

Data Visualization: A Study of Tools and Challenges

Ahmad Tasnim Siddiqui

Associate Professor - Visiting at BGI, Faizabad
Assistant Professor - Department of Computer Science & Applications
Sherwood College of Professional Management
Lucknow, India
tasnim5@yahoo.com

Abstract— Over the year's research, particularly scientific research and publication is getting more advanced and accurate. Publishing is also getting more professional and clear. Now a days everyone is talking about data visualization. Data visualization is very encouraging area. Almost every company is using data visualization tools and techniques to display beautiful and attractive data in the pictorial format. Same is the case with educational institutions. We can use visual features like graphs, maps, and charts. Data visualization helps to elucidate facts, figures and decide sequences of actions. Any field of study will be benefited which requires advanced techniques of presenting large and complex information. There are number of good data visualization tools available which provide an accessible technique to see, understand and explain trends and patterns in data. We are currently in the world of Big Data, where data visualization tools and technologies are of great importance to analyze huge amounts of data and make data-driven results.

Keywords – Data visualization, information visualization, data, tools, graphics, reports

I. INTRODUCTION

We live in a society of information technology and there has been the requirement of presenting enormous amounts of data in such a manner that can be is comprehensive and easily available. Visualization is pictorial representation of data which aims clear understanding to the user. Each and every organizations are producing large amount of data every day. Due to the reason, the amount of data storage on the web has increased intensely. It is very challenging for users to explore, visualize, and use this huge amount of data. Data visualization is directly related with the design implementation, development, and applications of computer made pictorial representation of the data. This permits management people to see the analysis in visualized format and makes it easy to understand the data. It also supports them to understand patterns, information, and make an opinion [3].

These days data visualization contributes a clear idea of what the data and information means by giving it visual look through chart, maps or graphs. This makes the data more understandable and natural for our mind. It makes it easier to find out the patterns and trends within huge data sets. Showing information in beautiful manner is an art. Our eyes are made to sees and identify colorful and beautiful patterns around us. Unfortunately, across the world many facts and figures are presented incorrectly or partially correct and they use data visualization practices [1]. The new data visualization techniques should come with better methods to analyze, process and visualize the data. In the recent times, technology has provided advancement to create better visual presentations in terms of quick and complex creation of data visualization.

Over the last couple of decades research and publication is getting more advanced, accurate and clear. Specially publishing is getting more professional and clear [8]. We are lucky enough that we have different type of data visualization tools available which are being used to create a beautiful and effective visuals such as charts, graphs, maps, dashboards and infographics etc. We can use Microsoft Excel as a data visualization tool. People are not aware that MS-Excel has enormous capability to visualize the data. It has properties like PivotTables,

Charts, slicers and more. But for more advanced visualization there are tools like Power-BI, Tableau, Infogram, ChartBlocks, Google Charts, Datawrapper, Plotly, SAP Analytics Cloud, D3.js, and NVD3 etc. These are the tools which can help you to get a clean and interactive visualization for all the data. They can produce an effective visual within a minute but we need to choose the correct tool according to our data [2].

II. DATA VISUALIZATION TECHNIQUES

Data visualization is achieved by computer supported tools to represent the visuals. Visualized data supports interactive data visualization which allows users to choose the format of their choice to display data. Commonly used type of data visualizations are [6]:

- Line graph: This type of technique displays relationships between items.
- Bar chart: It is used to compare the quantities
- Pie chart: Is used to compare portion of whole chart
- Tables: Represents normal and interactive table
- Maps: It displays maps such as heat map
- Gantt chart
- Bullet Graph
- Histogram
- Timeline etc.

So, the layout of charts and graphs can take the form of line graph, bar chart, scattered plot, pie chart etc. Data is abstracted and summarized in data visualization process. A visualization system should follow the steps like preparation, cleaning, organization and execution to get the dataset. It can visualize the result in the form of graphs and charts and display the output in a user friendly manner.

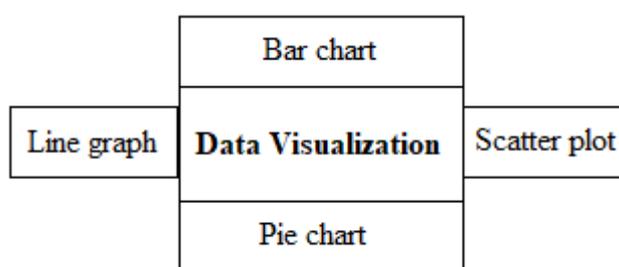


Fig. 1: Common techniques of Visualization

Most of the visuals are commonly used for decision making. It can be used in producing research related data, health data, energy consumptions, educational budgetary analysis, fraud detection representation and many more [3].

These tools include ManyEyes owned by IBM, SmartMoney for stock market, Insights owned by Facebook Corporation, and Thoth developed by California Institute of Technology, Power-BI developed by Microsoft, Google charts by Google etc. [4, 5]. These tools make data visualization very easy to understand, interpret and produce rapidly.

Apart from all the above tools we can also create visualization using Python programming. Python provides multiple great graph designing libraries that are packed with large numbers of different features. If you are well

versed in Python programming and you want to produce live, interactive and highly customized plots python has excellent libraries for you. Here are a few popular plotting libraries:

- Area: It is for area plots
- Bar: It is for vertical bar charting
- Matplotlib: It is low level and it provides lots of freedom
- Pandas Visualization: It is easy to use interface, built on Matplotlib
- ggplot: It is based on R's ggplot2. It uses Grammar of Graphics
- Hist: Can be used for histogram
- Plotly: It can create interactive plots

The code snippet below is responsible to generate the visualization:

A. Examples with Python

```
#Example to draw line plot
import sys
import matplotlib
matplotlib.use('Agg')
import matplotlib.pyplot as plt
import numpy as np
ypoints = np.array([3, 6, 1, 9])
bpoints = np.array([5, 5, 1.5, 8])
plt.title('Line by Ahmad Tasnim')
plt.plot(ypoints, color = 'g')
plt.plot(bpoints, color = 'b')
plt.show()
#These lines to make our compiler able to draw:
plt.savefig(sys.stdout.buffer)
sys.stdout.flush()
```

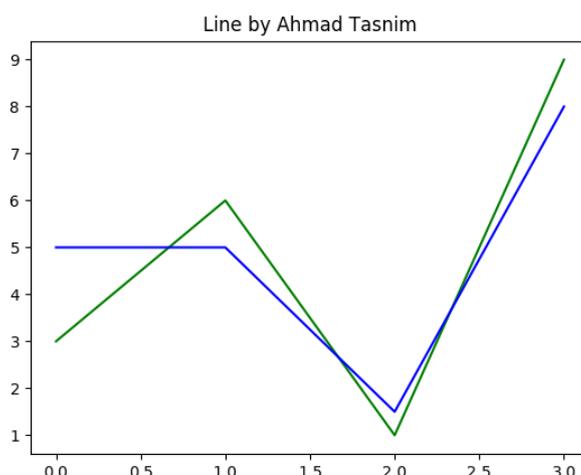


Fig. 2. Line plot created by Matplotlib using Python

We can use data visualizations for:

- Identifying trends and outliers inside a data set
- Making data attractive and easily understandable
- Emphasizing on important parts of a data set
- Telling a story found within the data
- Highlighting an argument or belief
- Make more interactive books, blog posts, reports etc.

B. Creating Report in Power-BI

After finishing with reports in Power BI Desktop, I would like to install the version of Power BI Desktop which is optimized for Power BI Report Server, because you know the server and the app are always in sync. We can have both versions.

- We can connect to a variety of data sources. From the welcome screen, select the Get Data option and import the file/files you want to. I'm using an On Street Crime dataset which is downloaded from Kaggle.
- Now, get to the Data View and make the necessary alterations in the data we've imported.
- We can now begin designing the report.
- After the things done, need to publish the report to the Power BI website.

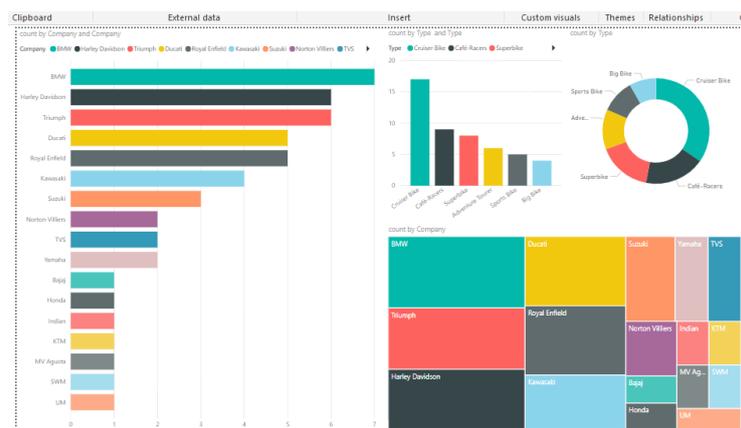


Fig 3. Report generated by Power-BI [10]

III. COMMON FEATURES IN BEST DATA VISUALIZATION TOOLS

We know that a lot of tools for data visualizations are available, but there are few things in common in best data visualization tools. First and foremost is ease of use. The best tools can handle massive data sets. They can provide output of different charts, graphs, and maps types. There are some exemptions to the range of output criteria, however some visualization tools emphasis on an explicit type of graph, chart or map and produce it superbly. These tools also referred among the “top” tools.

Finally, it is necessary to be cost effective. A higher price range doesn't means to eliminate a tool, but the higher price label should be justified in relations to better features, better support, and overall great value. Some crucial features any data visualization tool possess are:

- Decision-making Ability
- Integration Capability

- Highly-effective Infrastructure
- Prompt Discovery of Results or Insights
- Real-time Team Collaboration

IV. CHALLENGES IN DATA VISUALIZATION

“What anybody can think the most significant issue in data visualization is?”

The answers mostly landed in three major categories:

- Issues around data,
- Visualization in practice, and
- The general profession.

Data visualization has changed the scenario and it is facing great challenge due to the massive amount of data. Data visualization of large and complex data has become a demanding task [9]. From the very simple project to complex projects there are lots of things changed. Erroneous, missing and duplicate entry should be identified and corrected data sets should be available to process. We should also think about the size, speed, real-time scalability, interactive scalability, perceptual scalability and the variety of the data [3]. It is a huge challenge to cope with increasing data in static or dynamic form.

V. CONCLUSION

This paper provides information on the field of data visualization. After the discussion we can say that data visualization is the method to represent data in a pictorially, clearly and effectively. It emphasizes not only how to visualize but also how visualization can be helpful [7]. It has appeared as a very powerful, widely acceptable and applicable tool to analyze and understand huge and complex data. Data visualization has come up as a rapid, easy method of conveying ideas in a universally accepted format. It communicates complex ideas with ease, clarity, accurately, and effectively. These benefits make data visualization to be useful in various fields of study. Data visualization is the need of future to reduce the report generation effort.

REFERENCES

- [1]. Midway, S.R. (2020). "Principles of Effective Data Visualization", Cell Press USA. <https://doi.org/10.1016/j.patter.2020.100141>. Date Accessed: 09/03/2021
- [2]. L'Astorina, E. (2017). Review of 20 best big data visualization tools. <https://bigdata-madesimple.com/review-of-20-best-big-data-visualization-tools/> Date Accessed: 09/03/2021
- [3]. Sadiku, M. N. O. et al. (2016) "Data Visualization". International Journal of Engineering Research and Advanced Technology (IJERAT). ISSN: 2454-6135, 02(12)
- [4]. Chen, C. (2010), "Information visualization," WIREs Computational Statistics, vol. 2, July/August, pp. 387-403.
- [5]. Laher, R.R. (2016). "Thoth: software for data visualization and statistics," Astronomy and Computing, vol. 17, pp. 177-185.
- [6]. Data visualization beginner's guide: a definition, examples, and learning resources. URL: www.tableau.com/learn/articles/data-visualization Date Accessed: 10/03/2021
- [7]. Kaidi, Z. (2004). "Data Visualization" National University of Singapore, Document version 1.0
- [8]. Goodman, A.A.; Borkin, M.A and Robitaille, T.P. (2018). "New Thinking on, and with, Data Visualization". 2018 Sackler Colloquium.

- [9]. Aggarwal, R. et al. "Challenges and opportunities with big data visualization". ACM Proceedings of the 7th International Conference on Management of computational and collective intelligence in Digital EcoSystems
- [10]. Komal (2020). "Everything You Need To Know About Power BI Charts". URL: <https://www.edureka.co/blog/popular-power-bi-charts/> Date Accessed: 11/03/2021