

Data Discovery with Tableau: A Case Study Using Data Analytics Tool

¹Amresh Kumar, ²Sayali M. Baitule, ³Santosh Kumar Sahu

¹Assistant Professor, Dept. of Computer Science and Engineering, G.H. Rasoni College of Engineering, Nagpur, Maharashtra, INDIA

²Freelancer, Nagpur, Maharashtra, INDIA

³Assistant Professor, Dept. of Computer Application, Shri Ramdeobaba College of Engineering and Management, Nagpur Maharashtra, INDIA.

Email:

Email: ¹amresh.kumar@raisoni.net, ²sayali.baitule1992@gmail.com, ³sahusk1@rk nec.edu

Received: 09th July 2018, Accepted: 14th August 2018, Published: 31st August 2018

Abstract

We live in the era of big data, where large volume of data is continuously rapidly growing with high velocity. Looking into the various characteristics of big data we must have a concluding tool to get best insight of the collected data efficiently. Thus there is a need of the technology and support. Information and Communication Technology (ICT) stands for information technology (IT) that enables the person to access, accumulate, transmit, and control Big Data, Economic & Business Analysis and Management Model, Government Management Models, Cloud and Security and many more. Over the years IT companies or business is trying to take benefit of such information and communication technology. This Paper outlines the brief of case study and practices done on a dataset “Global Super Store 2016”. This data is a sample data set provided by Tableau Community. In this paper, data analytic tool (business intelligence) is applied for generating insights through data visualization. We tried to analyse the market with respect to various region, category and segment as dimensions and sales, profit, discount as measures in order to improve the growth in market and deciding business strategy to capture good market. Also we tried to facilitate any organizations to construct improved decisions as well as confirm and contradict active theories or models.

variety of forms, like text on paper or bytes stored in electronic memory.

C. Big Data

Big data [3] is defined as hug collection of data sets that qualifies characteristics like volume, veracity, verity, value, velocity. Big data are very too complex hat conventional data processing application software are insufficient to deal with them.

D. Big Data Analytics

Data Analytics [4] is the scientific way of investigate raw data with the aim of detection of patterns and drawing conclusions by applying an algorithmic or mathematical process to gain insights and to facilitate organizations to construct better decisions as well as confirm and contradict existing theories or models.

E. Data Analysis

It [3] is a business decision making process of investigating past and present data in order to discover useful information, suggest conclusions to run any business smoothly.

Keywords: ICT, Big Data, Data Analytics, Data Analysis, Data Visualization, Business Intelligence

Introduction

A. ICT

Information and communication technology (ICT) [1][3] is extensional word for information technology (IT) which stresses the function of unified communications and telephone lines and wireless signals, computers as well as necessary enterprise software, middleware, storage, and audio-visual systems, which enable users to access, store, transmit, and manipulate information.

B. Data

As per common individual view, data is the collection of raw fact that can be analyzed for making decisions. [2] Data is a set of values of qualitative or quantitative variables. Data exists in a

Data Analysis	Data Analytics
<ul style="list-style-type: none"> • Data Analysis as a subhead. • Data analysis is the practice that encompasses the use of Data Analytics. • Process of compiling and analysing data in order to present findings to management to help inform business decision making. • “DATA ANALYST” 	<ul style="list-style-type: none"> • Data Analytics is a broader term. • Analytics is basically the concepts used to do the analysis. • It involves data extract, transform, and load; specific tools, techniques, and methods, to achieve business objectives. • “DATA SCIENTIST”

Fig. 1: Summarizing Data Analysis & Analytics Data Analytics (Materials and Methods)

A. Process

a. How to start (Prerequisites)

- i) These add benefits to those who want to create, read, write, and modify Business Intelligence Reports.
- ii) It will also be quite useful for those who would like to become a Data Analyst or a Data Scientist.
- iv) Before going on we should have a fundamental understanding of Computer

Programming terminologies and Data analysis.

- v) We should also have some knowledge on various types of graphs and charts.
- vi) Familiarity with SQL will be an added advantage.
- b. *After we stop (Outcomes)*
 - i) Understood and applied the data analytics technique for visualization.
 - ii) Analyzed and demonstrated various real time problem statements of data analytics.
 - iii) Applied advanced concepts of data analytics.
 - iv) By the end we will be able to produce influential reports and dashboards that will facilitate people construct decisions and obtain action based on their business statistics.



Fig. 2: Data Analytics Process

- c. *Steps (How Data Analytics Works)*
 - i) Data is collected in raw form from various sources with different formats.
 - ii) Then raw format of data is translated to desired formats.
 - iii) After that cleaning of data is done by removing unwanted data and filling missing data.
 - iv) Thus, in the fourth stage of data analytics process Analysis is done.
 - v) After analysis we get visual graphs and thus we draw insights from that to make report.
 - vi) Incas analysed data is further processed with models and algorithm then it results into a data driven product.

B. Applications

1.) Policing Security	8.) Customer Interactions
2.) Transportation	9.) City Planning
3.) Fraud and Risk Detection	10.) Healthcare
4.) Manage Risk	11.) Travel
5.) Delivery Logistics	12.) Energy Management
6.) Web Provision	13.) Internet Web Search
7.) Paper Spending	14.) Digital Advertisement



Fig. 3: Data Analytics Applications [8]

Data Visualization and Data Discovery Data with Tableau

A. Data visualization

Data visualization [10] involves the creation and study of the visual representation of data. A primary goal of data visualization is to communicate information clearly and efficiently via statistical graphics, plots and information graphics.

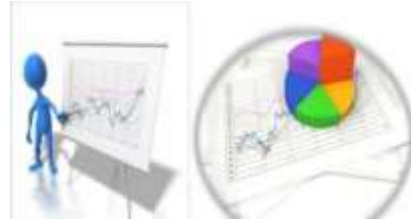


Fig. 4: Data Visualization

B. Data Discovery

Data discovery [9] is a process not a tool, it is a business centric (targeted) procedure for spotting the patterns and characterizes it by applying guided superior analytics. Data discovery is an iterative procedure which does not need widespread upfront model formation.

Data Discovery has three main categories:

- a) Data preparation;
- b) Visual analysis; and
- c) Guided advanced analytics.

A data discovery requirement is the skills to understand data relationships and data modeling.



Fig. 5: Data Visualization and Data Discovery All Around

C. Existing Technologies in Data Analytics are:

1. Excel
2. Tableau
3. Microstrategy
4. Informatica

D. Why Tableau? [5] [11]

- i) In upcoming years the world will generate much more amount of data.
- ii) Thus gathered data will provides huge number of opportunities and to turn these opportunities into reality, people need to use data to solve problems.
- iii) Tableau solves the problem of data visualization and also demonstrates best

practices for data visualization and data storytelling.

- iv) You will view examples from real world business cases.

E. Features in Tableau [11]

- i) The influential data discovery and investigation purpose allows us to answer significant questions in seconds.
- ii) Tableau has drag and drop interface for visualizing any data, look at different views, and even unite various databases simply.
- iii) It does not require any complex scripting for Data Analysis.
- iv) After analysis gets done, the sharing of output with others is as simple as publishing to Tableau Server.
- v) Feature: Dashboard and storytelling (data stories)

F. Getting start with Tableau [5]

- i) Download Tableau Desktop
 - a. <https://public.tableau.com/en-us/s/>
 - b. Provide Your Details (EMail ID)
- ii) Start the Installation Wizard
- iii) Verify the Installation
- iv) You are now ready to learn Tableau.

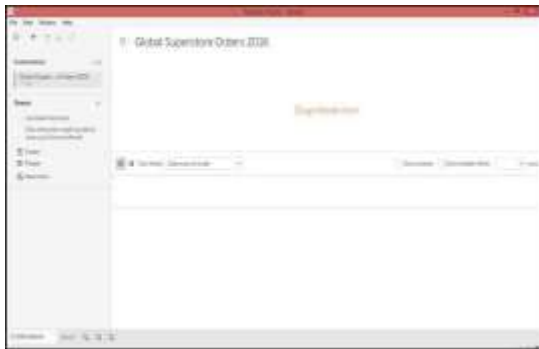


Fig. 6: First Screen of Tableau Public

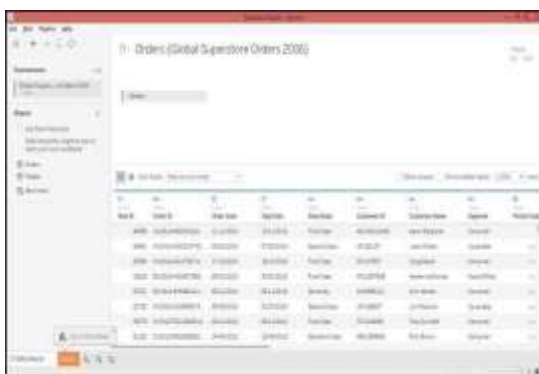


Fig. 7: Screen of Tableau Public after Loading Data Set

Case Study and Results Generation Using Data Analytics Tool Tableau

This section is the all about various cases and predicted data visualization on a data set “Global Superstore Orders 2016.

Case 01: Use proper data visualization to Show category wise total sales for CANADA regions using EXCEL and TABLEAU Individually.

Using EXCEL

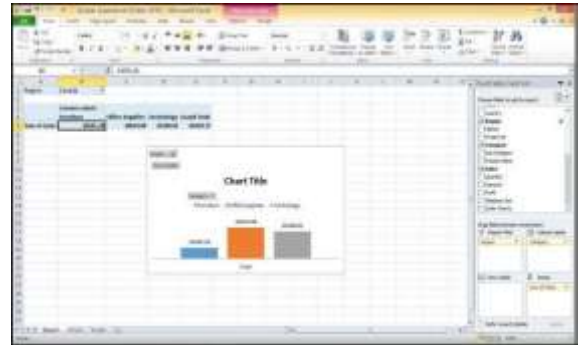


Fig. 8: Category Wise Total Sales Using EXCEL Using TABLEAU



Fig. 9: Category Wise Total Sales Using TABLEAU

Conclusion: Using Tableau we can generate visualization very efficiently and fast.

Case 02: Show segment wise growth in total sales in each category for all regions. Choose appropriate markets and emerging markets.



Fig. 10: Segment Wise Growth in Total Sales in Each Category for all Regions

Conclusion: For showing growth line chart gives proper picture of visualization.

Case 03: Show total sales Vs total profit in each category for Canada region in segment wise representation and Find a best market for investment as per good sales: profit margin.

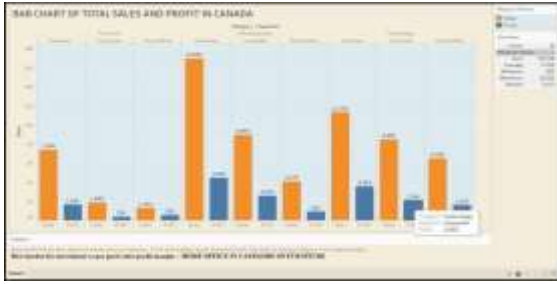


Fig. 11: Total Sales vs Total Profit in Each Category for Canada Region in Segment Wise Representation

Conclusion: sales Vs total profit is shown using side by side bar, this chart helps in comparing two different visualization properly.

Case 04: Forecast the growth of market for each category individually along with Justification.

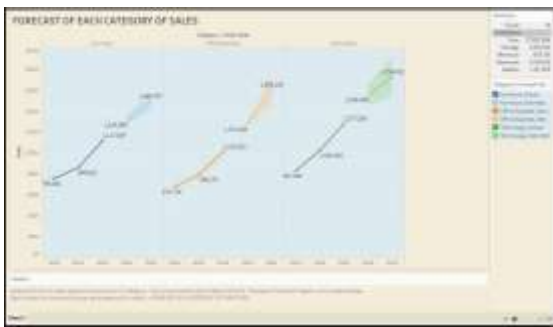


Fig. 12: Forecast the Growth of Market for Each Category

Conclusion: Here forecasting helps in knowing the future predicted values from previous stored/ gathered data.

Case 05: Analyse total sales, total profit and total discount for each category for all region and Find out the reason of UPs and DOWNs in the profit and Target any region and suggest appropriate business strategy to hold strong grip in market.

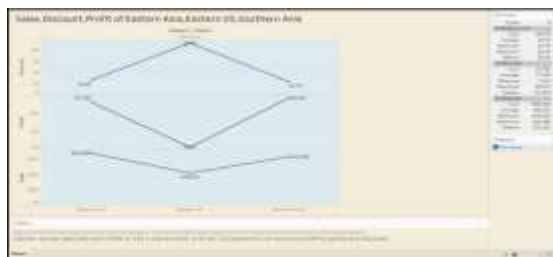


Fig. 13: Analyzed Total Sales, Total Profit and Total Discount

Conclusion: Here we have selected some region only for suggesting appropriate business strategy to hold strong grip in market.

Case 06: Apply analytics to Analyse total sales, total profit and total discount for Furniture or Office Supplies or Technology category for all regions and gets the best insight of Data Visualization using data analytic

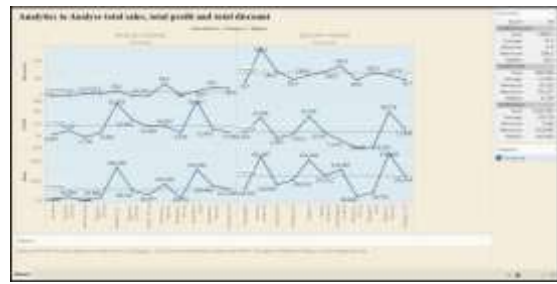


Fig. 14: Analyzed Total Sales, Total Profit and Total Discount using Analytics

Conclusion: Here we can generate parallel line charts for all regions to gets the best insight of Data Visualization using data analytic

Discussion and Conclusion

In this paper various case studies has been performed and tried to interpret the various data visualization patterns to observe the trends in sales in the given data set “Global super store 2016”. This case study helps in finding the trends of growth in sales and predicts the business best strategy in order to suggest and take some proper action or steps to improve business decision. Thus as a conclusion data discovery with Tableau as data analytics tool for generating various insights via data visualization has been done successfully. Hence in this way this case study was done to observe and decide various business strategies to capture good market and facilitate the organizations to formulate superior decisions.

References

1. https://en.wikipedia.org/wiki/Information_and_communications_technology
2. Kumar, Amresh & Kiran, M & Mukherjee, Saikat & Prakash G., Ravi. (2013). Verification and Validation of MapReduce Program model for Parallel K-Means algorithm on Hadoop Cluster. International Journal of Computer Applications. 72. 48-55. 10.5120/12518-9099.
3. Prof. Amresh Kumar, Ms. Sayali Baitule, “Big Data Analytics in the Cloud for Future Smart Cities” IJSRD International Journal for Scientific Research & Development| Vol.4, Issue02, 2016| ISSN (online): 2321 0613
4. <https://searchdatamanagement.techtarget.com/definition/data-analytics>
5. <https://community.tableau.com/>
6. <https://www.tableau.com/solutions/excel>
7. <http://ieeexplore.ieee.org/document/7748897>
8. <https://councilcommunity.com/>
9. <https://bi-survey.com/data-discovery>
10. <https://www.forbes.com/sites/bernardmarr/2017/07/20/the-7-best-data-visualization-tools-in-17/#5071440e6c30>