ocean academy

Career Advancement Program in

Data Analytics

Data Analyst - syllabus

Python Programming:

- Introduction
- Conditional statements
- Looping
- Control statements
- String manipulation
- Lists
- Tuple
- Dictionary
- Functions
- Modules
- File Handling
- Exceptional Handling
- Git & GitHub

Data Manipulation using numpy:

- Introduction to NumPy
- Creation of Array & Array Operations
- Numpy Methods & Attributes
- Array indexing and slicing
- Array Looping
- Arange & Reshape
- Random module
- Arithmetic & Matrix Operations using array
- Array Statistical Functions
- Array Mathematical Functions
- Array Creation Functions – identity, ones, zeros, full
- Copy and View
- Boolean operations and Filtering
- Array Sorting
- Join, Split, Stack

Data Manipulation using Pandas:

- Introduction to pandas
- Pandas Data Structures
- Series Operations
- DataFrame Operations
- Indexing, loc and iloc
- Data Loading and Summary Statistics
- Datatype conversion
- Missing Value Treatment
- Data Manipulation
  - Add/remove – (.assign(), .drop())
  - Renaming columns and Indexes – (.rename())
  - Sorting (.sort_values() & sort_index())
  - Filtering Data
- Creating New Columns
- Outliers, Skewness and Kurtosis
- Merging, Joining, Concatenating Data
- Group by, Reshaping and Pivot Tables
- Data Visualization with pandas

Data visualization using Matplotlib:

- Introduction to Matplotlib
- Basic Statistical Terms and Correlation
- Difference between Univariate, Bi-variate, Multivariate
- Setting Graph figure size
- Creating Basic Plots
  - Line Plot
  - Scatter Plot
  - Bar Plot
  - Histogram Plot
  - Pie Chart
- Customizing Plots
  - Titles and Labels
  - Legend
  - Ticks and grids
  - Lines and Markers
- Subplots

Data visualization using Seaborn:

- Introduction to Seaborn
- Import Library
- Univariate:
    - Displot (Distribution Plot)
    - Count Plot
    - Box Plot
- Bivariate:
    - Strip Plot
    - Swarm Plot
    - Violin Plot
    - Joint Plot
- Multivariate:
    - Pairplot
    - Heatmap

Exploratory Data Analysis:

- Introduction to EDA
- Loading Data
- Data Inspection
- Descriptive Statistics
- Handling Missing Values
- Data Types and Conversion
- Feature Engineering
- Data Visualization
    - Univariate, Bi-variate, multivariate
- Outlier Deduction and Treatment
- Scaling and Transformation
- Key Finding and Business Insights

Mathematics and stats:

- Introduction to Statistics
- Descriptive Statistics
- Inferential Statistics
- Probability & Distributions
- Bayes Theorem
- Normal Distribution
- Basics of Hypothesis Testing
- Scipy Library

SQL

- Introduction to DBMS and SQL
- Installation of MySQL
- SQL Commands
- DDL Commands
- DML Commands
- Functions
- Joins
- Sub-queries
- Windows function

Data Structure and Algorithms:

- Introduction to data structure
- Array and Lists
- Dictionaries and Hash Tables
- Sets
- Stacks and Queues
- Tree, Graphs and Hash tables
- Sorting
- Searching

Excel Basic to Advanced:

- Introduction to MS-Excel
- Data Import and Export
- Tabs and Ribbons
- Add, Edit and Delete Data
- Formatting
- Data Cleaning
- Formulas and Functions
- Pivot Table
- Charts and Graphs
- Keyboard Shortcuts
- Automation and Macros
- Dashboard Building