

NAME: Shresthi Singh

REG NO: 21BCE2682

BRANCH: CSE CORE

UNIVERSITY: VITVELLORE

**Task - 1 Create a pandas dataframe
(DataFrame name as 'df') (10 observations and 5 features)**

In [1]:

```
import pandas as pd
import numpy as np
d={"Maths" : [24,26,21,12,30,25,11,22,14,29], "English": [10,9,14,25,np.nan,
    "Hindi": [4,6,9,np.nan,15,18,21,30,24,19]}
df = pd.DataFrame(d)
df
```

Out[1]:

	Maths	English	Science	History	Hindi
0	24	10.0	11	12	4.0
1	26	9.0	22	23	6.0
2	21	14.0	14	14	9.0
3	12	25.0	16	16	NaN
4	30	NaN	17	25	15.0
5	25	26.0	19	9	18.0
6	11	19.0	24	4	21.0
7	22	20.0	26	13	30.0
8	14	27.0	27	26	24.0
9	29	16.0	1	22	19.0

Task- 2 Check the info of 'df'

In [2]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 10 entries, 0 to 9
```

```
Data columns (total 5 columns):
```

#	Column	Non-Null Count	Dtype
0	Maths	10 non-null	int64
1	English	9 non-null	float64
2	Science	10 non-null	int64
3	History	10 non-null	int64
4	Hindi	9 non-null	float64

```
dtypes: float64(2), int64(3)
```

```
memory usage: 528.0 bytes
```

Task 3- Check the descriptive statistics of 'df'

```
In [3]: df.describe()
```

```
Out[3]:
```

	Maths	English	Science	History	Hindi
count	9.000000	10.000000	10.000000	10.000000	9.000000
10.000000		0	0		
0					
mean	18.444444	17.700000	16.400000	16.222222	
21.400000		4	0	0	2
0					
std	6.866999	6.728876	7.860591	7.351493	8.599096
min	11.000000	9.000000	1.000000	4.000000	4.000000
25%	14.000000	14.500000	12.250000	9.000000	
15.750000		0	0	0	
0					
50%	19.000000	18.000000	15.000000	18.000000	
23.000000		0	0	0	0
75%	25.000000	23.500000	22.750000	21.000000	
25.750000		0	0	0	0
0					
max	30.000000	27.000000	27.000000	26.000000	30.000000
		0	0	0	0

```
In [4]: df.median()
```

```
Out[4]: Maths      23.0

English      19.0
Science      18.0
History      15.0
Hindi        18.0
```

```
dtype: float64
```

```
In [5]: df.mode()
```

```
Out[5]:
```

	Maths	Englis h	Scienc e	Histor y	Hind i
0	11	9.0	1	4	4.0
1	12	10.0	11	9	6.0
2	14	14.0	14	12	9.0
3	21	16.0	16	13	15.0
4	22	19.0	17	14	18.0
5	24	20.0	19	16	19.0
6	25	25.0	22	22	21.0
7	26	26.0	24	23	24.0
8	29	27.0	26	25	30.0
9	30	NaN	27	26	NaN

Task 4- check the 4th index observation with 'loc' slicing operator.

```
In [6]: df.loc[4]
Out[6]:
Maths      30.0
English    NaN
Science     17.0
History     25.0
Hindi       15.0
Name: 4, dtype: float64
```

Task 5 - Check the null values in your 'df'

```
df.isnull().any()

In [7]:
Out[7]:
Maths      False
English     True
Science    False
History     False
Hindi      Tru
```

```
df.isnull()
```

```
In [8]:
```

Out[8]:

	Math s	Englis h	Scienc e	Histor y	Hin di
0	False	False	False	False	Fals e
1	False	False	False	False	Fals e
2	False	False	False	False	Fals e
3	False	False	False	False	Tru e
4	False	True	False	False	Fals e
5	False	False	False	False	Fals e
6	False	False	False	False	Fals e
7	False	False	False	False	Fals e
8	False	False	False	False	Fals e
9	False	False	False	False	Fals e