

In [1]:

```
from UniAnalysis import Univariate
import pandas as pd
```

In [2]:

```
obj=Univariate()
```

In [3]:

```
dataset=pd.read_csv("Placement.csv")
```

In [4]:

```
qual,quan=obj.QuanQual(dataset)
```

In [5]:

```
qual
```

Out[5]:

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

In [6]:

```
quan
```

Out[6]:

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

In [7]:

```
obj.AnalysisTable(dataset,quan)
```

Out[7]:

	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
Mean	67.3034	66.3332	66.3702	72.1006	62.2782	198702
Median	67	65	66	71	62	240000
Mode	62	63	65	60	56.7	0
25th-Percentile	60.6	60.9	61	60	57.945	0
50th-Percentile	67	65	66	71	62	240000
75th-Percentile	75.7	73	72	83.5	66.255	282500
90th-Percentile	82	79	77.12	92	70.678	356000
99th-Percentile	87	91.86	83.86	97	76.1142	629000
100th-Percentile	89.4	97.7	91	98	77.89	940000
IQR	15.1	12.1	11	23.5	8.31	282500
IQR*1.5	22.65	18.15	16.5	35.25	12.465	423750
Lesser Boundary	37.95	42.75	44.5	24.75	45.48	-423750
Greater Boundary	98.35	91.15	88.5	118.75	78.72	706250
Min value	40.89	37	50	50	51.21	0
Max value	89.4	97.7	91	98	77.89	940000
Outlier	Not Present	Present	Present	Not Present	Not Present	Present

**Here outlier Present does not make any big difference  
so no worries about Outliers row**

In [8]:

```
obj.frequency(dataset, "hsc_p")
```

Out[8]:

	UniqueValues	Frequency	Relative-Freq	Cumulative-Freq
0	37.00	1	1.030928	1.030928
1	39.00	1	1.030928	2.061856
2	40.00	1	1.030928	3.092784
3	42.00	1	1.030928	4.123711
4	42.16	1	1.030928	5.154639
...	...	...	...	...
92	90.90	1	1.030928	217.525773
93	91.00	1	1.030928	218.556701
94	92.00	1	1.030928	219.587629
95	97.00	1	1.030928	220.618557
96	97.70	1	1.030928	221.649485

97 rows × 4 columns

In [ ]: