Module 1: Excel & Visualizations

Excel Basics

- Worksheet Basics

- Entering Values and Formulas

- Data Formats

- Data Handling Basics - Cut, Copy and Paste

- Saving and Printing - Basics

- Basic Formula Operations

- Mathematical Formulas

- Textual Formulas

- Logical Formulas

- Date-Time Formulas

- Lookup Formulas (V Lookup, Hlookup,

- Index-Match)

- Data Tools

- Formatting Data and Tables

- Pivot Tables

Data Visualization with Excel

- Importance of Data Visualization

- Categories of Messages to be Conveyed

- Elements of Charts

- The Easy way of Creating Charts

- Bar and Column Charts

- Formatting Charts

- Line Charts

- Area Charts

- Pie and Doughnut Charts

- Why We Should Avoid Pie Charts

- Scatter Plot or XY Chart

- Frequency Distribution and Histograms

- Waterfall Charts

- Sparklines

- Heatmaps

- Pivot Charts

Analyze data using Excel

Module 2: SQL

Introduction
- Course Introduction

- Course Resources

- Installation of PostgreSQL

- Case Study

Fundamentals of SQL
- What is SQL

- Tables and DBMS

- Types of SQL Commands

- PostgreSQL

- CREATE

- INSERT

- Import Data from the File

- SELECT Statement

- SELECT DISTINCT

- WHERE

- Logical Operators

- UPDATE

- DELETE

- ALTER

- Restore and Back-up

- Debugging Restoration

- Creating DB using CSV Files

Basic Data Operations using SQL

- Module Introduction

- IN

- BETWEEN

- LIKE

- Commenting in SQL

- ORDER BY

- LIMIT

- AS

- COUNT

- SUM

- AVERAGE

- MIN MAX

- Group By

- HAVING

- CASE WHEN

Joins and Subqueries

- Module Introduction

- Joining Concepts

- Preparing Data for Joins

- Inner Join

- Left Join

- Right Join

- Full Outer Join

- Cross Join

- Intersect, Intersect all, Union all

- Except

- Union

- Subquery in WHERE clause

- Subquery in FROM clause

- Subquery in SELECT clause

- Views

- Index

Important SQL Functions

- Module Introduction

- LENGTH

- UPPER LOWER

- REPLACE

- TRIM LTRIM RTRIM

- CONCATENATION

- SUBSTRING

- LIST AGGREGATION

- CEIL FLOOR

- RANDOM

- SETSEED

- ROUND

- POWER

- CURRENT DATE TIME

- AGE

- EXTRACT

- Pattern Matching Basics

- Regular Expressions

- Converting Numbers or Dates to String

- Converting String to Numbers or Dates

- User Access Control

Joins and Subqueries

- Module Introduction

- Introduction to Window Functions

- Introduction to Row Number

- Implementing Row Number in SQL

- RANK and DENSERANK

- NTILE function

- AVERAGE function

- COUNT

- SUM TOTAL

- RUNNING TOTAL

- LAG and LEAD

Advanced Interview-Related and Performance TipsWhat is SQL

- Module Introduction

- Tablespace

- Primary key and Foreign Key

- ACID compliance

- Truncate

- Normalization

- First Normal Form

- Second Normal Form

- Third Normal Form and BCNF

- Fourth and Fifth Normal Form

Develop a project using SQL

Module 3: Tableau

Getting Started and Setting up Data Models

- What is Tableau and Why is it Used?

- Tableau Products

- Installing Tableau Desktop and Tableau Public

- About the Data

- Connecting to Data

- Live vs Extract

- Combining Data from Multiple Tables

- Relationships in Tableau

- Joins in Tableau

- Types of Joins

- Union in Tableau

- Physical Logical Layer and Data-models

Data Visualization Basics

- The Visualization Screen - Sheet

- Types of Data in Tableau

- Changing Data Type in Tableau

- Bar Chart

- Line Chart

- Scatter Plot

- Marks Card

- Adding Dimensions and Measures on

- Marks Card

- Customizing Line Chart using Marks Cards

- Adding Marks in the Scatterplot

- Text Tables, Heat Map and Highlight Tables

- Pie Charts

- Area Charts

- Creating Custom Hierarchies

- Tree map

- Dual Combination Charts

Customizing the Charts

- Creating Bins

- Histogram

- Grouping Data

- Filtering Data

- Dimension Filters

- Measure Filters

- Date-Time Filters

- Filter Options

- Types of Filters and Order of Operation

- Customizing Visual Filters

- Sorting Options

Data Visualization- Advanced

- How to Make a Map Chart

- Considerations Before Making Map Chart

- Marks Card for Customizing Maps

- Customizing Maps using Map Menu

- Layers in Map

- Visual Toolbar on Map

- Creating Territories

- Adding a Custom Background Image

- Blending and Custom Geocoding

- Calculated Fields

- Table Calculations

- Functions in Tableau

- Level of Detail Expressions

- Analytics Pane

- Parameters

- Sets

- Creating a Dashboard

- Customizing the Dashboard

- Adding SQL Data Source

- Adding a Cloud-based Data Source

Develop a project using Tableau

Module 4: Machine Learning with Python

Basics of Python

- Introduction to Python

- Installing Python and Anaconda

- Opening Jupyter Notebook

- Introduction to Jupyter

- Arithmetic Operators in Python: Python Basics

- Strings in Python: Python Basics

- Lists, Tuples, and Directories: Python Basics

- Working with Numpy Library of Python

- Working with Pandas Library of Python

- Working with Seaborn Library of Python

- Types of Data

- Types of Statistics

- Describing Data Graphically

- Measures of Centers

- Measures of Dispersion

- Introduction to Machine Learning

- Building a Machine Learning Model

Data Preprocessing

- Gathering Business Knowledge

- Data Exploration

- The Dataset and the Data Dictionary

- Importing Data in Python

- Univariate Analysis and EDD

- EDD in Python

- Outlier Treatment

- Outlier Treatment in Python

- Missing Value Imputation- Theory

- Missing Value Imputation in Python- Practical

- Seasonality in Data

- Bi-variate Analysis and Variable Transformation

- Variable Transformation and Deletion in Python

- Non-usable Variables

- Dummy Variable Creation: Handling Qualitative Data

- Dummy Variable Creation in Python

- Correlation Analysis

- Correlation Analysis in Python

Linear Regression

- Basic Equations and Ordinary Least Squares

- (OLS) Method

- Assessing the Accuracy of Predicted Coefficients

- Assessing Model Accuracy: RSE and R squared

- Simple Linear Regression in Python

- Multiple Linear Regression

- The F- statistic

- Interpreting Results of Categorical Variables

- Multiple Linear Regression in Python

- Test-train Split

- Bias Variance Trade-off

- Test Train Split in Python

- Regression Models other than OLS

- Subset Selection Techniques

- Shrinkage Methods: Ridge and Lasso

- Ridge Regression and Lasso in Python

- Heteroscedasticity

Logistic Regression, LDA and KNN

- Three Classifiers and the Problem Statement

- Why can't we use Linear Regression?

- Logistic Regression

- Training a Simple Logistic Model in Python

- Result of Simple Logistic Regression

- Logistics with Multiple Predictors

- Training Multiple Predictor Logistic Model in Python

- Confusion Matrix

- Creating Confusion Matrix in Python

- Evaluating the Performance of a

- Model- Theory

- Evaluating Model Performance in

- Python- Practical

- Linear Discriminant Analysis

- LDA in Python

- Test-Train Split

- Test-Train Split in Python

- K-Nearest Neighbors classifier

- K-Nearest Neighbors in Python

- Understanding the Results of

- Classification Models

- Summary of the Three Models

Decision Trees and Ensemble Techniques

- Basics of Decision Trees

- Understanding a Regression Tree

- The Stopping Criteria for Controlling Tree Growth

- The Data Set for this part

- Importing the Data Set into Python

- Missing Value Treatment in Python

- Dummy Variable Creation in Python

- Dependent- Independent Data Split in Python

- Test-Train Split in Python

- Creating Decision Tree in Python

- Evaluating Model Performance in Python

- Plotting Decision Tree in Python

- Pruning a Tree

- Pruning a Tree in Python

- Classification Tree

- The Data Set for the Classification Problem

- Classification Tree in Python : Preprocessing

- Classification Tree in Python : Training

- Advantages and Disadvantages of Decision Trees

- Ensemble Technique 1 - Bagging

- Ensemble Technique 1 - Bagging in Python

- Ensemble Technique 2 - Random Forests

- Ensemble Technique 2 - Random Forests in Python

- Using Grid Search in Python

- Boosting

- Ensemble Technique 3a - Boosting in Python

- Ensemble Technique 3b - AdaBoost in Python

- Ensemble Technique 3c - XGBoost in Python

Develop a prediction model using ML

Module 5: Capstone Project

Develop a project based on Custom problem statement & dataset