DIMENSION REDUCTION TECHNIQUES

LECTURE 13

DR. GAURAV DIXIT
DEPARTMENT OF MANAGEMENT STUDIES

DIMENSION REDUCTION TECHNIQUES

- Large no. of variables
  - Subsets of variables might be highly correlated
  - Computational issues
  - Costs of data preparation, exploration, and conditioning
  - Dimensionality (Principle of Parsimony)
- Dimension Reduction is also called as factor selection or feature extraction is some domains

DIMENSION REDUCTION TECHNIQUES

- Dimension Reduction Techniques
  - Domain Knowledge
  - Data Exploration Techniques
  - Data Conversion Techniques
  - Automated reduction Techniques
  - Data Mining Techniques

DIMENSION REDUCTION TECHNIQUES

- Domain Knowledge
  - Identifying key variables for the data mining task
  - Removing redundant variables
  - Identifying erroneous variables
  - Measurement issues for variables

DIMENSION REDUCTION TECHNIQUES

- Data Exploration Techniques
  - Descriptive statistics
    - Summary statistics
    - Pivot tables
    - Correlation analysis
  - Visualization Techniques

- Open RStudio

DIMENSION REDUCTION TECHNIQUES

- Data Conversion Techniques
  - Combining categories
  - Converting a categorical variable into a numerical variable

- RStudio

- Automated reduction Techniques
  - Principal Component Analysis (PCA)

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...