

PANDAS

pandas is a python library it is used for data analysis and data manipulation

why it is used

pandas is used for storing,reading,cleaning, and analysing the data

installing pandas

```
In [1]: # pip install pandas
```

importing pandas

```
In [2]: import pandas as pd
```

Data structures in pandas

pandas provide two datastructures for manipulating the data

1) Pandas Series

A pandas series is a one dimensional labelled array it is capable of holding any data type(int,float,string)

creating a pandas series

```
In [3]: s=pd.Series([1,3,5,7,9])
print(s)
```

```
0    1
1    3
2    5
3    7
4    9
dtype: int64
```

```
In [4]: s=pd.Series(["Akshitha",19,50.5,21])
print(s)
```

```
0    Akshitha
1        19
2        50.5
3        21
dtype: object
```

```
In [5]: import pandas as pd
import numpy as np
```

```
ser=pd.Series()
print("pandas Series",ser)
data=np.array(['p','y','t','h','o','n'])
ser=pd.Series(data)
print("Pandas Series:\n",ser)
```

```
pandas Series Series([], dtype: object)
Pandas Series:
  0    p
  1    y
  2    t
  3    h
  4    o
  5    n
dtype: object
```

2)Pandas data frame

pandas dataframe is a two dimensional data structure it is used to store data in rows and columns just like an excel

```
In [6]: import pandas as pd
df=pd.DataFrame()
print(df)
lst=["python", "is","powerful","language"]
df=pd.DataFrame(lst)
print(df)
```

```
Empty DataFrame
Columns: []
Index: []
  0
0    python
1      is
2  powerful
3  language
```

creating an empty pandas series

```
In [7]: ser=pd.Series()
print(ser)
```

```
Series([], dtype: object)
```

creating a series from numpy array

```
In [8]: import pandas as pd
import numpy as np
data=np.array(['p','y','t','h','o','n'])
ser=pd.Series(data)
print(ser)
```

```
0    p  
1    y  
2    t  
3    h  
4    o  
5    n  
dtype: object
```

creating a series from list

```
In [9]: data_list=["data science",2018,13.5,"easy"]  
ser1=pd.Series(data_list)  
print(ser1)
```

```
0    data science  
1        2018  
2        13.5  
3       easy  
dtype: object
```

creating a series from a dectionary

```
In [10]: data_dict={"name":"india","age":25,'eligible':'yes'}  
ser2=pd.Series(data_dict)  
print(ser2)
```

```
name      india  
age        25  
eligible   yes  
dtype: object
```

common methods of series

s.describe():- describe function is used to provide summary statistics of the data

```
In [11]: s=pd.Series([1,3,5,7,9])  
print(s)
```

```
0    1  
1    3  
2    5  
3    7  
4    9  
dtype: int64
```

```
In [12]: s.describe()
```