Pandas

what is Pandas?.

→ Pandas is a Python library used for working with datasets.
it has functions for analysing, cleaning, exploring, and manipulating data.
The name "Pandas" has a reference to both "panel Data" and "Python Data Analysis and was created by wes Mckinney in 2008.

→ why we pandas?.
Pandas allows us to analyze big data and make conclusions based on ~~start~~ statistical theories.
Pandas can clean messy datasets, and make them readable and relevant.
Relevant data is very important in data science.

→ What can Pandas Do?.

• Is there a correlation between two or more columns.

• what is average value?

• max value?

• min value?

installation of Pandas.

if you have Python and PIP already installed on a system, then installation of Pandas is very easy

import Pandas

→ import Pandas.

Pandas as Pd

Pandas is usually under the pd alias.

create an alias with the as keyword while importing.

import Pandas as pd

→ checking Pandas version.

The version string is sorted under _version_ attribute.

```
import pandas as pd

Print ( pd. __version__ )
```

Pandas Series

what is a series ?

A pandas series is like a column in a table.

Functions

Locate Row

Pandas use the loc attribute to return one or more specified row(s)

Named Indexes.

with the index argument , you can name your own indexes.

locate named indexes

use the named index in the loc attribute to return the specified row.

[1]
```
import numpy as np
import pandas as pd
```

Basic data structures in pandas.

1. Series: a one-dimensional labeled array holding data of any type. such as integers, strings, Python objects etc...

2. Dataframe: a two-dimensional data structure that holds data like a two-dimensional array or a table with rows and columns.

Object creation

Creating a Series by Passing a list of values, letting Pandas create a default RangeIndex.

viewing data

See the Essentially basics functionality section.

Use DataFrame. head() and DataFrame. tail() to view the top and bottom rows of the frame respectively:

→ df. head()

→ df. tail()

Display the Dataframe. index or DataFrame. columns:

→ df. Index.

→ describe() show a quick statistic summary of your data.

→ df. describe()

Transposing your data:

df.T

→ DataFrame. sort_index()

→ DataFrame. sort_values()

Selection by Position.

DataFrame. iloc() or DataFrame. iat()

Merge

Concat

pandas provide various facilities for easily combining together Series' and DataFrame objects with various kinds of set logic for the indexes and relational algebra functionality in the case of join / merge - type operations.

concatenating pandas objects together row-wise with concat ()

join

merge () enables SQL style join types along spesific columns. See the Database style joining section.

→ Pd.merge (left, right, on = "Key")

Scanned with OKEN Scanner

The stack() method "compresses" a level in the DataFrame's columns.

→ stacked = df.stack(future_stack=True)

Plotting

→ matplotlib.

→ import matplotlib.pyplot as plt.