

Czech University of Life Sciences Prague

Faculty of Economics and Management

Department of information engineering



Diploma Thesis

**Converting raw data into useful information for taking Business
Decisions by using BI tools.**

Avinash Madala

© 2021 CLUS in Prague

CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

Faculty of Economics and Management

DIPLOMA THESIS ASSIGNMENT

Ing. Avinash Madala

Systems Engineering and Informatics
Informatics

Thesis title

Converting Raw Data into Useful Information for Taking Business Decisions by using BI Tools.

Objectives of thesis

The main objective of the thesis is to identify a significant visualization method for converting raw data into useful information in order to take effective business decisions. The partial objectives of the thesis are:

- To create the literature review in the field of the big data analytics and various visualization methods.
- To create the analysis of various tools used for data visualization.
- To examine the proposed visualization tool & provide recommendations.

Methodology

Methodology for the literature review will be based on data collection from specialized publications, websites, journals and other written or online sources. The master thesis will mainly consist of two parts: theoretical and practical.

The theoretical part will be based on literature review that provides understanding about the big data and data visualization methods. The practical part will be the utilizing the learned topics and implementing it in a best possible manner. The practical part will include analysis of current market of visualization tools and provide understanding about top applications. The analysis will also provide information about the advantages of applications and difficulties in using them.

The proposed extent of the thesis

60 – 80 pages

Keywords

Business Intelligence, Power BI, Data Transformation, Business Decisions.

Recommended information sources

ECKERSON, Wayne W. Performance dashboards: measuring, monitoring, and managing your business. John Wiley & Sons, 2010.

CHAUDHURI, Surajit; DAYAL, Umeshwar; NARASAYYA, Vivek. An overview of business intelligence technology. Communications of the ACM, 2011, 54.8: 88-98.

KIMBALL, Ralph; ROSS, Margy. The data warehouse toolkit: the complete guide to dimensional modeling. John Wiley & Sons, 2011.

TYRYCHTR, J. – VASILENKO, A. Business Intelligence in Agribusiness – Fundamental Concepts and Research. Brno: KONVOJ, spol. s r. o. 2015, 100s. ISBN 978-80-7302-170-2.

Expected date of thesis defence

2021/22 WS – FEM

The Diploma Thesis Supervisor

doc. Ing. Jan Tyrychtr, Ph.D.

Supervising department

Department of Information Engineering

Electronic approval: 1. 11. 2021

Ing. Martin Pelikán, Ph.D.

Head of department

Electronic approval: 23. 11. 2021

Ing. Martin Pelikán, Ph.D.

Dean

Prague on: 28. 11. 2021

Declaration

I declare that I have worked on my diploma thesis titled “Converting raw data into useful information for taking Business Decisions by using BI tools” by myself, and I have used only the sources mentioned at the end of the thesis. As the author of the diploma thesis, I declare that the thesis does not break the copyrights of any person.

In Prague on 29.11.2021

Acknowledgment

I would like to thank the supervisor, Ing. Jan Tyrychtr, Ph.D., for his advice and support during this thesis.

Converting raw data into useful information for taking Business Decisions by using BI tools

Abstract

This thesis deals with converting raw data into useful information using various data visualization tools. The study includes information about data analytics, and it describes multiple graphs to be used for data visualization. It also describes the top applications for graphical presentation such as Tableau, Power BI, Oracle, and many more. The data for analysis is collected from the annual report of the Reserve Bank of India. The study analyses data of foreign direct investment for the period from 2016 to 2020. It creates a dashboard based on the data available using Power BI. The dashboard helps readers to quickly and effectively understand the inflow of FDI in India in various sectors. It also indicates the most suitable visualization tool for international companies, top managers of companies, entrepreneurs, researchers, and economists. Further, the study provides recommendations based on the creation of a dashboard in Power BI.

Keywords: Raw data, Power BI, Tableau, Data visualization, Business decision, graphical representation, Dashboard, FDI

Table of content

1 Introduction.....	11
2 Objectives and Methodology.....	13
2.1 Objectives.....	13
2.2 Methodology.....	13
3 Literature Review.....	14
3.1 Overview of data analytics.....	14
3.1.1 Types of data analytics.....	15
3.2 Types of graph.....	17
3.2.1 Line chart.....	19
3.2.2 Area chart.....	20
3.2.3 Column Line chart.....	21
3.2.4 Bar chart.....	22
3.2.5 Stacked Area chart.....	23
3.2.5.1 Horizontal char.....	24
3.2.6 Waterfall chart.....	25
3.2.7 Trellis Layout of Multiple charts.....	26
3.2.8 Part-to-whole chart.....	27
3.2.9 The pie or Ringchart.....	28
3.2.10 Funnel chart.....	29
3.2.11 Box Plot chart.....	29
3.2.12 Scatter Plot chart.....	31
3.2.13 Choropleth Map chart.....	32
3.2.14 Geo Bubble chart.....	33
3.2.15 Custom chart Extension.....	34
3.2.16 Treemap chart.....	35
3.3 Description of the most used visualization tools on the market.....	36
3.3.1 Overview of Power BI.....	36
3.3.2 Overview of Tableau.....	38
3.3.3 Overview of Cognos.....	41
3.3.4 Overview of Qlik view.....	45
3.3.5 Overview of SAP Analytics.....	48
3.3.6 Multidimensional Schema.....	48
3.3.7 Online analytical processing.....	48
3.3.8 ETL (Extract, Transform and Load).....	48

3.4	Summary of the Business Analytics.....	58
4	Practical Part.....	60
4.1	Research Questions.....	60
4.2	The structure.....	60
4.3	Description of a case study.....	61
4.4	Sources of data.....	63
4.4.1	Database on Indian Economy-Reserve Bank of India’s Data Warehouse...	63
4.4.2	Description of the data used.....	65
4.5	Summary of creating the dashboard in Power BI.....	66
4.5.1	The types of chart.....	66
4.5.2	The process of importing the data in Power BI Software.....	67
4.5.3	The process for the creation of a dashboard.....	71
4.6	Difference between Power BI and Tableau.....	80
4.7	Comparative analysis of Power BI and Tableau.....	82
5	Results and Discussion.....	91
6	Conclusion.....	92
7	References.....	93
8	Appendix.....	95

List of figures

Figure 1: Visual perception in terms of accuracy.....	18
Figure 2: Sample of Line chart.....	19
Figure 3: Sample of Area chart.....	20
Figure 4: Sample of the column line chart.....	21
Figure 5: Sample of Bar chart.....	22
Figure 6: Sample of Stacked chart.....	24
Figure 7: Sample of Horizontal bar chart.....	24
Figure 8: Sample of Waterfall chart.....	25
Figure 9: Trellis layout of Multiple Charts sample).....	26
Figure 10: Part-to-whole chart sample.....	27
Figure 11: Sample of the Pie chart.....	28
Figure 12: Sample of Funnel chart.....	29
Figure 13: Sample of Box Plot of data from the Michelson-Morley experiment.....	30
Figure 14: Sample of Scatter plot chart.....	31
Figure 15: Sample of Choropleth map.....	32
Figure 16: Sample of Geo Bubble chart.....	34
Figure 17: Animated Custom Chart Extensions sample.....	34
Figure 18: Sample of Treemap chart.....	35
Figure 19: Power BI desktop (own picture).....	37
Figure 20: Image of tableau products offerings.....	38
Figure 21: Sample of Tableau Desktop.....	40
Figure 22: The possible chart types supported by Power BI.....	66
Figure 23: Power BI desktop (Own picture).....	67
Figure 24: Home menu of Power BI (Own photo).....	68
Figure 25: The Sources of importing the data (Own view).....	68
Figure 26: The Process of inserting the data (Own image).....	69
Figure 27: The transformation of importing data (Own photo).....	70
Figure 28: Power BI workspace desktop (own picture).....	71
Figure 29: Power BI workspace instruments for visualization.....	72
Figure 30: The data used for creating the Power BI dashboard (own photo).....	73
Figure 31: Interactive visualization Stacked Bar Chart in Power BI (own picture).....	74
Figure 32: Interactive visualization Line Chart in Power BI (own photo).....	74
Figure 33: Funnel chart in Power BI (own image).....	75
Figure 34: Pie chart in Power BI (own picture).....	76
Figure 35: Waterfall chart in Power BI (own photo).....	77
Figure 36: Treemap chart in Power BI (own image).....	78
Figure 37: Own dashboard created in Power BI.....	79
Figure 38: Modified raw data (own photo).....	95

List of tables

Table 1: Different between Power BI and Tableau.....	80
Table 2: Comparative analysis of Power BI and Tableau.....	82
Table 3: Characteristics of data.....	95

1 Introduction

Many different devices and software are available in the present world of modern information technology, where other devices and software can be used to register processes related to human activities. There is a considerable amount of statistical information generated, which is processed by various computer systems. Visual display of data is the most effective technique to show quantitative information.

The purpose of this Thesis is to analyze the data visualization tool Power BI with compare to Tableau. This will be very helpful in understanding the various benefits and functions of Power BI and how users can most effectively upgrade their software usage for visualizing the data by using this technology.

The statistical data can be displayed using visual and graphical methods such as box plots, histograms, spectrograms, pie charts, funnel charts, and many more. The visual and graphic methods are used for the same purpose: helping humans by describing the data, analyzing the data, and comparing the data in different ways. Besides this, the visual and graphical methods explain the dependency among variables.

The data stored with a high degree and made to detect knowledge is being hidden in data that may influence the decision-making process of an Individual. Data visualization refers to one of the most powerful ways to analyze data efficiently. Data visualization used to be the most effective & explicit data transformation method as it uses different types of graphics and charts. Visualization provides perspective and context which non-visual tools cannot offer. Further, Data Visualization is a modern way of converting raw data into meaningful data that is interactive and efficient at the same time.

People use data visualization for many purposes, like describing raw data, preparing and explaining business projects and reports formally to the viewer to analyze the data more efficiently. With the help of data visualization, any data can be presented visually in maps, charts, graphics, and other visual formats so that anybody can easily understand the data.

Data visualization is vital as it indicates the trends and patterns in a more particular way. Apart from this, there are enormous ways to search and identify hidden patterns of machines without losing sight of human capabilities for data analysis. The main aim of data visualization is to

represent a massive amount of data for humans so that users can see the data and understand the data quickly, which may be difficult for them to isolate algorithmically.

Nevertheless, if the presentation is interactive, the visual method is easy to present the data. The visual process helps deliver data in graphical forms, which makes users interact efficiently. For instance, it is possible to analyze the data quickly using either a mouse or a touch screen, which is a significant advantage for the international business Industry.

2 Objectives and Methodology

2.1 Objectives

The main aim of the thesis is to identify an effective visualization method for converting raw data into useful information to make effective business decisions.

- To create a literature review in the field of big data analytics and various visualization methods.
- To create the analysis of various tools used for data visualization.
- To examine the proposed visualization tool & provide recommendations.

2.2 Methodology

Methodology for the literature review is based on data collection from specialized publications, websites, journals, and other written or online sources. The master thesis mainly consists of two parts: theoretical and practical.

The theoretical part is based on the literature review that provides an understanding of various data visualization methods. The practical part is about utilizing the learned topics and implementing them in the best possible manner. The valuable part includes an analysis of the current visualization tools market and understanding top applications. The research is done by using Power BI. The study consists of information about the advantages of applications and difficulties in using them. It also defines the differences between Power BI and Tableau by creating the design of dashboards and the interactive visualizations of the specific data.

The method of analysis and synthesis is used to formulate the conclusion for the thesis. The overview of data analytics and various tools used for data visualization is included in the literary review of this thesis.

There are different methods used to compare Power BI & Tableau for this Thesis and explain the same given in the below-mentioned sections.

2.2.1 Expert evaluation - This, also called heuristic evaluation, reviews your product's user interface by two or more usability specialists. Working independently, these experts use published research data, industry-accepted usability principles (heuristics) and best practices, and years of experience observing users in lab and field settings to evaluate the product and identify usability problems. Evaluators also put on the user's hat, walking through the user interface based on task scenarios to assess workflow issues.

2.2.2 Multi-Criteria Decision-Making (MCDM) - This is also known as Multi-Criteria Decision Analysis, MCDA) is about methods – including software – for making decisions when multiple criteria (or objectives) need to be considered together, to rank or choose between the alternatives being evaluated.

2.2.3 Scoring methods are used in investment appraisal. They have two primary purposes. Firstly they are helpful where benefits are difficult to quantify objectively; secondly, they can be used to aggregate the results of multiple appraisal methods to provide an overall comparison. However, a more accurate picture is gained by scoring each factor according to how well it meets the criteria and weighing each factor according to its importance to the organization.

3 Literature Review

3.1 Overview of Data analytics

Data analytics is the science of raw data analysis to conclude that information. Data analytics techniques and processes have been developed into machine processes and algorithms that work and raw data for human use. Data statistics help a business improve its performance.

Data analytics is a broad term that encompasses many different types of data analysis. Any kind of information can be viewed with data analysis techniques to gain insight that can be used to improve things. Data analysis techniques can reveal styles and metrics that would be lost in the bulk of the information. This information can be used to add processes to increase the effectiveness of a business or program.

For example, manufacturing companies often record the operating time, downtime, and lines of operation of various machines and then analyze the data to load the load better so that the equipment can work closer to a higher power.

Data statistics are vital because they help businesses increase their performance. Applying it to a business model means that companies can help reduce costs by identifying more efficient ways of doing business and storing more data. The company can also use data analytics to make better business decisions and help analyze customer styles and satisfaction, leading to new - and better products and services.

The process involved in data analysis involves several different steps:

- The first step is to determine the data needs or how the data is collected. Information can be categorized by age, demographics, income, or gender. Data values may be priced or categorized.
- The second step in data analytics is the process of collecting it. This can be done through various sources such as computers, online sources, cameras, environmental sources, or by staff.

- Once the data has been collected, it should be edited for analysis. Editing may be in a spreadsheet or other type of software that can capture statistical data.
- The data is then cleaned before analysis. This means that it is cleaned and tested to ensure no duplication or error and is incomplete. This step helps to correct any errors before going to a data analyst for analysis.

3.1.1 Types of Data Analytics

Data statistics were divided into four basic categories.

Descriptive analytics describes what happened over some time. Has the number of ideas increased? Are sales stronger this month than last?

Diagnostic analytics focus on why something happened. This includes a variety of data entries and a bit of speculation. Has the weather affected the sale of beer? Has that recent marketing campaign contributed to sales?

Predictive analytics is likely to occur shortly. What happened to sales during the last summer we had a hot summer? How many weather models predict the hot summer this year?

Prescriptive analytics suggest a course of action. If the hot summer is rated at an average of these five types of weather above 58%, we should add an evening shift to the brewery and hire an additional tank to increase the outflow.

Other sectors that have taken over the use of data analytics include the tourism and hospitality industry, where turnaround can be faster. The sector can collect customer data and look at where problems lie and how they can be fixed.

Data analysis is a complicated process that includes several shifts among data formats, tools, models, and between symbolic & visual thinking. Description of how the design of improved tools facilitates people's exploration and understanding of data covering in terms of both interactive demos and principles combining concepts from machine learning, data visualization, and computer systems to design new interactive tools for analysis. Moreover, to understand it effectively, people should interact with it for

Understanding patterns, trends, and changes (Ohlhorst, 2013). In the present scenario, data visualization is used frequently in so many different areas of employment, like:

- The result of analysis of financial and stock market
- Data related to comets
- Magnetic resonance imaging of breast cancer for disease and therapy
- Modeling of a complex process

Graphic excellence provides the viewer to have an end no. of ideas in a short period with listed ink in the smallest space, as per Edward Tufte⁵, who is known for his work on information design. Visualization has become vital for all types of business (Philip Chen Zhang, 2014).

The patterns and trends become more understandable when information is represented graphically. It helps businesses identify the right priorities and focus on them while reducing variable data (The Data Visualization Catalogue, 2020). The main goals of visualization are:

- Provide the ability to choose and filter data
- Be able to change the way of presenting data
- Be able to change the way of encoding data
- Reflects interactions
- Increase areas of interest and indicates information about the same.

3.2 Types of graph

A graph is a mathematical diagram that depicts the relationship between two or more numerical data sets over some time. Primary data is mainly 2-dimensional with a focus on raw data represented through lines, curves, etc.

On the other hand, charts represent datasets to make the user understand the information better. Graphs are an excellent example of charts used for data visualization.

Graphs and charts cover large amounts of data in easy-to-understand formats to discuss the main points clearly and effectively. In choosing the best way to present your data, think about the purpose of your graph or chart and what you want to introduce. Then, decide what flexible things you want to do; also, it should be displayed as waves, percentages, or categories.

When deciding which type of graph or chart best shows your data, you must consider what kind of data you work with. The details of the section are as follows, grouped into irregular categories (e.g., rate, race, and answers to yes or no). Bar graphs, Line graphs, and pie charts help with the display category data. Continuous data is rated at scale or continuity (such as weight or test scores). Histograms help show continuous data.

Bar graphs, line graphs, and histograms with x- no axis. The x-axis is the horizontal part of the graph as well the axis is a straight part.

Graphs measure and represent numerical data assumed or aggregated as the sum of Revenue, Average Cost, Profit per capita, or non-numeric data calculated.

Measures are objects which indicate calculations and aggregate functions that are frequently applied to numeric data. The thing is accumulated for the column to be measured.

Sales revenue is a measure, but the total prices of the product list are not as it is a dimension. It forms actions from categories by counting their elements—for example, the number of countries visited by customers (Vasudev, 2006).

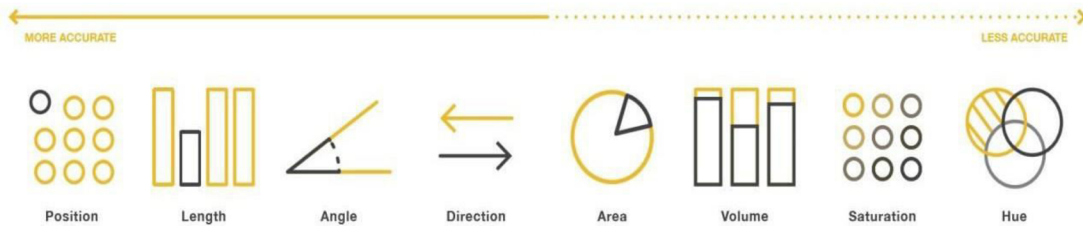


Figure 1: Visual perception in terms of accuracy (Source: own Source)

In above figure 2, visual perception in terms of accuracy is presented. The line indicates the less accurate types of charts at the right side of a picture and more precise types of charts at the left side of the picture. For example, visual acuity clarifies eyesight and can be measured by recognizing letters or numbers based on a standard eye chart from a specific distance. A dimension represents categorical data such as year, salary range, product, and country.

It indicates:

- **Categorical** (also called “nominal” for discrete values)
- **Ordinal** - It refers to the dimension members have a set default order
- **Interval** - It refers to each value in the dimension representing a range of values

3.2.1 Line chart

Line graphs are represented by a group of data points joined together by a straight line. Each of these data points describes the relationship between the horizontal and the vertical axis on the graph. The Y-axis shows the numeric value, and X-axis shows the key measurements or time dimensions such as a month, quarter, or year (Mitchell, 2012). Line and area charts are usually used to represent trends, Data over a specific period, Period-over-period, temporal patterns & correlation.



Figure 2: Sample of Line chart (Source: www.excel-easy.com)

Uses of a Line Graph

- It helps in studying data trends over some time.
- They are easy to read and plot.
- Disadvantages of a Line Graph
- It can only be used to visualize data over a short period.
- It is not convenient to plot when dealing with fractions and decimals.

3.2.2 Area chart

Area charts measure data trends over time by coloring the area between the line segment and the x-axis. In simpler terms, an area chart is an extension of the line chart.

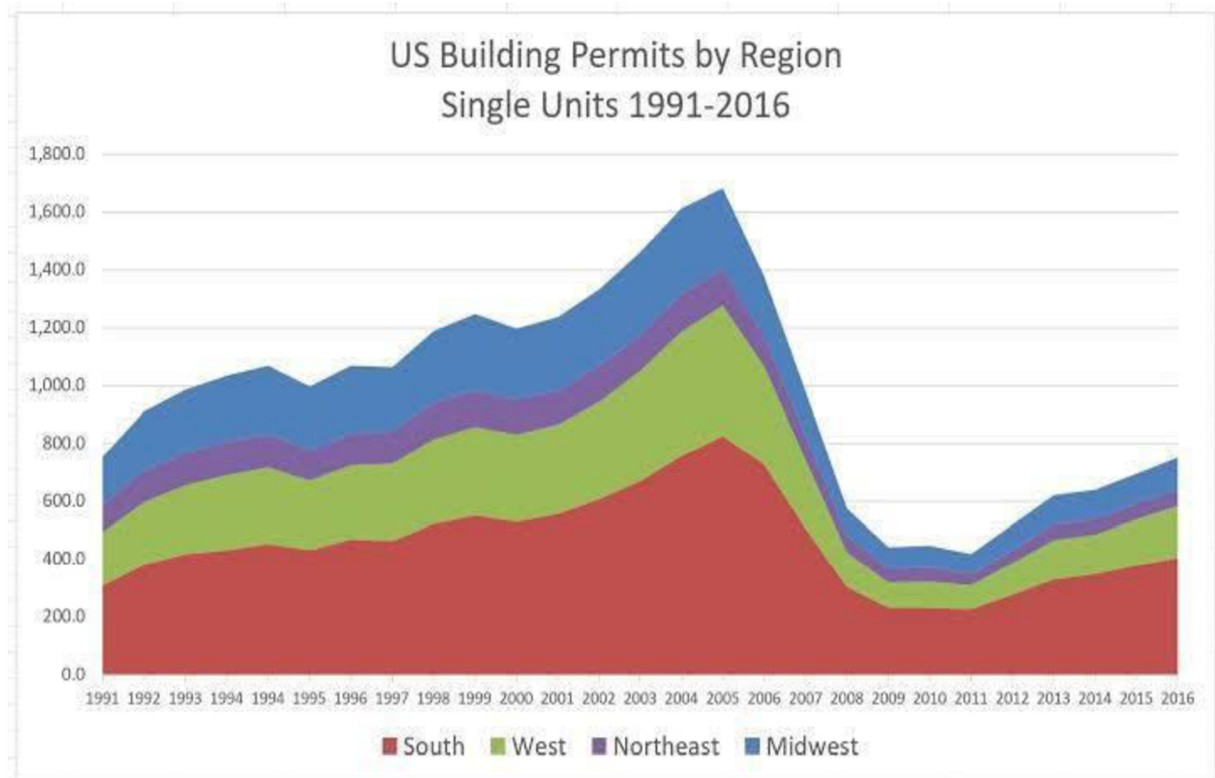


Figure 3: Sample of Area chart (Source: www.displayr.com)

Uses of Area Chart

- It is visually appealing.
- It gives a clear comparison of different groups of data.
- Cons
- It may be difficult to read when compared to other data types.

3.2.3 Column Line chart

The column line chart is a combination of line charts as well as a column chart. This type of chart indicates one measure as a column and another as a line. These two measures are shown under the time dimensions such as months, quarters, and years. This chart is more suitable for indicating the relationship between two steps: gross margin, net income, and sales revenue (Christensen, 2017).

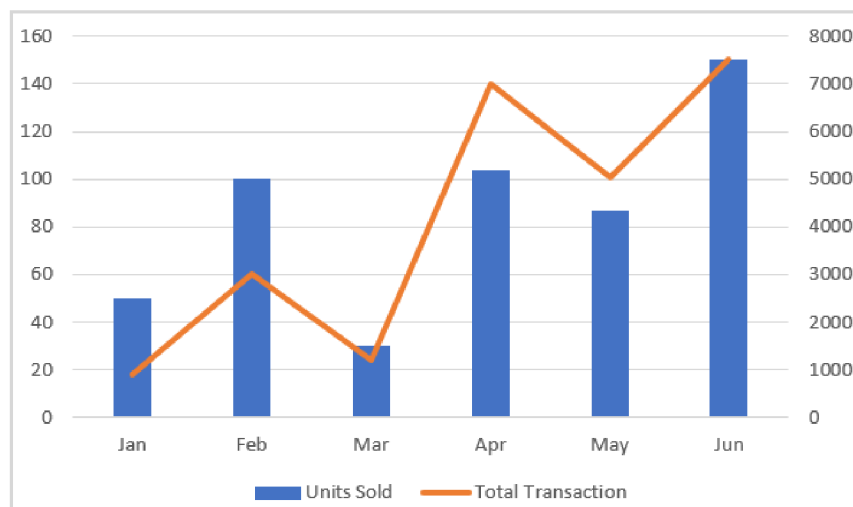


Figure 4: Sample of the column line chart (Source: www.microsoft.com)

In figure 4, a chart represents the number of units sold along with the entire transactions line.

The column line charts describe trends, data over time, temporal pattern & correlation for displaying two different data sets together and for displaying the sets of the same type of data that have extensive ranges between them.

Implementation suggestions:

1. Use a column line chart to display two trends of different types over a certain period. For example, we are returning customers and selling items.
2. This type of chart can also be used to indicate the changes over a specific period.

3.2.4 Bar Chart

Bar charts show comparisons and are easy to create. Take a moment to study the bar chart below. What does it show you? What do you conclude? What are the prompt chart questions? Depending on the audience, this bar chart may require more details of the methods used, such as the sampling process, data collection. Note that footnotes are defined in the same way in these two studies, making the findings similar. In the youth population, the results could not be compared. That's because a Williams County study describes teens as young between 11 and 18, and a nationwide survey describes teens as young as 21.

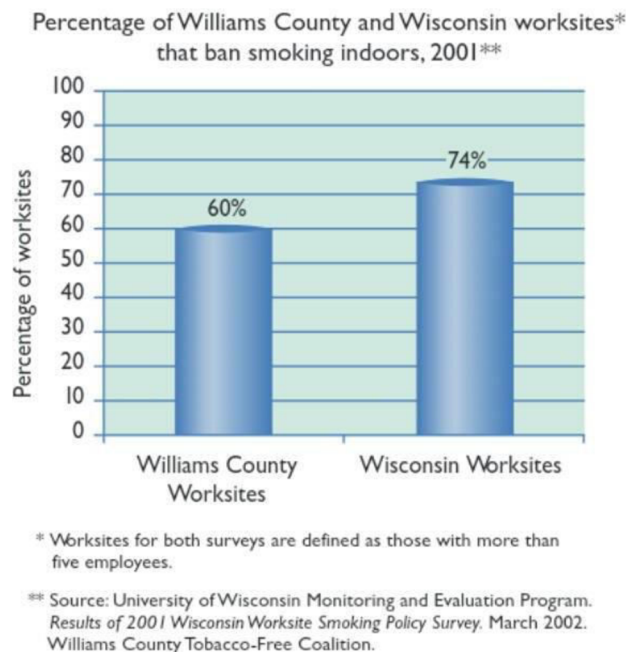


Figure 5: Sample of Bar chart (Source: www.istockphoto.com)

A few simple steps make the chart less congested:

- Value labels (percent written above bars) add clarity.
- The title also uses natural language; "Internal smoking control workshops" are less understandable than "workplace smoking restrictions."

- Grid lines add depth and size, which helps readers to see the differences between each data bar.
- Although the y-axis data label may seem unwanted, it ensures that readers know what the values mean.

The bar chart is the most frequently used chart among other chart types. Bar charts are majorly utilized for the following (Christensen, 2017):

- Positioning the data from most significant to most minor and vice versa
- Filtering the data which is not essential for the message
- For grouping data

Bar charts are usually used for comparing different values of categorical data, which includes Grouping the data by combining values in a chart – if there are various categories available. The less relevant absolute values should be grouped into another group.

3.2.5 Stacked Area chart

They stacked area chart extension of the base area chart. It indicates the appearance of some groups in the same image. The values of each group are displayed on top of each other, which allows you to look at the same picture the emergence of both the total of a numeric variable and the significance of each group.

The lines can be drawn, for instance, to track the change in the population of various states across particular periods. The area below each line can be colored with different colors to represent the state it depicts, resulting in a graph illustrating population trends and showing each state's data in order from least to most populous state (Christensen, 2017).

Popularity of American names in the previous 30 years

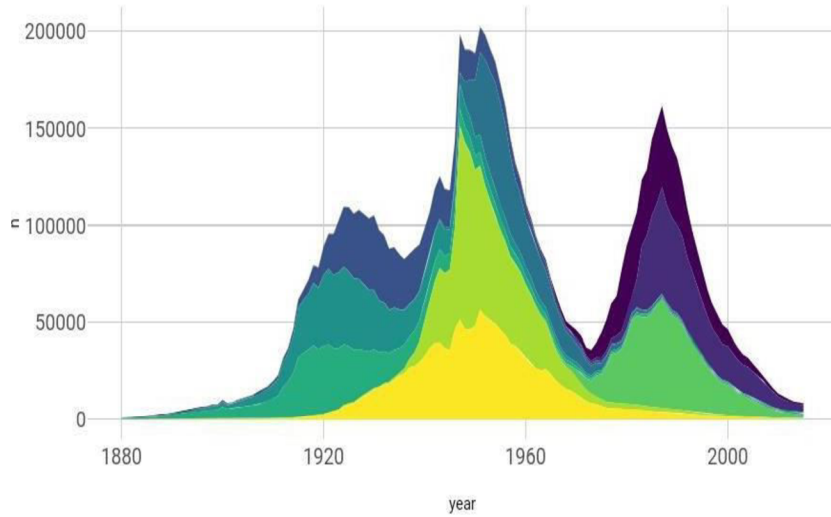


Figure 6: Sample of Stacked chart (Source: www.data-to-viz.com)

Figure 7 illustrates the emergence of baby names in the US between 1880 and 2015.

3.2.5.1 Horizontal chart

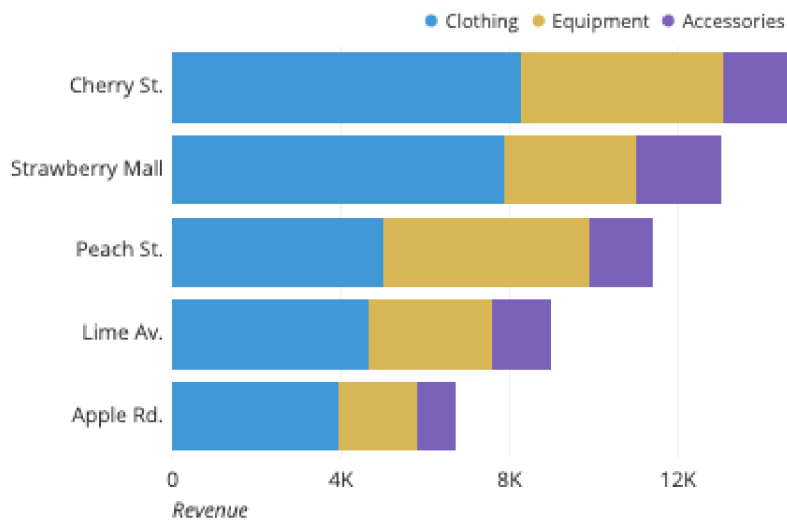


Figure 7: Sample of Horizontal bar chart (Source: www.alamy.com)

A horizontal bar chart is the best graphical tool for representing comparisons between categories of data. The presenter can show large data labels quickly with the help of a horizontal bar chart as the flat rectangles have enough room for displaying textual information. The example above shows the revenue data of clothing, equipment, accessories of different companies that indicate horizontal bar charts.

3.2.6 Waterfall chart

A waterfall chart shows data that helps understand the aggregation effect of positive or negative values. These intermediate values can be time-based or phase-based. The waterfall chart is also known as the flying brick chart or Mario chart due to the apparent formation of columns (bricks) in the middle of the air. Usually, in finance, it will be called a bridge. This type of chart helps visualize the fluctuation in positive and negative values (Christensen, 2017).

Waterfall charts are proposed by communications technology company McKinsey & Company in its presentations to customers.

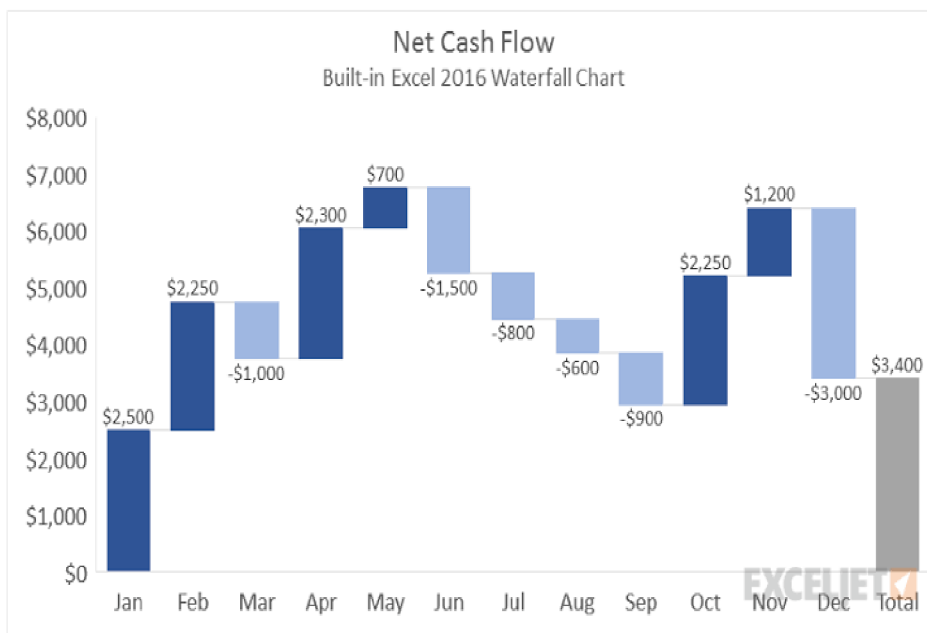


Figure 8: Sample of Waterfall chart (Source: en.m.wikipedia.org)

Difficulty can be added to the waterfall charts with many complete columns and axis crossing values. Additions and decreases in sufficient quantities may cause the accumulated value to fall above and below the axis in various locations. Middle footnotes, shown in full columns, can be added to the graph between floating columns. In figure 8, the net cash flow is indicated along with each month. Waterfall charts are used for:

- Cumulative effect
- Deviations and differences

Implementation suggestions:

1. Display how it arrived at a net value.
2. Break down the cumulative effect of positive as well as negative contributions.
3. Visualize an initial quantity.

3.2.7 Trellis Layout of Multiple Charts

A Trellis chart (also called multiple small charts) is a series of similar graphs or diagrams that use the same scale and axes, allowing them to be easily compared. This chart shows the same chart grid of values of different sizes and can compare metrics for other data groups. Edward Tufte loved the term ‘Trellis.’ These types of charts include vivid charts based on the same data set using the same axes to allow users for categorical comparisons of values within a dimension (Christensen, 2017).

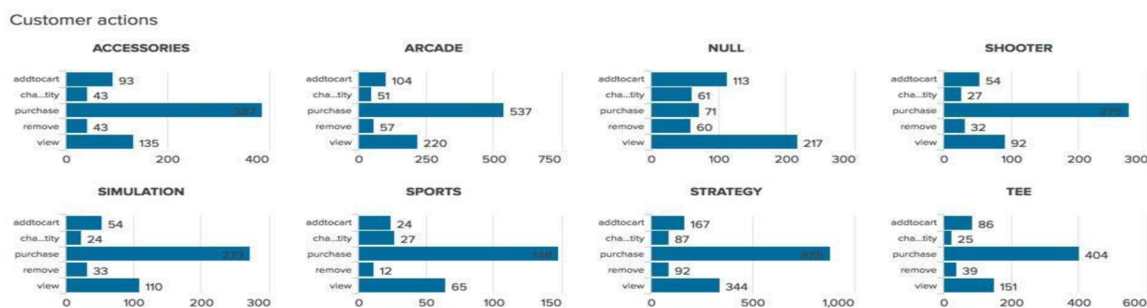


Figure 9: Trellis layout of Multiple Charts sample (Source: www.linkedin.com)

These types of charts are used for:

- Finding out patterns across multiple categorical values.
- For Comparing values within a category. For example, to display the sales values for each type in a different or separate chart.

3.2.8 Part-to-whole chart

This type of chart is used for a Part-to-Whole relationship which indicates how to measure values that form something whole. For example, No. of containers sold compared to one another, and each container compares itself with the total.

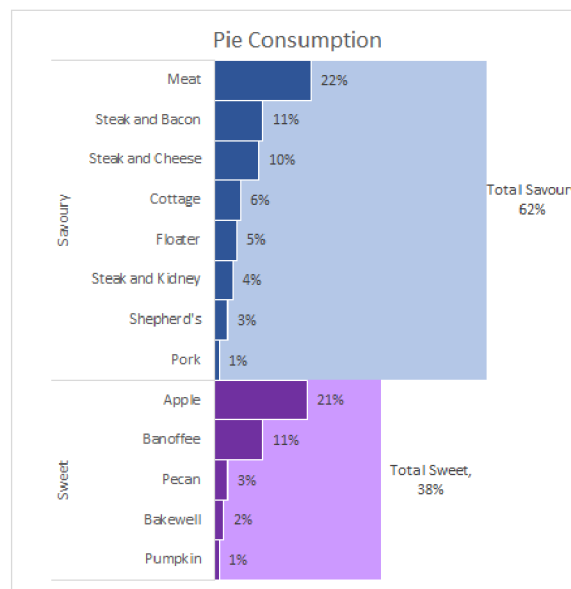


Figure 10: Part-to-whole chart sample (Source: www.anychart.com)

3.2.9 The Pie or Ring chart

If you are trying to figure out the percentage structure of the value, the first type of chart that often comes to mind is the Pie chart. These charts visualize the whole number as a circle (100%) and its parts as "pieces" related to their size. When using a pie chart, the main issues to keep in mind are maintaining the number of features more minor, seven max, and avoiding substantial and minimal differences between sizes, making the graph difficult to read.

Pie, Ring, and Funnel Charts are used to differentiate part to whole comparisons to either give importance to a portion of the data or examine values for other categorical values. It is suggested not to include many segments in this type of chart because the viewer may face problems distinguishing objects between numerous different colors (Christensen, 2017). Pie charts are a straightforward and effective visual tool for comparing parts of a whole.

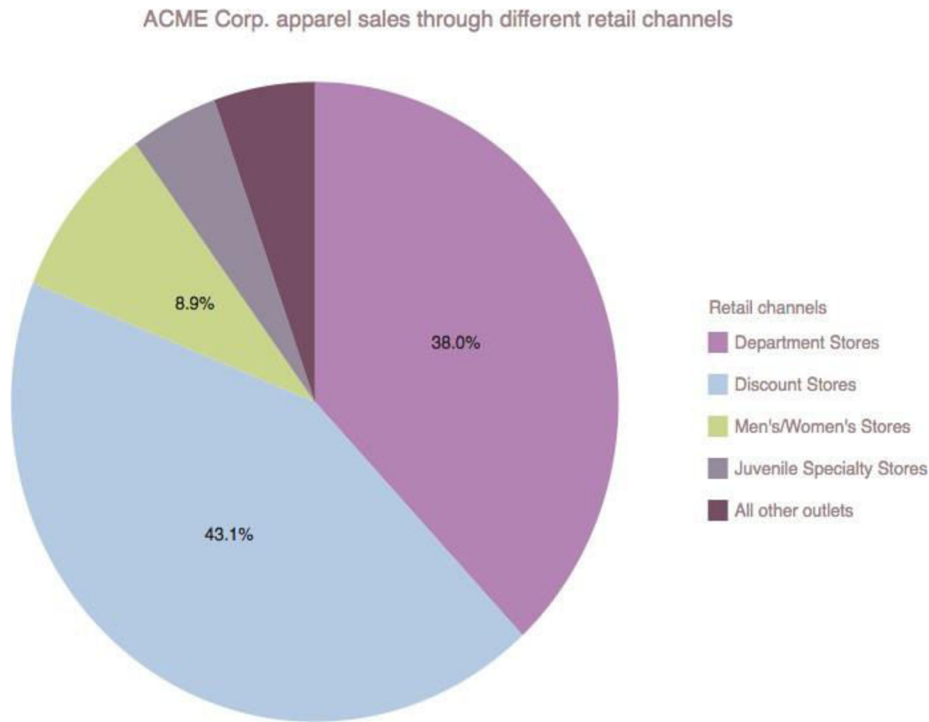


Figure 11: Sample of the Pie chart (Source: www.statisticshowto.com)

Figure 11 indicates the example of sales by channel. The pie chart can easily, quickly, efficiently, and effectively forecast numerous population segments. It can also respond to questions related to market research compared to ring and funnel charts (Christensen, 2017).

For example, a pie chart can indicate how the most renowned cell phone manufacturers are compared based on the size of their user-bases in the entire world.

3.2.10 Funnel chart

When you check out the process stages that start at 100% and end with a small percentage, see the steps, and point out the problems (the decline in which the stages occur and how much), you may like the Funnel chart type.

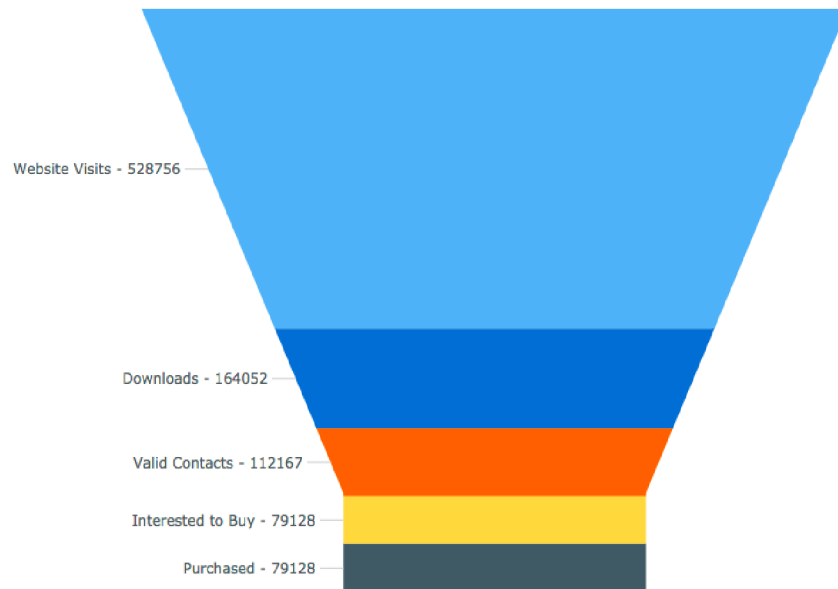


Figure 12: Sample of Funnel chart (Source: <https://charito.com>)

Figure 12 shows the conversion of website visitors into paying customers.

3.2.11 Box Plot chart

In descriptive statistics, a box or boxplot structure clearly shows groups of numerical data by their quartiles. Box plots can have lines from boxes (beards) showing distinction except for the upper and lower quartiles, hence the names of the box-and-whisker box and the box-beard drawing. Outliers can be constructed as individual points. Box plots are not parametric: they reflect the variability of mathematical samples without making any assumptions of primary mathematical distribution.

The spaces between the various parts of the box indicate the level of dispersion (spread) and the inclination of the data and suggest the vendors. In addition to the points themselves, they allow

One can visualize the various L dimensions, especially the interquartile, midhinge, middle-range, and Crimean range. Box sites can be drawn horizontally or vertically. Box sites found their name in the middle box and the actual structure.

Figure 13 illustrates how to create a box plot chart. Apart from this, the labels' parameter is used to give x-tick tags for every sample (Christensen, 2017).

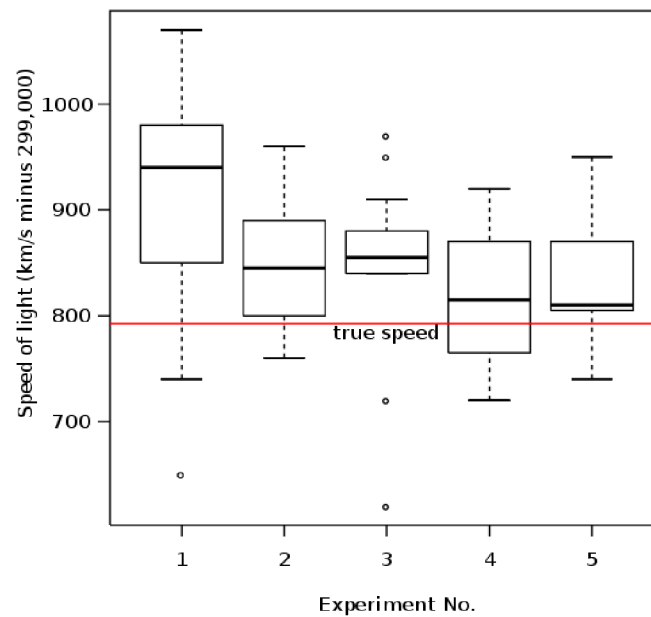


Figure 13: Sample of Box Plot of data from the Michelson–Morley experiment (Source: en.m.wikipedia.org)

The box plot chart is mainly used for:

- Comparison
 - Box plot chart compares data distribution for multiple categorical values.
- Distribution of values
 - The box plot chart indicates the distribution of medians in the data.
- Identifying outliers
 - Box plot adds a reference line to the overall median in data.

3.2.12 Scatter Plot chart

The scatter plot (scatter chart, scatter graph) uses dots to represent the numbers of different numeric variables. The position of each dot on the horizontal axis indicates the value of each data point. Scatter plots are used to view relationships between variables.

The scatter plot example above shows the width and length of the sample of the felled trees. Each dot represents one tree; the horizontal position of each point indicates the width of the tree (in inches), and the vertical position indicates the height of the tree (in meters). From the structure, we can see a good solid positive connection between the width of the tree and its measurement. We can also see the outer point, a tree that is much larger than the others. The tree appears to be too short for its distance, which may ensure further investigation.

The scatter plot chart indicates we scatter plots in a Trellis layout to compare various scatter plots in a single chart.

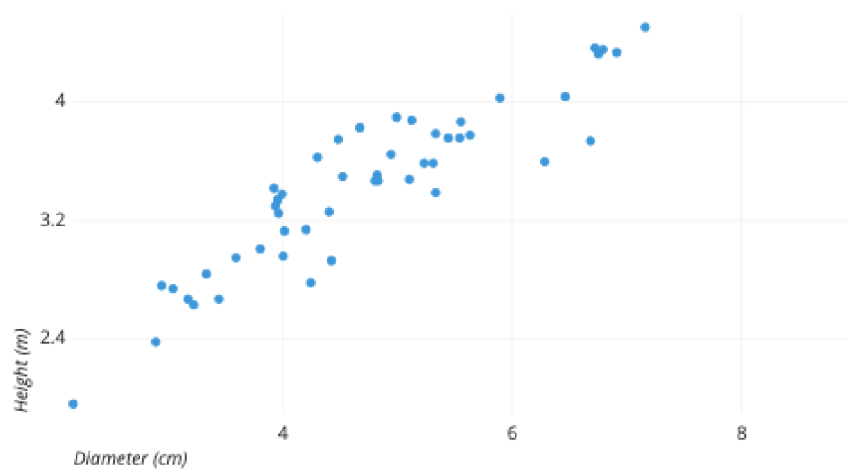


Figure 14: Sample of Scatter plot chart (Source: <https://chartio.com>)

There is a requirement to use the color to indicate groups of points, but the number of colors used to be minimum as too many colors or shapes will impact the readability of a chart. Creating Geo hierarchy on top of location data, for example, states, cities, etc., allows drilling up to higher levels of geography details (Christensen, 2017).

A **bubble chart** represents a scatter plot, but it enables visualization of a third measure because the bubble size shows it. The larger the standard is, the larger the bubble (Christensen, 2017).

3.2.13 Choropleth Map chart

Choropleth Map is a type of map chart used to show geographically segregated areas - countries, regions, etc. - that are colored, patterned, or inserted concerning dynamics. This way allows you to visualize data values in a particular area, understanding how they differ throughout the location shown on such a map.

When you create a large data map in Choropleth Maps, a specific color continuity should be used to display the data correctly. Choropleth Map uses different shading, coloring, and symbols within defined regions to show measured values in those areas (Christensen, 2017).

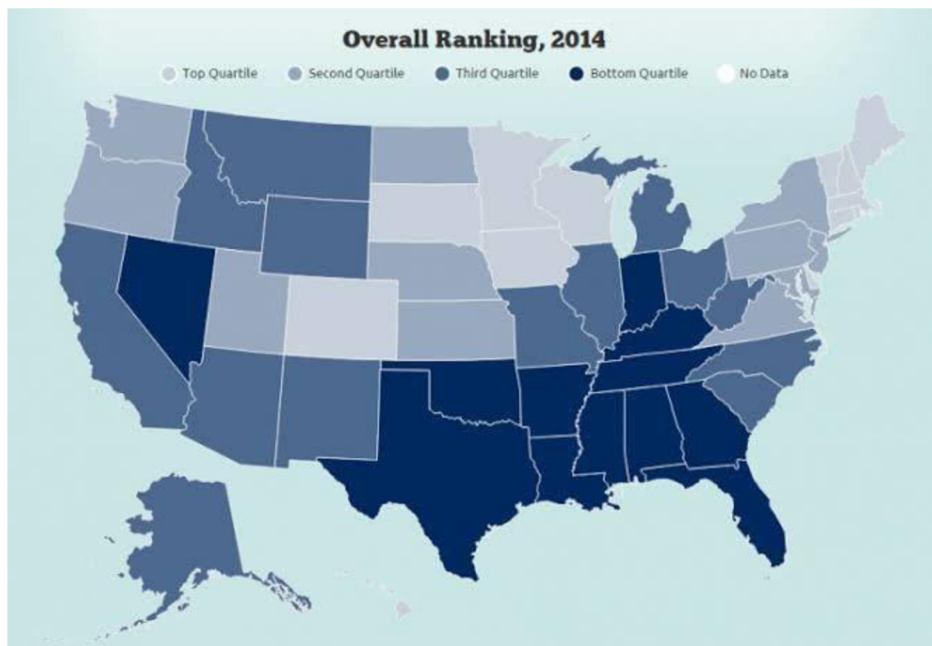


Figure 15: Sample of Choropleth map (Source: www.anychart.com)

Choropleth Map is used for supporting location-based comparisons of standardized data like rates, densities, percentages.

This type of map is usually used for similar locations as the size of the area colored may overemphasize larger areas. For example, Canada includes a larger area than Japan despite being smaller in the context of population. Choropleth Map may present visually clear information using various colors and shapes for the viewer. It may also contain disco and postcodes information related to all countries concerned about a particular industry or service.

3.2.14 Geo Bubble chart

Geo charts show map data by location. They are geo, geo bubble, and geo heatmap charts. And these geo charts can display six types of country data, depending on the site:

- Country
- State
- County
- Zipcode
- Point (latitude/longitude)
- Other sub-nation regions (for international countries)

This type of chart indicates measured values in the form of bubbles on a map. The more the measured values are, the more giant bubbles form on the map (Christensen, 2017). Geo bubble charts, like bubble charts, show the average value for a relative bubble size. Zip code data is a good choice for geo bubble charts.

Geo Bubble Chart is used for the following:

- Geo bubble chart is used for viewing measures that Country, Region, and City sort.
 - Emphasize the values on the map and build animation over time.
- Comparison of activities in different geographic areas.

- If the relative size of the areas covered does not match then, it provides a measurement value.



Figure 16: Sample of Geo Bubble chart (Source: <https://docs.thoughtspot.com>)

In figure 16, the bubble chart indicates an increase in sales, where it can be observed that the bigger the bubble, the larger the number of sales, the lower the drop, the smaller the number of sales.

3.2.15 Custom Chart Extensions

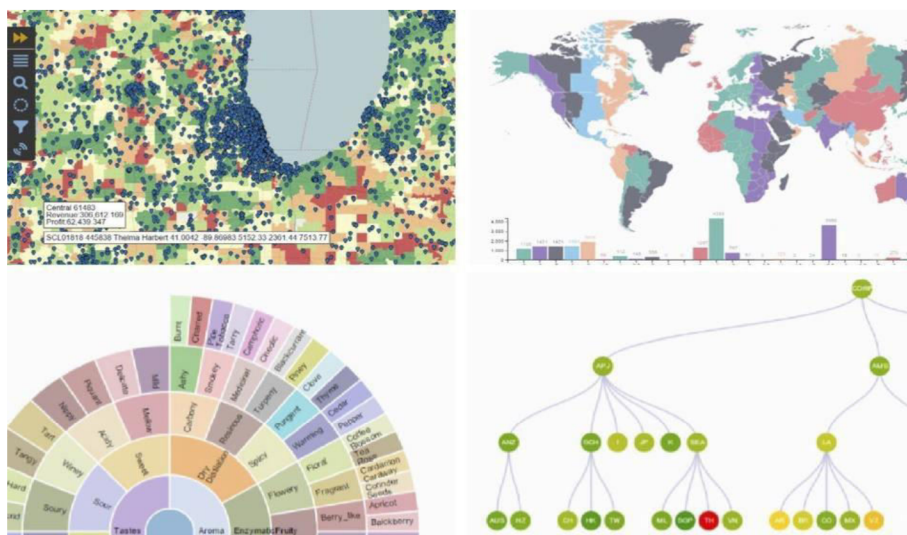


Figure 17: Animated Custom Chart Extensions sample (Source: own source)

This type of chart is prepared with a developer mind. There is always a requirement for creating custom visualizations for a particular object. Data can be obtained from various sources, and that could be an extraordinary one to report on. In both cases, the extensible framework forms custom charts and connectors (Christensen, 2017).

There is a need to start from scratch, adopt D3 charts, use open-source visualizations and data access extensions to build a custom extension. The D3 charts are JavaScript which is used for developing interactive data visualizations in web browsers. It allows users to use widely implemented scalable vector graphics 7, HTML5⁶, and Cascading Style Sheets⁸ standards.

3.2.16 Treemap Chart

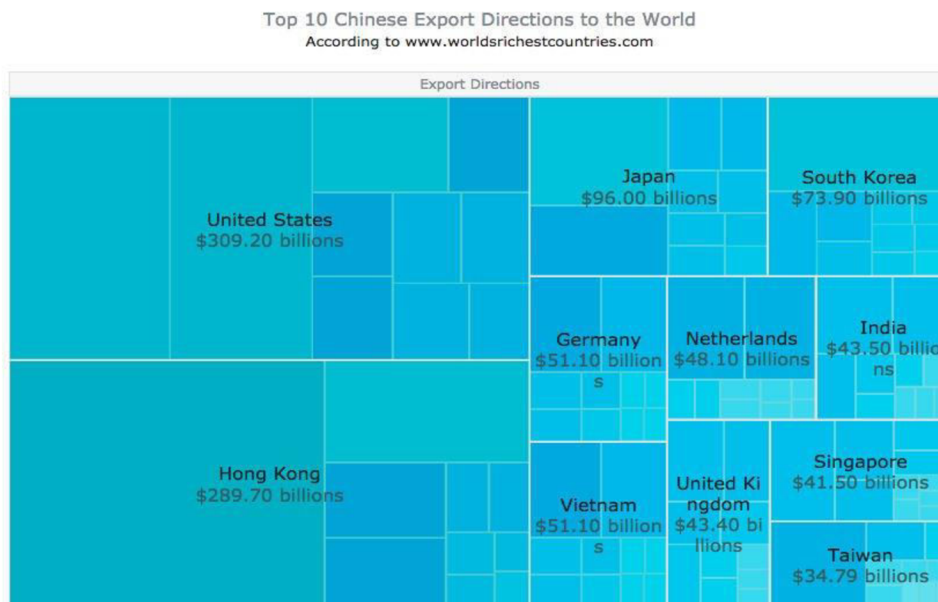


Figure 18: Sample of Treemap chart (Source: developers.google.com)

Looking at the sequence data, where the design is complex and organized by the tree, the Treemap chart is a great solution. With this type of chart, the whole (“tree”) is seen as a rectangular rectangle with smaller rectangles showing branches and smaller branches equally, making it easier to understand the structure and compare the shares. While Pie and Donut charts work with few points, Treemap charts are ideal for large data sets. Figure 18 indicates the example of export by country of destination.

3.3 Description of the most used visualization tools on the market

3.3.1 Overview of Power BI

Power BI, Microsoft, is an organization of business analytics tools used to analyze data and share information in reports and dashboards. User data in a variety of ways - spreadsheets, text files, databases, etc. Create Power BI input. Data sets are created by converting data provided by users. Users determine data conversion. This step is used to remove errors and unwanted data, correct formatting, and prepare additional analysis data by sorting them into appropriate standard forms, and so on. Depending on the report and the dashboard being created, filtering data to include only relevant pieces focuses only on essential data.

Once the database is ready, reports can be created from them by adding from the multi-item selection option. The visual objects in the Power BI range from the number one display to the gradient color map. These mirrors help to present data in a way that provides context and understanding. Filters can be used in reports so that relevant data is directed to users interested in data analysis. Such information can be created manually, or once the accounts have been published, report items or the entire message can be added to the dashboard.

Power BI dashboards display 360-degree data viewing by enabling users to keep their most important metrics in one place. They also allow users to participate in filtering or query reports and allow native language queries. However, it is limited to one page, thus only opening the relevant parts of the data to make it easier for users to gain insight. It is also possible to regularly update data and dashboard data in real-time and make it available on all devices such as PCs and smartphones (Vijay Krishnan, S Bharanidharan, G Krishnamoorthy, 2017).

Most users working on business intelligence projects use the Power BI Desktop to report, provide and analyze and then use the Power BI app to share their reports with others (Data Visualization, Microsoft Power BI, 2020). The key features of the Power BI Desktop are as follows:

- Connect to data
- Change data and build a data model

- Create visuals (charts or graphs) that provide visual data presentations
- Create visual reports on one or more of the report pages
- Share reports with others via the Power BI service

Power BI is free to use the app to visualize various data types that analyze and retrieve filed information, intelligently organizing chapters with a closely related group table into one category (Data viewing, Microsoft Power BI, 2020). It can create complex and visually rich reports, use data from multiple sources, and immediately shares it with others through the Power BI service thinking (Data viewing, Microsoft Power BI, 2020).

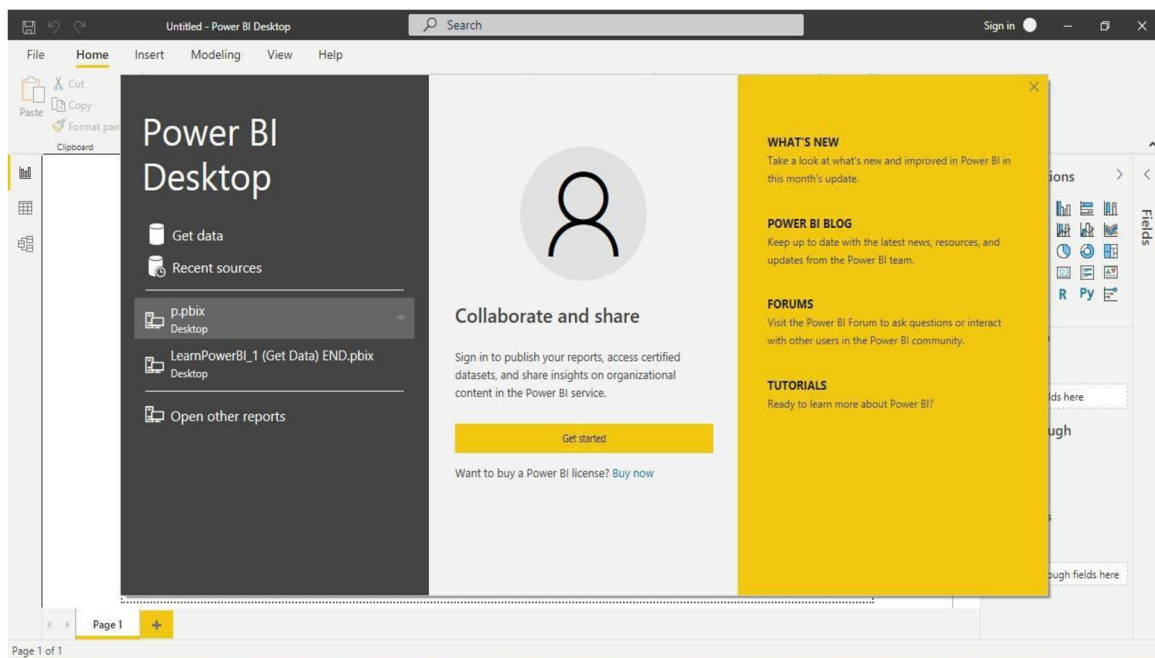


Figure 19: Power BI desktop (Source: own picture)

In figure 20, the first page of the Power BI Desktop report is represented as Home. It is the beginning of a visual tool. There are many features that Power BI offers to connect to other apps and devices. There are three ways to use Power BI, such as:

- Power BI online
- Power desktop
- Power BI Cell Phone

3.3.2 Overview of Tableau

Tableau Software is an American computer software company based in Seattle, WA, USA. Produces BI-based data consulting products. The company was established in the Department of Computer Technology at Stanford University between 1997 and 2002.

The various products made by the tableau are:

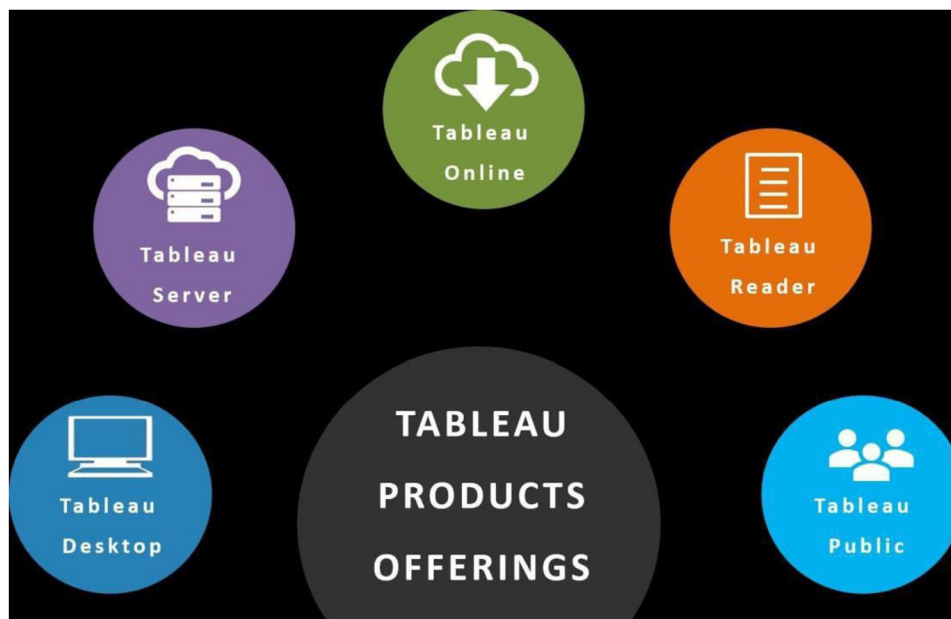


Figure 20: Image of tableau products offerings (Source: <https://intellipaat.com>)

Tableau Desktop

Tableau Desktop is a data recognition application to help you scan almost any type of structured data and generate highly interactive graphs, beautiful dashboards, and reports within minutes. Once installed, you can link to virtually any data source from spreadsheets to data repositories and display details in several clicks. Designed for easy use, you will be working much faster than before.

Tableau Server

It is a business spy app that provides browser-based analytics that anyone can use. It is the fastest way to fire at the slow pace of traditional BI software. It is an online solution designed for the

Sharing, distribution, and interaction of content created in Tableau. What makes Tableau different? Everyone is recommended. No script is required, so everyone can grow up to be a mathematician. You can increase your postage, as you need. Train online for free. Get answers in minutes, not months.

Tableau Online

Tableau Online is a secure, cloud-based solution used to share, distribute, and collaborate on Tableau views and dashboards. Tableau online puts the flexibility and simplicity of a powerful cloud-based data recognition solution for us - without servers, server software, or IT support. Tableau Online can measure as much as you need, and you can use it wherever you need. You do not need to purchase, set up or manage any infrastructure.

Tableau Public

Tableau Public is free software to help anyone connect to a spreadsheet or file and build interactive web visibility. It is provided as a service that allows the user to get up and work through the night. Tableau Public users can create amazing interactive views and publish them instantly, without the help of a program or IT.

Features of Tableau

- It is straightforward to use.
- You don't have to be familiar with programs of any kind; you need a few data and tables to produce visually appealing reports and tell a story that user should say to their supervisors or please their professor in class.
- It has a drag and drops feature where the user can generate stories or reports with pain and drop and a few thoughts.
- This may be because VizQL (VizQL is a visual query language that translates the action of drag and drops to data queries and renders that data visually Query and analysis.

Such a tool does not require any special technical or organizational skills to collaborate or create any feature. It has a collective interest among people from a variety of businesses, researchers or

Industries. Tableau Software has an intuitive, easy-to-use interface, which allows business users to discover, analyze and display the results of their data without technical and system capabilities. Tableau's capabilities include advanced forecasting, integration, default geocoding, and a visual formula editor. It allows users to do more research, uses their data faster and more efficiently than most competing platforms, and downgrade data distribution.

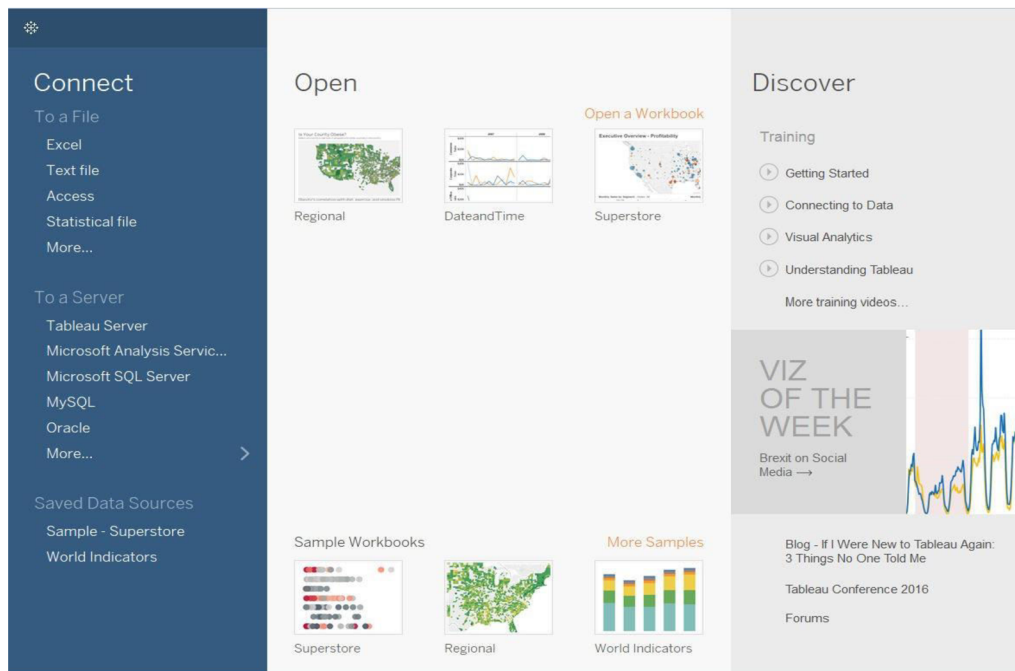


Figure 21: Sample of Tableau Desktop (Source: <https://intellipaat.com>)

There is also an online community called Tableau Public for development training and support. Some of the weaknesses are high product costs, difficulty modeling when combining data from different sources, low performance, and a large amount of data.

Dashboards can be shared with others as a static files. While launching Tableau, data connectors are available and allow you to connect to any database needed.

In summary, Tableau is a business platform designed for collaboration, distribution, management, security model, and automated features. End customers are better experience accessing files stored in all locations such as desktop, mobile phone, or email. (Tableau, 2020)

3.3.3 Overview of Cognos

IBM Cognos is a business intelligence tool for web-based reporting and analytics. This business software offers a variety of features to enable data integration and create detailed reports that are easy to use. Cognos also provides the option to send notifications in XML or PDF format and view reports in XML format.

Features of IBM Cognos

The following are the critical features of Cognos.

- Provides streaming statistics within the memory
- Provides real-time events, notifications, and notifications.
- Intuitive, attractive Web 2.0 interface
- Personal and progressive collaboration
- Drag and drop, free form integration, and search-assisted authorization
- External data operated by the wizard
- Automatic access to SAP BW queries
- Allows features such as status model creation, real-time monitoring, and forecasting
- Users can edit existing data.
- Drill-Through Power.
- Integration of Possible Image Documents.
- It is an independent, reliable, and dependable platform.
- Provides secure data as protected by a firewall.

Components in Cognos

Here are the key features of the Cognos software:

Cognos connections:

Cognos Connect is a website that allows users to access Cognos 10 and studios. Depending on your assigned role, you may use this section to locate, view, publish, manage, and edit company reports, scoreboards, and agents. The administrator also uses Cognos Connection to establish user roles and permissions and manage Cognos Connection content.

Cognos Business Insight:

The offer of Cognos Business Insight allows users to create their dashboards and use anything. All content that the user is allowed to view will be presented as an object. This can be used in your workplace to create a completely personalized dashboard.

Cognos Query Studio:

Cognos Query Studio helps business users find quick answers to business-related questions. It helps organizations better understand the product, customers, and organizational needs. It also helps them to respond faster and stay ahead of the competition.

Cognos Analysis Studio:

Cognos Analysis Studio helps businesses find and focus on business priorities. It is also helpful to understand the latest trends, compare data, and evaluate business performance with several analyzes.

Cognos Business Insight Advanced:

Cognos Business Understanding Improved by a new module installed in Cognos 10. Includes Cognos Questionnaire and Cognos Analysis Studio. It provides a robust environment for authorizing business people.

Cognos Report Studio:

Using the Cognos reporting tool, you can create complete pixel reports for your organization. Allows you to create charts, maps, lists, or any other type of report available using affiliate or multi-distribution data sources.

Cognos Event Studio:

This tool allows you to provide a specific event that sends a notification to a stakeholder in your organization. You can create agents that empower you at your events and borders; therefore, the agent who sends the information can reach an event or limit.

Cognos Metric Studio:

Cognos Metric Studio lets you view and analyze your organization's business metrics by creating a scorecard environment. It also helps you to initiate the process and monitor your organization to see how it responds as changes are made to the process.

Types of Cognos

Here are the key types of Cognos reporting tools:

Content Store:

Set with data tables used by the Content Manager to store Cognos application data.

Content Manager:

Cognos Manager helps you manage storage and retrieve reported data, configuration data, published packages from the content store database.

Cognos Content Database:

The Cognos Content Database is a self-contained database server used to host a content store database in demos locations where DMBS business is unavailable.

Other major significant BI vendors:

- Business Items (Managed by SAP)
- Information Builders - Focus and WebFocus.
- Microsoft - SSRS / SSIS / SSAS
- MicroStr Strategy
- OBIEE (Oracle)
- QlikTech - QlikView
- SAS

Pros of using Cognos

Here are the benefits/benefits of using Cognos software:

- You can also publish Cognos-enabled files on a secure BI site.
- Cognos BI provides an unlimited work environment to support the way people think and work.
- Provides a variety of BI skills such as query and reporting, analysis, and scorecards using a single construction.
- Easy views - combines and customize details.
- It helps you analyze the facts and anticipate the impact of strategies and strategies
- It allows you to participate in establishing networks for data sharing decisions and driving in integrated intelligence.
- It provides transparency and accountability to deliver alignment and consistency.
- It helps you to connect and coordinate activities involving the right people at the right time

- Allows you to integrate and integrate analytics into solutions and business process flow processes.
- Reduced decision time for data access without delay.

Cons of Cognos

- Cognos Bi was also not welcomed with much enthusiasm in the use of departments or sectors.
- There is no support for any Multi-dimensional analysis.

3.3.4 Overview of Qlik view

Click (formerly known as QlikTech Inc. provides a business statistics platform. SaaS software company was founded in 1993 in Lund, Sweden, and is now based in King of Prussia, Pennsylvania, United States. Qlik Sense and Qlik Replicate, Both cloud-based software for business intelligence and data integration.

QlikView is a business intelligence tool for data integration, chat statistics, and converting raw data into knowledge. QlikView provides ad query features and makes it easy to make decisions based on easily accessible information. This software works on the integration model and can be accessed in any way to search for answers.

Here are the benefits / great benefits of using QlikView.

- It is straightforward and intuitive to use the platform.
- Provides data visibility in a clear and professional way
- Quick response time without limit to the amount of data
- Allows you to search for all data directly and indirectly
- QlikView provides support for dynamic applications, dashboards, and analysis, etc.

- It facilitates the rapid integration of data from multiple sources into a single application.
- It helps you identify styles and information to make new decisions.
- Allows you to access, analyze, and retrieve data from mobile devices.
- It provides a quick return on investment due to the short start time.
- Rapid Development and Change Management
- Self Service Business Intelligence for end users.

Key features of QlikView

1) Data Integration Model: QlikView works on an in-memory integration model. So you don't need "IT professionals" If you can write SQL, select Question and understand your data.

2) Platform: QlikView is a windows-only product but is configured for 64 bit and many processors.

3) Costs effective: QlikView is a memory app, so it's an inexpensive way for your organization.

4) Economical: QlikView saves a lot of disk space as the amount of memory depends on the amount of data you want for each program.

QlikView Applications

- It helps you to analyze revenue data based on various filters of a large international organization.
- Allows you to create a detailed watch report based on social media feedback from customers around the world.
- The Qlikview tool lets you make decisions based on predicting cause and effect with drawings using its historical data.
- QlikView lets you analyze existing items, activity orders, sales, and staff usage - all focused on improving performance.

- Tools allow you to analyze information about sales, production, shipping, and product performance.
- You will efficiently and quickly create reports and analyses at the strategic, practical, and strategic levels.
- Canon, Panasonic, Shell, FILA, ING LEASE are some well-known companies that use Quickview in their businesses.

Disadvantages of using QlikView

- Since QlikView needs to carry a heavy load of all data in system RAM, reloading data requires a lot of workload and time.
- It does not provide any advanced ETL capabilities required for complex data integration.
- QlikView does not provide some of the advanced skills needed to create highly formatted reports.
- QlikView is a tool designed specifically for interactive analysis and not for report writing.
- If you want to create a formatted report on QlikView, you need to use macros and repeat and save QlikView objects.
- It is just a simple reporting and analysis tool that does not WRITE back to the database.

3.3.5 Overview of SAP Analytics

SAP Analytics Cloud, commonly referred to as SAC, is a tool for viewing cloud data from SAP.

SAP Analytics Cloud combines the following functionality into one device:

- Business Intelligence (BI);
- Predictability Analysis;
- Planning.

SAP provides a complete solution when it comes to business intelligence. Available online in SaaS mode, SAC is becoming increasingly crucial in SAP's environmental system. Finally, it is intended to be a reliable BI tool for the SAP Cloud suite.

What needs does SAC address?

- SAP Analytics Cloud is a new tool in the SAP portfolio to meet viewing data in the cloud. It is delivered as an all-in-one product based on SaaS. Includes requirements for data identification, budget planning, and forecasting analytics.
- Its principal function is to create data reports.
- This data can come from various sources: real-time data taken from your business operations, budget status data, or thousands of data lines formatted as an MS Excel table.
- The solution is accurate and easy to use. Allows the user to generate reports of various complexities with complete independence.
- SAC is part of a data analytics tool, similar to Microsoft's tool, Power BI.

The key features of SAC

Business Intelligence

Data Access

SAP Analytics Cloud allows data to be available in two ways: Importing Data and Direct reading from database / ERP.

In the case of direct reading, real-time data is updated only by connecting to the SAP system (HANA, BW, BPC, or another).

Data Modeling

This allows for data modification with additional features such as indexing, formulas, and converting values for acquired characteristics.

Data visualization

This feature allows you to view and highlight your data with graphs, tables, or other graphic objects.

Predictive Analysis

To highlight relevant metrics or predict future results, SAP has introduced artificial intelligence. A few examples of using these new features include automated content creation with intelligent detection, robust feature identification, uncompromising record acquisition, and much more.

Planning

The SAC planning process lets you consolidate your financial and operating statements with the same tool to have a complete and clear view of your current situation.

Types of budget planning can be developed and modified from existing financial statements. The machine learning tools provided and predictive tools can help you anticipate your money in the future. Also available is the famous "What If" mode of simulation.

What does the future hold for SAC?

SAP is focused on constantly improving SAC. Like a cloud tool, it is continuously updated and visible at no additional cost. SAP's future strategy for SAC can be defined by implementing the solution: BI allocation and planning.

BI offer

The integration of the SAC data recognition component into other SAP Cloud solutions furthers the synchronization of the reporting style and simplifies its use from one tool to another. Examples include SAP S / 4HANA Cloud Integrated Analytics and a new cloud-based data storage tool, SAP Data Warehouse Cloud.

In terms of SAP On-Premise solutions, SAP Business Objects remains the preferred option, although hybrid solutions with On-Premise ERP and cloud-based solutions are becoming increasingly available.

By planning

SAP strategy to prioritize cloud solutions for editing tools. The use of SAP Analytics Cloud is recommended for scheduling without the need for legal consolidation.

In addition, its collaborative planning tools provide an additional value for updating forecasts in each company category, and its integration with SAP S / 4HANA allows for easy data integration.

However, for customers using the On-Premise solution via SAP BPC, the extension is possible with SAC. Even improved data usage is recommended.

In short, SAP develops SAP Analytics Cloud across all segments offered, whether it reports on its new solutions or plans in conjunction with other tools in its category. With SAP's focus on improving the device, SAC has a bright future ahead of it.

3.3.6 Multidimensional Schema

Multidimensional Schema is specifically designed to model data storage systems. Schemes are intended to address the unique needs of more significant data intended for analysis (OLAP) (David Taylor, 2021).

The following are the three main types of broad range schemes with their unique benefits.

Star Schema

Snowflake Schema

Galaxy Schema

Star Schema

The schema in the database, where the star center can have one table of facts and many tables of the corresponding size. It is known as the star schema as its structure resembles a star. The Star Schema data model is the simplest type of Data Warehouse schema. Also known as the Star Join Schema and designed to query big data sets.

Snowflake Schema

Snowflake Schema in the database is a logical layout of tables in various databases, such as an ER diagram similar to a snowflake. Snowflake Schema is an extension of Star Schema and adds extra dimensions. Typical size tables divide data into additional tables.

Galaxy Schema

The Galaxy Schema contains a table of two facts that share size tables, also called the Fact Constellation Schema. The schema is considered a constellation hence the name Galaxy Schema.

Star Cluster Schema

The snowflake schema contains fully expanded hierarchies. However, this can add complexity to the Schema and requires additional joining. The star schema, on the other hand, contains

Completely broken hierarchies, which can lead to unemployment. So, the best solution would be to balance the two systems, which is the Star Cluster Schema design.

Passive sizes can be found as forks in hierarchies. Separation occurs when a business operates as a parent in two different categories. Fork organizations are then classified as relationships from one to many.

3.3.7 Online analytical processing

OLAP (online analytical processing) is a computer program that enables users to quickly and selectively extract and query data for analysis in different viewing environments. OLAP intelligence inquiries often assist in trend analysis, financial reporting, sales forecasting, budgeting, and other planning purposes (Jacqueline Biscobing, 2020).

For example, a user may request that the data be analyzed to show a spreadsheet showing all marine company products sold in Florida in July, compare revenue figures for similar products in September, and see comparisons of other product sales in Florida at the same time.

How OLAP programs work

To facilitate this type of analysis, data is collected from multiple sources, stored in data repositories, and then processed and sorted into data cubes. Each OLAP cube contains data categorized by size (customer, region, and time sales region) based on size tables in the data repositories. Members then fill the measure (such as customer names, countries, and months) sorted by category. OLAP cubes are often summed up before size to improve question time in the relationship database significantly.

Analysts can work on five types of OLAP analytics tasks with the following databases:

Roll up: - Also known as merging or drilling, this function summarizes the data in size.

Drill down: - This allows analysts to navigate deeper between data sizes, for example, drilling down from the “period” to “years” and “months” to change product sales growth.

Slice: - This allows the analyst to take one level of information to present, such as "selling in 2017."

Dice:-This allows the analyst to select data from a wide range of sizes for analysis, such as "the sale of Iowa blue balls in 2017."

Pivot: - Analysts can get a new view of the data by exchanging cube data axes.

The OLAP software then finds a combination of sizes, like all products sold in the Eastern region more than a certain amount at a particular time, and it shows. The effect is "average"; each OLAP cube has at least a few or a hundred steps found in the information stored in the fact tables in the database.

OLAP (online analytical processing) systems typically fall into one of three categories:

Multidimensional OLAP (MOLAP) is an OLAP that points directly to a multidimensional database.

Relational OLAP (ROLAP) is an OLAP that performs a robust analysis of data stored in a related database.

Hybrid OLAP (HOLAP) is a combination of ROLAP and MOLAP. HOLAP was developed to integrate a large amount of ROLAP data and MOLAP processing power.

Uses of OLAP

OLAP can be used for data mining or the discovery of previously unimaginable relationships between data objects. The OLAP website does not need to be as large as a database, as not all transaction information is required for practice analysis. Using Open Database Connectivity (ODBC), data can be imported from existing relationship information to create various OLAP data.

OLAP products include IBM Cognos, Oracle OLAP, and Oracle Essbase. OLAP features are included in tools such as Microsoft Excel and Microsoft SQL Server's Analysis Services). OLAP products are usually designed for multi-user locations at a software cost, depending on the number of users.

3.3.8 ETL (Extract, Transform and Load)

Definition of ETL

ETL is a process that extracts data from different source systems, and then converts data (such as data entry, solutions, etc.) and finally uploads data to the Data Warehouse system. The complete form of ETL is Extract, Transform and Load. (Jacqueline Biscobing, 2020).

It is tempting to assume that creating a Data Database simply removes data from multiple sources and uploads it to the database. This is far from the truth and requires a complex ETL process. The ETL process requires active input from various stakeholders, including developers, analysts, testers, senior managers, and is a technological challenge.

To maintain its value as a decision-making tool, the data storage system must adapt to business transformation. ETL is a regular (daily, weekly, monthly) data storage system and requires speed, automation, and precision.

ETL process in Data Warehouses

Step 1) Extraction

In this ETL architectural step, data is extracted from the source system at the stage. Any changes are made to the scene so that the functionality of the source system is not compromised. Also, retrieval will be challenging if corrupted data is copied directly from the source to the final database. Storage space allows you to verify the extracted information before entering the database.

The database needs to integrate systems with different DBMS, Hardware, Operating Systems, and Communication Protocols. Sources can include asset applications such as Mainframes, customized applications, point of communication devices such as ATMs, call switches, text files, spreadsheets, ERP, vendor data, partners, among others.

So one needs a logical data map before the data can be extracted and physically loaded. This data map describes the relationship between sources and targeted data.

Three Ways to Extract Data:

- Full Extraction
- Partial Extraction- without update notification.
- Partial Extraction- with an update notification

Regardless of the method used, extraction should not affect the performance and response time of source systems. These are live source information systems. Any drop or lock may affect the company's bottom line.

- Particular confirmations were made at the time of release:
- Link records with source details
- Make sure no spam / unwanted data is uploaded
- Type of data check
- Delete all types of duplicate/split data
- Check whether all buttons are present or not

Step 2) transformation

The data extracted from the source server is raw and unusable in its original way. It, therefore, needs to be cleaned, mapped, and modified. This is a crucial step where the ETL process adds value and converts data such as intelligent BI reports can be generated.

It is one of the essential ETL concepts in which you include a set of activities in the extracted data. Data that does not require a change is referred to as direct or indirect data transfer.

In the transition phase, you can perform customized tasks on data. For example, if a user wants sales revenue not in the database, or if the name and surname on the table are in different columns. It is possible to install them before uploading.

Validations are done during this stage

Sorting - Select only specific columns to

upload Data set-up rules, and tables are used

Character Set Conversion and encoding handling

Conversion of units of measurement such as daily time conversion, currency conversion, price conversion, etc.

Data limit verification testing. For example, the age may not be more than two digits—

data flow verification from the stage location to the central tables.

Required fields should not be left blank.

Cleaning (for example, making a NULL map to 0 or Gender Male to “M” and a woman to “F.” etc.)

Divide the column into duplicates and combine multiple columns into one column.

Rows and columns are passed,

Use the look to combine data

Any complex data verification is used (e.g., if the first two consecutive columns are empty and automatically reject the queue from processing)

Step 3) Loading

Uploading data to a targeted data database is the final step in the ETL process. In a typical data house, a large volume of data needs to be loaded in a short time (night). Therefore, the upload process should be designed to work.

In an upload failure, recovery methods should be configured to restart from the failure location without data loss. Data Warehouse administrators need to monitor, resume, cancel loads as existing server operations.

Types of loading

- Initial load - fills all Warehouse Data tables
- Incremental Load - to apply continuous changes as needed from time to time.
- Full Refresh - open one or more table contents and reload with new details.

3.4 Overview of the Business Analytics

Business Analytics is how businesses use technology and technology to analyze historical data to gain new insights and improve strategic decision-making. Business analytics, data management solution, and subset intelligence means using methods such as data mining, forecasting statistics, and statistical analysis to analyze and convert data into useful information, identify and anticipate trends and results, and ultimately make wise business decisions driven data.

Critical elements in the standard business analytics dashboard include:

Data aggregation: before analysis, data must first be collected, organized, and sorted, either by voluntary data or by-product records

Data mining: digging data for business statistics using big data sets using data, statistics, and machine learning to identify trends and build relationships

Association and Sequence Identification: identification of unexpected actions performed in conjunction with other activities or sequences

Text Mining: scans and organizes large, informal data sets for quality and quantity analysis

Forecasting: analyzes historical data from time to time to make informed predictable estimates of future events or behavior

Predictive Analytics: Predictive business forecasts use a variety of statistical methods to create predictive models, extract data from databases, identify patterns, and provide predictive results for multiple organizational outcomes

Optimization: once trends have been identified and forecasts have been created, businesses can engage in simulation strategies to test the best conditions

Data Visualization: provides visual presentations such as charts and graphs for easy and fast data analysis

Other examples of business statistics include operating and managing clinical information systems in the healthcare industry, tracking player spending and building casino maintenance efforts, and simplifying fast restaurants by monitoring customer hours and indicating that certain food items should be prepared based on meeting time.

Modern, state-of-the-art software analytics software and platforms are developed to capture and process large data sets that businesses encounter and use for efficient business operations. BI or data mining systems provide business users with simple tools for analyzing fixed, structured data. One of the advantages of the main things of such programs is the availability of visual and analytical data tools for multiple users (Jackson, 2018).

Business Analytics vs. Data Analytics

Data analytics is a broad umbrella term for the science of raw data analysis to convert that data into useful information where styles and metrics are found. While both business statistics and data analysis aims to improve efficiency, business analytics focused on business operations and data analytics are more focused - both business intelligence and online analytics reporting and processing (OLAP) are under the umbrella of data analytics.

Data scientists, data analysts, and data engineers work together in the data analytics process to collect, compile, and edit data for the development, testing, and review of analytical models, to ensure accurate results. Data analysis for business purposes is indicated by its focus on specific business performance questions.

4 Practical Part

4.1 Research questions

Q1: What is the best tool for building visual dashboards?

Q2: What are the significant difference between Tableau and Power BI?

4.2 The structure

Today, many BI platforms and data viewing tools make it easy to analyze data, provide statistics, display it on screen, and share it with other users.

In this part of the thesis, two popular and widely used methods of BI platforms are compared. The methods reached in theory include power BI and tableau. Further, the idea also compares their key parameters:

- Accessibility
- Usability and visibility
- Differences in dashboard design
- Features form recognition and their strategies

Power BI and tableau are cloud-based business analytics tools, both below Office 365 Enterprise E5 version. Both of these programs are good programs in their way. It is, therefore, interesting to note the pros and cons being used in studying the case study.

Basic argument: Ever since the decisions were made in business are complex as there is no ready-to-use shelve software available. The reason is that every user will lead it differently.

Operational efficiency: Better performance efficiency in construction, the design, and comparison of the following dashboards, can be very helpful to use questions and answers to provide better analytics because it is easier for the reader. According to a case study aimed at presenting inflow of foreign direct investment of India considering an important decision for companies, economists and researchers, therefore, Q is also related A's will be awarded.

4.3 Description of a case study

Case study: The development and preparation of the visualization dashboard for research scholars, economists, and managers of the company regarding making investment decisions and having a better understanding of the inflow of foreign direct investment in India.

Scenario: The researcher is required to present the foreign direct investment data of India in such an innovative way that it can be quickly and understood by research scholars, managers, economists, and any individual.

Research scholars, economists, and managers of well-established companies need sound-processed data that match their requirements and expectations. Therefore, it is required to formulate creative ideas for presenting the raw data and prepare a report based on data received from the annual information of the Reserve Bank of India concerning international investors, economists, researchers, companies that have already invested their money abroad. The data report should be informative and understandable for rapid analysis of the data, as international investors or companies often do not understand the data of FDI available in reports issued by the country's central bank. They invest based on the information or documentation that aims to help an individual better understand FDI inflows in India.

How to help these research scholars, economists, and managers? A modern and sophisticated process is used to create an interactive dashboard that explains the general data and is clear to the end viewer for analysis to interpret the data intelligently. The critical interest of the economists or researchers is obvious to get easily understandable FDI data that can be used to conduct further research or make predictions about future FDI inflows in India. Apart from economists and researchers, the managers of international companies will also know about FDI data for investing-related decisions in India. Many viewing tools can visualize big data, the most popular tool on the market described in literature reviews and the types, purposes, and methods of building such visuals.

The dashboards should rapidly explain to end-user the most critical indicators in general. For example, the inflow of FDI, the number of sectors having FDI inflow, and top deals taking place in several industries and other related information. The result of the

The thesis should provide pros and cons to the economists, researchers, and managers that whether it is good to invest in India or whether India is performing well in terms of FDI or not. For example, suppose any foreign company is planning to invest in India for business. In that case, it should make sure that it has potential growth opportunities and will generate good returns with lower risk.

This thesis will primarily compare Power BI and tableau dashboards, analyzing the data from authenticated database sources. It will also give suggestions for further improvements in creating visual dashboards. The result of dashboards will show visible differences and indicate the benefits of these used tools for analysis.

4.4 Sources of data

4.4.1 Database on Indian Economy - Reserve Bank of India's Data Warehouse

The RBI regulates the financial and other banking policies of the Indian government. The Reserve Bank of India (RBI) was established on April 1, 1935, following the Reserve Bank of India Act, 1934. The Reserve Bank has been situated in Mumbai since 1937.

The Reserve Bank of India (RBI) has a rich tradition of publishing information on various aspects of the Indian Economy through its few publications. Through this website (Database on Indian Economy), the information is presented in a series of reports formatted chronologically. These reports are organized into sectors and sub-sectors based on their periodicities. Messages can be saved as excel sheets for further analysis. The Database on the Indian Economy provides data on various aspects of the Indian economy, such as banking and finance.

RBI Annual publications

Reserve Bank of India's data warehouse provides information on various topics such as:

Annual Report - The annual report is the official report of the Reserve Bank of India issued annually. The report contains estimates of India's economic progress. An overview of the economy, the performance of the Reserve Bank for that year, and the vision and proposed plan for next year's RBI and the Reserve Bank's annual accounts.

Report on Trend and Progress of Banking in India - This report analyzes last year's financial policies and developments.

Lectures - The Reserve Bank of India creates three annual courses. Two of these speeches were made by the Governors of the Past Bank, and a well-known economist conducted one study.

Report on Currency and Finance - This report is written and presented by the Reserve Bank of India staff and focuses on a specific theme and gives a detailed economic analysis on the theme related to the article.

Handbook of Statistics on the Indian Economy - This report is an important step the Reserve Bank took to improve data distribution. It is a repository containing great mathematical details.

State Finances: A Study of Budgets - The report is an essential source of government-funded financial data and provides insights into data conducted on the state of the national financial system throughout India. This data entry is used to analyze specific issues of importance.

Statistical Tables Relating to Banks in India - This annual publication contains full-time details regarding the Commercial Banks of India (SCBs). The report also looks at the details of the balance sheets and performance indicators for each SCB in India. The journal also includes separate data sources on critical banking experience, banking group intelligence, and state intelligence level.

Basic Statistical Returns - This is another annual data-focused magazine that represents more detailed information on the number of offices, staff, deposits, and credit of Organized Commercial Banks with minute details such as regions, governments, and intellectuals reading the details. This information also applies to the needs of the people and the debts of each bank.

4.4.2 Description of the data used

The data used for the analysis includes information about the inflow of foreign direct investment in India for the monitored period of the preceding five years, such as from 2015 to 2020. In table 1, there is a table indicating the inflow of foreign direct investment in each sector (see table 2 in the Appendix).

It includes:

➤ **Sectors (US \$ million)**

- The name of the various sectors, e.g., "Manufacturing."

➤ **2015-16**

- The amount of foreign direct Investment in 2015-16 is mentioned in the US \$, e.g., 8,439\$ million.

➤ **2016-17**

- The amount of foreign direct Investment in 2016-17 is mentioned in the US \$, e.g., 11,972\$ million.

➤ **2017-18**

- The amount of foreign direct Investment is mentioned for 2017-18 in the US \$, e.g., 7,066\$ million.

➤ **2018-19**

- The amount of foreign direct Investment is mentioned for 2018-19 in the US \$, e.g., 7,919\$ million.

➤ **2019-20**

- The amount of foreign direct Investment in 2019-20 is mentioned in the US \$, e.g., 8,153\$ million.

4.5 Summary of creating the dashboard in Power BI

4.5.1 The types of chart

It is essential to define what types of charts are supported by Power BI based on the primarily used ones listed in the literature review. (Data Visualization, Microsoft Power BI, 2020)

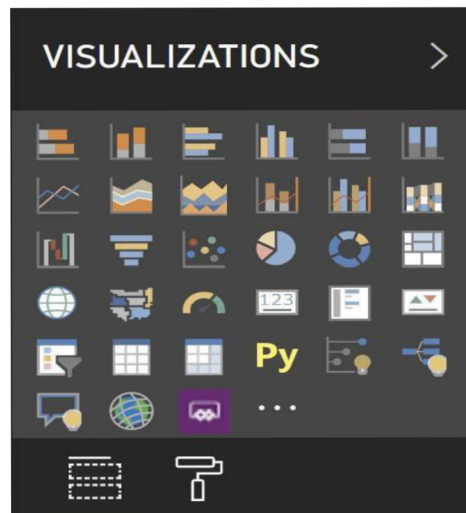


Figure 22: The possible chart types supported by Power BI

- Stacked Bar and Column Charts
- 100% Clustered Bar and Column charts
- Clustered Bar and Column Chart
- Line and Area Charts
- Line and Stacked / Clustered column charts
- Waterfall chart
- Scatter and Funnel Charts
- Pie and Donut charts
- Treemap chart
- Map and 3-D filled map
- Other items such as Gauge, Card, Multi-Row card, KPI11, Slicers, and Interactive Tables to describe the visuals effectively.

4.5.2 The process of importing the data in Power BI software

The first step for importing the data is to run the desktop application or online Power BI software. In the following figure, 25, the desktop of the Power BI application is shown. Power BI Desktop is a free program that installs your local computer to connect, modify, and visualize your data.

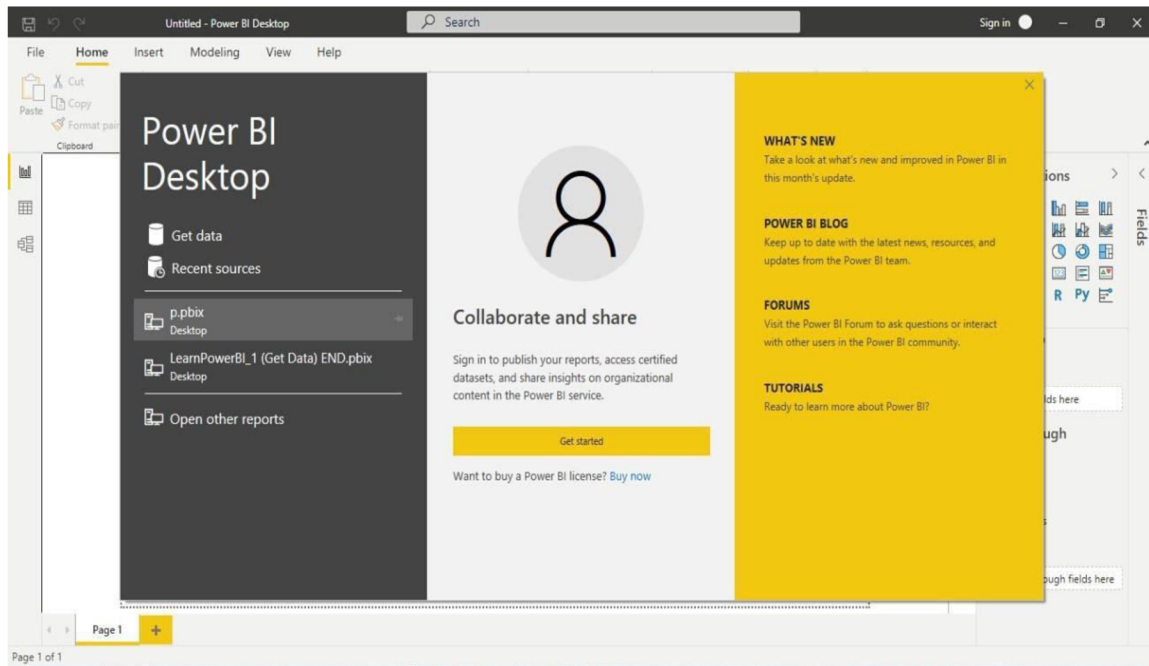


Figure 23: Power BI desktop (Own picture)

With Power BI Desktop, you can connect to various data sources and integrate them (commonly called modeling). This data model allows you to create visuals and collections that you can share as reports with other people within your organization. Many users working on business intelligence projects use the Power BI Desktop to create reports and then use the Power BI application to share their accounts with others.

After running the desktop, the home menu will get open. In figure 24, the home menu of the Power BI application is shown. The home menu of Power BI software allows you to create your visuals by inserting data. The data can be inserted using the get data option.

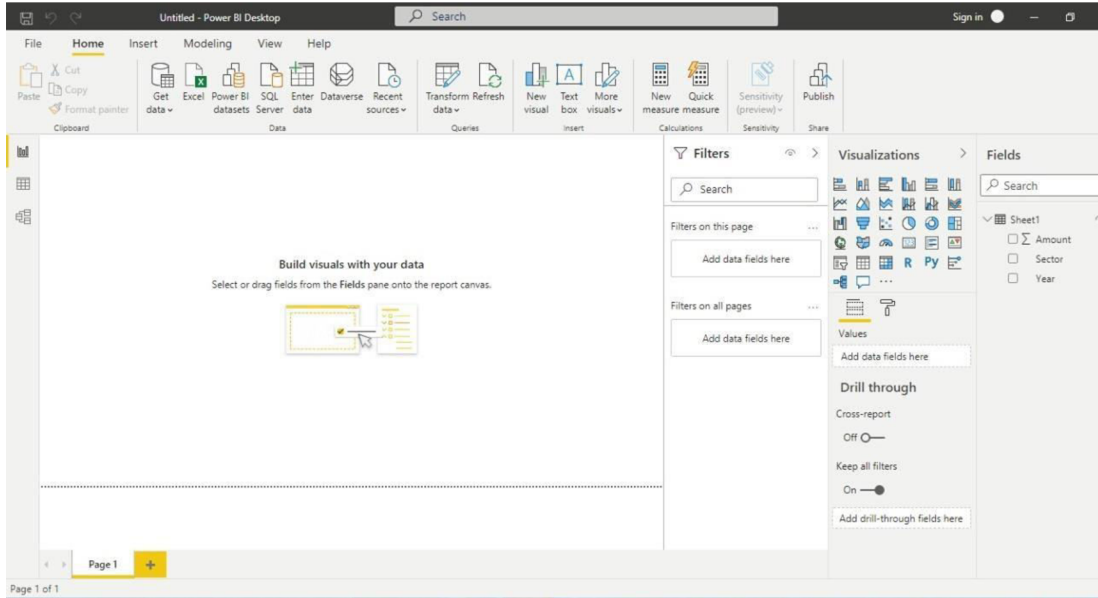


Figure 24: Home menu of Power BI (Own picture)

The home menu includes various options such as to get data, excel, Power BI datasets, and many more. It has different visualizations which can be used for presenting the data uniquely. It also enables the user to filter the data used for preparing the report.

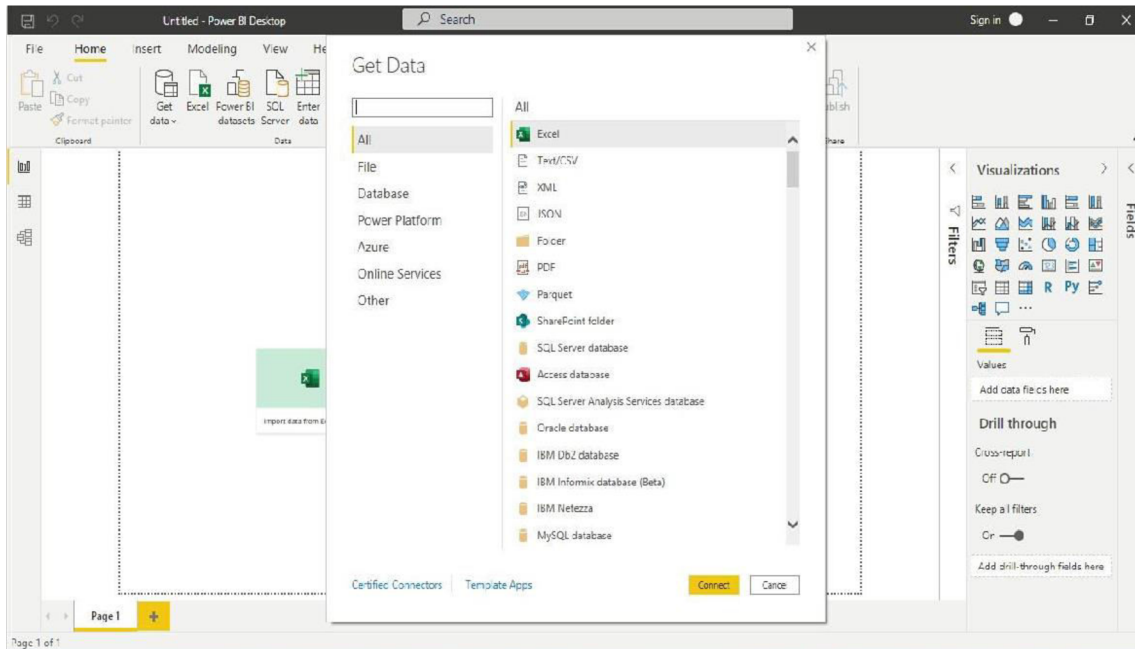


Figure 25: The Sources of importing the data (Own picture)

In figure 25, the sources for importing data are mentioned. There are many sources for inserting the data into Power BI software. From the Home ribbon, select get data for inserting the data.

The get data window will appear, indicating several categories to which Power BI desktop can connect. The data can be imported from sources like Excel, Text/CSV, and XML files. Besides this, it is also possible to import data through Power BI desktop using Azure SQL database and many more.

Note:

For example, intending to import excel data, the aggregate data is formatted in table form for Power BI to identify the data and bifurcate it into particular sections for providing visualizations. The data for analysis is imported by selecting all options from the get data window. After choosing all possibilities, Excel is selected, and then connect option has to be selected.

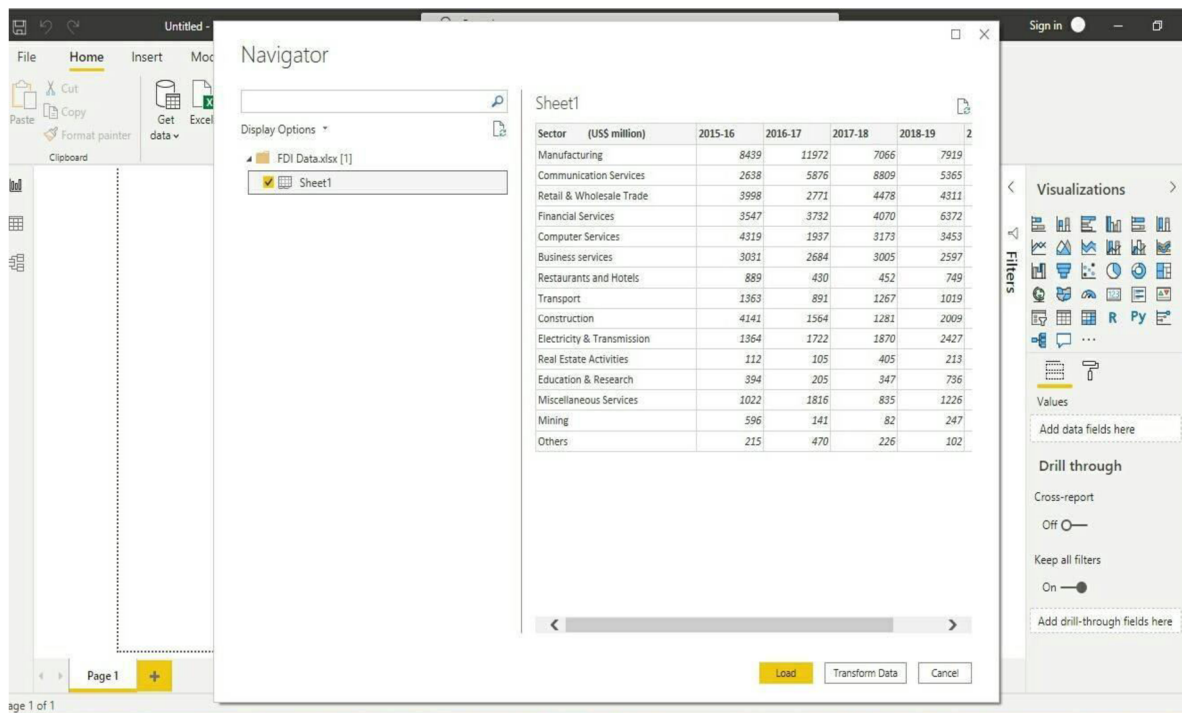


Figure 26: The Process of inserting the data (Own picture)

In figure 26, the process of inserting the data is shown. After clicking on the connect button, the FDI Data.xlsx file is selected to insert the data into the Power BI application.

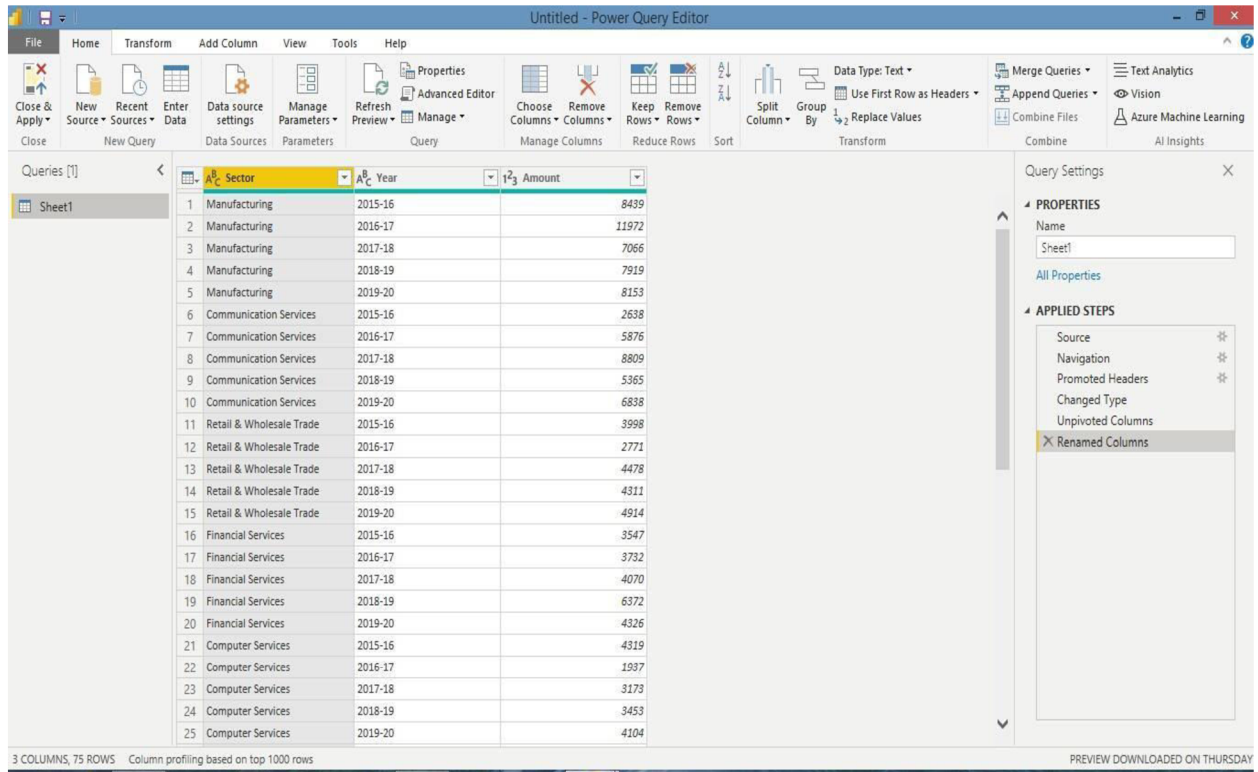


Figure 27: The transformation of importing data (Own picture)

Figure 27 indicates the transformation of imported data. Once the data file used for the analysis is selected, it will be transformed to create the dashboard.

4.5.3 The process for the creation of a dashboard

The next important thing is to help the Power BI application identify fields related to the inflow of foreign direct investment. Therefore, the data is required to be categorized and divided into particular areas.

The process of designing the dashboard

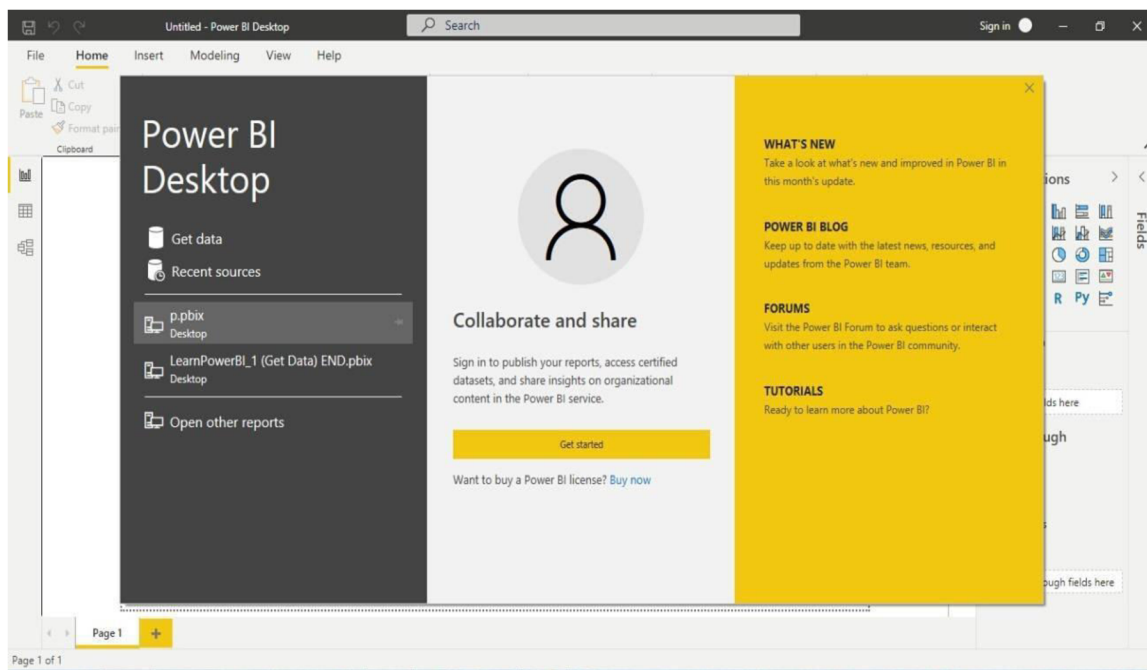


Figure 28: Power BI workspace desktop (own picture)

Figure 28 includes a report view, and there is a workspace in Power BI Software. Moreover, the figure also indicates the essential parts to be emphasized for performing analysis.

In the figure, there is a white-colored empty sheet where the dashboard is going to be created. The maps and charts are going to appear on a blank white sheet automatically. Further, it is possible to move it in any place on a sheet.

The white sheet appears in A4 format to understand the page's parameters and move the visual charts; however, he wants to.

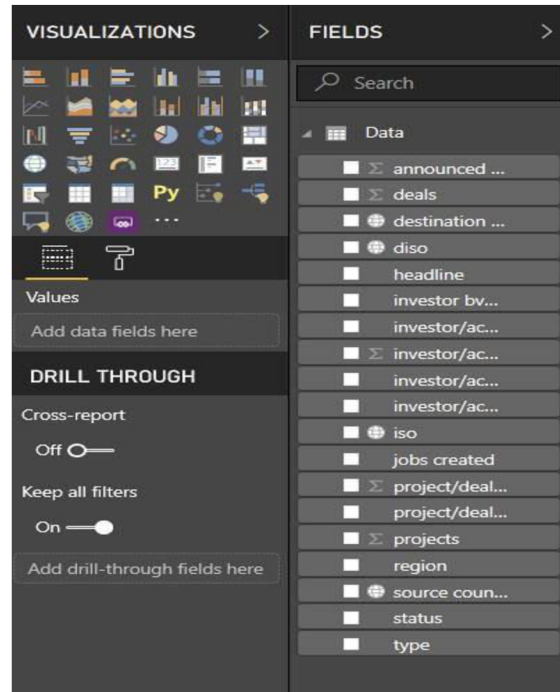


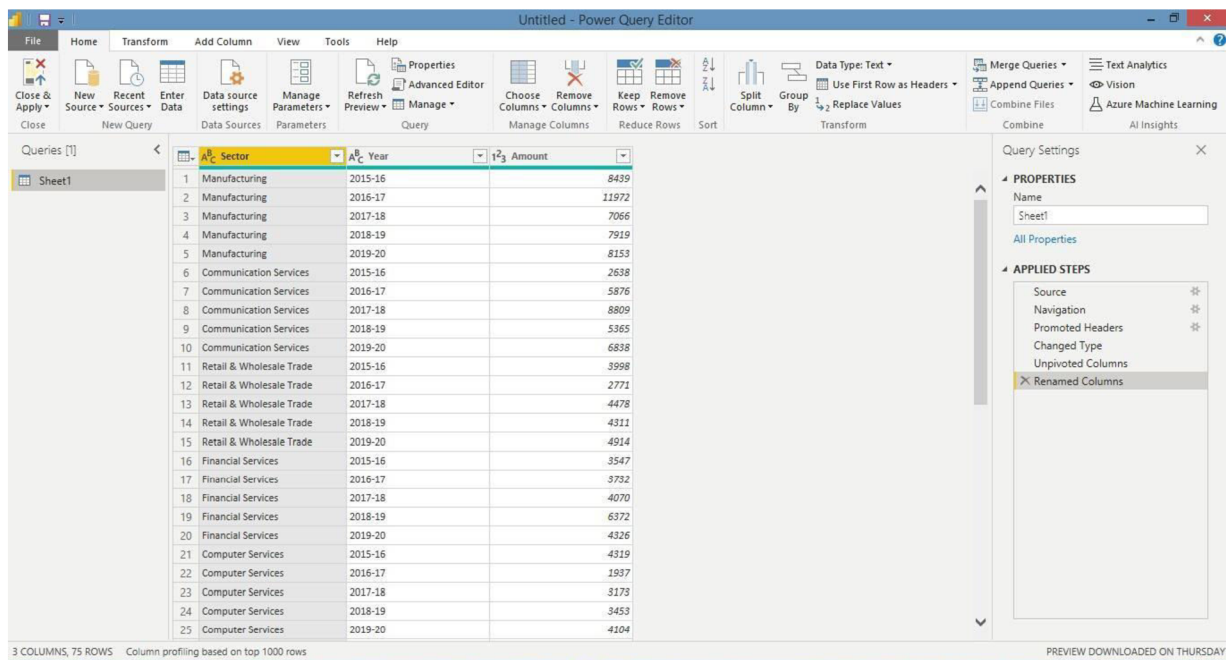
Figure 29: Power BI workspace instruments for visualization

In figure 29, it is possible to have a closer look at the board. There are icons such as fields and visualization.

The visualization icon includes different types of charts that can represent the data graphically in an interactive way.

There are fields situated on the right-hand side of the figure. In a visualization section, choose any chart and click on that chart to create a chart and then select one or more fields for running the connection between data and present it through a chart. Different charts represent the data differently in respective charts mentioned in the literature review like bar charts, maps, pie charts, and many more.

The steps for creating a chart



	Sector	Year	Amount
1	Manufacturing	2015-16	8439
2	Manufacturing	2016-17	11972
3	Manufacturing	2017-18	7066
4	Manufacturing	2018-19	7919
5	Manufacturing	2019-20	8153
6	Communication Services	2015-16	2638
7	Communication Services	2016-17	5876
8	Communication Services	2017-18	8809
9	Communication Services	2018-19	5365
10	Communication Services	2019-20	6838
11	Retail & Wholesale Trade	2015-16	3998
12	Retail & Wholesale Trade	2016-17	2771
13	Retail & Wholesale Trade	2017-18	4478
14	Retail & Wholesale Trade	2018-19	4311
15	Retail & Wholesale Trade	2019-20	4914
16	Financial Services	2015-16	3547
17	Financial Services	2016-17	3732
18	Financial Services	2017-18	4070
19	Financial Services	2018-19	6372
20	Financial Services	2019-20	4326
21	Computer Services	2015-16	4319
22	Computer Services	2016-17	1937
23	Computer Services	2017-18	3173
24	Computer Services	2018-19	3453
25	Computer Services	2019-20	4104

Figure 30: The data used for creating the Power BI dashboard (own picture)

The most significant thing in the data will be analyzed in the entire dashboard to create a chart. In this case, the amount of foreign direct investment in each sector for a particular year is shown in figure 30.

In Power BI, Stacked Bar Chart helps compare several dimensions against a single measure. The following step is to create an interactive chart to know about the amount of FDI inflow in each sector mentioned in the field. Therefore, select for making a stacked bar chart in Power BI, first drag and drop the sectors fields section to canvas region. It will automatically create a column chart.

After that, click on the stacked bar chart under the visualization section, which will automatically convert the column chart into a stacked bar chart. Next, add the no. of year's field to the legend section for converting it into a stacked bar chart. In the figure, the x-axis indicates the total amount of FDI inflow in the past five years, and the y-axis shows the name of sectors.

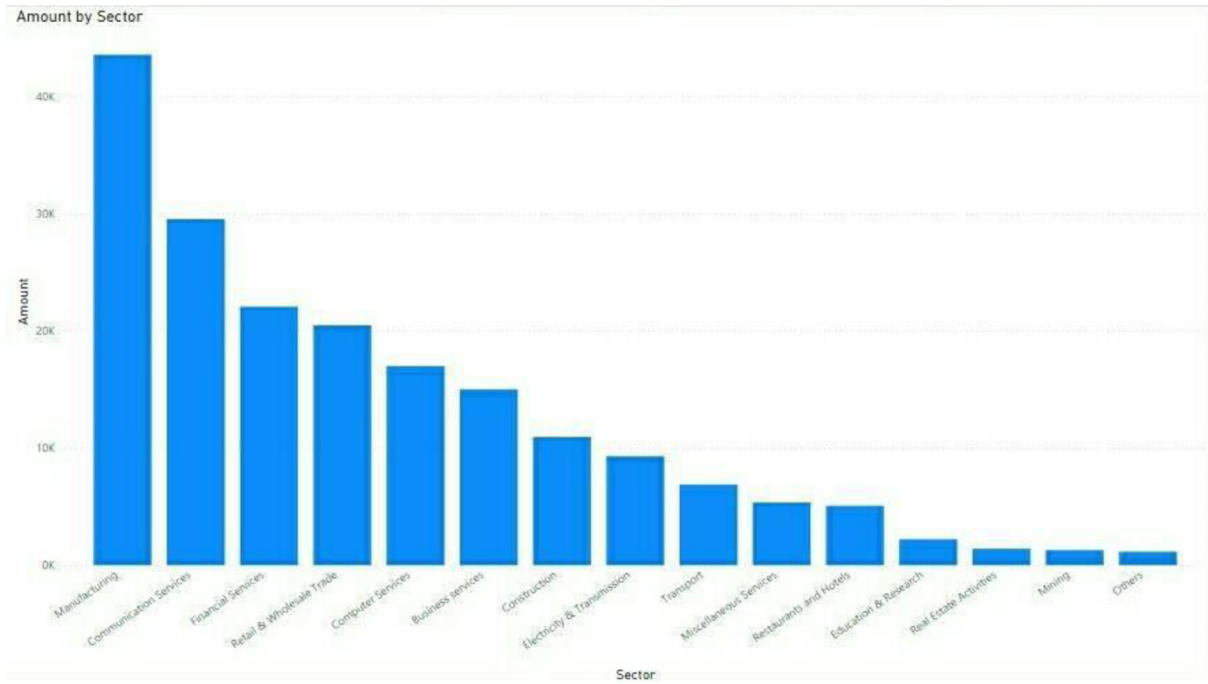


Figure 31: Interactive visualization Stacked Bar Chart in Power BI (own picture)

The following step is to create an interactive line chart table that provides information about FDI inflow sector-wise in the past five years.

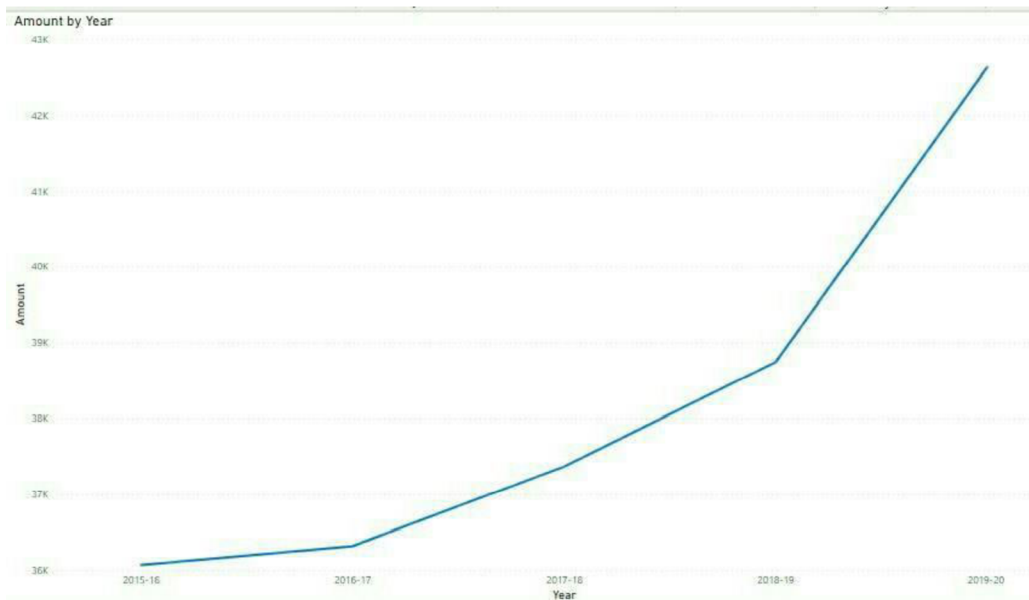


Figure 32: Interactive visualization Line Chart in Power BI (own picture)

Figure 32 indicates the line chart for FDI inflow in India. To create a line chart, drag the sectors field section to the canvas region, add the number of years field to the axis section, or simply checkmark the year's column.

After that, it creates a column chart automatically. Select a line chart option under the visualization section; it will automatically convert the column chart into a line chart.

Intending to provide an excellent analytical picture, it requires to use of the different types of charts like

- Funnel chart
- Pie chart
- Waterfall chart
- Treemap

Select the chart from the visualization section and move the units required to display in the funnel chart to create a funnel chart. The figure of the funnel chart indicates the amount of foreign direct investment made in each sector.

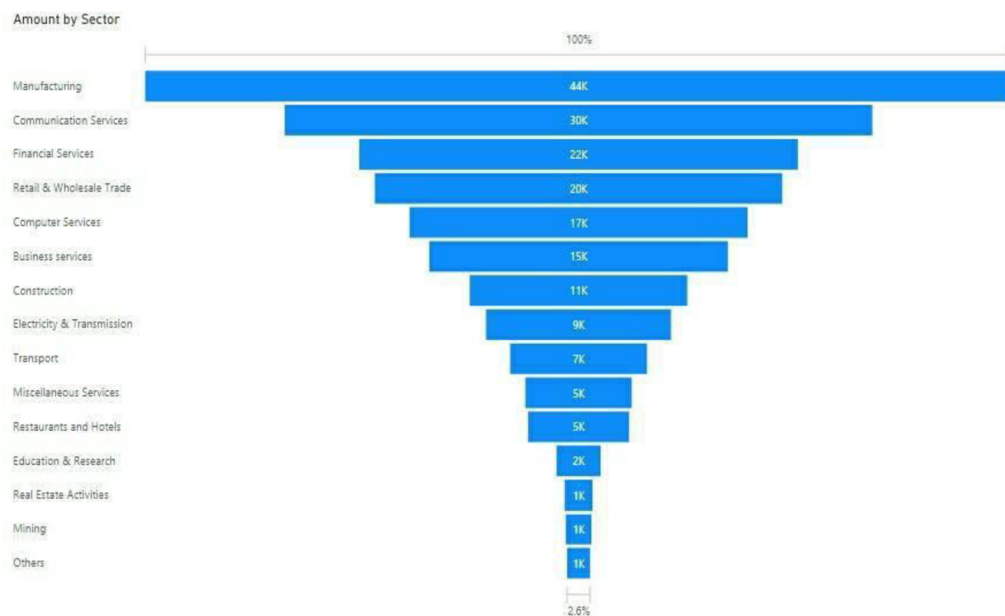


Figure 33: Funnel chart in Power BI (own picture)

Here, the funnel chart shows that the manufacturing sector has the highest foreign direct investment of around 44 million US \$ compared to other industries, followed by communication services, financial services, and so on.

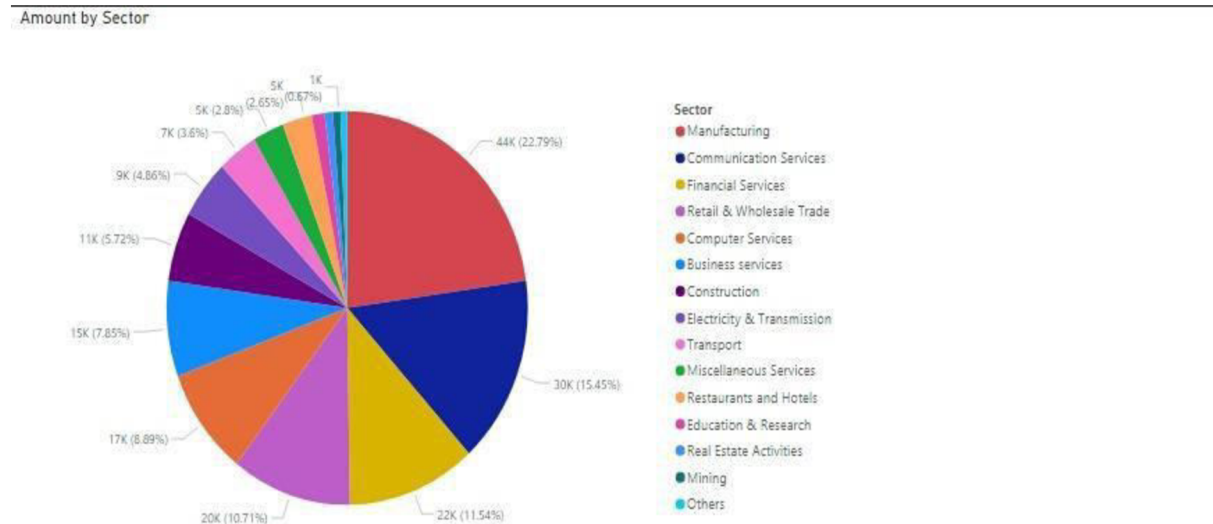


Figure 34: Pie chart in Power BI (own picture)

The following chart in the dashboard is the pie chart. Let's try to develop a chart that includes detailed information about the inflow of foreign direct investment for each sector in the preceding five years.

In figure 34, the pie chart indicates the total inflow of foreign direct investment for each sector is displayed in digits and percentages. In this type of chart, there are no details needed, so let's keep it empty.

Nevertheless, the pie chart shows the percentage of FDI inflow in each sector of India. Each industry is represented using different colors. On the right-hand side of the figure, the name of each initiative is mentioned.

Waterfall charts indicate a running total as Power BI adds and subtracts values. In the figure, the waterfall for foreign direct investment in India for each sector is represented.

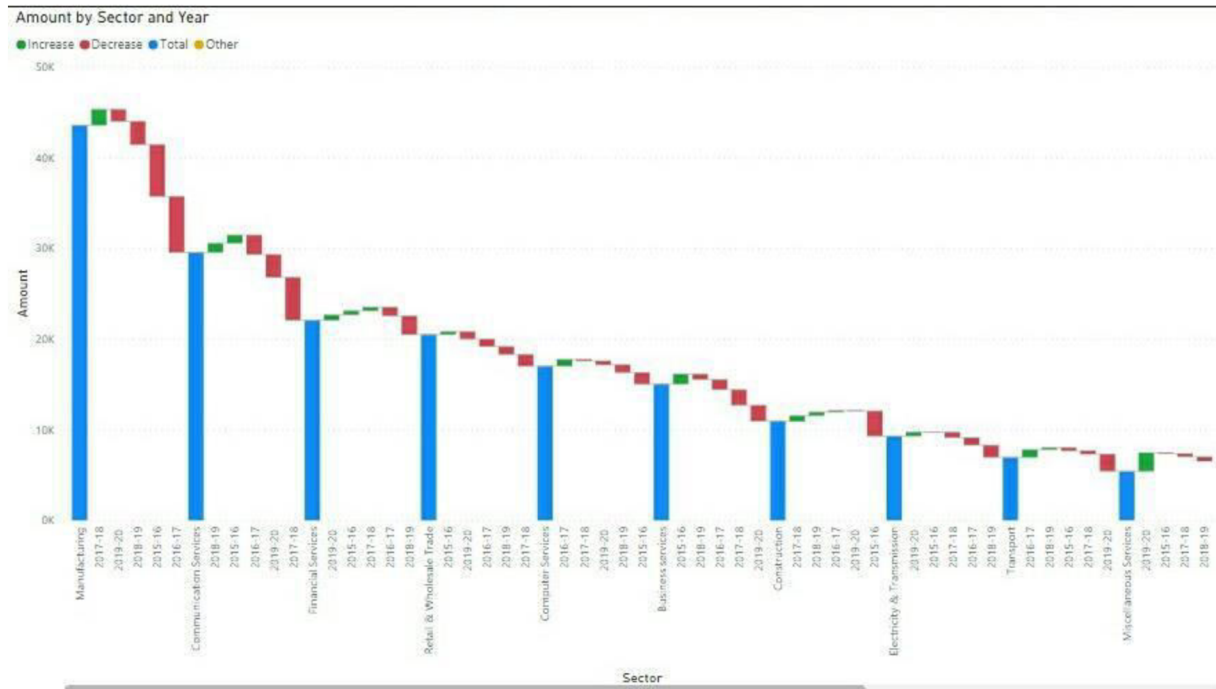


Figure 35: Waterfall chart in Power BI (own picture)

For creating a waterfall chart, first, select the name of the sectors field section and then select the no. of year's field section. Afterward, select the waterfall icon from the visualization section.

In the waterfall chart, the x-axis represents the name of sectors and years. In contrast, the y-axis displays the inflow of foreign direct investment in India for the respective year in each industry.

The last chart in this dashboard would be the treemap chart. Let us insert a chart that includes detailed information about the inflow of foreign direct investment in India for the previous five years respectively.



Figure 36: Treemap chart in Power BI (own picture)

In the figure, the treemap chart is displayed. The treemaps display sequence data as a set of rectangular structures. Each hierarchy level is represented by a colored rectangle (branch) consisting of small rectangles (leaves). Power BI supports the size of the space within each rectangle than the approximate value. Rectangles are arranged in length from top left (large) to bottom right (very small).

For example, the figure indicates a treemap chart. The treemap chart represents the different sizes of rectangles according to foreign direct investment in each sector. The sector with a higher foreign direct investment is represented with a big rectangle and vice versa.

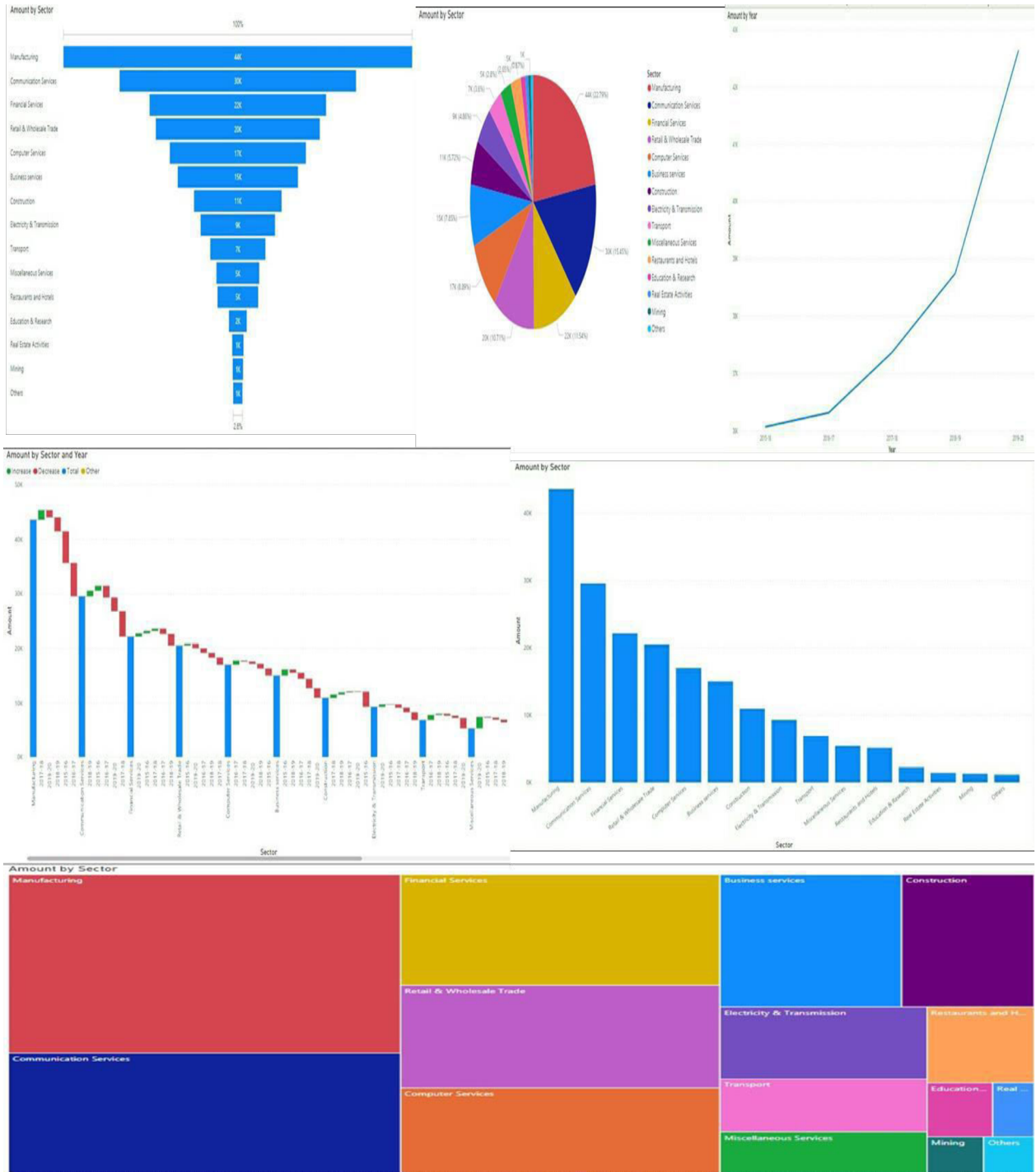


Figure 37: Own dashboard created in Power BI

4.6 Difference between Power BI and Tableau

Table 1: Different between Power BI and Tableau

S.R. No.	Basis of comparison	Power BI	Tableau
1.	Data Size & Data capability	All groups or workstations can handle up to 10 GB of data in Power BI and more than 10 GB of data in the cloud or Azure. If we are talking about a local database, then Power BI simply pulls data and does not import it.	In the case of Tableau, it operates in a column base structure that allows different values to be stored in each column. This type of organization will enable billions of records or lines to be downloaded.
2.	Data source Access	Power BI blocked access to other servers and data (SQL, SAP HANA, Oracle Database, etc.)	Tableau has access to a large number of different servers and data sources, including JSON, PDF, accurate statistics reports, Microsoft documents, SQL Server, MySQL, Oracle Redshift, the list goes on.
3.	Integration	Power BI can work seamlessly with a variety of data sources such as SharePoint, Azure power flow, power applications, 365 Excel, and many other Microsoft products.	Tableau is used by several release platforms, including Microsoft-based platforms and tools based on an easy-to-use method.
4.	Data Shaping	Power BI provides a multi-functional quiz editor that helps you do many things quickly.	Tableau does not offer this type of service.

5.	Data Modeling	Power BI provides Power Pivot dots that are easily accessible to the user.	While Tableau does not offer the facility of data modeling.
6.	Support Level	Customer support for users with free Power BI accounts is limited. The app firmly provides support documentation and resources.	Tableau provides its users with extensive assistance through the knowledge base.
7.	Infrastructure	Power BI is available in three stages. Desktop, mobile, and service. The basic setup is Azure Tenant.	Tableau makes it easy to share the results generated on the Tableau desktop over Tableau Online or Tableau Server.
8.	Speed	Power BI is easy to use. It is faster and more efficient when data volume is limited. Power BI tends to slow down while handling large amounts of data.	The tableau can handle large amounts of data quickly. It is fast and offers comprehensive features for visualizing data. Tableau does not measure the number of data points in view or enforce line or size limitations, giving you a 360-degree view of your data.
9.	Application type	Power BI is used for custom graphics on dashboards and in many open source solutions.	While Tableau is used for long-term performance and ad hot analysis.

4.7 Comparative analysis of Power BI and Tableau

Table 2: Comparative analysis of Power BI and Tableau

	Power BI	Tableau
Data Connectivity	✓	
Dashboards and Data visualization		✓
Reporting	✓	
Data Querying		✓
Data Security	✓	
Augmented Analytics	✓	
Embedded Analytics	✓	
IoT Analytics	✓	
Geospatial Analytics		✓
Natural Language Processing	✓	
Native Mobile App	✓	
Free Trial	✓	

Data connectivity

The Power BI supports connections to many types of data sets and data sources. Power BI ensures seamless connectivity to existing data analytics investments with native integration to Microsoft technologies like Excel and strong support for many other platforms. Nearly every program that works with Microsoft can integrate with Power BI, offering an expansive library of possible collaboration options. It also can mine data from programs like Google Analytics, MySQL, Oracle, Salesforce, MailChimp, Facebook, and Zendesk. Its strength lies in connections to both on-premise and cloud-based sources, with desktop and browser-based authoring, all a hybrid integration strategy based on the Microsoft Azure cloud APIs.

Tablet users can access hundreds of data sources via the native connections of many prominent business platforms such as Amazon Redshift, Cloudera, Google Analytics, Microsoft Excel, My SQL, and more. Users can connect to data already published on their Tableau site, stored in a cloud database, private server, or stored in Excel or text-based files. In addition, developers can build their connections to currently unsupported sources, or users may ask Tableau to add support for that link in future updates.

Power BI excels in data connectivity, offering multiple data integration and support for new and more data sources.

Dashboards and Data Visualization

Microsoft Power BI hosts one of the most widely used data viewing tools in the business intelligence market. The intuitive interface for downloading and dragging software and access to the library for viewing data and 85 other data recognition applications provides an intuitive experience that keeps you coming out with excellent and informative reports. Power BI data visual performance depends on its tiles; visual metrics serve as an access point for sub-reports and databases, allowing greater visibility in numerical significance. Creating dashboards is easy, as users can pin tiles from any report and adjust the look and feel of their dashboards in the platform toolbar. Users can set alerts for specific dashboard tiles and provide embedded or favorite dashboards to monitor critical metrics. Easy sharing capabilities make it easy for users to interact, share information with all groups, and access any device via the web. Power BI uses many of the same functions found in Excel, so users familiar with Office basics will increase the value of this field.

Tableau is well known in the industry as an easy-to-use data recognition tool. Its intuitive interface allows non-technical users to quickly and easily build and customize dashboards that provide comprehensive business information comprehension. These dashboards feature responsive layouts that fit a wide range of devices and screen sizes, allowing Table-wide users to feel compact design. Users can quickly create drawings by hand drag and drop fields or automatically ask solution questions in the native language with Ask Data. Users can interact with their dashboards through filters, the Show Me tool, highlighting, and more in-depth digging.

Tableau has a library of pre-built templates called Dashboard Starters, which automatically creates dashboards based on integration with popular business data sources. Users can make their customizations or use what other Tableau users have created in an online community.

While comparing Tableau and Power BI - both solutions include dashboard creation and drag. Still, Tableau offers out-of-the-box features such as animations and pre-built templates built into Power BI only for customization or extensions. Tableau also allows users to use any number of data points for data analysis. In contrast, Power BI users are limited to 3,500 data points for data mining, making Tableau dashboards very important for interaction. This one is close, but Tableau finally comes out ahead of Power BI. Tableau provides comprehensive and easy-to-use information that makes it a clear front-end runner for data and dashboard with its actual usability, customization options, and real-time analytics.

Reporting

Power BI users across all subscription levels can create and share reports. Users can create, customize, and view their interactive, responsive, and compatible accounts across all screen sizes and objects with the drag and drop screen. Users can publish reports in the cloud or on the premises or embed them in existing applications or websites. Additionally, users can send highlighted notifications in various ways, including Excel, Word, XML, CSV, PowerPoint, MHTML, and PDF files. Power BI reports are designed to fit perfectly on a smooth, complete pixel print page. Power BI Report Server is a business-based reporting solution that allows you to manage on-premises and report back-fire distribution, while it is a cloud-based solution. Users can share their messages with others within their organization with the Share button, and managing permissions and accessing their reports is accessible from within the app.

The Tableau Reporting Tool provides automatic, scheduled, and interim reporting. Users can create reports in spreadsheets and spreadsheet formats by visualizing them in graphs, charts, or histograms. These reports can be sent to various countries in forms, including images and PDF files, which users can access for regular email updates. Words can be converted to customized dashboards, accessible through a web browser.

Power BI facilitates reporting and customization while focusing on print-ready sharing and publishing - for these reasons, Power BI emerges as the winner of reporting.

Data Querying

Power BI's Power Query Editor allows you to link to different sources with queries. Its ribbon has five organization tabs: Home, Convert, Column Add, View, and Help. Users can create and convert data with questions or view several active queries in the Power Query Editor window.

Tableau uses a query language called VizQL and also supports custom SQL queries. VizQL or Visual Query Language translates drag and drop actions into data queries. With custom SQL queries, users can create a variety of applications. They can connect to a specific database. Its custom query allows users to add, compile and compile data. Users can create merged database joins or table unions, resize or reduce data size that you can use for analysis, and more.

Tableau stands out as a winner, with its built-in data query support in its query language and SQL, as well as its NLP data analysis capabilities.

Data Security

Managed by Microsoft standards and standards for security, Power BI has strict compliance with data security standards. Power BI's commitment to data security is reflected in its very basic design: each post contains a web front-end and back-end cluster collection. Data remains secure as the solution assures clients and transfers data between two sets. Administrators can select their preferred model for user authentication. Within the resolution, Power BI provides line-level security and column-level security that can be done to prevent data access for specific users. Whenever users share a dashboard, they can choose to grant recipients permission to read or reshare; they can later view and manage individual user permissions in each dashboard. Power BI platform security includes multi-person environmental protection, network security, and the ability to add additional security measures.

Tableau supports a variety of authentication methods, including location, SAML, Kerberos, mutual SSL, Open ID, trusted authentication, and personal access tokens. It promotes good server

Practices, comprehensive, encryption and network security. Users can protect live data sources, choosing whether to instruct users to provide a database when they click on a published view. Table administrators have three options that they can use collectively to manage data security: database login accounts, verification methods, and user filters. Site roles and permissions allow administrators to work data access and authorization. Tableau also provides line-level security, limiting certain lines of data that users can see within the workbook.

While Tableau and MS Power BI share similar security features, Power BI uses powerful Azure security features and built-in architecture.

Augmented Analytics

Power BI users can use Quick Insights to identify basic understandings, combinations automatically, and output obtained with a set of mirrors made of algorithms. AI-enabled data detection automatically displays vital drivers that affect metrics provided with segment data in groups using decision trees. With the Get Insights feature, users can simply click a button inside any dashboard to discover more about specific details; Power BI can use machine learning models to explain what is happening. While these types of ML work short-term to retrieve data quickly, users may choose to continue the analysis, allowing the AI to update the data as it learns more. Data analysts can also use predictable analytic. Using AutoML, users can train, validate and request machine learning models, simplifying the development and testing of machine learning algorithms. AutoML automatically removes the most relevant features, selects the appropriate algorithm, and tunes and validates the ML model. After model training, Power BI automatically generates a performance report that includes verification results. Users can also request an ML model for any new or updated data within the data flow. With customization, Power BI supports self-view animation and Siri shortcuts such as "Hey Siri, open my report."

As far as Tableau is concerned, Smart Analytics uses machine learning to automate data correction, allowing users to segment and separate data or index and group-related words with ambiguous comparisons. Tableau also uses machine readings to recommend database tables and join based on organizational data metrics. Define data, enabled by algorithms and mathematical

Models help users to define specific points in their data with a set of focused definitions. Tableau also supports predictive analytics, prediction and drag, and integration of statistical models with R and Python.

All in all, when it comes to augmented analytics, Power BI is the winner, with AI-enabled explanations available throughout the tool and built-in machine learning algorithms.

Embedded Analytics

Power BI users can embed Power BI dashboards and interactive tiles reports within the system using REST APIs and JavaScript APIs; they may view their data directly in Power BI or an embedded application. Additionally, it supports multiple multi-user-enabled rentals with a single login confirmation. Power BI users can create a dedicated capacity to build reliable services for their clients or workplaces. Users looking for a more robust solution can purchase the Power BI Embedded. This incredible Power BI version allows additional visibility and features, including a white label, customization, and Q&A support.

Tableau provides embedded analytics in multitenant quantity, white-label, customization, and security methods such as single login, line-level security, and user permissions. Tableau offers many BI embedding options for custom web applications with the JavaScript API and Rest API. It also provides a sample code for Mobile App Bootstrap to give the first point of mobile app development with Table embedded functionality.

While Tableau and Power BI are neck and neck for embedded analytics, Power BI is the winner; by comparison, it is easy to embed in other apps and supports safe retrieval with custom viewing, while Tableau only supports this with custom development.

IoT Analytics

Power BI integrates with Azure Stream Analytics, real-time analytics, and a sophisticated event analysis engine designed to simultaneously analyze and process high amounts of fast streaming data from multiple sources. It also integrates with Azure Event Hub and Azure IoT Hub to capture data from sources such as connected devices, sensors, social media feeds, click buttons, apps and log files. Power BI can identify patterns, can initiate action, and create workflows such

Users can create real-time dashboards and alerts for these IoT devices based on temporary or local patterns and randomness by making signs, feeding information to a reporting tool, or storing modified data for later use.

Tableau can collect data from sensors or devices and analyze and visualize it to improve comprehension based on collected data. The table can create dashboards from IoT data and detect inclusions or inconsistencies.

With Azure Stream Analytics, real-time data analytics is a cinch with Power BI; its awareness capabilities, dashboard, and seamless detection move the front of the pack.

Geospatial Analytics

Power BI integrates with ArcGIS maps to provide map layouts, which users can locate in dashboards. Users can also create map views based on TopoJSON maps. If their geospatial data is in other formats, Power BI users can convert frame files or GeoJSON files into TopoJSON files which they can transform into their dashboard displays.

Table users can import geospatial data from various file formats and analyze and visualize it directly on the platform. Tableau supports a variety of interactive map viewings, such as standardized map maps, choropleth maps, point distribution maps, temperature maps, flow maps, and more. In Tableau, users can search maps to find locations and view data related to those locations. When users start typing in the search box, map searches suggest potential places in their map view based on location names and text from their data sources. Users can search for local types such as continent, state, state or province, region, city, and postal code. Tableau also offers advanced geocoding and reverse geocoding and various location functions that allow users to perform advanced location analysis and integrate location files and data into other formats such as text files or spreadsheets. Tableau location connector enables users to connect and join ESRI Shapefiles, KML files, GeoJSON, MapInfo tables, and other types of geospatial data.

Tableau leads with its geospatial data integration into many different types, interactive map detection, and powerful map search function.

Natural Language Processing

Power BI provides a Q&A option to test data by asking questions in the native language and finding answers in charts and graphs. Users can type queries with ratios, formulas, and indicators by chat, and Power BI will generate automated responses and reports. These reports may include clever narratives developed by Narrative Science Quill, an advanced tool for producing natural language to detect hidden behaviors. This account is dynamic and continuously updated as users interact with their data. Users can edit the report and share it with others. Power BI supports natural and text-based language questions and uses automatic completion to show relevant and contextual suggestions. Power BI meets Cortana; users can ask Cortana questions, a built-in AI Assistant for Windows, and get an immediate answer. The Power BI mobile app also supports native language queries.

Tableau supports NLP queries with Ask Data, which uses algorithms to allow users to type queries in the native language and visually get feedback suggestions. Users can continue to refine their queries and view data. The result can be saved for later use and shared with others. Users can also add additional data sources, search maps, and create dashboards with native language queries.

With more comprehensive features, Power BI is a master of natural language processing.

Native Mobile App

Power BI offers native mobile applications for Windows, Android, and iOS devices that allow users to securely view and share real-time reports with dashboards and view all their essential data in one place. They can tap to check, filter, and focus on the essentials. Mobile users can create and share reports, send native language queries, and set up tile changes app notifications to receive personal data alerts sent to their devices. With access to 250 offline data storage, mobile users do not need to be connected to a network to view their data. By default, the mobile app updates data frequently; users can access and interact with dashboards they have previously accessed while offline, even if reports are overloaded and some tiles that require active server connections will not be available. The mobile app notifications tab shows users a customized feed for you in a series of messages about alerts set, new shared dashboards, updates to the

Workspace and more. Power BI Mobile also supports advanced interaction through annotation allowed for contact and sharing, geospatial analysis, and QR code scanning.

Tableau has traditional apps for Android and iOS users. Provides access to dashboards and KPIs, allowing users to interact with filters, browse and zoom actions, and drill-down functions. The table enables mobile users to tag, scroll, search and browse on dashboards. Tablet offline capabilities of Tableau include interactive previews that allow offline users to access data; when users are connected to a server, the mobile app downloads all their favorite workbooks and views, which they can access later on or offline. Mobile users can subscribe to workbooks and set mobile alerts when data reaches a given limit. Tableau mobile users have a home screen designed for you that provides a selected, consistent view of their most important metrics.

Power BI is the winner in this category, prioritizing Tableau in its leading collaborations, geospatial analysis, and QR code support, as well as generational and NLP performance questions.

Free Trial

Power BI offers a free 60-day Pro trial. It also has a free regular edition for individual users, the Power BI Desktop.

Tableau offers a 14-day free trial of Tableau Server or Tableau Online, including Desktop. It also has an ongoing free program, Tableau Public.

Microsoft offers the free and most potent trial to three vendors, so we offer this Power Supply.

5 Results and Discussion

The thesis analysis indicates that it is possible to convert raw data into meaningful information based on the case study analysis. From the research, it has been observed that applications like Power BI provided promising results in creating the dashboard. Tableau has less distinct benefits as compared to Power BI.

Power BI and Tableau are both applications used for serving different purposes. However, both applications develop the same results while the dashboard is created. It is challenging to develop a dashboard for large databases using Power BI as well as Tableau. Power BI and Tableau are very affordable as anyone can easily buy and install the respective applications.

While comparing Tableau with Power BI, it has been found that Tableau does not offer the facility of data shaping. At the same time, Power BI provides Power Pivot dots that are easily accessible to the user. Power BI includes all the necessary things required for creating an effective dashboard for securing business analytics and creating reports and dashboards that goods look in terms of visualization. Tableau is performing well in terms of speed compared to Power BI as it processes a considerable amount of information very quickly despite having the same techniques and characteristics. The literature review and practical part describe the significant differences and complexities in using both applications. As per the analysis, Power BI is suitable for any user, and it can create any unique dashboard as it has many advantages and benefits than Tableau. Power BI is used to present the data to catch both viewers' eyes and attention. In contrast, Tableau creates smooth, essential, and straightforward visualizations without uniqueness in the report and dashboard.

Moreover, the most important thing is the result of the analysis because well-designed, well-defined visualization best describes the information that explains the critical dignity, which may help transform raw data into useful information for easy understanding of data. For example, instead of telling the foreign direct investment figures of different sectors to an individual or managers of the foreign companies, it is better to explain data using visualization tools that can present the information in a better way effectively and effortlessly.

6 Conclusion

Nowadays, in the world of modern technologies, all the markets are competitive. If a company does not collect and analyze data, it cannot understand its current situation, costs, and revenues. It means that a company will not see what is outdated and redirect resources to another area if needed. If one company cannot explore prospects, it means that it “sees nothing” while competitors accumulate data and draw conclusions, providing themselves with a considerable advantage. Therefore, companies need to convert the raw data into useful information using data visualization tools. So that companies can make better decisions to compete with competitors.

As mentioned at the beginning of this thesis, the main goal of it is to know various methods and tools used for data visualization. The partial objectives and main objectives were covered both in the theoretical part and in the practical part.

Data visualization is an efficient technique to convert the raw data into effective graphic representations of any business or a project that will help make decisions related to future business actions. The thesis aims to analyze widely used visualization tools in the market, validate differences, and emphasize benefits. It is necessary to take two of the most popular applications of the market, design a dashboard and create an interactive visualization using detailed data. The study has used foreign direct investment data for the previous five years, collected from the Reserve Bank of India’s annual report. The study develops a dashboard for FDI data with the help of Power BI.

The study also suggests the best suitable tool among Power BI and Tableau for international companies, senior executives of such companies, and entrepreneurs to express their ideas, projects, and company moves professionally. Tools like Power BI and Tableau are less expensive today because of their price and availability for startups and first-time users. These tools were fun to use and create dashboards because they analyze data quickly and include analysis that automatically provides valuable and important aspects of data. It is challenging to choose one and lead one over the other because Power BI and Tableau have different goals; overall, one is a spreadsheet application that can create interactive visuals, while another aims to design compelling visuals. Overall, the results showed that Power BI is unique and provides seemingly relevant and professional-looking results that are very important to businesses and companies

worldwide.

References

1. [“Geo charts”]. (2021, July 20). Thoughtspot.github.io. <https://docs.thoughtspot.com/6.0/end-user/search/about-geo-charts.html>
2. AnyChart | Choose Right Chart Type for Data Visualization. Part 2: Data Composition, Parts to Whole. (n.d.). AnyChart. <https://www.anychart.com/blog/2017/04/20/data-composition-part-whole-chart-type/>
3. AnyChart. (n.d.). *Choropleth Map | Chartopedia*. AnyChart. Retrieved August 12, 2021, from <https://www.anychart.com/chartopedia/chart-type/choropleth-map/>
4. 5. Healy, Y. H., and C. (n.d.). *Stacked Area Graph*. Wwww.data-To-Viz.com. Retrieved August 12, 2021, from <https://www.data-to-viz.com/graph/stackedarea.html#:~:text=A%20stacked%20area%20chart%20is>
5. JACKSON, Thomas W. and LOCKWOOD, Steven, 2018. Business analytics. A Contemporary Approach. 176 p. ISBN 1137610603.
6. KIRK, Andy, 2016. Data visualization. Representing Information on Modern Web. Packt Publishing. 531 p. ISBN 1787125076.
7. LOSHIN, David, 2012. Business intelligence. 2 ed. 400 p. ISBN 0123858895. MITCHELL, Michael N, 2012. A visual guide to stata graphics. College Station: A Stata Press Publication. 499 p. ISBN 9781597181068.
8. Naveen. (2019, September 30). *Introduction to Tableau*. Intellipaat Blog; Intellipaat. <https://intellipaat.com/blog/tutorial/tableau-tutorial/introduction-tableau/>
9. POWER BI, 2020. Data Visualization: Microsoft Power BI, 2020 [online]. Available at: <https://powerbi.microsoft.com/en-us/>.

10. TABLEAU, 2020. Business Intelligence and Analytics Software, 2020. [online]. Available at: <https://www.tableau.com/>.
11. ULMAN, Milos and MUSTEEN, Martina, 2020. Decision-Making in International Business: A Big Data Approach. In: AIB West 2020 Annual Conference February 27-29, 2020, San Diego State University, California, USA [online]. Academy of International Business U.S. West Chapter. 2020. Available from: <http://www.aib-w.org/2020-annual-conference/>.
12. Reserve Bank of India - Annual Report. (n.d.). M.rbi.org.in. <https://m.rbi.org.in/Scripts/AnnualReportPublications.aspx?Id=1307>
13. Restori, M. (2018). *What is a Scatter Plot and When to Use It?* Chartio; Chartio. <https://chartio.com/learn/charts/what-is-a-scatter-plot/>
14. *Trellis Chart in Tableau*. (n.d.). Www.linkedin.com. Retrieved August 12, 2021, from <https://www.linkedin.com/pulse/trellis-chart-tableau-saloni-khatri>
15. VASUDEV, C, 2006. Graph theory with applications. New Delhi: New Age International (P) Ltd., Publishers. 466 p. ISBN 812241737X, 9788122417371.
 - a. WILEY, John 2017. The big book of the dashboard: Visualizing your data using real-world business scenarios. 448 p. ISBN 1119282713, 9781119282716.
16. DATA VISUALIZATION, 2020. The Data Visualisation Catalogue, 2020. [online]. Available at: <https://www.datavizcatalogue.com/>
17. INSIGHTS, SAS and INSIGHTS, BIG, 2020, Data Visualization: What it is and why it matters, 2020. [online]. Available at: https://www.sas.com/en_us/insights/big-data/data-visualization.html/.
18. *What is Business Analytics? Definition and FAQs | OmniSci*. (n.d.). Www.omnisci.com. <https://www.omnisci.com/technical-glossary/business-analytics#:~:text=Business%20Analytics%20is%20the%20process>

8 Appendix

	A	B	C	D	E	F	G
1	Sector (US\$ million)	2015-16	2016-17	2017-18	2018-19	2019-20	Total
2	Manufacturing	8,439	11,972	7,066	7,919	8,153	43,549
3	Communication Services	2,638	5,876	8,809	5,365	6,838	29,526
4	Retail & Wholesale Trade	3,998	2,771	4,478	4,311	4,914	20,472
5	Financial Services	3,547	3,732	4,070	6,372	4,326	22,047
6	Computer Services	4,319	1,937	3,173	3,453	4,104	16,986
7	Business services	3,031	2,684	3,005	2,597	3,684	15,001
8	Restaurants and Hotels	889	430	452	749	2,546	5066
9	Transport	1,363	891	1,267	1,019	2,333	6,873
10	Construction	4,141	1,564	1,281	2,009	1,937	10,932
11	Electricity & Transmission	1,364	1,722	1,870	2,427	1,906	9,289
12	Real Estate Activities	112	105	405	213	564	1399
13	Education & Research	394	205	347	736	528	2210
14	Miscellaneous Services	1,022	1,816	835	1,226	443	5,342
15	Mining	596	141	82	247	217	1283
16	Others	215	470	226	102	137	1150
17							

Figure 38: Modified raw data (own picture)

Table 3: Characteristics of data

Sector (US \$ million)
2015-16
2016-17
2017-18
2018-19
2019-20
Total