Project Report

Project Title: "Open Source Data Visualization".

Submitted By

Tribe M :

1. Aishwarya S Bhandari
2. Bhavani
3. Divyashree
4. Aishwarya
5. Abhilash
6. Bhavana
7. Pallavi
8. Shambhavi S K
9. Ambika
10. Mamata M Nanna
11. Aishwarya H Nandikoor
12. Sachinkumar
13. Shreeshail
14. Kaveri Patil
15. Vijayalaxmi Kalal

Introduction

Data visualization is the graphical representation of any information or data. By using visual elements like 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.

We need to visualize it to explore, understand, and explain our data fully, and that is the power of data visualization.

Open source Data visualization is visualizing and analyzing the data. In the world of Big Data, data visualization tools and technologies are essential to analyze massive amounts of information and make data-driven decisions.

Data visualization refers to the techniques involved in graphically representing data, using visual elements like charts and graphs to spot trends, patterns, and outliers, for quick insights, and to help in real-time decision-making. It's increasingly important in today's world to understand the overwhelming volume of data being generated by businesses every single day.

Figure(a)

Why is data visualization important?

Good data visualization removes noise from data, highlights useful information, and tells a story. Edward R. Tufte, a pioneer in data visualization says, "Graphical excellence is that which gives to the viewer the greatest number of ideas in the shortest time, with the least ink in the smallest space."

Data visualization software for businesses

In this era of big data, businesses need a versatile data visualization software that would solve all their visualization needs. The effectiveness of any data visualization software lies in the richness of data visualization capabilities that it offers.

A data visualization software which offers this excellence will become a vital component of any business and its processes.

Broad goals of data visualization

In a business environment, the speed at which decisions are made is key. It shouldn't take too long to find answers for known questions and indicators. Yet, while dealing with unknown scenarios, there should be sufficient scope to explore the available data in less time-consuming ways.

Hence, for any data visualization tool, the two major goals are:

- ✔ Explanation: Visuals that provide users with relevant information for standard, day-to-day needs (ex: volume of sales in a day across regions)
- ✔ Exploration: Offering a multi-dimensional view of a data-set for users to explore, ask questions as they reflect, and uncover insights along the way (e.g., Performance of a business measured by different parameters over a given time period)

Components of data visualization

Human minds process visuals at an incredibly high speed, and the same also applies to data visualizations. The reason is that it takes less than half a second for the eye and the brain to grasp what is called the pre-attentive visual properties of an image—the color, form, spatial positioning and movement that makes up a data visualization.

A wide range of visualizations can be created with these building components. They include:

- ➔ Time-series visualizations, such as line charts and area charts which depict how a variable, or multiple variables, change over time.
- ➔ Ranking & comparison of different types of bar-charts, like horizontal/vertical, grouped, and stacked.
- ➔ Part-to-whole views of data on a percentage basis using pie charts.
- ➔ Correlation between two or three variables can be effectively understood using scatter plots and bubble charts.
- ➔ Geo mapping visualizations like choropleth—a shaded or pattern-filled map—are useful when it's necessary to compare or segment a data set by geographic region

Types of Data Visualizations

- Charts
- Tables
- Graphs
- Maps
- Infographics
- Dashboards

Charts

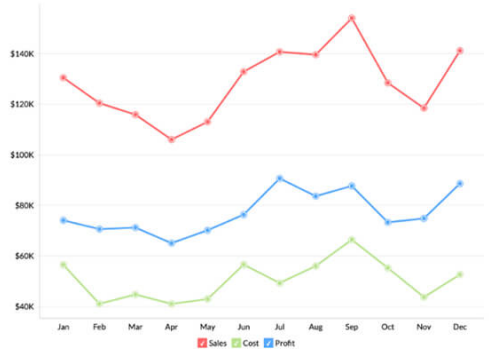
Be it discrete or continuous, charts can visualize any type of data for effective analysis and interpretation. There are a wide range of chart types like Area, Line, Bar, Stacked, Pie, Scatter,

Combination, Funnel, Web etc. Each chart facilitates a wide range of user interaction options for further deeper and contextual analysis.

Some notable and widely used chart types include:

*** Line Chart**

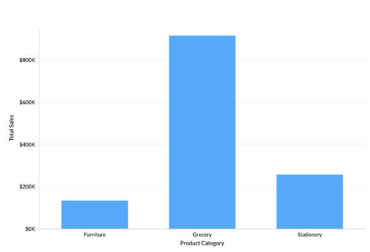
This can be used to visualize trends across any time-period. There are two variants of line chart: Smooth Line and Step chart.

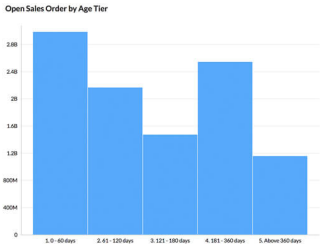


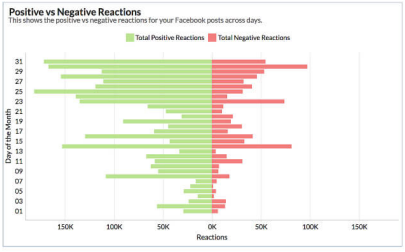
Figure(b): Line Chart

*** Bar Chart**

This displays data as individual bars whose height is proportional to the value plotted. Bar chart is useful to compare data classified into discrete groups.



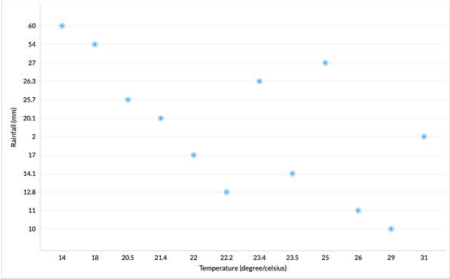




Figure(

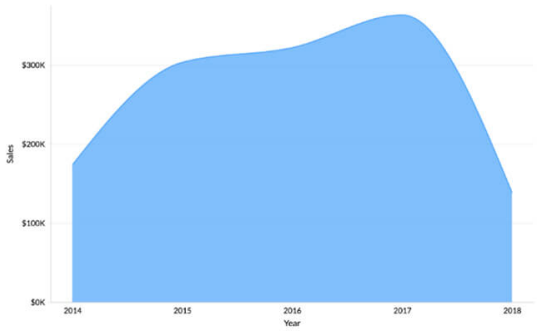
*** Scatter Chart**

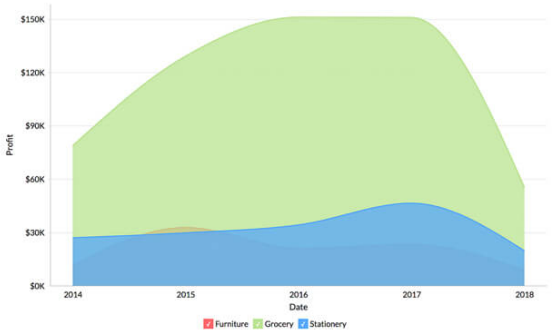
Typical use of the scatter chart is to plot sporadic data with uneven intervals. It is used for the comparison between two numerical axes unlike the line chart, where one axis is never numerical.



*** Area Chart**

Area charts shade the area beneath the lines and therefore help you more readily compare data magnitudes.





Real-time business use cases

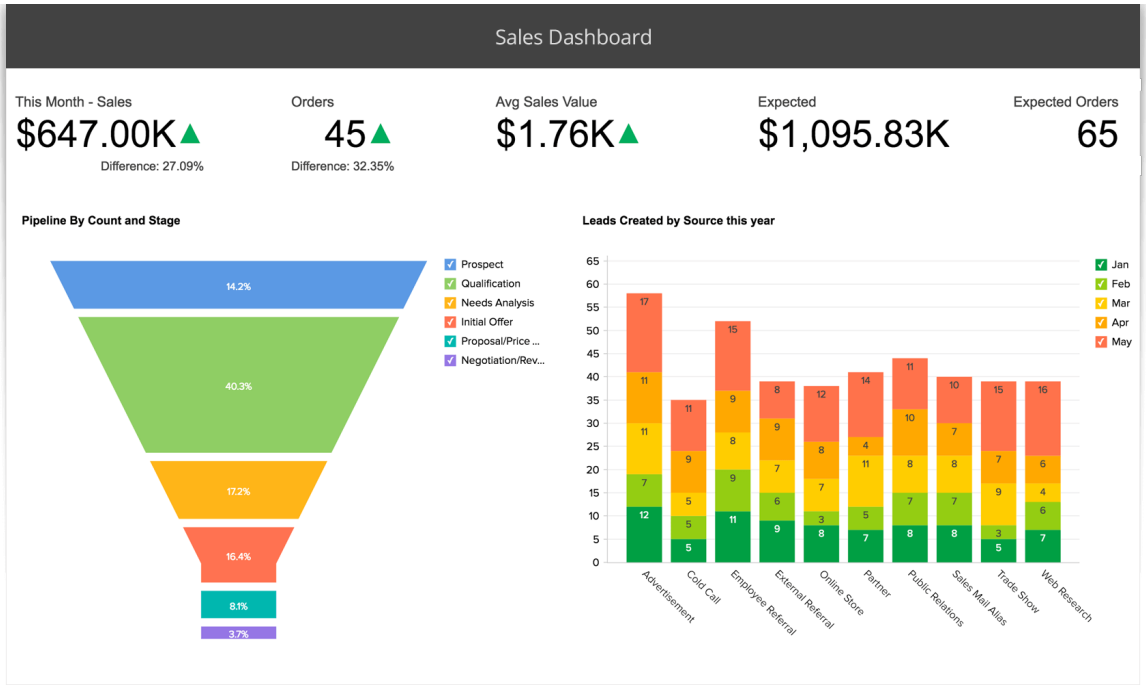
Faster decision-making in business can be better facilitated by precisely offering the context needed for stakeholders.

Choosing appropriate data visualizations for business can be made based on

1) the type of users, 2) the level of detail they're looking for, and 3) their frequency of usage.

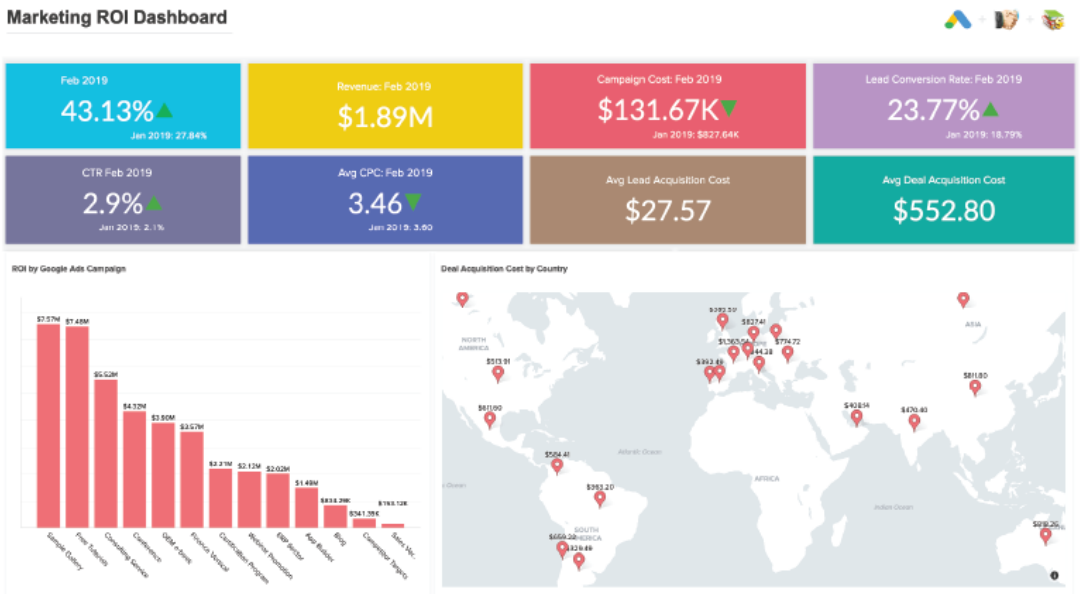
Data visualizations have their specific utilities in different disciplines and industries. Here are just a few examples:

*** Sales:** Visualizations help in evaluating the effectiveness of different channels in achieving the larger sales objectives by communicating data aggregated from multiple tools and sources. Learn more about sales data visualization.



*** Marketing**

Track the impact of your marketing initiatives in terms of engagement and conversion-rates using insightful visualizations. They can also be easily shared across the stakeholders involved for their up-to-date awareness. Learn more about marketing data visualization.



Why is data visualization important?

By now, you would have understood how data visualization simplifies the way information is presented. However, is that the only power of data visualization? Not really. As the world is changing, the need for information is changing as well.

Here are a few benefits of data visualization:

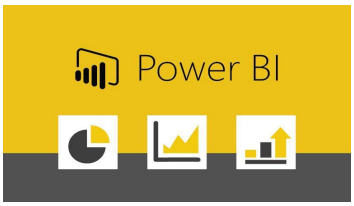
1. Easily graspable information – Data is increasing day-by-day, and it is not wise for anyone to scram through such a quantity of data to understand it. Data visualization comes handy then.
2. Establish relationships – Charts and graphs do not only show the data but also establish co-relations between different data types and information.
3. Share – Data visualization is also easy to share with others. You could share any important fact about a market trend using a chart and your team would be more receptive about it.

PROBLEM STATEMENT:

“ We have a client who wants to set up his own restaurant in Bangalore city. He wants to know how it can be profitable to him by setting up the restaurant.”

SOLUTION: We visualized the given dataset and analyzed the different restaurant’s ratings, total votes given, average bill for two people, availability of facilities like- online ordering, table-booking and we tried to build a dashboard so that our client can get a clear idea about setting up his restaurant.

Tools Used: Power BI(Business Intelligence)



Power BI is a Business Intelligence and Data Visualization tool for converting data from various data sources into interactive dashboards and analysis reports. Power BI offers cloud-based services for interactive visualizations with a simple interface for end users to create their own reports and dashboards.

With a free license, you can explore Power BI for personal data analysis and visualization using My Workspace, but you can't share with other users. A Power BI Pro or Power BI Premium Per User license is required to share content.

Dataset Used: zomato.csv file from Kaggle Website

The basic idea of analyzing the Zomato dataset is to get a fair idea about the factors affecting the establishment of different types of restaurant at different places in Bengaluru, aggregate rating of each restaurant, Bengaluru being one such city has more than 12,000 restaurants with restaurants serving dishes from all over the world. With each day new restaurants opening the industry has'nt been saturated yet and the demand is increasing day by day. Inspite of increasing demand it however has become difficult for new restaurants to compete with established restaurants.

With such an overwhelming demand for restaurants it has therefore become important to study the demography of a location. What kind of food is more popular in a locality? The entire locality loves vegetarian food. If yes then is that locality populated by a particular sect of people for eg. Jain, Marwaris, Gujaratis who are mostly vegetarian. These kind of analysis can be done using the data, by studying the factors such as

Location of the restaurant, Approx Price of food Theme based restaurant or not, Which locality of that city serves that cuisine with maximum number of restaurants, the needs of people who are striving to get the best cuisine of the neighborhood, Is a particular neighborhood famous for its own kind of food.

The dataset contains 17 columns: String 12, Boolean 2, Integer 1, Other 2.

Project Implementation:

Step 1: Load the dataset i.e the zomato.csv file into the power BI tool using the Get data option by selecting the type of file you want to load.

Step 2: Clean/remove the unnecessary columns from the dataset by transform data option. Before cleaning the data there were 17 columns in the dataset, after removing the unnecessary columns we have 11 columns to implement.

Step3: Implementing the required fields in our dashboard using appropriate visual forms or graphs.

i) Restaurant Types: It includes various themed restaurants such as Buffet, Cafes, Delivery, Desserts, Dine-out, Drinks & nightlife, Pubs and bars. We used slicer visual form to implement this.

ii) Average Bill Amount: It includes approximate cost for two people, we have used a gauge visual form to represent average bill amount.

iii) Total Votes: we have used card visual form for this which generally used to represent total of any field, eg: total sales etc.

iv) Average rating: it gives the average rating(out of 5 from our dataset) of a particular restaurant type.

v) Location: It contains various locations in bengaluru city with famous restaurants. Here a slicer visual form is used.

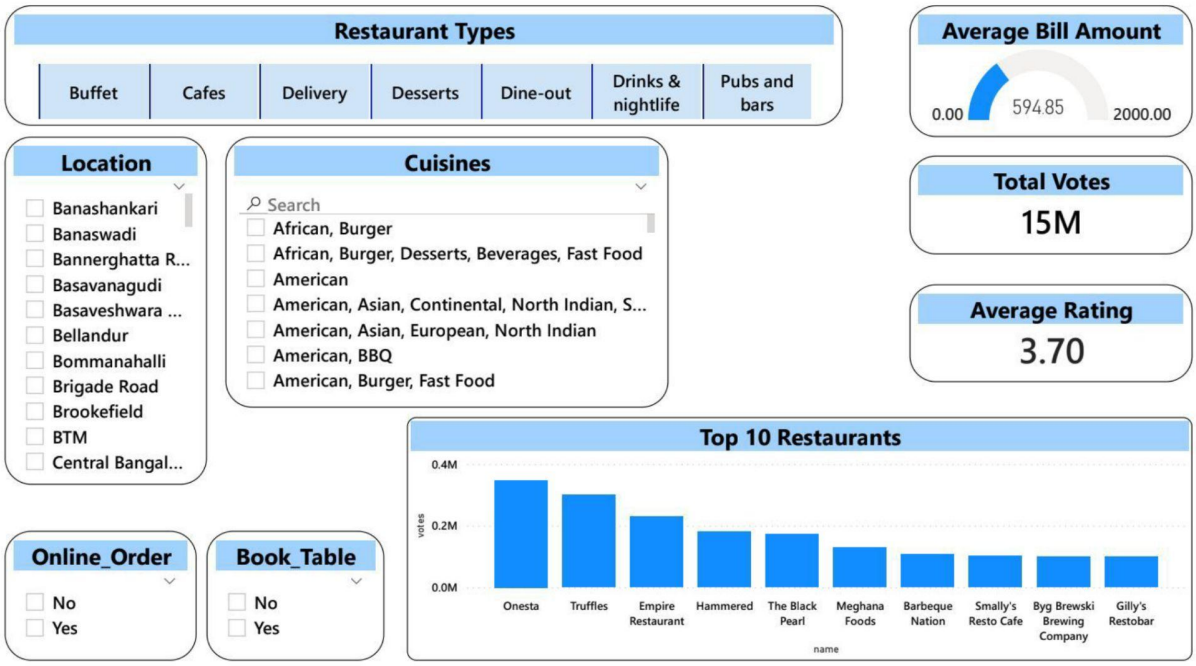
vi) Cuisines: It includes a list of food items from different parts of the world. Here a slicer visual form is used.

vii) Onile_Order: it represents whether a particular restaurant has a facility to order online through zomato or not, slicer visual form is used here.

viii) Book_Table: it represents whether a particular restaurant has a facility to book tables or not, a slicer visual form is used here.

ix) Top 10 Restaurants: It is represented by a stacked column chart with names of top 10 restaurants on x-axis and votes field on y-axis.

Result:


Restaurant Types
Buffet
Cafes
Delivery
Desserts
Dine-out
Drinks & nightlife
Pubs and bars
Average Bill Amount
0.00
594.85
2000.00
Location
Banashankari
Banaswadi
Bannerghatta R...
Basavanagudi
Basaveshwara ...
Bellandur
Bommanahalli
Brigade Road
Brookefield
BTM
Central Bangal...
Cuisines
Search
African, Burger
African, Burger, Desserts, Beverages, Fast Food
American
American, Asian, Continental, North Indian, S...
American, Asian, European, North Indian
American, BBQ
American, Burger, Fast Food
Total Votes
15M
Average Rating
3.70
Top 10 Restaurants
0.4M
0.2M
0.0M
votes
Onesta
Truffles
Empire Restaurant
Hammered
The Black Pearl
Meghana Foods
Barbeque Nation
Smally's Resto Cafe
Byg Brewski Brewing Company
Gilly's Restobar
name
Online_Order
No
Yes
Book_Table
No
Yes