# 📊 Sorting DataFrames in Pandas

### 🔹 1. Sorting by Index

Use **sort\_index()**

* Sorts rows/columns based on index labels.

**import pandas as pd**

**df = pd.DataFrame({"Name": ["Charlie", "Alice", "Bob"], "Age": [23, 25, 22]},**

**index=[3, 1, 2])**

**print(df.sort\_index()) # Sort rows by index**

**print(df.sort\_index(axis=1)) # Sort columns by labels**

### 🔹 2. Sorting by Column Values

Use **sort\_values()**

* Sorts rows based on values in one/more columns.

**print(df.sort\_values(by="Age")) # Sort ascending**

**print(df.sort\_values(by="Name", ascending=False)) # Descending**

# 📊 Joins in Pandas

### 🔹 1. What is a Join?

* **Join = combine two DataFrames** based on a common column/index.
* Similar to SQL joins.
* Done using pd.merge() or .join().

### 🔹 2. Types of Joins

#### (a) Inner Join → Keeps only matching rows.

**import pandas as pd**

**df1 = pd.DataFrame({"ID": [1, 2, 3], "Name": ["Alice", "Bob", "Charlie"]})**

**df2 = pd.DataFrame({"ID": [2, 3, 4], "Salary": [50000, 60000, 70000]})**

**print(pd.merge(df1, df2, on="ID", how="inner"))**

✅ Output:

**ID Name Salary**

**0 2 Bob 50000**

**1 3 Charlie 60000**

#### (b) Left Join → All rows from left + matching from right.

**print(pd.merge(df1, df2, on="ID", how="left"))**

✅ Output:

**ID Name Salary**

**0 1 Alice NaN**

**1 2 Bob 50000**

**2 3 Charlie 60000**

#### (c) Right Join → All rows from right + matching from left.

**print(pd.merge(df1, df2, on="ID", how="right"))**

✅ Output:

**ID Name Salary**

**0 2 Bob 50000**

**1 3 Charlie 60000**

**2 4 NaN 70000**

#### (d) Outer Join → All rows from both, fill NaN for missing.

**print(pd.merge(df1, df2, on="ID", how="outer"))**

✅ Output:

**ID Name Salary**

**0 1 Alice NaN**

**1 2 Bob 50000**

**2 3 Charlie 60000**

**3 4 NaN 70000**

### 🔹 3. Joining on Multiple Keys

**df3 = pd.DataFrame({"ID": [1, 2], "Dept": ["HR", "IT"], "Salary": [40000, 50000]})**

**print(pd.merge(df1, df3, on=["ID"], how="inner"))**

### 🔹 4. Using .join() (Index-based join)

**df1 = pd.DataFrame({"Name": ["Alice", "Bob"]}, index=[1, 2])**

**df2 = pd.DataFrame({"Salary": [50000, 60000]}, index=[1, 2])**

**print(df1.join(df2))**

✅ **Summary**

* inner → only common rows.
* left → all left + matching right.
* right → all right + matching left.
* outer → all rows from both.
* Use **merge()** for column-based joins, **join()** for index-based joins.