Advanced Analytics using Statistics
PG-DBDA September 2021

List of Books / Other training materials

Text Book:

1. Statics Using R by Sudha Purohit, Pub: Narosa

Reference:

1. Beginning R – The Statistical Programming Languageby Dr. Mark Gardener PUB: WILEY
2. Art of Programming in R, by Norman Matloff
3. Statistics for Management by Levin
4. Business Analytics: Methods, Models, and Decisions by James R Evans
5. Introductory Statistics with R (Statistics and Computing) by Peter Dalgaard
6. R in a Nutshell by Joseph Adler (O'REILLY)
7. R Cookbook by Paul Teetor (O'REILLY)
8. The R Book, Second Edition
9. Statistics Using R, Shailaja Deshmukh, Sudha Purohit, Sharad Gore, Pub: Narosa

Note:

- Each session mentioned is for theory and of 2 hours' duration. Lab assignments are indicatives; faculty needs to assign more assignments for better practice.
- Trainer has to teach the statistical and probability concepts involved here in detail
- Trainer must teach 'Scipy' package in detail.

Session 1:

- Introduction to Analytics
- Data analytics Life Cycle:
- Discovery,
- Data preparation
- Model planning
- Model building implementation
- Quality assurance
- Documentation
- Management approval
- Installation
- Acceptance and operation

Session 2:

- Intelligent data analysis,
- Nature of Data,
- Analytic Processes and Tools,
- Analysis vs. Reporting
- Modern Data Analytic Tools

Advanced Analytics using Statistics
PG-DBDA September 2021

Session 3:

- Visualization and Exploring Data

Session 4 & 5:

- Descriptive Statistical Measures
- Summary Statistics - Central Tendency & Dispersion (Mean, Median, Mode, Quartiles, Percentiles, Range, Interquartile Range, Standard Deviation, Variance, and Coefficient of Variation)

Session 6:

- Sample& population, Uni-variate and bi-variate sampling, re-sampling
- Sample Spaces and Events
- Joint, Conditional and Marginal Probability
- Bayes' Theorem

Session 7 & 8:

- Random Variable
- Probability Distribution and Data o Continuous and discrete distribution – (Normal, Binomial, and Poisson distribution)
- Central Limit Theorem

Session 9:

- Sampling and Estimation
- Statistical Interfaces

Session 10:

- Concepts of Correlation
- Covariance
- Outliers

Session 11 & 12:

- Predictive modelling and analysis
- Application
- Types
- Benefits and challenges
- The Future of predictive modelling
- The Limitations of Predictive modelling
- Predictive modelling Tools

Advanced Analytics using Statistics
PG-DBDA September 2021

Session 13 & 14:

- Predictive Modelling (From Correlation to Supervised Segmentation):
- Identifying Informative Attributes,
- Segmenting Data by Progressive Attributive, Models,
- Induction and Prediction,
- Supervised Segmentation,
- Visualizing Segmentations,
- Trees as Set of Rules,
- Probability Estimation;

Session 15:

- Prescriptive Modelling
- Difference between predictive and prescriptive modelling
- How prescriptive analytics works?
- Examples and use cases

Session 16:

- Regression Analysis
- Forecasting Techniques

Session 17:

- Simulation and Risk Analysis
- Optimization, Linear, Nonlinear

Session 18 & 19

- Overfitting and Its Avoidance:
- Generalization,
- Holdout Evaluation Vs Cross Validation;

Session 20:

- Decision Analytics:
- Evaluating Classifiers,
- Analytical Framework,
- Evaluation,
- Baseline,
- Performance and Implications for Investments in Data;

Session 21:

- Evidence and Probabilities:
- Explicit Evidence Combination with Bayes Rule,
- Probabilistic Reasoning;

Session 22:

- Business Strategy:
- Achieving Competitive Advantages,
- Sustaining Competitive Advantages

Session 23:

- Factor Analysis,
- Directional Data Analytics,