Analyzing Factors Affecting College Student Placement

Khushi Patel

MSBA 305: Business Intelligence and Decision Support

Instructor: [Instructor Name]

Date: August 2025

# Introduction

The transition from college to professional employment is a critical milestone for students. Understanding the factors that influence student placement outcomes can help educational institutions better prepare students for successful careers. This study investigates how academic performance, demographic attributes, and work experience contribute to college student placements. Using Python-based data analysis, the project explores patterns in the data and builds predictive models to estimate the likelihood of placement.

# Data Collection

The dataset used in this study is sourced from Kaggle and is titled 'College Student Placement Dataset.' It consists of student records containing academic and demographic information along with placement status. The features include gender, secondary and higher secondary education percentages (SSC%, HSC%), degree percentage, MBA percentage, specialization (e.g., Marketing, Finance), work experience, and placement outcome (Placed/Not Placed). These attributes provide a comprehensive view of how different factors affect placement results.  
  
Before analysis, the dataset underwent preprocessing to handle missing values, convert categorical variables into numerical representations, and standardize formats. These steps ensured that the data was clean, consistent, and ready for statistical analysis and modeling.

# Results of Data Analysis

The analysis began with descriptive statistics to summarize student characteristics and placement outcomes. Exploratory Data Analysis (EDA) was then conducted to visualize relationships between factors such as degree percentage, MBA performance, gender, and work experience with placement status.  
  
Correlation analysis was used to examine how strongly numerical features were related to placement. Statistical testing provided further insights into significant differences between placed and non-placed students.  
  
Two predictive modeling techniques were implemented: Logistic Regression and Decision Tree Classification. Model performance was evaluated using accuracy, precision, recall, and F1-score. The results revealed that academic performance and work experience were strong predictors of placement, while demographic variables such as gender had a less significant effect.

# Analysis and Interpretation

The findings indicate that higher SSC%, HSC%, and degree scores are positively associated with higher placement rates. MBA performance also contributes significantly, with students achieving strong MBA percentages being more likely to be placed. Work experience was identified as one of the most influential factors, demonstrating that practical exposure enhances employability.  
  
The logistic regression model achieved reliable accuracy, indicating its suitability for predicting placement outcomes. The decision tree model provided interpretability by highlighting which factors most strongly determined placement decisions. Together, these models emphasize that academic excellence combined with work experience is a key driver of successful placement.

# Conclusion

This study analyzed the factors affecting college student placements using statistical analysis and predictive modeling. The results suggest that academic performance, MBA specialization, and prior work experience significantly impact placement success. The logistic regression and decision tree models both demonstrated effectiveness in predicting placement outcomes, though with varying strengths.  
  
These insights can guide educators and policymakers in designing student development programs that emphasize both academic rigor and experiential learning. Future studies could expand on this analysis by incorporating larger datasets across multiple institutions and testing additional machine learning models.

# References

Islam, S. (2023). College Student Placement Factors Dataset. Kaggle. https://www.kaggle.com/datasets/sahilislam007/college-student-placement-factors-dataset  
  
OpenAI. (2025). ChatGPT (GPT-5) [Large language model]. OpenAI. https://chat.openai.com  
  
[Add any additional references here following APA guidelines]

# Appendix

The full Python code used for this project is provided in the accompanying Jupyter Notebook (HTML export). The appendix includes code for data preprocessing, visualization, correlation analysis, logistic regression, and decision tree modeling.