|  |  |  |  |
| --- | --- | --- | --- |
| **Core Module 5: Business Data Analytics methods and tools: 320 Hrs** | | | |
| **HOUR No.** |  |  |  |
| **Learning outcome Reference** | **Professional Skills (Trade Practical)** | **Professional Knowledge (Trade Theory)** |
| **With Indicative Hours** | |
|  | Able to understand business analytics and develop business intelligence | **Introduction to Business Analytics Overview (60 Hrs.)**   * Use Excel for understanding different types of data (Integer, double, text, date) * Perform operations on different data types. * Segregate data in different sheets. * Calculate arithmetic mean, geometric mean and Harmonic mean * Calculate median from raw & grouped data * Calculate mode for row & grouped data * Calculate standard deviation for set of data * Calculate standard variance for a set of data * Using VLOOKUP in excel for searching operation * Plot basic charts in excel over numeric data series * Plot uniform and binomial distributions in excel * Implement Central limit theorem in excel * Generate data table and find chi-square analysis | **Business Analytics (50 Hrs.)**   * Introduction to business analytics and concepts of business analytics. * Trends in business analytics. * Introduction to Big Data Analytics * Introduction to descriptive statistics and inferential statistics, measure of central tendency and spread. * Types of distributions-uniform, binomial, normal, log, exp * Sampling techniques, population * Probability theories * Bayes’ Theorem, Maximum Likelihood * Hypothesis Testing * Central limit theorem * Chi-square test |
|  | Able to analyze data using statistical and data mining techniques for business intelligence. | **Business Analytics Foundation (60 Hrs.)**   * Install NumPy, pandas, matplotlib, Seaborn, sklearn in python 3 * Creating arrays in NumPy * Creating multidimensional array in NumPy * Numpy Operations, methods and attributes * Numpy case studies * Understanding Pandas series and dataframe * Pandas ingestion of data from csv, json, html, excel, text files * Pandas functionalities for Series & Data Frames * Grouping, Merging, concatenating, joining, segregation * Python lambda function operations on series or data frames * Dealing with missing and noisy data * Finding outliers * Visualising your data through matplotlib under basic charts * Labels, legends and axes * Subplotting, grid, and 3D plots * Plot formatting- custom attribute values * Advanced charts in seaborn- countplot(), jointplot(), boxplot(), heatmap(), regression plot, etc | **Data Analytics using Python (30 hr)**   * Data mining, wrangling, data manipulation techniques * Data cleaning and pre-processing techniques * Data analytics project lifecycle * Numerical Computing using NumPy Library * Multidimensional data handling using Pandas Library * Data Visualization using Matplotlib * Advanced data visualization using seaborn * Pandas profiling for report generation * Need for data visualization |
|  | Able to use machine learning techniques to generate predictive analytics model | **Skills on predictive analytics using ML (80 hr)**   * Installing sklearn library * Simple linear regression using excel * OLS in sklearn * Train-test-split of data in sklearn * Methods of linear regression- fit(), predict(), coeff\_, intercept\_, score() * Creating linear regression model in python * Evaluating linear regression model * Performing minmax scaling and standard scaling * Implementing KNN in python using sklearn * Evaluation of KNN model in python, and visualizing results * Evaluating model using AUC, ROC curve * Implementing logistic regression for binary and multi-class classification * Sigmoid function in Logistic regressions * Predicting probability of classification models * Charting confusion matrix * Integration of analytics with django/Flask app | **Fundamentals of Predictive Analytics using Machine Learning techniques (40 hr)**   * Machine learning and its types & applications * Supervised machine learning techniques * Classification vs regression * Understanding Regression and types * Linear regression using OLS * Multi-Variate Linear Regression * Correlation concepts * Metrics- Loss function, MSE, RMSE, MAE, R2 Score * Residuals in Regression * Polynomial features * Classification techniques * Types of distance metrics * KNN Classification * Gradient Decent * Logistic Regression * Evaluation- Confusion Matrix, Precision, Recall, F1 Score, Accuracy * Python Library: Sci-Kit Learn |
|  | Project work / Industrial Visit | | |
|  | Revision | | |
|  | Examination | | |