Course Title: Mastering Data Analytics: Raw Data To Decision

Duration: 4-6 Week
Who it is for: Pursuing B.Tech/B.E./MCA/M.Tech/BCA/M.Sc(IT)/B.Sc(IT)
Mode of Training: Live, Online, Instructor Led
Access to Recording: Yes
Pre-Requisite: No Pre-requisite, Concepts would be built from Scratch
Teaching Methodology: No to Minimum theory, maximum Hands on. Object oriented; problem driven.
Doubt Session: Weekly Doubt Session to ensure no one is left behind
Certificate: Industry Endorsed Certificate (Jointly Signed By Our Primary Industry Partner)

Project:
1) Module Specific 3-4 Mini-Project.
2) Hand On Major Project
3) Sample Project
4) Capstone Project

Course Objectives:
- Learn Python programming tailored for data analytics tasks.

- Perform Exploratory Data Analysis (EDA) using Python libraries.

- Gain a foundational understanding of SQL for extracting and manipulating data.

- Build a solid grasp of statistical methods essential for data analytics.

Teaching Approach:
- **Interactive Lectures**: For concept clarity.

- **Hands-On Labs**: Real World Simulated datasets and exercises.

- **Capstone Project**: Integration of all tools and concepts.

- **Feedback Sessions**: Instructors review and guide project presentations.

Course Structure:

Module 1: Python Programming for Data Analytics (12 Hours)

Learning Objectives:

- Understand the fundamentals of Python and apply it to data analytics tasks.

- Work with key Python libraries like Pandas and NumPy.

Content:

1. **Python Basics for Data Analytics** (6 Hours)

   o Python Syntax, Variables, and Data Types.

   o Control Structures: Loops, Conditionals, and Functions.

   o File Handling: Reading and Writing Files.

2. **Data Manipulation with Pandas** (4 Hours)

   o DataFrames: Creation, Indexing, and Selection.

   o Filtering and Aggregating Data.

   o Data Cleaning: Handling Missing Values, Duplicates, and Formatting.

3. **Numerical Operations with NumPy** (2 Hours)

   o Working with Arrays: Creation, Indexing, and Manipulation.

   o Basic Mathematical Operations on Arrays.

Mini-Project: Analyze a small sales dataset to clean and manipulate data using Pandas and NumPy.

Module 2: Exploratory Data Analysis (EDA) (6 Hours)

Learning Objectives:

- Perform detailed analysis of datasets to discover patterns and insights.

- Visualize data to identify trends and anomalies.

Content:

1. **Introduction to EDA** (1 Hours)

   o Understanding the Purpose of EDA in Data Analytics.

   o Key Steps in EDA: Summary Statistics, Outlier Detection, Correlations.

2. **Data Visualization with Matplotlib and Seaborn** (2 Hours)

   o Creating Line, Bar, Scatter, and Histogram Charts.

   o Advanced Visualizations: Heatmaps, Pairplots, Boxplots.

3. **Handling Complex Datasets** (1 Hours)

   o Managing Large Datasets and Reducing Dimensionality.

   o Combining Multiple Datasets for Analysis.

4. **EDA Workflow and Automation** (2 Hours)

   o Automating Common EDA Tasks with Python.

   o Exporting EDA Results for Presentations and Reports.

Mini-Project: Conduct a full EDA on a given dataset (e.g., customer churn data), visualizing key patterns and summarizing insights.

Module 3: SQL for Data Analytics (8 Hours)
Learning Objectives:

• Learn how to use SQL to retrieve, manipulate, and analyze data from databases.

• Understand the integration of SQL with Python.

Content:

1. **SQL Basics** (2 Hours)

   o Understanding Databases and Tables.

   o Writing Basic Queries: SELECT, WHERE, GROUP BY, and ORDER BY.

2. **Data Manipulation with SQL** (3 Hours)

   o Combining Tables: JOINs (INNER, LEFT, RIGHT, FULL).

   o Aggregating Data: COUNT, SUM, AVG, MAX, MIN.

   o Writing Subqueries for Advanced Analysis.

3. **SQL Integration with Python** (3 Hours)

   o Setting Up and Connecting to Databases using MySQL.

   o Executing SQL Queries in Python.

   o Retrieving Data from SQL Databases into Pandas Data Frames.

Mini-Project: Create and analyze a database of employee records to answer business-related questions using SQL and Python.

Module 4: Basics of Statistics for Data Analytics (4 Hours)

Learning Objectives:

- Build a foundational understanding of descriptive and inferential statistics.

- Apply statistical concepts to analyze and interpret datasets.

Content:

1. **Descriptive Statistics** (2 Hours)

    o   Measures of Central Tendency: Mean, Median, Mode.

    o   Measures of Variability: Range, Variance, Standard Deviation.

    o   Data Distribution: Normal Distribution and Skewness.

2. **Inferential Statistics Basics** (2 Hours)

    o   Sampling Techniques and Sample Size.

    o   Confidence Intervals.

    o   Hypothesis Testing Basics: p-Value and t-Test Overview.

Mini-Project: Analyze a dataset to calculate summary statistics and test hypotheses (e.g., comparing product sales across regions).

Capstone Project (10 Hours)

Objective:

Combine Python, SQL, EDA, and Statistics to solve a real-world data analytics problem.

Structure:

1. **Problem Statement**: Analyze sales and customer data to determine trends, predict future outcomes, and provide actionable insights.

2. **Workflow**:

    o   **Python**: Clean and prepare the dataset.

    o   **EDA**: Identify patterns and anomalies.

    o   **SQL**: Retrieve and manipulate data from a database.

    o   **Statistics**: Apply statistical techniques for hypothesis testing.

3. **Deliverables**:

   o   A detailed report summarizing insights and recommendations.

   o   Visualizations and a Python script demonstrating the full analysis workflow.