having defined the concepts of Big Data and its benefits, it is also important to know the risks and the care that must be taken when using this data. If something very powerful, it is important that limits be established, and that there be a certain ethics regarding the use of data. This is because among all these data may be personal information, thus jeopardizing the safety of the person in general.

According to ISACA® (2013b), while Big Data can provide a competitive advantage and other benefits, this also entails significant risks. Now that companies have huge amounts of structured and unstructured data available, management must be asked: Where to store this data? How to protect the data? and, how to use the data in a safe and legal way?

For ISACA® (2013a), the risk related to Big Data can be categorized as operational or based on information technology. These risk categories can be reduced with strong governance. Thus, operational risk encompasses internal and external factors, which include geopolitical risk and the race to satisfy senior management and management, who want to get ahead of the competition. The IT risk is also business, specifically, the business risk associated with the use, ownership, operation, involvement, influence and adoption of IT in a company. IT risk occurs when security assurances are overcome.

Security and privacy play an increasingly important role in Big Data, and all stakeholders should be aware of the implications of storing and cross-checking large amounts of confidential data. (ISACA®, 2013b). The author further mentions that it is necessary to understand that some data should be
considered toxic, in the sense that loss of control over them could be detrimental to the company, such as: Private or custody information, such as credit card number, information such as social security numbers and personal health information; Strategic information, such as intellectual property, business plans and product designs; and Data as key performance indicators, sales results, financial and production metrics used to make crucial decisions.

The need to manage data risk within the enterprise may not be clearly communicated and understood at all levels of management. It is essential to show that addressing the risks and concerns of Big Data can not be seen merely as an exercise in information technology. The participation of the entire company, including legal department, finance, compliance, internal audit and other business departments, allows everyone to focus on business objectives in the planning stage. Companies can stay focused on both technical and business aspects of Big Data. (ISACA®, 2013B. P. 08)

Using Big Data is essential to ensure a higher level of competitiveness for companies, but this powerful concept can open doors to an illegal use of it. In order to avoid risks to the safety of the companies themselves and especially of the customers, it is essential that some measures be taken as mentioned above, with the aim of only using their beneficial side

2.2 BUSINESS PROCESSES

The companies in general are composed of administrative processes that help in the development of the activities in an organized and efficient way. Some examples of processes are: marketing, sales, logistics and decision making. All these processes are, in a way, interconnected and should work in a synchronized way so that the company has better results.

2.2.1 Marketing

Marketing is a key part of the retail market. The company needs to be in the head of the customers, so that it is remembered when the customer decides to make a purchase. In addition, there are customers who do not know of the existence of the company, brand or product offered, plus one of the marketing roles make them known to customers.

According to Kotler (2012, p.33), an objective definition of marketing is to 'meet needs generating profit'. Marketing involves the identification and satisfaction of human and social needs. The author further adds that marketing applies to several things, such as: goods, services, events, experiences, people, places, properties, organizations, information and ideas.

The American Marketing Association proposes the following definition: Marketing is the activity, the set of knowledge and the processes of creating, communicating, delivering and exchanging offers that have value to consumers, customers, partners and society as a whole.7 Handle these processes requires a good deal of work and skill. Marketing management happens when at least one party in a potential trading relationship looks for ways to get the desired responses from the other parties. We see, therefore, marketing management as the art and science of selecting target markets and capturing, maintaining and retaining customers by creating, delivering and communicating a superior value to the customer. (KOTLER, 2012. P. 03)
The layout is also an important point within the marketing, as it aims to provide better use of space and provide a certain comfort to the customer at the time of purchase. Las Casas (1994. apud CAMARGO, TOALDO, NIÑO, 2009) defines layout as the essential parts or elements that help a store achieve maximum productivity.

In this way, Underhill (1999. apud CAMARGO, TOALDO, SOBRINHO, 2009, p.03) bring important points to be observed when deploying the layout:

- catch rate: refers to how much the consumer sees what is exposed. The reliable zone goes from a little above the eyes to the height of the knees. For the remaining space should be displayed large products of easy visualization (ex: diapers);
- boomerang fee: this is how many times the customer fails to fully cover a corridor. One way to minimize this effect is to position popular products in the middle of the aisle or place decoys at both ends;
- goods should not interrupt the line of sight of customers: they should be able to see not only what is in front of them but also in other parts of the store;
- take the consumer to the shop floor: for this position in this area, destination categories, so that those looking for these items will buy others in the round trip.

Marketing reaches all areas, be it goods, services or people. That way its importance is obvious. Understanding how the market works is essential to apply marketing in a way that positively targets all of the targets, impacting them directly and indirectly in ways that improve bottom line results.

2.2.2 Sales

It is convenient to emphasize that marketing and sales are different things, although a certain connection is characterized. Weitz et al. (2004. apud CASTRO and NEVES 2005, p. 8) defines personal selling as a personal communication process in which a seller identifies and satisfies a buyer's needs for the long-term benefit of both parties. Already for Rogers (1993. apud CASTRO and NEVES 2005, p. 8), personal selling is the sharp side of marketing as it is when company representatives come face to face with potential buyers. The sales force acts as a link between the company and the customers.

According to Castro and Neves (2005, p.16), the sales process is a sequence of steps or steps through which sellers make the sale. This sequence was originally proposed by Dubinsky in 1980. The steps recognized as formators of the sales process are seven, being:

- Prospecting: Identification of potential clients;
- Pre-approach: Gathering information so the seller gets ready for the approach;
- Approach: Ensure good impression, in addition to arousing the interest and attention of the client;
- Sales presentation: Presentation of offers and benefits;
- Dealing with objections and overcoming resistance: Overcoming buyer resistance through responses and emphasis on product benefits;
- Closing: Request of the request according to the client’s wishes, in the most appropriate and
- After-sales service: Track the delivery process and customer satisfaction, providing greater credibility and a possible future sale

Sellers need to understand the behavior of the market and the customers so that they can get in the right direction and give the customer what they need and want in the best way possible. The sales process is a
set of steps that sellers need to go through to make the sale, however it is important to remember that the process does not end in the sale of the product, since it also covers after-sales.

2.2.3 Logistics

Logistics is one of the processes within companies and organizations that demand a lot of attention. It is the material flow planning process, which aims at delivering the needs in the right time and in the desired quality, optimizing the resources and increasing the quality in the services. (BALLOU, 1999. apud PAURA, 2012, p. 13). According to PAURA (2012), logistics, besides helping companies, helps the local quality of life in terms of developing the infrastructure for its operation. This theme is currently vital to business as it optimizes resources and increases quality, which means increased results and lower spending.

The business logistics professional studies how to efficiently provide profitability (1) in customer distribution services, (2) in the material flow within the company, (3) in purchasing planning, through the control and organization of inventories of raw materials and finished products, (4) in production control planning, and (5) in packaging transport control. (PAURA, 2012. P. 23)

PAURA (2012) explains that the logistics for commerce is considered fundamental by several factors, among them there is the dependence that the commerce has of the fulfillment of the terms of the industries so that they can offer the products to the final clients. In addition, managing the logistics of the trade itself is essential not only to control what goes in and out, but also to give information about when and when to request more orders. When this is done correctly, often this trade has its products in better conditions and with fairer prices.

Based on the above information, it is understood the importance that logistics currently has in commerce in general. Having a good structure and necessary information, the control done has a better quality, which provides better competitiveness in this market so disputed.

2.2.4 Decision Making

Decision is a choice, a judgment that takes into account the whole context and environment around it and that can directly or indirectly interfere with the final outcome. In order to achieve the best future result, decisions must be carefully formulated. According to Pereira and Fonseca (1997, apud MORITZ and PEREIRA, 2006, p. 31), the decision is a systemic process, paradoxical and contextual, and can not be analyzed separately from the circumstances that involve it. Knowledge of the characteristics, paradoxes and challenges of society is essential to understanding decision-making processes.

According to Simon (1963, apud Moritz and Pereira, 2006: 33), the decision is a process of analysis and choice among several available alternatives of the course of action that the person should follow. He also points out six classic elements in decision making:

- The decision maker: it is the person who makes the choice between the several others available;
- The objectives: what the decision-maker intends to achieve with his actions and decisions taken;
- The preferences: are the criteria used as the basis for the decision maker to make their choices;
- Strategy: the path that the decision-maker adopts to achieve his objectives according to the
available resources;

- The situation: external factors that can not be controlled by the decision maker, and that impact on the events and results, affecting their choices; and
- The result: final consequence. Result of strategies and options

For Bethlem (1987, apud MORITZ and PEREIRA, 2006, p.34), it suggests a generic decision-making model composed of four stages:

Step 1 - Decision to decide - to take a behavior leading to any decision is a decision.

Step 2 - Once you decide to start the decision making process, the next step is the definition of what we are going to decide. There are times when we work on solving problems that we do not define, but statistically their numbers are less significant;

Step 3 - formulation of alternatives. The various possible solutions to solve the problem or crisis, or the alternatives that will allow us to seize the opportunities; and

Step 4 - choose the alternatives that we think are most appropriate. It is decision making.

Decision making is vital for the management of organizations, and it is understood that the administrative process is essentially decision-making. The decision-making process is complex and contains several steps that, even if they are not fulfilled within a rigid order, require a certain order so that the decisions made are efficient and rational, directly impacting the final results of the company or organization.

3. METHODOLOGY

Scientific methodology is the study of the path that must be followed towards a goal. It is a way of thinking in order to arrive at the nature of a certain problem, aiding its explanation or its study.

For Prodanov and Freitas (2013, p.14):

The Methodolog[y consists of studying, understanding and evaluating the various methods available for conducting an academic research. The Methodology, at an applied level, examines, describes and evaluates research methods and techniques that enable the collection and processing of information, in order to guide and solve problems and / or research questions.

According to Roesch (1999) the methodology aims to describe how the project will be carried out. In this step, it will be defined, from the final objectives, which type of project is most appropriate.

3.1 PLANNING OR RESEARCH DESIGN

The research can be understood as a process of knowledge production, or also as a learning procedure, and is thus part of the reconstructive process of knowledge (DeMO, 2000, p.20, apud PRODANOV and FREITAS, 2013, p.42). Its purpose is, through scientific procedures, to solve problems and to solve doubts (BARROS; LEHFELD, 2000, p.14, apud PRODANOV and FREITAS, 2013, p.42).
Based on the information that will be presented below, this research is delineated qualitatively and quantitatively. For Roesch (1999), any type of project can be approached from both perspectives. It can use a more quantitative approach in Results Evaluation and a more qualitative approach in Formative Evaluation (which focuses on the teaching-learning process).

According to Oppenheim (1993, page 21, apud ROESCH, 1999, p. 130), the quantitative delineation is indicated to find associations and explanation, less oriented to the description and more to the prediction. According to the author, if the purpose of the project is to obtain information about a given population, the option is to use a descriptive study. For Staw (1977), quantitative research is appropriate to evaluate changes in large organizations. Thus, the research has a quantitative character as regards the part where there is separation and interpretation of the quantity and frequency of measurable data, in order to form statistical information for reading them.

Regarding the qualitative research design, it is argued that its methods of data collection and analysis are appropriate for an exploratory phase of the research (STAW, 1977, apud ROESCH, 1999, p.154). For Roesch (1999) there is also another way of looking at qualitative research, where it can be considered as a different paradigm of research, as in the case of action research, where the researcher's position is based on capturing the perspective of the interviewees differently of a pre-established model. From this information, this research can also be considered qualitative because of the amount of non-measurable data that will be compared and studied.

According to Gil (2007, apud GERHARDT; SILVEIRA, 2009, p. 35), based on the objectives, it is possible to classify the research into three groups, being: exploratory, descriptive and explanatory. Still according to the author, the exploratory research aims to provide greater familiarity with the problem, with a view to making it more explicit. Most of them involve a bibliographical survey, interviews with people who have had practical experiences with the problem researched and analysis of examples to stimulate understanding. They can be classified as case study and bibliographic research. This research has an exploratory character, since it deals largely with a bibliographical research, aimed at highlighting the problem and making it better known.

### 3.2 DEFINITION OF TARGET AREA OR POPULATION

This paper aims to expose possible opportunities not utilized by the retail trade by failing to read and interpret data that is generated in the daily life of the company and all its stakeholders. However, due to certain requirements to carry out the application of Big Data, this work is aimed at medium and large retailers, where they have sufficient information and necessary infrastructure.

According to Roesch (1999) the internship can be concentrated in one department of the company, or encompass the whole organization. In the latter case, it is important to describe the structure of the area, including the number of people who work on it.

### 3.3 COLLECTION PLANS AND INSTRUMENTS

According to Roesch (1999), if the collection is of primary data, it is important to specify in this section the source of the data, when they will be collected and through which instruments. For secondary data, it is necessary to report its nature and specifications.
In the case of this work only secondary data will be used, since all the information will be collected in books, magazines, scientific articles and internet sites. All of this data will be information on what Big Data is, how it works and how it can help retailers. Also will be cited real examples of large companies that have had good results after the use of systems to make the analysis of this data.

3.4 DATA ANALYSIS PLAN

According to Roesch (1999), specifying the type of data collected delimits the possibilities of analysis. In addition, this is where the author can predict the use of graphs, tables and statistics. Thinking about the analysis helps criticize your own data collection, which can provide better quality at work.

As already mentioned, the research will be based on secondary data. Thus, an analysis of data and information that has been used, tested and in great majority have already proven results will be performed in this work. It will also make an analysis of the positive and negative results that large companies had after the implementation of Big Data. These results will be discussed through dissertation and will rely on the aid of images and graphs for a better understanding of the same.

4. DEVELOPMENT

This part of the paper aims to explain how Big Data can process important data about customers and thus create a great advantage for companies that improve their administrative processes based on information generated from the same data collected. It is important to remember that this work will emphasize the administrative part, and not in terms of the technical operation of the tools responsible for collecting, storing, analyzing and transforming data into information.

4.1 IMPLEMENTATION OF TOOLS FOR THE USE OF THE BIG DATE IN RETAIL AND MEDIUM-SIZED RETAIL ENTERPRISES

The company that wants to use Big Data should follow some steps. These steps consist of decisions and choices to deploy a basic structure involving tools, services and people capable of storing, organizing and analyzing this data. Considering that the company already has its own management system the steps are listed below:

- **Qualified Personnel Recruitment:** First, in order to implement all the tools and systems required to use Big Data Analytics, the company must rely on specialized help. This involves hiring three types of professionals who will be able to give all the necessary help for such, being: Professionals responsible for technical support and sufficient infrastructure to support Big Data; Professionals capable of analyzing large volumes of data and extracting new business opportunities, from their qualification in statistics, mathematics and / or computer science; and business analysts capable of asking the right questions and analyzing the answers obtained from the database.

- **Database Location:** The company needs a storage location for the clients. The option of renting cloud servers is recommended, and preferably in several locations, because in this way the risk of data loss is reduced in cases of some unexpected accident such as a fire or natural disaster. In addition, the structure of these companies is usually prepared for such, so that these servers have constant monitoring and specific security, both hardware and software. There is also the possibility of the company acquiring its own server, but it is very likely that the costs will be
much higher.

- API creation:[5] The process of transferring the data obtained by the company to the Database is not so simple. In order for the data to be passed you need to create an API, which will be responsible for communicating between the Store Software and the Database. In this way, the API will be responsible for picking up the desired data from the store's software, and transferring it in an organized way to the Database. The creation of API requires technical knowledge, and must be performed by a specialized team, whether it is a member of the company or outsourced.

- Flowchart

**4.2 COLLECTION AND DATA ANALYSIS FLOWCHART**

- Defining the right questions;

IT professionals along with the management and administrative part of the company should discuss what are the right questions to ask in order to get the answers in the structure they want.

- Definition of data of interest;

Data that is interesting to the company and can be used to bring results, should be well defined. The storage of data has its cost, so any additional data that is stored and that does not have utility means loss to the company. In an opposite scenario the company also loses because, failing to collect important data for it, the gains can be reduced or null. It is extremely important to identify which data should be collected and which data should be ignored.

- Collection of data generated by customers;

Once the customer makes a purchase in the physical store or virtual store, various data is being generated. All this data is collected by the company. Examples of these data are: date and time of purchase, customer name, purchase value, products, etc.

- Send data to the database through the API;

Sending the data collected by the company to the database is performed automatically and in real time through the API, which is responsible for this process.
Structured data storage in the database;

The data collected will be stored in an organized way in the Database as the IT staff have defined it as the best way, so that the data can then be processed and analyzed with greater ease.

Processing and transformation of data into information;

With data stored and organized on the server, processing and analysis of this data is performed. Data processing consists of a series of activities performed in an orderly fashion, forming an arrangement of information, from which data and/or information is collected so that it can be organized and analyzed resulting in the objective that the system or user intends to use.

Analysis and study of information;

At this point, all information generated through the processing of collected data will be analyzed. Here the responsible professionals will look for answers to the questions asked in the beginning. Individual customer profiles or certain groups, consumer trends, behavior patterns and purchase patterns are often found.

Decision making based on the analysis of the information;

After analyzing and finding answers to the initial questions, the company will be able to find ideas for possible adaptations and improvements in its processes based on this information. Behavioral and buying patterns, for example, can directly influence the marketing process, making it much more efficient, but maintaining the same investment. In summary, at this stage, decisions about how to behave in front of the market and customers should be made based on the analysis made on the information collected.

Modification application and improvements;

In this stage, companies will actually apply the process improvements defined in the previous step. These improvements must be applied in an organized and planned manner, so that you generate the best possible result.

Cycle of collection, analysis and application of data. Source: Data produced by the research.

4.3 IMPACTS OF THE USE OF BIG DATA IN PROCESSES OF RETAILER COMPANIES

The use of Big Data in companies, if done correctly, can generate great results for the same as a whole. This is due to the fact that Big Data can generate improvements in many individual processes of the company, that is, it does not impact only one process the other. Among these processes that can be impacted by Big Data, this work will be limited to presenting only marketing, sales, logistics and decision-making, even as they are processes that suffer a great direct impact.

4.3.1 Impacts of Big Data on Company Marketing
Marketing is essential for the company and for the customers, so it is important that the company care for that point. In addition to seeking to discover the needs and demands of a particular target market, it is the marketing that makes the presentation of a certain product or service to the customers in order to attract them and arouse the consumer interest of them. With the use of Big Data, the marketing of the company can become much more efficient.

With Big Data, companies can get a lot more information about their target audience's buying and behavioral habits, so they can get the marketing process done more objectively and efficiently. In addition, it is possible to discover other groups of customers, with different habits or even different uses for certain products. With all this information, marketing can be customized.

A good example is that, instead of the store carrying out generalized promotions, the promotions can be customized for each client. They can be sent by e-mails, pamphlets delivered at home, or even through ads on each client's social networks. In this way, the company, investing the same value in marketing, can obtain better results since this product, brand or service will reach people who have a greater predisposition to acquire them.

Mayer-Schönberger and Cukier (2013) exemplify in their book, an American network of stores called Target realized the consumption habits of pregnant women in the first months of gestation, where consumption of moisturizing or similar products was very high. From there, the store began sending personalized promotional flyers and letters to these customers, where they presented product promotions for pregnant women and babies. In the end, Target discovered that a young woman was pregnant even before her father. Target, with this information, created a specific marketing plan for these people, because I was able to make projections of when the child would actually be born, and close to that date, offered promotional products like cribs, diapers, and baby supplies.

Recognizing these standards is key to Big Data, because with them it is possible to have a sense of customer behavior, and how to use it to make the company's processes more efficient. Overall, Big Data can help marketers in this way by providing information that directs where, when, and for whom certain promotions or information about products or services should be sent. It's worth remembering that Big Data often will not show why it works that way in specific ways, but in numbers it can prove that it really works.

4.3.2 Impacts of Big Data on Sales and Logistics

The sale is the main objective for companies, it is one of the main reasons to carry out this whole process of deployment and use of Big Data. Although all processes are interconnected, logistics has a greater relevance when it comes to sales, since to carry out the sale of the product or service, it is necessary for the company to be able to provide it to the customer. This involves not only availability, but also speed and quality of delivery, the latter applies mainly to orders that are sent by delivery services, different from when the customer takes it personally.

Just as buying and behavior patterns help Marketing, so do sales and logistics. Knowing how the customer will behave in certain situations, the company can anticipate, and organize the environment in a way that the client consumes more and that leaves him satisfied in the same way. Organizing the
environment is not just a matter of layout or exposure of products, but also ensuring that the product that
the customer seeks is available, and that the store is not surprised in certain situations, ceasing sales and
causing customer discontent that Did not find what you were looking for. One of the interesting examples
cited by Mayer-Schönberger and Cukier (2013, p.37) is that of Walmart:

In 2004, Walmart scoured the enormity of data from past transactions: what item each consumer bought
and the total cost, what else there was in their basket, schedule, and even the weather. In doing so, the
company noted that, prior to a hurricane warning, not only the sale of lanterns was increasing but also that
of Pop-Tarts, a typical American sweet. Thus, as a storm approached, Walmart stocked the Pop-Tarts
boxes at the front of the stores, near hurricane supplies, to make it easier for incoming and outgoing
consumers to live-and thus sales.

In this example, it may be noted that they already had knowledge about the increased consumption of
flashlights when hurricane forecasts were available. In addition, they were able to read that there was also
an increase in demand for a typical American candy called Pop-Tarts. That way, when there was a
hurricane warning in a city or state, Walmart was able to get ready so that these products could be found
easily within the store, and also ensure that they had enough products to meet the demand. It is worth
mentioning that in the livor, Mayer-Schönberger and Cukier report that they were not sure about the
reason for the increase of the sale of the sweet Pop-Tarts, but it worked.

Mayer-Schönberger and Cukier (2013, p.37) give another example of Walmart, which proves that Big Data
has already shown its power since the 1990s.

"In the 1990s, it[Walmart] revolutionized retailing by registering all products as data, through a system
called Retail Link, which led vendors to monitor rate and sales volume and inventory. The creation of this
transparency allowed the company to force suppliers to take care of the supply themselves. In many
cases, Walmart does not "appropriate" a product until a moment before the sale, thus reducing the risk
and cost of inventory. Walmart used data to become, in fact, the largest consigned goods store in the
world."

These examples show how Big Data has been able to help a large retailer so that it can reduce its costs,
increase its sales and consequently be more competitive in the market. The use of Big Data has made
these processes a great positive impact, increases its efficiency, and consequently impacting the overall
result of the store.

4.3.3 Impacts of Big Data on Decision Making

Decision making is a critical time in business. This is the moment when the company must decide one
way among many others. The weight of the decision is high, because in terms of the company's image and
even financial, it can mean gain or loss. Therefore, it is interesting that the company has as much
information as possible, whether it is about customers, internal or external environment. The more
information, the greater the ability to make a successful decision.

This is how Big Data helps companies make decisions by providing as much useful information as
possible. Information that can be used to get a preview of how the market may behave in the near future. Things are in fact uncertain, nothing can be guaranteed, but it is so for all, and those who have more information are ahead. By cross-referencing customer data, how they behave, what their habits are, along with other information that can be acquired about geolocation, or even political and economic situations, companies can have a good idea of how to behave in this market that suffers so many variations.

In a market like the current one, which presents extremely rapid changes, the process of collecting the data until its transformation into decisions can not be slow. As the market changes very fast, business responses should be as fast as. It is no use taking the right decision if it is held late. In this way, Big Data is definitely a great alternative, because with it, you can analyze an immense amount of data, and make it generate value for the company as a whole.

Final considerations

This work aimed to show the importance of using Big Data in the current retail market. This importance comes from the impact that the use of this enormous amount of data causes in the processes of the company and in the knowledge of the behavior of its clients and of the market. A few years ago the importance of this technology is recognized, and it is becoming more and more essential for business.

According to the development of the research, it can be observed that deploying the necessary structure in the companies to carry out the reading and interpretation of Big Data is not so simple. It is necessary to hire people who have knowledge in the area to give due support. The investment in equipment is also not low, since it requires an immense capacity of data processing to read and to analyze this great amount of data. Therefore, it is necessary a feasibility study of the implementation of the system, because although the return is high, a high investment is also necessary, which for small and medium enterprises may be impracticable.

The use of Big Data allows companies to analyze data beyond the physical space of their stores or their own websites and social networks. Any data that can be collected and that can generate any kind of information about the user, are collected constantly, allowing companies to follow the behaviors, habits and ways of thinking of people in general, whether they are consumers of their products or not. That way, the level of preparation that the company can have for certain situations is extremely high, well beyond what it once was.

Examples of processes that may be more efficient are presented and commented in the research development, among them are: better application of the capital invested in marketing, better management of the logistics of the products, better exposure of the products and organization of the stores. All this impacts positively, reducing waste and consequently generating a profit increase for companies.

As data is generated faster and in greater quantity, there is always more data to be analyzed. It is interesting that the data analysis and interpretation system is a cycle. In this way, after the company applies changes and improvements based on the data that has already been analyzed, it will collect data generated in response to these applications, improving processes more and more as changes in the behavior of markets and customers. From this line of reasoning was created the model of implementation and application of Big Data, which was presented in the work. Finally, it can be concluded that the objectives of the work were achieved
Final Suggestions

The present study was limited to producing information about Big Data and its functionality in organizations, without, however, presenting its application and practical deployment in the routine of an organization. Thus, it is suggested that before carrying out the application of the model in any company, the practical test of the same, in order to prove in practice the efficiency of the presented model.

It is also suggested that the work be continued so that the practical part of this model is tested, which would allow to identify its real functionality, importance and added value for the retail companies.

References


[1] Academic of the Administration course of CEULJI / ULBRA

[2] Ms. UFSC, professor of the Administration course at CEULJI / ULBRA


[5] API is a set of programming routines and standards for access to a software application or Web-based platform. The following API refers to the English term "Application Programming Interface" which stands for "Application Programming Interface“ in Portuguese.