NAME: Aditya Kumar Gupta

REG NO: 21BCE2344

BRANCH: CSE CORE

UNIVERSITY: VIT VELLORE

Task - 1 Create a pandas dataframe (DataFrame name as 'df') (10 observationand 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
     float64dtypes: float64(2),
int64(3)
memory usage: 528.0 bytes
```

Task 3- Check the descriptive statistics of'df'

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

Out[3]:

	Maths	Maths	English	Science	History	Hindi
count	9.00000 10.000 00		10.0000 0	10.0000 00	10.0000 00	9.00000 0
mean	18.4444 21.4000 0		17.70000 44	16.4000 0	16.4000 00	16.22222 2
std	6.86699		6.72887 6	7.860591	7.351493	8.59909 6
min	11.00000		9.00000 0	1.00000 0	4.00000 0	4.00000 0
25%	14.0000 15.750 00		14.5000 00	12.25000 0	9.00000 0	
50%	19.0000 23.00000		18.0000 00	15.0000 00	18.0000 00	
75%	25.0000 25.7500 0		23.5000 00	22.7500 00	21.0000 00	
max	27.0000 30.00000		27.0000 00	26.0000 00	30.0000 00	

In [5]:

In [4]:

Out[4]:

```
df.median()
```

```
Maths      23.0

English   19.0
Science   18.0
History   15.0
Hindi     18.0
dtype:
float6
4
```

```
df.mode()
```

	Maths	English	Science	History	Hindi
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 observationwith '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]:

	Maths	English	Science	History	Hindi
0	False	False	False	False	False
1	False	False	False	False	False
2	False	False	False	False	False
3	False	False	False	False	True
4	False	True	False	False	False
5	False	False	False	False	False
6	False	False	False	False	False
7	False	False	False	False	False
8	False	False	False	False	False
9	False	False	False	False	False