**Day 6 (12-06-25)**

**What is Pandas?**

**Pandas** is a **Python library** used for data analysis and manipulation. It provides **easy-to-use structures** like:

* Series (1D)
* DataFrame (2D – like a table)

It is widely used in **data science**, **machine learning**, **finance**, and many more domains.

**✅ Key Features of Pandas**

* Fast and efficient data manipulation
* Handling of missing data
* Powerful group-by functionality
* Data filtering, sorting, reshaping
* Merging and joining datasets
* Reading/writing CSV, Excel, JSON, etc.

**🧱 Core Structures in Pandas**

**1. Series (1D)**

import pandas as pd

data = [10, 20, 30]

s = pd.Series(data)

print(s)

**2. DataFrame (2D Table)**

import pandas as pd

data = {

'Name': ['Alice', 'Bob', 'Charlie'],

'Marks': [85, 90, 95]

}

df = pd.DataFrame(data)

print(df)

**🔧 Common Pandas Operations**

**🔹 Load CSV**

df = pd.read\_csv('filename.csv')

**🔹 View Data**

df.head() # First 5 rows

df.tail() # Last 5 rows

df.shape # Rows x Columns

df.columns # Column names

**🔹 Accessing Data**

df['Name'] # Single column

df[['Name', 'Marks']] # Multiple columns

df.loc[0] # Row by label

df.iloc[0] # Row by index

**🔹 Filtering Rows**

df[df['Marks'] > 90]

**🔹 Sorting**

df.sort\_values(by='Marks', ascending=False)

**🔹 Grouping**

df.groupby('Name')['Marks'].mean()

**🔹 Add/Update Column**

df['Grade'] = ['A', 'A+', 'A']

**🔹 Drop Column**

df.drop('Grade', axis=1, inplace=True)

**🔹 Handle Missing Data**

df.dropna() # Remove rows with nulls

df.fillna(0) # Replace nulls with 0

**📘 Use Cases of Pandas**

| **Domain** | **Use Case Example** |
| --- | --- |
| 📊 Data Analysis | Analyzing sales, user data, trends |
| 📉 Finance | Stock prices, risk modeling |
| 🏥 Healthcare | Patient data, statistics |
| 🛒 Retail | Sales forecasting, customer segmentation |
| 🤖 ML/AI | Preprocessing data for training models |

**🔁 File Formats Supported**

| **File Type** | **Function** |
| --- | --- |
| CSV | read\_csv() |
| Excel | read\_excel() |
| JSON | read\_json() |
| SQL | read\_sql() |

**Why Use Pandas?**

Pandas makes handling structured data **easy, powerful, and fast**, compared to regular Python lists and dictionaries.