

DETAILED REPORT ON DATA EXPLORATION ANALYSIS

E.No.: 23314065

Name: Megh Bhavesh Shah

Data Type: The data being analysed is the placement data of randomly selected 200 students of a college. It includes students' academic details, work experience and placement status.

Report on the Data Analysis:

Starting with observing the data given, the data contains 215 rows and 15 columns. First 5 columns are shown for initial understanding.

Then, null values are checked. The null values were only in the salary column, which were because those students didn't get placed. Those null values were imputed by 0. The duplicates were placed. The outliers are handled by replacing them with the nearest boundary. At last, the cleaned dataset is checked again.

Summary statistics for each column is described. Skewness of each column is calculated. Frequency distribution of each categorical variable is shown. From this, it is inferred that 65% of the students were male, and 69% of the students got placed.

Histogram and Box Plots were made for the numerical variables. From them, it is indicated that maximum students scored 65% in their degree course and 55-65% in the e-test. Approximately 70

people didn't get placed (= 0 salary) and 270000 was the mode salary.

From the correlation matrix, it is inferred that salary was largely determined by the SSC percentage and Degree percentage. The scatter plots implied that those who got good percentile in both degree and SSC were likely to get placed.

Bar, Violin and Box Plots were plotted for every pair of a continuous variable and a numerical variable. ['GENDER', 'WORKEX', 'STATUS'] were taken as the categorical variables while ['DEGREE_P', 'MBA_P'] were taken as the numerical variables. Similarly, every other variable pair can be plotted.

Similarly, pair plots were drawn for every pair of numerical variables.

Grouped comparison of Salary was done based on Specialisation, gender and work experience. The comparison implies that for males, work experience mattered both for HR and Finance firms. But for females, work experience mattered only in Finance firms. It didn't matter that much for HR firms.

To conclude, data analysis on placement statistics provided that placements were hugely dependent on degree percentage and ssc percentage.