Contents

7. PANDAS – LOADING DIFFERENT FILES

1. Loading files

- ✓ We can load different files in pandas.
- ✓ Whenever we are loading the files then pandas return a DataFrame.
- ✓ Once after DataFrame is returned then based on requirement we can perform several operations.

2. Loading files

- ✓ Csv file
- ✓ Json file
- ✓ Excel file
- ✓ Tsv file
- ✓ Table from webpage

2.1. csv file

- ✓ Csv file stands for comma separated file.
- ✓ A csv file contains data like rows and columns.
- ✓ Whenever we are loading csv file then pandas returns DataFrame.
- ✓ We can load csv file by using **read_csv("files name with path")** function

Syntax

pd.**read_csv("file name with path")**

Program Name: Loading csv file
Input file: demo1.py
sales1.csv

```
import pandas as pd

df = pd.read_csv("sales1.csv")
print(df)
```

Output

```
     Order ID Customer Name                Product  Quantity
0      166837         Veeru  34in Ultrawide Monitor         2
1      166838         Tarun             Samsung m10         3
2      166839         Kedar            20in Monitor         1
3      166840       Lavanya               iPhone 11         3
4      166841          Venu     Macbook Pro Laptop         2
..        ...           ...                     ...       ...
595    167403        Balaji     Macbook Pro Laptop         1
596    167404       Lavanya        ThinkPad Laptop         1
597    167405          Venu          Flatscreen TV         1
598    167406        Siddhu            Samsung m20         2
599    167407         Tarun     LG Washing Machine         1

[600 rows x 4 columns]
```

Note

- ✓ Above program, sales1.csv file should exist in current directory otherwise we will get an error like **FileNotFoundError**
- ✓ Let's load file from the folder in this case we need to provide file name with folder.

Program Name	Loading csv file demo2.py
Input file	If file name not exist then we will get an error

```
import pandas as pd

df = pd.read_csv("sales123.csv")
print(df)
```

Output

```
FileNotFoundError: No such file or directory: 'sales123.csv'
```

Program Name: Loading csv file from **files** folder
Input file: demo3.py
files\sales1.csv

```
import pandas as pd

df = pd.read_csv('files\sales1.csv')
print(df)
```

Output

```
     Order ID Customer Name                Product  Quantity
0      166837         Veeru  34in Ultrawide Monitor         2
1      166838         Tarun            Samsung m10         3
2      166839         Kedar           20in Monitor         1
3      166840       Lavanya              iPhone 11         3
4      166841          Venu     Macbook Pro Laptop         2
..        ...           ...                    ...       ...
595    167403        Balaji     Macbook Pro Laptop         1
596    167404       Lavanya        ThinkPad Laptop         1
597    167405          Venu          Flatscreen TV         1
598    167406        Siddhu            Samsung m20         2
599    167407         Tarun     LG Washing Machine         1

[600 rows x 4 columns]
```

Program Name: Loading csv file from D drive

Input file: demo4.py

D:sales1.csv

```
import pandas as pd

df = pd.read_csv('D:\sales1.csv')
print(df)
```

Output

```
     Order ID Customer Name                Product  Quantity
0      166837         Veeru  34in Ultrawide Monitor         2
1      166838         Tarun            Samsung m10         3
2      166839         Kedar           20in Monitor         1
3      166840       Lavanya              iPhone 11         3
4      166841          Venu     Macbook Pro Laptop         2
..        ...           ...                    ...       ...
595    167403        Balaji     Macbook Pro Laptop         1
596    167404       Lavanya        ThinkPad Laptop         1
597    167405          Venu          Flatscreen TV         1
598    167406        Siddhu            Samsung m20         2
599    167407         Tarun     LG Washing Machine         1

[600 rows x 4 columns]
```

2.2. json file

- ✓ JSON stands for JavaScript Object Notation
- ✓ In json file data is like key-value pairs and arrays.
- ✓ It is commonly used for transmitting data in web applications
- ✓ Whenever we are loading json file then pandas returns DataFrame.
- ✓ We can load csv file by using **read_json("file name with path")** function

Syntax

pd.**read_json("files name with path")**

Program Name: Loading json file
Input file: demo5.py
sales1.json

```
import pandas as pd

df = pd.read_json("sales1.json")
print(df)
```

Output

```
     order_id     cust_name                product  quantity
0    16278939       Lavanya        ThinkPad Laptop         2
1    16278966         Kedar          Flatscreen TV         1
2    16278993   Jaya Chandra     Macbook Pro Laptop         2
3    16279020    Mallikarjun              iPhone 11         3
4    16279047        Shahid     LG Washing Machine         1
..        ...            ...                    ...       ...
495  16292304         Sagar              iPhone 7          1
496  16292331     Chaithanya            Samsung m20         1
497  16292358        Siddhu               LG Dryer          2
498  16292385        Siddhu   AA Batteries (4-pack)         1
499  16292412         Sagar              iPhone 11         2

[500 rows x 4 columns]
```

2.3. excel file

- ✓ Excel is a spreadsheet from Microsoft.
- ✓ It is using mainly to store business applications data
- ✓ Whenever we are loading excel file then pandas returns DataFrame.
- ✓ We can load csv file by using **read_excel("file name with path")** function

Run below command

```
pip install xlrd
```

Syntax

```
pd.read_excel("files name with path")
```

Program Name: Loading excel file
Input file: demo6.py
sales1.xlsx

```
import pandas as pd

df = pd.read_excel("sales1.xlsx")
print(df)
```

Output

```
     Order ID Custer Name                Product  Quantity
0      166837       Veeru  34in Ultrawide Monitor         2
1      166838       Tarun            Samsung m10         3
2      166839       Kedar           20in Monitor         1
3      166840     Lavanya              iPhone 11         3
4      166841        Venu     Macbook Pro Laptop         2
..        ...         ...                    ...       ...
595    167403      Balaji     Macbook Pro Laptop         1
596    167404     Lavanya        ThinkPad Laptop         1
597    167405        Venu          Flatscreen TV         1
598    167406      Siddhu            Samsung m20         2
599    167407       Tarun     LG Washing Machine         1

[600 rows x 4 columns]
```

2.4. TSV file

- ✓ TSV stands for **T**ab **S**eparated **F**ile
- ✓ Whenever we are loading tsv file then pandas returns DataFrame.
- ✓ We can load csv file by using **read_table("file name with path")** function

Syntax

pd.**read_table**(path and file name)

Program Name: Loading tsv file
Input file: demo7.py
sales1.tsv

```
import pandas as pd

df = pd.read_table("sales1.tsv")
print(df)
```

Output

```
     Order ID Custer Name                 Product  Quantity
0      166837       Veeru  34in Ultrawide Monitor         2
1      166838       Tarun             Samsung m10         3
2      166839       Kedar            20in Monitor         1
3      166840     Lavanya               iPhone 11         3
4      166841        Venu      Macbook Pro Laptop         2
..        ...         ...                     ...       ...
595    167403      Balaji      Macbook Pro Laptop         1
596    167404     Lavanya         ThinkPad Laptop         1
597    167405        Venu           Flatscreen TV         1
598    167406      Siddhu             Samsung m20         2
599    167407       Tarun      LG Washing Machine         1

[600 rows x 4 columns]
```

2.5. Table from webpage

- ✓ We can load table from webpage.
- ✓ Whenever we are loading table from webpage then it pandas returns DataFrame.
- ✓ We can load csv file by using **read_html("url")** function

Syntax

pd.**read_html("url")**

Program Name: Loading table from website
demo8.py

Input: Loading from web page

```
import pandas as pd

url = 'https://en.wikipedia.org/wiki/The_World%27s_Billionaires'
df_list = pd.read_html(url)

print(df_list[2])
```

Output

	No.	Name	Net worth (USD)	Age	Nationality	Source(s) of wealth
0	1	Elon Musk	$219 billion	50	United States	Tesla, SpaceX
1	2	Jeff Bezos	$177 billion	58	United States	Amazon
2	3	Bernard Arnault & family	$158 billion	73	France	LVMH
3	4	Bill Gates	$129 billion	66	United States	Microsoft
4	5	Warren Buffett	$118 billion	91	United States	Berkshire Hathaway
5	6	Larry Page	$111 billion	49	United States	Alphabet Inc.
6	7	Sergey Brin	$107 billion	48	United States	Alphabet Inc.
7	8	Larry Ellison	$106 billion	77	United States	Oracle Corporation
8	9	Steve Ballmer	$91.4 billion	66	United States	Microsoft
9	10	Mukesh Ambani	$90.7 billion	64	India	Reliance Industries