

Data Descriptive Analysis Report for Placement Data Percentiles

[42]: descriptive

[42]:	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary
Mean	108.0	67.303395	66.333163	66.370186	72.100558	62.278186	288655.405405
Median	108.0	67.0	65.0	66.0	71.0	62.0	265000.0
Mode	1	62.0	63.0	65.0	60.0	56.7	300000.0
Q1:25%	54.5	60.6	60.9	61.0	60.0	57.945	240000.0
Q2:50%	108.0	67.0	65.0	66.0	71.0	62.0	265000.0
Q3:75%	161.5	75.7	73.0	72.0	83.5	66.255	300000.0
Q5:99%	212.86	87.0	91.86	83.86	97.0	76.1142	671200.0
Q4:100%	215	89.4	97.7	91.0	98.0	77.89	940000.0

To provide a descriptive analysis of the placement data and calculate the difference between the highest and lowest values within each percentile range, we can follow these steps:

Descriptive Analysis Report

1. Summary Statistics:

- **Mean:** The average value of each metric.
- **Median:** The middle value separating the higher half from the lower half.
- **Mode:** The most frequently occurring value.
- **Quartiles (Q1, Q2, Q3):** These values split the data into four equal parts.
- **99th Percentile (Q5):** The value below which 99% of the observations fall.
- **100th Percentile (Q4):** The maximum value.

2. Detailed Analysis of Percentiles:

Here's a breakdown of each percentile and the differences:

- **Q1 (25th Percentile):** This is the value below which 25% of the data falls.
- **Q2 (50th Percentile or Median):** This is the middle value of the dataset.
- **Q3 (75th Percentile):** This is the value below which 75% of the data falls.
- **Q4 (100th Percentile):** This is the maximum value in the dataset.
- **Q5 (99th Percentile):** This is slightly less than the maximum value but represents 99% of the data.

3. Calculation of Differences Between Percentile Ranges:

For each metric (ssc_p, hsc_p, degree_p, etest_p, mba_p, salary), calculate the range difference between the highest and lowest values for each quartile.

Summary Table

Metric	Q1 - Q2 Difference	Q2 - Q3 Difference	Q3 - Q4 Difference	Q4 - Q5 Difference
ssc_p	$54.5 - 108.0 = 53.5$	$108.0 - 161.5 = 53.5$	$161.5 - 215 = 53.5$	$215 - 212.86 = 2.14$
hsc_p	$60.6 - 67.0 = 6.4$	$67.0 - 75.7 = 8.7$	$75.7 - 89.4 = 13.7$	$89.4 - 87.0 = 2.4$
degree_p	$60.9 - 65.0 = 4.1$	$65.0 - 73.0 = 8.0$	$73.0 - 97.7 = 24.7$	$97.7 - 91.86 = 5.84$
etest_p	$61.0 - 66.0 = 5.0$	$66.0 - 72.0 = 6.0$	$72.0 - 91.0 = 19.0$	$91.0 - 83.86 = 7.14$
mba_p	$60.0 - 71.0 = 11.0$	$71.0 - 83.5 = 12.5$	$83.5 - 98.0 = 14.5$	$98.0 - 97.0 = 1.0$
salary	$240000 - 265000 = 25000$	$265000 - 300000 = 35000$	$300000 - 940000 = 640000$	$940000 - 671200 = 268800$

Explanation:

- **ssc_p:** The differences between quartiles are relatively stable except between Q3 and Q4, where the difference is smaller.
- **hsc_p:** Differences between quartiles increase as we move towards the higher percentiles, indicating more variability in higher percentiles.
- **degree_p:** Shows the highest variability between Q3 and Q4, suggesting significant differences in degree percentages in the higher range.
- **etest_p:** Differences increase at higher percentiles, indicating larger variability in higher test scores.
- **mba_p:** Differences are more stable with a significant drop between Q4 and Q5, showing smaller variability at the highest MBA percentages.
- **salary:** The largest range difference is between Q3 and Q4, reflecting significant variability in higher salary ranges.

This analysis helps in understanding how the data is distributed across different percentiles and identifies where significant changes occur in each metric.