

## Central Tendency: Percentile

Placement:

```
for eachField in quan:
    descriptive[eachField]["Mean"] = dataset[eachField].mean()
    descriptive[eachField]["Median"] = dataset[eachField].median()
    descriptive[eachField]["Mode"] = dataset[eachField].mode()[0]
    descriptive[eachField]["0%"] = ny.percentile(dataset[eachField], 0)
    descriptive[eachField]["1%"] = ny.percentile(dataset[eachField], 1)
    descriptive[eachField]["Q1:25%"] = ny.percentile(dataset[eachField], 25)
    descriptive[eachField]["Q2:50%"] = ny.percentile(dataset[eachField], 50)
    descriptive[eachField]["Q3:75%"] = ny.percentile(dataset[eachField], 75)
    descriptive[eachField]["99%"] = ny.percentile(dataset[eachField], 99)
    descriptive[eachField]["Q4:100%"] = ny.percentile(dataset[eachField], 100)
```

descriptive

	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
0%	1	40.89	37	50	50	51.21	NaN
1%	3.14	43.14	40.28	50.968	50.0672	51.4682	NaN
Q1:25%	54.5	60.6	60.9	61	60	57.945	NaN
Q2:50%	108	67	65	66	71	62	NaN
Q3:75%	161.5	75.7	73	72	83.5	66.255	NaN
99%	212.86	87	91.86	83.86	97	76.1142	NaN
Q4:100%	215	89.4	97.7	91	98	77.89	NaN

### 1. SSC Percentage (ssc\_p):

- **Q4 (100%):** The maximum score is **89.4**, indicating all scores are below this value.
- **Insight:** The distribution shows a maximum of 89.4, with most scores below this, confirming that Q4 captures the entire range.

### 2. Higher Secondary Percentage (hsc\_p):

- **Q2 to Q3 Observation:**

- Scores increase by approximately **8 marks** (from **Q2: 65** to **Q3: 73**).
- This corresponds to an increase of **66 students** moving from the Q2 to Q3 range (based on **161.5 - 105** total population count).
- **Insight:** Suggests a gradual upward trend in scores, with significant representation between the median and upper quartile.

### 3. SSC Percentage (ssc\_p):

- **Missing Values and Nulls:**
  - If there are missing values in the SSC column, percentile calculations such as **Q1 (25%)** cannot be determined, resulting in NaN values.
- **Insight:** Ensure all missing or null data is handled (e.g., replacing NaN with a placeholder like the mean or median) to allow proper computation.

### 4. MBA Percentage (mba\_p):

- **Observation:** Between **Q3 (75%)** and **Q4 (100%)**, the number of scores increases compared to other intervals.
- **Insight:** Indicates that more candidates have scored at the higher end (above Q3), suggesting a performance increase toward the top percentile.