TECHNOCATION FREELANCING TRAINING INSTITUTE & SOFTWARE HOUSE

Professional Pandas Course Outline

Module 1: Introduction to Pandas

- What is Pandas? Why is it Important?
- Installing and Setting Up Pandas (`pip install pandas`)
- Understanding Pandas Data Structures:
  - Series (`pd.Series`)
  - DataFrames (`pd.DataFrame`)
- Loading Data into Pandas (`read_csv()`, `read_excel()`, `read_json()`)

Module 2: Data Manipulation with Pandas

- Indexing, Selecting, and Slicing Data (`loc[]`, `iloc[]`)
- Adding and Removing Columns
- Filtering Data Using Conditions
- Sorting and Ranking Data (`sort_values()`, `rank()`)

Module 3: Data Cleaning and Preprocessing

- Handling Missing Data (`isnull()`, `dropna()`, `fillna()`)
- Removing Duplicates (`drop_duplicates()`)
- Handling Incorrect Data Types (`astype()`, `to_datetime()`)
- Applying String Operations on Data (`str.lower()`, `str.contains()`)

Module 4: Data Aggregation and Grouping

- Grouping Data with `groupby()`
- Aggregation Functions (`sum()`, `mean()`, `count()`, `min()`, `max()`)
- Pivot Tables and Crosstabs (`pivot_table()`, `crosstab()`)
- Multi-Indexing and Hierarchical Indexing

Module 5: Merging and Joining Datasets

- Concatenating DataFrames (`concat()`)
- Merging Data (`merge()`)
- Joining Data (`join()`)
- Handling Duplicates and Overlapping Data

Module 6: Data Visualization with Pandas

- Introduction to Pandas Built-in Visualization (`plot()`)
- Line Plots, Bar Charts, Histograms, and Scatter Plots
- Customizing Plots (Titles, Labels, Legends)
- Integrating Pandas with Matplotlib and Seaborn

Module 7: Time Series Analysis with Pandas

- Working with DateTime Indexes
- Resampling and Frequency Conversion (`resample()`)
- Rolling and Expanding Window Functions
- Time Zone Handling (`tz_localize()`, `tz_convert()`)

Module 8: Advanced Pandas Operations

- Using `apply()`, `map()`, and `applymap()` for Custom Functions
- Working with Large Datasets and Optimizing Performance
- Working with JSON, SQL, and APIs
- Exporting Data (`to_csv()`, `to_excel()`, `to_json()`, `to_sql()`)

Final Module: Capstone Project & Certification

- Implementing Real-World Data Analysis Tasks Using Pandas
- Cleaning and Analyzing Large Datasets
- Building a Data Pipeline with Pandas
- Final Presentation and Course Completion Certification