IIT ROORKEE

NPTEL ONLINE CERTIFICATION COURSE

VISUALIZATION TECHNIQUES Part-2

LECTURE 08

DR. GAURAV DIXIT

DEPARTMENT OF MANAGEMENT STUDIES

VISUALIZATION TECHNIQUES

- Scatterplot
  - Useful for prediction tasks
    - Focus is on finding meaningful relationships between numerical variables
  - Useful for unsupervised learning tasks such as clustering
    - Focus is on finding information overlap
  - Both the axis are used for numerical variable
  - Open RStudio

VISUALIZATION TECHNIQUES

- Distribution Plots
  - Histogram and Boxplot
    - Distribution of a numerical variable
    - Directions for new variable derivations
    - Directions for binning of a numerical variable
  - Useful in supervised learning, specifically prediction tasks
    - Variable transformation in case of a skewed distribution
    - Selection of appropriate data mining method

VISUALIZATION TECHNIQUES

- Boxplots
  - Display entire distribution
  - Side-by-side boxplots for comparing groups
    - Importance of numerical predictors in classification tasks
  - Series of boxplots for changes in distributions over time
  - Open RStudio

VISUALIZATION TECHNIQUES

- Histograms
  - Display frequencies covering all the values
  - Vertical Bars are used
  - Open RStudio
- Heatmaps
  - Display numeric variables using graphics based on 2-D tables
    - Color schemes are used to indicate values
  - Useful to visualize correlation and missing values
    - Specially, in case of large no. of values

VISUALIZATION TECHNIQUES

- Multidimensional Visualization
  - Multiple panels
  - Color
  - Size and shape
  - Animation
  - Aggregation, rescaling, and Interactivity
  - Main idea is to help build visual perception to support the subsequent analysis
- Open RStudio

VISUALIZATION TECHNIQUES

- Multidimensional Visualization
  - Trend Lines
  - In-plot labels
  - Scaling Up
  - Multivariate plots
- Specialized Visualization
  - Network graphs (Network data)
  - Treemaps (Hierarchical data)
  - Map Charts (Geographical data)

- Open RStudio

Key References

- Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data by EMC Education Services (2015)
- Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XLMiner by Shmueli, G., Patel, N. R., & Bruce, P. C. (2010)

Thanks…