

# Python Programming

## Unit 06 – Lecture 03 Notes

### Pandas Basics (Series and DataFrame)

Tofik Ali

February 14, 2026

## Contents

<b>1</b>	<b>Lecture Overview</b>	<b>1</b>
<b>2</b>	<b>Setup (If Needed)</b>	<b>2</b>
<b>3</b>	<b>Core Concepts</b>	<b>2</b>
3.1	Series . . . . .	2
3.2	DataFrame . . . . .	2
3.3	Inspecting Data . . . . .	2
3.4	Column Selection . . . . .	2
3.5	Add/Delete Columns . . . . .	2
3.6	Boolean Filtering . . . . .	3
3.7	Iteration (Use Carefully) . . . . .	3
<b>4</b>	<b>Demo Walkthrough</b>	<b>3</b>
<b>5</b>	<b>Interactive Checkpoints (with Solutions)</b>	<b>3</b>
<b>6</b>	<b>Practice Exercises (with Solutions)</b>	<b>3</b>
<b>7</b>	<b>Exit Question (with Solution)</b>	<b>4</b>

## 1 Lecture Overview

Pandas is a powerful library for working with tabular data (CSV, Excel-like tables). It provides:

- **Series** (1D labeled data),
- **DataFrame** (2D table).

This lecture focuses on creating and manipulating these structures.

## 2 Setup (If Needed)

If Pandas is not installed:

```
pip install pandas
```

## 3 Core Concepts

### 3.1 Series

A Series is like a column with an index.

```
import pandas as pd

s = pd.Series([10, 20, 30], index=["a", "b", "c"])
print(s)
```

### 3.2 DataFrame

A DataFrame is a table of columns (each column can be considered a Series).

```
import pandas as pd

df = pd.DataFrame({
    "name": ["Asha", "Bilal", "Charu"],
    "marks": [88, 76, 91]
})
```

### 3.3 Inspecting Data

- `df.head()` shows first rows
- `df.tail()` shows last rows
- `df.info()` shows dtypes and missing data summary
- `df.describe()` gives statistics for numeric columns

### 3.4 Column Selection

```
marks = df["marks"]
```

### 3.5 Add/Delete Columns

```
df["cgpa"] = df["marks"] / 10
del df["cgpa"]
```

### 3.6 Boolean Filtering

```
passed = df[df["marks"] >= 50]
```

### 3.7 Iteration (Use Carefully)

Pandas is designed for vectorized operations. Iteration is slower but sometimes necessary. Common ways:

- `df.iterrows()`
- `df.itertuples()`

## 4 Demo Walkthrough

File: `demo/pandas_basics_demo.py`

Observe:

- creating a DataFrame,
- adding new derived columns,
- filtering using boolean expressions.

## 5 Interactive Checkpoints (with Solutions)

### Checkpoint 1 Solution

**Question:** difference between Series and DataFrame?

**Answer:** Series is 1D labeled data; DataFrame is 2D tabular data with columns.

### Checkpoint 2 Solution

**Question:** how to select "marks" column?

**Answer:** `df["marks"]`

## 6 Practice Exercises (with Solutions)

### Exercise 1: Create a DataFrame

**Task:** Create a DataFrame with columns `name` and `age`.

**Solution:**

```
import pandas as pd
df = pd.DataFrame({"name": ["A", "B"], "age": [18, 19]})
print(df)
```

## Exercise 2: Create cgpa Column

**Task:** Create cgpa from marks.

**Solution:**

```
df["cgpa"] = df["marks"] / 10
```

## 7 Exit Question (with Solution)

**Question:** one line to create cgpa from marks?

**Answer:** `df["cgpa"] = df["marks"] / 10`