

Pandas Basics

1. What is Pandas?

Pandas is a Python library used for data analysis and data manipulation. It helps in handling structured data easily.

Example:

```
import pandas as pd
```

2. What is a DataFrame?

A DataFrame is a 2-dimensional table-like data structure with rows and columns in Pandas.

Example:

```
import pandas as pd

data = {
    "Name": ["Anu", "Rahul"],
    "Age": [20, 22]
}
df = pd.DataFrame(data)
print(df)
```

3. Difference between Series and DataFrame.

Series	DataFrame
One-dimensional	Two-dimensional
Single column	Multiple columns
Like a list	Like a table

Example:

Series (1D)

```
import pandas as pd
s = pd.Series([10, 20, 30])
print(s)
```

DataFrame (2D)

```
df = pd.DataFrame({
    "Marks": [80, 85, 90],
    "Grade": ["A", "A+", "A"]
})
print(df)
```

4. What is CSV?

CSV stands for Comma Separated Values.
It is a file format used to store tabular data in plain text.

Example CSV file (data.csv):

```
Name, Age
Anu, 20
Rahul, 22
```

Reading CSV using Pandas

```
import pandas as pd
df = pd.read_csv("data.csv")
print(df)
```

5. Why is Pandas important for AI?

Pandas is important for AI because it:

- Handles large datasets
- Cleans and preprocesses data
- Makes data ready for machine learning models

Example: Data Cleaning

```
import pandas as pd
```

```
df = pd.DataFrame({  
    "Age": [20, None, 22]  
})
```

```
df["Age"].fillna(df["Age"].mean(), inplace=True)  
print(df)
```