Visualize your data with Tableau

Sowmya Sri Indukuri
Computer Science
Portland State University
Portland, Oregon, USA
indukuri@pdx.edu

Sukanya Kothapally
Computer Science
Portland State University
Portland, Oregon, USA
sukanya2@pdx.edu

ABSTRACT

In this paper, we represent the most widely used software "Tableau" and explore different ways to use it in different fields. Our goal is to provide a basic introduction to Tableau's Desktop and Tableau's Server environment along with Tableau Reader and give in detail significance of each one .Tableau is a platform for business intelligence and analytics that allows analysts and data scientists to analyze the data. It helps in simplifying raw data into the very easily understandable format. The data generated with Tableau can be understood at all levels of an organization by professionals. It can even build a customized dashboard from a non-technical user. This research report mainly helps the novice users who want to perform new analysis, to get introduced with the features of tableau, what it does with meaningless and large data.

KEYWORDS

Tableau, Tableau public, Tableau Desktop, Tableau Reader, business intelligence, dashboard, filters

ACM Reference format:

Sowmya Indukuri, and Sukanya Kothapally 2020. Visualize your data with Tableau. In *Proceedings of CS 410/510: Explorations of Data Science. Portland, OR, USA, 4 pages.*

1 Introduction

Over many years Excel has been the go-to tool for many companies for visualizing the data, but later there comes a powerful tool with more features which is named Tableau, a data visualization tool that debuted in 2005, as part of everyday operations. In terms of speed and connectivity, Tableau is much better at connecting live to external sources. As a result of the big data issues faced by the organizations, Tableau entered the industry as a sensation. Excel was a limited resource. The rows and columns of Excel are restricted unlike Tableau.

Tableau was founded by Pat Hanrahan, Christian Chabot, and Chris Stolte from Stanford University in 2003. The main idea behind its creation is to make the database industry interactive and comprehensive. Tableau has seen a considerable growth of 82% in its annual sales from $18 million in 2009 to $654 million in 2015. A report by Forbes in 2016 shows that the total income of Tableau grew 32% in the first quarter to $172 million, with foreign income up to 52%. The company closed 268 transactions greater than $100,000, up to 8% per year. Many researchers say that if Tableau continues to perform with the same growth, its net worth will be in the $3 billion counted as one of the top three BI companies in the world. In 2008, Tableau was named a Codie award winner for "Best Business Intelligence Solution" by the Software and Information Industry Association.

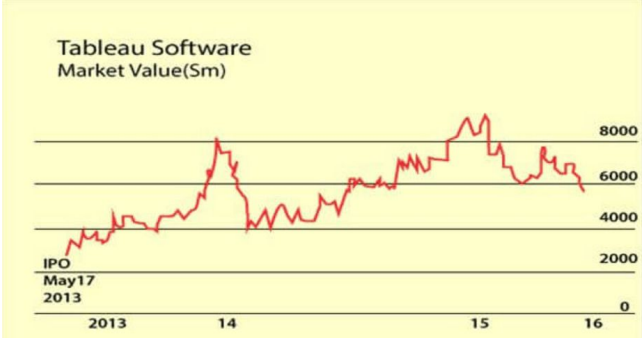

Tableau provides a wide range of products and services which includes Tableau Desktop, Tableau Online/Server, Tableau Prep Builder, Tableau Vizable (Consumer data visualization mobile app released in 2015), Tableau Public, Tableau Reader, Tableau Mobile. Among these Tableau Desktop and Tableau Public come under development tools while Tableau Server, Tableau Reader and Tableau Online come under sharing tools.

2 Tableau Desktop

Tableau Desktop is a software that needs to be installed in our machine where we can connect to a variety of data sources and start visualising data. We primarily have two versions in Tableau Desktop: Professional Edition and Personal Edition. Professional edition is more often used because of its growing layer of features. It is a paid desktop application with a 14 days trial period and developers can create charts, formulas,

dashboards, actions and everything. To publish/save our dashboards they can be stored on your Local Drive, incase of global share we can do it with the help of Tableau Server. Tableau Desktop uses an AJAX-based JavaScript to create a vision. Tableau Desktop could connect to data on-prem or in the cloud even if it's big data, SQL database, a spreadsheet, or cloud apps. It will access and combine disparate data without writing code where the power users can pivot, split, and manage metadata to optimize data sources. It's main advantage is it keeps no count on the number of rows of the data it can store, process or share. Tableau Desktop is intended for someone who has a background in business information and data management such as analysts and BI professionals.

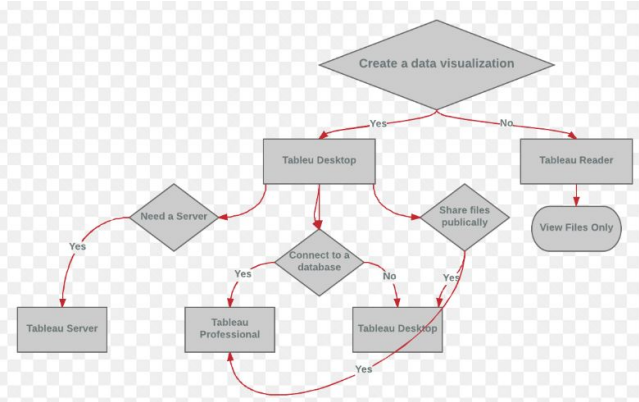

Figure 1 : Tableau products

2.1 Desktop Professional vs Personal

Professional edition can process, process and store enormous volumes of data that are responsible for all the decisions of an organization driven by data. Desktop Personal connects only to files such as CSVs or Excel documents and can save them only locally where you can't publish globally using Tableau Server/Online. Tableau Desktop Professionals appreciate the opportunity to exchange vizzes with Tableau Online or Tableau Server securely over 65 native data connectors. There is a huge subscription price difference between the two editions. Desktop Personal costs $35 (with annual maintenance $999)
Desktop Professional costs $70 (with annual maintenance $1,999)

2.2 Features

Tableau Dashboards is a graphical user interface which provides a wholesome view of your data by the means of visualizations, visual objects, text, etc. They are highly interactive as they can display data, allow multiple views and artifacts to be added, provide a wide range of designs, formats, templates and allow users to install correct filters. You can also copy a dashboard or its unique elements conveniently in one workbook. Tableau allows collaboration with other users and shares data in the form of visualizations, sheets, dashboards, etc. It facilitates secure data sharing from different data sources, such as on-premise, on-cloud and hybrid. etc. It gives the user the flexibility to use data from more than one type of data source without any restrictions which is meant by it supports connectivity to both live data sources or data extraction from external data sources as in-memory data.

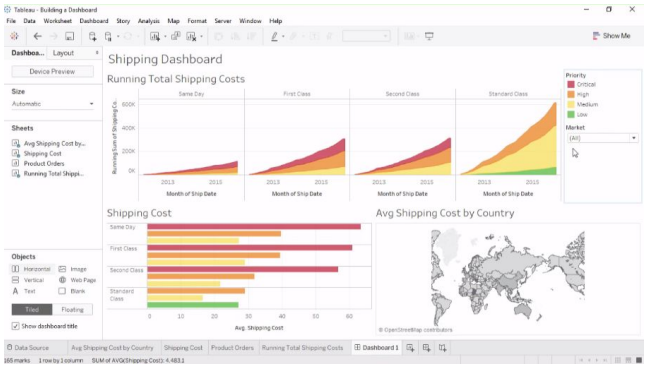

Figure 2: Building Dashboard

In Tableau we can visualize the data in a wide range of formats for which it is famous for. Different formats include Bar chart, Pie chart, Histogram, Gantt chart, Bullet chart, Motion chart, Motion chart, Treemap, etc. we can select and create any kind of visualization we need easily by selecting the visualization type from the Show Me tab. The Tableau contains a lot of information pre-installed on maps like cities, mailing codes, administrative frontiers etc. The maps produced on Tableau are therefore very comprehensive and informative. You can create informative maps in Tableau with your information by adding different levels of geology to the map according to your requirements. The types of maps provided by Tableau are Heat map, Flow map, Choropleth maps, Point distribution map, etc. With fast geocoding Tableau automatically turns the location information you already have. We can discover the world with the data and share our work in just a few clicks in the visual environment of Tableau.

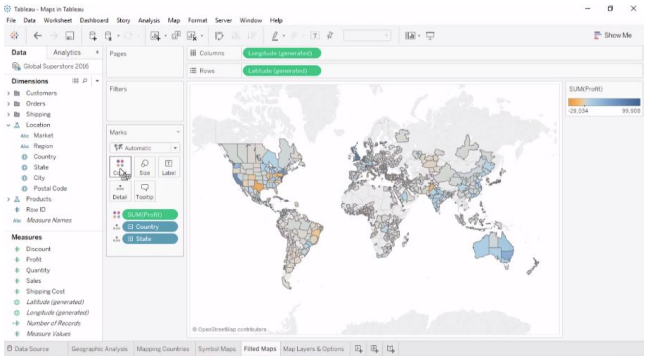

Figure 3: Geo Maps with Tableau

3 Tableau Server

Tableau server is essentially an online hosting platform to hold all your tableau workbooks, data sources and more. Tableau server is a product from Tableau, it means we can use the functionality of Tableau, without needing to always be downloading and opening workbooks to open with Tableau desktop. The main advantage of using a Tableau server is

COLLABORATION. Tableau Server is for customers that need to control where and how the data is stored and managed. The software is specifically used to share the workbooks, visualizations that are created in the Tableau Desktop application across the organization. To share dashboards in the Tableau Server, you must first publish your work in the Tableau Desktop. Once the work has been uploaded to the server, it will be accessible only to the licensed users. It can be installed on hardware behind your firewall or on cloud services that you contract with directly.

Tableau Server provides core-based licensing and the ability to easily add more servers to handle user load. Server is often more affordable than providing Desktop licenses to all your dashboard developers.Tableau Server runs well within a variety of hardware configurations. It can be deployed for small organizations on a relatively inexpensive single system. It can also be deployed for large organizations with thousands of users on clusters containing many powerful machines.

3.1 Tableau Server Architecture

Tableau Server is a robust environment that provides technology managers with the tools to secure and maintain the environment while also providing information consumers with fast access to the information they need. One of the basic characteristics of Tableau is to support your choice of data architecture. It does not need your data to be stocked in any single system, proprietary or otherwise. In general all organizations have a heterogeneous data environment, data warehouses live alongside databases and Cubes, and flat files like Excel are still very much in use.

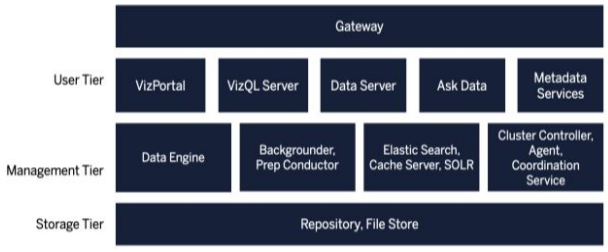

Figure 4: Tableau server architecture

The application server handles requests to the web application such as searching, browsing, logging in, generating static images and managing subscriptions. The VizQL server handles the task of loading and rendering requested views. The data engine receives queries made to Tableau Data Extracts present on the server. These queries come from the VizQL processes. To service these queries, the Data Engine loads the Tableau Data Extracts into memory and returns the requested record set. The backgrounder runs maintenance tasks and data extract refreshes. The data server handles requests to Tableau Data Sources. These requests can come from the Tableau Server or from Tableau Desktop users.

3.2 Deployment in Tableau Server

Server's architecture provides the flexibility to scale from a single box to large multi-server deployments. User-friendly tools are provided for setup and maintenance of access rights, scheduling, and notification. A new feature in Tableau Server Version 8 provides users with the additional flexibility of editing or creating new reports and analysis. This functionality isn't a replacement for Tableau Desktop, but allows staff that doesn't have a Tableau Desktop license to build and modify reports. There are three primary reasons to deploy Tableau Server Data governance, Efficiency, Flexibility.

Tableau Server is a scalable system that is capable of meeting the demands of the most intense enterprise environments. Proper planning is an important first step before you settle on the appropriate hardware configuration and licensing options. At a minimum, you should consider the following details when planning your deployment like user count,user concurrency rate, workbook complexity, user locations, database locations, database size, extract usage.

3.2.1 Deploying Tableau Server in High Availability Environments

As Tableau Server usage increases, the need to ensure its continuous availability also rises. Strategies to guarantee constant availability are broadly referred to as high availability. These strategies necessitate that core components of Tableau Server be redundant to minimize the chance of unplanned downtime. Realizing this goal requires deployment in a distributed environment and running redundant critical processes on separate servers.

3.2.1.1 The Three Node Cluster

In this configuration, one node hosts the Gateway, which routes requests to the other two Worker servers. The two Worker servers both run all of the server processes. Even though all of Tableau Server's processes should be made redundant, the three processes that must be made redundant are the gateway, the data engine process, and the repository process. The loss of a Worker machine can occur without making the cluster inaccessible. However, since there is only a single Gateway machine, should that server go offline the cluster will be inaccessible to users. To have complete fault tolerance, a four-node cluster is required.

3.2.1.2 The Four Node Cluster

In a four-node cluster, a second Gateway machine is added to make that critical node redundant. However, this standby Gateway server must be promoted to active status manually. There is currently no automatic fail-over for Gateway machines.

3.3 Performance Recorder

Prior to Tableau Version 8, this data had to be collected and analyzed manually from log files or via a third party application that was created by InterWorks. The Performance Recorder basically creates a Tableau workbook of your Tableau workbook's performance. Information about the following events

is captured and displayed visually as query execution, geocoding, connections to data sources, layout computations, extract generation, data blending, server rendering.

Performance Recorder is disabled on Tableau Server by default. To begin using it you must enable it on a per site basis. To activate Performance Recorder on the server, navigate to the Administration-Sites page and check the site you wish to enable. Performance Recording will continue capturing data about interactions with the view until the user navigates away or removes the string from the URL.

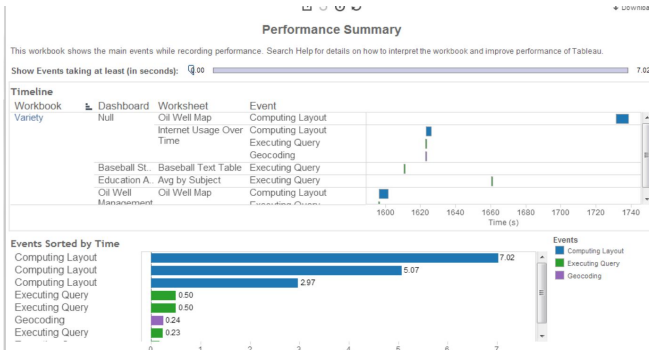

Figure 5: Performance summary workbook

4 Tableau Reader

The most secure way to distribute Tableau workbooks is via Tableau Server. But,when you want to share a workbook with someone that doesn't have access to Tableau Server, Tableau Reader provides a free alternative. Distributing content with Tableau Reader requires that you save the Tableau workbook file as a packaged workbook. Tableau Packaged Workbooks require local file sources such as excel, access, text files, tableau data extract files. Tableau Reader does not need to connect to any source, but it can read reports built on top of any and all of the listed data sources that Tableau Supports. Tableau Reader is for somebody that only needs to monitor the analytics. The CEOs & CFOs of the organizations wouldn't be making dashboards for meetings but are definitely going to be interested to drill down.

If the datasource for the workbook one wants to share with Tableau Reader comes from a server-based database then they must extract the source data first saving extracted data as a Tableau Data Extract and then saving the workbook as a Tableau Packaged Workbook.

4.1 Creating and Using Filters

There are a few different ways to add filtering to your visualization. Dragging any dimension or measure on to the filter shelf provides filtering that is accessible to the designer. Make that filter accessible to more people by turning it into a quick filter. One can also create conditional filters that operate according to the rules defined. Tableau Reader needs to be exposed to filter control on the desktop.

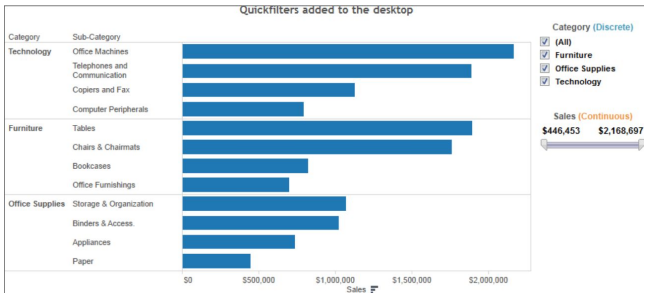

Figure 6: Adding quick filters

5 Conclusion

Tableau is a powerful and fastest growing data visualization tool used in the Business Intelligence Industry. Unparalleled capabilities of visualizing information is on top the list of Tableau software benefits. Even the products of traditional business intelligence vendors, such as Oracle Data Visualization or IBM's products for data rendition, cannot compete with the illustration and design quality that Tableau provides.It converts unstructured statistical information into comprehensive logical results, which are fully functional, interactive and appealing dashboards. The tool's intuitive manner of creating graphics and a user-friendly interface allow non-dev users to utilize the basic app's functionality to the fullest.

Apart from its high visualization functionality, users rate its overall performance as robust and reliable. The software supports establishing connections with many data sources, such as HADOOP, SAP and DB Technologies, which improves data analytics quality and enables the creation of a unified, informative dashboard. The number of Tableau fans who invest their expertise and skills in the community increases steadily. Business users can beef up their knowledge on data parsing and reporting and get many useful insights in this community. Tableau Desktop has a rich feature set and allows you to code and customize reports. Tableau server is specifically used to share the workbooks, visualizations that are created in the Tableau Desktop application across the organization. Tableau Reader is a free tool which allows you to view the workbooks and visualizations created using Tableau Desktop or Tableau Public.

ACKNOWLEDGMENTS

We would like to thank Professor Kristin Tufte, the instructor for this course, for providing advice, ideas, and resources for this project.

REFERENCES

[1]https://www.javatpoint.com/history-of-tableau#:~:text=%E2%86%92%20%E2%86%90%20prev-,History%20of%20Tableau,from%20Stanford%20University%20in%202003.

[2] https://www.guru99.com/what-is-tableau.html#5

[3] https://www.tableau.com/support/help

[4]https://intellipaat.com/blog/tutorial/tableau-tutorial/tableau-architecture/

[5]Tableau your data by Daniel G.Murray

[6]https://www.edureka.co/blog/tableau-desktop-vs-tableau-public-vs-tableau-reader/

https://data-flair.training/blogs/tableau-features/