

Experiment -8

Aim :- To perform Graph analytics and visualization using Tableau

Theory: Data visualization is graphical representation of information & data by using visual elements like graphs and maps, data visualization tools provide an accessible way to see and understand trend outlines and pattern in data. In the world of big data, data visualization tools & technologies are essential to analyze massive amount of information & make data-driven decisions.

Advantage

our eye are drawn to color & pattern we can quickly identify red from blue & blue from circles - Data visualization is another form of visual art that grabs our interest & keeps our eye on message.

Some other advantages of data

- easily sharing information
- I can quickly explore opportunities
- visualize pattern & relationship.

Disadvantage:-

While there are many advantages, some of disadvantages may seem less obvious, for ex. when viewing a visualization with many different data part, it's to make an acute assumption.

Some other disadvantages

- Biased or inaccurate information
- correlation doesn't always mean causation
- core message can get lost in translation

General type of visualization

Chart :- information presented in tabular graphical form with data displayed along two axes.

Table : A set of figure displayed in row

Graph : - A diagram of points, lines segment, curve or area that represent certain variable in comparison to each

Infographic: A combination of visuals & words that represents data.

More Specific Examples

- Area Map : A form of geospatial visualization map are used to show specific value set over a map of country state, country or any other geographic location.
- Bar chart :- Bar charts represent numerical values compared to each other.
- Box & whisker plot :- These show a selection of ranges across a set measure.
- Bullet graphic:- A bar marked against a background to show progress or performance against a goal.
- Gantt chart : Typically used in project management Gantt charts are bar chart depiction of timeline.
- Heat map: A type of graphical visualization in map from which display specific data
- Highlight table:- A form of table that uses color categories similar data

pie chart:- A circular chart with triangular segments that shows data as percentage of whole

Tree map: A type of chart that shows different related values in form of rectangle nested together

Conclusion - Successfully performed graph analytics & visualization using tableau

Experiment - 9

aim:- To implement basic function & command in R programming better visualization then data table.

Theory :-

R-charts & graph

R language is mostly used for Statistics and data analytics purpose to represent the data graphically in software. There are hundred of chart and graphs present in R.

Type of R-charts

- Bar plot
- piece diagram
- Histogram
- Scatter plot
- Box plot.

Bar plot or Bar chart

Bar-plot or bar chart in R is used to represent the value in data vector as height of bars. The data vector passed to function is represented over Y-axis of graph. Bar chart can behave like histogram by using table () function instead of data vectors.

Pie Diagram or Pie chart

Pie Chart is Circular chart divided into different segments according to ratio of data provided. The total value of pie is 100 and segments tell the fraction of whole pie. It is another method to represent statistical data in graphical form and `pie()` function is used to perform

Histogram

Histogram is graphical representation used to create a graph with bars representing frequency of grouped data in vector. Histogram is same as bar chart but only difference between them is histogram represents frequency of grouped data rather than data itself.

Conclusion - Successfully implemented basic function of command in R programming better visualization then data table



Experiment - 10

Aim : To use following platforms for solving any big data analytic problem of your choice

- 1) Amazon web services
- 2) Microsoft Azure
- 3) Google App engine

Theory :

- 1) Amazon web services.
 - offers a wide range of services for big data analytics, including managed Hadoop services, real-time streaming data processing & data warehousing solution.
 - provides flexible pricing option, allowing user to pay only for what they use
 - offers a high level of scalability & reliability with the ability to handle large volume of data & traffic
 - has large & active user community, with wealth of resources & support available

Aws is cloud computing platform provided by Amazon. It offers a variety tools & services to help to solve big data analytics problem including,

- 1) Amazon elastic map reduce
- 2) Amazon Athena
- 3) Amazon Redshift
- 4) Amazon Kinesis.

2) Microsoft Azure

- Provide a range of big data services, including managed hadoop services, real-time streaming data processing, & machine learning capabilities.
- Offers a variety of pricing options, including pay-as-you-go and reserved in the instance.
- Integrates well with other Microsoft products and services, such as Windows Server & Office 365.
- Provides strong security features, including encryption, access controls & compliance certifications.
- Has a large and growing user community, with a wealth of resources & support available.

Microsoft Azure is cloud computing platform provided by Microsoft. It offers a range of services to help solve big analytics problem including

- 1) Azure HDInsight
- 2) Azure Data Factory
- 3) Azure Stream Analytics
- 4) Including Azure HDInsight

3) Google App Engine

- offers a range of big data services, including cloud-based data storage, data warehousing, & real-time & batch data processing.
- provides flexible pricing option, including pay-as-you-go & sustained use discount.
- Integrates well with other Google products & services, such as Google Analytics & Google Ads.
- provides strong security features, including encryption, access controls, & compliance certifications.
- Has a large & growing user community with wealth of resources & support available.

Google App Engine is (PaaS) provided by Google. It provides a range of services to help solve big data analytics problems, including

- 1) Google Cloud Dataflow
- 2) Google BigQuery
- 3) Google Cloud Dataproc

Conclusion - Successfully used AWS, Microsoft Azure & Google App Engine platforms for solving big data analytic problem.