UNIT – V Big Data Visualization

Introduction to Data visualization, Challenges to Big data visualization, Conventional data visualization tools, Techniques for visual data representations, Types of data visualization, Visualizing Big Data, Tools used in data visualization, Propriety Data Visualization tools, Open – source data visualization tools, Analytical techniques used in Big data visualization, Data visualization with Tableau, Introduction to: Pentaho, Flare, Jasper Reports, Dygraphs, Datameer Analytics Solution and Cloudera,Platfora, NodeBox, Gephi, Google Chart API, Flot, D3, and Visually.

Data Visualization

- Visualization : a pictorial or visual representation Technique
- Anything which is represented in pictorial or graphical form with the help of diagrams, charts, pictures ,flowcharts etc.
- Data visualization is pictorial or visual representation of Data.

Types

Name	Description	Tool
1 D /Linear	List of Items	No tool is used
2D /Planer	Data represented in the form of charts ,diagrams, cartograms	Google Maps ,Tableau Public, Many eyes etc
3D/Volumetric	3 D Computer Modeling ,surface rendering ,Computer Simulations	AC3D,TrueSpace
Temporal	Timeline, time series, Gantt charts,	Tableau public ,Google charts, TimeFlow, Timeline JS , Excel
Multidimensional	Pie charts ,histograms, bar chart, scatter plot	Tableau Public, Many eyes ,Google charts etc
Tree/ Hierarchical	Dendograms, Hyperbolic Tree	D3,Google charts and Network Workbench/Sci2
Network	Matrix,node linkk plots etc	GUESS, Many eyes,Sci2 etc

Visualizing Big Data

- Heterogeneity of data sources ,data streaming and real time data difficult to handle by traditional tools
- Traditional tools
  - Use relational model
  - More suitable for static interaction
  - Not suitable for Big data due to response time
- Need for more robust analytical tools

- Analytical tools should be
  - Simple so that non technical users can operate it
  - Interactive to connect with different data sources
  - Competent to create visual interpretations
  - Able to interpret big data

Open source Tools

- VTK
- Cave5D
- Tulip
- IBM OpenDX
- Tableau Public

Propriety Tools

- Tablaue
- NVivo
- Platfora
- Sas
- Lightening Chart

Analytical techniques used in Big data visualization

- Commonly used analytical techniques
  - Regression Analysis
  - Grouping Methods
  - Multiple equation models

- Regression Analysis
  - Statistical tool used for prediction
  - Types
    - Ordinary Least Squares Regression
    - Logistic Regression
    - Hierarchical Linear Modeling

- Grouping Methods
  - Categorizing observations into significant or purposeful groups
- Multiple Equation models
  - Used to analyze pathways
    - Path analysis
    - Structural equation Modeling
-

Tableau

- Tableau Desktop
- Tableau Server
- Tableau Reader
- Tableau Public

Tableau

- Tableau is a Business Intelligence tool for visually analyzing the data.
- Users can create and distribute an interactive and shareable dashboard, which depict the trends, variations, and density of the data in the form of graphs and charts.
- Tableau can connect to files, relational and Big Data sources to acquire and process data.
- allows data blending and real-time collaboration, which makes it very unique.
- Used by businesses, academic researchers, and many government organizations for visual data analysis.

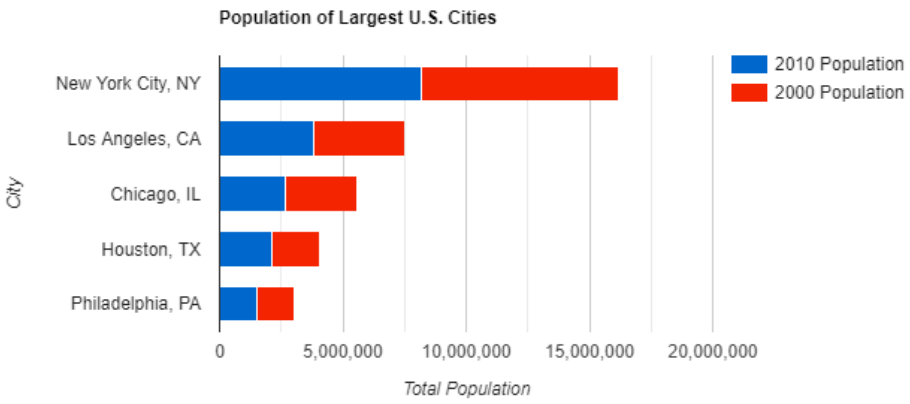
Features

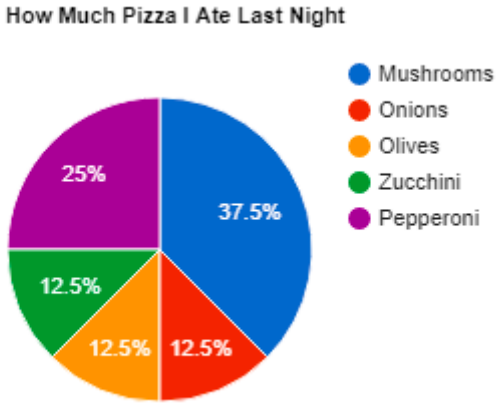
- Single click data analytics in visual form
- In-depth statistical analysis
- Management of Meta data
- In built data engine
- Quick and accurate data discovery
- Business dashboard creation
- Easy and quick integration of R
- Analysis of data from surveys

And many more

Google Charts API

- Allows user to create dynamic charts embedded in a web page
- Chart obtained from data and HTTP request is converted to PNG image by Google





Flare

- Interactive Data Visualization tool
- Action Script Library which run in the Adobe Flash Player
- Toolkit Supports
  - Basic Charts to Complex interactive Graphics
  - Data Management
  - Visual encoding

D3

- Data-Driven Documents
- A powerful visualization tool
- Used through the web
- open-source JavaScript library developed by Mike Bostock to create custom interactive data visualizations browser on the Web

D3

- The following web standards are heavily used in D3
- HTML
- DOM
- CSS
- SVG
- Javascript

Visualization tools

- Jasper Reports
- Pentaho
- Dygraphs
- Datameer Analytics Solution
- Cloudera
- Platfora
- NodeBox
- Gephi,
- Flot
- Visually