**Pandas (adv. Python)**

# Pandas is a data manipulation and analysis tool that is built on Numpy.

# Pandas uses a data structure known as DataFrame (think of it as Microsoft excel in Python).

# DataFrames empower programmers to store and manipulate data in a tabular fashion (rows and columns).

# Series Vs. DataFrame? Series is considered a single column of a DataFrame.

**pandas** –

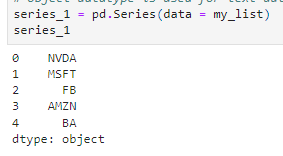
1. Series- 1D ( sirf 1 column hoga pure data me )
2. dataFrame- 2D ( 1 ya 1 se jada column hote hain)

**Importing pandas**



**Creating list to series**

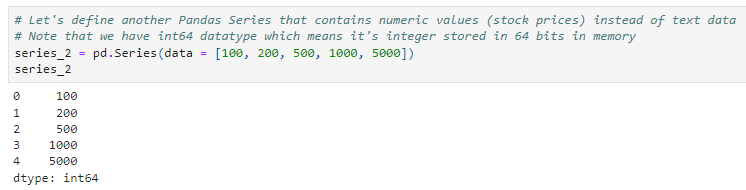




Obeject data type (string) hota hai pandas me

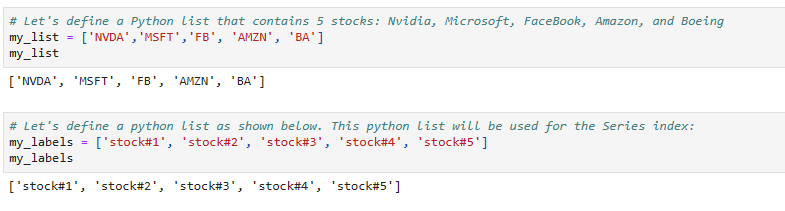
1. Note that series is formed of data and associated index (numeric index has been automatically generated)
2. Check Pandas Documentation for More information: https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.Series.html#pandas.Series
3. Object datatype is used for text data (String)

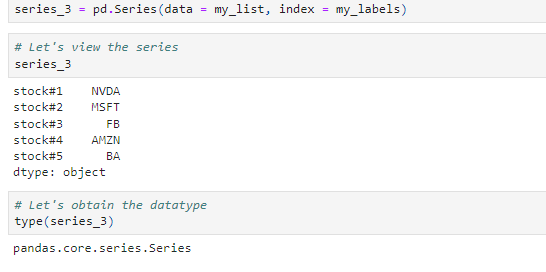
**Another way to creating list to series :-**

****

32 bit operating system hoga to int32 aayaga or 64 bit hoga to int64 aayaga ye os batata hai.

**DEFINE A PANDAS SERIES WITH CUSTOM INDEX**

****

****

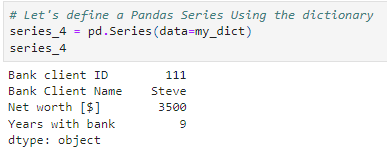
# DEFINE A PANDAS SERIES FROM A DICTIONARY

**A Dictionary consists of a collection of key-value pairs.**

Or key as a indexing kaam karta hai value as a value karat hai kaam.

****

**Now converting above dictionary into series**

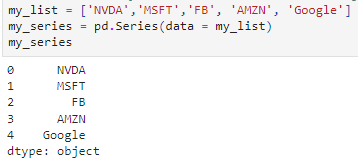
****

* **Attributes –** values shape,size {jike piche parenthese () , nahi lagte vo attributes hote hain}
* **Methods-** tail(), head(),etc. {jike piche parenthese () , lagte vo Methods hote hain}
* **Indexers-** [] {core python ki tarah same indexing}

# PANDAS ATTRIBUTES

**Attributes/Properties: do not use parantheses "()"**

* **values**

****

Values output hamesha array me deta hai

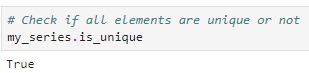
****

* **index**

# 

# dtype

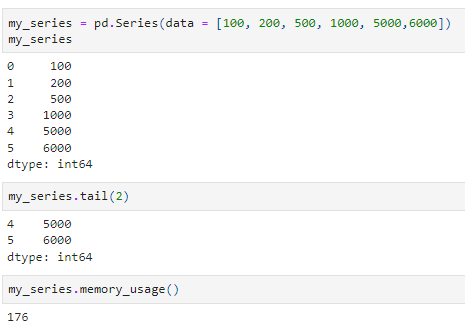
# 

* **is\_unique**
* ****

****

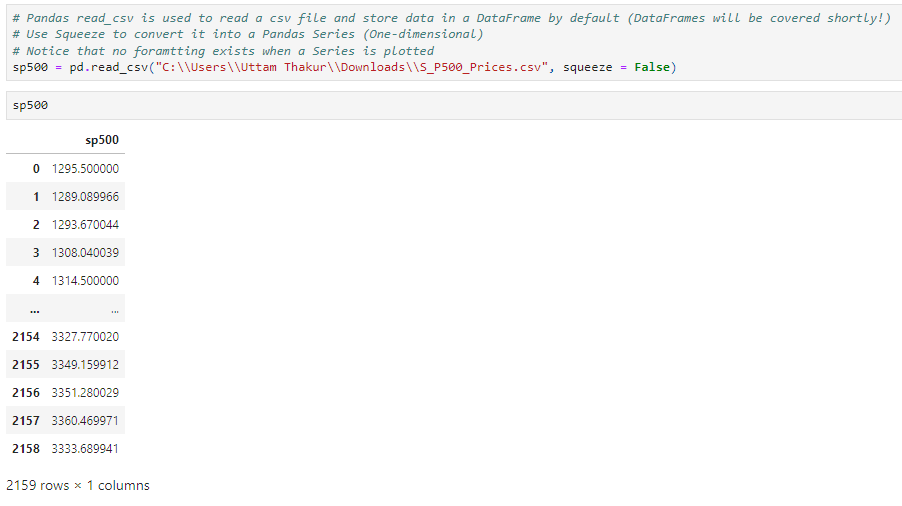
**Head () –** bina k 5 rows dikhata hai suru k or iski max limit 60 rows hoti hai

**Tail() –** bina k 5 rows dikhata hai last k or iski max limit 60 rows hoti hai

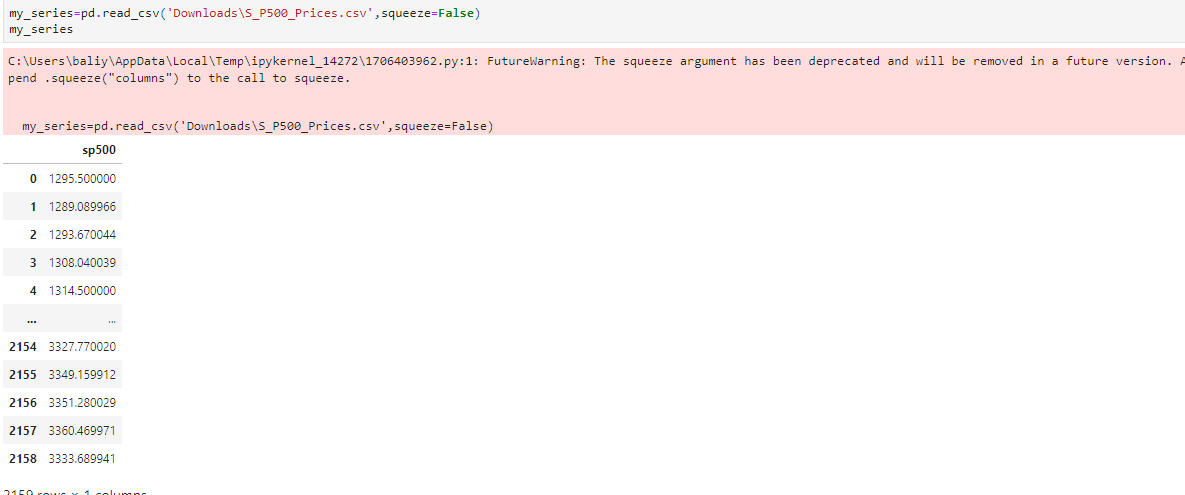
****

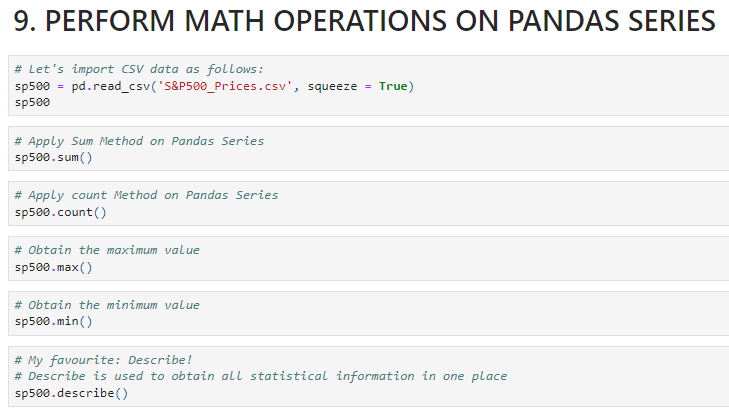
# IMPORT CSV DATA (1-D) USING PANDAS

**Data file used in note 🡪**  [**click here**](https://github.com/naveenkrch/pandas_files_4archivers/blob/main/S_P500_Prices%20(1).csv)

****

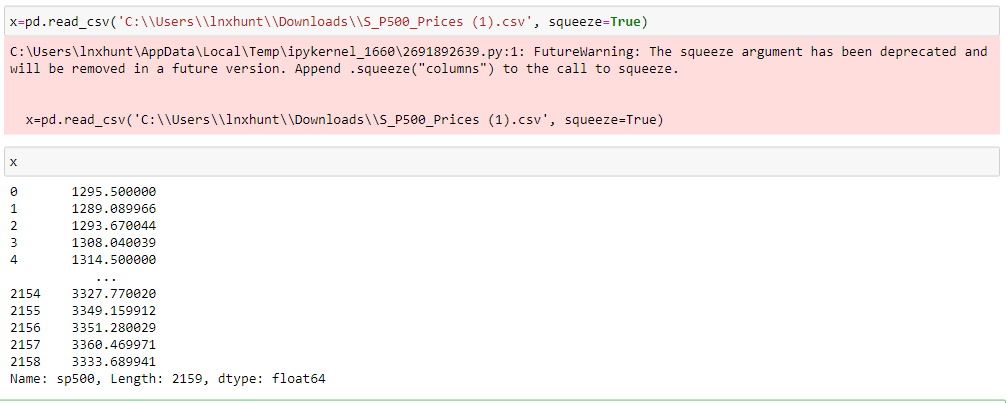
Squeeze= true hoga toh series me datatype aayaga or false hoga toh data frame me datatype aayaga

****

****

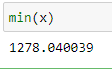
**Shorting pandas series**

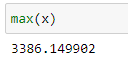
* **sort\_values( )**

****

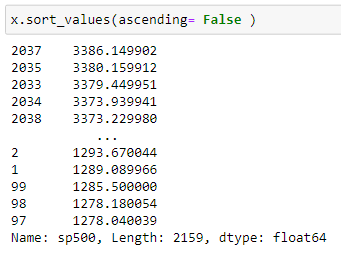
Isme ye assending to descending hain values

Sabse minimum value check karne k liye min function ka use kar sakte hain

****

****

**Agar value descending se ascending me chiye toh :-**

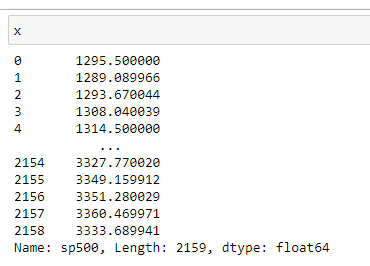
****

Sabse highest value

Isse actual data me change nahi ho raha hai.

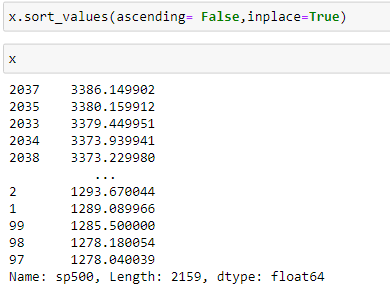
Agar actual data me changes karna hai toh

{ inplace = True } command dalna hoga

****

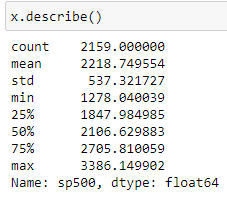
Original data me change nahi hua hai

**Make change in original data**

****

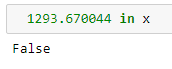
ab original data me change ho gaya hai , { inplace } ki help se

**Statical information of data like(std, sum,count,mean ,25%,50%,75)**

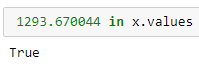
****

**(in) function :-**

ye values dekhne k kaam aata hai ki value data me exist karti hai ni nahi



Ye value exist kar rahi hai fir bhi false ara hai kyu ki ye by default indexing dekhta hai



Agar values check karna hai ki exist karta hai ki nahi{ values } attribute ka use karna hoga

**[ Dataframe ]**

1**. Meta Data-->info()**

**2. Statcial information-->describe()**

**3. find out the null values-->isna().sum,isnull().sum()**

**4. drop column-->need axis=1**

**5. add new column**

**6. rename column**

**7. Drop null values**

**8. Fill null values**

**9. groupby**

**10. concat**

**11. Merge**

**12. chnge the datatype**

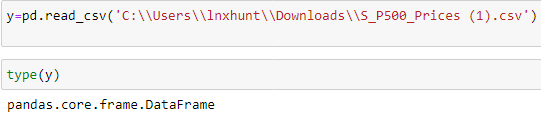
**13. loc,iloc**

**14. sort\_values**

**15. Set index**

**16. reset index**

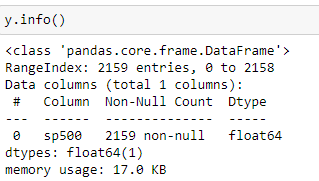
**17. indexing**



Metadata

Column ka datatype

Batata hai



Column me null values hai ki ni ye batata hai.

Describe

