Pandas

Agenda

Key Takeaways-

- **What** is **Pandas**?

- **Advantages** of using **Pandas** library

- **Pandas Series** and multiple **ways** to create them

- **DataFrames** and various **data manipulation** operations

- **Merging,** **Joining** and **Concatenation** of DataFrames

- **GroupBy** operation on **DataFrames**

- Handling **Missing** data

Pandas

- **Pandas** is a fast, powerful, flexible and easy to use open source data analysis and manipulation tool, built on top of the Python programming language.

- **Pandas** has a rich set of functions and features to analyse and understand the data better.

- **Pandas** allows to customize, subset, edit and pivot the data, in order to bring most value from the data as needed.

- **Pandas** has two main data structures for data storage - Series and DataFrame.

Advantages of Pandas library

- Allows data import and export to various formats such as CSV, Excel, SQL, JSON etc.

- Allows various data manipulation operations such as merging, reshaping, sorting, subsetting, selecting etc.

- Provides multiple convenient methods for data filtering and selection.

- Suitable for all kinds of data cleaning/preprocessing steps and techniques.

- Better performance through optimized and efficient built-in functions.

- Built-in data visualization capabilities

- Built-in Time Series functionality - Date range generation, date shifting, lagging, moving window statistics etc.