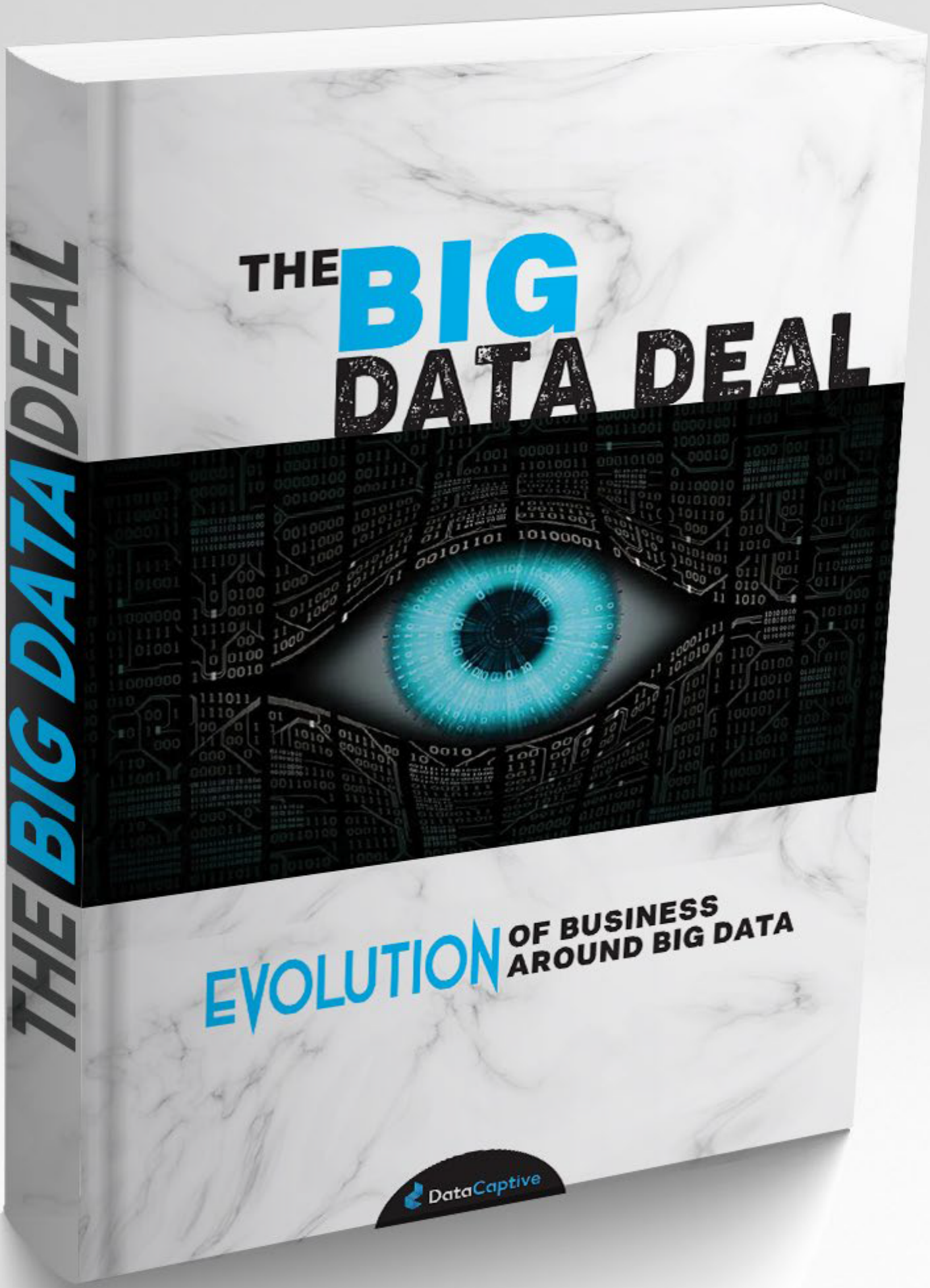

THE BIG DATA DEAL
THE BIG DATA DEAL
EVOLUTION OF BUSINESS AROUND BIG DATA
DataCaptive

TABLE OF CONTENT

INTRODUCTION

CHAPTER 1 — 1 A Short History of Big Data

CHAPTER 2 — 12 Evolution of Data Collection and Analytics

CHAPTER 3 — 20 How Big Data Is Empowering Artificial Intelligence & Machine Learning

CHAPTER 4 — 25 Application of Database Management Systems

CHAPTER 5 — 31 How Companies Are Using Big Data and Analytics?

CHAPTER 6 — 44 Synopsis

1st CHAPTER

A Short History of Big Data

"The world is now awash in data and we can see consumers in a lot clearer ways."

Max Levchin, PayPal co-founder

Big data- This word has become a hype these days. "Data is growing at a rapid pace. By 2020 the new information generated per second for every human being will approximately amount to 1.7 megabytes". Everywhere you go, every business enterprise you go to, works around one fundamental principle and that is "Big Data". Let us find out what is this "Big data", how it started and let us look into a short history of Big Data.

> "Big Data is a collection of data sets so large and complex that it becomes difficult to process using on-hand database management tools or traditional data processing application."
>
> –Wikipedia.

Big data is a term that describes the large volumes of data – both structured and unstructured – that inundates a business on a day-to-day basis. For Big data, it's not the amount of data that's important. It's what organizations do with the data that matters. Big data can be analyzed for insights that lead to better decisions and strategic business moves.

The term "Big data" is relatively new but the act of collecting and storing the information for the purpose of analyzing has been in practice from a very long time.

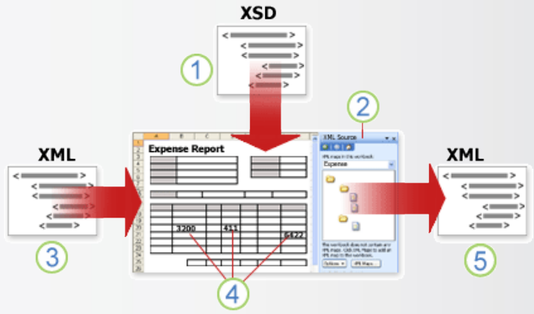
It can be classified into three forms- unstructured, structured, and semi- structured.

Structured data- fixed format

ID	NAME	CLASS	MARK	SEX
1	John Deo	Four	75	female
2	Max Ruin	Three	85	male
3	Arnold	Three	55	male
4	Krish Star	Four	60	female
5	John Mike	Four	60	female
6	Alex John	Four	55	male
7	My John Rob	Fifth	78	male
8	Asruid	Five	85	male
9	Tes Qry	Six	78	male
10	Big John	Four	55	female

Unstructured data

Semi- structured data



Characteristics of Big data

Volume

The word Big Data makes it all so clear that the volume of information involved with Big Data is huge.

Variety

Along with spreadsheets and databases, emails, photos, videos, PDFs, audios, etc. are considered in analysis applications.

Velocity

Velocity refers to the speed of generation of data. Big Data deals at the speed at which data flows in business processes, mobile devices.

Variability

This refers to the inconsistency shown by the data at times that hampers the process of handling and managing the data effectively.

History of Big Data

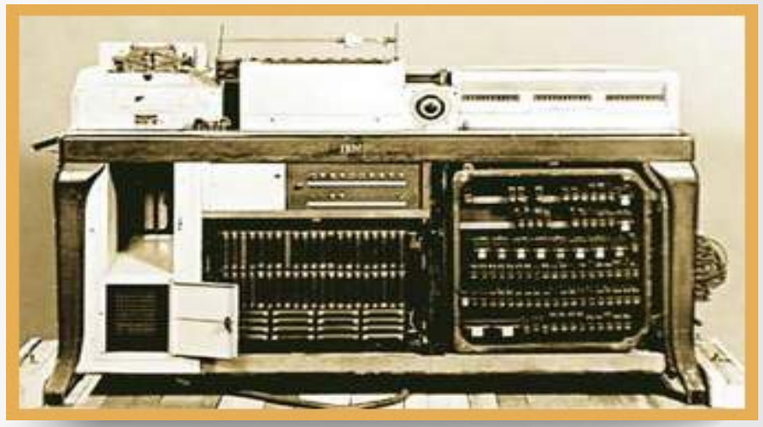
"You can have data without information, but you cannot have information without data."

– Daniel Keys Moran

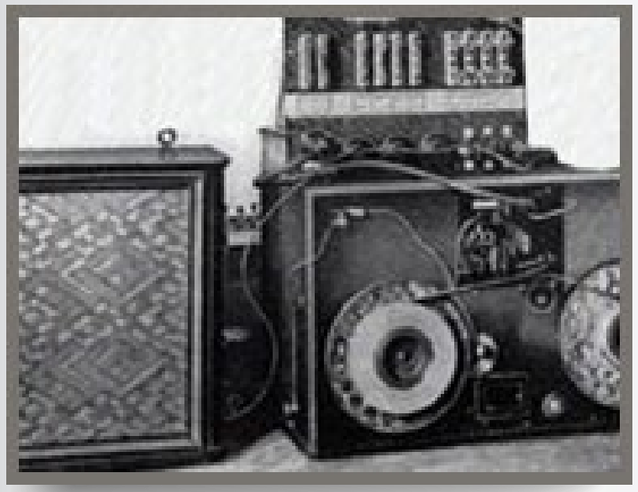
Evolution of Big Data

It is very important to understand the past of any given subject as it can provide you with a proper framework to understand and analyze whether the particular subject is a trend or just a fad. When looking up for the history of Big Data, I found tons of information available (thanks to Google!) what is Google, again? Big Data. Big Data can easily be considered the soul of every business enterprise in the modern day business world!! When some-thing is so important, learning about its history is even more important! The earliest record of using data to track and control businesses date back from 7000 years ago when accounting was introduced in Mesopotamia in order to record the growth of crops and herds.

In 1663, John Graunt recorded and examined all information about mortality roles in London. He wanted to create and spread awareness regarding the ongoing bubonic plague. He gathered his findings in the book "Natu-ral and Political Observations Made upon the Bills of Mortality". Through his book, he was able to provide great insights into the causes of death in the 17th century. He is even considered the Father of Statistics!



In 1880, U.S. Census Bureau data became a problem and it was estimated it would take 8 years to handle that data, and predicted the data from the 1890 census would take more than 10 years to process. Fortunately, in 1881, Herman Hollerith created the Hollerith Tabulating Machine. This machine was based on the punch cards designed for controlling the patterns woven by mechanical looms. His tabulating machine reduced ten years of labor into three months of labor. Herman Hollerith's impressive results earned him the contract to process and tabulate 1890 census data. Modified versions of his technology would continue to be used at the Census Bureau until replaced by computers in the 1950s. This was the beginning of the modern day data processing.



The 20th Century

Fritz Pfleumer, an Austrian-German engineer developed a means of storing information on tape. After various experiments with variety of materials, he settled on a very thin paper, striped with iron oxide powder and coated with lacquer for his patent in 1928.
After the Social Security Act came into law in 1937, the government had to keep track of contribution from 26 million Americans and more than 3 million employers. IBM got the contract IBM got the contract to develop punch card-reading machine for this project. It is considered to be the first major data project.

1943 marks the year of first data- processing machine named Colossus which was developed by the British to decipher Nazi codes during the World War II. It searched for patterns in intercepted messages at a rate of 5000 characters per second. This machine reduced the task from weeks to merely few hours.
National Security Agency (NSA) was created in 1952 and within 10 years contracted more than 12000 cryptologists. During the Cold War they were confronted with information overload as they were collecting and processing intelligence signals automatically.

In 1961, Derek Price published "Science since Babylon", in which he charted the growth of scientific knowledge by taking into consideration the growth in number of scientific journals and papers. He concluded that the number of new journals had grown exponentially rather than linearly. Price calls this the "law of exponential increase". He also explained that "each scientific advance generates a new series of advances at a reasonably constant birth rate, so that the number of births is strictly proportional to the size of the population of discoveries at any given time."

On Oct 29, 1969 ARPANET began when a message was sent from UCLA's host computer to Stanford's host computer. People were with a transatlantic satellite, linking it to the Norwegian Seismic Array. Infrastructure of ARPANET started to age by 1989. Other networks like NSFNET were more efficient and fast and a lot of organizations started to move to these networks. In 1990, the ARPANET project was shut down due to a combination of age and obsolescence. However, the creation of ARPANET led directly to the Internet.

United States Government decided to build the first data centre in 1965 to store over 742 million tax returns and 175 million sets of fingerprints by transferring all of the records onto a magnetic computer tape that had to be stored in a single location. Due to the fear of "Big Brother" the project was dropped, but it is generally accepted that was the beginning of the electronic data storage era.
With the introduction of Personal Computers in 1977, personal computers came on the market and became

a stepping stone in the evolution of the internet, and subsequently, Big Data. A personal computer could be used by a single individual, as opposed to mainframe computers. The introduction of microprocessor, prices of personal computers lowered significantly. Many of the early personal computers were sold as electronic kits, designed to be built by hobbyists and technicians. Eventually, personal computers would provide people worldwide with access to the internet.

A British computer scientist Tim BernersLee invented the World Wide Web in 1989. He wanted to facilitate the sharing of information via a 'hypertext' system. His invention had a huge impact on Big data in the following years. By the fall of 1990, he was working for CERN. He wrote three basic IT commands that are the foundation of today's web:

- **HTML:** HyperText Markup Language. The formatting language of the web.
- **URL:** Uniform Resource Locator. A unique "address" used to identify each resource on the web. It is also called a URI (Uni**formResource Identifier).**
- **HTTP:** Hypertext Transfer Protocol. Used for retrieving linked resources from all across the web.

In 1995, the first super computer was built, which enabled us to do as much work in a second as a calculator operated by a single person can do in 30000 years. This was the time when more and more people became aware of the speed at which tasks could be finished by machines and how convenient machines actually are.

The 21st Century

2005 was the year when the term "Big data" actually came into existence. O'Reilly Media coined this term and only after a year they created the term Web 2.0. The term refers to a large set of data that is almost impossible to manage and process using the old traditional business intelligence tools. It was is the same year when Hadoop was created by Yahoo! which was built on Google's MapReduce. The goal was to index the entire World Wide Web. Open source Hadoop is used by a lot organizations to crunch through huge amounts of data.

More and more social networks started to appear and the web 2.0 took flight which led to the creation of more and more data on a daily basis. Government started to work on Big Data Projects and a lot of innovative startups slowly started to dig into the massive amount of data. In 2009, the Indian Government decided to take an iris scan, fingerprint and photograph of all of the 1.2 billion inhabitants. **This data is stored in the largest biometric database in the world.**

Eric Schmidt spoke at the Techonomy conference in Lake Tahoe in California in 2010. He stated that, "there were 5 exabytes of information created by the entire world between the dawn of civilization and 2003. Now that same amount is created every two days."

2011 McKinsey report on Big Data: the next frontier for innovation, competition, and productivity stated that by 2018 the USA alone will face a shortage of 140000- 190000 data scientist as well as 1.5 million data managers.

There has been a massive increase in the Big Data startups in the past few years, all of which have been trying to deal with Big Data and helping organizations to understand the concept. A growing number of companies are moving towards Big Data. It is as far as the internet itself.

This was a brief history on Big data, and it is safe to say that the "Big Data Era" has begun. The large revolution is still ahead of us so a lot will change in the coming years.

The future of Big Data

Data is changing our world and the way we live at an unprecedented rate.

We are dependent on Big Data for so many things, almost every Big organization is working around it. If Big Data is capable of all this today, the future of only looks brighter. The amount of data available to us is only going to increase and analytics technology will become more advanced.

The ability to leverage Big Data is going to become increasingly critical in the coming years for businesses. In fact it is found by a recent study that "by 2020, the accumulated volume of Big data will increase from 4.4 zettabytes to roughly 44 zettabytes or 44 trillion GB."

To survive and thrive, it is particularly important for companies to view data as a strategic asset and also to ensure that your company is ahead in the market and is not left behind, you must keep the data revolution in mind!

2nd CHAPTER

Evolution of Data collection and analytics

"If you torture the data long enough, it will confess."

Ronald Coase, Economist

As we are well aware of the history of data, let us now consider the evolution of Data collection and analytics. Data in itself will be of no use until it is collected and analyzed. Data is everywhere and it is being used in production, marketing and sales of every kind of product/ service but it is important for us to recognize what kind of data is important enough to be collected and then analyzed. In this chapter we shall discuss the importance of Data collection and analytics and how it has evolved over the years.

Importance of Data collection and analytics

"Without Big data analytics, companies are blind and deaf, wandering out onto the web likedeer on a freeway."

–Geoffrey Moore, author and consultant

Data collection is the systematic approach for gathering and measuring information from a variety of sources to get a complete and accurate picture of an area of interest. Data collection enables an organization:

→ Answer relevant questions

→ evaluate outcomes

→ make predictions about future trends

Data analysis is a primary component of data mining and Business Intelligence (BI) and is key to gaining the insight that drives business decisions. The benefits of Data analysis are countless, to name afew:

→ creating more effective marketing campaigns

→ getting the right information for your business

→ gaining a better understanding of customers

→ evaluate outcomes

→ make predictions about future trends

Data analysis is a primary component of data mining and Business Intelligence (BI) and is key to gaining the insight that drives business decisions.

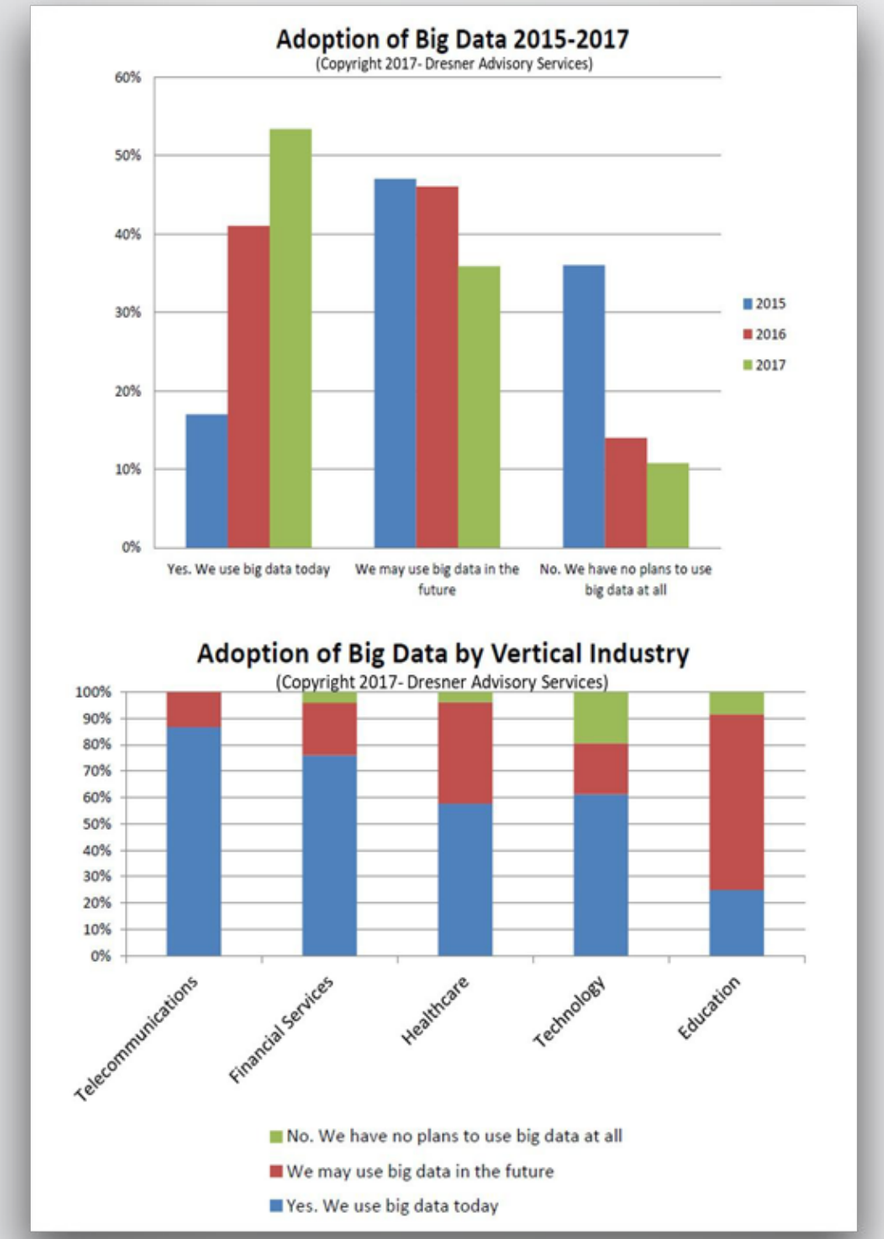
The benefits of Data analysis are countless, to name the few:

→ creating more effective marketing campaigns

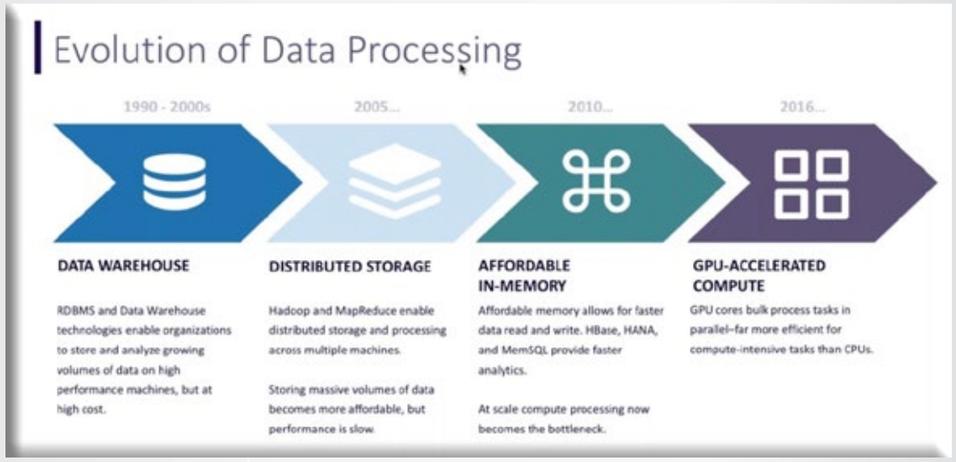
→ getting the right information for your business

→ gaining a better understanding of customers

When it comes to the growth of Big Data analytics, it is found that 53% of companies are using it today. Among all the industries, Telecom and Financial Services industries have been the fastest adapters! Technolo-gy and Healthcare comes next whereas Education has the lowest adoption as 2017 comes to a close. North America and EMEA have the maximum level of Big Data analytics adoption while Asia- Pacific respondents report 44% current adaption.



The evolution of Data collection and analysis



There has been tremendous amount of changes in Data over a period of time, particularly the amount, collection, types, and analysis of it. The future of the market greatly depends on the future of the data. As a result, it is important to understand how Data collection and analysis has evolved over the years.

According to the Harvard Business Review, there has been a shift in the way companies work over the past decade. The people who have been in business define their careers as BBD and ABD, or Before Big Data and After Big Data. The science of analytics has evolved to keep pace with the massive troves of available data. Here's an explora-tion of the field's evolution.

The history of Data collection and analysis is very long and tiring. There are various fields under section in today's time, but things were not the same when everything started.

Data collection was primarily a manual job. Analysts spent more time in collecting the data more than analyzing that data. The things were relatively easy when the amount of data was small in size, but as the data started growing it became difficult to process it. This need was felt the most during the 1880 U.S. Census. The data was huge and it was estimated it would take 7 years to process the Big data and to avoid this problem in the 1890 Census, Herman Hollerith came up with a solution. He invented a Tabulating Machine and It took only 18 months with the help of this machine to tabulate the given data. Data has been analyzed using different tactics and machines over the years.

1. Data Analysis and Statistics

Data analysis is rooted in statistics, which has a pretty long history. The beginning of statistics can be marked in ancient Egypt as it took a periodic census for building pyramids. Statistics has played an important role for governments across the world throughout history. it was used for creation of census, which were used for various governmental planning activities. The first step obviously was collection of data, as they should then be analyzed in order to discover beneficial information from that data. For example, analyzing the population of a particular area can help you understand the different problems faced by those people and determine the kind of solution that you must provide.

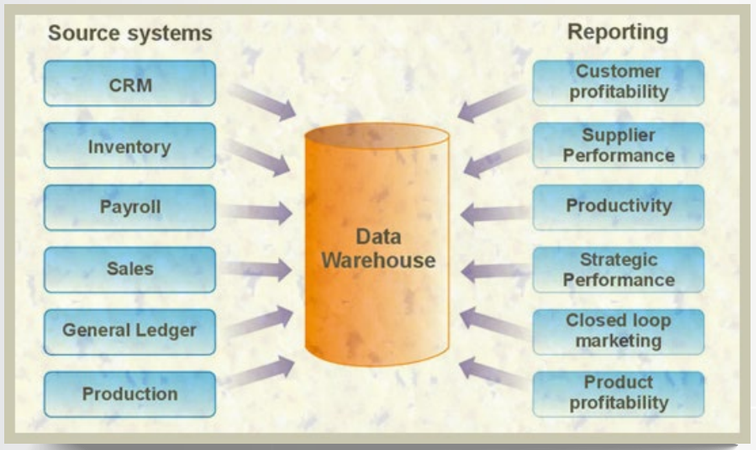
2. Data Analysis and Computing

The "Tabulating Machine" by Herman Hollerith helped a lot in systematically processing data recorded on punch cards. This machine helped in tabulating the results of 1890 US census with in a smaller budget.

The invention of Neumann architecture led to more emphasis on processing of data for data analysis. The appearance of RDB (relational database) in the 1980s allowed users to write Sequel (SQL) to retrieve data from a database. The users enjoyed the advantage of RDB and SQL and were able to analyze their data on demand. The process of getting the data became easy and helped in spreading the database use. The combination of easier/ cheaper data collection with cheaper/faster data storage/retrieval technology has pushed the boundaries of what we can do with data.

3. Data warehouse and Business Intelligence

The amount of data collected has increased significantly after the costs of hard disks drives decreased. It was around this time, William H. Inmon proposed a "data warehouse", which is a system optimized for reporting and data analysis. Data warehouse are optimized to align response times to queries. The data that is stored with a timestamp and operations such as DELETEs and UPDATEs are used less frequently.



For example, the sales trends of a business change each month, with these transactions can be stored timestamps within a data warehouse, and queried based on this timestamp.

The term "BI (Business Intelligence)" was proposed by Hoard Dresner at Gartner in 1989. It comprises the strategies and technologies used by enterprises for the data analysis of business information. The birth of the concept of B.I. helped in analyzing the customer data more systematically while making business decisions.

4. Data mining

Data mining appeared around 1990s. It is the computational process to discover patterns in large datasets. It is a different way of analyzing data, unexpected but beneficial results are expected. Database and data warehouse technologies made the development of data mining possible. which enables the companies to store more data and analyze it in a reasonable manner. Companies started to "predict" customers' potential needs bases on the analysis of their historical purchasing patterns.

5. Google Web Search

The internet is the biggest invention of our lifetime! It changed the way data is analyzed as well. Larry Page and Sergey Brin developed the Google search engine which assists in processing and analyzing Big data. Google engine responded with the results which you most likely wanted to see in just a few seconds. The key points of google search are that, it is "automated", "scalable", and "high performer. In the late 2000s, many open source software projects like Apache Hadoop and Apache Cassandra were created to analyze the Big data. Apache Hadoop, an Apache Software Foundation Project, is a Big data analytics framework that focuses on near-time and batch-oriented analytics of historical data. Hadoop helps run analytics on high volumes of historical line of business data on commodity hardware whereas Cassandra on the other hand is a very scalable and resilient database that is easy to master and simple to configure, providing neat solutions for quite complex problems.

6. Big Data Analysis on the Cloud

> "Everything is going to be connected to cloud and data… All of this will be mediated by software."
>
> – Satya Nadella, CEO of Microsoft

Amazon Redshift is a cloud based data warehouse and Google BigQuery, which processes a query in thousands of Google servers, were released in the early 2010s. According to a study by Deloitte, the key influen-tial reasons for the companies to move to a cloud was for include faster payback times (30%) and improved agility (29%).

These were cheaper and they lowered the hurdle to process Big data. These days, every company is able to get an infrastructure for Big data analysis within a reasonable budget. It is estimated that by 2020, a minimum one-third of all data will be stored and analyzed using cloud computing. Big data tools like Amazon Redshift has made it possible even for start-up companies to conduct data analysis within their reasonable budget.

The future of Data collection and analysis

As for the future, Data collection and analysis is revolutionizing with the help of machine learning and Artificial Intelligence. 76% of decision-makers surveyed foresee significant changes in the domain of storage systems because of the "Big Data" phenomenon. Companies like Netflix, Amazon, Facebook, etc. are making incredible profits with their top class tools and tactics in Data Collection and Data Analysis. As for now, only 0.5% of all accessible data is analyzed and used. There is a huge potential in the field of Data Analysis. More and more companies have started to believe in the power of data.

3rd CHAPTER

How Big Data Is Empowering Artificial Intelligence & Machine Learning?

How Big Data Is Empowering Artificial Intelligence & Machine Learning?

"Everyone is exposed to machine learning every single day. Do you use an Amazon Echo? iPhone's Siri? Netflix? Google? If you said no to Google, we know you're lying."

The world is taken by a storm from Artificial Intelligence and Machine Learning. Google and Facebook have placed huge bets on Artificial Intelligence and Machine Learning and are already using it in their products. Similarly, there are a lot of other companies that are heavily dependent on AI and machine learning. Let us first understand what they actually mean:

"Artificial intelligence (AI) is intelligence demonstrated by machines, in contrast to the natural intelligence displayed by humans and other animals".

"Machine learning is a field of computer science that uses statistical techniques to give computer systems the ability to "learn" (e.g., progressively improve performance of a specific task) with data, without being explicitly programmed."

AI and machine learning constitutes set of technologies that empower connected machines and computers to learn, evolve and improve upon their own learning by updating the data bank through recursive experiments and human invention in a consistent basis. Most of the companies these days use these two for improving their customer relationship as well as improving their product quality.

So, how really is Big Data empowering Artificial Intelligence and Machine Learning?

1. Big data enhances Artificial Intelligence

Machine learning is considered as an advanced application of AI in interconnected machines and peripherals by granting them access to databases and making them learn from it. As the size of the data increases, there is more information being fed to the machines. Big data analysis helps machines to have a more meaningful and contextually relevant ideas of their functions. Machines are automated with better and more insightful data.

For example, Walmart is on the edge of finding ways to transform retail and provide a better service to its customers. They use the Scan and Go feature on the app, Pick-up Towers and are also experimenting with facial recognition technology to determine if a customer is happy or sad. Clearly, Big data enhances Artificial Intelligence and that increases the customer satisfaction as well!

2. Improves machine learning

Millions of bits of information can be processed in seconds with the help of modern processors. Additionally, CPS (Calculations per second) rate of processors have raised with the aid of graphics processors. The Big Data

aids in processing of big volumes of data and renders rules for machine learning.

3. Helps in expanding AI and Machine Learning Workforce

One of the major concerns about having AI and machine learning is that it will minimize human requirements in all the job sectors. The truth though, is far from it when observed with the role to be played by Big data in the picture, it was found that the sentimental and emotional Big data analysis will always require human intelligence as the machines lack emotional intelligence and decision making abilities based on sentiments. The increasing collaboration between AI, machine learning and Big data will only pave a way for talented and capable human scientists to consistently evolve and rise in the market. More applications of these technologies only means more employment to more people.

4. Improved recognition algorithms

Technological advancements has made it possible to program AI machines in such a way that they can make sense of whatever we tell them almost as if they are humans. Big data is capable of empowering AI in the same way as it can form algorithms that human brain possesses. The voluminous data renders a broad base for building algorithmic analysis, which enhances the quality of AI perception. The best examples of improved recognition in AI are Alexa, HomePod, and Google Home.

5. Better AI and Machine Intelligence Solutions

The overall size of global markets for machine learning and artificial intelligence based solutions used to be limited, but with the increasing proliferation of Big Data Analysis devices and machines getting smarter they are able to perform in a better manner. There is a consistent improvement, enhancement and advancement in the AI and machine intelligence solutions. James Canton was a well- known Big Data and AI expert whoes analytical piece suggested that in the near future even the network nodes, chips, sensors and the software programs that will run IoT networks will be AI enabled via the cloud or at the chip or infrastructural level. Big Data enables humans to run this network with a digital brain or AI based solution that is smart enough to process, distribute and collect massive data.

6. Promote Open- Source Programming Languages

Big data has helped with scripting languages such as Python or Ruby which is excellent for statistical data analy-sis. Additional programming tools for data are utilized because of Big data.

The new developments in technology such as Big Data has enabled the scope and the future of AI soaring in new dimensions. The merging of Big Data analytics and AI has enabled us to create a highly efficient, reliable and dependable AI defined infrastructure.

7. Global diversification of AI and Machine Learning Offerings

The advent of technologies and increasing scales of production has lead to the fall in the prices of AI and machine learning based solutions. The third world nations can adopt these devices due to the fall in the prices. Considering the diverse cultural, religious, ethnic, linguistic and political affiliations and inclinations in these countries the AI and machine learning solutions that is provided to these markets will have to be trained and skilled differently from each other.

By taking into account the Big data analysis of different factors from region to region, the AI and machine learning solutions will be able to be trained and molded in a way that doesn't hurt the sentiments of different demographics.

8. Cost Effective and Highly Reliable Memory devices

Memory and storage determine the overall strength of computers and are also considered the most essential A Short History of Big Data components of a computing machine. For any device or even AI to work more smartly, efficient storage and quick retrieval of data plays a prominent role.

As per the latest technology, data is being accessed from cloud itself and it is not centralized in one computer. Memory devices like Dynamic RAMs and flash memories make use of information merely for processing and not for storage. Better surface realities can be attained with the aid of Big Data with more precise knowledge. With the better memory of information, higher AI learning will be imparted with reduced costs.

9. Big Data backs Market Analysis to boost AI and machine learning adoption

The market sector of artificial intelligence and machine learning based product is in its nascent stage right now. As the market grows, organizations will come to realize all sorts of product offerings, features and functionalities which are preferred by the customers. Big data acts as a formidable resource to enterprises developing and selling AI and machine learning based solutions.

You can gather the feedback, sales patterns and suggestions of a large section of customers through a marketwide and deep with Big data analysis. This kind of information will allow the companies to develop solutions and services that would be more suitable for the requirements of their customers.

Big data is like the backbone of machine learning and AI.

You must ensure that the quality of your Big data is of top notch quality, if you want to achieve the best results in your AI and machine learning tools.

Getting that best data all by yourself can be very hectic and time consuming, why not choose a partner for it? It is wiser and will get you ahead in the competition.

To get the most effective data, choose DataCaptive as your partner and forget about all of your data related worries!

4th CHAPTER

Application of Database Management System

"I think you can have a ridiculously enormous and complex data set, but if you have the right tools and methodology then it's not a problem."

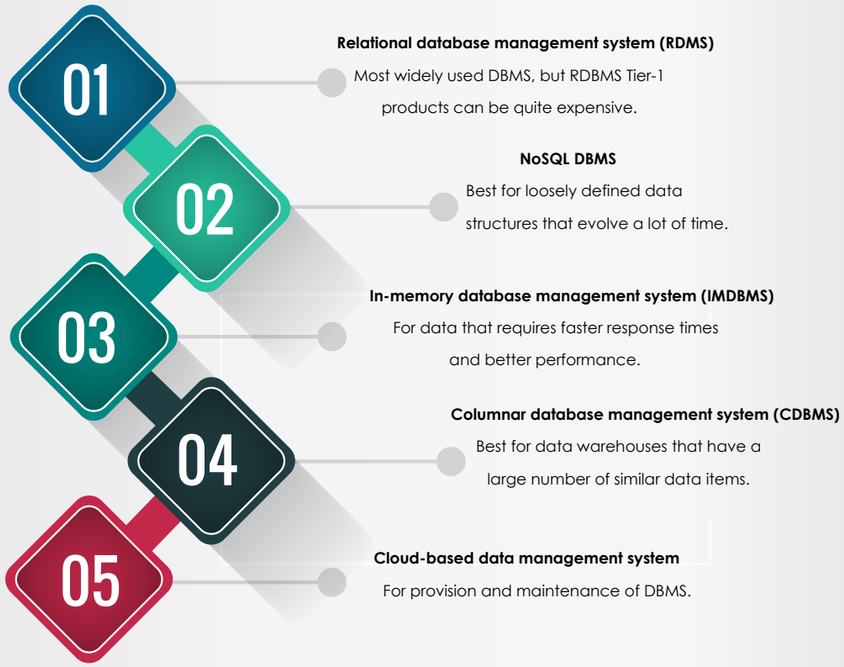
Aaron Koblin, Entrepreneur in data and digital technologies

We are well aware of different types of data, the evolution of data and the evolution of data analysis and storage, we have even covered about how machine learning and artificial intelligence are impacted by Big data. It is all essential for our knowledge, now it is time for us to cover the most important thing and that is "Database Management System". What is DBMS you may ask?

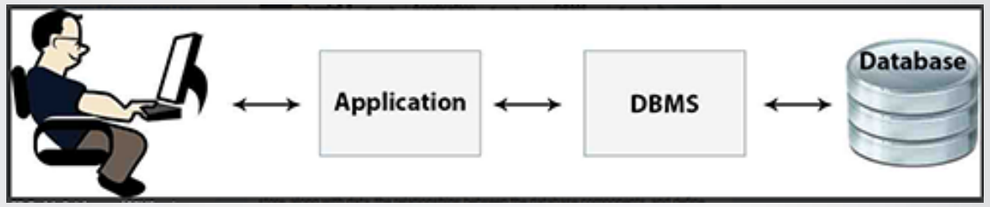
Well,

A database management system (DBMS) is system software for creating and managing databases. The DBMS provides users and programmers with a systematic way to create, retrieve, update and manage data.

The most popular types of DBMS



Applications of DBMS



1. Banking

Thousands of transactions take place through banks daily. People do not have the time and energy to visit the banks for every transaction, and the option of online banking and mobile banking has made the work so easy and fast. DBMS manages all the bank transactions .

2. Universities and colleges

Student registrations details, results, courses and grades all the relevant information are stored in institution's database. DBMS has made it possible for universities and colleges to carry examinations online and has helped a lot of students with distant learning options.

3. Transportation

Railway reservation and airline reservation is carried on by DBMS. These management tools keep the record of ticket book-ing arrival anddeparture station. It also informs about any delays that may happens before the journey or during the duration of the journey.

4. Human Resource Management

Big firms install these kind of DBMS applications in their firms which allows the Human Resource management department to keep the records of each employee's salary, tax, etc.

5. Manufacturing companies

It can be very difficult to keep the records of all the details about the product regarding their quality, bills, purchase, sales, etc. manually on a daily basis. A DBMS is the best solution!

6. Social media sites

Social media is so huge today! Every day, millions of people use social media sites like Facebook, Instagram, Snapchat, Twitter, etc. Lakhs of accounts are created on a daily basis and DBMS makes it possible for the users to store their information and connect to people from across the world!

7. Military

Millions of soldiers work for military and their data needs to be kept secured and safe. With the help of DBMS, one can easily search for all the information about any soldier within seconds.

8. Tele-communications

All the information about the customer's call details, pre- paid customers, post- paid customers, the bill payment records, etc. are stored in a DBMS. The tele communication cannot function at all without DBMS.

9. Elections

Elections are possible only because of DBMS applications. Everything works smoothly as it is automated. DBMS has made the hectic process of counting the votes so easy and fast.

10. Online shopping

Online shopping has become so popular these days and it is possible only because of an effective DBMS. Purchase information, invoice bills and payments, all of these are done with the help of DBMS.

These are just the few of the industries where DBMS is used but it is utilized in almost all the industries. Though 85% of companies are trying to be data-driven, only 37% of that number say they've been successful. There is a huge scope for Database Management System and it can easily be considered as the most important component of an enterprise.

5th CHAPTER

How Companies Are Using Big Data Analytics

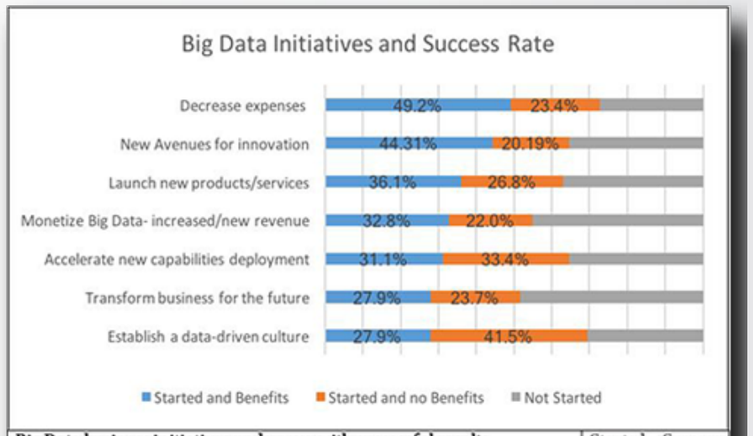
How Companies Are Using Big Data Analytics

"Before Google, and long before Facebook, Bezos had realized that the greatest value of an online company lay in the consumer data it collected."

George Packer, author for the New Yorker

Companies can harness the power of Big Data and find the set of tools to use it in an efficient manner to generate the maximum value out of it. Adopting a data- driven mindset and clearly identifying the business challenges in addressing the data analytics can provide the success that businesses aim from Big Data.

According to a report from NewVantage Venture Partners, Big Data delivers the most value to enterprises by decreasing expenses to almost 50% and creating new avenues for innovation and disruption to 44.3%.



Big Data business initiatives underway; with successful results.	Started	Success
Decrease expenses through operational cost efficiencies	72.6%	49.2%
Establish a data-driven culture	69.4%	27.9%
Create new avenues for innovation and disruption	64.5%	44.3%
Accelerate the speed with which new capabilities and services are deployed	64.5%	31.1%
Launch new product and service offerings	62.9%	36.1%
Monetize Big Data through increased revenues and new revenue sources	54.8%	32.8%
Transform and reposition your business for the future	51.6%	27.9%

Businesses use information derived from data to increase its efficiency and success in automating processes, gaining in-depth knowledge of the target market, etc. In this chapter we shall cover the five businesses who have gotten the maximum out of their database management systems.

Companies that use Big Data Analytics

Amazon

"Amazon.com Inc is a leader in collecting, storing, processing and analysing personal information from you and every other customer as a means of determining how customers are spending their money. The company uses predictive analytics for targeted marketing to increase customer satisfaction and build company loyalty."

– Jennifer Wills: Owner of JDW Writing

Amazon is an online retail giant that has mastered the art of ecommerce. Amazon manages a marketplace platform that leverages Big data and a customer-centric focus to improve the customer experience and internal operations. The information is obviously put to use in advertising algorithms, however, Amazon also uses the information to improve customer relations.

The next time you contact the Amazon help desk with a query, don't be surprised when the employee on the other end already has most of the pertinent information about you in hand. This allows for a faster, more efficient customer service experience that doesn't include having to spell out your name three times.

Amazon is the one of the most successful companies in the worldand the founder Jeff Bezos is the richest man alive, lleaving behind the Microsoft founder Bill Gates. Let us check some of the most interesting and fascinating facts about Amazon

1995 — 1.Jeff Bezos founded Amazon in 1995. It was initially founded to sell books

2.The logo has an arrow that moves from A to Z, meaning this company sells all the products from A to Z. to sell books

3.Jeff Bezos chose the name Amazon after the largest river in the world Amazon.

2000 — 4.The company faced a lot pf problems and growth was very slow until the early 2000s.

5.Bezos is known to be a demanding boss and could explode at employees. Rumor has it, he hired a leadershipcoach to help him tone it down.

1999 — 6.The firm expanded into selling toys, electronics,tools,and hardware; Bezos was named Time Magazine's "Person of the Year" in the year 1999.

7.Amazon introduced kindle. The device has undergone series of design changes and it is considered the world's best e- reader device

2008 — 8.In 2008 Amazon.com purchased Audible.com, AbeBooks.com, BoxOfficeMojo.com, Shelfari.com, Reflexive Entertainment, etc.

2009 — 9.In 2009 further acquisitions occurred such as Zappos, Lexcycle.com, and SnapTell.com.

2017 — 10.In 2017, Amazon introduced Alexa which revolutionized the game of Artificial Intelligence. It is avoice assistant and can help us with almost everything!

How was all of this possible? Well, Big data. Are you tempted to try your luck with Big Data Analytics? Visit https://www.datacaptive.com/ and get your hands on the best and most updated data available in the market.

Netflix

> “I don't have a television. All I have is Netflix and Apple TV and Hulu.”
>
> Bojana Novakovic

Netflix has over 100 million subscribers and with that comes a wealth of data they can analyze to improve the user experience. Big Data has helped Netflix massively in becoming the king of on-line streaming. The entertainment streaming service has abundant data and analytics providing insight into the viewing habits of millions of international consumers.

“From predicting the kind of content that would garner high viewership to recommending content to specific users, Netflix uses data everywhere. In fact, since its days of being a DVD-by-mail service, Netflix placed prime importance on collecting user data and building a recommendation system. Cinematch was the first algorithm behind their recommendation system. After launching their streaming media service in 2007, it took them 6 years to collect enough data to predict the sure-shot success of their first original production ‘House of Cards’. Data accumulated from numerous sources influence decisions regarding shows. Not only user data, Netflix also observes data generated by piracy sites. “Prison Break” is a hit show on that front.” – Toai Chowdhury: Author at upX Academy.

According to Netflix, its algorithms save $1 billion a year from its customer retention. This figure is eye opening for the biggest household names in entertainment. The algorithms are constantly evolving to give the best possible user expe-rience.

1.Netflix is the world's leading internet entertainment service with 130 million memberships in over 190 countries. with almost everything!.

2.Reed Hastings and software executive Marc Randolph co-found Netflix started to offer online movie rentals in 1997.

3.Netflix's initial business model included DVD sales and rental by mail, although Hastings jettisoned DVD sales about a year.

4.It was originally called Kibble Directpix.com, Replay.com, and others were considered.

5.Netflix held a million-dollar contest in 2006. The challenge was to create a new, perfected video content recommendation algorithm.

6.Netflix introduced streaming, which allows members to instantly watch television shows and movies on their personal computers in 2007.

7.The first video content ever put on Netflix was a test film created to inspect standard film frames and was known as "Example Show".

8.The most watched show on Netflix are Breaking Bad, Orange is the New Black, The Crown, The Walking Dead, Stranger Things.

9.Netflix has its own awards — The Flixies. It's an award show where users are invited to vote on their favorite shows.

10.As of July 2018, Netflix had 130 million total subscribers worldwide, including 57.38 million in the United States.

Facebook

"I get a 100% open rate,' said nobody ever. All the more reason to place Facebook ads targeting your own email list."

– Mari Smith, social media expert

With more than two billion monthly active users worldwide, Facebook is easily the most popular social media network! It was estimated that "Every minute Facebook users send roughly 31.25 million messages and watch 2.77 million videos."

It is also the fifth most valuable public company in the world, with a market value of approximately $321 billion. Facebook is a massive data wonderland and it knows what we are doing, our likes, our dislikes, and so much more.

Have you observed how Facebook shows a "flashback" of posts, likes, or images? This is one of the example of how Facebook uses Big Data. Apart from Google, Facebook is probably the only company that possesses this high level of detailed customer information. Let us look at ways Facebook determines the user behavior apart from the analysis of user data:

1. Tracking cookies: Facebook tracks a user who is logged into Facebook and simultaneously browses other websites.

2. Facial recognition: Facial recognition and image processing capabilities is one of the and it latest investment.

3. Analyzing the Likes: One study found that by analyzing the user's Facebook Likes one can accurately predict a range of personal attributes of a user. According to a research conducted at Cambridge University and Microsoft Research, Likes can accurately predict your sexual orientation, emotional stability, alcohol and drug use, etc.

"Big Data is crucial to the company's very being. Facebook relies on a massive installation of Hadoop, a highly scalable open-source framework that uses clusters of low-cost servers to solve problems. Facebook even designs its own hardware for this purpose. Hadoop is just one of many Big Data technologies employed at Facebook."

Facebook Inc. analytics chief Ken Rudin

1.Facebook is an American online social media and social networking service company launched on February 4, 2004.

2. Founders are Mark Zuckerberg and his roommates Eduardo Saverin, Andrew McCollum, Dustin Moskovitz and Chris Hughes.

3. The actor Al Pacino was the first "face" on Facebook. It was a header image featuring Al Pacino's face obscured behind binary code.

4. Peter Thiel, co-founder of PayPal, was the first major investor to back Facebook. He invested $500,000 into theyoung company in 2004.

5.Videos are the most shared content type on Facebook. On an average, Facebook users watch a collective 4 billion videos per day.

6. Sean Parker, co-founder of now-defunct music sharing site Napster, originally acquired the facebook.com domain name for $200,000.

7. Facebook stores approximately 300 PETABYTES of user data on its servers. There are 1 million gigabytes in apetabyte.

8. Posting just once or twice per day on Facebook yields an average of 40% more engagement than posting three or more times per day.

9. Facebook's user base grows by eight people per second, or 28,984 people every hour.

10. Facebook has grown by acquiring the likes of WhatsApp and Instagram, adding features that broaden the company's reach and appeal.

Want to reach your market on Facebook? Well, get the contact lists of your targeted industries and go ahead!

APPLE

"I think if you do something and it turns out pretty good, then you should go do something else wonderful, not dwell on it for too long. Just figure out what's next."

"Think different" has been the slogan for Apple. It holds so true with the policies and tactics that Apple has used in its company. While Google had been using Big data way before Apple, it didn't take much time to catch with the Big Data trend.

Apple has loads of data on how customers use their iPhones, MacBook, and iPods, etc. which gives them valuable information when designing new products or the latest versions of existing devices. The Apple Watch in particular has the potential to revolutionize data gathering. It has sensors to collect information such as how many steps a person walks, their temperature, how much they eat, and numerous other factors. There is a lot of potential in this technology.

Apple has joined hands with IBM to get the most out of digital health information. It hopes to cure diseases and help improve the lifestyle of people using the Big Data analytics. The partnership is also looking to use Big data to create health-related mobile apps, which can be used with Apple devices. Recent studies predict that

healthcare could save up to $300 billion a year by improving the integration of Big Data, on an average can reduce the costs by $1000 a year for each person.

There is use of Big Data in iTunes and iCloud. One can safely say that Apple is the company that depends vastly on Big Data for most of its strategies and technical developments.

1. Apple was founded by Steve Jobs, Steve Wozniak, and Ronald Wayne in April 1976 to develop and sell Wozniak's Apple I personal computer.

2. The name Apple was chosen because it was the favorite fruit of the late Steve Jobs.

3. Apple's first company logo featured the Father of Physics, Sir Isaac Newton.

4. If Apple was a country, it would be the 27th richest country in the world between Venezuela and Belgium.

5. SIRI, a voice-activated personal assistant that now lives in all of our iPhones has been acquired by Apple in February 2010.

6. It cost $666.66 to build the first Apple computer. There has been speculation of conspiracy theories behind the pricing but none have been proven.

7. Apple overtook Google in the list of the World's most valuable brands in 2015.

8. Apple has patent for 'slide-to-unlock' feature and it sued Samsung for infringing on this patent.

9. The list of Apple products that are going to be launched are: Mac Pro, iPhone XI, Apple Watch Series 4, AirPower, HomePod.

10. Apple is known for it's famous slogans. Few of them are:

- **i phone: It's not an iPhone if it's not aniPhone.**
- **MacBook: Light Years Ahead.**

Starbucks

"Starbucks represents something beyond a cup of coffee."

Howard Schultz

A Short History of Big Data

With around 90 million transactions a week in 28,039 stores (as of January 2018) of this coffee giant worldwide is in many ways on the cutting edge of Big Data and Artificial Intelligence to help sales and decisions related to product modification and product development.

Starbucks launched its reward programs and mobile apps to further amplify its customization. The mobile app has more than 17 million users and the reward program has 13 million active users. These users create an overwhelming amount of data for Starbucks! The app also has a feature that suggests the users about what kind of coffee they might be interested in trying. The app also allows you to place an order through voice command using artificial intelligence algorithms behind the scenes.

When Starbucks is in the process of determining its location, it taps into the power of data intelligence through Atlas, a mapping and business intelligence tool developed by Esri. This tool evaluates all the necessary details and helps Starbucks in its decision making.

When the company decided to expand and provide its customers with their products in grocery stores they used the power of Big Data to determine the kind of product they would develop and modify in order to achieve the maximum customer satisfaction.

The experience is really the result of all the data, Artificial Intelligence they have amassed in so many years.

1. Starbucks Corporation is an American coffee company and coffeehouse chain. Starbucks was founded in Seattle, Washington in 1971.
2. Starbucks is the largest coffeehouse chain worldwide, with revenues of 22.39 billion U.S. dollars in 2017.
3. Starbucks purposely have round table in order to make the customers feel less lonely.
4. Pumpkin Spice Latte which 101 shots of espresso is the most expensive drink ever with cost $93.58.
5. Starbucks cofounder Gordon Bowker said he and fellow cofounders were to go with Piquod and Cargo House before they finally settled on Starbucks.
6. There are 87,000 combinations of beverages available in Starbucks according to Starbucks spokesperson Lisa Passe (as of 2017
7. Aroma is so crucial to the Starbucks experience, Schultz banned perfumes and smoking in the baristas.
8. The siren of the famous Starbucks logo is intended to represent the seductive power of coffee.
9. Starbucks baristas undergo over 30 hours of training on everything from Frappuccino technique to the origins of coffee bean.
10. Few of their catchy slogans are:
    - You are what you drink. So what do you want to be today?
    - Because that's your name right next to ours
    - It's not just a coffee. It's Starbucks
    - Coffee that inspires

How exactly can you get the maximum benefit out of Big Data?

The stories of these companies are so inspiring, right? The main factor of these companies has been Big Data. The White House has invested a $200 million in Big data projects. This figure in itself is enough to prove the potential of Big Data, right? However, all of these companies put their customers first and they try to find as much as they can about their tastes and preferences and strive to give their tastes and company a better experience with their brand.

Do you want to get the maximum out of your Big Data? If so, get your data game strong. Let me help you with few essential steps:

1.Decide on the kind of industries you want to target, set objec-tives for yourself and decide on the kind of industry you want to target at! It can be under SIC code, NAICS code, Technology users or any kind of industry like the agriculture, education, etc.

2. Choose your targeted audience

Be clear about the type of audience you want to target and accordingly you can formulate your market strategies.

3. Choose a Data vendor

In this fast paced world, it is rather effective to choose a third party for your contact lists. Choose the data vendor that provides the best and most effective solutions!

DataCaptive

4. Choose DataCaptive

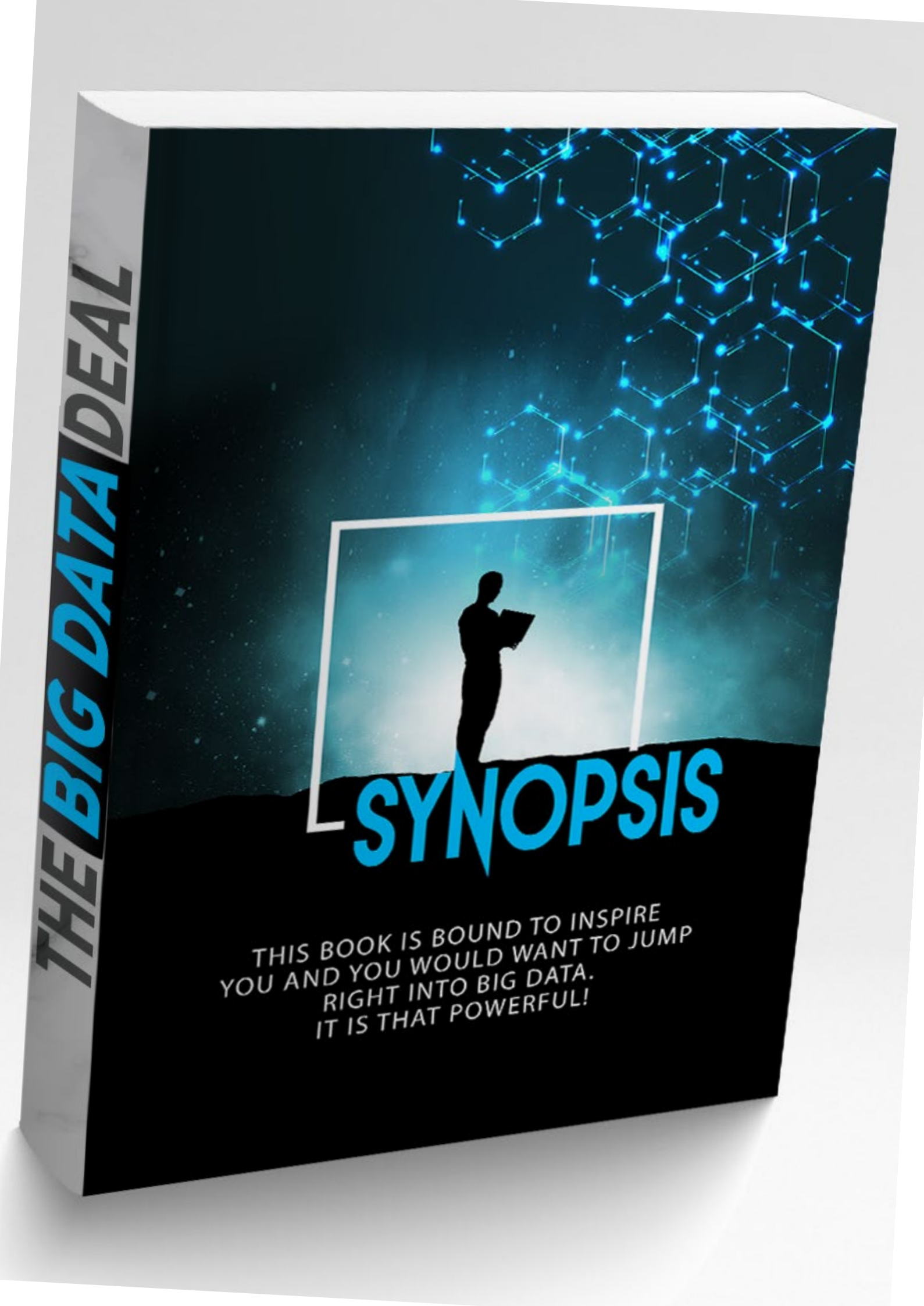
To get the best in the market, choose the best in the market. Say goodbye to your data related worries forever. Join hands with us, we are proud to help you!

Synopsis

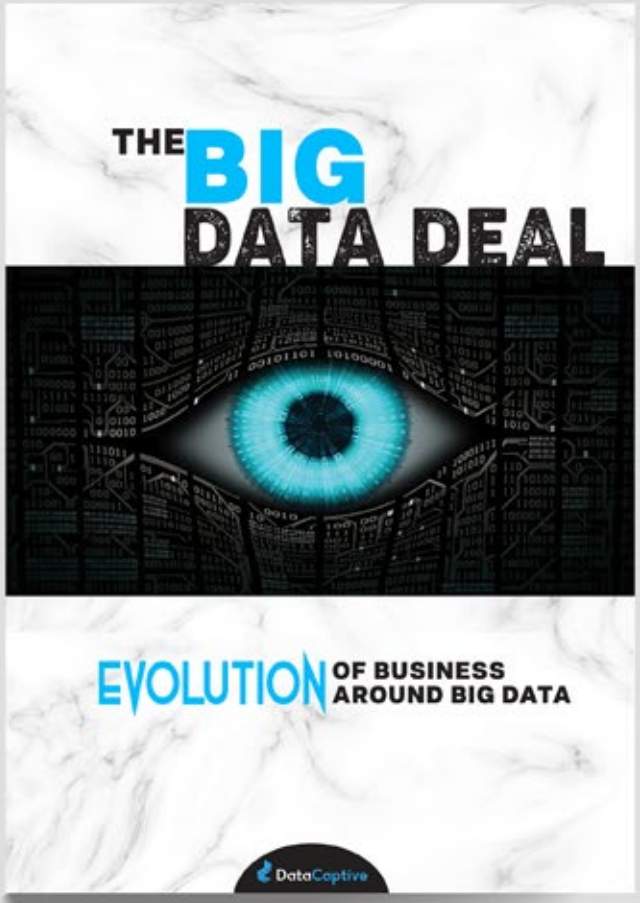
The purpose any E-book is to provide a thorough knowledge about a particular topic. This book is specially designed with great delight and patience to give its reader a thorough knowledge of evolution of businesses around Big Data. From various sources, I have collected the most relevant and updated information on this topic and wish to give my readers the most important information on the same.

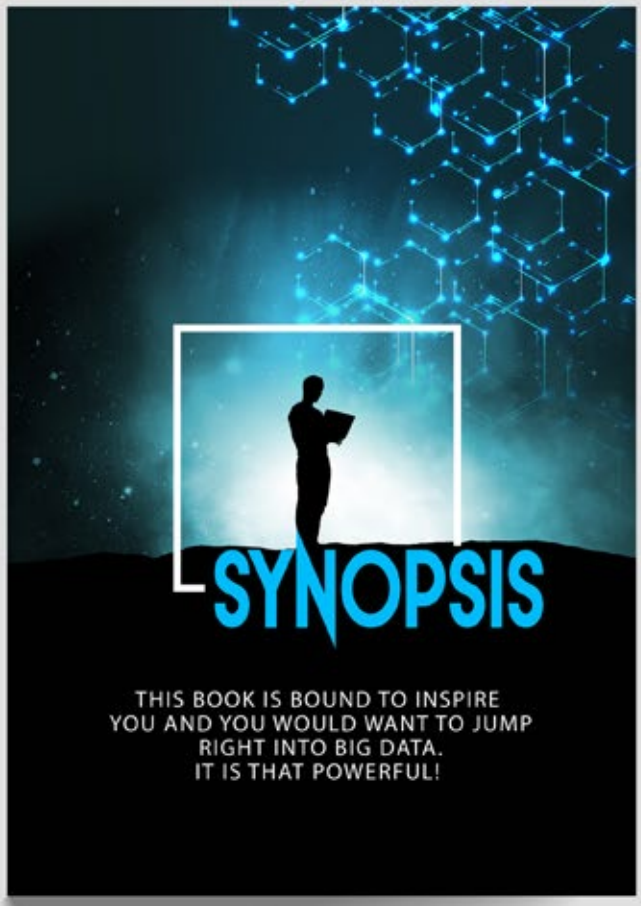
It clearly defines and provides relevant information:

- The real meaning of Big Data, its importance, the different types of Big Data
- The history of Big Data: right from how it started and how it is used in the present moment
- The Big Data storage and analysis: the past, the present, the future
- The different uses of Database Management System
- The different companies that revolutionized the uses of Big Data and grew into massive unimaginable empires!


THE BIG DATA DEAL
SYNOPSIS
THIS BOOK IS BOUND TO INSPIRE
YOU AND YOU WOULD WANT TO JUMP
RIGHT INTO BIG DATA.
IT IS THAT POWERFUL!

THE BIG DATA DEAL
EVOLUTION OF BUSINESS AROUND BIG DATA
DataCaptive


THE BIG DATA DEAL
EVOLUTION OF BUSINESS AROUND BIG DATA
DataCaptive


SYNOPSIS
THIS BOOK IS BOUND TO INSPIRE
YOU AND YOU WOULD WANT TO JUMP
RIGHT INTO BIG DATA.
IT IS THAT POWERFUL!