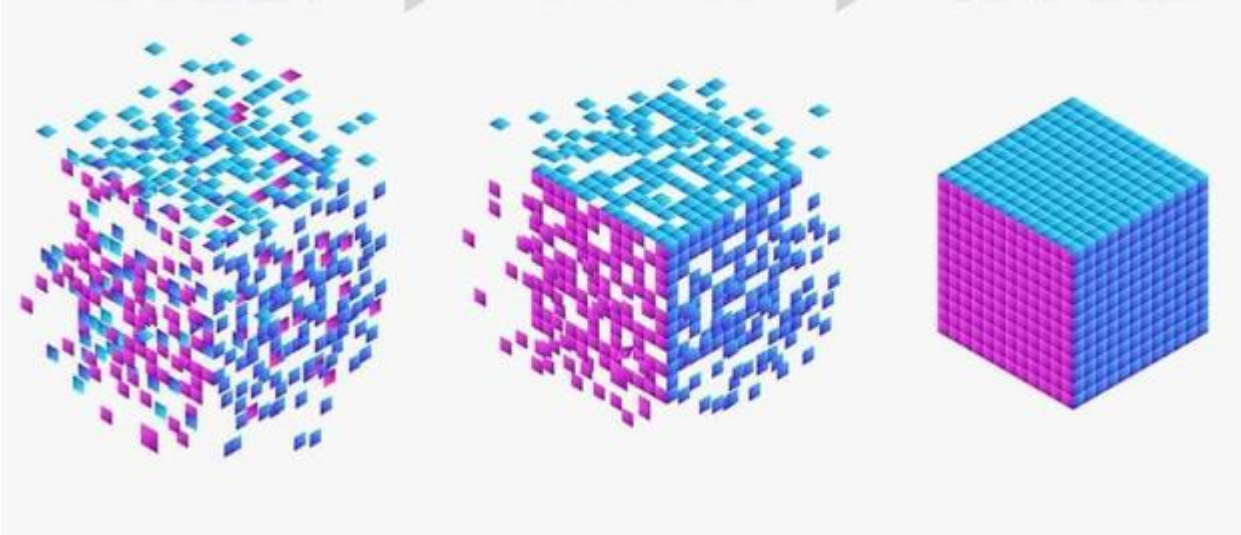
Pandemic Analytics

2.3

Working with Data in Python

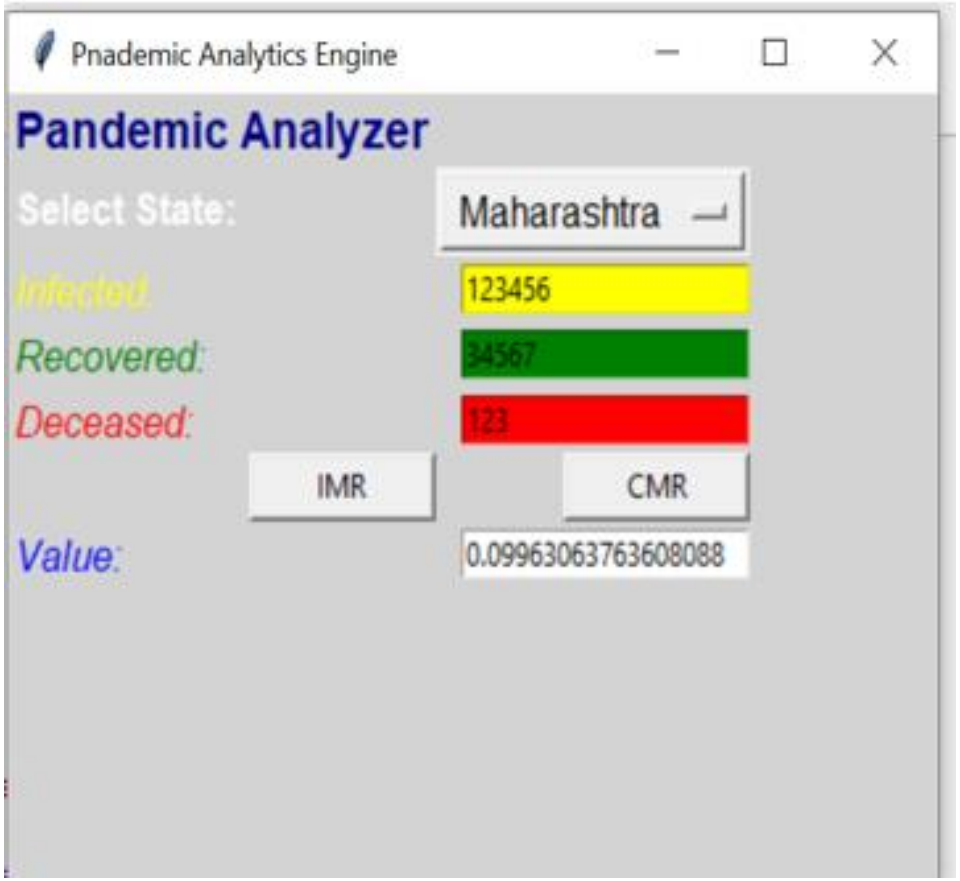


Day: **2**

Session: **3**

ETD: **40 min**

Recap


Pnademic Analytics Engine
Pandemic Analyzer
Select State:
Maharashtra
Infected:
123456
Recovered:
34567
Deceased:
123
IMR
CMR
Value:
0.09963063763608088

Reading files with open()

File Object

File object

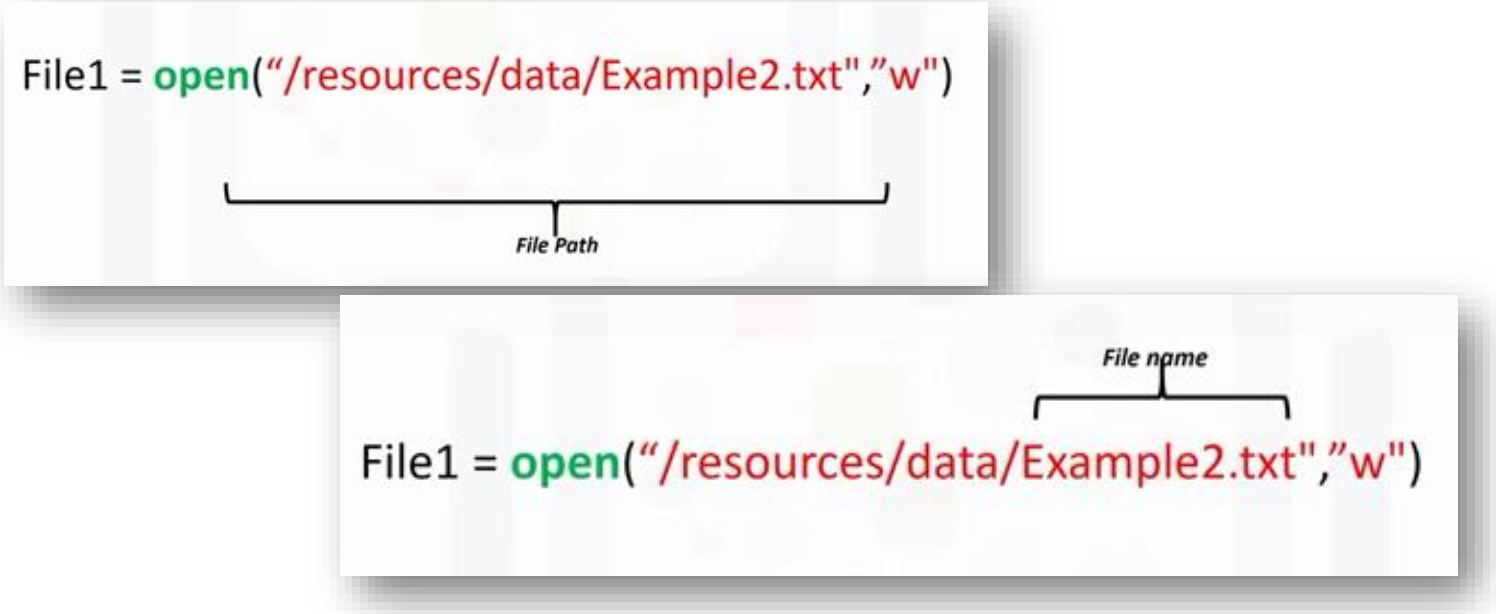
Reading files with open()

File Object

File object

.read()

Continue..



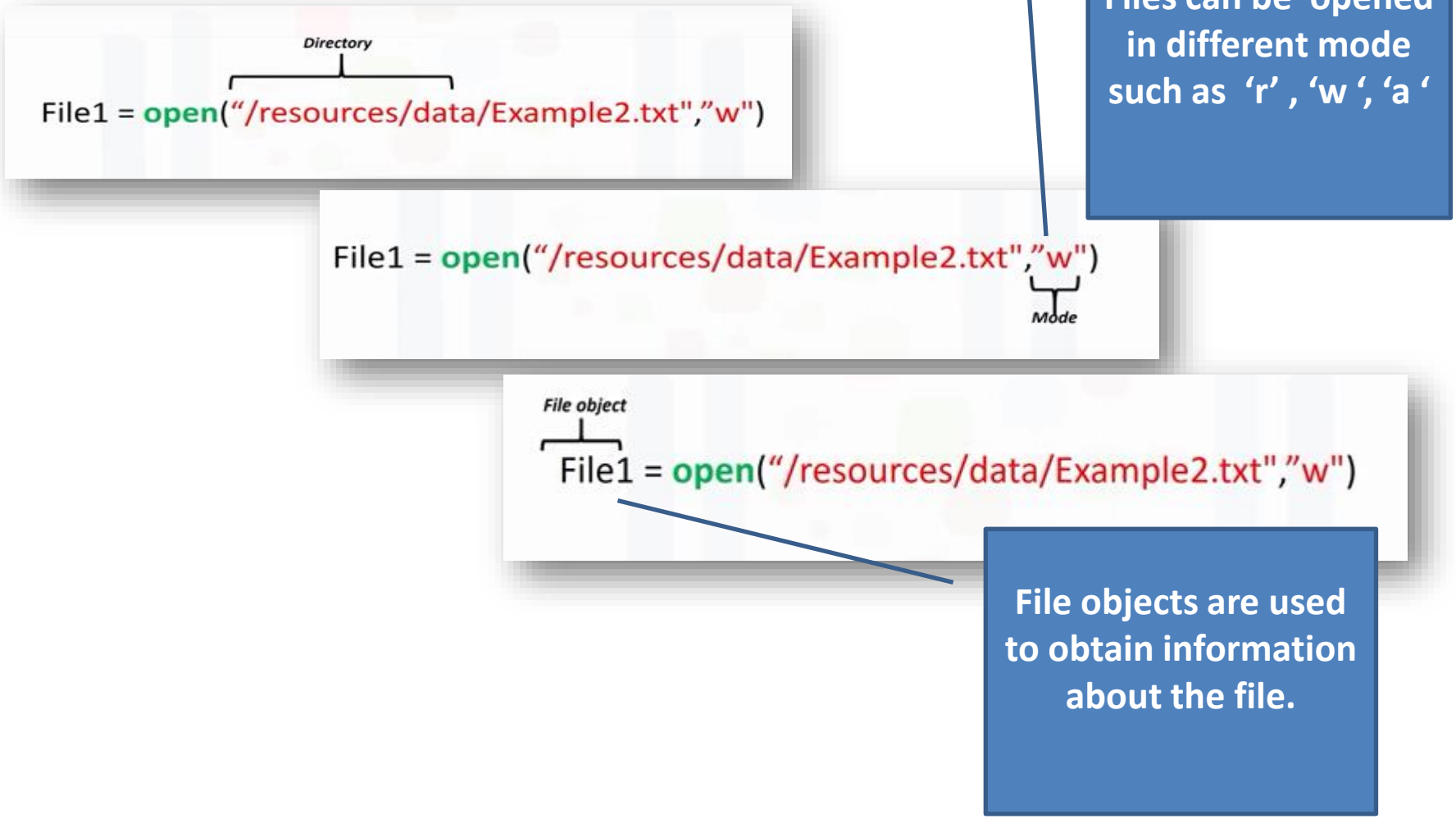
If the file resides in ROOT then no need to mention the path

File1 = open("myfile.txt")

If accessing through remote location then complete path is required

File1 = open("https://mygov.in/.../covid_19_india.csv"

Continue..

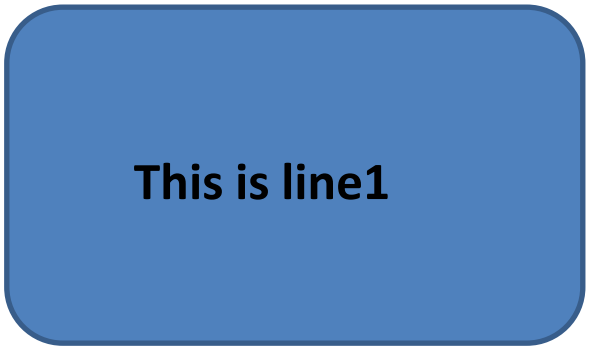

Directory
File1 = open("/resources/data/Example2.txt","w")
File1 = open("/resources/data/Example2.txt","w")
Mode
File object
File1 = open("/resources/data/Example2.txt","w")
Files can be opened in different mode such as 'r' , 'w ', 'a '
File objects are used to obtain information about the file.


DEMO

Writing files with open

File Object

File object



.write(“This is line 1 ”)

Continue..

```
File1 = open("/resources/data/Example2.txt","w")
```

Mode

```
File1.write ("This is line A\n" )
```


DEMO

Additional Data Structures

Base Python package has built in data structures as:

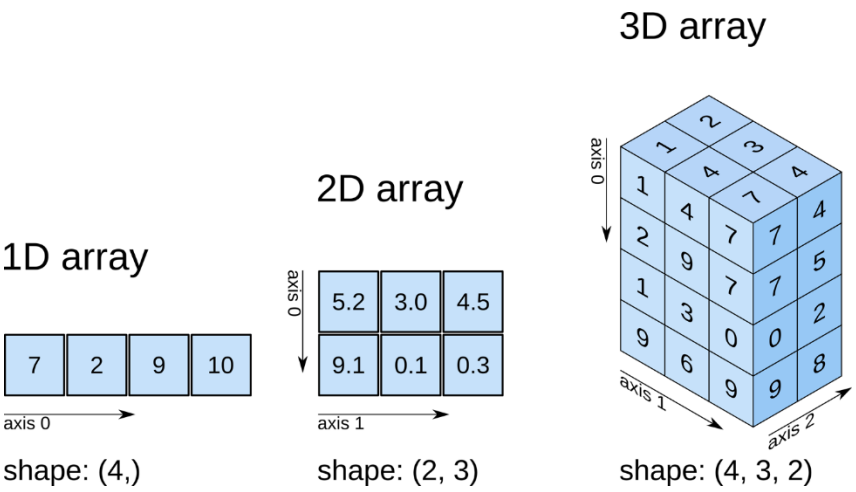
Lists, Tuples, Sets, Dictionaries

These are not sufficient for numeric, scientific and analytic advances hence additional data structures were made available as:

Arrays - Numpy

Series - Pandas

Data frames - Pandas



Pandas

pandas

pandas is a fast, powerful, flexible and easy to use open source data analysis and manipulation tool, built on top of the Python programming language.

Importing Library

```
import pandas
```

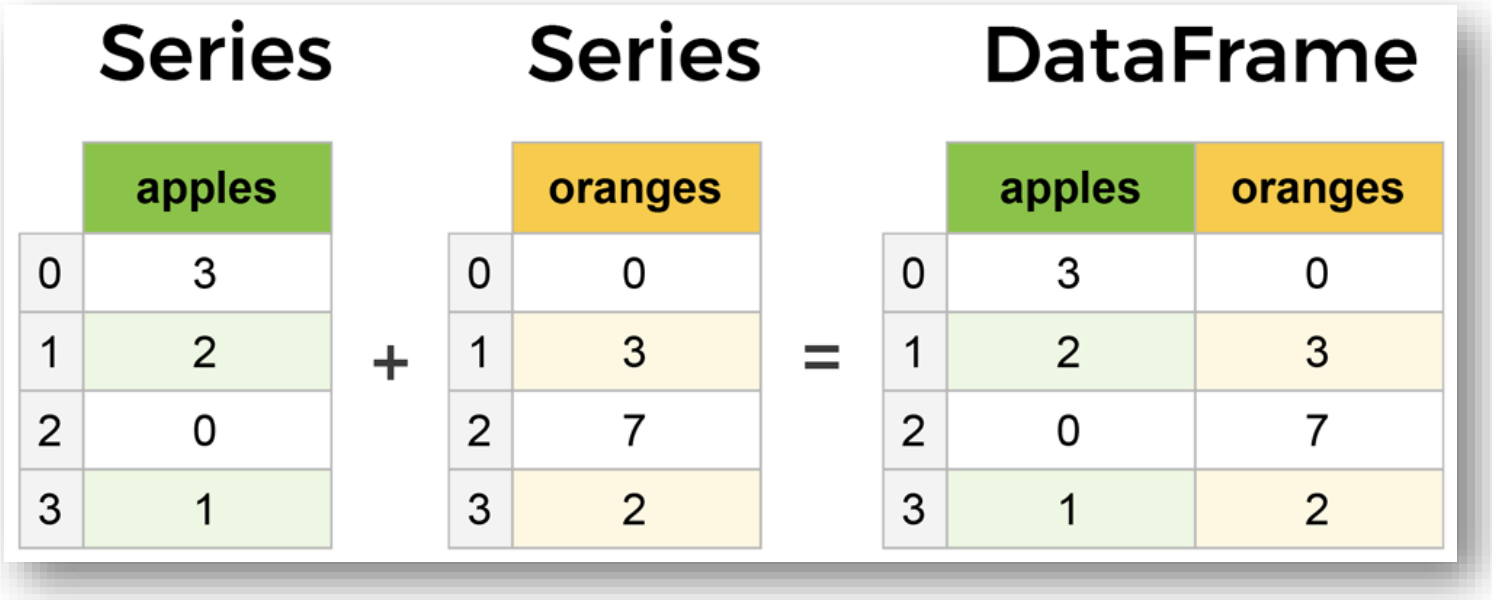
Python library is a collection of functions and methods that allows you to perform many actions without writing your code

Once the library is imported it provides access to a large number of pre-built classes and functions.

Creating Data

There are two core objects in pandas:

- DataFrame
- Series



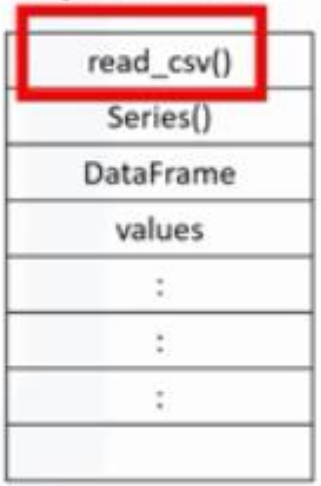

DEMO

Loading data with Pandas

```
import pandas

csv_path='file1.csv'

df=pandas.read_csv(csv_path)
```



DataFrame

```
csv_path='file1.csv'
df= pd.read_csv(csv_path)
```

```
xlsx_path='file1.xlsx'
df= pd.read_excel (xlsx_path)
```

Save Dataset

Pandas enables us to save the dataset to different methods

Read/Save Other Data Formats

Data Formate	Read	Save
csv	pd.read_csv()	df.to_csv()
json	pd.read_json()	df.to_json()
excel	pd.read_excel()	df.to_excel()
hdf	pd.read_hdf()	df.to_hdf()
sql	pd.read_sql()	df.to_sql()
...	...	...

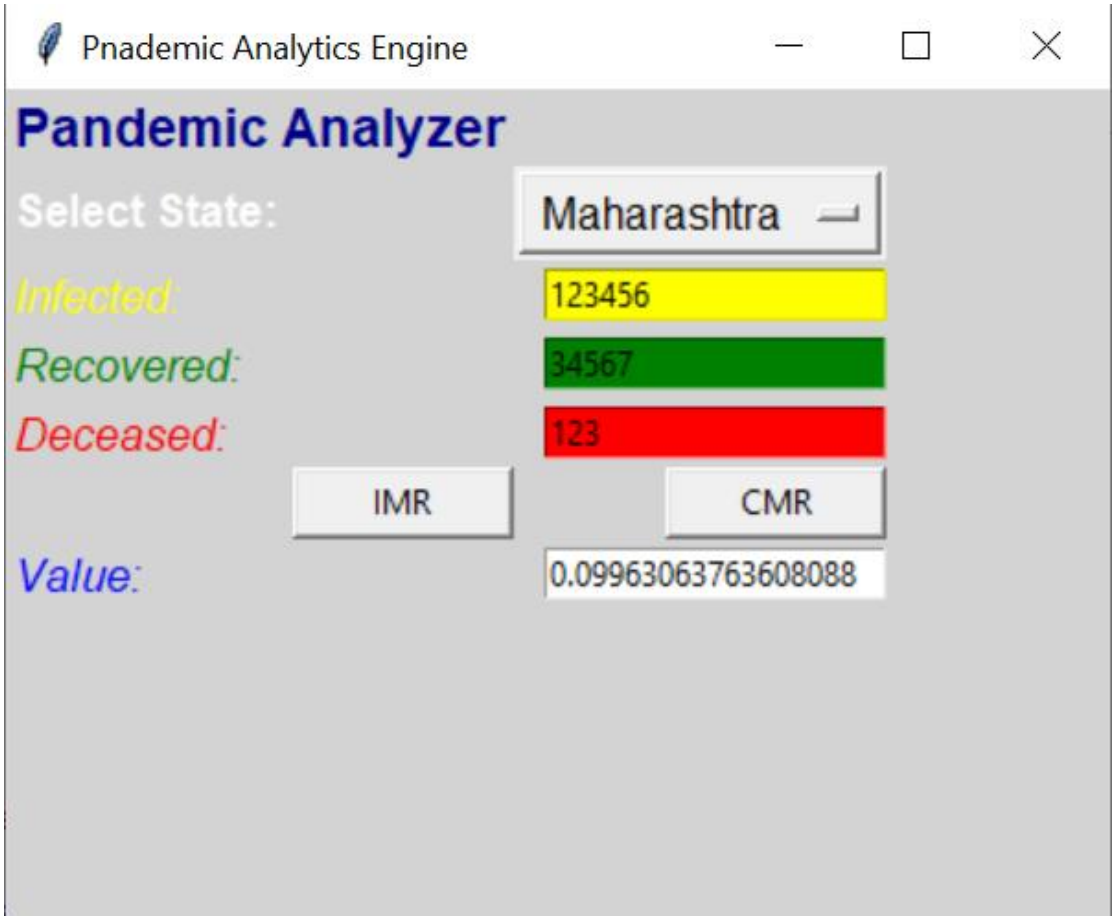
Take away…

- Importing Libraries
- Loading Data with pandas
- Saving Dataset

Hands in grease

Let’s get our hands wet

Take Home Exercise: THE #2



Day: 3
Session: 1

❑ Importing Datasets