**Inferential Analysis**

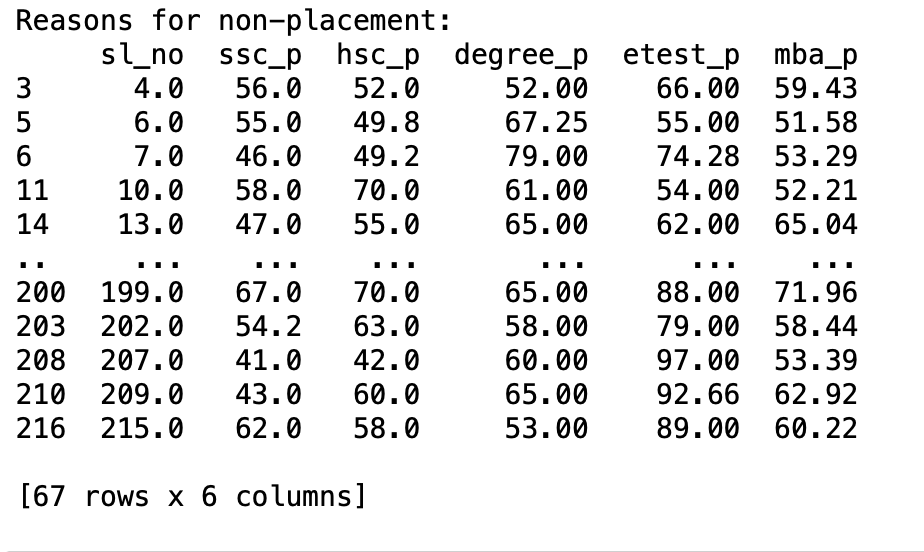
1)Replace the NaN values with correct value. And justify why you have chosen the same.

The dataset, named 'df,' has been checked for NaN values using the isnull().sum() method. Fortunately, there are no missing values in any of the columns. This implies that all data points are complete, allowing for robust analysis without the need for imputation. The absence of NaN values ensures the integrity of the dataset for subsequent analyses.

2) How many of them are not placed?

The number of individuals who are not placed: 67

3) Find the reason for non placement from the dataset?



The provided statistics describe the academic performance of non-placed students in various metrics. On average, non-placed students have relatively lower scores in secondary education (ssc\_p: 57.54), higher secondary education (hsc\_p: 58.40), and undergraduate studies (degree\_p: 61.13). Additionally, their average scores in the employment test (etest\_p: 69.59) and MBA program (mba\_p: 61.61) are included. The common factors among non-placed students suggest a tendency towards lower academic performance compared to their placed counterparts.

4)What kind of relation between salary and mba\_p

The correlation coefficient between salary and MBA percentage (mba\_p) is approximately 0.18. This positive correlation value suggests a weak positive linear relationship between the two variables. However, the correlation is not very strong, indicating that the MBA percentage alone may not be a highly predictive factor for salary in the given dataset. Further analysis or exploration of additional features may be necessary to better understand the factors influencing salary in this context.

5)Which specialization is getting minimum salary?

The specialisation with the minimum salary is

Marketing and Finance with a salary of 200000.0.

6) How many of them getting above 500000 salary?

The number of individuals getting above 500,000 salary is: 3

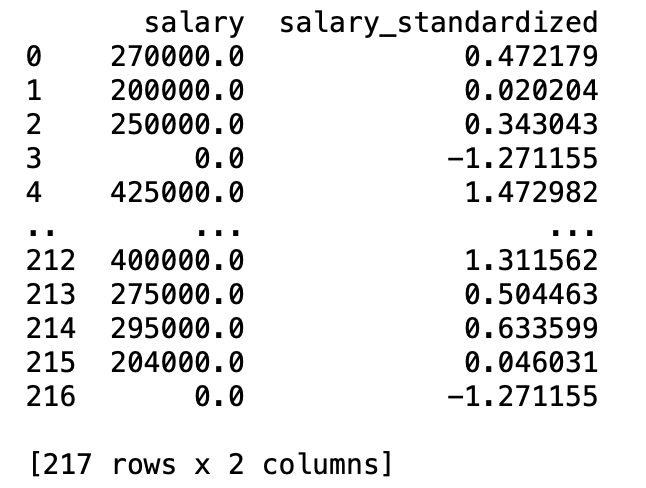
7)Test the Analysis of Variance between etest\_p and mba\_p at signifance level 5%.(Make decision using Hypothesis Testing)

Reject the null hypothesis. There is a significant difference between etest\_p and mba\_p.

8)Test the similarity between the degree\_t(Sci&Tech) and specialisation(Mkt&HR) with respect to salary at significance level of 5%.(Make decision using Hypothesis Testing)

Reject the null hypothesis. There is a significant difference in salary between 'Sci&Tech' and 'Mkt&HR'.

9) Convert the normal distribution to standard normal distribution for salary column.



The original 'salary' values alongside their corresponding Z-scores in the 'salary\_standardized' column. Each Z-score quantifies how many standard deviations a salary is from the mean of the original distribution. Positive Z-scores indicate salaries above the mean, while negative Z-scores indicate salaries below the mean in the standardized distribution.

10)What is the probability Density Function of the salary range from 700000 to 900000?

The probability density function for the salary range

from 700000 to 900000 is approximately: 0.0006

11) Test the similarity between the degree\_t(Sci&Tech) with respect to etest\_p and mba\_p at significance level of 5%.(Make decision using Hypothesis Testing)

Fail to reject the null hypothesis. There is no significant correlation between 'etest\_p' and 'mba\_p' within 'Sci&Tech' degree.

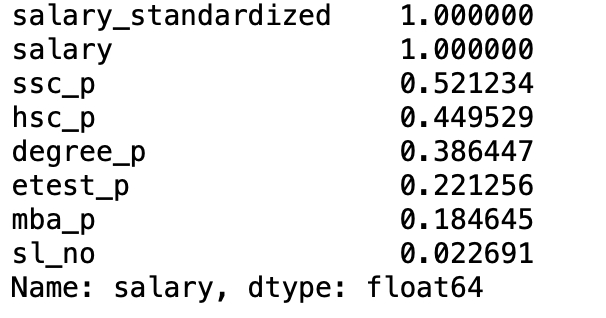
Reject the null hypothesis. There is a significant mean

difference in 'etest\_p' within 'Sci&Tech' degree.

Reject the null hypothesis. There is a significant mean

difference in 'mba\_p' within 'Sci&Tech' degree.

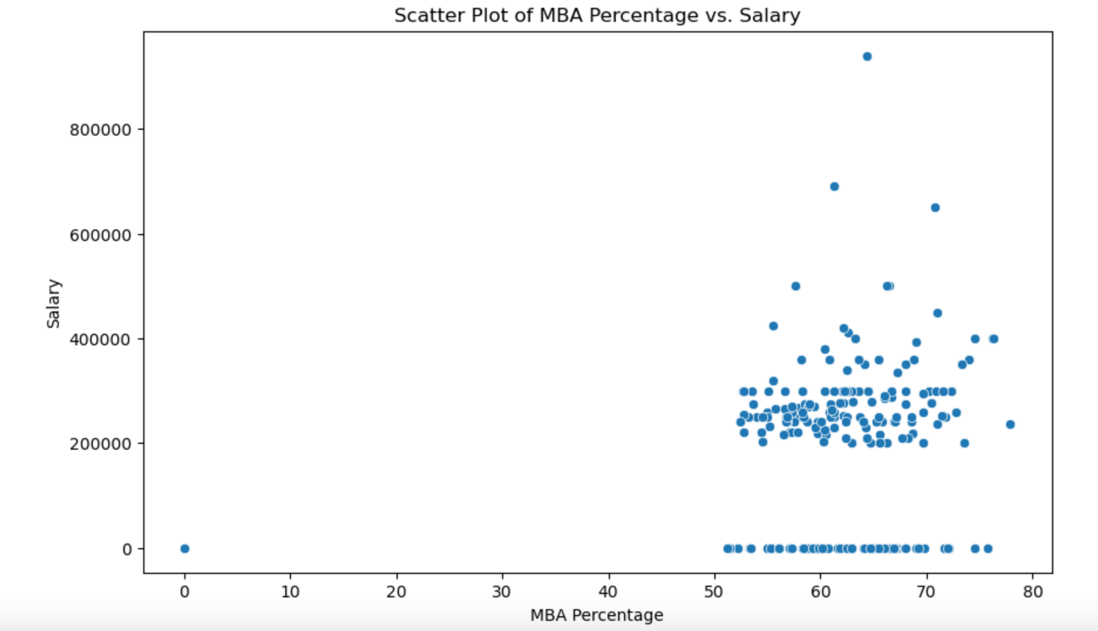
12)Which parameter is highly correlated with salary?



Based on the provided correlation coefficients, the parameter most highly correlated with 'salary' is 'ssc\_p' (SSC Percentage) with a correlation coefficient of approximately 0.521. This positive correlation suggests that there is a moderate positive relationship between SSC Percentage and Salary.

The correlation coefficients are sorted in descending order, and 'ssc\_p' has the highest positive correlation among the listed parameters.

13) plot any useful graph and explain it.



The scatter plot shows that as MBA percentages increase, there is a general trend of higher salaries, especially in the 50 to 80 percentage range. However, the scattered distribution suggests that the relationship is not strictly linear, indicating that other factors may also influence salary outcomes.

Git hub Link:<https://github.com/Vidhyavino/Inferential-Analysis/blob/5c70fdb358b2a943f8cfe0493bb37336691d48fc/%20Inferential%20Analysis.ipynb>