

# 1. Creating DataFrames & Series

## **pd.DataFrame()** - Create a DataFrame

```
python
CopyEdit
import pandas as pd
df = pd.DataFrame({'A': [1, 2, 3], 'B': [4, 5, 6]})
print(df)
```

## **pd.Series()** - Create a Series

```
python
CopyEdit
s = pd.Series([10, 20, 30], name="Numbers")
print(s)
```

---

# 2. Viewing & Inspecting Data

## **df.head(n)** - View First n Rows

```
python
CopyEdit
df.head(3)  # Shows first 3 rows
```

## **df.tail(n)** - View Last n Rows

```
python
CopyEdit
df.tail(2)  # Shows last 2 rows
```

## **df.info()** - Get DataFrame Information

```
python
CopyEdit
df.info()
```

## **df.describe()** - Summary Statistics

```
python
```

CopyEdit

```
df.describe() # Shows mean, count, min, max, etc.
```

### **df.shape - Get Dimensions of DataFrame**

python

CopyEdit

```
rows, cols = df.shape # Returns (rows, columns)
```

### **df.columns - Get Column Names**

python

CopyEdit

```
print(df.columns)
```

### **df.dtypes - Get Column Data Types**

python

CopyEdit

```
print(df.dtypes)
```

---

## **3. Selecting & Filtering Data**

### **df['column\_name'] - Select a Single Column**

python

CopyEdit

```
df['A'] # Returns a Series
```

### **df[['col1', 'col2']] - Select Multiple Columns**

python

CopyEdit

```
df[['A', 'B']]
```

### **df.iloc[] - Select by Index Position**

python

CopyEdit

```
df.iloc[0, 1] # First row, second column
```

```
df.iloc[1:3] # Rows 1 to 2
```

### **df.loc[ ] - Select by Label**

python

CopyEdit

```
df.loc[0, 'A'] # First row, 'A' column  
df.loc[:, ['A', 'B']] # All rows, columns 'A' and 'B'
```

### **df[df['col'] > value] - Filter Rows**

python

CopyEdit

```
df[df['A'] > 2] # Select rows where 'A' > 2
```

---

## **4. Adding, Modifying & Removing Columns**

### **df['new\_col'] = values - Add Column**

python

CopyEdit

```
df['C'] = [7, 8, 9]
```

### **df.insert(position, column\_name, values) - Insert Column at Specific Position**

python

CopyEdit

```
df.insert(1, 'D', [10, 20, 30])
```

### **df.drop(columns=['col\_name']) - Remove Column**

python

CopyEdit

```
df.drop(columns=['B'], inplace=True) # Removes 'B' column
```

### **df.rename(columns={'old\_name': 'new\_name'}) - Rename Columns**

python

CopyEdit

```
df.rename(columns={'A': 'Alpha'}, inplace=True)
```

---

## 5. Handling Missing Values

### **df.isnull().sum()** - Count Missing Values

python

CopyEdit

```
print(df.isnull().sum()) # Counts NaN in each column
```

### **df.fillna(value)** - Fill Missing Values

python

CopyEdit

```
df.fillna(0, inplace=True) # Replace NaN with 0
```

### **df.dropna()** - Drop Rows with Missing Values

python

CopyEdit

```
df.dropna(inplace=True)
```

---

## 6. Sorting & Ordering Data

### **df.sort\_values(by='col\_name')** - Sort by Column

python

CopyEdit

```
df.sort_values(by='A', ascending=False, inplace=True)
```

### **df.sort\_index()** - Sort by Index

python

CopyEdit

```
df.sort_index(inplace=True)
```

---

## 7. Grouping & Aggregating Data

### **`df.groupby('col').agg()` - Group by Column**

python

CopyEdit

```
df.groupby('A').agg({'B': 'sum'})
```

### **`df.value_counts()` - Count Unique Values**

python

CopyEdit

```
df['A'].value_counts()
```

### **`df.pivot(index, columns, values)` - Pivot Table**

python

CopyEdit

```
df.pivot(index='A', columns='B', values='C')
```

---

## **8. Merging & Joining DataFrames**

### **`pd.concat([df1, df2])` - Concatenate DataFrames**

python

CopyEdit

```
df_combined = pd.concat([df1, df2])
```

### **`df1.merge(df2, on='common_column')` - Merge DataFrames**

python

CopyEdit

```
df_merged = df1.merge(df2, on='ID', how='inner')
```

---

## **9. Working with Dates**

### **`pd.to_datetime(df['date_column'])` - Convert to Datetime**

python

CopyEdit

```
df['date'] = pd.to_datetime(df['date'])
```

**`df['date'].dt.year / .dt.month / .dt.day` - Extract Year, Month, Day**

python

CopyEdit

```
df['year'] = df['date'].dt.year
df['month'] = df['date'].dt.month
df['day'] = df['date'].dt.day
```

---

## 10. Exporting & Importing Data

**`df.to_csv('file.csv')` - Save as CSV**

python

CopyEdit

```
df.to_csv('output.csv', index=False)
```

**`df.to_excel('file.xlsx')` - Save as Excel**

python

CopyEdit

```
df.to_excel('output.xlsx', index=False)
```

**`pd.read_csv('file.csv')` - Load CSV**

python

CopyEdit

```
df = pd.read_csv('file.csv')
```

**`pd.read_excel('file.xlsx')` - Load Excel**

python

CopyEdit

```
df = pd.read_excel('file.xlsx')
```

---

## 11. Statistical Functions

**`df.mean()`, `df.median()`, `df.std()` - Compute Mean, Median, Standard Deviation**

```
python
CopyEdit
df.mean()    # Column-wise mean
df.median()  # Median
df.std()     # Standard deviation
```

### **df.corr()** - Correlation Matrix

```
python
CopyEdit
df.corr()
```

### **df.cumsum()** - Cumulative Sum

```
python
CopyEdit
df.cumsum()
```

---

## **12. Applying Functions**

### **df.apply(function, axis)** - Apply Custom Function

```
python
CopyEdit
df['A'] = df['A'].apply(lambda x: x * 2) # Double values in column
'A'
```

### **df.map(function)** - Apply Function to Series

```
python
CopyEdit
df['B'] = df['B'].map(lambda x: x + 1)
```

---

## **13. Creating Dummy Variables**

### **pd.get\_dummies(df['category\_col'])** - One-Hot Encoding

```
python
CopyEdit
pd.get_dummies(df['Category'])
```

---

## Summary

Category	Function
Creating Data	<code>pd.DataFrame()</code> , <code>pd.Series()</code>
Viewing Data	<code>df.head()</code> , <code>df.tail()</code> , <code>df.info()</code>
Selecting Data	<code>df.iloc[]</code> , <code>df.loc[]</code> , <code>df['col']</code>
Modifying Data	<code>df['new_col']</code> , <code>df.drop()</code> , <code>df.rename()</code>
Handling Missing Values	<code>df.isnull()</code> , <code>df.fillna()</code> , <code>df.dropna()</code>
Sorting & Filtering	<code>df.sort_values()</code> , <code>df[df['col'] &gt; value]</code>
Grouping & Aggregating	<code>df.groupby()</code> , <code>df.value_counts()</code>
Merging DataFrames	<code>pd.concat()</code> , <code>df.merge()</code>
Exporting & Importing	<code>df.to_csv()</code> , <code>pd.read_csv()</code>
Statistical Functions	<code>df.mean()</code> , <code>df.corr()</code> , <code>df.std()</code>

This covers **most Pandas functions** you'll need in **data analysis and machine learning**.

 Let me know if you need any specific explanations!