PYTHON PANDAS TUTORIAL

Notes by Sharad sir

***What is Pandas?

Pandas is a Python library used for working with data sets.

It has functions for analyzing, cleaning, exploring, and manipulating data.

The name "Pandas" has a reference to both "Panel Data", and "Python Data Analysis"

*** Why Use Pandas?

Pandas allows us to analyze big data and make conclusions based on statistical theories.

Pandas can clean messy data sets, and make them readable and relevant.

Relevant data is very important in data science.

PYTHON PANDAS TUTORIAL

Notes by Sharad sir

***What Can Pandas Do?

Pandas gives you answers about the data.
Like:

- Is there a correlation between two or more columns?

- What is average value?

- Max value?

- Min value?

Pandas are also able to delete rows that are not relevant, or contains wrong values, like empty or NULL values. This is called cleaning the data.

***Where is the Pandas Codebase?

The source code for Pandas is located at this github repository

https://github.com/pandas-dev/pandas

PYTHON PANDAS TUTORIAL

Notes by Sharad sir

***Installation of Pandas

If you have Python and PIP already installed on a system, then installation of NumPy is very easy.

Install it using this command:

C:\Users\Your Name>pip install pandas

***Import pandas

Once pandas is installed, import it in your applications by adding the import keyword:

import pandas

***Checking NumPy Version

import pandas as pd
print(pd.___version___)