Report: Analysis of Placement Data (BrainDead, Revelation 2k23)

Team Members: Dibyarup Dutta, Anuvab Sen, Mritunjay Haldar, Abhiroop Mukherjee, Aritra Bandyopadhyay

University: Indian Institute of Engineering Science and Technology, Shibpur

Overview

Challenge Description

The main goal of this project is to analyze the factors affecting the factors that affect the placement and salary of the students. The dataset contains information about the placement records of students in a MBA college.

Dataset Description

The dataset includes variables such as secondary and higher secondary school percentages, degree specializations, work experience, and the salary offered to the students.

Relevant Links:

1. **Data Analysis Link**
   https://www.kaggle.com/code/anuvabsen123/thedatadorks/notebook

Dataset Exploration

- **Data Exploration on the Entire Dataset:** We first used pairplot and regression analysis on the entire dataset to derive insight of the entire dataset

- **Cluster Exploration based on Placement Status:** We divided the dataset into two parts based on placement status and derived insights from the two datasets

Cluster Exploration based on Specialization: We then divided the dataset into two parts based on specialization and then derived insights from the two divided datasets

Data Analysis Methods

We used various data analysis methods, including:

Pairplot:
Pairplot visualizes given data to find the relationship between them where the variables can be continuous or categorical.

Exploratory Data Analysis:
Various histograms were plotted to find out trends in Data so that these trends can be analyzed further

Correlation Matrix:
Linear Bivariate analysis method to find out correlations between various attributes.

Pie Charts:
Pie Charts were plotted that show trends among placed and non-placed students for various attributes

Bar Graphs:
Bar Graphs were plotted that helped in finding trends in min, max, median and average salaries and how they relate to various attributes

Box Plots:
Box Plots included show the capture trends and show information about which percentile of a certain class gets more salary, it is a standardized way of displaying the dataset based on the five-number summary: the minimum, the maximum, the sample median, and the first and third quartiles

Results of Analysis

Observation: Salaries distribution are on the lower side of spectrum (Expected):

Analysis:

The salary range of the students in the dataset is between 200,000 and 940,000, with a mean salary of 288,655. The median salary is 265,000, which indicates that the salary distribution is slightly skewed towards the lower end of the spectrum.

This observation is in line with the expectations, as fresh MBA graduates usually start with lower salaries and gradually progress in their careers.

```
max          940000.000000
min          200000.000000
mean         288655.405405
median       265000.000000
Name: salary, dtype: float64
```

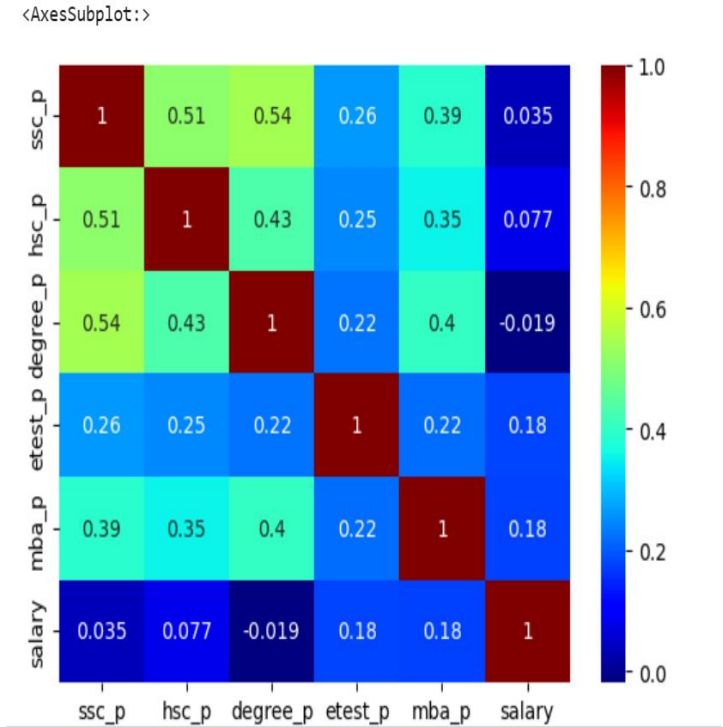
Observation: Correlation exists between degree and high school percentages

The dataset analysis revealed a moderate level of correlation between the high school and degree percentages of the students, indicating that students who performed well in high school tend to perform well in their degree programs (undergraduate as well as specialization).

Analysis:

However, the correlation coefficient is low, suggesting that factors such as work experience, specialization, and performance in the interview rounds are more important in determining the placement and salary of the students.

The salary is only slightly correlated to the MBA percentage as well as the employability test score, but no other inference for salary can be extracted from the data, perhaps there is some non-linear relationship that the correlation matrix cannot capture.

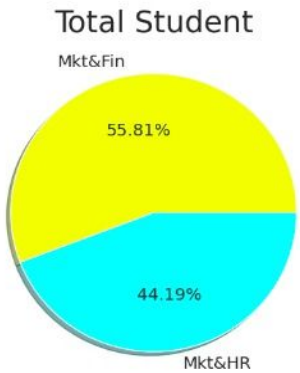


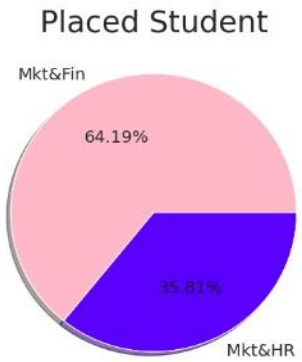
Observation: Placement Disparities Among MBA Specializations

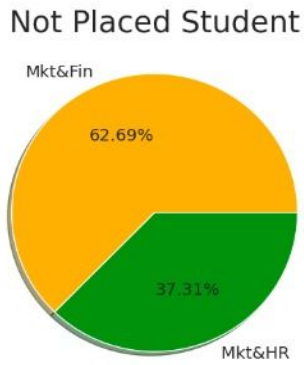
The analysis of the dataset revealed that students with specializations in Marketing and Finance had a higher number of placed students relative to their representation in the total student population.

Analysis: This suggests that these specializations are in high demand in the industry.
On the other hand, students with specializations in Marketing and HR had a lower number of placed students compared to their representation in the total student population, indicating that these specializations may not be as much in demand in the industry.

The findings suggest that students with specializations in Marketing and Finance may have a competitive edge in the job market compared to those with specializations in Marketing and HR. make it more concise





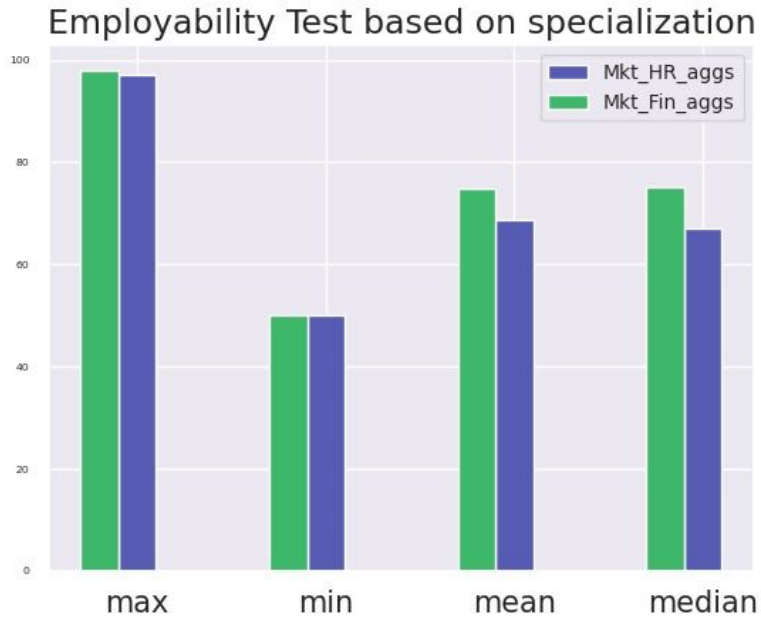


Observation: Employability Test Percentage based degree Specialization

Observation 1: The etest_p (Employability Test percentage) of Marketing and Finance students is higher compared to Marketing and HR students.

Observation 2: Marketing and Finance students received better job offers in general compared to Marketing and HR students.

Analysis: Based on the given observations, we can analyze that Marketing and Finance students have a better employability rate and also received better job offers compared to Marketing and HR students.

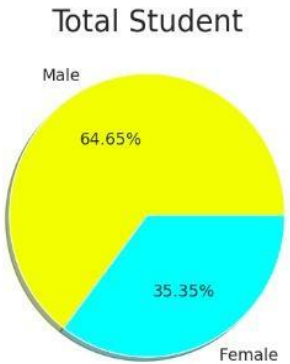


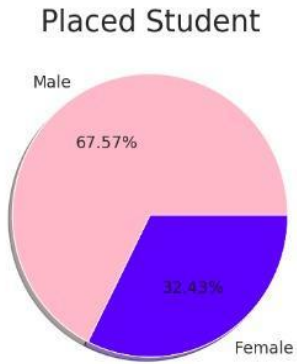
Observation: Analysis of Placements based on Gender and Total Population Representation

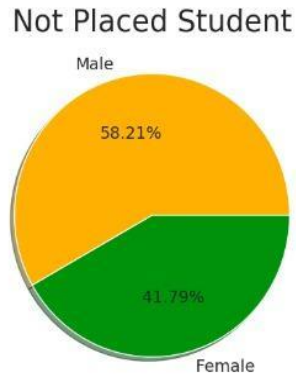
The placement rate is higher for students who have a higher representation in the total population, while female students have a lower placement rate relative to their representation in the total population.

Analysis: From the above pie charts we can conclude the following points-

- Students who have have a higher number of placed students in the aspect of their representation in the total population.
- Female Students have a lower number of placed students in the aspects of their representation in the total population. Their max salary is also 44.61% lower than Male.







Observations: Salary Analysis based on degree specializations

The mean, min, and median salaries were approximately the same for both Marketing and Finance students as well as Marketing and HR students,

However the maximum salary for Marketing and Finance students is significantly higher than Marketing and HR students.

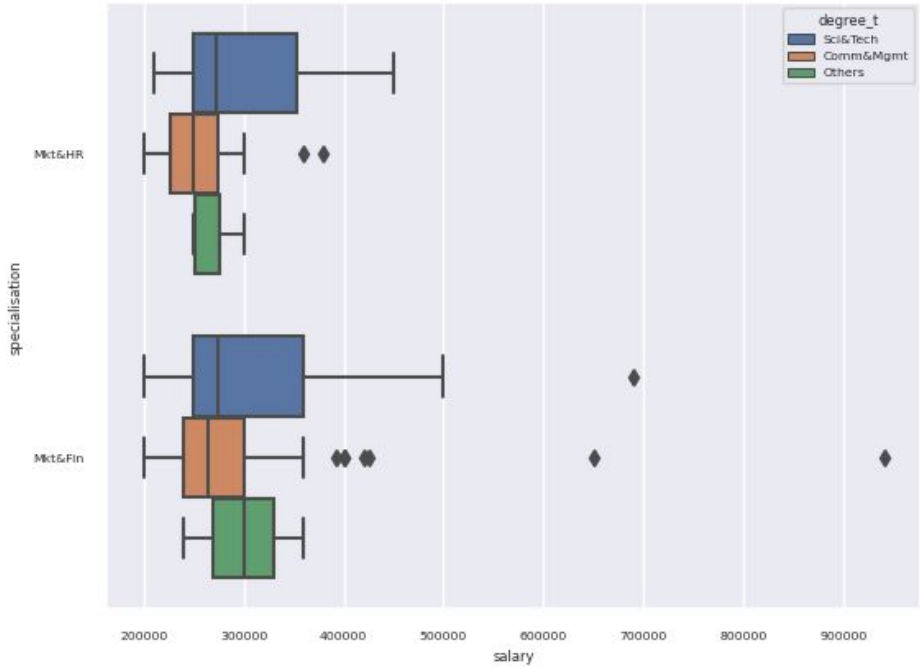
The significantly high outliers are responsible for the slight increase in mean salary for the Marketing and Finance students.

Observations: Salary Analysis based on undergraduate specializations

It is evident from the box plots that students with a specialization in Science and Technology in their undergraduate studies had a higher overall salary compared to students who pursued other degrees

Analysis:

One possible explanation for this correlation could be that Science and Technology degrees provide individuals with specialized skills and knowledge that are in high demand in the job market.



Dataset Analysis Link

Team Name: TheDataDorks
Link: https://www.kaggle.com/anuvabsen123/thedatadorks

Revelation 23