HOPE AI

Percentile Analysis Document

Github link of Data:

https://github.com/JayachandraPrabha/3.Data-Science/blob/main/Placement.csv

Percentile is nothing but the comparison between a particular score and the scores of the rest of the data.

Example: if the score is 75 points on a test, and ranked in the 85 th percentile, it means that the score 75 is higher than 85% of the scores.

Let us consider the given data, some of the basic informations about the given data are,

1) The overall shape of the dataset: **215 rows × 15 columns.**
2) The data basically deals with the survey of the candidates who have **placed and not placed** in the placement and the columns in the data are,
    - 'sl_no', 'gender', 'ssc_p', 'ssc_b', 'hsc_p', 'hsc_b', 'hsc_s', 'degree_p', 'degree_t', 'workex', 'etest_p', 'specialization', 'mba_p', 'status', 'salary'
    - In the above mentioned columns there were qualitative columns / variables (Categorical data) and quantitative columns / variables (Numerical data), hence both were separated.
3) From the above data, the percentile is calculated as follows:

	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	\
Mean	108.0	67.303395	66.333163	66.370186	72.100558	62.278186	
Median	108.0	67.0	65.0	66.0	71.0	62.0	
Mode	1	62.0	63.0	65.0	60.0	56.7	
min	1.0	40.89	37.0	50.0	50.0	51.21	
Q1:25%	54.5	60.6	60.9	61.0	60.0	57.945	
Q2:50%	108.0	67.0	65.0	66.0	71.0	62.0	
Q3:75%	161.5	75.7	73.0	72.0	83.5	66.255	
99%	212.86	87.0	91.86	83.86	97.0	76.1142	
Q4:100%	215.0	89.4	97.7	91.0	98.0	77.89	

	salary
Mean	288655.405405
Median	265000.0
Mode	300000.0
min	200000.0
Q1:25%	240000.0
Q2:50%	265000.0
Q3:75%	300000.0
99%	NaN
Q4:100%	940000.0

1) From the obtained insights from the data let us calculate the percentile from the differences:

Column Name	0th Q: 0-25%	1-Q: 25-50%	2-Q: 50-75%	3-Q: 75-99%	4-Q: 99-100%
SSC pass %	60.6-40.89 =19.71	67-60.6 =6.4	75.7-67 =8.7	87-75.7 =11.3	89.4-87 =2.4
HSC pass %	60.9-37 =23.9	65-60.9 =4.1	73-65 =8	91.86-73 =18.86	97.7-91.86 =5.84
Degree pass %	61-50 =11	66-61 =5	72-66 =6	83.86-72 =11.86	91-83.86 =7.14
ETest pass %	60-50 =10	71-60 =11	83.5-71 =12.5	97-83.5 =13.5	98-97 =1
MBA pass %	57.94-51.21 =6.73	62-57.94 =0.73	66.25-62 =4.25	76.11-66.25 =9.86	77.89-76.11 =1.78
Salary	40,000	25,000	35,000	-	-

Conclusion:

- SSC pass % lies higher in the 0th Quadrant (0-25%) → 19.71
- HSC pass % lies higher in the 0th Quadrant (0-25%) → 23.9
- Degree pass % lies higher in the 3rd Quadrant (75-99%) → 11.86
- Entrance test pass % lies higher in the 3rd Quadrant (75-99%) → 13.5
- MBA pass % lies higher in the 3rd Quadrant (75-99%) → 9.86
