PANDAS

Presentation By: Prof.Prachi Pancholi

Ganpat University

- *derived from the term "panel data“*
- acquainted with your data by cleaning, transforming, and analyzing it.
- built on the Numpy package
- high-performance data manipulation and analysis tool
- Installation:

pip install pandas(on windows)

Check :

>>>import pandas as pd

Core components of pandas (data structures)

- Series - a column
  - 1D labeled homogeneous array, size immutable, values of data mutable
- DataFrame - a multi-dimensional table made up of a collection of Series.
  - general 2D labeled, size-mutable tabular structure with potentially heterogeneously typed columns, data mutable
- Panel- General 3D labeled, size-mutable array
  - Container of Data frame, size mutable, data mutable
  - Labeled axes (rows and columns)
  - Can Perform Arithmetic operations on rows and columns

Series

A series can be created using various inputs like −

- Array
- Dict
- Scalar value or constant

Dataframe

A pandas DataFrame can be created using various inputs like −

- Lists
- dict
- Series
- Numpy ndarrays
- Another DataFrame

Panel

- 3D container of data
- 3 axis:
  - **items** − axis 0, each item corresponds to a DataFrame contained inside.
  - **major_axis** − axis 1, it is the index (rows) of each of the DataFrames.
  - **minor_axis** − axis 2, it is the columns of each of the DataFrames.

```
>>>data = np.random.rand(2,4,5)
>>>p = pd.Panel(data)
>>>print(p)
```

Output:

```
<class 'pandas.core.panel.Panel'>
Dimensions: 2 (items) x 4 (major_axis) x 5 (minor_axis)
Items axis: 0 to 1
Major_axis axis: 0 to 3
Minor_axis axis: 0 to 4
```

Note: this slide is just for informative purpose. (Panel wont be asked in exam)