PART A - INTRODUCTION TO STATISTICS

1. Fundamentals of Math and Probability

- Fundamentals of Probability
- Probability distributed function and cumulative distributed function
- Conditional Probability
- Bayesian Theorem
- Permutations & Combinations
- Associative, Commutative and Distributive Laws
- Matrices

➢ Class Hand-on

- Problem solving using R for vector manipulation Problem solving for probability assignments

2. Descriptive Statistics

- Describe or summaries a set of data Measure of central tendency and measure of dispersion.
- The mean, median, mode, Standard deviation, Variance, Range, kurtosis and skewness.
- Histograms, Bar chart, Box plot

➢ Class Hands-on

- 5 Point summary Box Plot, Histogram and Bar Chart Exploratory analytics R Methods

3. Inferential Statistics

- What is inferential statistics Different types of Sampling techniques
  Central Limit Theorem
- Univariate & Bivariate Analysis
- Correlations
- Least Square Regression

- Normal Distribution
- Binomial Distribution
- Point estimate and Interval estimate
- Creating confidence interval for population parameter Characteristics of Z-distribution and T-Distribution Basics of Hypothesis Testing
- Bias & Variance trade-offs
- Type of test and rejection region
- Type of errors in Hypothesis resting, Type-I error and Type-II errors
- False Positive & False Negative
- P-Value and Z-Score Method
- Crosstabulation & Chi-square test
- T-Test, Analysis of variance(ANOVA)
- Regression analysis in ANOVA

➢ Class Hands-on

- Problem solving for C.L.T Problem solving Hypothesis Testing
  Problem solving for T-test, Z-score test
- Case study and model run for ANOVA

4. Hypothesis Testing

- Hypothesis Testing Basics
- Null Hypothesis
- Alternate Hypothesis
- p-Value
- Type of test and Rejection Region
- Type o errors-Type 1 Errors, Type 2 Errors P value method, Z score Method

PART D – PYTHON PANDAS

1. Intoduction to Python

- Introduction to the Python Data Science Tool
- Introduction to the Python Data Science Environment
- Some Miscellaneous IPython Usage Facts
- Online iPython Interpreter

2. Introduction to Python Pre-Requisites for Data Science

- Rationale Behind This Section
- Different Types of Data Used in Statistical & ML Analysis
- Different Types of Data Used Programmatically
- Python Data Science Packages To Be Used

3. Introduction to Numpy

- Numpy: Introduction
- Create Numpy Arrays
- Numpy Operations
- Matrix Arithmetic and Linear Systems
- Numpy for Basic Vector Arithmetic
- Numpy for Basic Matrix Arithmetic
- Broadcasting with Numpy
- Numpy for Statistical Operation

4. Introduction to Pandas

- Data Structures in Python
- Read in Data
- Read in CSV Data Using Pandas
- Read in Excel Data Using Pandas
- Read in JSON Data
- Read in XML Data
- Working with Databases using Python

5. Data Pre-Processing/Wrangling

- Removing NAs/No Values From Our Data
- Basic Data Handling: Starting with Conditional Data Selection

- Drop Column/Row
- Subset and Index Data
- Basic Data Grouping Based on Qualitative Attributes
- Cross tabulation
- Reshaping
- Pivoting
- Rank and Sort Data
- Concatenate
- Merging and Joining Data Frames

6. Introduction to Data Visualizations

- What is Data Visualization?
- Some Theoretical Principles Behind Data Visualization
- Histograms-Visualize the Distribution of Continuous Numerical Variables
- Boxplots-Visualize the Distribution of Continuous Numerical Variables
- Scatter Plot-Visualize the Relationship Between 2 Continuous Variables
- Barplot
- Pie Chart
- Line Chart

➢ Python ML Libraries – Introduction

- Pandas
- sci-kit learn
- numPy
- sciPy
- Seaborn, bokeh, plotly

➢ Real Time Project

➢ Resume Preparation Tips

➢ Interview Guidance and Support

➢ Job Support and Placement Assistance