CONTENTS

ABSTRACT

Data visualization is a general term that describes any effort to help people understand the significance of data by placing it in a visual context. Patterns, trends and correlations that might go undetected in text-based data can be exposed and recognized easier with data visualization software. Data visualization is thepresentation of quantitative information in a graphical form. Data visualizations turn large and small data-sets into visuals that are easier for the human brain to understand and process.

Data visualizations are surprisingly common in our everyday life, but they often appear in the form of well-known charts and graphs. Good data visualizations are created when communication, data science, and design collide. Data visualizations done right offer key insights into complicated data-sets in ways that are meaningful and intuitive. In this we would like to discuss about data visualization, importance of data visualization, data visualization tools. The second is designed to display the solution in 3D to aid spatial recognition and data navigation to enhance the presentation of performance data obtained as output from simulation programs has remained almost unexplored. Data visualization is concerned with the design, development, and application of computer-generated graphical representation of the data

CHAPTER 01

THEORY OF PROJECT

1.1 INTRODUCTION:

In the mid-1980s the area of visualization separates from computer graphics to discern graphics added from scientific and engineering data. A foster partition took place in the early1990sto discern scientific, or physically established, data from intellectual ―information visualization.The leading publishing locales for research in visualization are the IEEE Visualization. Visual theatrical performance can help figuring out problem and find by rendering a structure for showing and conveying intending of extremely abstract data. Speaking about graphics, we ought to cue graphical attributes and entities. While visualizing, usually we only get the complying graphical attributes and entities to choose from (while not fixed to). Entity: line (curve), glyph, solid, surface, point, image, polyline, text Attribute: strength/color, style, size, location, and corresponding position/motion. In reality ―data‖ have approximately particular type. To produce visual images visualization, utilize computer graphics which support in the translating of composite, frequently theatrical performance of data. Data Visualization: Visualization-founded data find answer thatprovide extremely interactive and graphical user interfaces, are constructed on in-memory architectures. By a broader range of employees than traditional business analysis tools these answer generally enable users to search data without practically training and making them reachable. Data visualization usual techniques are Line Charts, Bubble Charts, Bar Charts, Scatterplots, Pie Charts, Images and maps

1.2 Overview:

1.2.1 What is Data Visualization?

Data visualization is but by definition it means "much **more"**. Data visualization is a graphical representation of any data or information. Visual elements such as charts, graphs, and maps are the few data visualization tools that provide the viewers with an easy and accessible way of understanding the represented information.

Data visualization enables you or decision- makers of any enterprise industry to look into analytical reports and understand concepts thatmight otherwise be difficult to grasp. When you think of data visualization, your first thought probably immediately goes to simple bar graphs or pie charts. While these may be an integral part of visualizing data and a common baseline for many data graphics, the right visualization must be paired with the right set of information. Simple graphs are only the tip of the iceberg. There's a whole selection of visualization methods to present data in effective and interesting ways.

Data visualization is one of the steps of the data science process which states that after data has been collected, processed and modeled, it must be visualized for conclusions to be made. Data visualization is also an element of the broader data presentation architecture (DPA) discipline which aims to identify locate manipulate format and deliver data in the most efficient way possible.

AI-driven visual analytics and dashboards can be used to quickly and easily generate and deliver high-value insights. These dashboards enable everyone in your organization to use insights for better decision-making with smart data discovery, data wrangling, geospatial analytics, and one-click predictive analytics.

Embedded analytics is the ability to deliver insights in the context that business applications, employees, and customers already use is critical. To accomplish this, the system must be able to integrate interactive reports and visualizations seamlessly into your product or service.

The system should modernize your enterprise reporting platform and cost-effectively deliver pixel-perfect reports. By generating reports that connect your data from multiple sources, you can deliver critical insights throughout the organization.

Empower everyone to transform data into insights quickly and easily. Data visualization allows easy discovery of trends and faster identification of outliers. The information helps build an understanding of how the business is performing and what opportunities and risks are developing.

Organizations can make fast data-driven decisions with visualizations that display real-time data in valuable, informative ways. By comprehending information and collaborating with others to spark insights and spot patterns in data faster, the days of relying on gut instinct to make quick decisions is long gone. Analyzing reports helps business stakeholders focus on the areas that require attention. The visual mediums help analysts understand the key points needed for their business. Whether it is a sales report or a marketing strategy, a visual representation of data helps companies increase their profits through better analysis and better business decision.

1.2.2 Why Is Data Visualization Important?

Data visualization provides a quick and effective way to communicate information in a universal manner using visual information. The practice can also help businesses identify which factors affect customer behavior pinpoint areas that need to be improved or need more attention; make data more memorable for stakeholders understand when and where to place specific products and predict sales volumes.

Other benefits of data visualization include the following:

1. The ability to absorb information quickly, improve insights and make faster decisions.
2. An increased understanding of the next steps that must be taken to improve the organizations.
3. An improved ability to maintain the audience's interest with informationthey can understand.
4. An easy distribution of information that increases the opportunity to shareinsights with everyone involved.
5. Eliminate the need for data scientists since data is more accessible and understandable.
6. An increased ability to act on findings quickly and, therefore, achieve successwith greater speed and less mistakes.

1.2.3 Why do we do data visualization?

To see and understand pictures is one of the natural instincts of human, and to understand numerical data is a year's training skill from schools, and even so, a lot of people are still not good with numerical data. From a well-drawn picture, one is much easier to find the trends and relations. Because visual presentation of information takes advantage of the vast, and often underutilized, capacity of the human eye to detect information from pictures and illustrations. Data visualization shifts the load from numerical reasoning to visual reasoning. Getting information from pictures is far more time-saving than looking through text and numbers – that's why many decision makers would rather have information presented to them in graphical form, as opposed to a written or textual form. Another thing we should mention is that: data visualization is not scientific visualization.

Scientific visualization uses animation, simulation, and sophisticated computer graphics to create visual models of structures and processed that cannot otherwise be seen, or seen in sufficient detail. While data visualization is a way that present and display information in a way that encourages appropriate interpretation, selection, and association. It utilizes human skills for pattern recognition and trend analysis, and exploits the ability of people to extract a great deal of informationin a short period of time from visuals presented in a standardized format.

Visualization is the graphical presentation of information, with the goal of providing the viewer with a qualitative understanding of the information contents. It is also the process of transforming objects, concepts, and numbers into a form that is visible to the human eyes. When we say "information", we may refer to data, processes, relations, or concepts. Here, we restrict it to data.

Data visualization is essential to assist businesses in quickly identifying data trends, which would otherwise be a hassle. The pictorial representation of data sets allows analysts to visualize concepts and new patterns. With the increasing surge in data every day, making sense of the quintillion bytes of data is impossible without Data Proliferation, which includes data visualization.

CHAPTER-2

LITERATURE SURVEY

To understand data by diagrams and maps visualization used in China as early as 1137. In all fields there has been vast Development in visualization techniques. To examine information and data visualization help to visualize and express ideas in architecture. With the coming of computer simulation visualization pertinence has been foster strengthened. A wide assortment of computer-based tool in constructing design into CAD (Computer Aided Design) design is provided by T. Honget al. (2000). The use of CAD for seeing design has been adopted exceedingly quickly by Professional throughout the world. Information visualization used to present design data with the aid of drawings and diagrams and data is usually conceptual or special, we require scientific visualization techniques like charts and graph.

There is a conterminous motive to further inquire the subsisting and new methods of visualization that efficaciously present multidimensional data. The visualization scheme should gain from manual methods where potential to aid designers make a conversion from their practice. Visualization should have power to present multidimensional data and it must be synergistic and permit efficacious communication. Using color coding and layering site analysis data is presented on the drawing with the power of controlling the visibility of layers as craved by the designers.

The domain of visualization grows, the instrument is striving users originate in our research laboratories. In dictate to confront accomplishable manifest of assessable gain that will promote more far-flung acceptance of visualization in which the serviceability ascertained experiments and studies reports are useful merely there is an arising want as substitute method of rating. Information visualization commonly part of some originative action that needs user to construct hypotheses, searches patterns and exclusion, and the polish their hypothesis. User frequently require to view the similar data from dissimilar linear perspective and above a year. They might require a kind of instruments to attain their aims, insistently importing and exportation of data.

The visualization community has latterly found dissever above the measures and effect of inordinate graph decoration and notation. Visualization experts such as Stephen Few and Edward Tufte encouraged the conventional view, contains that the visualization should present the data clearly without any perturb and should not include chart junk [8, 9]. Psychology lab studied has also been supported this view, which present that simple and clear visualization are easy to interpret. Memorability experiment result shows that visualization is as such memorable with consistency over people. Visualization is less unforgettable than instinctive scenes but like to pictures of faces, which might clue at general nonfigurative, characteristics of human retention. Not astoundingly, ascribes such as comprehension and color of a human recognizable aim increase memory power. Creating a visualization unforgettable intends creating the visualization ―stick‖ in the spectator minds. We require the most significant applicable facets of data the writer is attempting to transmitto stick. The expedition of big data band is significant but hard trouble.

Visualization of information technic might assist to figure out the trouble. Expedition of visual data has lots of applications like data mining and fraud detection utilize visualization of information technique for mended data analysis. Expedition of visual data commonly permits a quicker data expedition frequently supplies more beneficial result, particularly in instances where reflex algorithms flunks. In the determinations of the exploration visual data expedition technics furnish an often-higher level of assurance. This information extends to an eminent requirement for visual expedition technics and constructs them essential in colligation with automatic expedition technics. Information visualization concenters on data sets deficient underlying Two-Dimensional (2D) or Three-Dimensional (3D) substance and therefore as well as deficient a measure representation of the nonfigurative data onto the forcible screen. For visualization there a lot of long- familiar methods for data sets are x-y plots, histograms, and line plots. These methods are utile for data expedition but are restricted to comparatively minor and small dimensional data bands.

Data visualization instantly gain from development in technology that extend innovational ways of showing complex data. Data visualization cover geographical information system, graphical user interface, digital images, graphs, multidimensional tables, virtual reality and three-dimensional Data visualization instantly gain from development in technology that extend innovational ways of showing complex data.

Visualization technique construct vast and complex information understandable. Information visualization is an ocular interface that allow for insight of information to the exploiter. To constructthings easy to translate and construe and easy to use. In order to accomplishing the visual representation all serviceability emerge are significant to think about. The access of contrivingsubstantially visualization the procedure can be classified into dissimilar steps such as selection,presentation, mapping, usability, evaluation and interactivity, which describes main actions regarding visualization to direct exact and perfect design. Visualization techniques are classified differently,there is three classes of visualization i.e., scientific visualization, information visualization and software visualization. There are many formal data visualization techniques such as table, pie chart,bar chart, histograms, bubble chart, area chart and line chart.

To start the basic of visualization for the next generation, keeping in mind the meaning of visualization. The business sector even generally dismisses misinterpret, utilized in ineffective data andinformation visualization, which extends to the overall less output. To transmit the determination of vast or very vast data set that is not easy to interpret under custom approaches animated and interactive visualization is more advantages.

Visualization can be used for decision making and data analysis. People interaction with visualization tool has strongly affect on the understanding of data and system functions. Therefore, human interaction contribute significantly role in the valuation and design of visualization tool.

Visualization help human to display data as a graphics. And this can also call cognitive support. In a table there are different mechanisms summarized which explore the advantages of perception of human such as increased resources, reduced search, enhanced recognition, perceptual monitoring andmanipulate medium. The first one is a review on the recognized methodology of human factor especially focused on visualization research. Second, described the review on the current human factor in visualization research that provide base for the future research. In the third one describes the area in which future research is possible [15]. While there are number of frameworks exist of information visualization, different visualization techniques and data analysis strategies, but still datavisualization is process.

CHAPTER 3

SYSTEM ANALYSIS

3.1 PROBLEM STATEMENT:

Create a Data visualization system using open-source services and methods, where users can upload a file and expect different charts or graphs based on the data.

Misleading Color Contrast: In data visualization, high degrees of color contrast may cause viewersto believe that value disparities are greater than they really are

Too much Data: When data visualizations include too much data, information overwhelms, and data melts into a graphic soup that most viewers can't stomach.

Omitting Baselines and Truncating Scale: Data varies, sometimes widely, like when measuring levels or voting habits according to geographical regions. In an effort to make visualizations more dramatic or aesthetically pleasing, designers may choose to manipulate scale values on graphs. Common examples are omitting the baselines or starting Y-axis somewhere above zero to make data differences more pronounced. Another example is truncating the x value of a data series tomake it seem comparable to lower-value series.

Designer Takeaways:

1. Color is more than a way to differentiate between data series.
2. High-contrast color pairings cause viewers to perceive greater degree of data disparity. 3.Information overload applies to data visualization. If too much is present at once, viewers zone out.4.It can be more effective to communicate data with multiple data visualizations.
5. Aesthetic appeal is subordinate to accurate data representation.
6. Omitting baselines and truncating scales to intentionally exaggerate or minimize data disparities isunethical.

3.2 Implementation:

3.2.1 Aapproach:

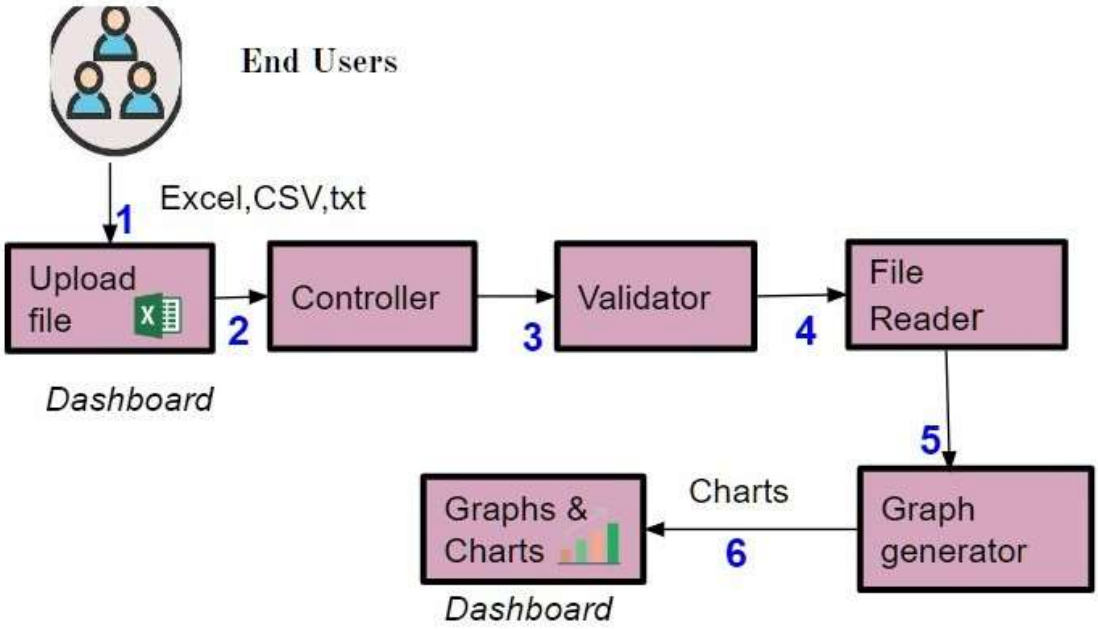


Fig. 3.2.1 Work Flow

3.2.2 Initial Flow:

- The user will provide a file in the dashboard.
- Triggers the controller.
- Controller will send the file to the validator.
- Validator checks the file and sends the request to the file reader.
- File Reader reads the data from the excel file and sends the data to the graph generator forgenerating the graph.
- By using that data, the Graph generator creates charts. (It may be pie chart, bar graph, linegraph). The chart will be shown in the dashboard.

3.2.3 Flow Chart:

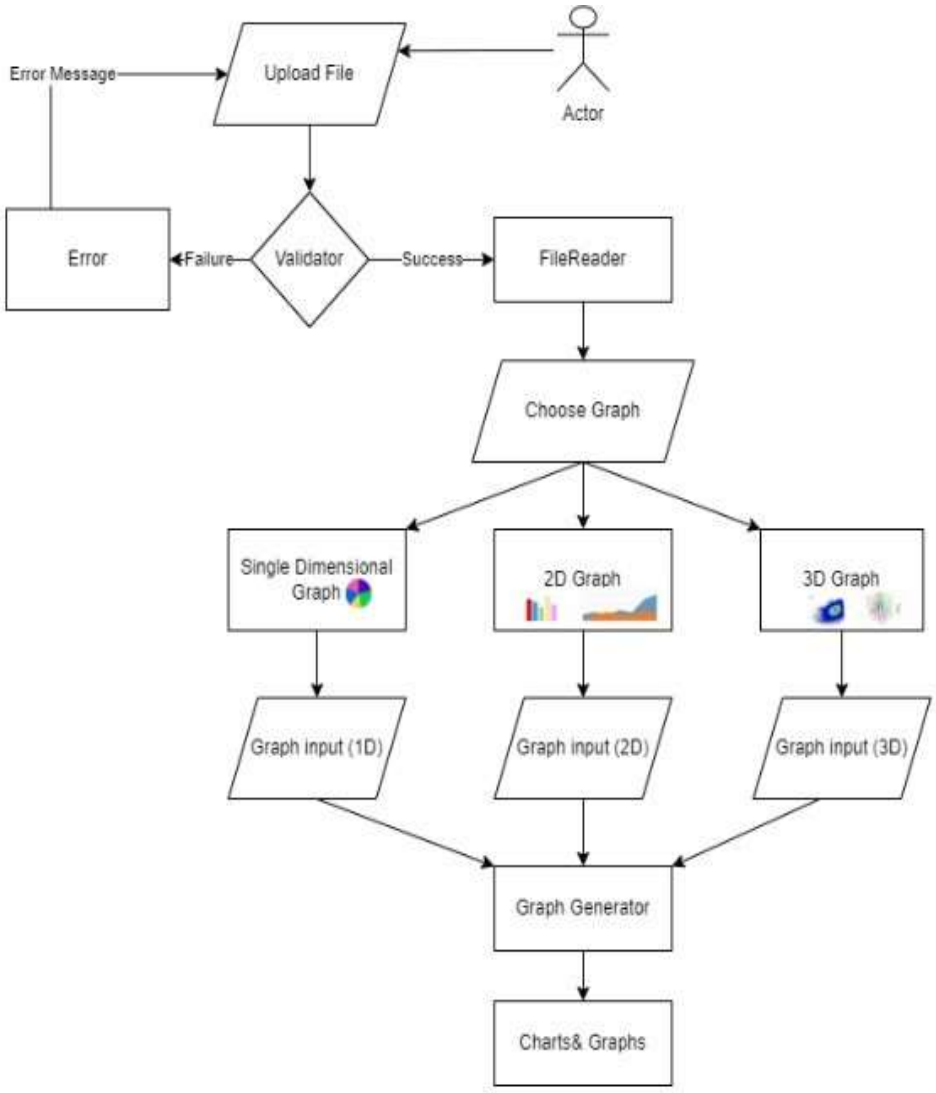


Figure 3.2.3 Detail Flow

3.3 Exceptions:

3.3.1 Differ in Type Format:

If user uploads dataset file other than csv, excel and txt than UI will throw error as file is notin correct format please Upload csv, excel or txt Format files.

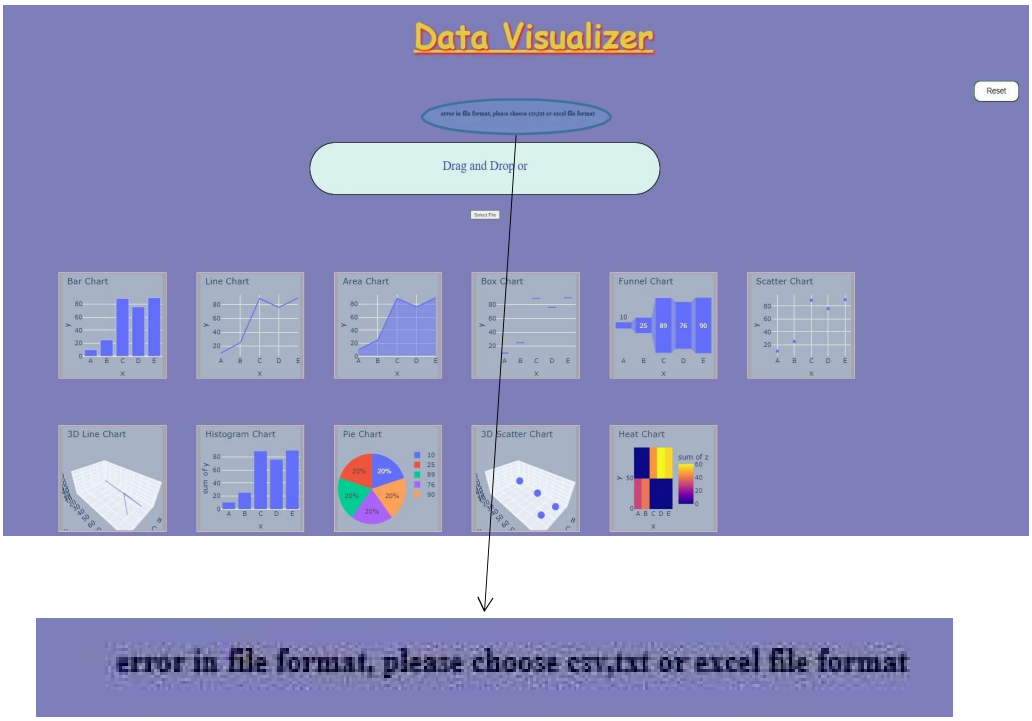


Fig. 3.3.1 Error in file format

3.3.2 File Not Uploaded:

If User Inputs Chart or Graph Before Uploading Dataset Than UI will Throw Error as “file isnot uploaded”.

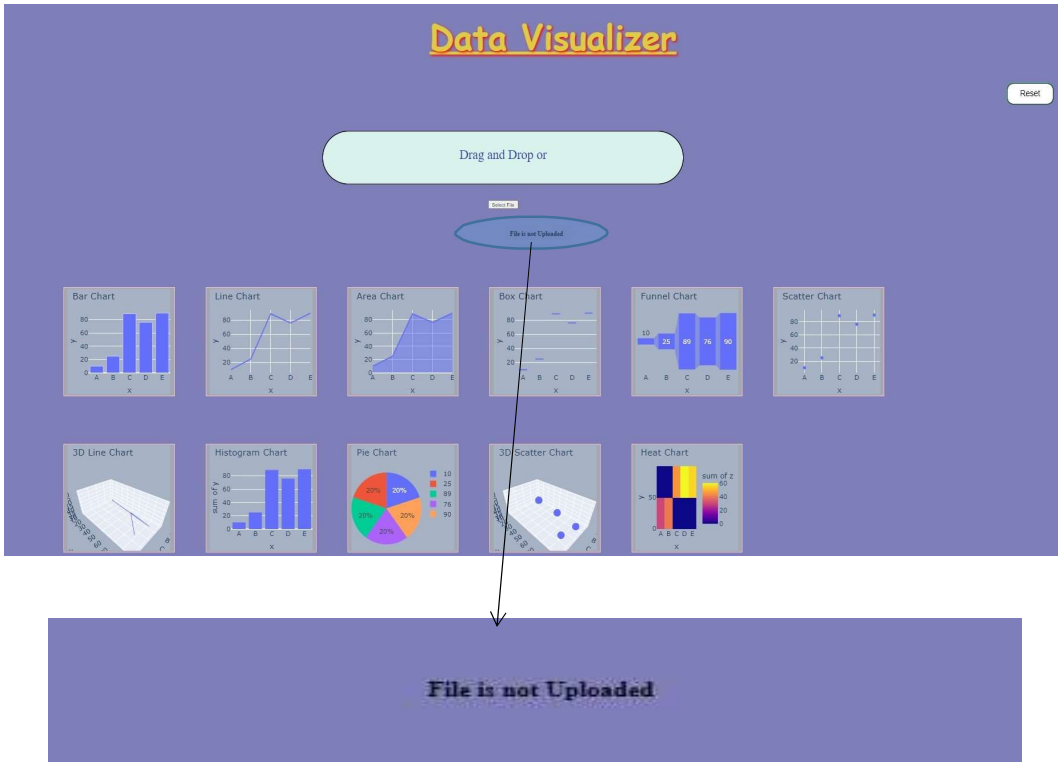


Fig. 3.3.1 File is not uploaded

3.4 Types of charts or graphs Used:

3.4.1 Bar Chart:

A bar chart or bar graph is a chart or graph that presents categorical data with rectangular bars with heights or lengths proportional to the values that they represent. The bars can be plotted vertically or horizontally. A vertical bar chart is sometimes called a column chart.

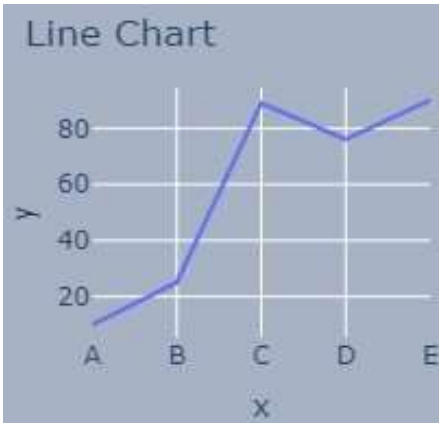
A bar graph shows comparisons among discrete categories. One axis of the chart shows the specific categories being compared, and the other axis represents a measured value. Some bar graphs present bars clustered in groups of more than one, showing the values of more than one measured variable.

3.4.2 Funnel Chart:

Funnel charts are a type of chart, often used to represent stages in a sales process and show the amount of potential revenue for each stage. This type of chart can also be useful in identifying potential problem areas in an organization's sales processes. A funnel chart is similar to a stacked percent bar chart.

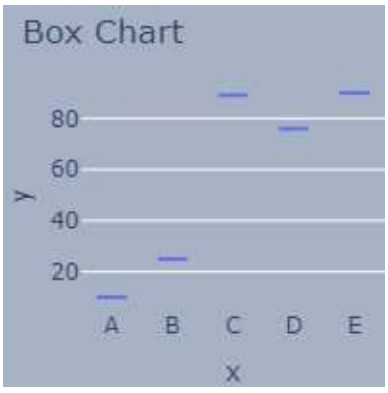
3.4.3 Line Chart:

line chart or line plot or line graph or curve chart is a type of chart which displays information as a series of data points called 'markers' connected by straight line segments. It is a basic type of chart common in many fields. It is similar to a scatter plot except that the measurement points are ordered (typically by their x-axis value) and joined with straight line segments. A line chart is often used to visualize a trend in data over intervals of time – a time series – thus the line is often drawn chronologically. In these cases, they are known as run charts.



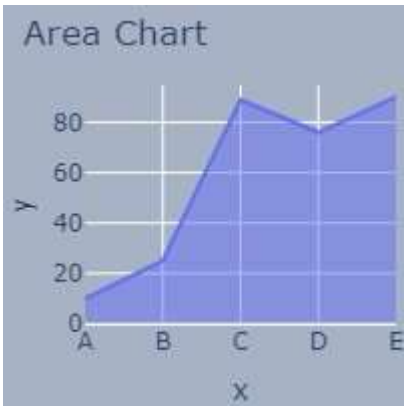
3.4.4 Box Chart:

A box plot or boxplot is a method for graphically demonstrating the locality, spread and skewness groups of numerical data through their quartiles. In addition to the box on a box plot, there can be lines (which are called *whiskers*) extending from the box indicating variability outside the upper and lower quartiles, thus, the plot is also termed as the box-and-whisker plot and the box-and-whisker



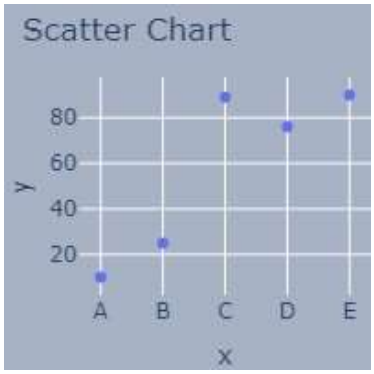
3.4.5 Area Chart:

An area chart or area graph displays graphically quantitative data. It is based on the line chart. The area between axis and line are commonly emphasized with colors, textures and hatchings. Commonly one compares two or more quantities with an area chart.



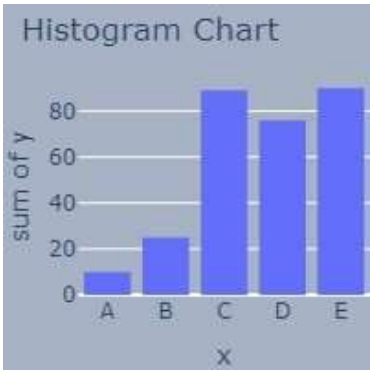
3.4.6 Scatter Chart:

A scatter plot (also called a scatterplot, scatter graph, scatter chart, scattergram, or scatter diagram) is a type of plot or mathematical diagram using Cartesian coordinates to display values for typically two variables for a set of data. If the points are coded (color/shape/size), one additional variable can be displayed. The data are displayed as a collection of points, each having the value of one variable determining the position on the horizontal axis and the value of the other variable determining the position on the vertical axis.



3.4.7 Histogram Chart:

A histogram is an approximate representation of the distribution of numerical data. The term was first introduced by Karl Pearson. To construct a histogram, the first step is to "bin" (or "bucket") the range of values—that is, divide the entire range of values into a series of intervals—and then count how many values fall into each interval. The bins are usually specified as consecutive, non-overlapping intervals of a variable. The bins (intervals) must be adjacent and are often (but not required to be) of equal size.

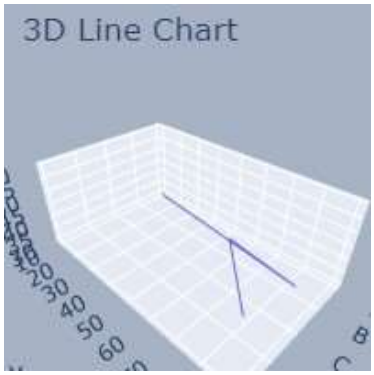


3.4.8 Pie Chart:

A pie chart (or a circle chart) is a circular statistical graphic, which is divided into slices to illustrate numerical proportion. In a pie chart, the arc length of each slice (and consequently its central angle and area) is proportional to the quantity it represents. While it is named for its resemblance to a pie which has been sliced, there are variations on the way it can be presented.

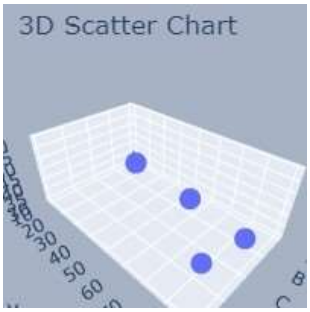
3.4.9 3D Line Chart:

A line chart or line plot or line graph or curve chart [1] is a type of chart which displays information as a series of data points called 'markers' connected by straight line segments.[2] It is a basic type of chart common in many fields. It is similar to a scatter plot except that the measurement points are ordered (typically by their x-axis value) and joined with straight line segments. A line chart is often used to visualize a trend in data over intervals of time – a time series – thus the line is often drawn chronologically. In these cases, they are known as run charts.



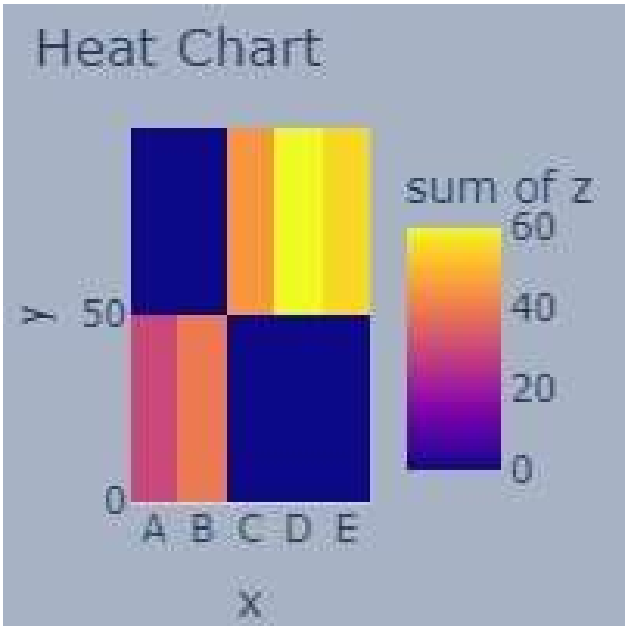
3.4.10 3D Scatter Chart:

A scatter plot (also called a scatterplot, scatter graph, scatter chart, scattergram, or scatter diagram) is a type of plot or mathematical diagram using Cartesian coordinates to display values for typically two variables for a set of data. If the points are coded (color/shape/size), one additional variable can be displayed. The data are displayed as a collection of points, each having the value of one variable determining the position on the horizontal axis and the value of the other variable determining the position on the vertical axis.



3.4.11 Heat Map:

A heat map (or heatmap) is a data visualization technique that shows magnitude of a phenomenon as color in two dimensions. The variation in color may be by hue or intensity, giving obvious visual cues to the reader about how the phenomenon is clustered or varies over space. There are two fundamentally different categories of heat maps: the cluster heat map and the spatial heat map. In a cluster heat map, magnitudes are laid out into a matrix of fixed cell size whose rows and columns are discrete phenomena and categories, and the sorting of rows and columns is intentional and somewhat arbitrary, with the goal of suggesting clusters or portraying them as discovered via statistical analysis. The size of the cell is arbitrary but large enough to be clearly visible. By contrast, the position of a magnitude in a spatial heat map is forced by the location of the magnitude in that space, and there is no notion of cells; the phenomenon is considered to vary continuously.



CHAPTER 4

SYSTEM REQUIREMENT

4.1 HARDWARE REQUIREMENTS:

1. Processor: Intel P4 3.0 GHz
2. RAM : 2 GB
3. Disk : 1GB

4.2 SOFTWARE REQUIREMENTS:

1. Operating System: Windows, Linux
2. Browser: Google Chrome
3. Coding Language: Python
4. Libraries: Plotly, Dash, Pandas
5. Software Development Kit: PyCharm Community
6. Version Control: Git & GitHub

CHAPTER 5

BRIEF OVERVIEW OF SOFTWARE TOOLS:

5.1 Python:

What is Python? Executive Summary

Python is an interpreted, object-oriented, high-level programming language with dynamicsemantics. Its high-level built-in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can be freely distributed.

Often, programmers fall in love with Python because of the increased productivity it provides. Since there is no compilation step, the edit-test-debug cycle is incredibly fast. Debugging Python programs is easy: a bug or bad input will never cause a segmentation fault. Instead, when the interpreter discovers an error, it raises an exception. When the program doesn't catch the exception, the interpreter prints a stack trace. A source level debugger allows inspectionof local and global variables, evaluation of arbitrary expressions, setting breakpoints, stepping through the code a line at a time, and so on. The debugger is written in Python itself, testifying to Python's introspective power. On the other hand, often the quickest way to debug a program is toadd a few print statements to the source: the fast edit-test-debug cycle makes this simple approach very effective.

Python is a widely used general-purpose, high level programming language. It was created by Guido van Rossum in 1991 and further developed by the Python Software Foundation.It was designed with an emphasis on code readability, and its syntax allows programmers to express their concepts in fewer lines of code.

Python is a programming language that lets you work quickly and integrate systems moreefficiently.

Below are some facts about Python Programming Language:

1.Python is currently the most widely used multi-purpose, high-level programming language.

2.Python allows programming in Object-Oriented and Procedural paradigms.

3.Python programs generally are smaller than other programming languages like Java. Programmers have to type relatively less and indentation requirement of the language, makes them readable all the time.

4.Python language is being used by almost all tech-giant companies like – Google, Amazon,Facebook, Instagram, Dropbox, Uber… etc.

The biggest strength of Python is huge collection of standard libraries which can be usedfor the following

5.2 Libraries:

5.2.1 PLOTLY:

Polly is a Montreal based technical computing company involved in development of data analytics and visualization tools such as Dash and Chart Studio. It has also developed open-source graphing Application Programming Interface (API) libraries for Python, R, MATLAB, Java script and other computer programming languages. Library is an open-source library that can be used for data visualization and understanding data simply and easily. Polly supports various types of plots like line charts, scatter plots, histograms, cox plots, etc.

There are three main modules in Polly. They are:

plotly. plotly -- acts as the interface between the local machine and Plotly. It contains functionsthat require a response from Plotly's server.

plotly. graph objects -- module contains the objects Figure, layout, data, and the definition of theplots like scatter plot, line chart that are responsible for creating the plots.

plotly. graph objects. Figure -- and these are serialized as JSON before it gets passed to plotly.js.

Plotly's Python graphing library makes interactive, publication-quality graphs. Examples of how to make line plots, scatter plots, area charts, bar charts, error bars, box plots, histograms, heatmaps, subplots, multiple-axes, polar charts, and bubble charts.

Plotly.py is free and open source.

5.2.2 Dash:

Dash is an open-source Python framework used for building analytical web applications. It is a powerful library that simplifies the development of data-driven applications. It's especially useful for Python data scientists who aren't very familiar with web development. Built on top of Plotly.js, React, and Flask, Dash ties modern UI elements like dropdowns, sliders and graphs directly to your analytical python code. Dash apps consist of a Flask server that communicates with front-end Reactcomponents using JSON packets over HTTP requests. Dash applications are written purely in python, so NO HTML or JavaScript is necessary. Dash is a python framework created by plotly for creating interactive web applications. Dash is written on the top of Flask, Plotly.js and React.js. WithDash, you don't have to learn HTML, CSS and JavaScript in order to create interactive dashboards, you only need python. Dash is open source and the application build using this framework are viewedon the web browser.

Dash applications are made up of 2 building blocks:

1. Layouts
2. Call Backs

Layout describes the look and feel of the app, it defines the elements such as graphs, dropdowns etc. and the placement, size, color etc. of these elements. Dash contains Dash HTML components using which we can create and style HTML content such as headings, paragraph, images etc. using python. Elements such as graphs, dropdowns, sliders are created using Dash Core components.

Callbacks are used to bring interactivity to the dash applications. These are the functions usingwhich for example we can define the activity that would happen on clicking a button or a dropdown.

5.2.3 Pandas:

Pandas is an open-source library that is made mainly for working with relational or labeled data both easily and intuitively. It provides various data structures and operations for manipulating numerical data and time series. This library is built on top of the NumPy library. Pandas is fast andit has high performance & productivity for users. Pandas were initially developed by Wes McKinney in 2008 while he was working at AQR Capital Management. He convinced the AQR to allow him to open source the Pandas. Another AQR employee, Chang She, joined as the second major contributor to the library in 2012.

Pandas generally provide two data structures for manipulating data, they are:

1. Series
2. Data Frame

Series is a one-dimensional labeled array capable of holding data of any type (integer, string, float, python objects, etc.). The axis labels are collectively called indexes. Pandas Series is nothing but a column in an excel sheet. Labels need not be unique but must be a hash able type.

Pandas Data Frame is a two-dimensional size-mutable, potentially heterogeneous tabular data structure with labeled axes (rows and columns). A Data frame is a two-dimensional data structure, i.e., data is aligned in a tabular fashion in rows and columns. Pandas Data Frame consists of three principal components, the data, rows, and columns.

Pandas are generally used for data science. This is because pandas are used in conjunction withother libraries that are used for data science. It is built on the top of the **NumPy** library whichmeans that a lot of structures of NumPy are used or replicated in Pandas.

5.3 Software Development Kit:

PyCharm is a dedicated Python Integrated Development Environment (IDE) providing a widerange of essential tools for Python developers, tightly integrated to create a convenient environment for productive Python, web, and data science development.

5.4 Version Control:

At a high level, GitHub is a website and cloud-based service that helps developers store and manage their code, as well as track and control changes to their code. To understand exactly what GitHub is, you need to know two connected principles:

- Version control
- Git

CHAPTER 6

6.1 User Interface:

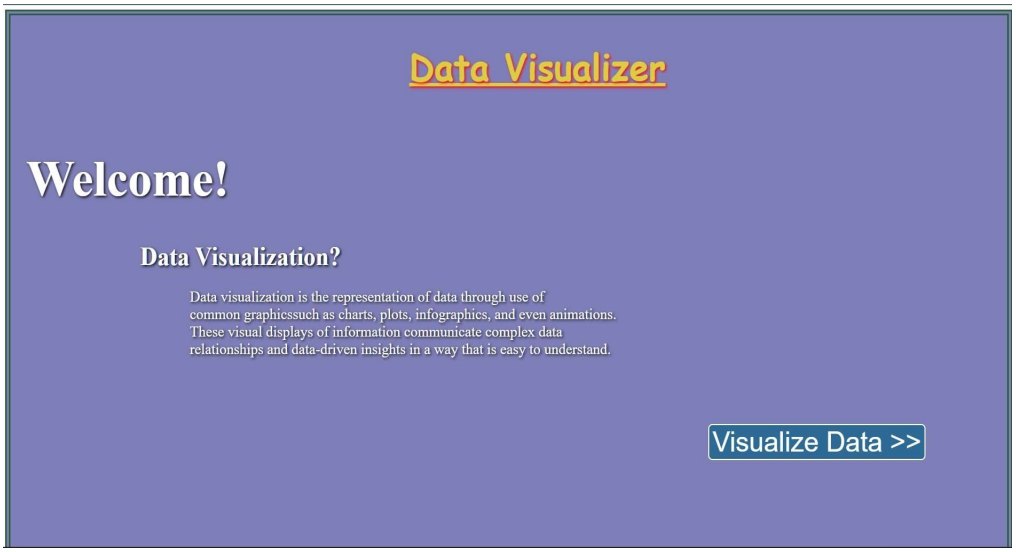


Fig. 6.1(a) Home Page

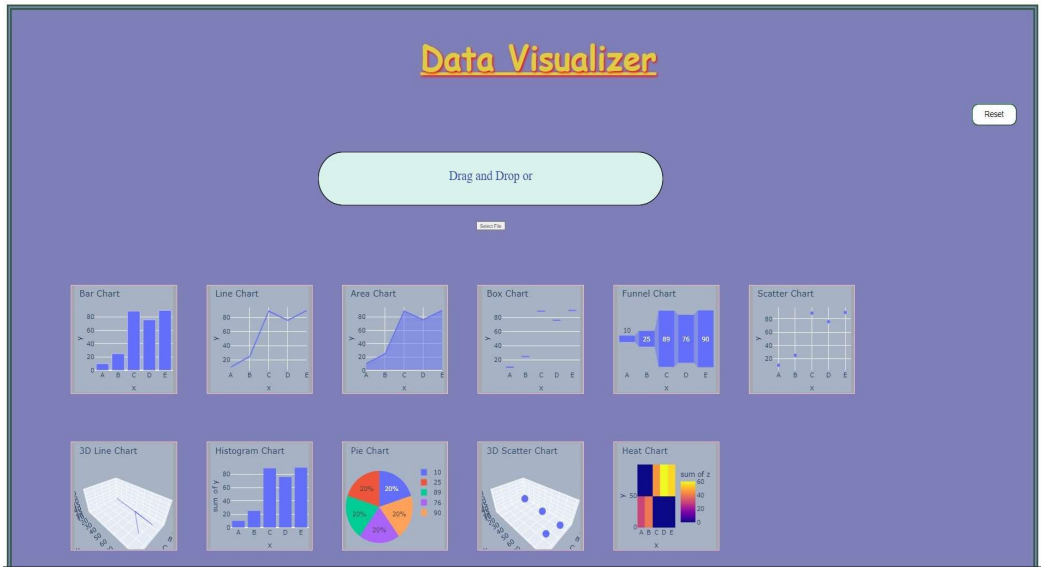


Fig.6.1(b) Main Page

6.2 Testing:

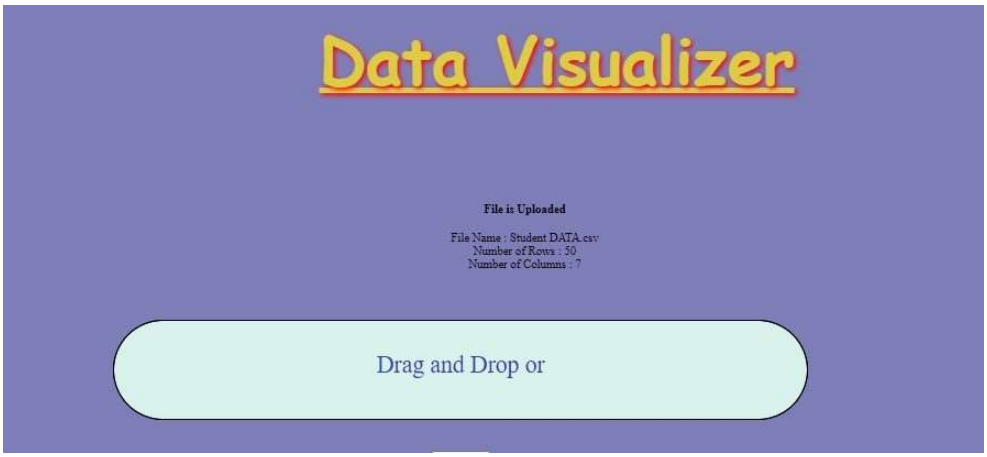


Fig.6.2(a)

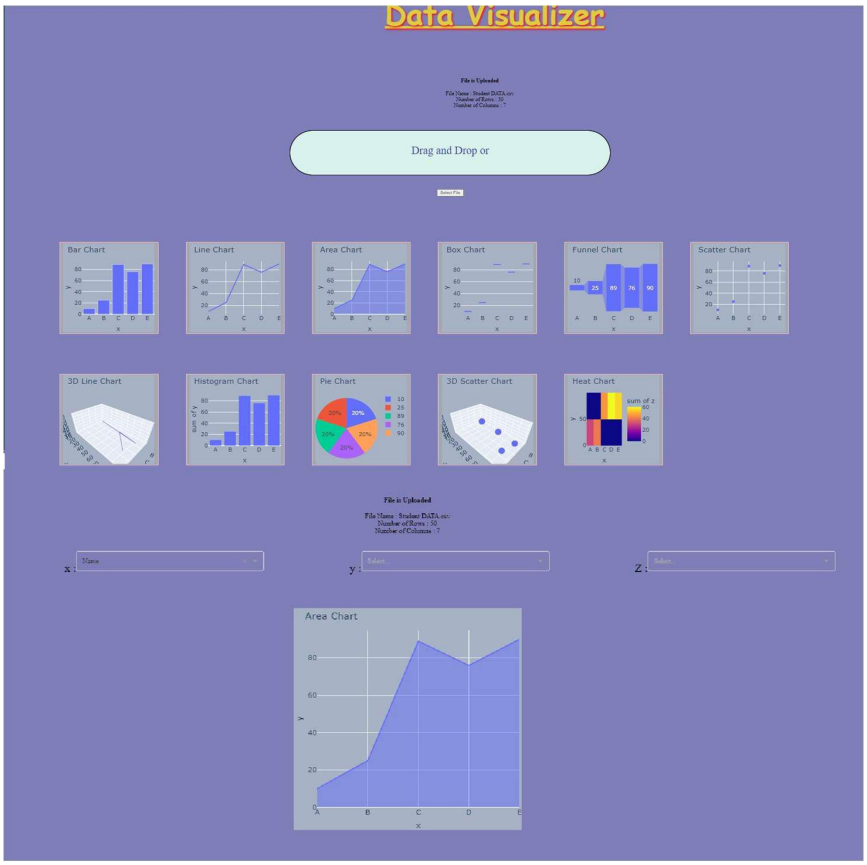


Fig.6.2(b)

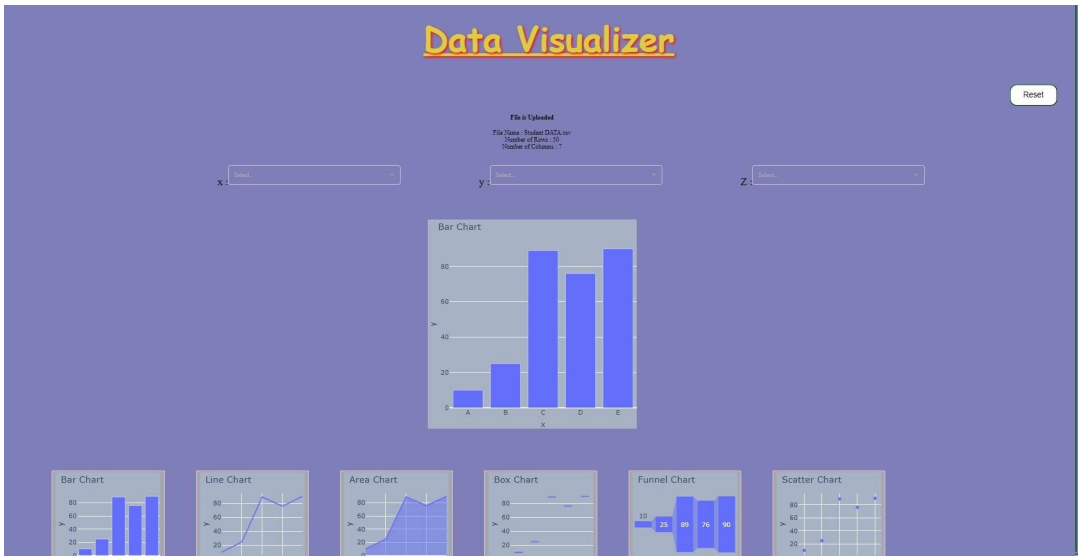


Fig.6.2(c)

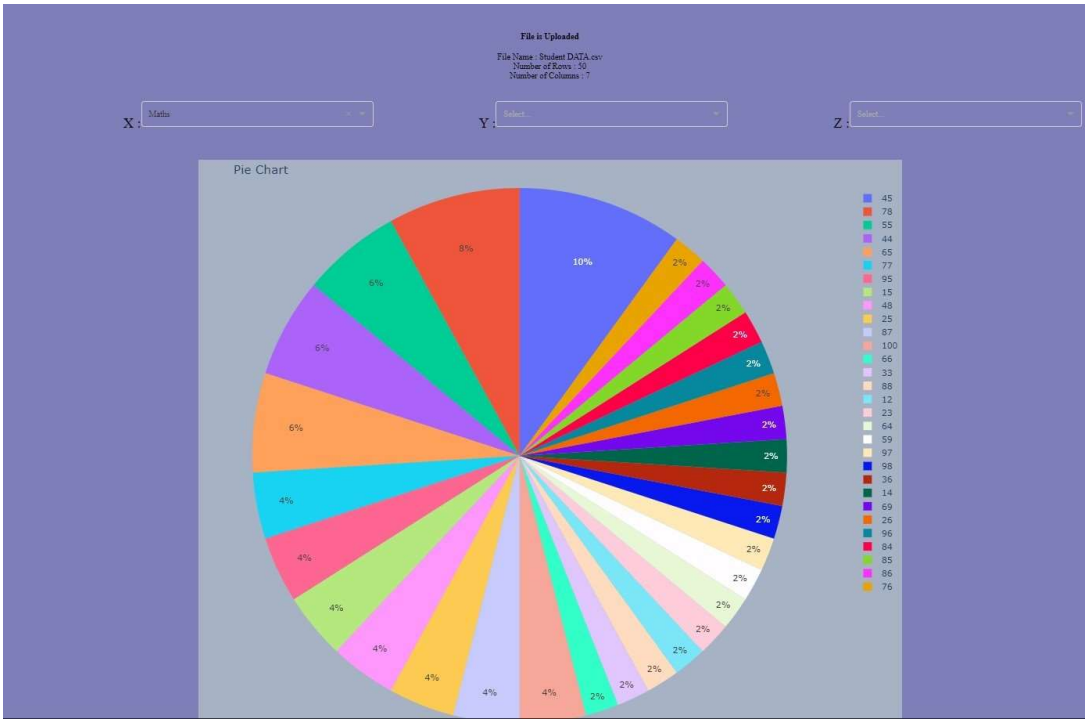


Fig.6.2(d)

Fig.6.2(e)

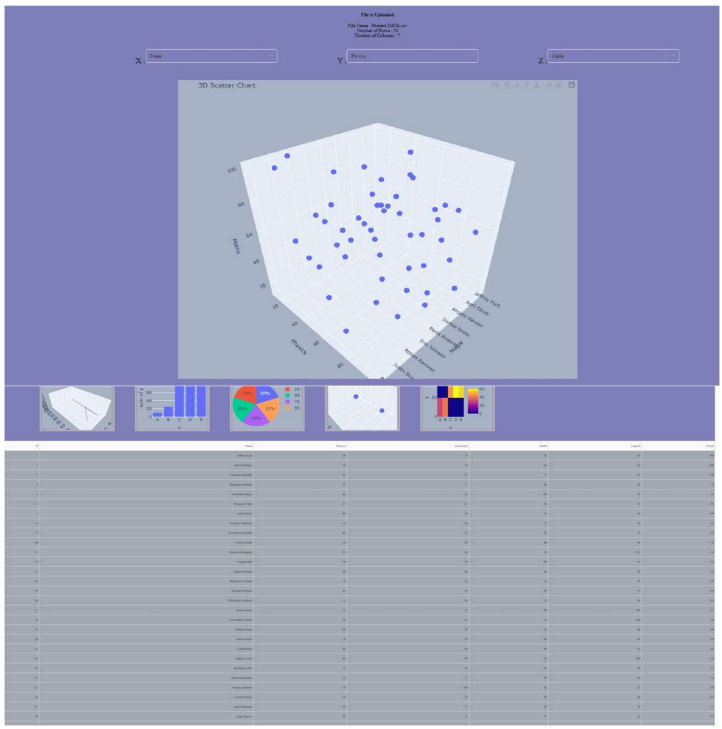


Fig.6.2(f)

CONCLUSION

Data visualization is displaying data in the graphics form so that it can be easily understandable. Different data visualization techniques are used to display data in visual form. This paper defines the literature review on data visualization techniques. This research was made by filled questionnaire. This research study provides best knowledge to the beginners who are willing to work on data visualization techniques. This paper provides best understanding about the concept of data visualization techniques. This paper define which visualization techniques is best. The result of research study is that new visualization technique should be an image form.

REFERENCES

[1] Maureen Stone, Information Visualization: Challenge for the Humanities

[2] Matthew Ward, Overview of Data Visualization‖, fromwww.cs.wpi.edu

[3] Jerzy Stefanowski, Instytut Informatic, Data Visualization or Graphical Data Presentation‖.

[4] Paul Kent, data visualization: ―Making Big Data Approachable and Valuable‖.

[5] Shweta Srivastav, Simon Lannon, Donald k. Alexander, and Phil jones, ―A Review and Comparison of Data Visualization Techniques Used in Building Design and in Building Simulation‖, Eleventh international Ipas conference Glasgow, Scotland, 2009

[6] Plaisant, C. ―The Challenge of Information Visualization Evaluation‖, Proc. of Conf. on Advanced Visual Interfaces, ACM, New York (2004), 109-116.

[7] Ben Shneiderman, Catherine Plaisant, Strategies for Evaluating Information Visualization Tools: Multi- dimensional In-depth Long-term Case Studies, Proceedings of the BELIV '06 workshop Advanced Visual Interfaces Conference 2006, Venice

[8] S. Few. Benefitting infovis with visual difficulties? Provocation without a cause. ―Visual Business Intelligence‖ Newsletter, 2011.

[9] S. Few. The chartjunk debate: ―A close examination of recent findings. Visual Business Intelligence Newsletter, 2011.