International Journal of Data Science and Analysis
2020; 6(6): 183-203
http://www.sciencepublishinggroup.com/j/ijdsa
doi: 10.11648/j.ijdsa.20200606.13
ISSN: 2575-1883 (Print); ISSN: 2575-1891 (Online)

Case Report

Developing and Implementing Big Data Analytics in Marketing

Dina Darwish

Computing and Digital Technology School, ESLSCA University, Giza, Egypt

Email address:
dina.g.darwish@gmail.com

To cite this article:
Dina Darwish. Developing and Implementing Big Data Analytics in Marketing. *International Journal of Data Science and Analysis.* Vol. 6, No. 6, 2017, pp. 183-203. doi: 10.11648/j.ijdsa.20200606.13

Received: October 30, 2020; **Accepted**: November 11, 2020; **Published**: November 19, 2020

Abstract: Big Data represents the greatest game-changing chance and change in outlook for marketing since the creation of the telephone or the Web going standard. Big Data alludes to the ever-expanding volume, velocity, variety, variability and multifaceted nature of data. Big Data is the key result of the new promoting scene, conceived from the computerized world we currently live in for marketing associations. The expression "big data" doesn't simply allude to the information itself; it additionally alludes to the difficulties, capacities and skills related with putting away and examining such gigantic data sets to help a degree of decision-making that is more precise and timely than anything recently endeavored. Because of the many benefits of big data, the big data applications have appeared, and they can play important roles especially in making companies take informative business decisions in different fields, such as, healthcare, banking, manufacturing, media and entertainment, education and transportation and many others. This paper concentrates on the importance of Big Data Analytics nowadays, especially in the marketing process inside companies, as well as challenges and obstacles facing Big Data analytics, and a case study of a bank wanting to market a new financial tool to its customers is studied using R tool.

Keywords: Big Data, Analytics, Marketing

1. Introduction

We are now in the age of big data, and there is a rising demand for instruments that can process and analyze it. Big data analytics deals with the extraction of useful data from complicated data that conventional data mining techniques do not manage. The term "big data" refers to any set of such data and resources for data processing. The first academic paper on 'Big Data' was written by Diebold [1] in 2000. Social networking sites, e-commerce websites, sensors (smart devices / IoT), etc., are the big data sources. 3 V's [2] describe Big Data: Volume, Variety, and Velocity. There is no fixed size to identify a dataset as big data or not, but the dimension of 'volume' refers to a dataset that is large enough to go beyond the method. While conventional data analytics are focused on periodic data processing, in real-time or near real-time, big data is processed and analyzed. The third dimension of 'velocity' has also been included, therefore. In addition to these 3 V's, two more V's have been added, value and veracity. As we can understand from the example of data generated on Twitter consisting of abbreviations, typos and colloquial expression, 'Veracity' refers to the lack of consistency and precision. The 'value' factor was introduced because big data must provide the company with useful knowledge after being analyzed that can be used to make critical business decisions, policies and strategies. It is not possible to scale up standard data processing tools (e.g. Excel, SPSS, etc.) to the size of increasing datasets. For example, analysts can't perform analysis on more than 1 million rows with Microsoft Excel 2007.

To deal with the increasing datasets, a tool that can scale up should therefore be used. Social networking sites such as Facebook, Twitter, etc. produce, at very high speeds, large amounts of useful social data every day. Such data must be analyzed in real-time to forecast the outcomes of elections, stock market actions, etc. To do this, we need instruments that can perform streaming data analysis. MPP (Massively Parallel Processing) relational databases such as Greenplum, Vertica, etc. have the ability to store and handle petabytes of data, where the data is partitioned across several nodes with each node requiring processors / memory to process data. MPP databases, however, have an upper limit on storage space as well as the same restrictions on data processing as SQL. Semi-structured data is data stored in a form other than tables (e.g. XML, JSON, etc.), and the NoSQL data store (NoSQL stands for Not Only SQL, implying that SQL-like query languages may also be supported) is a database management system that offers a framework for storing and retrieving such data. Cassandra, HBase, MongoDB, and so on are examples of NoSQL data stores. NoSQL data stores do not have a set static schema for data to fit in, such as RDBMS, but they can accommodate diverse data from various sources. BigTable [3] is a distributed storage system designed to handle massive semi-structured data volumes, or it can be said that BigTable is a distributed NoSQL data system.

While the amount of big data appears to attract the most attention, a more acceptable description of big data is usually given by the variety and velocity of the data. (Big Data is often defined as having 3 Vs: volume, variety, and velocity.) Big Data cannot be efficiently analyzed using only conventional databases or methods due to its size or structure. Big Data concerns require new techniques and technology for the storage, management and realization of business advantages. These new tools and technologies allow large datasets and the storage environments that house them to be created, manipulated, and managed. The McKinsey Global Report from 2011 offers another concept of Big Data: Big Data is data whose size, distribution, variety, and/or timeliness require the use of new technological architectures and analytics to allow insights that unlock new business value sources. Big Data: McKinsey & Co.; McKinsey 's concept of Big Data suggests that companies would need new data structures and analytical sandboxes, new software, new analytical methods, and the incorporation of multiple skills into the data scientist's new position. Several origins of the Big Data deluge are illustrated in Figure 1 [4].

The Analytics Practice for Current business challenges provide companies with many opportunities to become more analytical and data oriented, as seen in Table 1 [4]. Table 1 describes four types of popular business challenges faced by companies where they have the potential to exploit advanced analytics to generate competitive advantage. Organizations should apply advanced analytical methods to optimize processes and extract more value from these common activities, rather than just doing regular reporting on these fields. For several years, companies have been attempting to minimize customer turnover, boost revenue, and cross-sell clients. What is new is the chance to combine modern computational methods with Big Data in order to deliver more successful studies for these conventional business problems.

For decades, several enforcement and regulatory laws have been in place, but every year new provisions are introduced, which reflect increased operational complexity and data requirements. Anti-money laundering and fraud prevention laws require sophisticated analytical methods and techniques for proper enforcement and management. To approach them properly, the four business drivers shown in Table 1 involve a variety of analytical techniques.

While much is written about analytics in general, it is important to differentiate between BI and Data Science. There are many ways of comparing these classes of analytical methods, as seen in Figure 2 [4]. The time horizon and the kind of methodological methods being used are one way to determine the kind of research being conducted. BI appears to include documentation, dashboards, and queries on business questions for the current period or in the past. BI systems make it easy to answer questions about quarter-to-date sales, progress towards quarterly objectives, and understand how much of a given product was sold in the previous quarter or year. These questions appear to be closed-ended and describe present or past actions, generally by somehow aggregating and grouping historical data. BI usually addresses questions related to "when" and "where" incidents took place. By comparison, in a more forward-looking, exploratory way, Data Science tends to focus on analyzing the present and allowing informed decisions about the future. A team may use Data Science techniques such as time series analysis, Advanced Analytical Theory and Methods: Time Series Analysis, instead of aggregating historical data to look at how many of a given product sold in the previous quarter, to forecast future product sales and revenue more accurately than to deal with more open-ended questions. Furthermore, Data Science appears to be more exploratory in nature and can use optimization scenarios to answer more open-ended issues.

This approach offers insight into current behavior and foresight into future events, while concentrating broadly on topics related to "how" and "why" events occur. Where BI issues tend to involve highly structured information arranged for accurate reporting in rows and columns, Data Science initiatives tend to use many kinds of data sources, including large or unconventional datasets. Depending on the objectives of an entity, if it reports and produces dash, it can choose to embark on a BI project. Current Analytical Architecture As mentioned above, Data Science projects need workspaces with scalable and agile data structures that are purpose-built to experiment with data. Most companies do have data warehouses that provide outstanding support for conventional reporting and basic tasks for data processing, but sadly have a tougher time supporting more rigorous analyses. A typical data architecture and some of the challenges it poses to data scientists and others attempting to do advanced analysis are shown in Figure 3 [4]. The data flow to the Data Scientist is shown, and how this person integrates into the process of collecting data to evaluate projects.

Big information is making important new opportunities for organizations to derive new price and make competitive advantage from their most worthy asset: info. For businesses, massive information helps drive potency, quality, and customized product and services, resulting in improved levels of client satisfaction and profit. For scientific efforts, massive information analytics modify new avenues of investigation with doubtless richer results and deeper insights. In several cases, massive information analytics integrate structured and unstructured information with real-time feeds and queries, gap new ways to innovation and insight. This paper provides information about the importance of big data analytics for business managers and information analytics for a good decision making, and also, a use case of bank customers is illustrated. The rest of the paper is divided as follows; section 2 represents importance of Big Data analytics. Section 3 represents Big Data Analytics implementation in companies, section 4 importance of Big Data analytics in marketing, section 5 represents opportunities and challenges for Big Data, section 6 concentrates on a case study of Big Data analytics in marketing using R, and finally, comes the conclusion.

Table 1. *Advanced Business Drivers.*

Business Driver	Examples
Optimize business operations	Sales, pricing, profitability, efficiency
Identify business risk	Customer churn, fraud, default
Predict new business opportunities	Upsell, cross-sell, best new customer prospects
Comply with laws or regulatory requirements	Anti-money laundering, Fair Lending, Basel II-III, Sarbanes-Oxley (SOX)

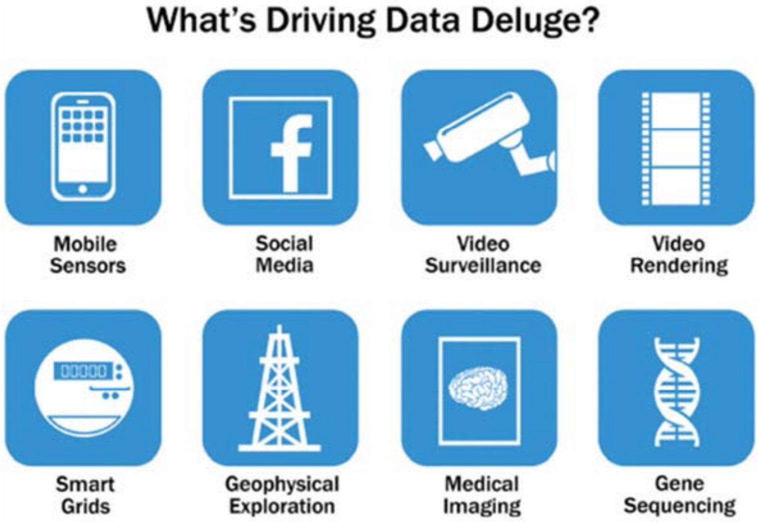

Figure 1. *What's driving the data deluge.*

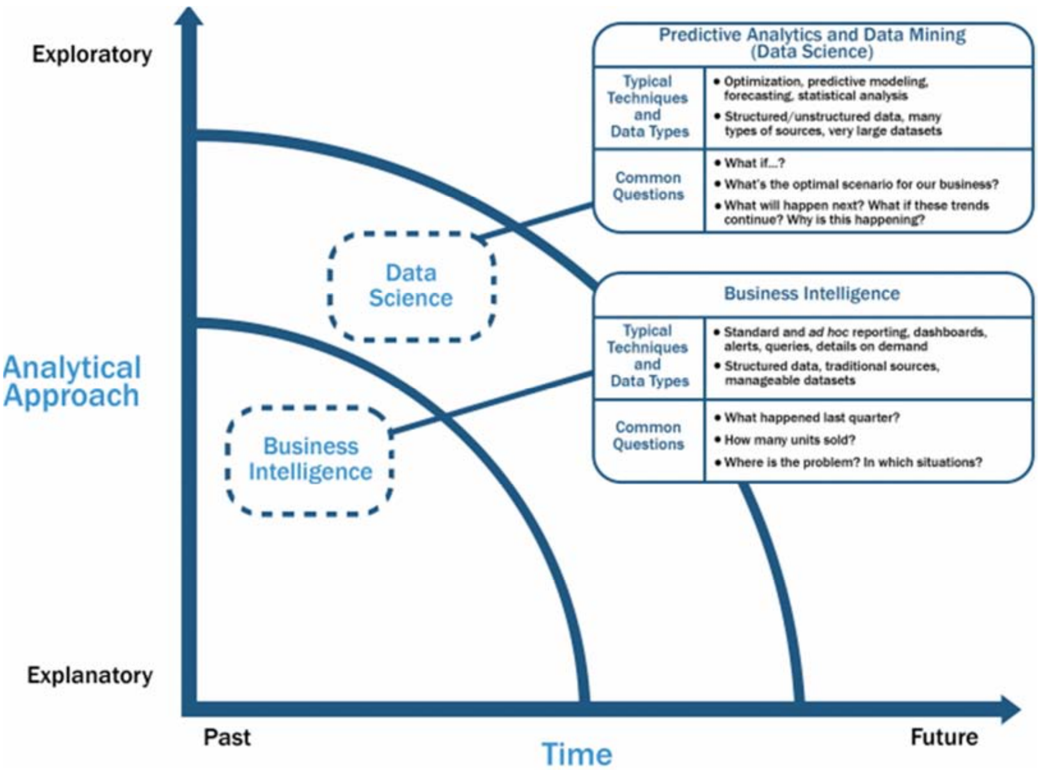

Figure 2. *Comparing BI with Data Science.*

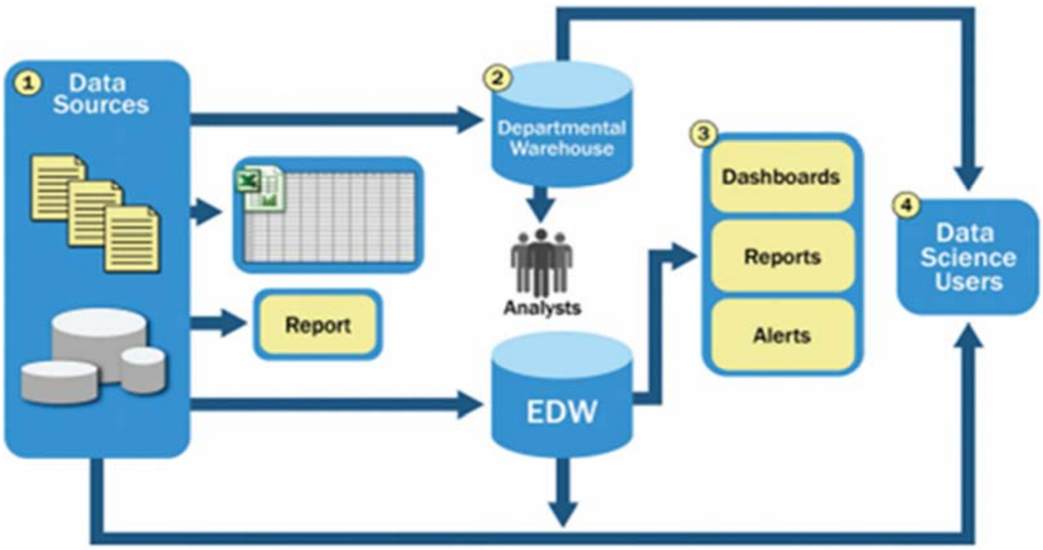

Figure 3. *Typical Analytical Architecture.*

2. Importance of Big Data Analytics

Big Data is continuously, and at an ever-pace, being generated. To establish a medical diagnosis, cell phones, social media, imaging technology, all these and more produce new data, and that must be processed for some reason somewhere. Diagnostic information that needs to be stored and processed in real time is automatically created by devices and sensors. It is difficult to simply keep up with this enormous data flow, but it is significantly more difficult to analyze vast quantities of it, particularly when it does not adhere to conventional conceptions of data structure, to recognize meaningful trends and to extract useful information. The ability to change industry, government, research, and daily life poses these challenges for the Big Data.

Several industries have led the way in improving their capacity to capture and manipulate information; Credit card companies track every payment made by their customers and can use rules derived by analyzing billions of transactions to detect fraudulent purchases with a high degree of accuracy. Mobile phone providers examine the calling habits of customers to determine, for instance, whether the regular contacts of a caller are on a competing network. The cell phone provider should proactively give the subscriber an opportunity to stay in its contract if the rival network offers an enticing promotion that might lead the subscriber to defect. Data itself is their primary commodity for companies such as LinkedIn and Facebook. These companies' valuations are heavily derived from the information they collect and host, which, as the information increases, includes more and more intrinsic value.

Three characteristics stand out as defining Big Data features:

1. Massive data volume: Big Data can be billions of rows and millions of columns rather than thousands or millions of rows.
2. Complexity of forms and systems of data: Big Data represents the range of emerging data sources, formats, and systems, including digital traces left for subsequent review on the web and other digital repositories.
3. Speed of new data production and growth: high velocity data can be represented by Big Data, with rapid data ingestion and near real time analysis. From many new sources, the Big Data trend is producing an immense amount of knowledge. In order to take advantage of these opportunities, this data deluge requires advanced analytics and new industry players, and new business trends.

Evolving Big Data Ecosystem require a new analytics approach for organizations and data collectors to understand that the data they can collect from individuals contains intrinsic value, and a new economy is evolving as a result.

The market is seeing the emergence of data vendors and data cleaners that use crowdsourcing (such as Mechanical Turk and GalaxyZoo) to test the effects of machine learning techniques as this modern digital economy continues to grow. By repackaging open source software in a simplified way, and taking the software to market, other vendors provide added value. For the open source platform Hadoop, vendors such as Cloudera, Hortonworks, and Pivotal have provided this value-add. In this interconnected network, there are four major groups of players as the new environment takes shape. These are shown in Figure 4 [4]. Data instruments and the "Sensornet" capture data from various locations and produce new information about this data on an ongoing basis. For each gigabyte of new data generated, an additional petabyte of data is produced about that data. For instance, consider playing an online video game through a PC, game console or smart device.

The types of data and the associated market dynamics differ greatly, as demonstrated by this evolving Big Data ecosystem. Sensor data, text, organized databases, and social media can be included in these databases. With this in mind, it is worth noting that in conventional EDWs (enterprise data warehouses), which were architected to streamline and centrally manage reporting and dashboards, these datasets would not work well. Instead, to succeed, big data challenges, initiatives need distinct approaches. To get the data they need inside an analytical sandbox, analysts need to partner with IT and DBAs (database administrators). Raw data, aggregated data, and data with various kinds of structure are included in a standard analytical sandbox. The sandbox allows for rigorous data exploration, and needs a savvy user in the sandbox environment to exploit and take advantage of data.

Why data analytics efforts bog down before they get big as recently as two or three years ago, data analytics leaders' main challenges were getting their senior teams to recognize their potential, finding enough talent to create models, and creating the right data fabric to connect within and outside the company databases together. But fresh challenges have arisen as these practitioners have pressed for size. In several ways, many senior executives concentrated on open-ended attempts to learn new knowledge from big data. Analytics vendors and data scientists who were willing to take data and run all sorts of tests in the hope of discovering diamonds, which fueled these efforts. Many managers have heard the claim "just give us your information". Figure 5 [5] represents how to get big impact from Big Data Analytics for companies to benefit from it.

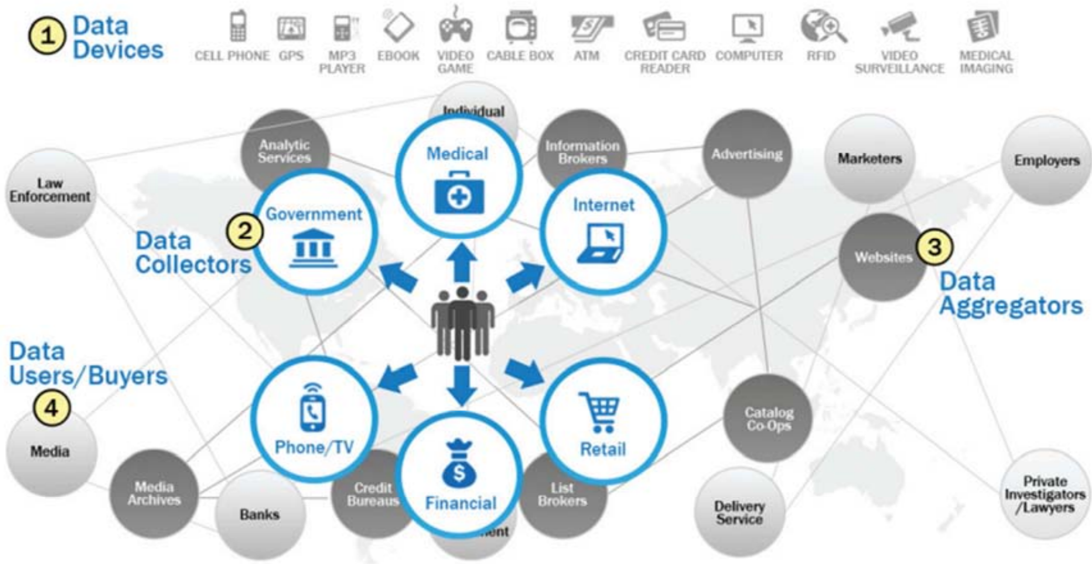

Figure 4. Emerging Big Data Ecosystem.

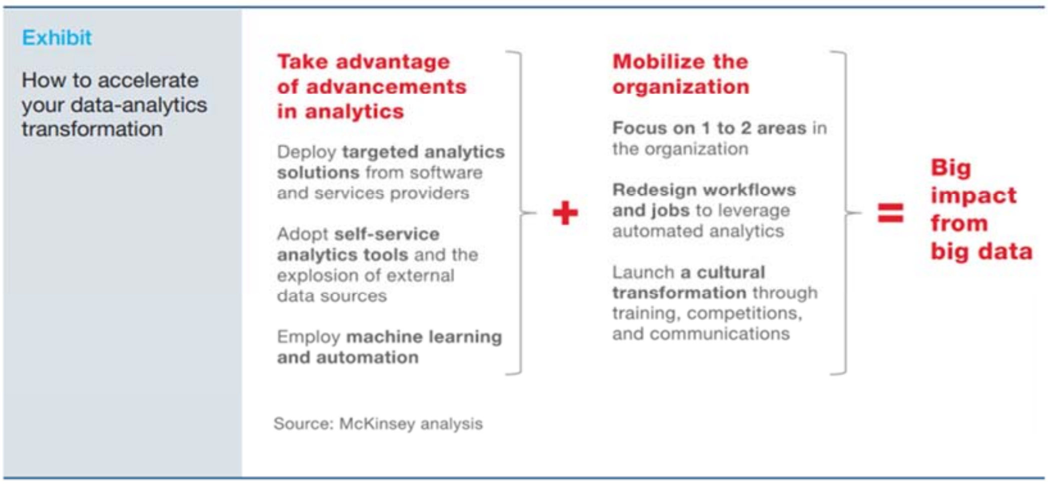

Figure 5. Big Impact from Big Data.

3. Big Data Analytics Implementation in Companies

Big data exploded onto the scene in the first decade of the 21st century, and the first organizations to adopt it were online and start-up companies. Firms like Google, eBay, LinkedIn, and Facebook were probably designed around big data from the beginning. They didn't have to reconcile or incorporate big data with more conventional data sources, and analyses performed on them, since they didn't have such traditional forms. They did not have to combine big data technology with their conventional IT infrastructures, because they did not exist. Big data could stand alone, big data analytics could be the only focus of analysis, and big data technology architectures could be the only architecture. Consider, however, the status of big, well-established enterprises. Big data in these environments should not be isolated, but must be integrated with everything else that's going on in the business. Big data analysis must coexist with other forms of data analysis. The Hadoop clusters have to operate alongside the IBM mainframes.

Data scientists must somehow get along and work together with mere quantitative analysts. To understand this coexistence [6], some 20 large organizations were interviewed in 2013 about how big data blends into their global data and analytics environments. Overall, the anticipated co-existence is noticed; in not one of these large organizations, big data was handled separately from other forms of data and analytics. In fact, integration has led to a new approach to analytics management, which we will call "Analytics 3.0."

Some managers appreciate the creative value of big data, but they find it more "business as usual" or part of a constant progression towards more data. They have been introducing new types of data to their systems and models for many years, and they see nothing groundbreaking about big data. Put another way, a lot of people were chasing big data before big data was big.

Since the three aspects of big data are: the lack of structure, the possibilities raised and the low cost of the technologies involved. This is consistent with the findings of a survey by New Vantage Partners of more than 50 major corporations in 2012 [6]. It was found, according to the survey summary: It's About Variety, not Volume: the survey shows that businesses concentrate on the variety of data, not on its volume, both today and in three years. The most important purpose and possible benefit of Big Data projects is the opportunity to evaluate different data sources and emerging data forms and not to handle very large data sets.

Companies that have long managed large data volumes are starting to enthuse themselves about the potential to manage a new form of data, voice or text or log files or images or video. For example, a retail bank is getting a handle on its multi-channel customer interactions for the first time by reviewing log files. The hotel business analyses the customer lines with video analytics.

In the 2017 Big Data Analytics Market [7], analysis is focused on a cross-section of data covering geographical, functional, organizational and vertical industries. It assumed that, unlike other industry analysis, this supports a more representative sample and a better predictor of real market dynamics. We have developed cross-tab analyses using these demographics to define and explain significant developments in the industry. Geography North America, which includes the United States, Canada and Puerto Rico, constitutes 66% of the respondents, EMEA accounts for the second largest party (24%) followed by Asia Pacific and Latin America.

Functions IT (36 per cent), Business Intelligence Competency Center (BICC) (18 per cent) and Executive Management (12 per cent) are the main categories of functions in our Big Data Analytics study [7] (Figure 6). Examining patterns and actions by role lets us compare and compare plans and goals in various areas of the organization.

In Vertical Industries, Technology (14 per cent), Healthcare (12 per cent) and Financial Services (12 per cent) are the industries most represented in research [7], led by Telecommunications, Education and Consulting (Figure 7). Responses from consultants, who also engage more closely with projects, and expand business awareness than many consumer peers do.

In Big Data Analytics, value of Big Data among technologies and initiatives considered to be strategic for business intelligence, Big Data Analytics ranks 20 out of the 33 topical areas are currently studied (Figure 8) [7]. It should be noticed that 2016 was a landmark year for increased adoption and perceived value of big data.

Although we continue to assume that the interest in big data differs widely from organization to organization, a wider trend has clearly emerged over the last 24 months. Contextually, we also note that the perceived value of conventional BI activities, such as monitoring, dashboards, and end-user self-service, is still a long way from big data. Figure 9 shows technologies and initiatives to business intelligence, such as, Innovations and Strategies, Strategic Business Intelligence, Video Analytics, Edge Computing, Internet of Things (IoT), Complex Event Processing (CEP), Open Source Applications, Social Network Monitoring (Social BI), Natural Language Analytics, Text analytics, Cognitive BI (e.g., Artificial Intelligence), Prepackaged vertical / functional analytical, Streaming data analysis, Ability to write transactional applications, Location intelligence / analytics Big Data (e.g., Hadoop), In-memory analysis, Cloud computing, Search-based framework, Data catalogue, Collaboration support for group-based.

For the first time in 2017 [7], existing users of big data (which we describe as "systems that enable end-users to access and analyze data contained and managed within the Hadoop ecosystem") have exceeded 50% (Figure 9). A further 36 percent of respondents say they may use big data in the future. Just 11 percent of respondents have "no plans to use big data at all" (an all-time low). Unexpectedly, larger companies (those with more than 5,000 employees) are embracing big data technology such as Hadoop much quicker than smaller companies. You would assume that smaller, younger businesses would be more nimble and able to adopt new technology, but when it comes to big data, the opposite is the case. It have been found [8] more than 300 major corporations that have made serious investments in Hadoop. On the other hand, there are only 300 other businesses with 5,000 or less workers that are mature Hadoop consumers. Since there are ten times as many small companies as there are, this means that in smaller companies, Hadoop has less than one-tenth of the reach it does in a large company. Many of the smaller

Hadoop companies are high-tech, data-driven companies themselves. But, they are lagging behind perhaps because they can't afford Hadoop and related technology, or it's because they can't pay the high wages ordered by data scientists and data engineers, Or perhaps they just don’t have as much data. Figure 10 [9] shows adoption of Big Data tools.

Data Scientists Need to Explore a large number of use cases for Hadoop and Spark in and around data science. Figure 11 shows the money spent on various Big Data cases [8]. A fast search [8] for "Data Scientists" on LinkedIn reveals that there are 16,000 workers in the US. What is surprising is that there are a huge 5,000 open opportunities for people with data science expertise. You will understand the challenging disparity if you equate it to clinical scientists in pharmaceutical companies. There are 3,200 medicinal scientists in the U. S., according to LinkedIn, but there are just 200 work openings for that position around American companies. Figure 12 [8] shows the number of data scientists versus the number of job openings for data scientists, especially in the pharmaceutical industry.

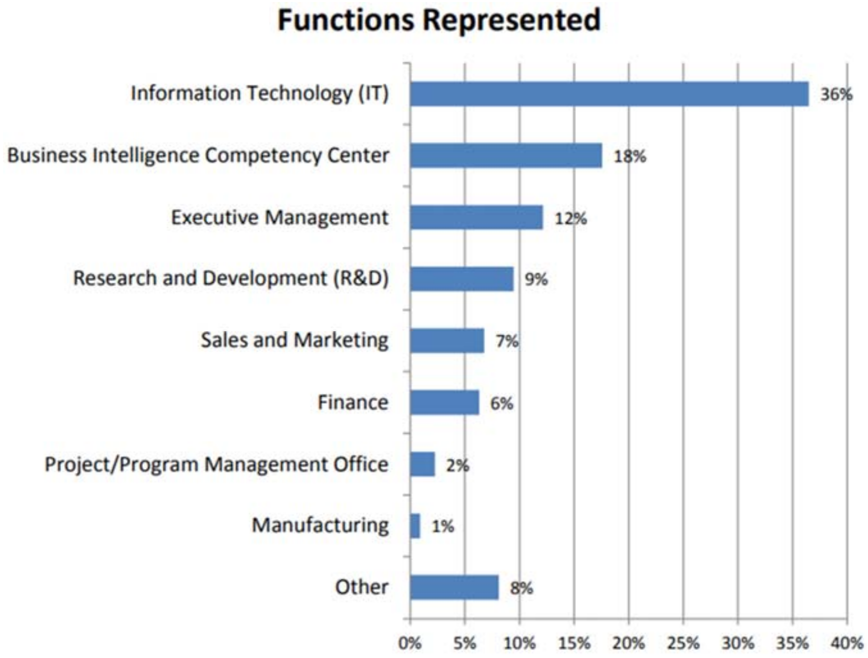

Figure 6. *Main categories of functions in Big Data Analytics.*

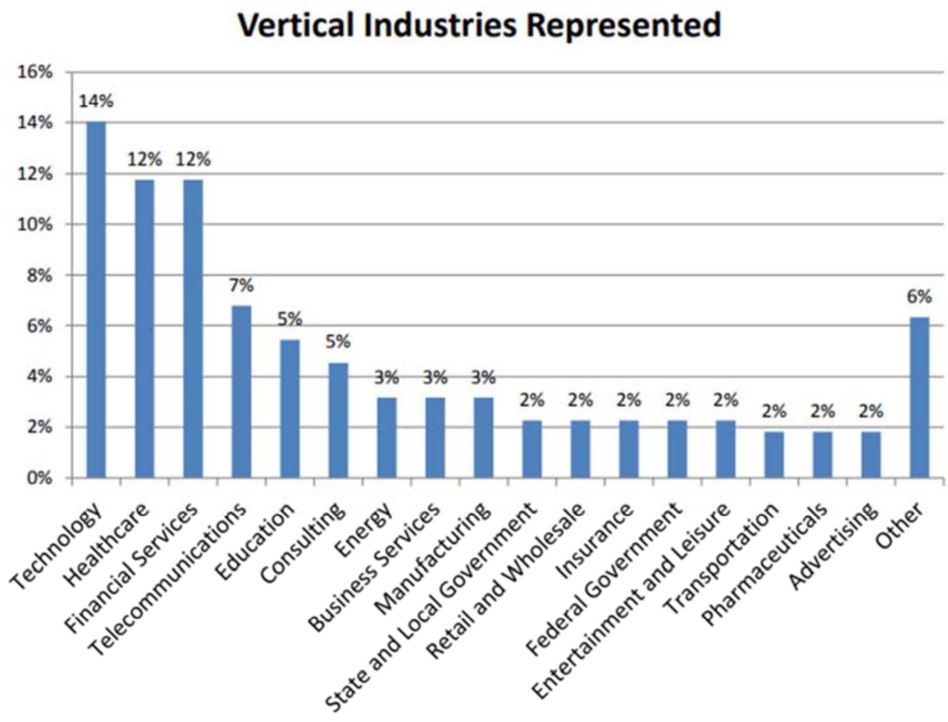

Figure 7. *Industries represented using Big Data Analytics.*

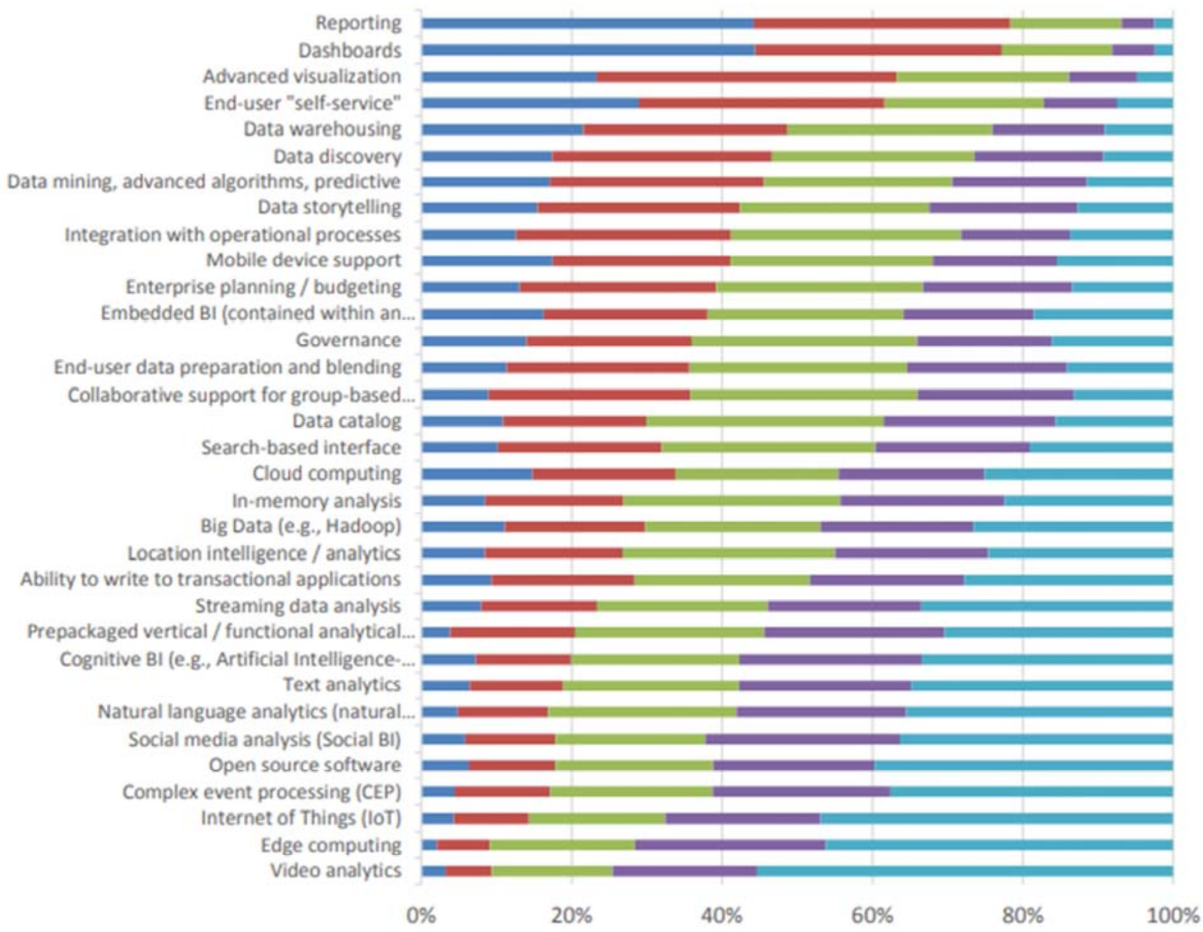

Figure 8. *Technologies and initiatives strategic to business intelligence.*

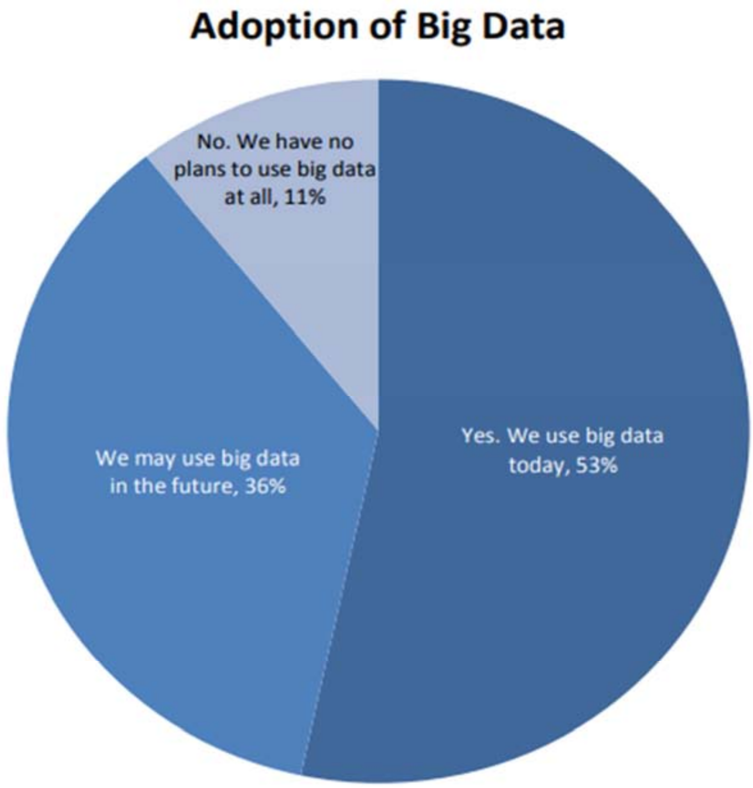

Figure 9. *Adoption of Big Data.*

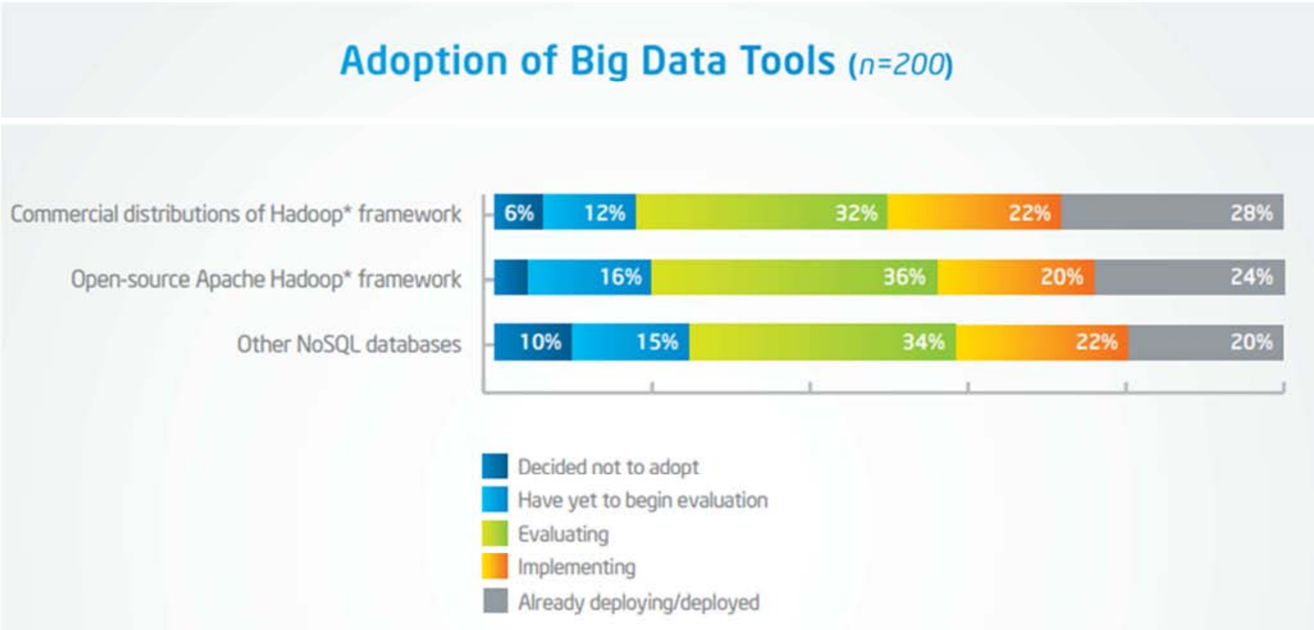

Figure 10. *Adoption of Big Data tools.*

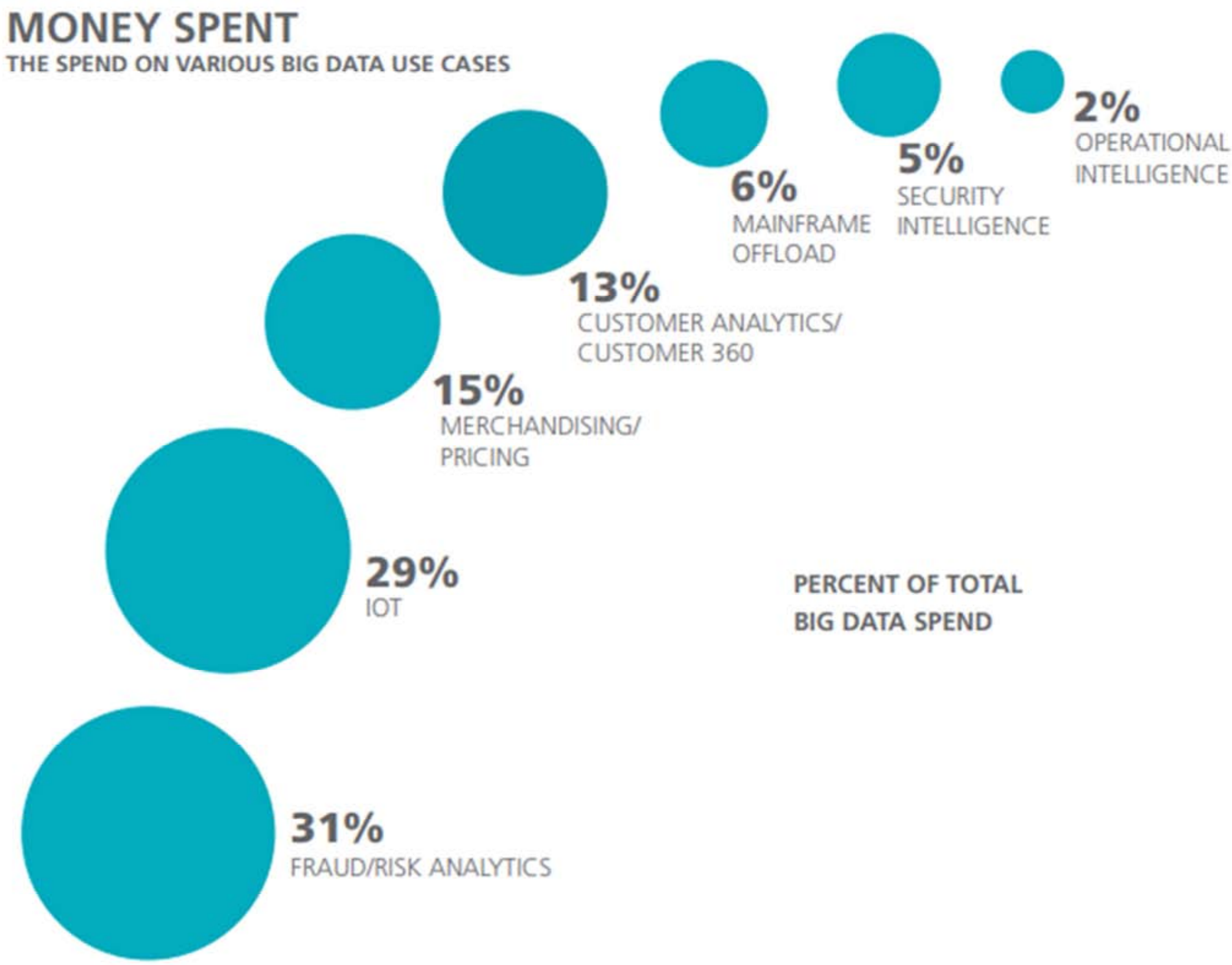

Figure 11. *Money spent on various Big Data cases.*

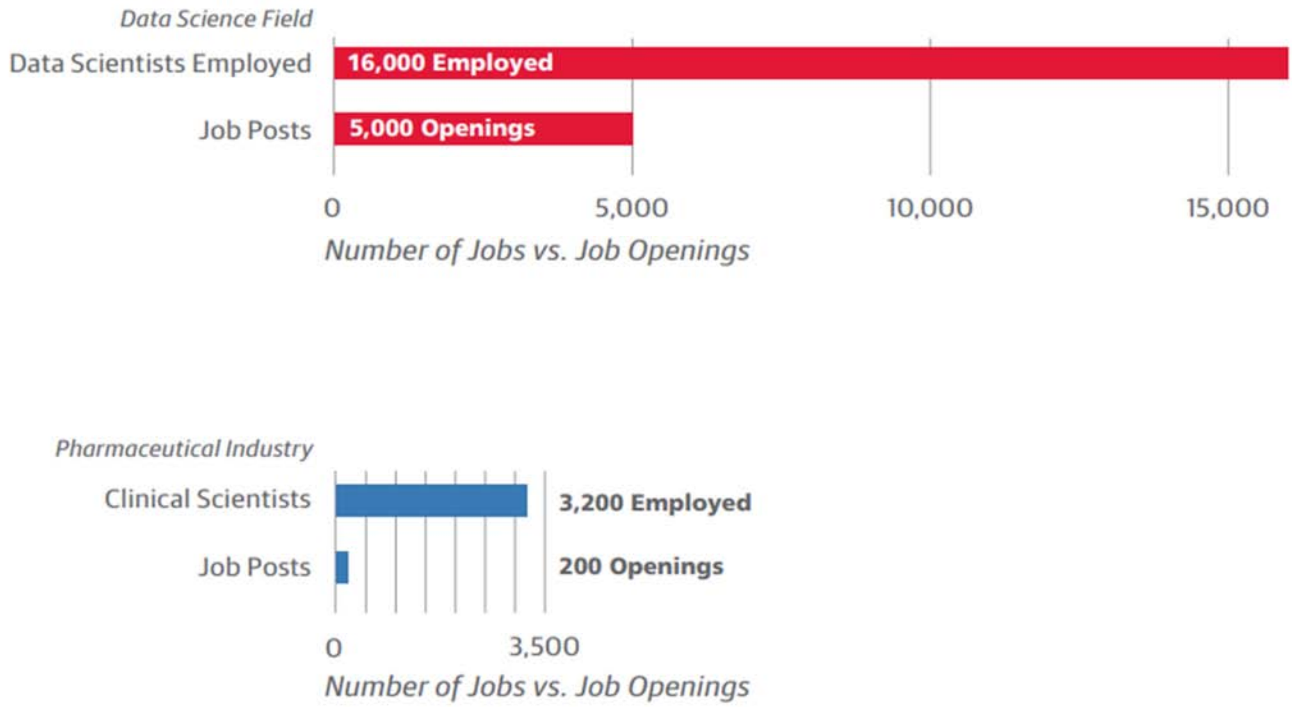

Figure 12. *The number of data scientists versus the number of job openings.*

4. Importance of Big Data Analytics in Marketing

Big data is the greatest game-changing opportunity and paradigm change in marketing since the advent of the popular phone or the Internet. Big data refers to the ever-increasing volume, velocity, variety, value and veracity of information. Big data is a crucial consequence of the modern marketing environment, born from the digital world we now live in, for marketing organizations. The term " big data "does not only refer to the data itself; it also refers to the difficulties, skills and competencies associated with storing and processing such large data sets to enable a level of decision-making, that is more reliable and timely than anything previously attempted; large data-driven decision-making.

Organizations today face vast volumes of data, operational complexity, rapidly evolving consumer behavior, and intensified competitive pressures. New technologies, as well as rapidly expanding networks and platforms, have created a vastly complex environment.

Data worldwide is growing at 40% per year, a growth rate that is good for any marketing and sales leader. Many marketers may believe that data has always been huge, and in some ways it has always been huge. But think of the consumer data hat companies gathered 20 years ago, sales transaction data point, responses to direct mail promotions, coupon redemption, etc. Customer data collected today, can be online purchasing data, click-through rates, browsing habits, social media interactions, mobile device use, geolocation data, etc. Big data can be thought of as a hidden ingredient, a raw material and an integral element. It's not the individual data that's so relevant. Instead, it is the knowledge gained from big data, the choices we make and the actions we take that make all the difference. Three types of big data are the key to marketing:

1. Consumer: The most common big data category for marketing which include behavioral, attitudinal and transactional indicators from such sources, as marketing strategies, points of sale, websites, customer surveys, social media, online communities and loyalty programs.
2. Operational: This broad data category usually contains quantitative indicators that assess the efficiency of marketing processes related to marketing activities, resource utilization, asset management, budgetary controls, etc.
3. Financial: Usually located in the financial systems of the organization, this big data category may include sales, expenses, income and other objective data forms that assess the financial health of the organization.

Having big data does not automatically lead to better marketing Organizations that wish to benefit in marketing from Big Data should do the following things:

1. Effective exploration of new possibilities. Effective exploration involves the development of a data advantage by drawing the related data sets from both within and outside the business. Analytics leaders need to step beyond general objectives such as "increase wallet share" and to a degree of detail that is relevant. They need to use digital knowledge to improve target buyers and use analytics to understand more about target buyers. Modern marketers should shed light on a more different level of information, such as: which websites are most commonly used by users, which social media accounts they have and use, and also which buttons they press on a website. With big data, "ideal customer

profiles" can be easily targeted.

2. Understand the journey of customer decision-making. Today's channel-surfing consumer, is comfortable using a variety of devices, tools and technologies that fulfil their tasks. Understanding that the decision journey is key to the discovery of battlegrounds, either to attract new customers, or to prevent current ones from finding competitors.

Marketing and sales executives need to build full photographs of their customers, so that they can create advertisements and items that are important to them. Through integrating big data with an integrated marketing management approach, marketing companies can have a major effect in these main areas:

1. Customer engagement: Big data will provide insight not only into who your customers are, but where they are, what they want, how they want to be contacted, and when.
2. Customer retention and loyalty: Big data will help you discover what affects customer loyalty and what prevents them coming back again and again.
3. Marketing Optimization / Performance: with big data, you can assess optimum marketing spending through multiple platforms, as well as continuously refine marketing programs through monitoring, evaluation and analysis.
4. Track Google Trends to help educate your global / local strategy. Google Trends is perhaps the most approachable way of using big data. Google Trends introduces trending patterns by quantifying how much a specific search word is entered compared to the overall search amount. Global marketers can use Google Trends to determine the popularity of certain topics across countries, languages or other parties that they might be interested in, or to keep up-to - date on topics that are cool, new, top-of-the-mind, or important to their buyers.
5. Build real-time choice for customers. Marketers ought to deliver the correct message at the right time. Timeliness and relevance are also the cornerstone of effective marketing campaigns, e-mail click-through rates, and customer interaction with your brand. Big data provides advertisers with timely insights into who is interested or involved in their product or content in real time. Tying digital buyer behavior to your CRM (customer relationship management) systems and marketing automation tools helps you to track the topics that your customers are most interested in and give them content that makes the most sense to grow certain ideas.
6. Identify the unique material that drives customers down the sales enclosure. How successful has a unique blog or social post been in generating revenue? It was an unanswerable issue before the big data. We did social media campaigns, and content development, because we thought it was working, but we didn't have a way to back the argument. Marketers can now drive the success of a marketing drive down to a tweet. Tools like content rate illuminate which individual content assets have been successful in a closed / won deal, and which have been inefficient.
7. This allows marketers to tailor their strategies to content themes or types that most resonate with their buyers, and truly push them to purchase.
8. Make it fast and easy. Companies need to invest in an automated "algorithmic marketing" strategy that enables large quantities of data to be processed via a "self-learning" mechanism, in order to build better and more meaningful experiences with customers. This may include predictive statistics, machine learning, and the mining of natural languages. For example, these systems will automatically monitor keywords, and update them every few seconds on the basis of changing search terms used, ad costs or consumer behavior. It can make price adjustments on the fly through thousands of items based on consumer expectations, price comparisons, inventory and predictive analysis.

Various business functions will benefit from analytics. The most common functional groups are:

1. Customer analysis: This category covers marketing applications (customer profiling, segmentation, social network analysis, brand credibility analysis, marketing mix optimization) and customer experience.
2. Supply Chain Analytics (SCA): This involves forecasting of demand and optimization of inventory, pricing, scheduling, transport and storage, as well as mitigation and danger. In The SCA, Human Capital Analytics and Labor Analytics, concern service industries where human resources are the most significant means of production.
3. Public domain analysis: motivated by natural resource constraints, governments use analytics for tasks such as detecting water pollution in delivery networks, making electricity grids and traffic structures smarter, and improving public safety.
4. Fraud and risk analysis: This involves the evaluation of different forms of risk (market, operational, credit) primarily in the financial sector.

5. Opportunities and Challenges of Big Data Analytics

In order for data sources to be loaded into the data warehouse, the data needs to be well known, organized and standardized with acceptable data form. While this form of centralization allows the security, backup and failure of highly sensitive data, it also means that data usually needs to undergo substantial pre-processing and checkpoints before it can reach this form of managed environment that does not lend itself to data discovery and iterative analysis. As a result of this level of control on the EDW (enterprise data warehouse), additional local systems may emerge in the form of departmental warehouses and local data marts created by business users to meet their need for flexible analysis.

In order to better understand the business drivers related to Big Data, it is important to first understand some of the historical history of data stores and the types of repositories and resources used to administer these data stores. The amount of information was mostly estimated in terabytes in the 1990s. Some companies analyzed structured data in rows and columns, and used relational databases and data warehouses to operate massive business information stores. The following decade saw a proliferation of various types of data sources, mainly efficiency and publishing tools such as content management repositories and networked attached storage systems, to handle this sort of information, and data began to grow in size and began to be measured at petabyte scales. In the 2010s, the information that companies are attempting to handle has broadened to include several other forms of data. Figure 13 [9] shows interest in software for Big Data analytics.

In this century, everyone and everything is leaving a digital footprint. The applications of Big Data, which produce data volumes that can be calculated on an exabyte scale, create opportunities for new research and provide new value for organizations.

When IT managers [9] were asked about their major big data problems, it is found that most of them were ranked equally in terms of their importance. This suggests that respondents face more than one significant data challenge. The key problem cited was not surprising: the rise in data and the expense and effort needed to contain or store it. Forty-eight per cent referred to this as one of their top three obstacles. However, this is accompanied closely by: Concern about data infrastructure and the capacity of their data center to provide the scalability, low latency and efficiency required for big data projects (41 per cent), Data governance and policy challenges for identifying data that will be processed, analyzed and accessed, along with deciding its importance (41 per cent) The findings are very simple. IT managers face major big data problems from a variety of directions. Figure 14 [9] shows challenges for Big Data.

When asked about clear barriers to big data analytics, IT managers see protection trumping them all. We believe that this is motivated by IT managers who want to collaborate with third-party cloud service providers, but are also worried about the capacity of these providers to protect their data. What do IT managers need to go on with? Interestingly, the IT managers in the sample [9] are comfortable knowing the needs of their organizations for big data analytics. They have plans to implement technology to meet these needs and are quite interested in using third-party cloud providers. IT managers are looking for support in a variety of ways, including creative ideas for particular applications and realistic tools to assist with implementation. The creation of industry standards for data protection and privacy as well as interoperability are also considered essential. Figure 15 [9] shows obstacles to Big Data analytics. When asked if they understand the extent of demand for big data from business users and data analytics experts in their organizations, 80% of the sample [9] thought they did. The vast majority of this community (89 per cent) also have plans to implement technology and applications to support big data requirements, including data volume, speed and variety of IT managers, which display very strong interest in learning more about infrastructure and software built to support big data analytics. Figure 16 [9] shows IT mangers understand of users/experts Big Data demands.

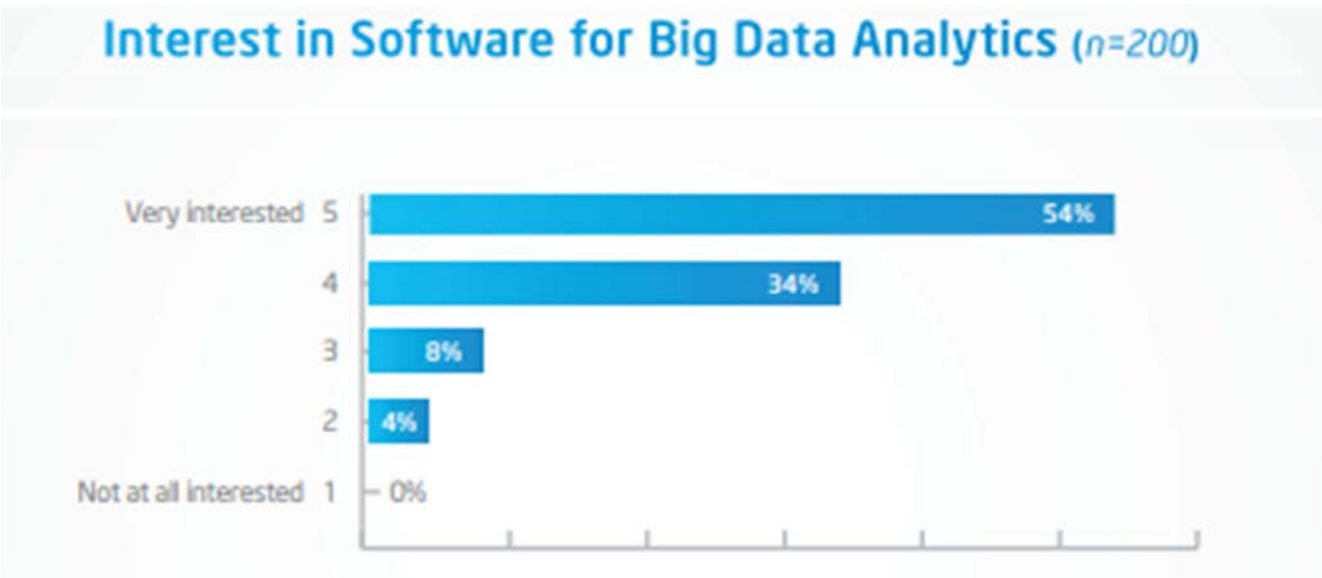

Figure 13. *Interest in software for Big Data Analytics.*

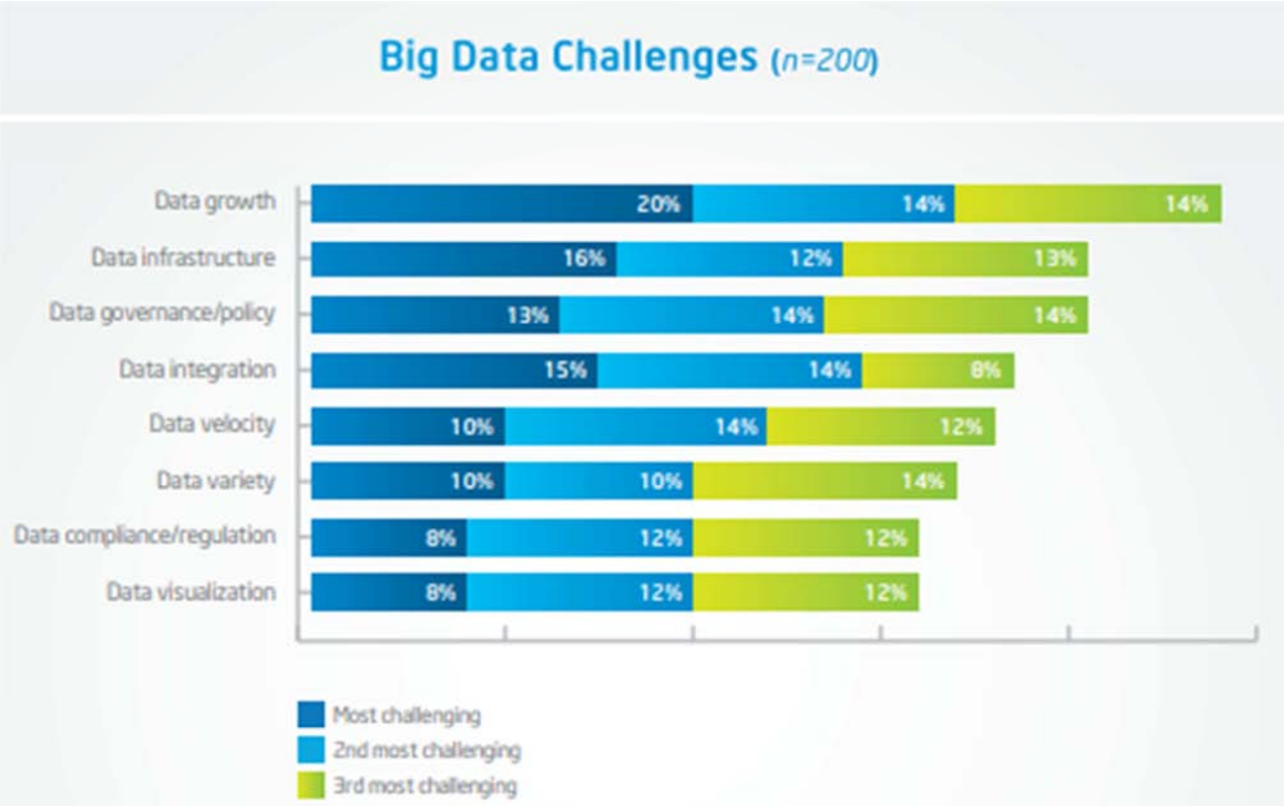

Figure 14. Big Data challenges.

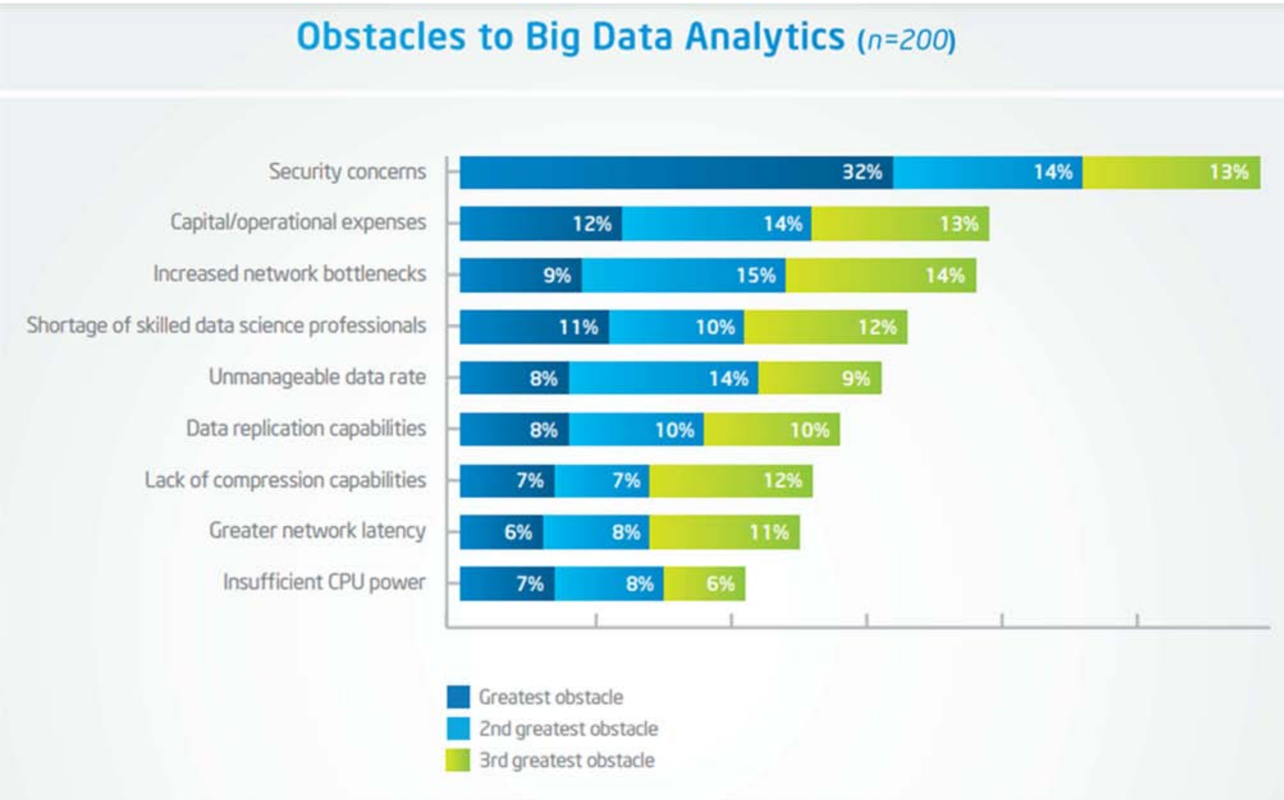

***Figure 15.** Obstacles to Big Data Analytics.*

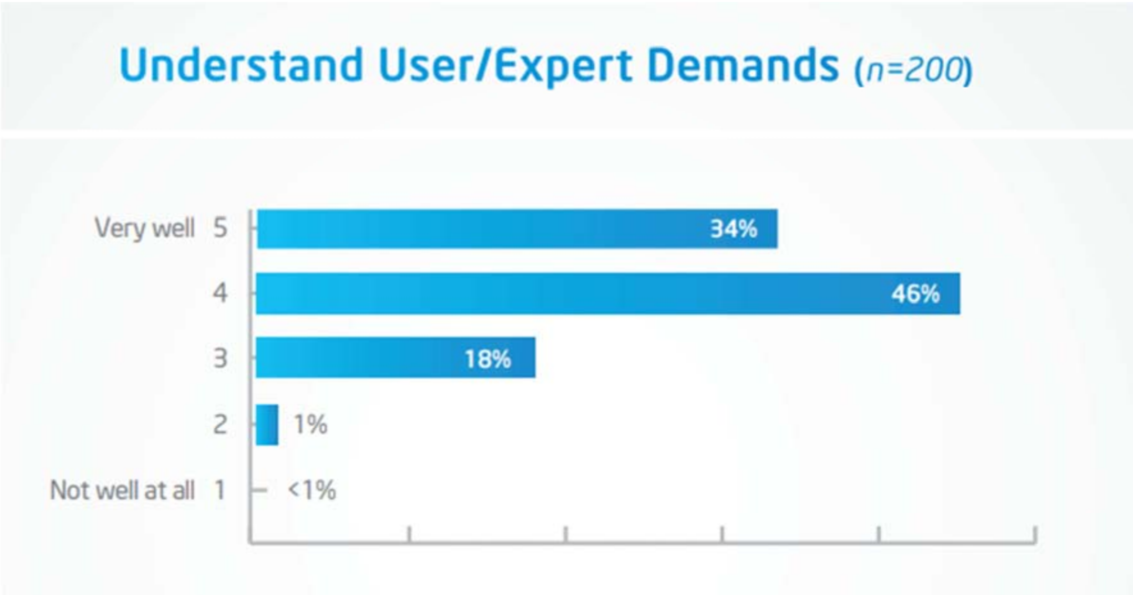

Figure 16. *IT managers understand user/expert Big Data demands.*

6. Case Study of Big Data Analytics in Marketing Using R Tool

In this use case, an example of providing new financial tool through a bank to its customers, and there is a data of all bank customers containing 24 variables, including their age, marital status, Education status, job, have a house or not, have obtained loan or not before to be studied before presenting a new financial tool, and how to market the new tool based on bank's customer needs, for example, if a huge portion of bank's customers does not have homes, a new financial tool directed for buying homes could be provided. Taking into consideration that the number of customers are more than 41 thousands persons.

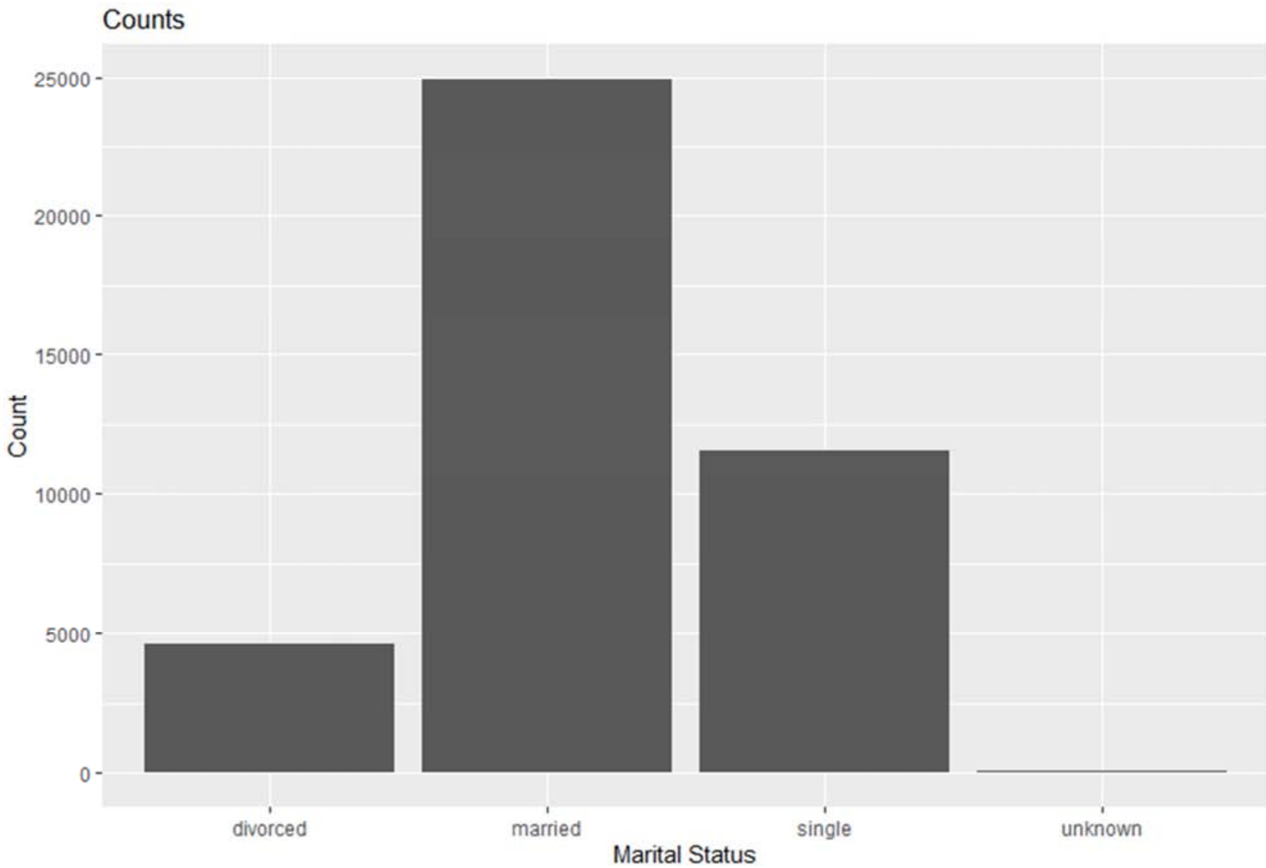

Figure 17. *Showing Marital Status.*

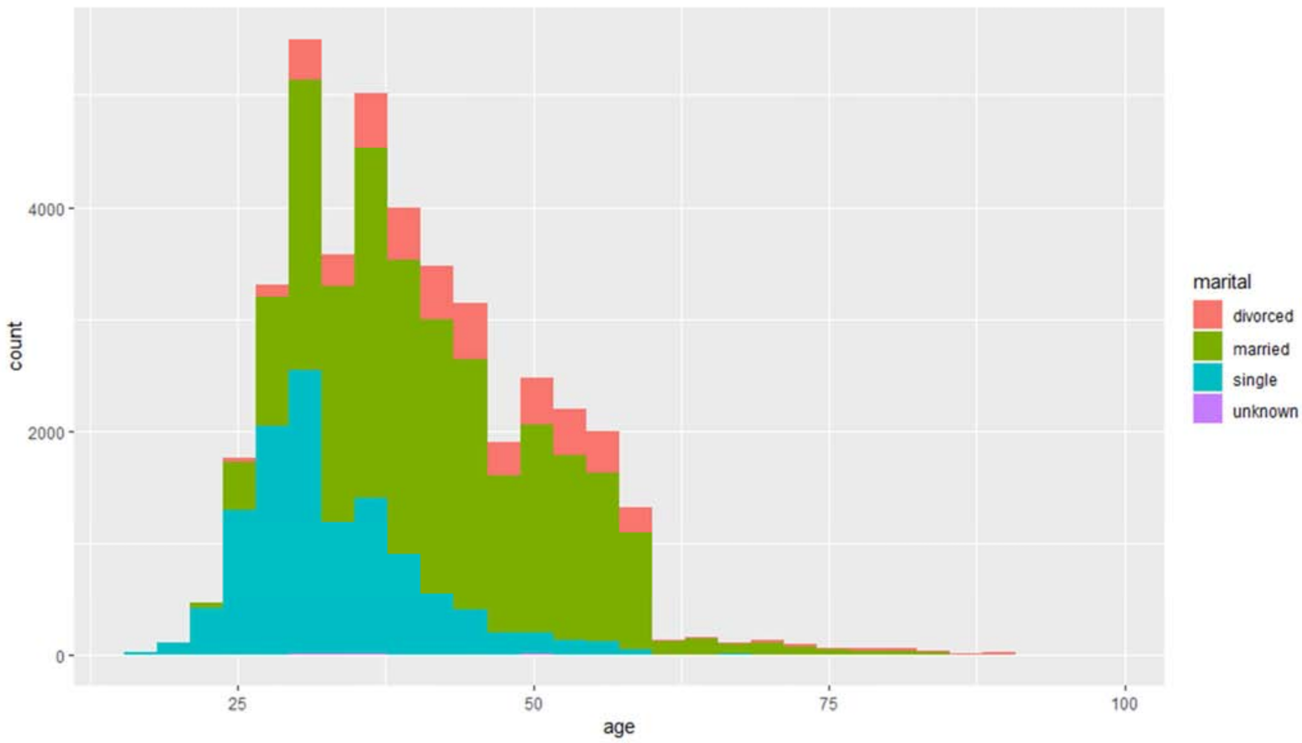

Figure 18. *Showing Marital Status versus Age.*

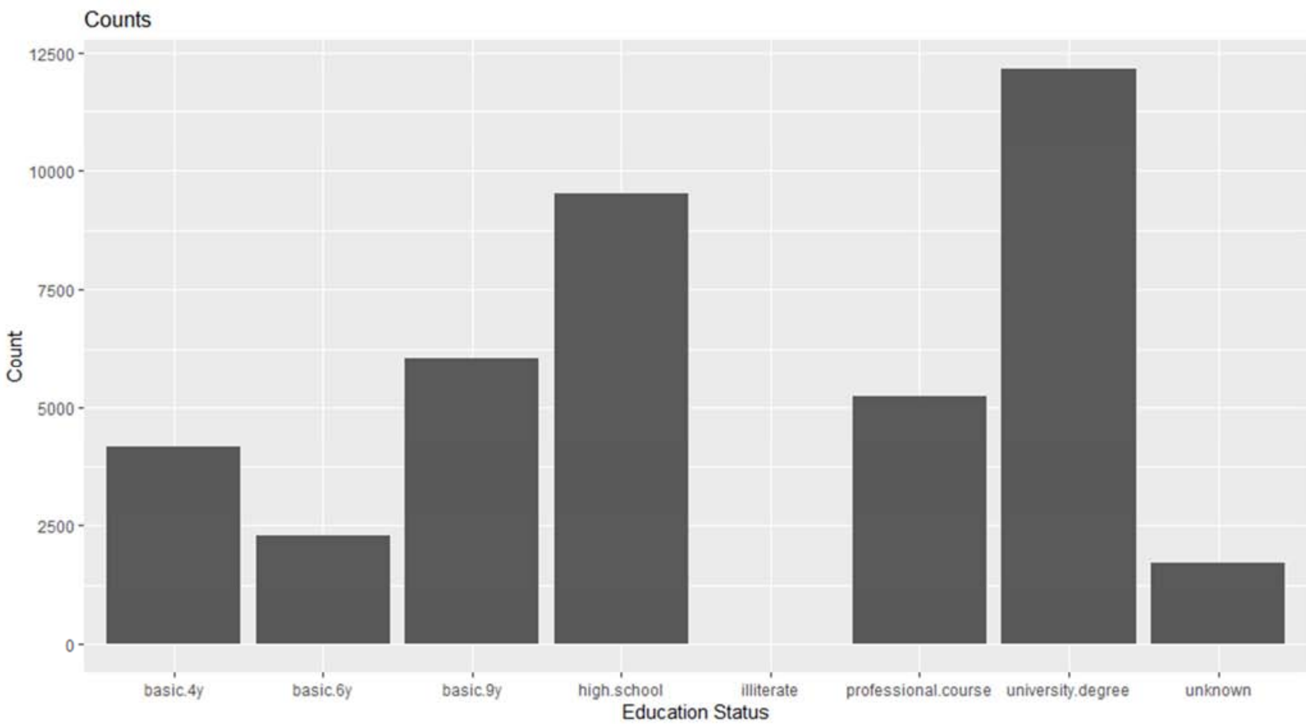

Figure 19. *Showing Education Status.*

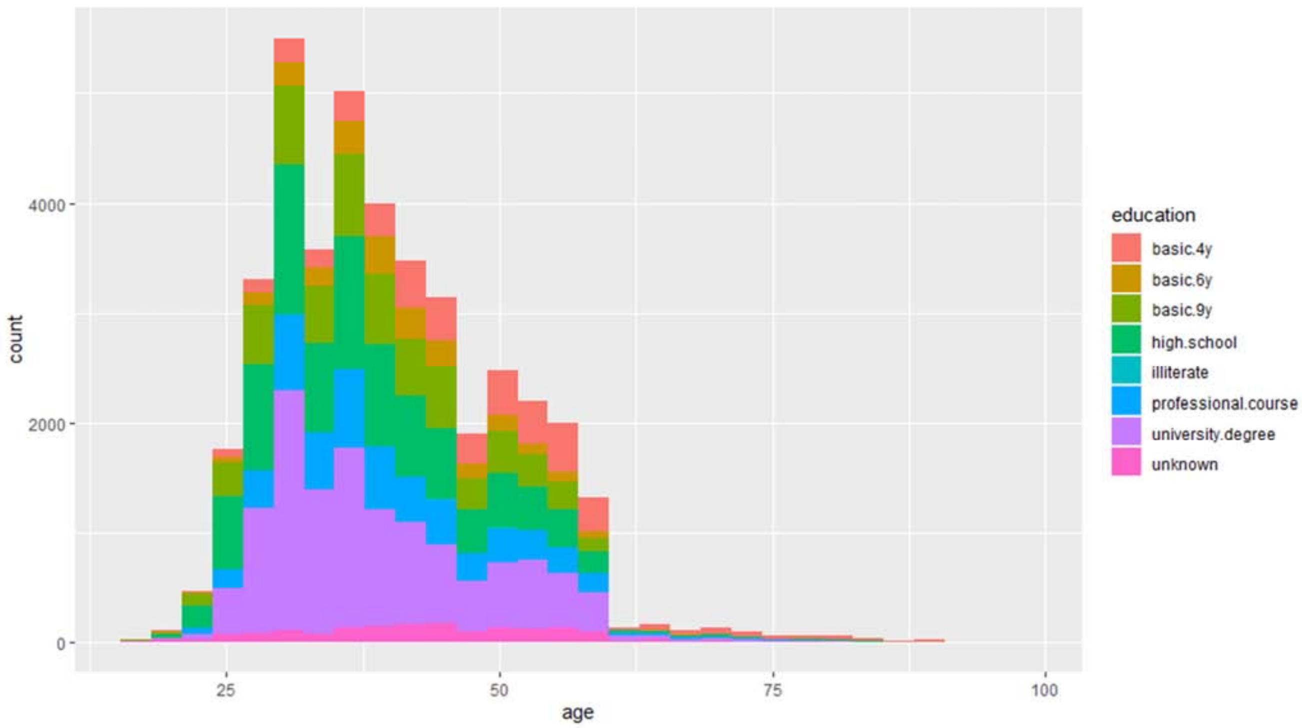

Figure 20. *Showing Education Status versus Age.*

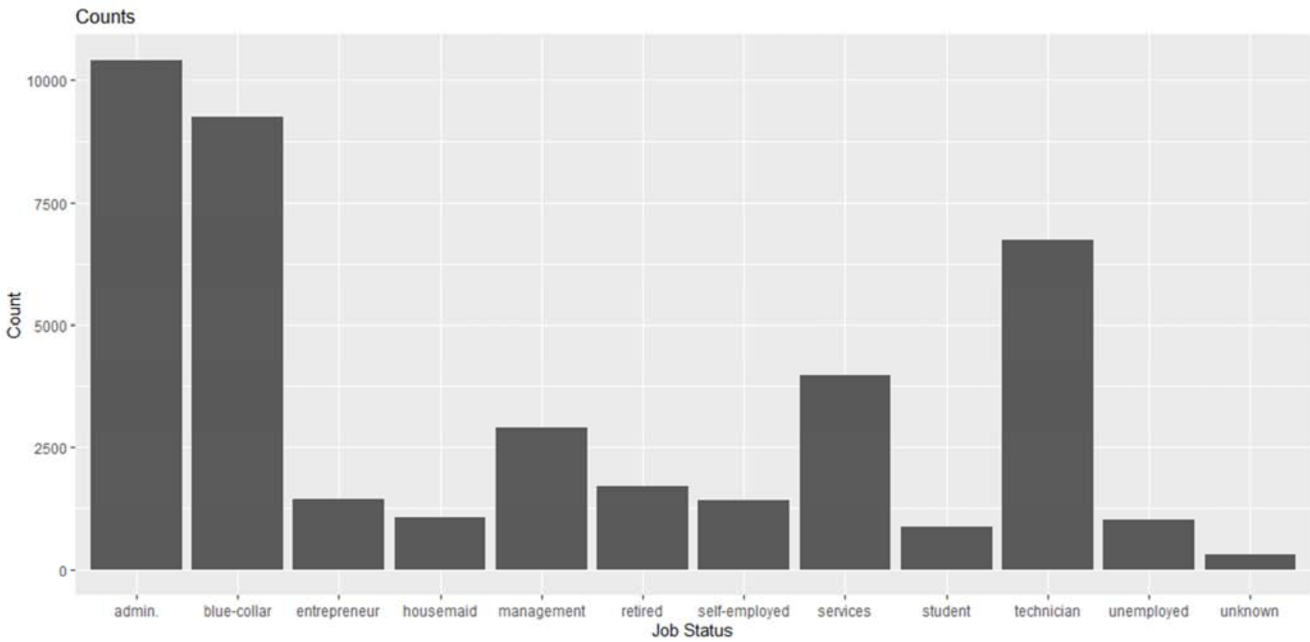

Figure 21. *Showing Job Status.*

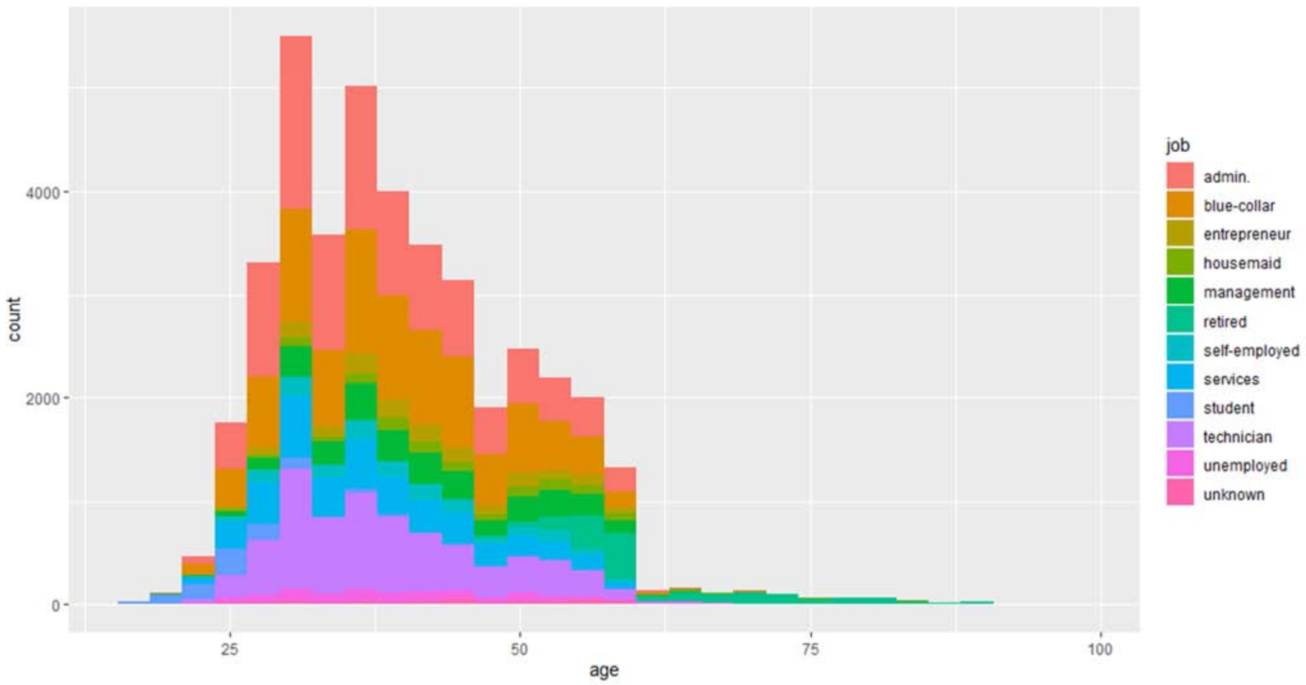

Figure 22. *Showing Job Status versus Age.*

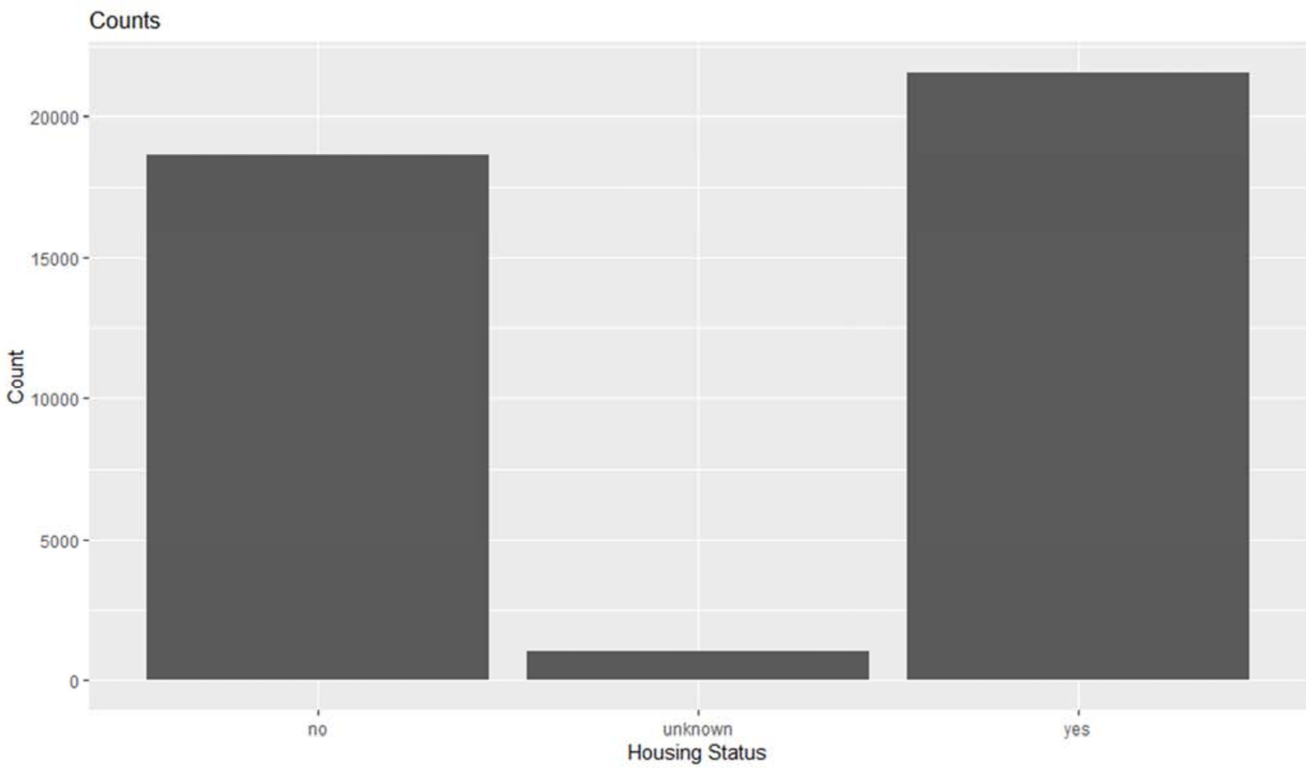

Figure 23. *Showing housing Status.*

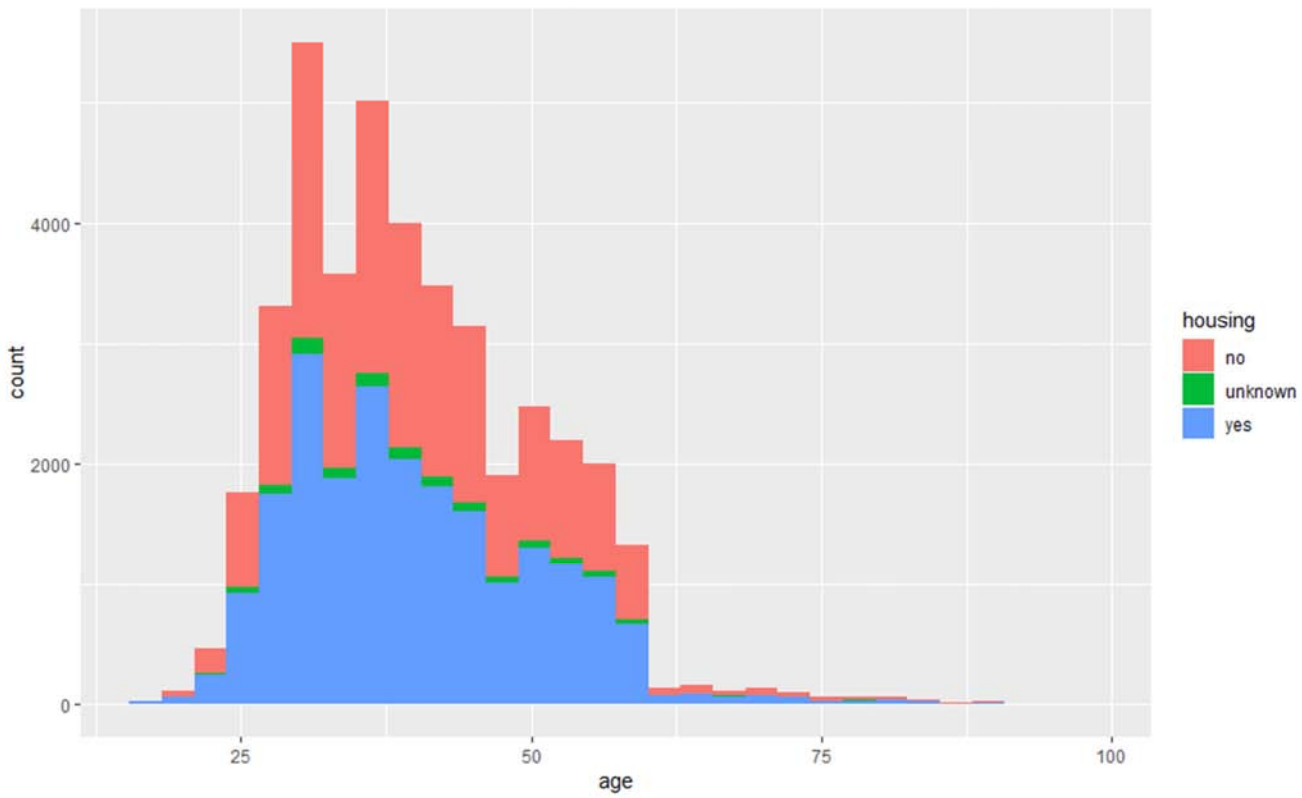

Figure 24. *Showing housing Status versus age.*

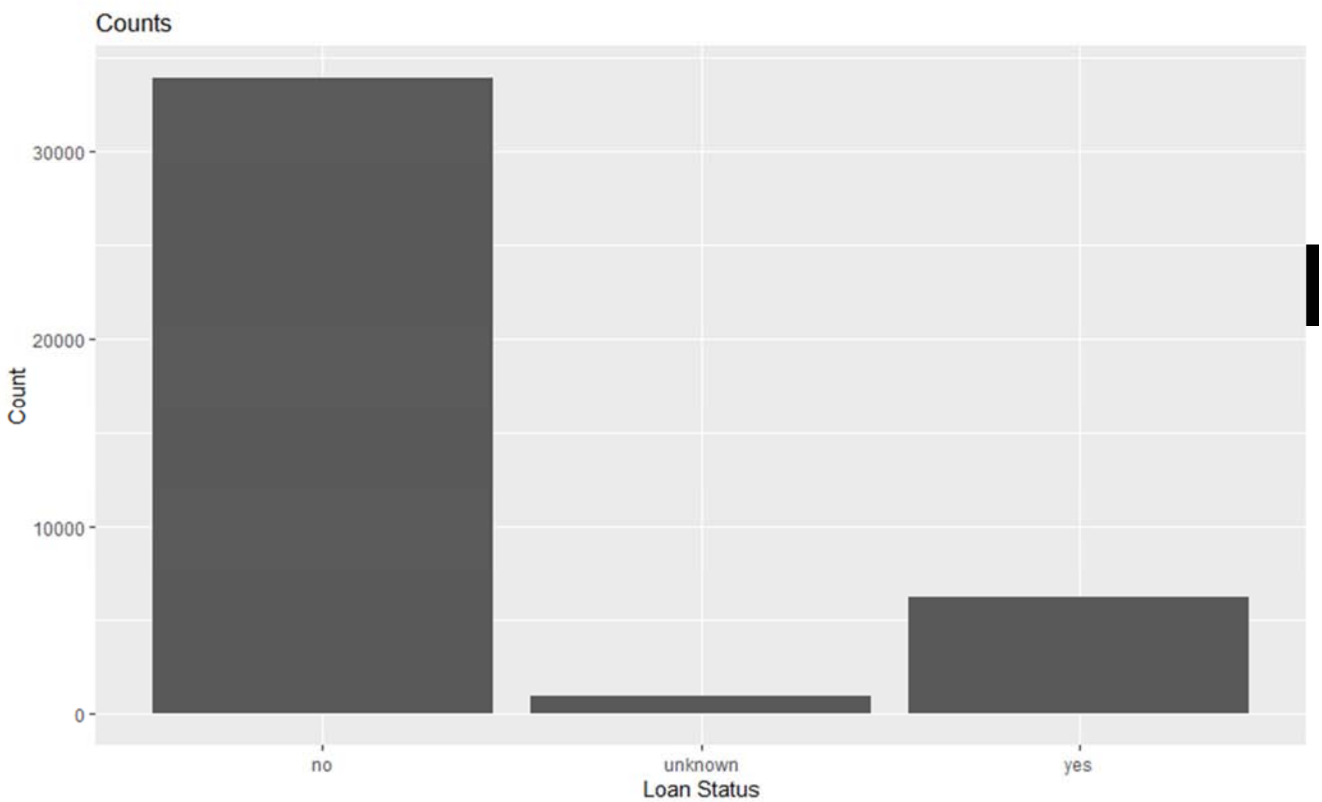

Figure 25. *Showing loan status.*

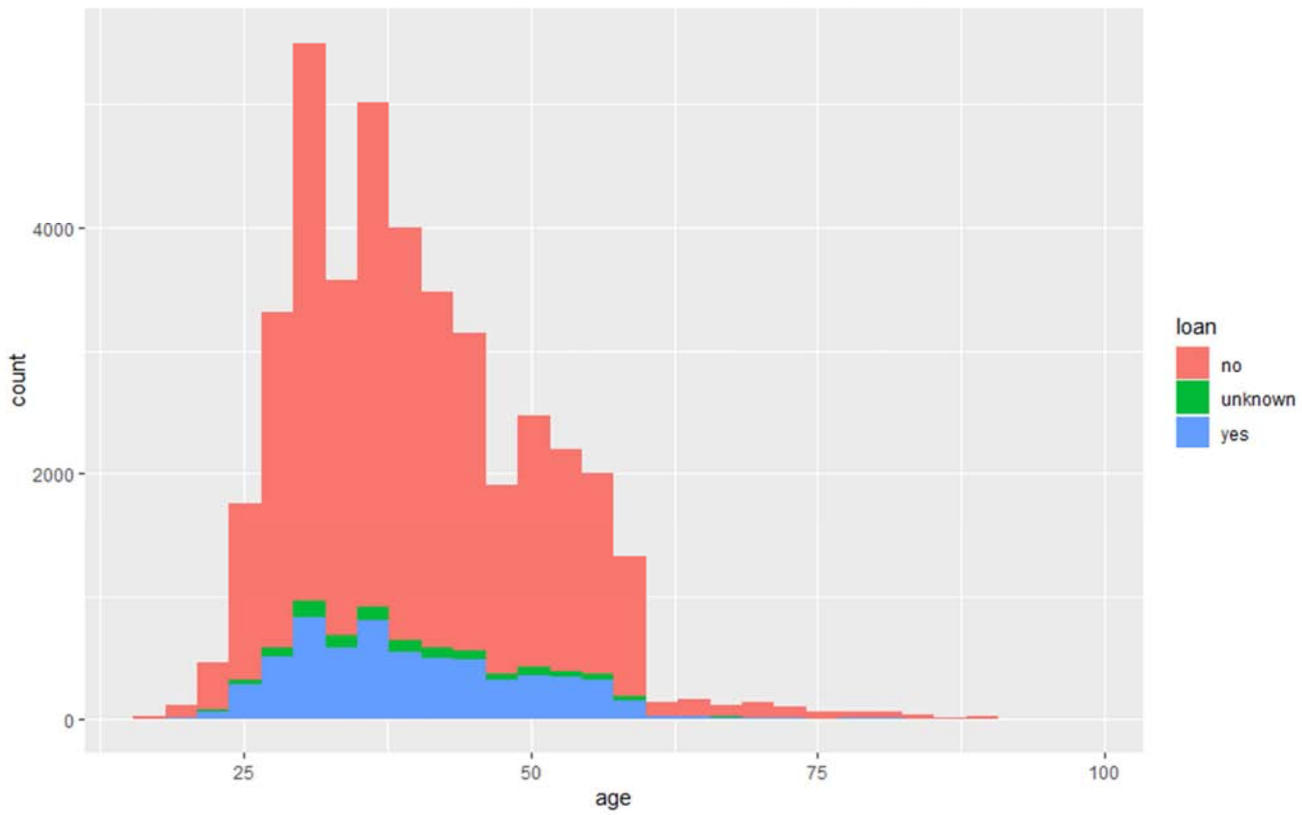

Figure 26. *Showing loan status versus age.*

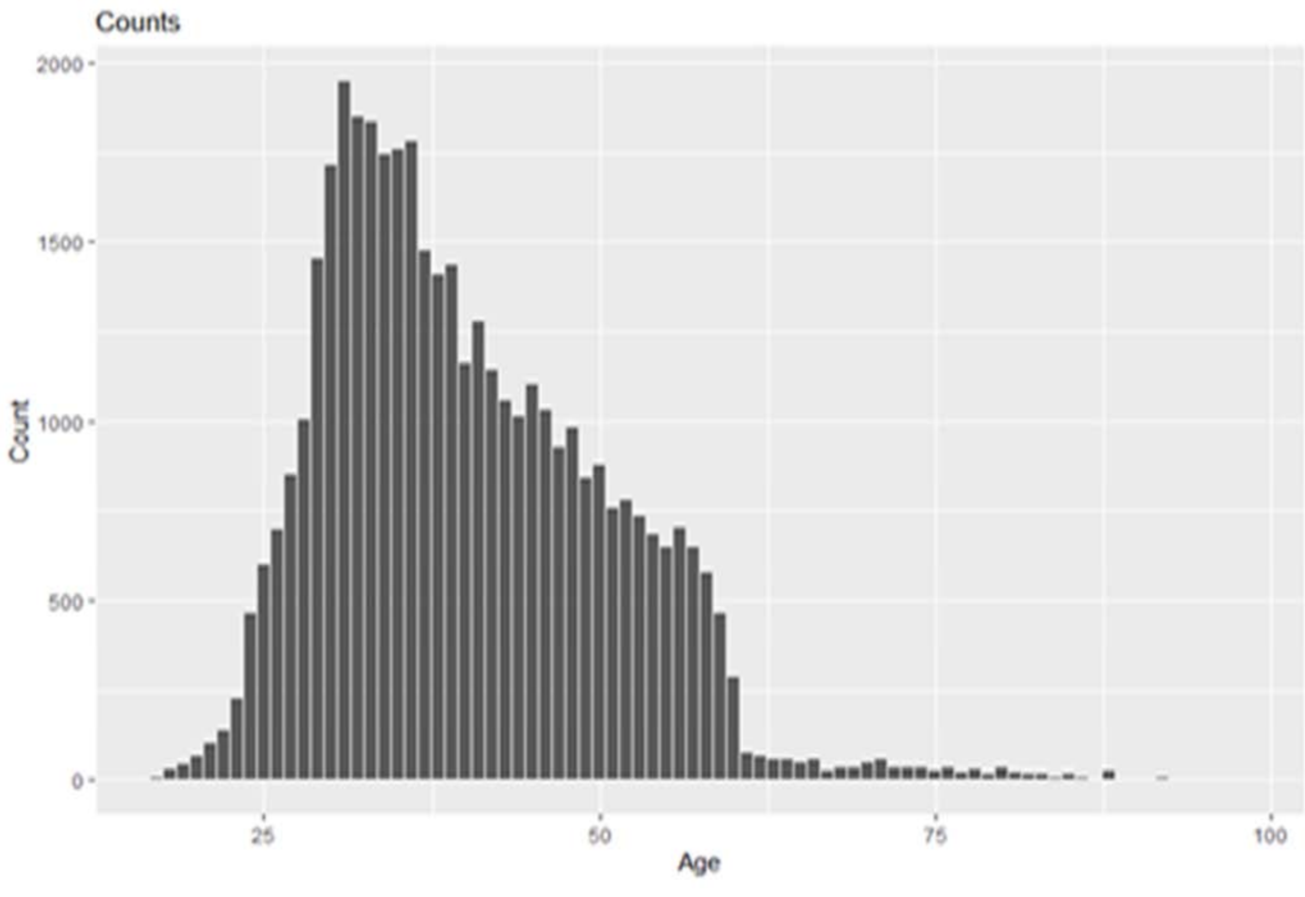

Figure 27. *Showing age of bank customers.*

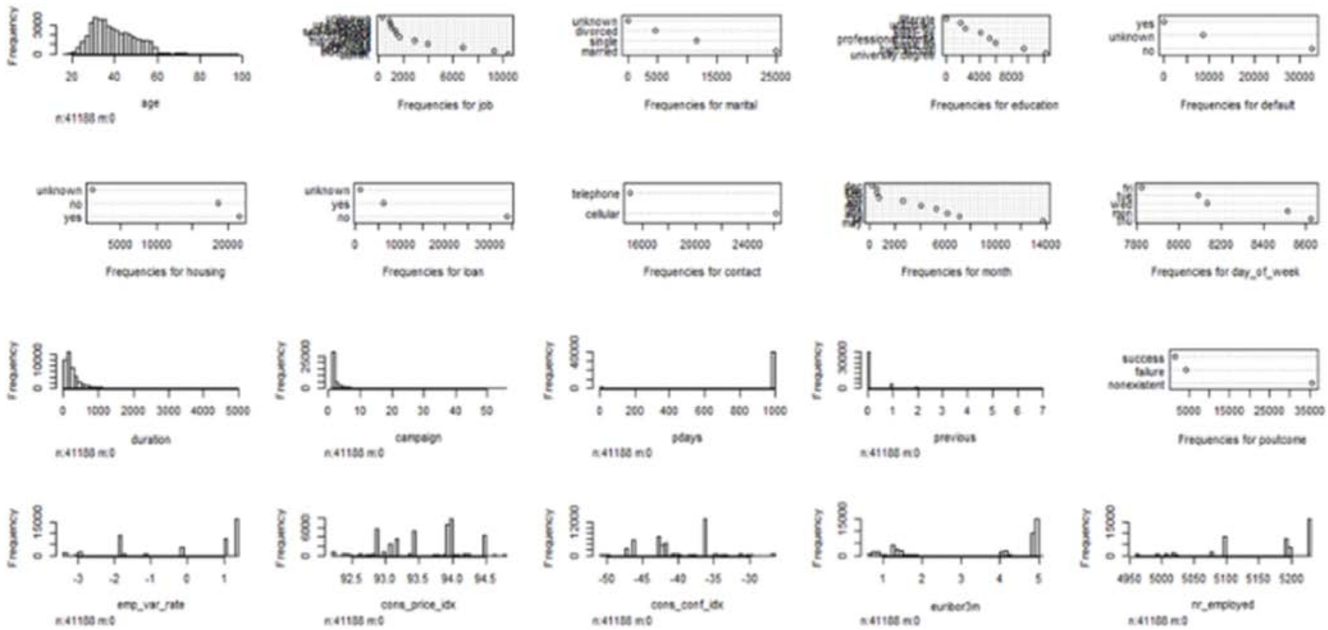

Figure 28. *Showing total 24 parameters.*

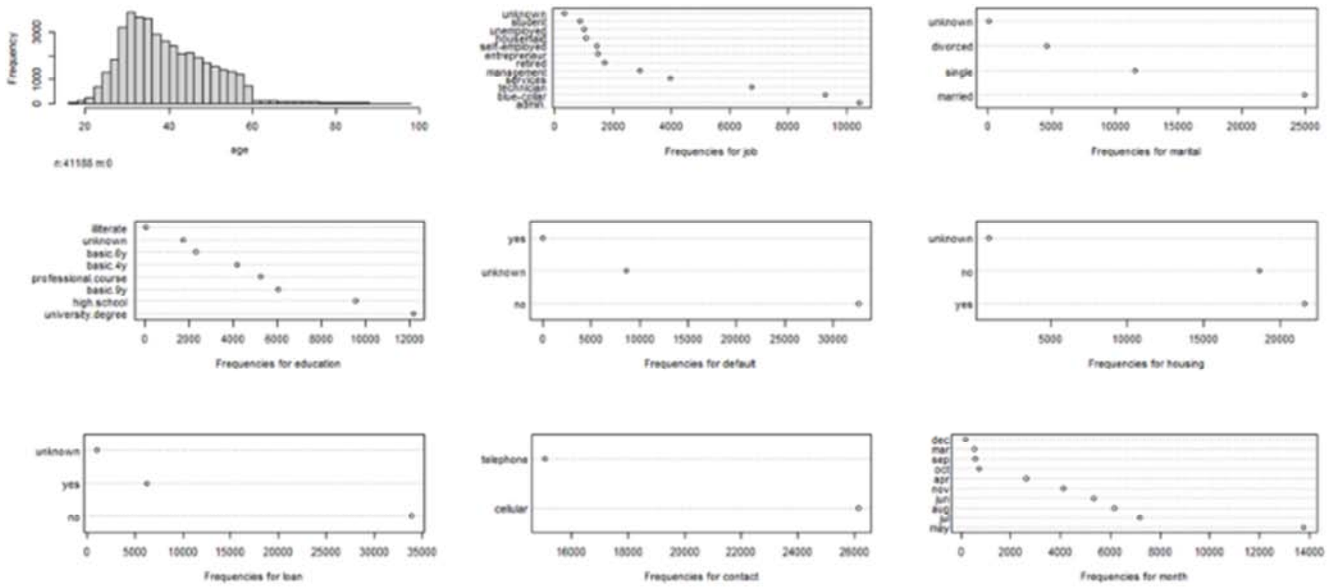

Figure 29. *Displaying the first 9 parameters.*

The R tool was used for simulating results. The R was created by Ross Ihaka and Robert Gentleman in the Department of Statistics at the University of Auckland in 1991. A major advantage that R has over many other statistical packages, and is that it's free in the sense of free software. The copyright for the primary source code for R is held by the R Foundation [10] and is published under the GNU General Public License version 2.0 [11]. R functionality is divided into a number of packages:

1) The "base" R system contains, among other things, the base package which is required to run R and contains the most fundamental functions.

2) The other packages contained in the "base" system include utils, stats, datasets, graphics, grDevices, grid, methods, tools, parallel, compiler, splines, tcltk, stats4.

3) There are also "Recommended" packages: boot, class, cluster, codetools, foreign, KernSmooth, lattice, mgcv, nlme, rpart, survival, MASS, spatial, nnet, Matrix. The R tool can be downloaded from the CRAN (Comprehensive R Archive Network) project website [12]. The R tool was chosen due to its many functionalities and ease of use [13, 14]. So, many experiments could be done using the R tool [15]. The following figures from 17 to 29 are created using R tool, and represent respectively marital status, marital status related to age, education status, education status related to age, job status, job status related to age, housing status, housing status related to age, loan status, loan status related to age, and finally age distribution of the bank's customers. Then, all 24 parameters representing the customers' information are represented in one graph, as well as the first 9 parameters in another graph. We

can derive from these statistics that a large portion of bank customers does not have a house, also, and that portion is mainly between the age of 25 and 55. Also, there could be other forms of financial tools based on these statistics, such as providing loans, since a big number of bank customers does not have loans.

7. Conclusions

This paper concentrates on the importance of big data analytics today, and how it represents a useful tool for decision making inside organizations and companies today. Also, the paper focused on the use of big data analytics in marketing and a use case of a bank customers was studied for the purpose of providing a new financial tool for these bank customers especially in housing. There is a lot of valuable information companies can get from using big data analytics. So, it will be beneficial for all companies to implement big data analytics in the future.

References

[1] Diebold, F. (2003). Big Data: Dynamic Factor Models for Macroeconomic Measurement and Forecasting. Advances in Economics and Econometrics, Eighth World Congress of the Econometric Society. Cambridge University Press, 115-122. doi: 10.1017/CBO9780511610264.005.

[2] Laney, D. (February 6, 2001). 3-D Data Management: Controlling Data Volume, Velocity and Variety. META Group Research Note. Retrieved 10 November 2020, from https://idoc.pub/documents/3d-data-management-controlling-data-volume-velocity-and-variety-546g5mg3ywn8.

[3] Chang, F., Dean, J., Ghemawat, S., Hsieh, W. C., Wallach, D. A., Burrows, M., Chandra, T., Fikes, A. and Gruber, R. E. (2008). Bigtable: A Distributed Storage System for Structured Data. ACM Transactions on Computer Systems, 26 (2), 1-26. doi: 10.1145/1365815.1365816.

[4] EMC education services (2015). Data Science and Big Data Analytics, Discovering, Analyzing, Visualizing and Presenting Data. John Wiley & Sons. doi: 10.1002/9781119183686.

[5] McKinsey & Company (March 2015). Marketing & Sales, Big Data, Analytics, and the future of Marketing & Sales. Retrieved 10 November 2020, from https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/Marketing%20and%20Sales/Our%20Insights/EBook%20Big%20data%20analytics%20and%20the%20future%20of%20marketing%20sales/Big-Data-eBook.ashx.

[6] H. Davenport, T., and Dyché, J. (May 2013). Big Data in Big companies. International Institute for analytics. Retrieved 10 November 2020, from https://docs.media.bitpipe.com/io_10x/io_102267/item_725049/Big-Data-in-Big-Companies.pdf.

[7] Dresner Advisory Services (December 20, 2017). Big Data Analytics Market Study Wisdom of Crowds Series Licensed to MicroStrategy. Retrieved 10 November 2020, from https://www3.microstrategy.com/getmedia/cd052225-be60-49fd-ab1c-4984ebc3cde9/Dresner-Report-BigData_Analytic_Market_Study-WisdomofCrowdsSeries-2017.pdf.

[8] Naimat, A. (2016). The Big Data Market, A data driven analysis of companies using Hadoop, Spark, Data Science & Machine Learning. Oreilly. Retrieved 10 November 2020, from http://ixion.pld-linux.org/~arekm/free-books/data/the-big-data-market.pdf.

[9] Intel IT Center Peer research (August 2012). Big data analytics, Intel's IT Manager Survey on How Organizations Are Using Big Data. Retrieved 10 November 2020, from https://www.intel.com/content/dam/www/public/us/en/documents/reports/data-insights-peer-research-report.pdf.

[10] Project, R (2020). R project. Retrieved 10 November 2020, from https://www.r-project.org/about.html.

[11] Project, GNU (2020). GNU project. Retrieved 10 November 2020, from http://www.gnu.org/gnu/gnu.html.

[12] Project, CRAN (2020). CRAN project. Retrieved 10 November 2020, from https://cran.r-project.org/bin/windows/.

[13] Anurag (18 Apr. 2020). 6 Reasons: Why Choose R Programming for Data Science Projects? Retrieved 10 November 2020, from https://www.newgenapps.com/blog/6-reasons-why-choose-r-programming-for-data-science-projects/.

[14] Agrawal, V. (25 February, 2016). Applications Of R Programming In Real World. Retrieved 10 November 2020, from https://elearningindustry.com/applications-r-programming-r-eal-world.

[15] Lawson, J. (2015). Design and Analysis of Experiments with R. Brigham Young University Provo, Utah, USA, Taylor & Francis Group. doi: 10.1201/b17883.