



IMPORTANCE OF DATA SCIENCE IN VARIOUS FIELDS

Navneet kaur (1)

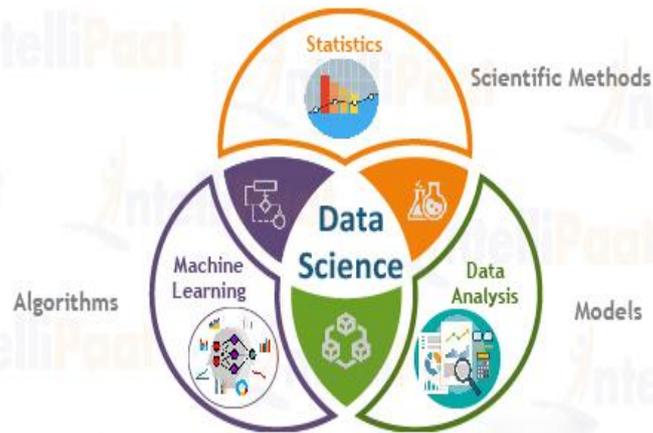
Gurpreet Kaur(2)

ABSTRACT

There is a lot of importance of data science in today's era. Data Science is something which is related to data. Today, Data science is used in various fields. Such As It is used in fields like Business, Medical Field, Retail, Sports And Much more. Most of the big companies use Data science to analyse the data such as Google, Amazon Etc. Various companies use and collect data various kind such as Raw Data, Structured Data, Unstructured Data. After that there is proper analyse of that data by the company. After by Graphs, Diagrams, Pie Charts they represent the data. It is very difficult to Handle a lot of big data. Also, It is not possible to analyse the whole of the data individually Because data can be raw, Structured, Non Structured, In the form of pictures, Mails Etc. So to handle such a Big data, Various companies use Various Analytical Tools to analyse that big Data. Also, there is a difference between Data science and Data Analysis.

INTRODUCTION

Data science is defined as the detailed and deep study of the data of the big data or information collected by the company in the form of Raw, Structured and Unstructured data to get the useful data by using various methods, tools, Skills and techniques. It is a multidisciplinary field.



It includes a lot of skills such as Mathematics, Statistics, and Knowledge of business domain Which helps to find out the different ways of:

- Reduction in Cost.
- Launch of the New product Or the service by the company.
- Coming into the new market.

Different companies such as Google, Amazon and Visa are using Data science to get the Fruitful Result. People says that Data science Is a complex field But according to me there is no as such thing that data science is complex. Data science is basically A kind of study which deals with Interpretation, Analysing as well as representation of the Meaningful data that have been collected from Raw data which is used by big Business companies For more efficiency of their business. There are various fields where Data science is being used such as Healthcare, Credit and insurance, Marketing , Sales, Automation etc and many more. Data science help a lot in all these fields. Data science help the various companies to solve out their problems as well to do a better decision making. This will help to increase in the profitability.

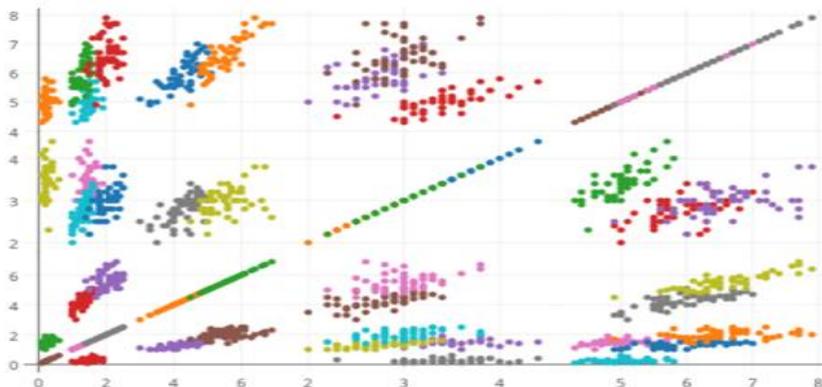
Data science is being used by most of the companies. The reason is that it help in making better decisions as well as in empowerment. Also it help in analyse the data which is being collected and this data help in taking the relevant solution of that particular problem. This will lead to increase in the Profitability as well increase in the performance of the business. Also, It help in Productivity within small as well in MNC enterprises. The data collected By the companies is in the form of Raw Data, Discrete Data, Qualitative as well as Quantitative Data.

Today, There are various analytical tools which are required for analyse as well to use that data for future. These tools are used when there is huge data collected by the companies. The

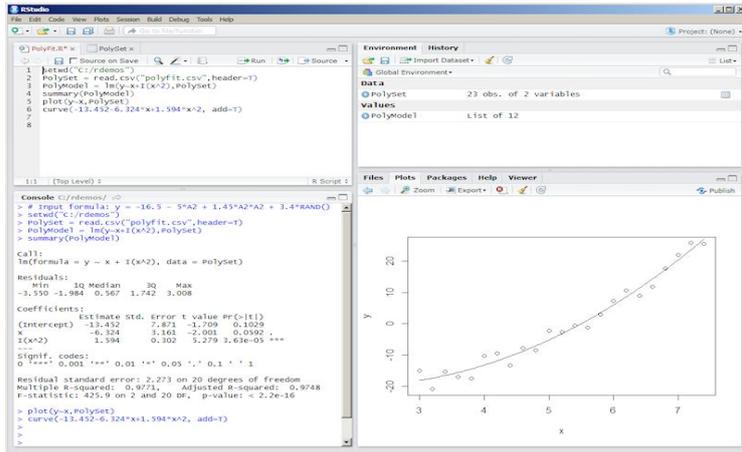
data they collect can be in the form of images , Locations, Objects, messages, Texts Or voices also. Furthermore, also there are various Methodologies of the big data analytics Such as Data Collection and Identification, Storage of Data, Classification of data and Extraction (Removal of the useless Data and decrease the volume), Data Cleansing as well as aggregation, Processing of data and analyse that data. The Last Phase is the Data visualisation. There are various tools which are used to analyse this Big data Such as web, cloud, Excel, Hadoop, Oracle, MangoDB, Map Reduce, Talend, Flink, Tableau. These are various tools which are used for above Methodologies.

Companies collect a lot of data from various sources and it is very difficult to Know about which data is valuable. So, Data scientists use various tools to solve the problems and come up with a accurate solution. These tools as be open source as well as paid tools. These Following are the best analytical tools as:

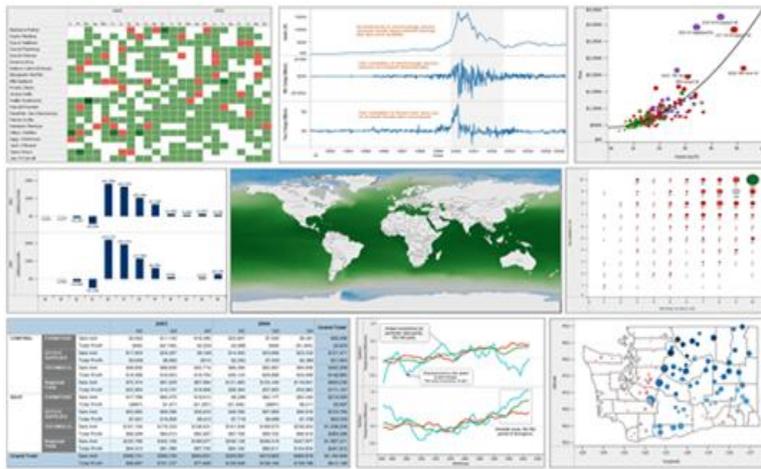
1) Python



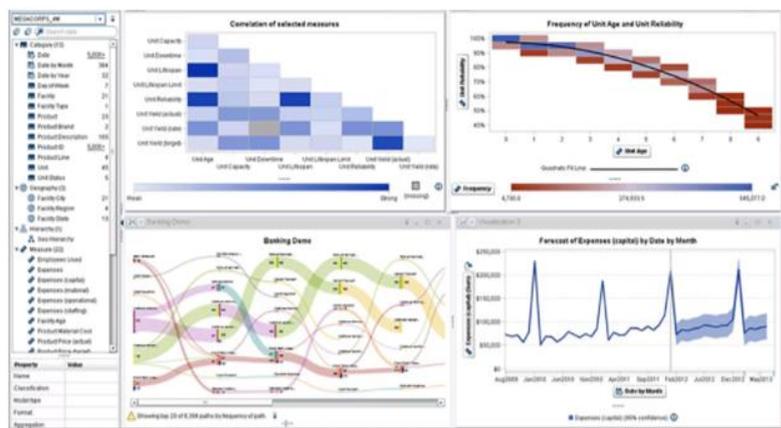
2) R Programming



3) Tableau



4) SAS



Review of literature

- The Paper “Big Data Analytics Tools” By Ms. Komal Tells about the various analytical tools which are required for analyse as well to use that data for future. These tools are used when there is huge data collected by the companies. The data they collect can be in the form of images , Locations, Objects, messages, Texts Or voices also. Furthermore, It also tells about various Methodologies of the big data analytics Such as Data Collection and Identification, Storage of Data, Classification of data and Extraction (Removal of the useless Data and decrease the volume), Data Cleansing as well as aggregation, Processing of data and analyse that data. The Last Phase is the Data visualisation. Thereare various tools which are used to analyse this Big data Such as web, cloud, Excel, Hadoop, Oracle, MangoDB, Map Reduce, Talend, Flink, Tableau. These are various tools which are used for above Methodologies.
- The Paper “Applications or Uses of Data Science Today” By Vidhya Tells us about the most common used applications are used in our daily life. Today companies are enough Intelligent to Sell up their products according to the needs and demands of the customers.
- The paper “Challenging tools on Research Issues in Big Data Analytics”By Althaf Rahaman , Sai Rajesh., .Girija Rani Tells that in this modern World, There is huge amount of data which is collected by the companies. The Big Data which is collected by the companies is of various varieties. It is very difficult to handle such a varieties of data by using the technical tools as well as methods. With the rapid increase in the growth of these data sets, solutions are required to be studied so as to come with the correct value as well as Knowledge out of these data bases. During the Analysis of these data, a lot of knowledge as well as a lot of efforts are required to come up with a correct decision making. So There are various Challenges and problems which are faced during The analyses of the Big data. Also Different Techniques are used to solve these problems.
- The paper “Data Science and Prediction” By Vasant Dhar Examins What is Data science. As Big Data, Data science also become increasingly common. Data Science is not only focus on the data but also study the various aspects of the organisation, Its properties as well roles. So this paper also tells about Predictive Model with its Limitations as well as their advantages.

- The Paper “Data science Applications and Uses cases” By Ekope Tells about Understanding the various challenges of big Data. Further It tells About what Data Science actually mean and what a data scientist do? It involves various case studies and cases such as Cancer Research, Health Care Research, Elections Case Study and Customer Analysis. There are some Important point which are mentioned in this paper and which are very useful and which will be further discussed.
- The paper “ Importance of Data Science” By Bobby Horvath Tells about The importance of the Data Science. It basically examines three Major Importance of the Data Science. First is It helps in making Decision by the management. Second is It helps in fulfilling the Actions on the basis of the latest Trends so that the organisations can achieve their goals. Third one is that the past data help in making decisions.

OBJECTIVES OF THE STUDY

- To know about what is data science.
- To study about the various tools and techniques used in various organisations.
- To study about the importance of data science in various fields.

Research Methodology

Data source

- **Primary Data**

Primary Data is the data Which is originally collected rather than studying or reading something which is already available. The data which is obtained from the such kind of resources are called Primary data. The data we collected is in the Raw form. You have to make it suitable for yourself by various processes so that it can be used further for analyses and gives us a great conclusion out of that. It includes surveys, Questionnaire, Interviews etc.

- **Secondary data**

I have collected information through secondary data sources. I have read E- Research papers, Weblogs, General Websites etc. I have also collected my data from various websites. These Sources Help Me a lot to get information. The major advantage of this data was that it was easily assessable and It is inexpensive. Also, It provided me various alternatives

DETAIL OF THE STUDY

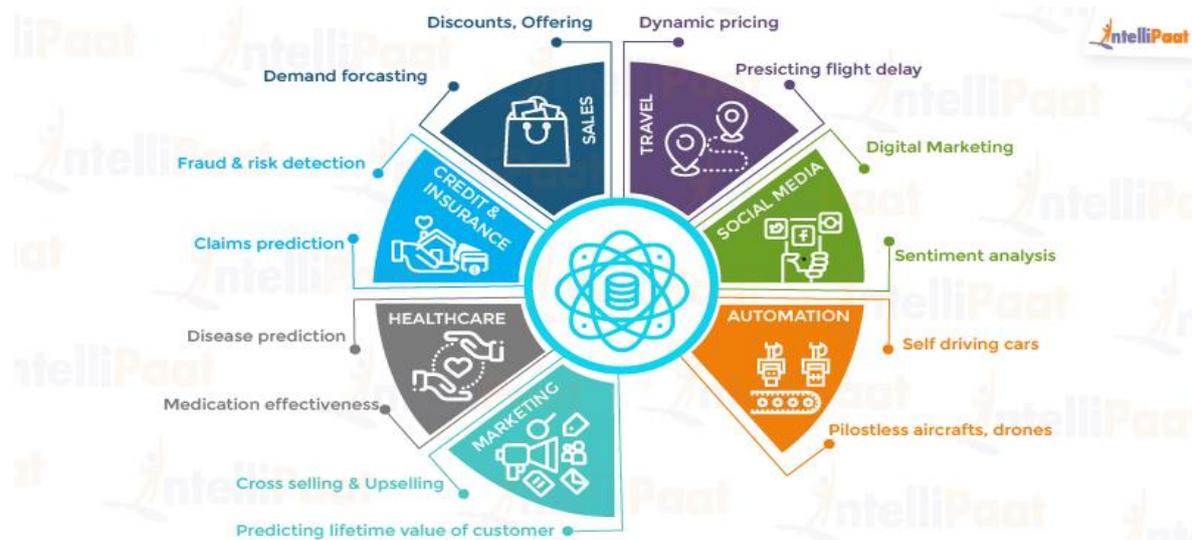
Data science is defined as the detailed and deep study of the data of the big data or information collected by the company in the form of Raw, Structured and Unstructured data to get the useful data by using various methods, tools, Skills and techniques. It includes a lot of skills such as Mathematics, Statistics, and Knowledge of business domain. Data is also known as big data. Companies are carrying a huge amount of data. So it is necessary to know that what they do with that data and how they use that Data.

So, At this point of time the data science various roles had came into picture. For this you required the skills and by take all the skills together like Mathematics, statistics and knowledge of business domain Helps to find out the way by:

- Reduction in Cost.
- Launch of the New product Or the service by the company.
- Coming into the new market

So with regard to this, Data Science plays an important role in the success of the organisation in the future.

So look at the info graph which is given below which will help us to understand how the Data science Creating its Impression(Effects):



1) Credit and insurance

In this competitive market, the most competitive industry is the insurance industry. This industry is one of the risk related industry. Therefore This industry is always dependent on the statistics. Because of Data science its Dependence has been changed forever.

Now , These Insurance companies has a large number of sources of information For their Relevant Risk Assessment. Big data Technologies are being used For Predicting the risk and also analyse them to develop the effective Strategies For the attraction of the customer and retain their attention. Also the insurance companies have got a huge profit from the data science application Inside the field of their great interest.

Following are the activities in the insurance industry:

- Fraud Detection

There is a huge financial loss due to the Insurance fraud every year. With the help of data science Software's and platforms make it possible for finding the Fraudulent activities, Suspicious links With the help of using the using various techniques. Basically, The insurance companies are using the statistical model for the Proper Fraud Detection. Also, Preddective models rae used here. These Predective models rae used to analyse as well as Seperate out the frauds. It also helping in the Recognition of the fraud schemes which are not notice previously.

Claims Prediction

2) SALES

Data science is known as one of the biggest trends in the business. There is a direct link between Data science and sales. Data science programs are helping to the sales leaders for running their business more efficiently and effectively and focus their major efforts on the Right sales aspect (Probability of future success) as well as on the Non covered missed opportunities. Mainly the data science helps in increasing the sales Processes and help in making the good strategic decisions.

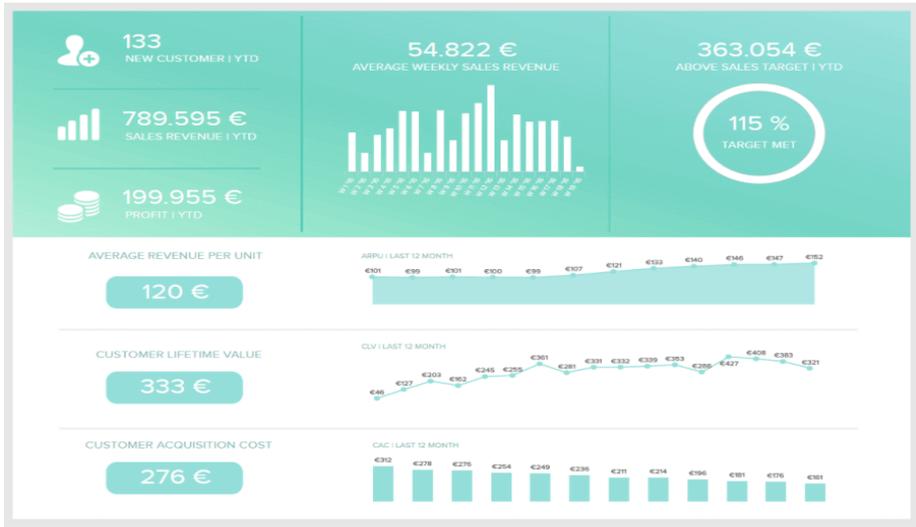


Table-1

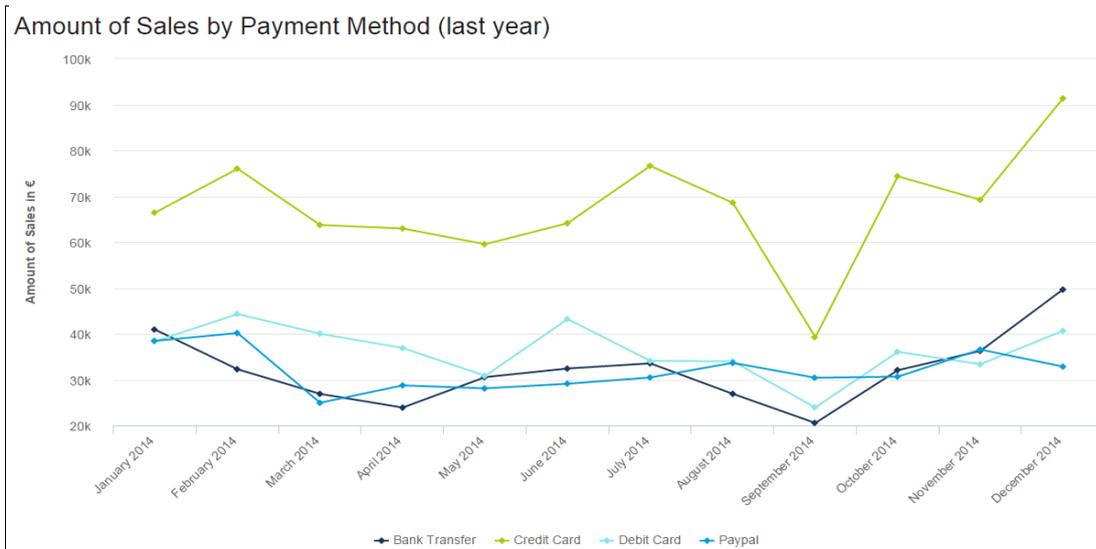


Table-2

So the figures i.e. Figure 1 and 2 shows the graphical representation of the data came after analyses. In figure 1, There is an overview of the sales performance as well as shown that how much customers you have got in this year. Also it includes the revenue(in Euro) You have been generated. Figure 2 Represents that the amount of the sales by the payment methods in the last year. This line graph represent the Amount of sales done by the various payment methods such as PayPal, Banking transfer, Credit card, and debit card.

The techniques in the data science can help the organisation with the information to made a more better decisions By providing sales forces.

3) HEALTHCARE

Healthcare is such area Where Data science is applied to make the changes. Healthcare analytics have such a potential Which can reduce the costs of the treatments , Help to avoiding the preventable diseases And improves the life quality Due to the data science analytics applications within the healthcare Has come with large number of positive as well as life saving outcomes. With the use of data science in health care, There is use of specific health data which helps to cure the diseases as well as cutting down of the costs etc. So, The basic purpose of the healthcare data analysis to find out the solutions of the problems and predict by using the past data before it will become too late.

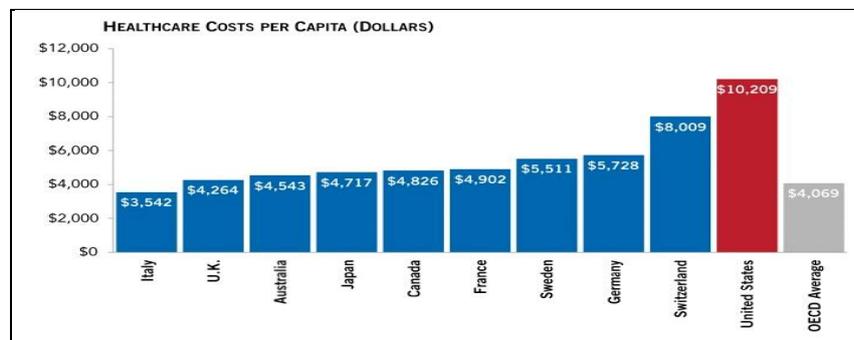


Table - 3

The figure-3 Shows the US per capita Spending Is more than the twice the average of the other developed countries.

There are various applications of Healthcare data analysis as:

- Predictions of the patients for improving the staff
- Reduce frauds
- Use predictive analysis.

4) Travel

Data science has raised a lot of opportunities To many of the industries. With this, It has came up with various changes and challenges. For the satisfaction of the customers and a process great data Chunks, The algorithms of data science are being viral. For teh Airlines, Hotels, Booking Websites and reservations and many more are making efforts for improving their services day by day.



Table- 4

Table- 5

So Figure 4 and 5 are the examples of travel in the form of graphical representation

5) Automation

- Driver- Less Cars:- These cars are made successfully and in reality with the help of applications of the rising technologies such as artificial intelligence, Data science etc.

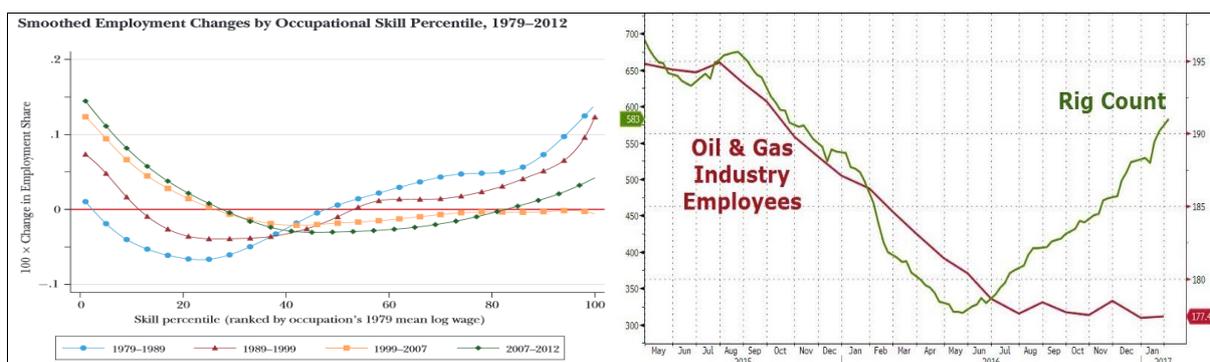


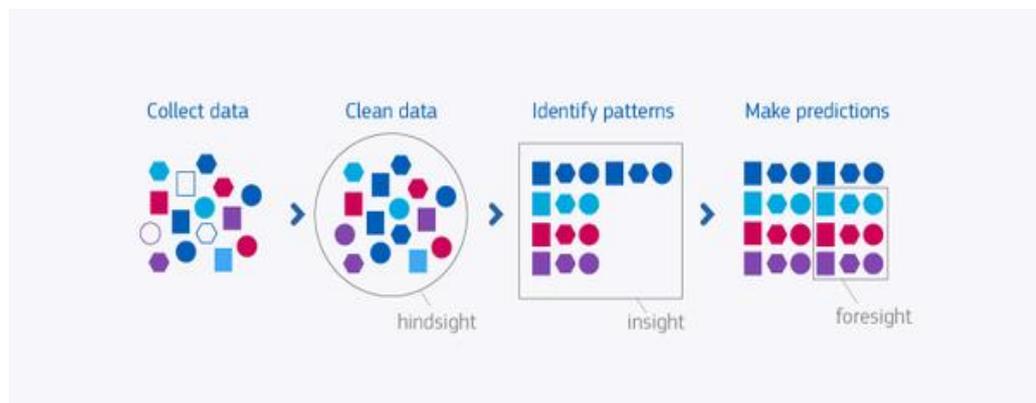
Table-6 Table- 7

So figure 6 and 7 are the examples of the Automation. Figure 6 represents the Skill percentage with the changes in the Employment share. Whereas in figure 7 We have see the chart which is plot of the oil rings within the united states with the number of employes in the oil and gas industry

✚ Importance of the Data Science.

Data science is a very broader concept. It first time came into existence in 1996. It includes the various aspects such as Statistics, Mathematics etc. Data science helps the individual to take the right decision making to achieve their organisation goals. It is basically a kind of multidisciplinary field. Everywhere data science is used. On the large scale, There are various data science models such as predictive model , Machine learning to solve out the problems of the large companies or industries. Data science is very essential part for the industries such as agriculture, Market analytics etc. Basically There are three Major Importance of the Data Science.

- **Helps in making Decision by the management.**



Data scientists are the one to who the highest priority is given and they are highly trusted advisor in every organisation because they are helping in taking the accurate decision making in the organisation. Data science helps the organisations or management to take the accurate and actionable decision. The decision taken by the organisation depends on the tracking, Recording as well as other various aspects or parameters of the organisation.

- **Helps in fulfilling the Actions on the basis of the latest Trends so that the organisations can achieve their goals.**

The data scientists in an organisation Check and explore carefully the data and other aspects of the organisation so far, and he takes the actions according to that which them to know about the short as well as the long term goals of the organisation.



This help in improving the performance of the company and attract the customers towards itself which will ultimately lead to increase in profits.

- **Past data help in making decisions.**

Past data will also used as an evidence that it have been worked in the past or will it be profitable. So the decisions taken by the Data scientists is based on the past or the present data evidence. No guess works are done by these people Like the previously used traditional methods.

So that to avoid the wastage of the resources with huge amount and well as risk and loss.

- **How the top most players are using the data science?**

There are various organisations which are required to handle their Expanding as well as complex data environment for the identification of the New sources so that to achieve the future opportunities, and also for the growth as well as work efficiently. The factors which differentiate the various organisations is that What is the value that they have took out from the storage of the data By using Analytics and in what way they will present it.

Further there are some of the best as well as the largest companies which are taking individuals as Data Scientists with highest salary packages are as:

- 1. Google**

Google is one of the biggest company which is hiring the Data scientists with Highest salary package. Today most of the Google is being Driven by the Data scientists as well as By machine learning and artificial intelligence also.

Data science plays an important role in the Success of the Google maps.

Google's biggest quality is their information, and far of it! It collects their information on the daily basis from an oversized range of reliable sources, along with your own phone. If you've got your location services on, Google are visiting be fed with a stream of anonymous bits that talk about Their location, Their relative speed and also itinerary.

With the help of the data analytics and so the information which is being collected from drivers, passengers, and pedestrians, the algorithms of the Google's machine learning are being able to predict about the traffic jams, facilitate your notice related to your optimum route and even verify that areas which must be avoided and the reason behind this is road works or accidents. Monumental(Large) amounts of historical info providing Google the ability to anticipate traffic designs.

Putting that data to use

SAP And Google are operational in the machine undertaking applications with the sort of the Client instruments that will enhance variation in people day by day. This can be regularly, Frequently done when Exploring the quality of the Huge data analyses and Data science and the growth in the volume and Alternative or gathering speed of information.

2. Amazon

Amazon is one of the biggest and global E- Commerce company Which is hiering the data scientists on the very large scale with a good salary package. For the finding of of the customer mindset, To know about the geographical Reach of the both Cloud as well as the E- commerce Domain among with the other Business goals, They required the Data Scientists. For this, They pay high salaries.

3) VISA

Visa is the one of the largest Financial Gateway for the various organisations or industries. It do the various transactions in the Number of the hundreds of millions on the day to day basis. With the healp of data science they Are generating a huge amount of revenue. Also it helps them to check the various Fraudulent activities or transactions. Also it is very helpful in making the products as well as services according to Requirements of the customers In various categories.

DATA ANALYSIS AND INTERPRETATIONS

Table 1 showing the sales performance and Table 2 shows The amount of sales by the payment methods

Interpretation

So the Tables i.e. Table 1 and 2 shows the graphical representation of the data came after analyses. In figure 1, There is an overview of the sales performance as well as shown that how much customers you have got in this year. Also it includes the revenue(in Euro) You have been generated. Figure 2 Represents that the amount of the sales by the payment methods in the last year. This line graph represent the Amount of sales done by the various payment methods such as PayPal, Banking transfer, Credit card, and debit card.

Table 3 shows the healthcare cost per capita(in dollars)

Interpretation

The Table-3 Shows the US per capita Spending Is more than the twice the average of the other developed countries

Table 4 Represents the various motives for the adventure travel. Whereas table 5 Represents that who rules the online travel market?

Interpretation

In Table 4 The pie chart shows the various motives in the form of percentage. Highest percentage motive is that people feel stress so to refresh them self they make adventure travel. Other motives are Birthdays, Have some fun with friends , Challenging myself etc. Table 5 Represents a chart which shows that how much the online travelling booking market has been become.

Tables 6 and 7 are the examples of the Automation. Table 6 represents the Skill percentage with the changes in the Employment share. Whereas Table 7 represents the graph between oil rings and oil & gas industry employees

Interpretation

In figure 6 the Skill percentage with the changes in the Employment share. In figure 7, the chart Represents the plot of the oil rings within the united states with the number of employees in the oil and gas industry. In this graph red colour represents the oil and gas industry and Green represents the Rig court. In 2013 the rig court and oil & gas industry employees are at the same point But in 2017, There is lot of increment in rig court as compared to Oil & gas industry.

CONCLUSION

At the end we can say that Data science is a very large field. It is a multidisciplinary field. It plays very important role in various fields. Today, in every sector data science is used. Companies use data science so that they can make the correct decision So that they can improve the performance of the company as well as increase the profitability of the company. Many companies Use data science for detecting various frauds which lead to a great loss of the company .It helps to take right decision making to achieve their goals. It used in various fields such as marketing, Sales, Travel, Healthcare, Credit and insurance etc. Various companies are using data science such as Google, Microsoft, Visa etc. . Many companies are generating a lot of revenue due to data science.

REFERENCES/ BIBLIOGRAPHY

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2071041

<https://www.annualreviews.org/doi/pdf/10.1146/annurev.publhealth.25.102802.124410>

<https://datajobs.com/what-is-data-science>

<https://www.quora.com/Why-is-data-science-important>

<https://blog.alexia.com/know-data-science-important/>

<https://intellipaat.com/blog/7-big-data-examples-application-of-big-data-in-real-life/>

<https://intellipaat.com/blog/what-is-data-science/>

<https://www.kdnuggets.com/2018/05/top-7-data-science-use-cases-finance.html>

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5565421/>

https://sdsi.stanford.edu/sites/g/files/sbiybj10211/f/data_science_research_at_stanford_2017-18_0.pdf

<https://www.quora.com/Why-is-data-science-important>

https://www.researchgate.net/publication/326111866_A_Review_Paper_on_Big_Data_Analytics_Tools/download

<file:///C:/Users/Dell/Downloads/bigdataanalytics.pdf>

<file:///C:/Users/Dell/Downloads/IJEDR1801110.pdf>

<https://towardsdatascience.com/the-future-of-data-science-14653afb52f5>

<https://www.quora.com/What-is-the-future-of-data-science-1>

<https://www-forbes->

com.cdn.ampproject.org/v/s/www.forbes.com/sites/quora/2018/07/16/how-will-data-science-evolve-over-the-next-decade/amp/?usqp=mq331AQCCAE%3D&_

js_v=0.1#referrer=https%3A%2F%2Fwww.google.com&_tf=From%20%251%24s&_share=https%3A%2F%2Fwww.forbes.com%2Fsites%2Fquora%2F2018%2F07%2F16%2FHow-will-

<data-science-evolve-over-the-next-decade%2F>

<https://blog.coursera.org/ds-academy-the-future-belongs-to-data-science/>

<https://www.analyticsvidhya.com/blog/2015/09/applications-data-science/>

<https://towardsdatascience.com/10-data-science-tools-i-explored-in-2018-d8226f99f342>

<https://towardsdatascience.com/data-science-simplified-hypothesis-testing-56e180ef2f71>

<https://acadgild.com/blog/data-science-driverless-cars>

<https://medium.com/activewizards-machine-learning-company/top-10-data-science-use-cases-in-insurance-8cade8a13ee1>

© Association of Academic Researchers and Faculties (AARF)

A Monthly Double-Blind Peer Reviewed Refereed Open Access International e-Journal - Included in the International Serial Directories.

