Certification on Data Analyst

Syllabus

Duration: 60 Hours

1. **Introduction**
   - **Why Python Programming**
   - **Course Overview**
   - **Data Types and Operators**
   - **Arithmetic Operators**
   - **Variables and Assignment Operators**
   - **Integers and Floats**
   - **Booleans, Comparison Operators, and Logical Operators**
   - **Strings**
   - **Type and Type Conversion**
   - **Lists and Membership Operators**
   - **List Methods, Tuples, Sets**
   - **Dictionaries and Identity Operators**
   - **Compound Data Structure**
   - **Discussion and Doubt Clear**
   - **Assignment and Test**

2. **Control Flow**
   - **Conditional statements**
   - **Iteration/looping statements**
   - **Break, Continue**
   - **Zip and Enumerate**
   - **Discussion and Doubt Clear**
   - **Assignment and Test**

3. **Function**
   - **Defining Functions, Variable Scope**
   - **Documentation, Lambda Expressions**
   - **Iterators and Generators**

4. **Scripting**
   - **Python Installation**
   - **Scripting with Raw Input**
   - **Errors and Exceptions**
   - **Reading and Writing Files, Importing Local Scripts**
   - **The Standard Library, Techniques for Importing Modules**
   - **Discussion and Doubt Clear with Assignment and Test**

5. NumPy
- **Introduction to NumPy, Why Use NumPy?**
- **Creating and Saving NumPy ndarrays**
- **Accessing, Deleting, and Inserting Elements Into ndarrays**
- **Slicing ndarrays**
- **Boolean Indexing, Set Operations, and Sorting**
- **Manipulating ndarrays**
- **Discussion and Doubt Clear with Assignment and test**

6. Pandas
- **Introduction to Pandas, Why Use Pandas?**
- **Creating Pandas Series and DataFrames, Accessing and Deleting Elements in Pandas Series, DataFrames**
- **Arithmetic Operations on Pandas Series**
- **Manipulate a Series, DataFrame**
- **Dealing with NaN**
- **Discussion and Doubt Clear with Assignment and Test**

7. Statistics
- **Probability**
- **Binomial Distribution, Conditional Probability , Bayes Rule**
- **Sampling Distributions and Central Limit Theorem**
- **Confidence Intervals**
- **Hypothesis Testing , T-Tests and A/B Tests**
- **Regression , Multiple Linear Regression , Logistic Regression**
- **Discussion and Doubt Clear with Assignment and Test**

8. Data Collection and Cleaning
- **Introduction to Data Collection**
- **Data Accessing**
- **Introduction to data cleaning**
- **Discussion and Doubt Clear with Assignment and Test**

9. Data Visualization
- **Introduction to Data Visualization**
- **Exploratory vs. Explanatory Analyses**
- **Visualization in Python**
- **What Makes a Bad Visual?**
- **Data Types (Continuous vs. Discrete)**
- **Identifying Data Types**
- **Chart Junk , Data Ink Ratio**
- **Bad Visual Quizzes ¸Using Color**
- **Designing for Color Blindness**

- **Tidy Data , Bar Charts**
- **Absolute vs. Relative Frequency, Pie Charts, Histograms**
- **Descriptive Statistics**
- **Outliers and Axis Limits**
- **Scales and Transformations**
- **Bivariate Exploration of Data - Scatterplots and Correlation, Overplotting, Transparency, and Jitter , Heat Maps , Violin Plots, Box Plots, Clustered Bar Charts, Faceting, Line Plots**
- **Multivariate Exploration of Data -Non-Positional Encodings for Third Variables, Color Palettes, Faceting in Two Directions, Adaptations of Bivariate Plots, Plot Matrices, Feature Engineering**
- **Explanatory Visualizations - Revisiting the Data Analysis Process, Polishing Plots, Creating a Slide Deck with Jupyter**
- **Discussion and Doubt Clear**
- **Assignment and Test**

10. Git & Github

- **What is Version Control?**
- **Create A Git Repo**
- **Review a Repo's History**
- **Add Commits To A Repo**
- **Tagging, Branching, and Merging**
- **Undoing Changes**
- **Working On Another Developer's Repository**
- **Staying In Sync With A Remote Repository**

11. Project Work

12. Placement Assistance Sessions

- **Mock Interviews**
- **GDs**
- **Preplacement Talks**
- **Industry Exposer Sessions**

13. Internship

*In Assignment there is Mini Projects.