Excel and Python for Data Analysis and Visualization

Python is widely used in the corporate sector for performing data analysis. Python is one of the best languages used by data scientists for various data science projects/applications. It provides great functionality to deal with mathematics, statistics and scientific function.

Excel is also a very powerful data analysis tool used extensively by businesses for decision making. This course is designed to make students understand advanced features available in Excel and Python for data analysis. The course takes you from basic operations such as reading data into excel using various data formats, organizing, and manipulating data, to some of the more advanced functionality of Excel.

Learning Objectives:

- Gain hands-on experience in Python and Excel for data analysis.
- Data manipulation and analysis to solve business problems.
- Master the art of presenting data visually through a variety of visualization tools and techniques.

Learning Outcomes:

- Learn to leverage Python libraries such as Pandas and NumPy for efficient data handling and manipulation.
- Develop advanced skills in Excel, exploring features for data organization, analysis, and visualization.
- Harness the power of Excel functions and formulas to extract insights from complex datasets.
- Explore the entire data analysis process, from data cleaning and preprocessing to exploratory data analysis (EDA)
- Use industry-standard tools like Matplotlib and Seaborn to create compelling and informative data visualizations.

Module I- Introduction to Excel   **(2 Hours)**

Introduction to spreadsheets, basic spreadsheet operations and functions. In this session, you will learn the how to prepare well-formatted reports using sort / filter operations and advanced formatting techniques. Performing Basic Formatting, Conditional Formatting & Advanced Formatting

Module 2-Data Analysis in Excel – I   **(2 Hours)**

In this session you will learn to use some powerful Excel techniques used to understand and manipulate data. Excel offers a wide range of charts like line charts, bar charts, pie charts, scatter plots and boxplots.

The learning objectives of this session are:

- Use csv and delimited files to read data
- Bucketing & Segmenting
- Complex, Text Functions, logical formula in MS Excel. Conditional Formatting
- Create visualisations to analyse and present data / important insights

Module 3- Data Analysis in Excel – II   **(3 Hours)**

In this session, you will learn to use some powerful Excel techniques used to understand and manipulate data. The learning objectives of this session are:

- Handling of missing data, Data cleaning and Transformation using Excel

- Use Pivot Tables to analyse data
- Use VLOOKUP to combine data from multiple sources
- Working on Case study

Module 4 - Introduction to Pandas, NumPy, SciPy: **(9 Hours)**

Quick review of Pandas DataFrames, Numpy multidimensional arrays, and SciPy libraries to work with different datasets

Module 5 - Import and Export of Data: **(1 Hour)**

Installing, loading and using packages for importing and exporting data in Python

Module 6 - Data Preprocessing and Transformation: **(2 Hour)**

Handling of missing data, Data cleaning and Transformation in Python

Identify and Handling Duplicate and Missing Data: Find and remove duplicate rows and impute missing values. Performing data cleaning by identifying the data types, fixing rows and columns, standardizing values, fixing invalid values and filtering the data

Module 7 - Data Exploration   **(1 Hour)**

Exploring data using statistical methods: mean, median, mode, quantiles. Building contingency table. Basics of grouping data, merge, and pivot tables. Finding correlations among numeric fields.

Module 8 -Data Visualization

Visualizing data with the help of matplotlib and seaborn library. Performing univariate and bivariate analysis.

Scatter Plot, line plot, displot, histogram, boxplot, heatmaps, barplot, countplot ,word clouds, exporting plots as images. Creating Pairplots and jointplots

Working on a case study using Python   **(10 Hours)**