Suggested Teaching Guidelines for
Data Analytics PG-DAI February 2025

Duration: 44 Classroom hours **+ 46** Lab hours

Objective: To introduce the student to Data Analytics

Prerequisites: Knowledge of Programming language of Python Programming and basics of statistics.

Evaluation method: Theory exam– 40% weightage
Lab exam – 40% weightage
Internal exam– 20% weightage

List of Books / Other training material

Courseware: Business Analytics James R Evans 2nd Edition

Reference Book:

1. Business Analytics: Data Analysis &Decision-Making Hardcover – Widescreen
2. Data Science for Business, Foster Provost & Tom Fawcett
3. Data Science and Analytics Paperback by V. K. Jain (Author)
4. Data Analytics for Beginners: A Beginner's Guide to Learn and Master Data Analytics, by Robert J. Woz
5. Data Analytics: Become A Master in Data Analytics by Richard Dorsey

Note:
- **Each session mentioned is of 2 hrs of Theory and 2 hours of Lab duration, unless indicated otherwise. T for Theory & L for Lab.**

Session 1 & 2:
- Introduction to Business Analytics using some case studies
- Data analytics Life Cycle:
- Discovery
- Data preparation
- Model planning
- Model building implementation
- Quality assurance
- Documentation
- Management approval
- Installation
- Acceptance and operation

Data Analytics PG-DAI February 2025

Assignment –Lab: Setup workspace in either in Google Colab or Jupyter and go through the documentation of Pandas, Numpy, Sklearn libraries.

Session 3:
- Intelligent data analysis oNature of Data
- Analytic Processes and Tools Analysis vs. Reporting
- Modern Data Analytic Tools

Assignment –Lab: Explore the nature of data.

Session 4:
- Visualization and Exploring Data
- Case studies: Making Right Business Decisions
based on data

Assignment –Lab: Load a sample dataset and explore the data and draw some insights. Also use Matplotlib and seaborn libraries.

Session 5:
- Descriptive Statistical Measures
- Summary Statistics - Central Tendency & Dispersion (Mean, Median, Mode, Quartiles, Percentiles, Range, Interquartile Range, Standard Deviation,
- Variance, and Coefficient of Variation)

Assignment –Lab: Load any dataset and find out the mean, median mode and other central tendencies of the dataset.

Session 6:
- Sample& population, Uni-variate and bi-variate sampling, re-sampling
- Sample Spaces and Events
- Joint, Conditional and Marginal Probability
- Bayes' Theorem

Assignment –Lab: Load any dataset and apply Naive Bayes function and predict the output.

Session 7:
- Random Variable
- Probability Distribution and Data
- Continuous and discrete distribution – (Normal, Binomial, and Poisson distribution)
- Central Limit Theorem

Assignment –Lab: Generate random numbers and check if they are in normal distribution using scipy libraries.

Suggested Teaching Guidelines for
Data Analytics PG-DAI February 2025

Session 8:
- Sampling and Estimation
- Statistical Interfaces

Assignment –Lab: Load any dataset and explore sampling techniques

Session 9:
- Concepts of Correlation
- Covariance
- Pearson Correlation
- Outliers

Assignment –Lab: Load any dataset and find out the covariance between two fields and also find the correlation and determine how two fields are correlated. Also handle the outliers in the data.

Session 10 & 11:
- Hypothesis Testing
- Z-Test
- Chi-Square test
- Skewness

Session 12:
- Predictive modelling and analysis
- Application
- Types
- Benefits and challenges
- The Future of predictive modelling
- The Limitations of Predictive modelling
- Predictive modelling Tools

Session 13, 14 & 15:
- 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
- Prescriptive Modelling
- Difference between predictive and prescriptive modelling
- How prescriptive analytics works?
- Examples and use cases

Data Analytics PG-DAI February 2025

Session 16: (2T + 4L)
- Regression Analysis
- Forecasting Techniques

Session 17:
- Simulation and Risk Analysis
- Optimization- Linear, Nonlinear

Session 18:
- Overfitting and Its Avoidance:
- Generalization
- Holdout Evaluation Vs. Cross Validation

Session 19:
- Decision Analytics
- Evaluating Classifiers
- Analytical Framework
- Evaluation
- Baseline,
- Performance and Implications for Investments in Data

Session 20:
- Evidence and Probabilities
- Explicit Evidence Combination with Bayes Rule
- Probabilistic Reasoning

Session 21:
- Factor Analysis
- Directional Data Analytics
- Functional Data Analysis

Session 22:
- Introduction to KNIME