

In [30]:

```
import numpy as np
import pandas as pd
import math
```

In [11]:

```
df=pd.read_csv("raw_data.csv")
```

In [12]:

```
df
```

Out[12]:

	math score	reading score	writing score	Placement Score	placement offer count
0	72.000000	72.00	74.0	78.000	1
1	69.000000	90.00	88.0	62.125	2
2	90.000000	95.00	93.0	74.000	2
3	47.000000	57.00	77.0	78.000	1
4	52.285714	78.00	75.0	81.000	3
5	71.000000	72.25	78.0	70.000	4
6	12.000000	44.00	52.0	12.000	2
7	52.285714	65.00	67.0	49.000	1
8	5.000000	77.00	89.0	55.000	0

In [17]:

```
sum_values=df['math score'].sum()
```

In [20]:

```
count_values=df['math score'].count()
```

In [36]:

```
mean=sum_values/count_values
```

In [37]:

```
mean
```

Out[37]:

```
52.28571422222222
```

In [25]:

```
np.sort(df['math score'])
```

Out[25]:

```
array([ 5.      , 12.      , 47.      , 52.285714, 52.285714, 69.
        , 71.      , 72.      , 90.      ])
```

In [31]:

```
middle_index=math.floor(count_values/2)
```

In [34]:

```
median=np.sort(df['math score'])[middle_index]
```

In [38]:

```
median
```

Out[38]:

```
52.285714
```

In [39]:

```
max_value=np.sort(df['math score'])[-1]
```

In [40]:

```
max_value
```

Out[40]:

```
90.0
```

In [41]:

```
min_value=np.sort(df['math score'])[0]
```

In [42]:

```
min_value
```

Out[42]:

```
5.0
```

In [44]:

```
def max_value(variable):  
    return np.sort(df[variable])[-1]
```

In [45]:

```
max_value('math score')
```

Out[45]:

90.0

In [46]:

```
def min_value(variable):  
    return np.sort(df[variable])[0]
```

In [47]:

```
min_value('math score')
```

Out[47]:

5.0

In [48]:

```
def median_value(variable):  
    middle_index=math.floor(count_values/2)  
    return np.sort(df[variable])[middle_index]
```

In [49]:

```
median_value('math score')
```

Out[49]:

52.285714

In [79]:

```
# lst=[]  
# for i in range(count_values):  
#     sigma=(df['math score'][i]-mean)**2  
#     lst.append(sigma)
```

In [60]:

```
std_dev
```

Out[60]:

26.515045145117497

In [53]:

```
df['math score'].std()
```

Out[53]:

28.123452338370043

In [116]:

```
def std_dev(variable):
    lst=[]
    for i in range(count_values):
        sigma=(df[variable][i]-mean)**2
        lst.append(sigma)
    return np.sqrt(sum(lst)/count_values)
```

In [117]:

```
std_dev('math score')
```

Out[117]:

26.515045145117497

In [84]:

```
df.describe()
```

Out[84]:

	math score	reading score	writing score	Placement Score	placement offer count
count	9.000000	9.000000	9.0000	9.000000	9.000000
mean	52.285714	72.250000	77.0000	62.125000	1.777778
std	28.123452	15.698328	12.5499	21.791268	1.201850
min	5.000000	44.000000	52.0000	12.000000	0.000000
25%	47.000000	65.000000	74.0000	55.000000	1.000000
50%	52.285714	72.250000	77.0000	70.000000	2.000000
75%	71.000000	78.000000	88.0000	78.000000	2.000000
max	90.000000	95.000000	93.0000	81.000000	4.000000

In [124]:

```
def variance(data,ddof=0):
    n = len(data)
    mean = sum(data) / n
    return sum((x - mean) ** 2 for x in data) / (n - ddof)

def stdev(data):
    var = variance(data)
    std_dev = math.sqrt(var)
    return std_dev
```

In [125]:

```
variance(df['math score'])
```

Out[125]:

703.047619047619

In [126]:

```
stdev(df['math score'])
```

Out[126]:

26.515045145117497

In [128]:

```
stdev(df['reading score'])
```

Out[128]:

14.800525516195549

In [129]:

```
stdev(df['placement offer count'])
```

Out[129]:

1.1331154474650633

In []: