

# Analysis of PLACEMENT DATASET

## logic to separate Quan & Qual - Automation logic using Function

```
def quanQual(dataset):
    quan=[]
    qual=[]
    for columnName in dataset.columns:
        if (dataset[columnName].dtypes == "O"): # "O" - Object Data type
            qual.append(columnName)
        else:
            quan.append(columnName)
    return quan,qual

quan,qual = quanQual(dataset) #calling quanQual() function
print("quan:",quan) # numerical data
print("qual:",qual) # categorical data

quan: ['ssc_p', 'hsc_p', 'degree_p', 'etest_p', 'mba_p', 'salary']
qual: ['gender', 'ssc_b', 'hsc_b', 'hsc_s', 'degree_t', 'workex', 'specialisation', 'status']
```

## Measure of central tendency

```
for columnName in quan:
    univariate_analysis.loc["Mean",columnName]=dataset[columnName].mean()
    univariate_analysis.loc["Median",columnName]=dataset[columnName].median()
    univariate_analysis.loc["Mode",columnName]=dataset[columnName].mode()[0]

univariate_analysis
```

	<b>ssc_p</b>	<b>hsc_p</b>	<b>degree_p</b>	<b>etest_p</b>	<b>mba_p</b>	<b>salary</b>
<b>Mean</b>	67.303395	66.333163	66.370186	72.100558	62.278186	288655.405405
<b>Median</b>	67.0	65.0	66.0	71.0	62.0	265000.0
<b>Mode</b>	62.0	63.0	65.0	60.0	56.7	300000.0

## MY ANALYSIS IN BUSINEES POV:

### Analysis of Placement Dataset Using Measures of Central Tendency

I analyzed the placement dataset by first separating the **quantitative** and **qualitative** columns. For the quantitative variables, I applied measures of central tendency—**mean**, **median**, and **mode**—to understand the overall academic and placement performance of the students.

## ◆ Mean Analysis

The mean values provide an overview of the average performance of students across different stages of education:

- The **mean SSLC pass percentage** is **67%**, indicating that students have, on average, performed at an **average level** in their SSLC board examinations.
- The **mean HSC pass percentage** is **66.3%**, which also suggests an **average performance** among students.
- The **mean degree pass percentage** is **66.3%**, showing that the overall degree performance of the batch is **average**.
- The **mean E-test score** is **72.1%**, which reflects a **good performance** in the entrance test.
- The **mean MBA pass percentage** is **62%**, indicating that students have performed at an **average level** during their MBA.
- The **mean salary offered during placements** is **₹2,88,655**, which gives an idea of the average compensation received by students.

From the mean analysis, it can be inferred that most students demonstrate **average academic performance**, with relatively better results in the E-test.

## ◆ Median Analysis

The median values help understand the central tendency **without the influence of outliers**:

- The **median SSLC percentage** is **67%**
- The **median HSC percentage** is **65%**
- The **median degree percentage** is **66%**

These values indicate that the majority of students have performed at an **average level** academically.

- The **median E-test score** is **71%**, suggesting that a significant group of students performed **well** in the test.
- The **median MBA percentage** is **62%**, again reflecting an **average performance**.
- The **median salary** is **₹2,65,000**, which is lower than the mean salary, indicating the presence of **higher salary outliers** in the dataset.

Since the median is less affected by extreme values, it provides a more **realistic representation** of the typical student's performance and salary.

## ◆ Mode Analysis

The mode identifies the most frequently occurring values:

- **SSLC mode:** 62%

- **HSC mode:** 63%
- **Degree mode:** 65%
- **E-test mode:** 60%
- **MBA mode:** 56.7%
- **Salary mode:** ₹3,00,000

This indicates that most students tend to score around these values, and the most common salary package offered is ₹3 LPA.

### ◆ Overall Conclusion

From the analysis using measures of central tendency:

- The dataset represents a group of students who have generally **average academic performance**.
- **Mean** includes all values, including extreme scores (outliers).
- **Median** provides a better estimate of typical performance by excluding the influence of outliers.
- **Mode** helps identify the most common scores and salary offered.

Overall, this analysis shows that the batch consists predominantly of **average-performing students**, with comparatively better performance in the E-test and a commonly offered salary package of ₹3 LPA.