



Diploma Program

Data Analytics

Subject Name: Statistical Analysis

Total Units: 14

UNIT I

- Unit 1: Permutation and Combination
- Unit 2: Repetition and Constrained Repetition
- Unit 3: Binomial Coefficients
- Unit 4: Binomial Theorem

UNIT II

- Unit 5: Frequency Distributions
- Unit 6: Histograms and Frequency Polygons
- Unit 7: Measures of Central Tendency

UNIT III

- Unit 8: Elementary Probability theory
- Unit 9: Theoretical Distribution
- Unit 10: Relation between binomial and normal distribution

UNIT IV

- Unit 11: Linear correlation
- Unit 12: Measure of Correlation
- Unit 13: Least Square
- Unit 14: Regression Lines

Subject Name: Data Analytics

Total Units: 14



UNIT I

- Unit 1: Data Analytics Lifecycle - Discovery
- Unit 2: Data Preparation, Model Planning, Model Building, Communicate Results, Operationalize
- Unit 3: Statistics Analysis of Data, Descriptive Statistics, Multivariate Analysis

UNIT II

- Unit 4: Visualisation Before Aanalysis, Dirty Data
- Unit 5: Examining Multiple Variables, Hypothesis Testing
- Unit 6: Difference of Means, Wilcoxon Rank-Sum Test
- Unit 7:Hypothesis Testing, Chi Square Test, T-test, Statsictical Descision

UNIT III

- Unit 8: Distance Measures, K-Means Clustering - Centroids and Distance Measures, How K-Means works
- Unit 9: DBSCAN - Linear Regression -User Case, Model Description
- Unit10: Logistic Regression - User case, Model Description

UNIT IV

- Unit 11: Binary Classification, Decsion Trees
- Unit 12: Overview of Decision Trees, Decision Tree Algorithm
- Unit 13: Naïve Bayes - Bayes Theorem , Naïve Bayes Classifier
- Unit 14: K- Nearest Neighbour Algorithms

Subject Name: Data Analytics using Excel

Total Units: 14

UNIT I



- Unit 1: Overview, Types of Data Analysis
- Unit 2: Data Analysis With Excel
- Unit 3: Functions in Excel
- Unit 4: Calculating the MEAN and MEDIAN, Finding the Mode

UNIT II

- Unit 5: Measuring Variation: VARP and VARPA, Sample Variance
- Unit 6: Population Standard Deviation - STDEVP and STDEV, Sample Standard deviation STDEV and STDEVA
- Unit 7: Data Analysis tool: Rank and Percentile

UNIT III

- Unit 8: IF statements
- Unit 9: Sensitivity Analysis with data tables, the Goal Seek command
- Unit 10: Using the Scenario Manager for sensitivity Analysis, sorting in excel , validating data

UNIT IV

- Unit 11: Summarizing data by using histograms
- Unit 12: Summarizing data by using descriptive statistics, Summarizing data with database statistical functions
- Unit 13: Filtering data and removing duplicates, Consolidating data, Estimating straight line relationships
- Unit 14: Modeling Exponential growth, Using correlations to summarize relationships

Subject Name: Python Programming

Total Units: 14

UNIT I

- Unit 1: Python Basics Keywords and Identifiers



- Unit 2: Python Variables and Definitions
- Unit 3: Python Data Types
- Unit 4: Python Operators

UNIT II

- Unit 5: Python Control Statements
- Unit 6: Looping Statements – I
- Unit 7: Looping Statements – II

UNIT III

- Unit 8: Python Lists
- Unit 9: Python Tuples
- Unit 10: Python Sets
- Unit 11: Python Dictionary

UNIT IV

- Unit 12: Python Functions
- Unit 13: Arrays in Python
- Unit 14: Exception Handling in Python

Semester 2

Subject Name: Fundamentals of Machine Learning

Total Units: 14

UNIT 1

- Unit 1: Introduction to Machine Learning
- Unit 2: Supervised Learning Fundamentals
- Unit 3: Dimensionality Reduction Techniques

UNIT 2



- Unit 4: Introduction to Regression
- Unit 5: Overfitting in Regression
- Unit 6: Classification Performance Measures

UNIT 3

- Unit 7: Linear Methods for Classification
- Unit 8: Decision Tree Algorithm
- Unit 9: Support Vector Machines (SVM)

UNIT 4

- Unit 10: Parameter Estimation Basics
- Unit 11: Unsupervised Learning Fundamentals
- Unit 12: Advanced Clustering Techniques
- Unit 13: Practical Applications using Scikit-Learn - Part 1
- Unit 14: Practical Applications using Scikit-Learn - Part 2

Subject Name: NumPy and Pandas

Total Units: 14

UNIT 1

- Unit 1: Introduction to NumPy Arrays - Narray Object, Array Attributes, Array Creation Routines, Array from Numerical Ranges, Basic Operations, Advanced Indexing, Broadcasting
- Unit 2: Array Manipulation Routines - Changing array shape, Joining arrays, Splitting arrays, Adding and removing elements, Rearranging elements, Unique()
- Unit 3: Advanced Array Functions - NumPy - Iterating over Array, RAVEL(), SWAPAXES(), SPLIT Function, HSPLIT(), VSPLIT(), LEFT Shift and RIGHT Shift Functions

UNIT 2

- Unit 4: Binary Operations - Elementwise bit operations, Bit packing
- Unit 5: String Functions - Basic String operations, Comparison, String information



- Unit 6: Mathematical Functions - Basic Mathematical Functions, Rounding, Power and Reciprocal Functions, Power and Mod Functions, Trigonometric functions, Arithmetic Operations, Extrema Finding, Statistical Functions, Sort, Search & Counting Functions, Copies & Views, Matrix Library
- Unit 7: Getting Started with Pandas - Why Pandas, Features of Pandas

UNIT 3

- Unit 8: Data Structures in Pandas - Series, Data Frame, Panel
- Unit 9: Series Creation - Using ndarray, Using dict, Using scalar values, Using list
- Unit 10: Accessing Elements of Series - Using indexing, Using slicing, Using ranging, Using iloc method, Using loc method

UNIT 4

- Unit 11: Data Loading - Storage and File Formats, Reading and Writing Data in file
- Unit 12: DataFrame Creation - Using list, Using dict, Using ndarray, Using series, Using DataFrame
- Unit 13: Viewing DataFrame Elements - Using describe function, Using column name, Using iloc method, Using iat method, Using head(), Using tail(), Using index method
- Unit 14: Data Cleaning and Preparation - Handling Missing Data, Data Transformation, String Manipulation, Combining and Merging Data Sets, Reshaping and Pivoting

Subject Name: Data Visualization

Total Units: 14

UNIT 1

- Unit 1: Data Visualization in Matplotlib: Features of Matplotlib
- Unit 2: Anatomy of a Matplotlib Plot, Creating a Plot and Figure, Changing Figure Sizes.
- Unit 3: Chart properties: Creating a chart, Labeling the axes, Formatting line style and color, Saving the chart in a file, Styling the chart, Adding annotations, Adding legends.
- Unit 4: Scatter plots, Heat maps, Bubble chart, Bar chart, Histogram, Pie chart, XKCD style
- Unit 5: 3D chart, Box and whisker plots, Time series plot, Graph data / line graph, Geographical data



UNIT 2

- Unit 6: Data Visualization in Seaborn: Installing seaborn; create histograms using seaborn, KDE plots
- Unit 7: Combining plot styles, combine histograms, rug plots, box and violin plots, regression plots, heat maps with seaborn.

UNIT 3

- Unit 8: Data Visualization in R: Introduction to R; ggplot2 foundations- geometries, facets, statistics, export plot
- Unit 9: Data wrangling- data transformation, grouping, piping, pivoting, transform and visualize data
- Unit 10: Exploratory data analysis- histogram and density plot, frequency polygon
- Unit 11: Area plot, bar plot; scatter plot, rug plot, bivariate distribution, boxplot, violin plot, matrix plots;

UNIT 4

- Unit 12: Advanced Data Visualization in R: Size and shape of points- facet wrap, facet grid, rmarkdown
- Unit 13: Pie chart, donut chart, time series visualization, waterfall chart, radar chart, parallel coordinates plot, heat map, mosaic plot;
- Unit 14: Plot customization- themes, annotations and labels.

Subject Name: Data Analytics using R

Total Units: 14

UNIT 1

- Unit 1: Introduction to R Basics
- Unit 2: R Data Structures and Numeric Vectors
- Unit 3: Advanced Vector Operations
- Unit 4: Foundations of Programming in R

UNIT 2



- Unit 5: Matrices and Matrix Operations
- Unit 6: Higher Dimensional Arrays and Avoiding Dimension Reduction
- Unit 7: Characters, Strings, and Lists
- Unit 8: Advanced R Operations with Lists and Matrices

UNIT 3

- Unit 9: Data Frames and Basic Operations
- Unit 10: Factors, Tables, and Advanced Data Frame Operations
- Unit 11: Math, Simulations, and Data Handling

UNIT 4

- Unit 12: Input/Output and String Manipulation
- Unit 13: Statistical Analysis and R Functions
- Unit 14: Graphics and Visualisation