

Process of Data Analysis using Data Visualization

Arpita Singh

B.Tech, Information Technology, Motilal Nehru National Institute of Technology (MNNIT, Allahabad)

ARTICLE DETAILS

Article History

Published Online: 02 April 2018

Keywords

Data visualization, Modern, Computer Graphics

ABSTRACT

Data visualization includes in the visual form due to which understanding of data in easier manner. This can benefit all field of study which requires multiple inventing and complicated data. The advent of computer graphics has shaped modern visualization. The paper presents complete details about this topic.

Introduction

There is a great use of data and it should be managed in such a way that can be easily understandable and can be accessed very easily. Due to this a large number of data is circulation on the websites and online platform. Due to this a huge amount of data is uploading as well as used by the scholars and other persons on the daily basis.

Data Visualization:

Data visualization is very much useful and helpful tools for qualitative benefits. This is very much helpful in finding the correct qualitative database.

Why Data Visualization Matters:

Data visualization generates a system so that common people can easily gain large amount of data. When the data is clearly visually it becomes obvious. And the Visualizations helps the person to reach to the conclusion. Data visualization helps in gaining better knowledge and understanding. Without data visualization, it is not favourable for the users and researchers.



Benefits of good data visualization

1. Our eyes are attracted to different color and patterns. With that we can right away find red from blue, square from the circle. We have a culture following visuals and expressions describing the characteristics of data in whole, including everything related to the sources of entertainment.
2. When we see a chart, we find trend outliers and pattern detection. If we can see something, we make it internal with more specific meaning to gather understanding about it.

The uses of Data Visualization as follows:

1. Strong way to find the data with systematic results.
2. The Primary use is the data mining process.
3. Keep up the data cleaning process by finding incorrect and missing values.
4. Which variable to be included and which variable to be removed.

Data Visualization Tools:

Tableau is the greatly-sized facts picture apparatus for making or put right things for united, as a body. It lets you make come into existence maps, graphs, maps and many other giving clear, full pictures.

Info-gram

Info-gram lets you link their (examples of seeing things in your mind) and graphs or charts that give information to actual data. (Examples that should be copied), decorate (with a personal touch) them with added/more (examples of seeing things in your mind)

Chart Blocks

ChartBlocks is a very easy tool which can be used for online purpose that needs/demands (examples of seeing things in your mind) from spreadsheets, (computer files full of information)... (examples of seeing things in your mind) will be (able to reply or react/quick to respond).

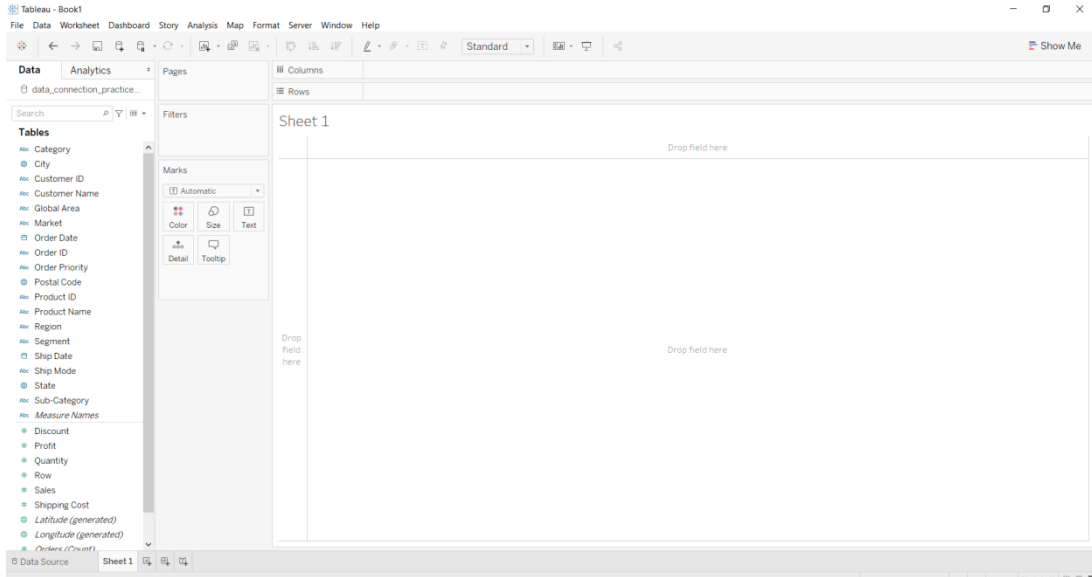
Data Visualization Technique:

- Plots related to box
- Bar diagram
- Maps which are related to heat

Plots related to box

Box plot includes value distribution and shows the value distribution.

All box indicates the half of the data. On both sides, the remaining data represents the clear lines.

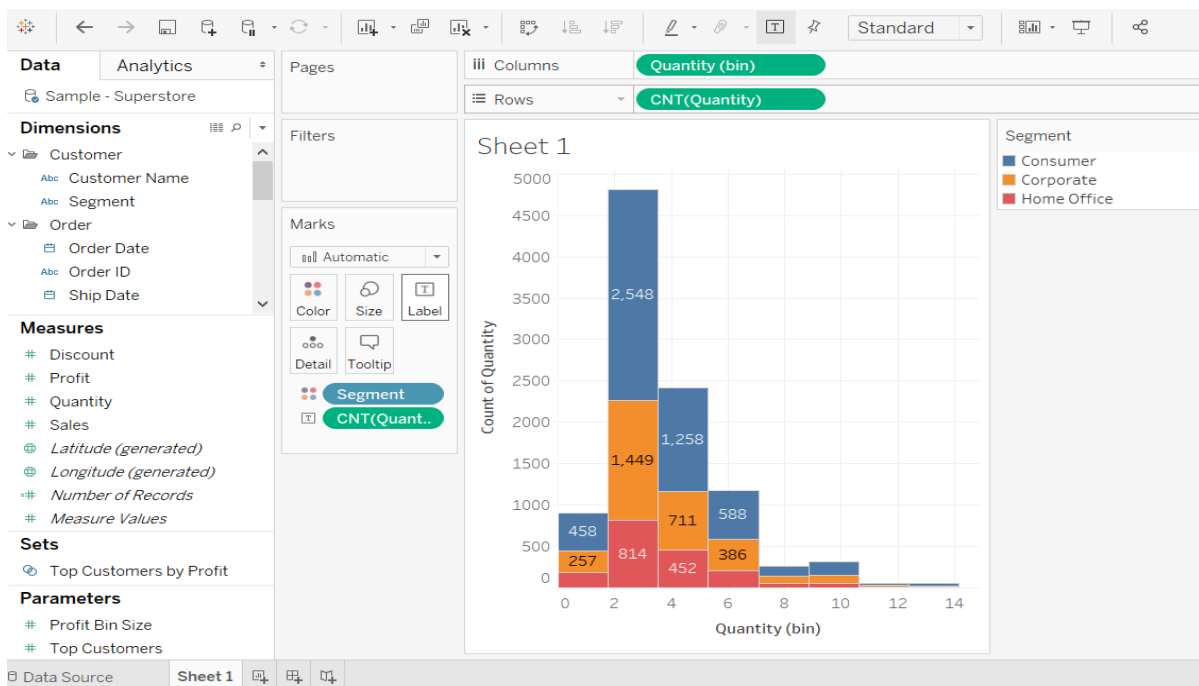


Given above the Framework of the box plot.

To draw a Box of Plot select a only three attributes by drag and drop then select the chart option as box of the plot. A plot of gives you a good idea of how the values changes in the data are changed. plot of box are more valuable and supreme to a. For some datasets, you will find that you required more details than central tendency (median, mean, and mode).

Histogram

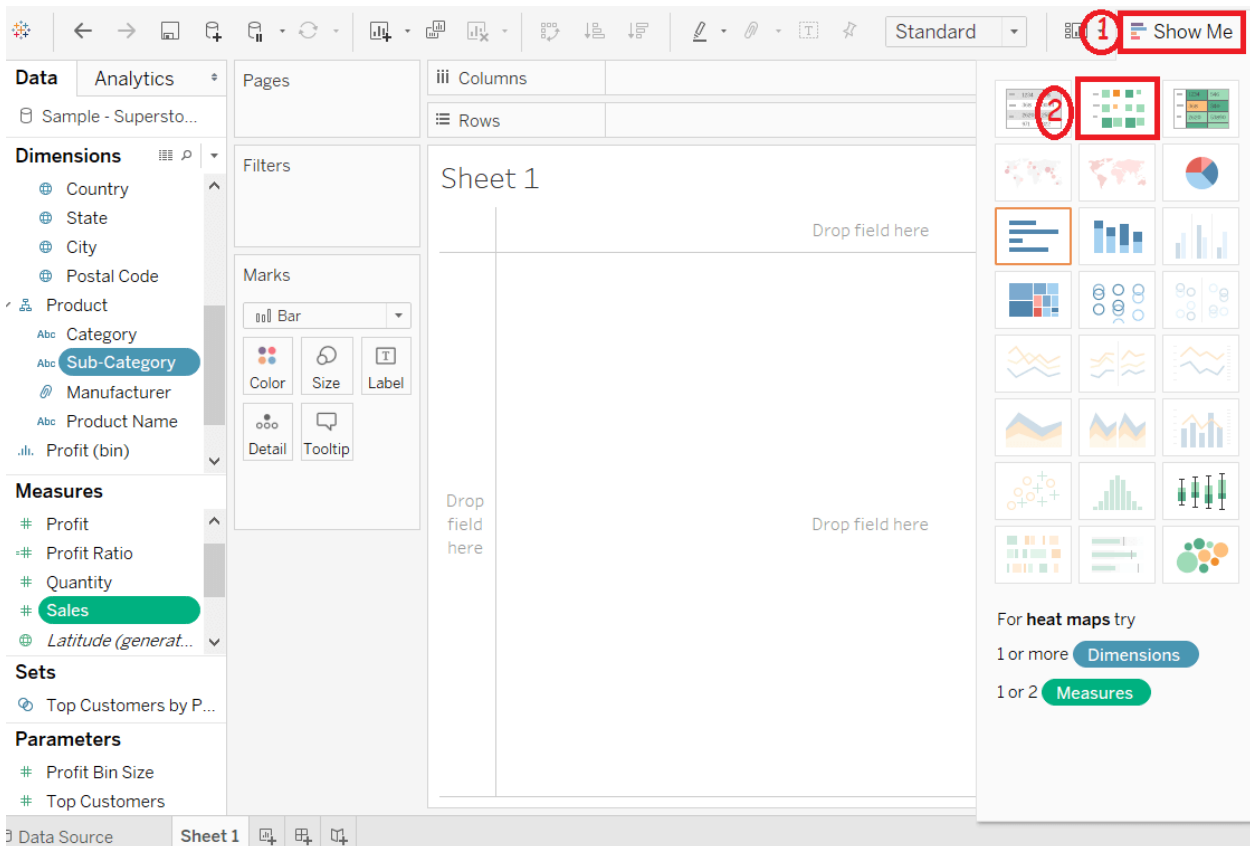
A histogram chart defines the distribution. A histogram looks like a bar chart. In the histogram, each bar defines the height of the number of values present in that range. In the creation of histogram, we required one measure.



Given above is the histogram representation for one category to the relatable attribute:

Heat Map

The heat map is useful to see the data and its size and color on different platforms. Two kinds measures are seen simultaneously with using map of the heat.

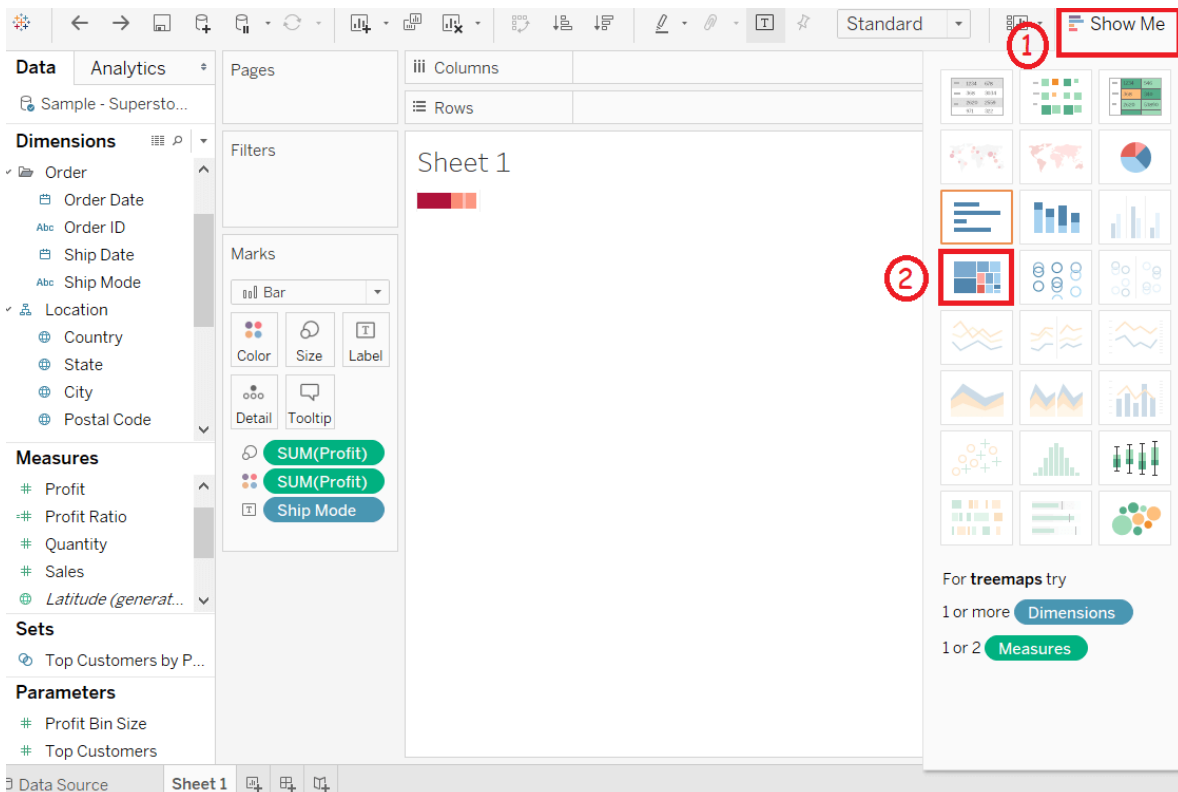


Given above creates the maps related to heat, which used to visualize two relatable fields across different the dimension.

Tree Map

The map of tree shows the data to build rectangular figures. This system shows the system of map of tree and scales determine color ,size and discrete square.

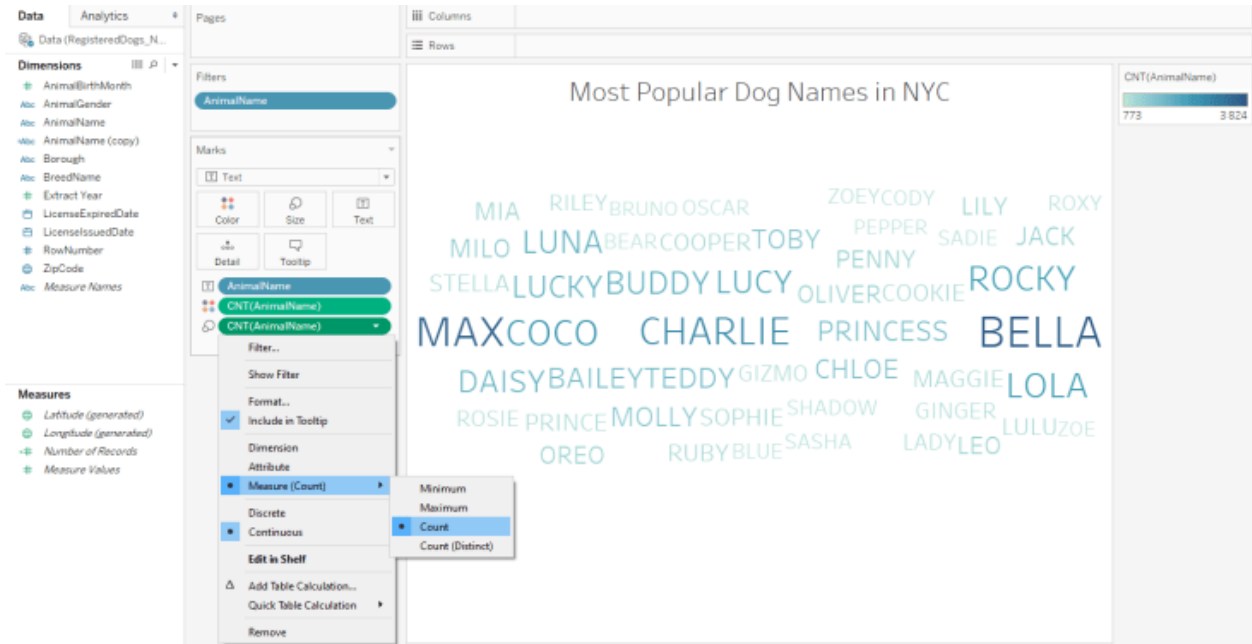
The squares are very much easy to seen as the size of the color.



The cloud of word

A cloud of word may be a clear representation data, which want to illustrate keyword on internet . Word clouds are very

famous type of graphical information with the benefit of which we can represent the frequency of words in data.



Given above is the word cloud representation, and for better visualization.

Methods to Visualize Data:

- **Column Chart:** It is also called an up and down bar chart where each category is given in a rectangular form. The height of the rectangular form is connected with values.
- **Bar Graph:** length and values are properly represented in rectangular bars.
- **Area Chart:** it joints number bars and lines.
- **Line Graph:** The data points are connected through a straight line; therefore, creating a representation of the changing popular thing.
- **Pie Chart:** It is a chart where different parts of a data set are presented in the form of a pie which represents their proportion in the whole data set.
- **Waterfall Chart:** With the help of this chart, the increasing effect of (one after the other) introduced values which can be understood.
- **Scatter Plot Chart:** It is also called a scatter chart or scatter graph. Dots are used to represent values for two different number based (numbers that can be changed).
- **Bullet Graph:** It is used to swap dashboard gauges and meters.
- **Funnel Chart:** The charts decide the flow of users with the help of a business or sales process.

Histograms

- A histogram is a display which is related to the graphical representations of data using bars of various heights. In a histogram, all the bar graphs are

mentioned in proper number. larger bars in height show that more data comes in range.

- It helps in discover, and show, the underlying. This allows the checking of the data for distribution. It is best representation of the division of numeric data, it relates only one variable.

Histograms are based on area, not height of bars

In a histogram, it is the product of multiplication of height and width of the bin that shows the frequency of repetition within that bin. One reason of wrong assessment of the height of the bars due to the presence of equal space bars.

Histogram Vs bar Chart

The major difference is that a histogram is only used to plot the frequency of score occurrences in a continuous data set that has been divided into classes, called bins. Bar charts, on the other hand, can be used for a lot of other types of variables including ordinal and nominal data sets.

Line Chart

The line chart is very much helpful in showing the relation and interdependence between two variables. And this is called the plot functioning.

Bar Charts

Bar charts are generally used for the purpose of measuring and comparing quantity of various several groups bars are represented as value categories.

Pie Chart

It is a graph having circle which helps in showing numerical values. Due to which the arc length of the slice is proportional

to the quantity given. Due to this, these are used to compare the parts.

Scatter Charts

The other common visual technique is a two-dimensional which represents two variation of two different data. Each represents an observation. The marker position shows the value for each observation and calculation. It shows correlations, between X and Y variables.

Bubble Charts

It involves the chart in which the data points are replaced with bubbles, and a different dimension of the data is represented in the size of the bubbles.

Timeline Charts

Timeline charts is the example of such events, which are in the chronological order — for example the progress of a report, — in whatever unit of time the data was recorded. It shows the

perfect systematic sequence of past or future events according to time.

Tree Maps

A map of tree is a visual which shows systematically arranged data as a set of rectangles, with their child elements. The sizes and colors of rectangles are related to the values of the data points which are given. They make efficient use of space, thus display various of items on the display at a particular time.

Conclusion

In this paper, we have discussed open issues related to Data analysis using data visualization. It is revealed that, although these articles have been investigated for almost a decade, a number of important challenges remain. In addition, Data analysis is necessary in the field of research because it makes the study of the data very easy. It helps the researchers to explain the data so that researchers can be helped.

References

1. SAS, Data Visualization: What it is and why it matters, www.sas.com/ensg/insights/big-data/datavisualization.html.
2. Childs H, Geveci B, Schroeder W, Meredith J, Moreland K, Sewell C, Kuhlen T, and Bethel EW, 2013. "Research challenges for visualization software," *Computer*, 2013 May 1(5), 34-42.
3. E.Y. Gorodov and V.V. Gubarev, 2013. "Analytical Review of Data Visualization Methods in Application to Big Data," *Journal of Electrical and Computer Engineering*, 2013, Jan
4. *Fundamentals of Data Visualization* by Claus O. Wilke (book preview available online at <https://serialmentor.com/dataviz/>)
5. *The Visual Display of Quantitative Information* (2nd ed.) by E.R. Tufte: one of the foundational book on visualization. It contains a rich set of historical visualization, thoughtful discussion on visualization principles.