Information Visualization for Data Analytics

Visualization is the way of representing facts in graphical form rather than traditionally keeping data as text. This writing focuses on information visualization in specific in contrary to typically explaining its types. Information visualization is the study of visual representation of data, meaning “information that has been abstracted in some schematic form, including attributes or variables for the units of information”, Friendly 2008 describes.

Information Visualization
It is an interesting area to talk about which promises role on Data analysis algorithms that enable extraction of patterns and trends in data, temporal, geospatial, topical, and network visualization and systems that drive research and development. In so many occasions we would infer the importance of visual effects of data, traditionally, in MS-Excel to represent data and results in terms of various charts namely bar chart, pie chart etc. and from those days the application of visualization still remains the same and ever promising on representing data effectively enough to be interpreted by end users and analyst in no time.

Importance of Information Visualization
Presenting data or the derived results in pictorial or graphical form is the core idea behind Information visualization. In crunch situations, the amount of data to be considered for analysis will be of huge volume as people talks in data warehouse rather than less capable data stores, flat files etc as it allows more data to derive accurate result than to be dependent on small amount of data to derive sensitive results very much in areas such as medical data analysis. As a result, the information derived may not be directly useful for decision makers and it may require analysts to compare to find relevance among results derived. So analysts have been dependent on visualization, presenting data in graphical form to easily understand and interpret information. Typical needs, briefly saying, of visualization are spreadsheets are difficult to visualize, toughness of interpretation of data shown in distinct way and so forth.

Visualization in Machine Learning
Traditionally, systems for analysis were developed to have well defined instructions. In converse to that, Machine learning gives an interesting way to develop systems which can derive instructions from given dataset. Instructions in many ways can be referred to as patterns that are highly useful in decision support systems. Normally, it will be of If...then rules, which is extracted from given dataset rather than defining it explicitly is said to be a process of learning. Data Mining is the vast area which provides powerful algorithms such as association rule mining, classification, clustering etc. for predictive analysis. Area of applications of data mining is unlimited but includes stock market analysis, weather forecasting, credit card fraudulent analysis and heart attack prediction. We would normally see that to represent data charts, maps, graphs being used. Before to visualize, the raw data needs to be processed and analysed in systematic way.

Steps in Data Mining Process to Visualize Information
Information visualization is a phased approach rather that just showing data in some manner. Here the primary concern is to visualize data which should be meaningful of some sort. Raw data needs to be cleaned, analyzed and visualized in the end. Data cleaning includes removing noisy data, reducing redundancy, filling up missing values and listed here comes under pre-processing as process. The whole approach tells how to take raw data, extract meaningful information, then use statistical tools and make visualizations. Even though reporting tools and some OLAP (Online Analytical Processing) queries can be used to retrieve data from data warehouse usually a data store which is capable of holding huge volume of data, it is of applying highly effective data mining algorithms on Data Warehouse to better extract useful information that will be visualized as an end result and which can be used in DSS (Decision Support Systems) also. Any analytics has no/fewer roles in decision support needs to undergo a lot of changes to effectively support forecasting and predictive analysis.

Decide on which Visual is Best
There are few criteria to be considered for generating the best visuals for presenting your data:

- Keen on data to visualize and its attributes, dimensions etc.
- Fix on what to be visualized and the kind of information to communicate.
- Select a visual type which express the information in the best and simplest form possible.

The biggest challenge in producing visualization is deciding on what visualization will best suit to display the information accurately. For an instance, Visualization tools like SAS Visual analytics has intelligent autocharting which is to create best suited visual for the data being selected. Especially for non-technicians and business people this will be a handy solution for decision making and much to understand information. The below figure is an example bar-chart generated from SAS which displays information by means of bar chart to quickly assess and interpret information.

It is not the case that autocharting always produces visualizations that you want. In specific case, user can select one from the list to create the intended visual type. It produces bar chart, scatter plot, map frequency chart for geographic data, histogram, multiscape, parabox, data constellation and these are the few you would opt for.

Visualization to Deal with Big Data
Case Studies show that human brain can process image faster than text. The final step in big data analytics workflow as just in data mining process, the big data analytics visualization is a graphical/pictorial (visual) representation of the knowledge patterns gained from analysis. There is a tool in name of Datameer, for...
big data visualization, presentation/representation is tied to the analysis, so as the data changes the visualization will also change consequently.

**Conclusion**

Seeing one picture can be of worth reading 100 lines of text; better wordings cannot be there to express the importance of visualization. There are many visualization tools available for distinct purposes as single methodology or tool cannot survive on specific needs of different application data. This article may not be a big deal for getting to know everything about information visualization in a single place, yet it can provide some useful facts and information about visualization as a concept and the key application areas of concern. More importantly this can take you for a small tour of visualization concepts.

**References**


---

**About the Author**

S.Palani Mururgan is an Assistant Professor, Department of Information Technology of E.G.S Pillay Engineering College, Nagapattinam. He is passionate about technical writing. He loves to teach Data Mining and Cloud Computing. His research interests are Soft Computing in Data Mining & Predictive Analysis for Cloud Security. He can be reached at suyambhu@gmail.com.