Contents

1. PANDAS - INTRODUCTION

1. Pandas

- ✓ Pandas is an open-source Python Library.
- ✓ Pandas have powerful data structures.

2. Main usage

- ✓ Main usage of pandas is analysing the data.
- ✓ By using pandas we can do below things easily,
  - Data loading
  - Data preparation
  - Data manipulation
  - Data analysis & etc

3. Creator

- ✓ Pandas created by Wes McKinney in the year of 2008

4. Data Analysis

- ✓ It is a process of cleaning, transforming, and modeling data to discover useful information for business decision-making.
- ✓ The purpose of Data Analysis is to extract useful information from data and taking the decision based upon the data analysis.
- ✓ Data Analysis became very easy after pandas was introduced to this world.

5. Key Features

- ✓ It's very easy to load different file formats of data.
- ✓ Handling missing data.
- ✓ Reshaping the data.
- ✓ Grouping the data
- ✓ Indexing and label-based slicing.
- ✓ Easily insert and delete columns from data
- ✓ Operations like aggregation and transformations.
- ✓ High performance merging and joining of data and many of others boss.

6. Open source

- ✓ Pandas is an open source means it's free.

7. Environment setup/pandas installation

- ✓ Open command prompt and execute below command

```
pip install pandas
```

8. pip command in python

- ✓ pip stands for **p**ython **i**nstaller **p**ackage
- ✓ Pip is a package management system.
- ✓ It is used to install and manage software packages.
  - pip install package_name
  - pip install pandas

9. Check installed python library.

- ✓ We can check installed python library by using below command.

```
pip show pandas
```

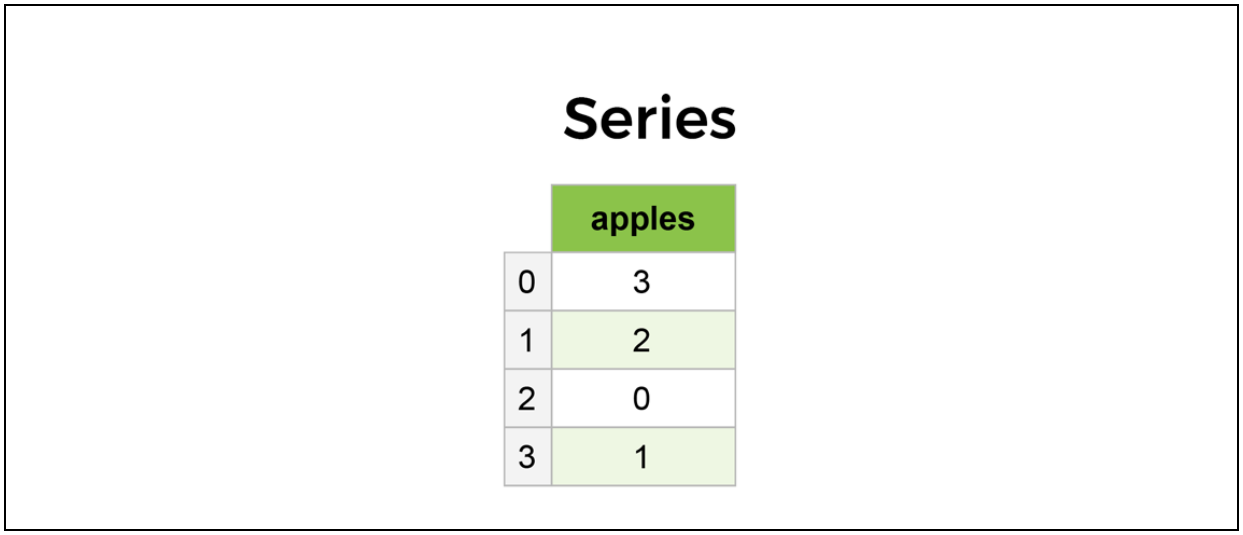
10. Pandas Data Structures

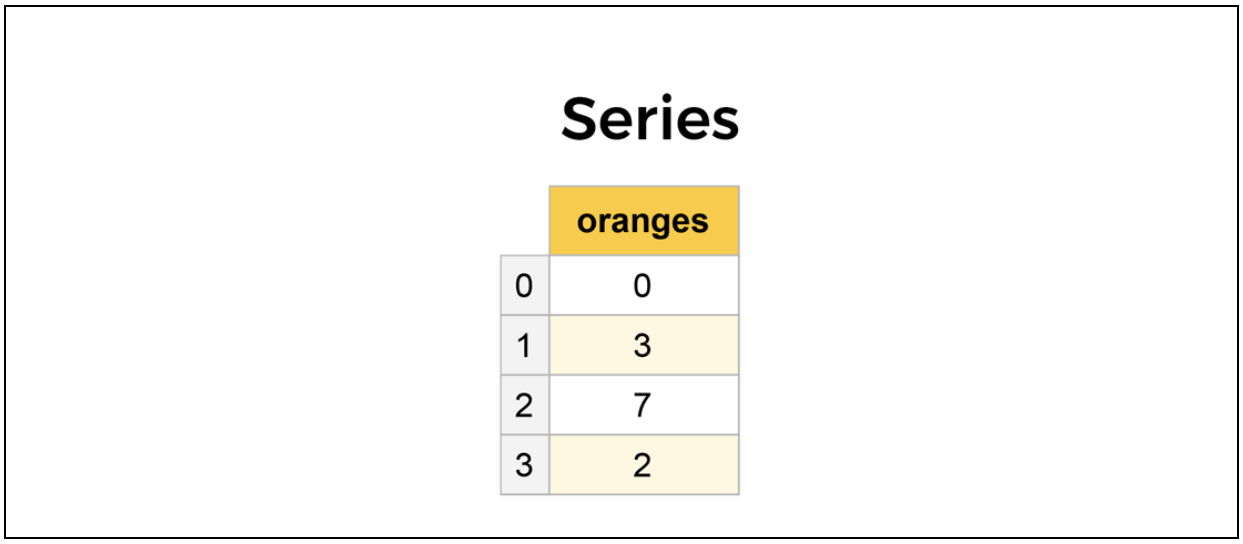
- ✓ One of the important features in pandas is data structures.
- ✓ There are mainly 3 types of data structures in pandas

Data Structure	Dimentionality	
1. Series	1D	Column
2. DataFrame	2D	Rows & Columns
3. Panel	3D	Group of Dataframes

10.1. Series

✓ A Series is similar to a single column of data.





10.2. DataFrame

✓ A DataFrame is similar to a table which contains rows and columns of data.

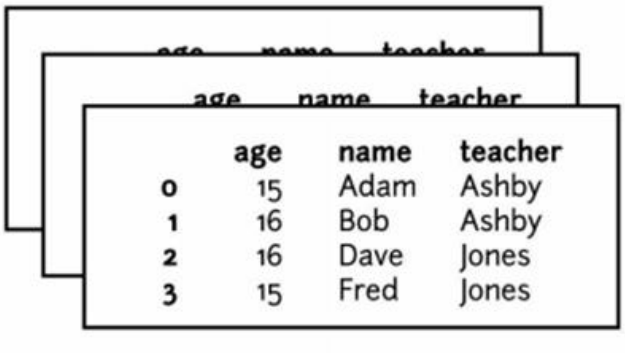
DataFrame

	apples	oranges
0	3	0
1	2	3
2	0	7
3	1	2

	Organization	Salesperson Name	Sales
0	Google	Sam	200
1	Google	Charlie	120
2	Salesforce	Ralph	125
3	Salesforce	Emily	250
4	Adobe	Rosalynn	150
5	Adobe	Chelsea	500

10.3. Panel

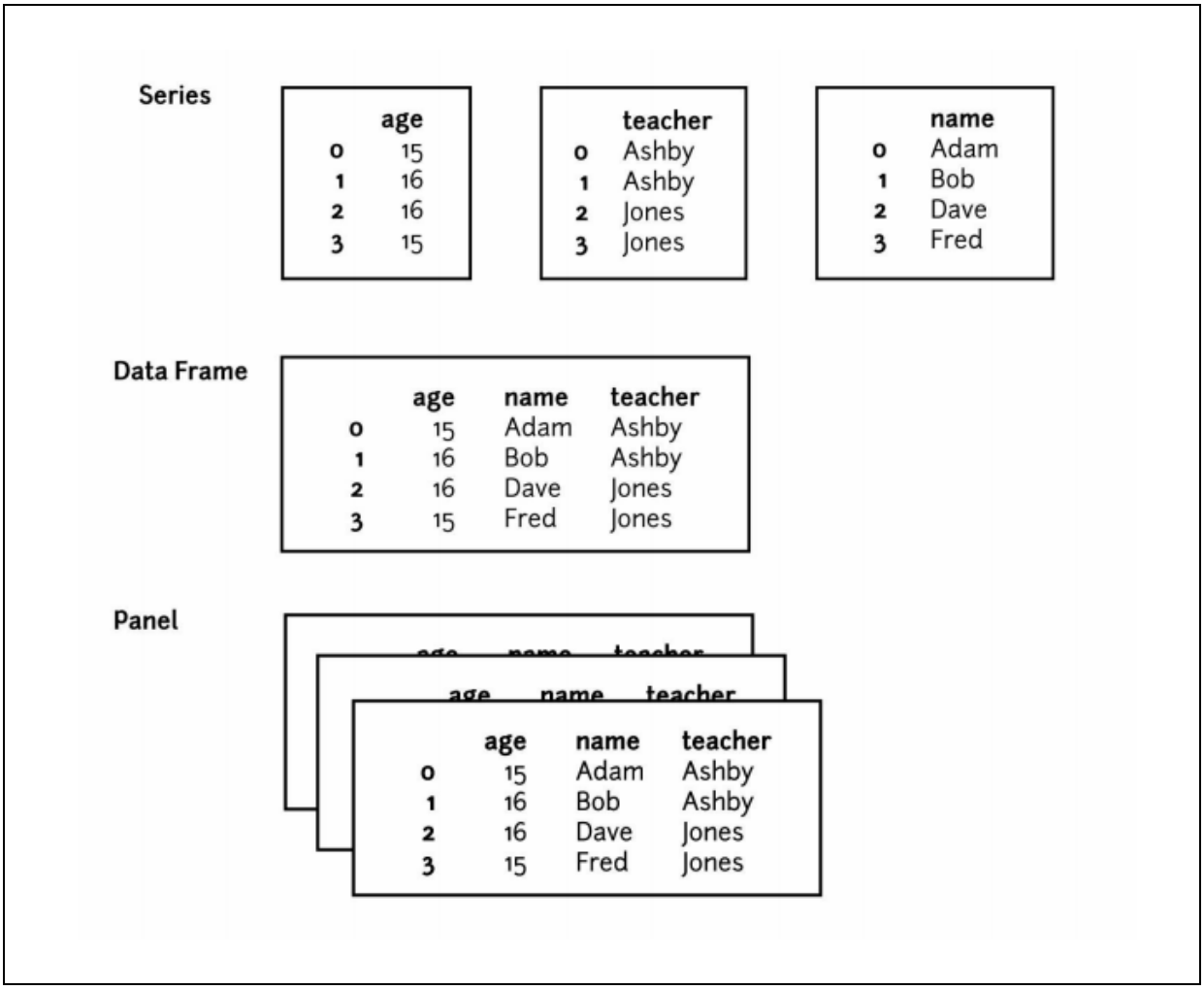
- ✓ A Panel is a group of DataFrames of data.



	age	name	teacher
0	15	Adam	Ashby
1	16	Bob	Ashby
2	16	Dave	Jones
3	15	Fred	Jones

Mostly used

- ✓ The most widely used data structures are the Series and the DataFrame.
- ✓ Series and DataFrames deal with array data and tabular data respectively



Relationship…

- ✓ Series having single column
- ✓ DataFrame can have multiple Series
- ✓ Panel has multiple DataFrames

11. Importing pandas

- ✓ We can import pandas package by using import keyword

Program Name: importing pandas
demo1.py

```
import pandas

print("pandas imported successfully")
```

Output

```
pandas imported successfully
```

12. Aliasing or renaming pandas

- ✓ We can alias or rename pandas name.
- ✓ As per python syntax we need to use as keyword to alias pandas name

12.1. Why aliasing Boss?

- ✓ Let me give one best example.
- ✓ My name is: Kasagani N Daniel, you can call by Daniel instead of full name for simplicity

12.2. Famous Alias name to pandas

- ✓ We can give alias name to pandas.
- ✓ Note this name can be any name but the famous alias name is **pd**

Program Name: aliasing pandas name as pd
demo2.py

```
import pandas as pd

print("pandas imported successfully")
print("pandas renamed to pd")
print("Now onwards we can use pd instead of pandas name")
```

Output

```
pandas imported successfully
pandas renamed to pd
Now onwards we can use pd instead of pandas name
```

12.3. Specialty of alias name as pd

- ✓ Here in our example we have given alias name as pd.
- ✓ We can provide any name as alias name.
- ✓ Real time developers habituated to give alias name as pd.
- ✓ So, its universal alias name