**Data Analysis and Visualization of Educational Dataset using Tableau: A Review**

Aditya Vidyarthi

Professor, Department of Information Technology

Institute of Technology & Management, Gwalior-India

 Jitendra Singh Kushwah

Associate Professor, Department of Information Technology

Institute of Technology & Management, Gwalior-India

Vineet Shrivastava

Assistant Professor, Department of Electronics & Communication Engineering

Institute of Technology & Management, Gwalior-India

Ashish Gupta

Assistant Professor, Department of Information Technology

Institute of Technology & Management, Gwalior-India

 Saurabh Shrivastava

Assistant Professor, Department of Information Technology

Institute of Technology & Management, Gwalior-India

**Abstract:** Data visualization, which is a component of data analytics, is the process of using visual data to make sense of the meaning of data once it has been corrected. Data visualisation and analysis have been increasingly valuable in corporate intelligence and analysis across many domains in recent years. Various libraries are available for visualising datasets, which can be dynamic or interactive in nature. This study focuses on the types of data sources that are available, the analysis and cleaning of datasets, and the interactive presentation of educational datasets utilizing Tableau Tool and Pandas Libraries.

**Keywords:** *analysis, visualization, business intelligence, dataset, libraries.*

**I. Introduction**

Plots, graphs, slicers, stacked column charts, histograms, bar charts, tables, matrices, and other visual representations are examples of visualization techniques. In this paper, we concentrated on interactive data visualization [2] using the Tableau tool, which is a suite of data visualization tools that facilitates the sharing of data and the rapid acquisition of answers through interactive data visualization using various visualization charts. Data visualization is the process of understanding the clarity of data through visual representation and is a component of data analysis [1] The user or individual can more easily and conveniently comprehend the data with the use of visuals and analysis, and it has its own architecture for visually analyzing the data. The primary component of the architecture is Tableau Desktop, which is where dashboards and published reports are shared around the company utilizing data sources. Another key component is Tableau Server, which manages real-time data operations from various data sources via a browser. Additionally, it contains desktop worksheet and report repositories that the Tableau server can distribute. Tableau Public is an architecture component that Tableau Desktop offers, however you are unable to save your worksheet to your own computer. Tableau reader is the component in Architecture where the worksheet and reports can be view and analysed.

Tableau tools are essential for users to view and access dashboards and worksheets using programs like Tableau Desktop, Tableau Server, Tableau Public, and Tableau Reader.

The following are Tableau's general functions [3]:
1) Retrieve the Information from the Necessary Data Source
2) Examine the information
3) Create the worksheets and report using various visual aids.
4) Use Tableau Server to publish the worksheet and report online.
5) Make any necessary edits to the report and publish it online using Tableau.
6) Use Tableau Reader to get the report data.
7) Update the dashboard by refreshing the data through various organizational gates.

**II. Literature Review**

Studying Tableau: A Tool for Data Visualization Paul Tomolonis, Oskar R. Harmon, and Steven Batt [4] This article outlines a tutorial for learning Tableau Public, a popular, robust, and free data visualization program. An essential element of undergraduate research projects is data literacy. The exercise presented in this article covers the basic Tableau ideas and instructions required to generate charts and graphs, put them together, and describe patterns found in data. We use a Tableau file to embed the explanations and instructions. This allows the student to drag and drop items within the same file while reading instructions and explanations.

Tableau-Based Interactive Healthcare Data Visualization According to this research by Inseok Ko, MS, and Hyejung Chang [5], big data analysis is becoming more and more popular across a variety of industries, including healthcare. Visualization is crucial to the entire process of gathering, cleaning, analyzing, and disseminating data as well as to the intuitive display of data analysis outcomes. Using Tableau as a business intelligence platform, this article outlines a process for interactively visualizing and analyzing healthcare data.

Rapid Analytics and Data Visualization: Using Tableau Desktop to Aid in Library Decision-Making Murphy, Sarah Anne [6] This study examines the Librarians can better organize, examine, and display information gathered by different people across a library organization with the use of data visualization. The Ohio State University Libraries' experiences with Tableau, an advanced data visualization and quick analytics program, are covered in this article. Transaction logs, Google Analytics, and e-resource usage reports are just a few of the diverse sources of data that librarians can combine and use using Tableau. In addition to sharing examples of visualizations made for two data analysis projects, the article offers background information on integrating data visualization into the Ohio State University Libraries' evaluation program. The advantages of combining and displaying data visualizations from several sources at once are discussed and examined. A brief discussion of prospective future initiatives for Tableau Desktop data visualization for libraries is included in the article's conclusion.

Visualization of data According to Matthew N. O. Sadiku, Adebowale E. Shadare, Sarhan M. Musa, and Cajetan Akujuobi [7], it entails displaying facts in a way that is easy to comprehend through graphics or pictures. Explaining the facts and choosing a plan of action are aided by it. Any field of study that needs creative ways to communicate vast amounts of complex material will profit from it. Modern visualization has been influenced by the development of computer graphics. An overview of data visualization is given in this paper.

Seeing Library Data: A Prototype Data Visualization Application for Librarians, Mark Eaton [8] in this paper, Raw library data can benefit from visualizations. By providing librarians with visual tools to study their datasets, solutions that programmatically create interactive visualizations of this type might further enhance the value of these data. A Web application that shows the libraries' book and e-book holdings is used to illustrate these advantages. In order to produce Web visualizations, this article explains how to get data from Kingsborough's discovery layer API (Application Programming Interface) and convert it.

**III. Process Model of Tableau**

The process model, Tableau tool visuals, and interactive data visualization techniques for the analysis and design of educational data visualization using Tableau tool were covered in this study. As explained below and illustrated in the picture, the seven processes that make up Tableau's process model are GET DATA, FETCH, PROCESS, ANALYZE, VISUALIZE, EDITING, and WEB.



Figure 1: Process model of Tableau tool

a) Get Data: The end user will obtain data in this stage from many data sources, including CSV, TXT, and XLSX files.
b) Fetch: In this stage, the data is chosen by either visiting the data location or utilizing queries to choose different kinds of data sources.
c) Procedure: The Python Pandas package is used to clean and edit the data in this stage.
d) Analyze: In this stage, the data is examined using conditional queries, filters, and other procedures like conditionally adding columns and removing inaccurate data.
e) Visualize: In this phase, the data that has been processed and analyzed is shown using visual aids like boxplots, graphs, and plots. Following this procedure, the report and spreadsheet are uploaded to Tableau.
f) Editing: This stage involves finalizing the published report by fixing any mistakes, such as whether any filters or graphics need to be fixed. Once the editing process is finished, Tableau Server is used to publish the report online.
g) Web: As mentioned in the previous phase, this stage involves converting the report into a dashboard that can be shared via websites, URLs, and other means.

**IV. Details of Data Sources**

**Data Format Data Sources**

Files Excel, Text/CSV, XML, JSON, Folder, Share Point Folder

Database SQL SERVER, ACCESS, ORACLE, IBM DB2, IBM INFORMIX, IBM NETEZZA, MySQL, Postgre, Sysbase, Teradata, SAP, Google Bigquery

Azure Azure SQL database, Azure SQL Data Warehouse, Azure Blob Storage, Azure Table Storage, Azure HD Insight(HDFS)

Online Services Microsoft Exchange Online, Salesforce, Google Analytics, Facebook, Github, Kaggle

Other Web, SharePoint list, OData Feed, Active Directory, Microsoft Exchange, Hadoop File(HDFS), R Script, ODBC, OLE DB

**V. Results and Discussion**

The dashboard looks like the figure below after it has been published to Tableau. If there are any filters or visuals that are not fully accessible, they can be fixed here before being published online. In addition to the modules that were previously discussed, we may view dashboards of various modules from each educational institution. These modules may pertain to personnel, infrastructure, and other facilities offered by the university, among other things.
The dashboard results for the educational institutions' discussed courses are shown below.



Figure 2: View of Student of Campus Placement Salary Package on the basis of Gender

In the above figure 2, we represent educational data on the basis of campus placement salary package and data categories by Gender (Male, Female). This is representation of data using Pie chart which is most useful data visualization chart in Tableau tool.



Figure 3: View of Student Salary Package on the basis of Specialization

In the above figure 3, Data Visualization of student salary package on the basis of specialization area using Bar plot in Tableau tool. So that we can understand that which specialization is more in demand? This type of visualization is helping to student to opt the subject and strong the specialization area according to market demand.



Figure 4: View of Student Selected/Not Selected on the basis of Specialization

In the above figure 4, Data Visualization of student selection on the basis of specialization area using Pie Chart in Tableau tool.

**VI. Conclusion and Further Work**

In every organization and business development, data visualization is crucial since it makes it easier and more convenient for the end user to view and comprehend the data. Through the analysis and correction of data collected from various sources, we have visualized the educational institution database in this article into three modules: Placement Progression, Students Internal Marks, and a module for viewing purposes from the student's perspective. Mobile applications can be used to view dashboards built with the Tableau tool. As further effort, we can create modules on staff-related concerns like feedback, work status, and other difficulties, as well as infrastructure-related amenities like the caliber of the infrastructure offered.

**References**

[1] Srivastava, Deepmala. (2023). An Introduction to Data Visualization Tools and Techniques in Various Domains. International Journal of Computer Trends and Technology. 71. 125-130. 10.14445/22312803/IJCTT-V71I4P116.

[2] Siddiqui, Ahmad. (2021). Data Visualization: A Study of Tools and Challenges. 11. 18-23.

[3] Srungaram, Vasundhara. (2021). DATA VISUALIZATION VIEW WITH TABLEAU. 25. 10.

[4] Batt, S., Grealis, T., Harmon, O., & Tomolonis, P. (2020). Learning Tableau: A data visualization tool. The Journal of Economic Education, 51(3–4), 317–328. https://doi.org/10.1080/00220485.2020.1804503

[5] Ko I, Chang H. Interactive Visualization of Healthcare Data Using Tableau. Healthc Inform Res. 2017 Oct;23(4):349-354. doi: 10.4258/hir.2017.23.4.349. Epub 2017 Oct 31. PMID: 29181247; PMCID: PMC5688037.

[6] Murphy, S. A. (2013). Data Visualization and Rapid Analytics: Applying Tableau Desktop to Support Library Decision-Making. Journal of Web Librarianship, 7(4), 465–476. https://doi.org/10.1080/19322909.2013.825148

[7] Sadiku, Matthew & Shadare, Adebowale & Musa, Sarhan & Akujuobi, Cajetan & Perry, Roy. (2016). DATA VISUALIZATION. International Journal of Engineering Research and Advanced Technology (IJERAT). 12. 2454-6135.

[8] Eaton, M. (2016). Seeing Library Data: A Prototype Data Visualization Application for Librarians. Journal of Web Librarianship, 11(1), 69–78. https://doi.org/10.1080/19322909.2016.1239236