

Welcome to **INTERNSHIP STUDIO**

Module 04 | Lesson 03

Pandas

Introduction to Pandas

What is Pandas?

- Pandas is an open-source data manipulation and analysis library in Python.
- It provides data structures and functions to efficiently handle and analyze structured data.
- Pandas is built on top of NumPy, another popular Python library for numerical computations.



Key Features of Pandas

1.Data Structures:

1. Series: A one-dimensional array-like object with labeled indices.
2. DataFrame: A two-dimensional table with labeled rows and columns, similar to a spreadsheet.

2.Data Manipulation:

1. Pandas offers powerful tools for data cleaning, transformation, and merging.
2. It supports operations like filtering, sorting, grouping, and aggregating data.

3.Missing Data Handling:

1. Pandas provides methods to handle missing or incomplete data, allowing for efficient data cleaning.

4.Data I/O:

1. Pandas supports reading and writing data in various formats, including CSV, Excel, SQL databases, and more.

How is Pandas used in Data Analytics

1.Data Exploration:

1. Pandas helps in exploring and understanding the data through descriptive statistics, data visualization, and summarization.

2.Data Cleaning and Pre-processing:

1. Pandas facilitates data cleaning by handling missing values, removing duplicates, and transforming data into usable formats.

3.Data Transformation:

1. Pandas enables data transformation by applying functions, mapping values, pivoting, and reshaping data for analysis.

4.Data Analysis:

1. Pandas provides efficient tools for slicing, filtering, grouping, and aggregating data, enabling in-depth analysis.

5.Data Visualization:

1. Pandas integrates with other libraries like Matplotlib and Seaborn to create informative plots and visualizations.

Installation and importing Pandas

- **Installation**: Pandas can be installed using pip, the Python package manager.

```
pip install pandas
```

- **Importing**:

```
import pandas as pd
```

SUMMARY

You got

this

- Pandas is a powerful library for data manipulation and analysis in Python.
- It offers versatile data structures and functions for data exploration, cleaning, transformation, and analysis.
- Pandas is widely used in data analytics, research, finance, and other domains for its ease of use and efficiency.

Next

session

Data Structures in Pandas