



# OPEN DATA VISUALIZER

**GORIJALA KIRAN**

School of computer science and engineering  
Lovely professional University Phagwara, Punjab, India

**Guided by: ATUL MALHOTRA**  
Assistant professor

School of Computer Science and Engineering Lovely Professional University Phagwara,  
Punjab, India

**Abstract-** Data visualization is the process of presenting the raw data in graphical form which is easy to understand and interpret. Data visualization tools provide us multiple ways of understanding outliers and patterns. Visualization is the key tool for the developers who were working on trillion rows of data which is being generated every day. It helps in any department of study which requires innovative ways of representing the data and also helps in enhancing the new skill and facilitates better task performance. So for understanding the raw data we provide you a web application that helps us in converting the data into a graphical way.

**Index Terms-**visualization,plots,application,feasible

## I. INTRODUCTION

There have been lots of uses by representing the massive amount of data in graphical ways that are easily understandable and accessible. In this fast-growing world, trillions of data have been generated every day, as a result, the amount of data on the web is increasing day by day so by using visualization tools we can have the proper meaning of data. Data visualization tools like charts, graphs, maps make data more interactive and understandable. In this Big Data-driven world,

data perception enables you or decision-makers of any business or industry to look at analytical reports and understand concepts that may be difficult to understand. The ability to visualize data is essential to scientific research. Today, systems are being used to process the bulk data. Data recognition is concerned with the design, development, and use of computer-generated graphical data. It Provides an effective representation of data from various sources.

This allows decision-makers to visualize analytics more easily and makes it easier for them to make sense of the data. It helps them to find patterns, to understand details, and then form an opinion.

By this time, you would have understood how data recognition simplifies the way data is presented.

However, is that the only power of data viewing? Not really. As the world changes, so do the need for knowledge. Here are a few benefits of looking at data:  
1.comprehensible information - Data is increasing day by day, and it is not wise for anyone to manipulate large amounts of data to understand it. Data recognition is helpful at that time.

2.Establish relationships - Charts and graphs not only show data but also establish relationships between different types of information.

3data visualization speeds up the decision making skills because data visualization provides more comfortable to the brain to visualize than the raw data.

## II. DATA VISUALIZATION TECHNIQUES

Visualization of large data can be done using some of the techniques these techniques can be used to understand the data more precisely. some of the techniques were : 1.pie chart

2.tables 3.charts 4.graphs 5.Dashboard 6.correlation 7.bar chart 8.cartogram

9.dot distribution map

Data Visualization uses computer graphics to show patterns, styles, and relationships between data objects. It can generate pie charts, bar charts, distribution sites, and other types of data graphs with simple drop-down and mouse click menus. Colors are carefully chosen for certain types of visualization for representing the data, we must choose colors that work to distinguish between data objects.

In the data visualization, the data is extracted and summarized. Location variations such as size and shape represent key elements in the data. The detection system should perform data reduction, modify and visualize the actual data on the screen. Without a visual representation of data, it's hard for people to grasp the true meaning of the data.

## III. PROPOSED MECHANISM AND TECHNIQUE

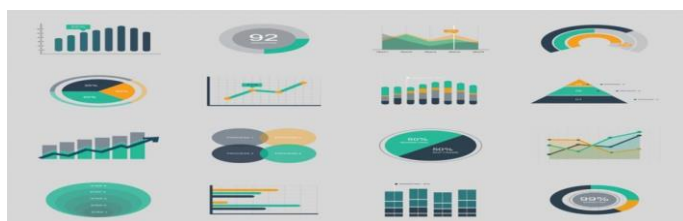
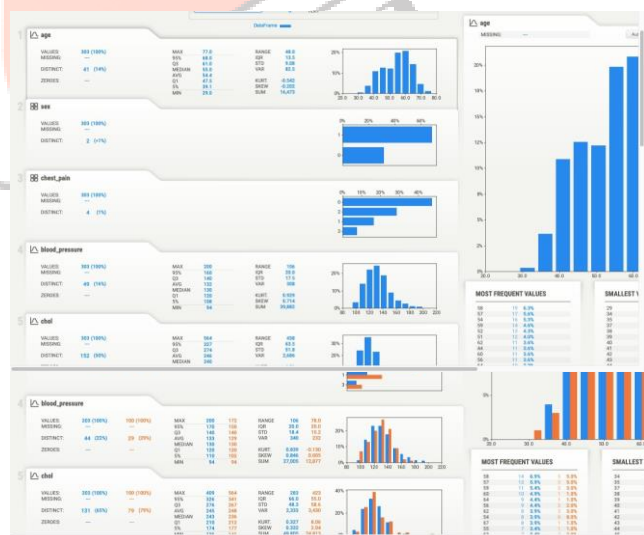
In this project, we used a web application for converting the data in tables to visual form which helps in understanding the data deeper. So using this application users can upload the datasets and they can get a visual representation of the data in the form of images and users can upload any size of data in the web applications and they can get their desired graphs in seconds.

In this application, data is going to be represented in the histogram, violin plots, heatmap of a continuous variable, Scatter plot of variables, and so on.

Techniques used for data visualization:

### SWEETVIZ:

Exploratory data analysis is the process where we analyze the complete dataset and gathering the main characteristics of the complete data by using different methods.EDA is very important because it helps us in understanding the big raw data. but EDA actually takes a lot of time in processing the data but using sweetviz makes our work easier in visualization.



## AUTOVIZ:

Visualization is a method used to visualize data using various graphs and episodes. In data science, we often use data viewing techniques to understand data and find relationships between data. Views can help find a pattern in the database used for further analysis.

There are different python strategies/libraries used for Data Identification such as Matplotlib, Seaborn, Plotly, etc. But while using all of these libraries we need to define the type of graph we want to visualize and the issues we need to visualize.



## DJANGO:

Django is a state-of-the-art Python web framework that enables the rapid development of secure websites. Built by experienced engineers, Django takes great care of the hassle of web development, so you can focus on writing your app without the need to refresh the wheel. It is free and open-source, with a thriving and active community, great documentation.

Characteristics of good data:-

**Validity** - the degree to which our data conforms the constraints.

**Accuracy**- comparing the actual values with the predicted

values

**Completeness**- the degree to which complete data is known.

**Consistency**- we should ensure that our data is consistent in our datasets.

**Uniformity** - the degree to which the data is specified using the same unit of measure.

## IV.APPLICATIONS:

We need data recognition because a visual summary of information makes it easier to identify patterns and trends than to look at thousands of lines in a spreadsheet. It's the only way how human brain works. As the purpose of data analysis is to gain understanding, data is very important when viewed. Although a data analyst can draw comprehension from data without being seen, it will be very difficult to communicate the meaning without seeing it. Charts and graphs make data acquisition easier even if you can identify patterns without them.

In undergraduate schools, students are often taught the importance of presenting data acquisition visually.

Without visual recognition of the details, it can be difficult for the audience to understand the true meaning of the findings. For example, holding a number for your supervisor will not tell them why they should care about the data, but will show them a graph of how much money the details can save/make sure will be considered. Data recognition is an important branch of computer science and has a wide range of applications. Many tools have been developed specifically for the system analysis of individual datasets for most

fields of medicine and science. Data visualization uses:-

**Changes over time:** this is the most basic one for data visualization. the reason is every data in a dataset has an element of time involved. So most of the data analysis the common one is the time.

**Determining frequency:** it is also often the most basic use of data visualization because it works and in data that involves time. When time is involved, it makes sense that you should determine how often the appropriate events occur over time.

Correlations: Correlation is the most important use of data display. It is very difficult to determine the relationship between two variables without realizing it, but it is important that you know the relationship in the data. This is a good example of the importance of data visualization in data analysis.

Analyzing value and risk: Determining values and risk for a dataset needs multiple variables to be involved and it is almost impossible to determine the accurate values for a dataset. Data visualization helps us solve this problem. visual representation of data in graphs makes it easier to find the values and the risks.

## V. CHALLENGES:

Data visualization is changing our society considerably.it also makes work easier for big data users. large,time-varying datasets pose a great problem for data visualization because of the large data volumes. Data visualization can make the user more active by responding to the issues that arise.

The image-generating method is used for the process of testing data interactions that vary from time to time. It visualizes temporary events by mimicking the design of storytelling techniques.

Visual reality will have a profound effect on data visibility, allowing people to interact with data in the third grade for the first time. Consider being able to select a set of data and move it to any axis to compare it with another, not too far away. According to SAS, we can only process 1 kilobit of information per second on a flat screen, which can be greatly increased when analyzed in the world of 3D VR.

Big data structured and unstructured data gives us a different challenge in the development of data visualization. This is due to the size and speed of visualizing the dataset. It is difficult and time-consuming to build a large simulation data set. It is also difficult to determine which visuals would be better to use.

## VII.CONCLUSION:

Data recognition is the process of representing data in a graphical or visual way in a clear and effective way. It has emerged as a powerful and comprehensive tool for analyzing and interpreting big and complex data. It has become a fast, easy way to convey ideas in a universal format.

It must communicate complex ideas clearly, accurately, and efficiently. These benefits have allowed data perception to be useful in many fields of study.

## REFERENCES

- [1] J. L. V. Sancho, J. C. Dominguez, and B. E. M. Ochoa, "An approach to the taxonomy of data visualization," *Revista Latina de Communication Social*, vol. 69, 2014, pp. 486-507.
- [2] "Data visualization techniques," SAS, [http://www.sas.com/en\\_us/offers/sem/data-visualizationtechniques2332568.html?keyword=data+visualization+techniques&matchtype=p&publisher=google&gclid=COycgoCbutACFcolgQodqwgliA](http://www.sas.com/en_us/offers/sem/data-visualizationtechniques2332568.html?keyword=data+visualization+techniques&matchtype=p&publisher=google&gclid=COycgoCbutACFcolgQodqwgliA)
- [3] J. Wolfe, "Teaching students to focus on the data in data visualization," *Journal of Business and Technical Communication*, vol. 29, no. 3, 2015, pp. 344-359. [4] T. Kalimba, G. Nimako, and K. Herbst, "Data everywhere: an integrated longitudinal data visualization platform for health and demographic surveillance sites," *Proceedings of the 6th ACM Conference on Bioinformatics, Computational Biology and Health Informatics*, Sept. 2015, pp. 551,552.
- [5] O. Kumar and A. Goyal, "Visualization: a novel approach for big data analytics," *Proceedings of the Second International Conference on Computational Intelligence & Communication Technology*, 2016, pp. 121-124.
- [6] S. Grainger, F. Mao, and W. Buytaert, "Environmental data visualization for non-scientific contexts: Literature review and design framework," *Environmental Modelling and Software*, vol. 85, 2016, pp. 299-318.

- [7] W. N. Dilla, R. L. Raschke, "Data visualization for fraud detection: Practice implications and a call for future research," *International Journal of Accounting Information Systems*, vol. 16, 2015, pp. 1-22.
- [8] S. A. Murphy, "Data visualization and rapid analytics: applying tableau desktop to support library decision-making," *Journal of Web Librarianship*, vol. 7, no. 4, 2013, pp. 465-476.
- [9] T. J. Brigham, "Feast for the eyes: an introduction to data visualization," *Medical Reference Services Quarterly*, vol. 35, no. 2, 2016, pp. 215-223.
- [10] C. Chen, "Information visualization," *WIREs Computational Statistics*, vol. 2, July/August 2010, pp. 387-403.
- [11] R. R. Laher, "Thoth: software for data visualization and statistics," *Astronomy and Computing*, vol. 17, 2016, pp. 177-185.
- [12] L. Yu et al., "Automatic animation for time-varying data visualization," *Pacific Graphics*, vol. 29, no. 7, 2010, pp. 2271 -2280. [13] X. Li et al., "Advanced aggregate computation for large data visualization," *Proceedings of IEEE Symposium on Large Data Analysis and Visualization*, 2015, pp. 137,138.

