Education & Career Success

Tran Dang Huy, Sia Chawla Chawla, Jiaqi Xu

Table of contents

1 Executive summary

(Maximum of 4 sentences)

2 Introduction

(Maximum 10 sentences)

Graduates today face many questions about how their experiences influence their careers. Grades alone may not determine who gets the best job opportunities or feels happiest at work. This report explores two main questions: Which student experiences are linked to receiving more job offers? And what makes graduates feel satisfied with their careers and achieve a good work-life balance?

To answer these questions, we use a dataset of 5,000 recent graduates from Kaggle. It includes information about students' academic background, personal demographic, and career-related outcomes. Rather than testing predefined theories, this project takes an open-ended, pattern-oriented approach. The aim is to explore which types of experiences appear most consistently linked to job outcomes and personal satisfaction.

```
Rows: 5,000
Columns: 20
$ Student_ID            <fct> S00001, S00002, S00003, S00004, S00005, S00006, ~
$ Age                   <int> 24, 21, 28, 25, 22, 24, 27, 20, 24, 28, 28, 25, ~
$ Gender                <fct> Male, Other, Female, Male, Male, Male, Male, Mal~
$ High_School_GPA       <dbl> 3.58, 2.52, 3.42, 2.43, 2.08, 2.40, 2.36, 2.68, ~
$ SAT_Score             <int> 1052, 1211, 1193, 1497, 1012, 1600, 1011, 1074, ~
$ University_Ranking    <int> 291, 112, 715, 170, 599, 631, 610, 240, 337, 138~
$ University_GPA        <dbl> 3.96, 3.63, 2.63, 2.81, 2.48, 3.78, 3.83, 2.84, ~
$ Field_of_Study        <fct> Arts, Law, Medicine, Computer Science, Engineeri~
$ Internships_Completed <int> 3, 4, 4, 3, 4, 2, 0, 1, 2, 1, 2, 2, 2, 0, 1, 3, ~
$ Projects_Completed    <int> 7, 7, 8, 9, 6, 3, 1, 5, 3, 5, 7, 2, 0, 4, 2, 5, ~
$ Certifications        <int> 2, 3, 1, 1, 4, 2, 3, 5, 0, 3, 5, 3, 5, 3, 3, 2, ~
$ Soft_Skills_Score     <ord> 9, 8, 1, 10, 10, 2, 3, 5, 5, 10, 8, 2, 2, 8, 1, ~
$ Networking_Score      <ord> 8, 1, 9, 6, 9, 2, 3, 1, 5, 2, 1, 9, 9, 6, 8, 9, ~
$ Job_Offers            <int> 5, 4, 0, 1, 4, 1, 2, 2, 2, 0, 5, 5, 2, 2, 1, 3, ~
$ Starting_Salary       <dbl> 27200, 25000, 42400, 57400, 47600, 68400, 55500,~
$ Career_Satisfaction   <int> 4, 1, 9, 7, 9, 9, 7, 2, 2, 4, 9, 7, 9, 4, 9, 7, ~
$ Years_to_Promotion    <int> 5, 1, 3, 5, 5, 2, 4, 3, 2, 2, 1, 4, 4, 3, 3, 4, ~
$ Current_Job_Level     <ord> Entry, Mid, Entry, Mid, Entry, Entry, Mid, Entry~
$ Work_Life_Balance     <ord> 7, 7, 7, 5, 2, 8, 3, 3, 2, 2, 2, 6, 8, 3, 6, 3, ~
$ Entrepreneurship      <fct> No, No, No, No, No, Yes, No, No, No, No, No, Yes~
```

3 Methodology

(Maximum 300 words; Should include a figure and a table and those must be referenced in the text and have adequate captions)

This study uses an exploratory approach to investigate how students' academic and experiential attributes may relate to early career outcomes. The dataset includes 5,000 graduate records collected from Kaggle.

We began by examining the structure of the input variables. As shown in Table 1, the dataset contains a mix of continuous (e.g., `SAT_Score`), ordinal (e.g., `Certifications`), and categorical (e.g., `Field_of_Study`) features, with a wide range of unique values. This diversity limits the applicability of simple linear models and supports the use of flexible methods that can accommodate mixed data types and non-linear relationships.

We also analysed the outcome variables: `Job_Offers`, `Career_Satisfaction`, and `Work_Life_Balance`. As visualised in Figure 1, all three outcomes are relatively well-distributed and show minimal skewness.

Given that `Job_Offers` is a count variable, we selected random forest regression, a non-parametric model capable of capturing complex relationships without strong distributional assumptions. For the ordinal outcomes, `Career_Satisfaction` and `Work_Life_Balance`, we applied ordinal logistic regression to model the ordered nature of the response while preserving interpretability.
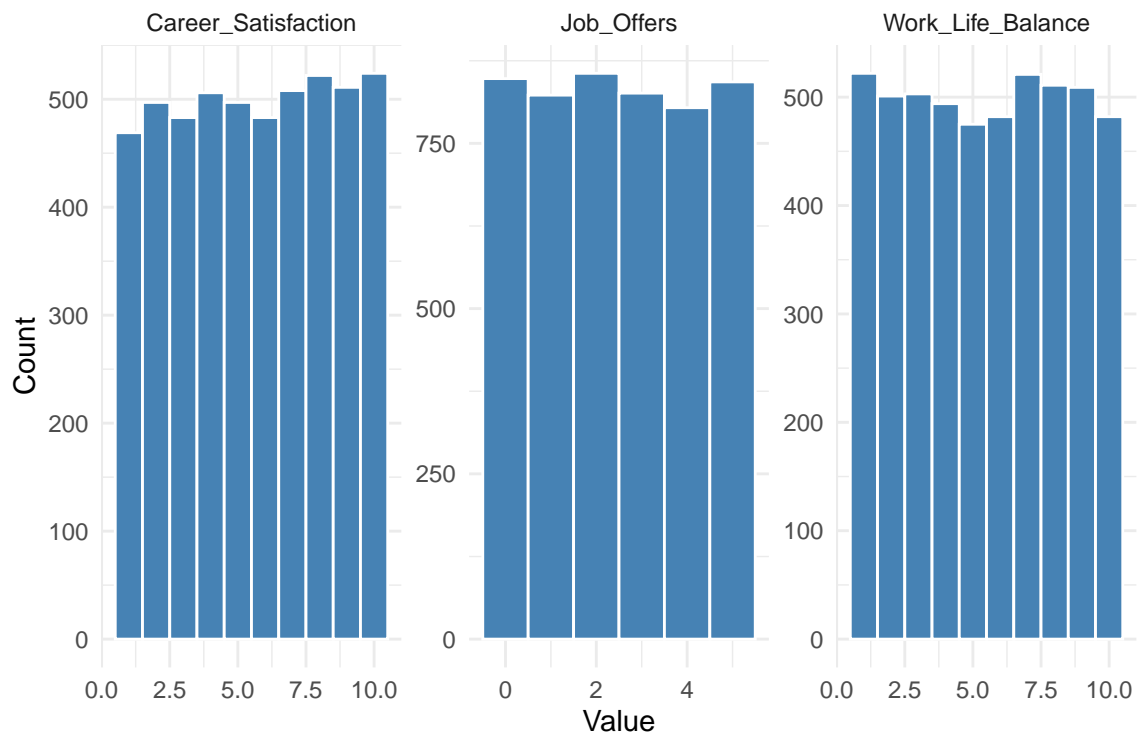
Figure 1: Distributions of Outcome Variables

4 Results

(Maximum 200 words. Should include either a figure or a table.)

5 Discussion, conclusion and recommendations

6 Reference section

(Include at least 1 reference.)