Assigment on Data_Science concepts
1)Quan,Qual,
2)Measure of Loacation of Data- Frequency, Relative Frequency, Cumulative Frequency, Percentile
3)Measure of central tendency - Mean, Median, Mode

Submitted by ABDHAHEER @ ABU - DATED: 17.05.22

In [13]:
```
import pandas as pd
```

In [14]:
```
dataset=pd.read_csv('Placement.csv')
```

In [15]:
```
dataset["ssc_p"].mean()
```

Out[15]:
```
67.3033953488372
```

In [16]:
```
dataset["ssc_p"].median()
```

Out[16]:
```
67.0
```

In [17]:
```
dataset["ssc_p"].mode()
```

Out[17]:
```
0    62.0
dtype: float64
```

In [18]:
```
from univariant import Univariate
```

In [19]:
```
obj=Univariate()
```

In [20]:
```
Quan,Qual=obj.QuanQual(dataset)
```

```
sl_no
gender
ssc_p
ssc_b
hsc_p
hsc_b
hsc_s
degree_p
degree_t
workex
etest_p
specialisation
mba_p
status
salary
```

In [93]:
```
Quan
```

Out[93]:
```
['sl_no', 'ssc_p', 'hsc_p', 'degree_p', 'etest_p', 'mba_p', 'salary']
```

In [22]:
```
Qual
```

Out[22]:
```
['gender',
 'ssc_b',
 'hsc_b',
 'hsc_s',
 'degree_t',
 'workex',
 'specialisation',
 'status']
```

In [23]:
```
unitable=pd.DataFrame(index=["Mean"],columns=Quan)
```

In [24]:
```
unitable
```

Out[24]:

	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
Mean	NaN	NaN	NaN	NaN	NaN	NaN	NaN

In [25]:
```
unitable["sl_no"]["Mean"]
```

Out[25]:
```
nan
```

In [26]:
```
for columnName in Quan:
    unitable[columnName]["Mean"]=dataset[columnName].mean()
```

In [27]:
```
unitable
```

Out[27]:

	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
Mean	108	67.3034	66.3332	66.3702	72.1006	62.2782	288655

In [28]:
```
unitable=pd.DataFrame(index=["Mean","Median","Mode"],columns=Quan)
for columnName in Quan:
    unitable[columnName]["Mean"]=dataset[columnName].mean()
    unitable[columnName]["Median"]=dataset[columnName].median()
    unitable[columnName]["Mode"]=dataset[columnName].mode()
```

In [29]:
```
unitable
```

Out[29]:

	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
Mean	108	67.3034	66.3332	66.3702	72.1006	62.2782	288655
Median	108	67	65	66	71	62	265000
Mode	0 1 1 2 2 3 3 4 4 ...	0 62.0 dtype: float64	0 63.0 dtype: float64	0 65.0 dtype: float64	0 60.0 dtype: float64	0 56.7 dtype: float64	0 300000.0 dtype: float64

In [30]:
```
dataset["ssc_p"].mode()
```

Out[30]:
```
0    62.0
dtype: float64
```

In [31]:
```
dataset["ssc_p"].mode()[0]
```

Out[31]:
```
62.0
```

In [32]:
```
unitable=pd.DataFrame(index=["Mean","Median","Mode"],columns=Quan)
for columnName in Quan:
    unitable[columnName]["Mean"]=dataset[columnName].mean()
    unitable[columnName]["Median"]=dataset[columnName].median()
    unitable[columnName]["Mode"]=dataset[columnName].mode()[0]
```

In [33]:
```
unitable
```

Out[33]:

	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
Mean	108	67.3034	66.3332	66.3702	72.1006	62.2782	288655
Median	108	67	65	66	71	62	265000
Mode	1	62	63	65	60	56.7	300000

In [34]:
```
dataset["ssc_p"].value_counts()
```

Out[34]:
```
62.00    11
63.00    10
67.00     9
52.00     9
73.00     9
         ..
80.60     1
65.20     1
80.40     1
59.96     1
78.50     1
Name: ssc_p, Length: 103, dtype: int64
```

In [35]:
```
dataset["ssc_p"].value_counts().sort_index()
```

Out[35]:
```
40.89    1
41.00    1
43.00    1
44.00    1
45.00    1
        ..
85.80    1
86.50    1
87.00    3
88.00    1
89.40    1
Name: ssc_p, Length: 103, dtype: int64
```

In [36]:
```
dataset["ssc_p"].value_counts().sort_values()
```

Out[36]:
```
48.00     1
61.08     1
55.68     1
81.70     1
51.57     1
         ..
73.00     9
52.00     9
67.00     9
63.00    10
62.00    11
Name: ssc_p, Length: 103, dtype: int64
```

In [37]:
```
freqtab=pd.DataFrame(columns=["unique_values","Frequency","Relative_Freq","Cum_Freq"])
```

In [38]:
```
freqtab
```

Out[38]:

unique_values	Frequency	Relative_Freq	Cum_Freq

In [39]:
```
dataset["ssc_p"].value_counts().sort_values().index
```

Out[39]:
```
Float64Index([ 48.0, 61.08, 55.68,  81.7, 51.57,  56.6,  75.4, 67.16,  69.7,
                41.0,
              ...
               77.0,  58.0,  69.0,  74.0,  65.0,  73.0,  52.0,  67.0,  63.0,
               62.0],
             dtype='float64', length=103)
```

In [40]:
```
freqtab["unique_values"]=dataset["ssc_p"].value_counts().sort_values().index
```

In [41]:
```
freqtab
```

Out[41]:

	unique_values	Frequency	Relative_Freq	Cum_Freq
0	48.00	NaN	NaN	NaN
1	61.08	NaN	NaN	NaN
2	55.68	NaN	NaN	NaN
3	81.70	NaN	NaN	NaN
4	51.57	NaN	NaN	NaN
...	...	...	...	...
98	73.00	NaN	NaN	NaN
99	52.00	NaN	NaN	NaN
100	67.00	NaN	NaN	NaN
101	63.00	NaN	NaN	NaN
102	62.00	NaN	NaN	NaN

103 rows × 4 columns

In [42]:
```
dataset["ssc_p"].value_counts().sort_values().values
```

Out[42]:
```
array([ 1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,
        1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,
        1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,
        1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,
        1,  1,  2,  2,  2,  2,  2,  2,  2,  2,  2,  3,  3,  3,  3,  3,  3,
        3,  3,  3,  3,  4,  4,  4,  4,  6,  6,  7,  7,  7,  9,  9,  9, 10,
       11], dtype=int64)
```

In [43]:
```
freqtab["Frequency"]=dataset["ssc_p"].value_counts().sort_values().values
```

In [44]:
```
freqtab
```

Out[44]:

	unique_values	Frequency	Relative_Freq	Cum_Freq
0	48.00	1	NaN	NaN
1	61.08	1	NaN	NaN
2	55.68	1	NaN	NaN
3	81.70	1	NaN	NaN
4	51.57	1	NaN	NaN
...	...	...	...	...
98	73.00	9	NaN	NaN
99	52.00	9	NaN	NaN
100	67.00	9	NaN	NaN
101	63.00	10	NaN	NaN
102	62.00	11	NaN	NaN

103 rows × 4 columns

In [45]:
```
(freqtab["Frequency"]/103)*100
```

Out[45]:
```
0       0.970874
1       0.970874
2       0.970874
3       0.970874
4       0.970874
          ...
98      8.737864
99      8.737864
100     8.737864
101     9.708738
102    10.679612
Name: Frequency, Length: 103, dtype: float64
```

In [46]:
```
freqtab["Relative_Freq"]=(freqtab["Frequency"]/103)*100
```

In [47]:
```
freqtab
```

Out[47]:

	unique_values	Frequency	Relative_Freq	Cum_Freq
0	48.00	1	0.970874	NaN
1	61.08	1	0.970874	NaN
2	55.68	1	0.970874	NaN
3	81.70	1	0.970874	NaN
4	51.57	1	0.970874	NaN
...	...	...	...	...
98	73.00	9	8.737864	NaN
99	52.00	9	8.737864	NaN
100	67.00	9	8.737864	NaN
101	63.00	10	9.708738	NaN
102	62.00	11	10.679612	NaN

103 rows × 4 columns

In [48]:
```
freqtab["Cum_Freq"]=freqtab["Relative_Freq"].cumsum()
```

In [49]:
```
freqtab
```

Out[49]:

	unique_values	Frequency	Relative_Freq	Cum_Freq
0	48.00	1	0.970874	0.970874
1	61.08	1	0.970874	1.941748
2	55.68	1	0.970874	2.912621
3	81.70	1	0.970874	3.883495
4	51.57	1	0.970874	4.854369
...	...	...	...	...
98	73.00	9	8.737864	170.873786
99	52.00	9	8.737864	179.611650
100	67.00	9	8.737864	188.349515
101	63.00	10	9.708738	198.058252
102	62.00	11	10.679612	208.737864

103 rows × 4 columns

In [50]:
```
freqtab=pd.DataFrame(columns=["unique_values","Frequency","Relative_Freq","Cum_Freq"])
freqtab["unique_values"]=dataset["ssc_p"].value_counts().sort_values().index
freqtab["Frequency"]=dataset["ssc_p"].value_counts().sort_values().values
freqtab["Relative_Freq"]=(freqtab["Frequency"]/103)*100
freqtab["Cum_Freq"]=freqtab["Relative_Freq"].cumsum()
```

In [51]:
```
freqtab
```

Out[51]:

	unique_values	Frequency	Relative_Freq	Cum_Freq
0	48.00	1	0.970874	0.970874
1	61.08	1	0.970874	1.941748
2	55.68	1	0.970874	2.912621
3	81.70	1	0.970874	3.883495
4	51.57	1	0.970874	4.854369
...	...	...	...	...
98	73.00	9	8.737864	170.873786
99	52.00	9	8.737864	179.611650
100	67.00	9	8.737864	188.349515
101	63.00	10	9.708738	198.058252
102	62.00	11	10.679612	208.737864

103 rows × 4 columns

In [52]:
```
def FreqTable(dataset,columnName):
    freqtab=pd.DataFrame(columns=["unique_values","Frequency","Relative_Freq","Cum_Freq"])
    freqtab["unique_values"]=dataset[columnName].value_counts().sort_values().index
    freqtab["Frequency"]=dataset[columnName].value_counts().sort_values().values
    freqtab["Relative_Freq"]=(freqtab["Frequency"]/len(freqtab))*100
    freqtab["Cum_Freq"]=freqtab["Relative_Freq"].cumsum()
    return freqtab
```

In [53]:
```
FreqTable(dataset,"ssc_p")
```

Out[53]:

	unique_values	Frequency	Relative_Freq	Cum_Freq
0	48.00	1	0.970874	0.970874
1	61.08	1	0.970874	1.941748
2	55.68	1	0.970874	2.912621
3	81.70	1	0.970874	3.883495
4	51.57	1	0.970874	4.854369
...	...	...	...	...
98	73.00	9	8.737864	170.873786
99	52.00	9	8.737864	179.611650
100	67.00	9	8.737864	188.349515
101	63.00	10	9.708738	198.058252
102	62.00	11	10.679612	208.737864

103 rows × 4 columns

In [54]:
```
FreqTable(dataset,"hsc_p")
```

Out[54]:

	unique_values	Frequency	Relative_Freq	Cum_Freq
0	40.00	1	1.030928	1.030928
1	64.20	1	1.030928	2.061856
2	90.90	1	1.030928	3.092784
3	65.66	1	1.030928	4.123711
4	73.20	1	1.030928	5.154639
...	...	...	...	...
92	64.00	8	8.247423	176.288660
93	60.00	9	9.278351	185.567010
94	67.00	9	9.278351	194.845361
95	62.00	12	12.371134	207.216495
96	63.00	14	14.432990	221.649485

97 rows × 4 columns

In [56]:
```
import numpy as np
np.percentile(dataset["ssc_p"],25)
```

Out[56]:
```
60.599999999999994
```

In [57]:
```
np.percentile(dataset["ssc_p"],50)
```

Out[57]:
```
67.0
```

In [58]:
```
np.percentile(dataset["ssc_p"],75)
```

Out[58]:
```
75.7
```

In [59]:
```
np.percentile(dataset["ssc_p"],80)
```

Out[59]: 77.0

In [60]:
```
np.percentile(dataset["ssc_p"],99)
```

Out[60]: 87.0

In [61]:
```
np.percentile(dataset["ssc_p"],100)
```

Out[61]: 89.4

In [62]:
```
np.percentile(dataset["ssc_p"],5)
```

Out[62]: 49.7

In [65]:
```
unitable=pd.DataFrame(index=["Mean","Median","Mode","25th","50th","75th","99th","100th"],columns=Quan)
for columnName in Quan:
    unitable[columnName]["Mean"]=dataset[columnName].mean()
    unitable[columnName]["Median"]=dataset[columnName].median()
    unitable[columnName]["Mode"]=dataset[columnName].mode()[0]
```

In [66]:
```
unitable
```

Out[66]:

	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
Mean	108	67.3034	66.3332	66.3702	72.1006	62.2782	288655
Median	108	67	65	66	71	62	265000
Mode	1	62	63	65	60	56.7	300000
25th	NaN	NaN	NaN	NaN	NaN	NaN	NaN
50th	NaN	NaN	NaN	NaN	NaN	NaN	NaN
75th	NaN	NaN	NaN	NaN	NaN	NaN	NaN
99th	NaN	NaN	NaN	NaN	NaN	NaN	NaN
100th	NaN	NaN	NaN	NaN	NaN	NaN	NaN

In [67]:
```
unitable=pd.DataFrame(index=["Mean","Median","Mode","25th","50th","75th","99th","100th"],columns=Quan)
for columnName in Quan:
    unitable[columnName]["Mean"]=dataset[columnName].mean()
    unitable[columnName]["Median"]=dataset[columnName].median()
    unitable[columnName]["Mode"]=dataset[columnName].mode()[0]
    unitable[columnName]["25th"]=np.percentile(dataset[columnName],25)
```

```
C:\Users\Lenovo\Anaconda3\lib\site-packages\numpy\lib\function_base.py:3826: RuntimeWarning: Invalid value encountered
in percentile
  interpolation=interpolation)
```

In [68]:
```
unitable
```

Out[68]:

	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
Mean	108	67.3034	66.3332	66.3702	72.1006	62.2782	288655
Median	108	67	65	66	71	62	265000
Mode	1	62	63	65	60	56.7	300000
25th	54.5	60.6	60.9	61	60	57.945	NaN
50th	NaN	NaN	NaN	NaN	NaN	NaN	NaN
75th	NaN	NaN	NaN	NaN	NaN	NaN	NaN
99th	NaN	NaN	NaN	NaN	NaN	NaN	NaN
100th	NaN	NaN	NaN	NaN	NaN	NaN	NaN

In [69]:
```
unitable=pd.DataFrame(index=["Mean","Median","Mode","25th","50th","75th","99th","100th"],columns=Quan)
for columnName in Quan:
    unitable[columnName]["Mean"]=dataset[columnName].mean()
    unitable[columnName]["Median"]=dataset[columnName].median()
    unitable[columnName]["Mode"]=dataset[columnName].mode()[0]
    unitable[columnName]["25th"]=np.percentile(dataset[columnName],25)
    unitable[columnName]["50th"]=np.percentile(dataset[columnName],50)
    unitable[columnName]["75th"]=np.percentile(dataset[columnName],75)
    unitable[columnName]["99th"]=np.percentile(dataset[columnName],99)
    unitable[columnName]["100th"]=np.percentile(dataset[columnName],100)
```

In [70]:
```
unitable
```

Out[70]:

	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
Mean	108	67.3034	66.3332	66.3702	72.1006	62.2782	288655
Median	108	67	65	66	71	62	265000
Mode	1	62	63	65	60	56.7	300000
25th	54.5	60.6	60.9	61	60	57.945	NaN
50th	108	67	65	66	71	62	NaN
75th	161.5	75.7	73	72	83.5	66.255	NaN
99th	212.86	87	91.86	83.86	97	76.1142	NaN
100th	215	89.4	97.7	91	98	77.89	NaN

In [71]:
```
dataset.isnull()
```

Out[71]:

	sl_no	gender	ssc_p	ssc_b	hsc_p	hsc_b	hsc_s	degree_p	degree_t	workex	etest_p	specialisation	mba_p	status	salary
0	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False	False	False	False	False	False	True
4	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
210	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
211	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
212	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
213	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
214	False	False	False	False	False	False	False	False	False	False	False	False	False	False	True

215 rows × 15 columns

In [72]:
```
dataset.isnull().sum()
```

Out[72]:
```
sl_no             0
gender            0
ssc_p             0
ssc_b             0
hsc_p             0
hsc_b             0
hsc_s             0
degree_p          0
degree_t          0
workex            0
etest_p           0
specialisation    0
mba_p             0
status            0
salary           67
dtype: int64
```

In [73]:
```
dataset.isna().sum()
```

Out[73]:
```
sl_no             0
gender            0
ssc_p             0
ssc_b             0
hsc_p             0
hsc_b             0
hsc_s             0
degree_p          0
degree_t          0
workex            0
etest_p           0
specialisation    0
mba_p             0
status            0
salary           67
dtype: int64
```

In [74]:
```
dataset
```

Out[74]:

	sl_no	gender	ssc_p	ssc_b	hsc_p	hsc_b	hsc_s	degree_p	degree_t	workex	etest_p	specialisation	mba_p	status	salary
0	1	M	67.00	Others	91.00	Others	Commerce	58.00	Sci&Tech	No	55.0	Mkt&HR	58.80	Placed	270000.0
1	2	M	79.33	Central	78.33	Others	Science	77.48	Sci&Tech	Yes	86.5	Mkt&Fin	66.28	Placed	200000.0
2	3	M	65.00	Central	68.00	Central	Arts	64.00	Comm&Mgmt	No	75.0	Mkt&Fin	57.80	Placed	250000.0
3	4	M	56.00	Central	52.00	Central	Science	52.00	Sci&Tech	No	66.0	Mkt&HR	59.43	Not Placed	NaN
4	5	M	85.80	Central	73.60	Central	Commerce	73.30	Comm&Mgmt	No	96.8	Mkt&Fin	55.50	Placed	425000.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
210	211	M	80.60	Others	82.00	Others	Commerce	77.60	Comm&Mgmt	No	91.0	Mkt&Fin	74.49	Placed	400000.0
211	212	M	58.00	Others	60.00	Others	Science	72.00	Sci&Tech	No	74.0	Mkt&Fin	53.62	Placed	275000.0
212	213	M	67.00	Others	67.00	Others	Commerce	73.00	Comm&Mgmt	Yes	59.0	Mkt&Fin	69.72	Placed	295000.0
213	214	F	74.00	Others	66.00	Others	Commerce	58.00	Comm&Mgmt	No	70.0	Mkt&HR	60.23	Placed	204000.0
214	215	M	62.00	Central	58.00	Others	Science	53.00	Comm&Mgmt	No	89.0	Mkt&HR	60.22	Not Placed	NaN

215 rows × 15 columns

In [75]:
```
dataset["salary"]=dataset["salary"].fillna(0)
```

In [76]:
```
dataset.isnull().sum()
```

Out[76]:
```
sl_no             0
gender            0
ssc_p             0
ssc_b             0
hsc_p             0
hsc_b             0
hsc_s             0
degree_p          0
degree_t          0
workex            0
etest_p           0
specialisation    0
mba_p             0
status            0
salary            0
dtype: int64
```

In [79]:
```
unitable=pd.DataFrame(index=["Mean","Median","Mode","25th","50th","75th","99th","100th"],columns=Quan)
for columnName in Quan:
    unitable[columnName]["Mean"]=dataset[columnName].mean()
    unitable[columnName]["Median"]=dataset[columnName].median()
    unitable[columnName]["Mode"]=dataset[columnName].mode()[0]
    unitable[columnName]["25th"]=np.percentile(dataset[columnName],25)
    unitable[columnName]["50th"]=np.percentile(dataset[columnName],50)
    unitable[columnName]["75th"]=np.percentile(dataset[columnName],75)
    unitable[columnName]["99th"]=np.percentile(dataset[columnName],99)
    unitable[columnName]["100th"]=np.percentile(dataset[columnName],100)
```

In [80]:
```
unitable
```

Out[80]:

	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
Mean	108	67.3034	66.3332	66.3702	72.1006	62.2782	198702
Median	108	67	65	66	71	62	240000
Mode	1	62	63	65	60	56.7	0
25th	54.5	60.6	60.9	61	60	57.945	0
50th	108	67	65	66	71	62	240000
75th	161.5	75.7	73	72	83.5	66.255	282500
99th	212.86	87	91.86	83.86	97	76.1142	629000
100th	215	89.4	97.7	91	98	77.89	940000

In [81]:
```
unitable=pd.DataFrame(index=["Mean","Median","Mode","25th","50th","75th","99th","100th","IQR","1.5IQR","Lesser","Greater
for columnName in Quan:
    unitable[columnName]["Mean"]=dataset[columnName].mean()
    unitable[columnName]["Median"]=dataset[columnName].median()
    unitable[columnName]["Mode"]=dataset[columnName].mode()[0]
    unitable[columnName]["25th"]=np.percentile(dataset[columnName],25)
    unitable[columnName]["50th"]=np.percentile(dataset[columnName],50)
    unitable[columnName]["75th"]=np.percentile(dataset[columnName],75)
    unitable[columnName]["99th"]=np.percentile(dataset[columnName],99)
    unitable[columnName]["100th"]=np.percentile(dataset[columnName],100)
```

In [82]:
```
unitable
```

Out[82]:

	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
Mean	108	67.3034	66.3332	66.3702	72.1006	62.2782	198702
Median	108	67	65	66	71	62	240000
Mode	1	62	63	65	60	56.7	0
25th	54.5	60.6	60.9	61	60	57.945	0
50th	108	67	65	66	71	62	240000
75th	161.5	75.7	73	72	83.5	66.255	282500
99th	212.86	87	91.86	83.86	97	76.1142	629000
100th	215	89.4	97.7	91	98	77.89	940000
IQR	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1.5IQR	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Lesser	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Greater	NaN	NaN	NaN	NaN	NaN	NaN	NaN

In [83]:
```
unitable=pd.DataFrame(index=["Mean","Median","Mode","25th","50th","75th","99th","100th","IQR","1.5IQR","Lesser","Greater
for columnName in Quan:
    unitable[columnName]["Mean"]=dataset[columnName].mean()
    unitable[columnName]["Median"]=dataset[columnName].median()
    unitable[columnName]["Mode"]=dataset[columnName].mode()[0]
    unitable[columnName]["25th"]=np.percentile(dataset[columnName],25)
    unitable[columnName]["50th"]=np.percentile(dataset[columnName],50)
    unitable[columnName]["75th"]=np.percentile(dataset[columnName],75)
    unitable[columnName]["99th"]=np.percentile(dataset[columnName],99)
    unitable[columnName]["100th"]=np.percentile(dataset[columnName],100)
    unitable[columnName]["IQR"]=unitable[columnName]["75th"]-unitable[columnName]["25th"]
```

In [84]:
```
unitable
```

Out[84]:

	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
Mean	108	67.3034	66.3332	66.3702	72.1006	62.2782	198702
Median	108	67	65	66	71	62	240000
Mode	1	62	63	65	60	56.7	0
25th	54.5	60.6	60.9	61	60	57.945	0
50th	108	67	65	66	71	62	240000
75th	161.5	75.7	73	72	83.5	66.255	282500
99th	212.86	87	91.86	83.86	97	76.1142	629000
100th	215	89.4	97.7	91	98	77.89	940000
IQR	107	15.1	12.1	11	23.5	8.31	282500
1.5IQR	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Lesser	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Greater	NaN	NaN	NaN	NaN	NaN	NaN	NaN

In [85]:
```
unitable=pd.DataFrame(index=["Mean","Median","Mode","25th","50th","75th","99th","100th","IQR","1.5IQR","Lesser","Greater
for columnName in Quan:
    unitable[columnName]["Mean"]=dataset[columnName].mean()
    unitable[columnName]["Median"]=dataset[columnName].median()
    unitable[columnName]["Mode"]=dataset[columnName].mode()[0]
    unitable[columnName]["25th"]=np.percentile(dataset[columnName],25)
    unitable[columnName]["50th"]=np.percentile(dataset[columnName],50)
    unitable[columnName]["75th"]=np.percentile(dataset[columnName],75)
    unitable[columnName]["99th"]=np.percentile(dataset[columnName],99)
    unitable[columnName]["100th"]=np.percentile(dataset[columnName],100)
    unitable[columnName]["IQR"]=unitable[columnName]["75th"]-unitable[columnName]["25th"]
    unitable[columnName]["1.5IQR"]=1.5*unitable[columnName]["IQR"]
    unitable[columnName]["Lesser"]=unitable[columnName]["25th"]-unitable[columnName]["1.5IQR"]
    unitable[columnName]["Greater"]=unitable[columnName]["75th"]-unitable[columnName]["1.5IQR"]
```

In [86]:
```
unitable
```

Out[86]:

	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
Mean	108	67.3034	66.3332	66.3702	72.1006	62.2782	198702
Median	108	67	65	66	71	62	240000
Mode	1	62	63	65	60	56.7	0
25th	54.5	60.6	60.9	61	60	57.945	0
50th	108	67	65	66	71	62	240000
75th	161.5	75.7	73	72	83.5	66.255	282500
99th	212.86	87	91.86	83.86	97	76.1142	629000
100th	215	89.4	97.7	91	98	77.89	940000
IQR	107	15.1	12.1	11	23.5	8.31	282500
1.5IQR	160.5	22.65	18.15	16.5	35.25	12.465	423750
Lesser	-106	37.95	42.75	44.5	24.75	45.48	-423750
Greater	1	53.05	54.85	55.5	48.25	53.79	-141250

In [89]:
```
unitable=pd.DataFrame(index=["Mean","Median","Mode","25th","50th","75th","99th","100th","IQR","1.5IQR","Lesser","Greater
for columnName in Quan:
    unitable[columnName]["Mean"]=dataset[columnName].mean()
    unitable[columnName]["Median"]=dataset[columnName].median()
    unitable[columnName]["Mode"]=dataset[columnName].mode()[0]
    unitable[columnName]["25th"]=np.percentile(dataset[columnName],25)
    unitable[columnName]["50th"]=np.percentile(dataset[columnName],50)
    unitable[columnName]["75th"]=np.percentile(dataset[columnName],75)
    unitable[columnName]["99th"]=np.percentile(dataset[columnName],99)
    unitable[columnName]["100th"]=np.percentile(dataset[columnName],100)
    unitable[columnName]["IQR"]=unitable[columnName]["75th"]-unitable[columnName]["25th"]
    unitable[columnName]["1.5IQR"]=1.5*unitable[columnName]["IQR"]
    unitable[columnName]["Lesser"]=unitable[columnName]["25th"]-unitable[columnName]["1.5IQR"]
    unitable[columnName]["Greater"]=unitable[columnName]["75th"]-unitable[columnName]["1.5IQR"]
    unitable[columnName]["Min"]=unitable[columnName].min()
    unitable[columnName]["Max"]=unitable[columnName].max()
```

In [90]:
```
unitable
```

Out[90]:

	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
Mean	108	67.3034	66.3332	66.3702	72.1006	62.2782	198702
Median	108	67	65	66	71	62	240000
Mode	1	62	63	65	60	56.7	0
25th	54.5	60.6	60.9	61	60	57.945	0
50th	108	67	65	66	71	62	240000
75th	161.5	75.7	73	72	83.5	66.255	282500
99th	212.86	87	91.86	83.86	97	76.1142	629000
100th	215	89.4	97.7	91	98	77.89	940000
IQR	107	15.1	12.1	11	23.5	8.31	282500
1.5IQR	160.5	22.65	18.15	16.5	35.25	12.465	423750
Lesser	-106	37.95	42.75	44.5	24.75	45.48	-423750
Greater	1	53.05	54.85	55.5	48.25	53.79	-141250
Min	-106	15.1	12.1	11	23.5	8.31	-423750
Max	215	89.4	97.7	91	98	77.89	940000

In [91]:
```
def uniAnalysis(dataset,quan):

    unitable=pd.DataFrame(index=["Mean","Median","Mode","25th","50th","75th","99th","100th","IQR","1.5IQR","Lesser","Gre
    for columnName in Quan:
        unitable[columnName]["Mean"]=dataset[columnName].mean()
        unitable[columnName]["Median"]=dataset[columnName].median()
        unitable[columnName]["Mode"]=dataset[columnName].mode()[0]
        unitable[columnName]["25th"]=np.percentile(dataset[columnName],25)
        unitable[columnName]["50th"]=np.percentile(dataset[columnName],50)
        unitable[columnName]["75th"]=np.percentile(dataset[columnName],75)
        unitable[columnName]["99th"]=np.percentile(dataset[columnName],99)
        unitable[columnName]["100th"]=np.percentile(dataset[columnName],100)
        unitable[columnName]["IQR"]=unitable[columnName]["75th"]-unitable[columnName]["25th"]
        unitable[columnName]["1.5IQR"]=1.5*unitable[columnName]["IQR"]
        unitable[columnName]["Lesser"]=unitable[columnName]["25th"]-unitable[columnName]["1.5IQR"]
        unitable[columnName]["Greater"]=unitable[columnName]["75th"]-unitable[columnName]["1.5IQR"]
        unitable[columnName]["Min"]=unitable[columnName].min()
        unitable[columnName]["Max"]=unitable[columnName].max()
    return unitable
```

In [95]:
```
uniAnalysis(dataset,Quan)
```

Out[95]:

	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
Mean	108	67.3034	66.3332	66.3702	72.1006	62.2782	198702
Median	108	67	65	66	71	62	240000
Mode	1	62	63	65	60	56.7	0
25th	54.5	60.6	60.9	61	60	57.945	0
50th	108	67	65	66	71	62	240000
75th	161.5	75.7	73	72	83.5	66.255	282500
99th	212.86	87	91.86	83.86	97	76.1142	629000
100th	215	89.4	97.7	91	98	77.89	940000
IQR	107	15.1	12.1	11	23.5	8.31	282500
1.5IQR	160.5	22.65	18.15	16.5	35.25	12.465	423750
Lesser	-106	37.95	42.75	44.5	24.75	45.48	-423750
Greater	1	53.05	54.85	55.5	48.25	53.79	-141250
Min	-106	15.1	12.1	11	23.5	8.31	-423750
Max	215	89.4	97.7	91	98	77.89	940000

In [ ]:
```

```