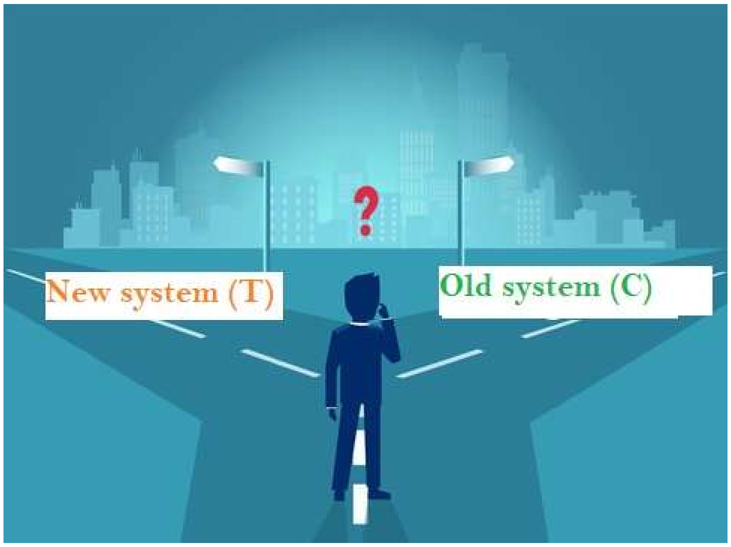

Data Project on FBI Interview Pipeline

PROJECT REPORT ON DATASET - 5

Uma Rajagopalan | DS-853 | Fall-21

Executive Summary

The purpose of this report is to provide insights on which of the two systems New (T) or Old (C) is better for the FBI recruitment process.

We find the best system by conducting Exploratory Data analysis on the dataset and verifying the results with hypothesis test. To get the overall picture, this process is repeated on comprehensive as well as granular level i.e., for each individual department.

In the comprehensive level, I find that the Old system (C) is better whereas in the granular level, the New system (T) is better. But, since every applicant ends up in a single department, the granular data analysis is a better approach on choosing between the two systems.

It is seen that the New system (T) is the best for the FBI recruitment process.

Introduction

The FBI Pipeline dataset consists of data to compare between 2 systems used for their recruitment process: Old (C) and New (T). The candidates are assigned to both these systems and the objective is to determine which system is better for the recruitment process. The system chosen should have more candidates successfully completing the 5 steps in the recruitment process which can be seen in Fig 1.

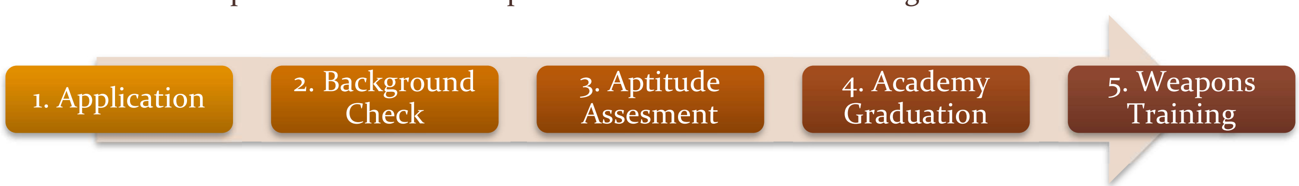

Fig 1: Steps in the FBI Recruitment Process

Fig 2 represents the recruitment lifecycle of one of the recruits from the dataset.

Fig 2: Order of the steps involved in the recruitment process of FBI

There are 6 departments within the FBI where the experiment is conducted.

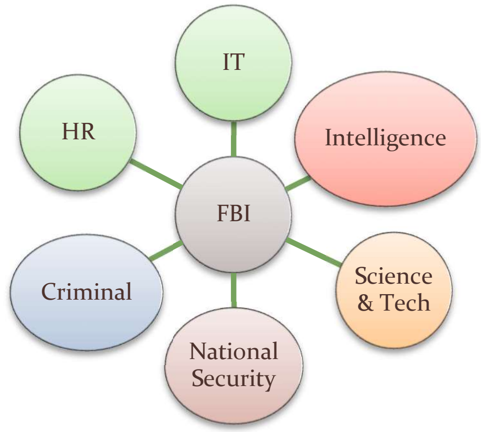

Fig 3: Various departments in the FBI

Methodology

The objective of this experiment is to conduct an **A/B Test** between the two systems of recruitment. The applicants in the dataset are already divided between the two systems: old system (C) and new system (T). We will conduct an Exploratory data analysis (EDA) on our dataset and prove the results of the EDA with Hypothesis Testing to determine which of the two systems is better.

We will conduct the process of A/B Test for the Overall Recruitment Process in the FBI and then further classify the experiment to department-wise and derive our conclusion.

Exploratory Data Analysis on the Overall Recruitment Process

To perform the EDA, we compare the percentage of candidates that completed weapons training which is the final step in the recruitment process of FBI in each of the two systems.

The completion percentage is calculated by the formula:

Completion Percentage = Count of candidates that completed Weapons Training in a system/ Count of candidates that applied in a system

Comparison of Candidates that completed Weapons Training Step of Recruitment Overall

Fig 4: Comparison of Percentage of Candidates that completed the final step of recruitment process

It is evident from Fig 4 that the percentage of candidates that completed Weapons Training is greater when they use the **Old system (C)** than the New system (T).

Old System (C) > New system (T)

Let's conduct hypothesis testing to support our Exploratory Data Analysis.

Hypothesis Test on the Overall Recruitment Process

We determined that the Old system performed better than the New system through the EDA. Now, let's verify this statement with hypothesis testing.

Step 1: Determine the Null and Alternate Hypothesis

Null Hypothesis is a statement that shows no difference between two groups.

Null Hypothesis Ho: New system is just as good as the Old System: P(T) = P(C)

Alternate hypothesis is the statement that shows there is a difference between the two groups.

Alternate Hypothesis Ha: New system is lower than the Old System: P(T) < P(C)

Step 2: Set the significance level (α)

I am setting the significance level as 0.05

This means that there is a 5% chance that I accept the alternate hypothesis when null hypothesis is true

Step 3: Calculate the test statistic and corresponding p value

There are four parameters required to calculate the p-value

Number of Trials(n): The total count of applied candidates in the new system (T) = 4416

Sample Size(S): The count of candidates that completed weapons training in the new system (T) = 1622

Probability(P): The percentage of the candidates that completed weapons training in the old system (C) = 41%

The p-value describes the probability of obtaining the test statistic of an extreme outcome when null hypothesis is true.

Since this is a left tailed test, p-value = BINOM.DIST(n, S, P, TRUE) = BINOM.DIST (1622, 4417, 41%, TRUE) = **0.00000036%**

Step 4: Draw conclusions

We can draw conclusions based on the following conditions:
1. If P-value <= significance level (α) => Reject your null hypothesis in favor of alternative hypothesis. The result is statistically significant.
2. If P-value > significance level (α) => Fail to reject your null hypothesis. The result is not statistically significant.

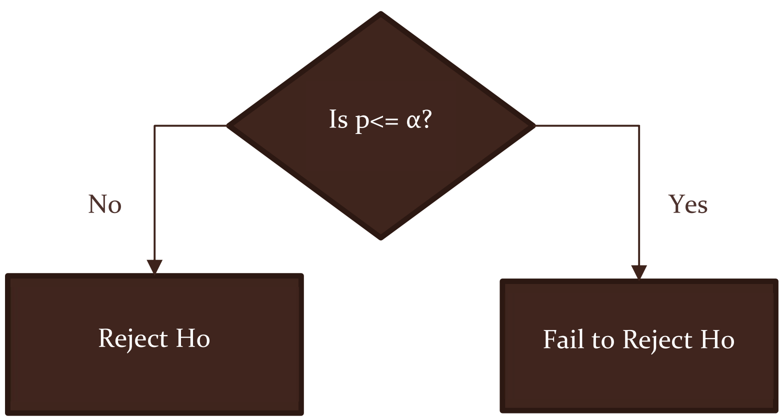

Fig 5: Condition to test if the test statistic is significant

We have our p-value (0.00000036%) < significance level (α = 5%).

Thus, we can reject null hypothesis and conclude alternate hypothesis.

This proves that the New system (T) is lower than the Old system (C).

We can conclude that the Old system (C) is better at the comprehensive level.

EDA for Individual Departments

Now, let's include the departments within FBI and compare the completion percentages.

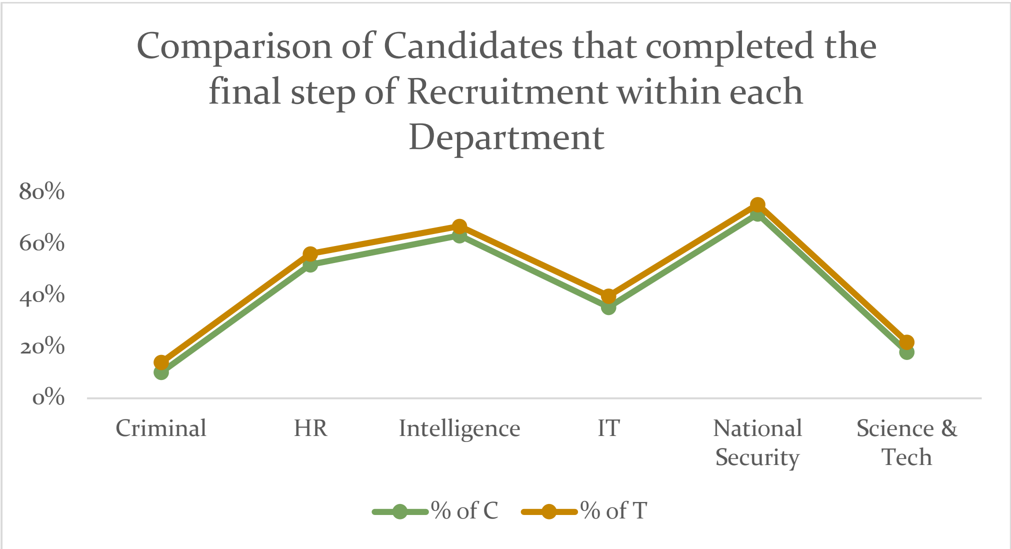

Fig 6: Comparison of Percentage of Candidates that completed the final step of recruitment process (Weapons Training) within departments

It is evident from Fig 6 that the **New system (T)** performs better than the Old system (C).

New System (T) > Old system (C)

Let's conduct hypothesis test to support our Exploratory Data Analysis.

Hypothesis Testing for Individual Divisions

We determined that the New system (T) is performing better than the old system (C) with our EDA.

Step 1: Determine the Null and Alternate Hypothesis

Null Hypothesis Ho: New system (T) is similar to the Old system (C) for all individual departments: P(T) = P(C)

Alternate hypothesis Ha: New system (T) is better than the Old system for each of the individual departments: P(T) > P(C)

Step 2: Set the significance level (α)

I am setting the significance level as 0.05

Step 3: Calculate the test statistic and corresponding p-value

Number of Trials(n): The count of candidates that completed weapons training step with the New system (T)

Sample Size(S): The count of candidates that applied with the New system (T)

Probability(P): Percentage of candidates that completed the weapons training step in the Old system (C)

Since this is a right tailed. Thus, p-value = 1-BINOM.DIST(n, S, P, TRUE)

Divisions	% Completed in C (P)	Count completed in T (n)	Count applied in T (S)	P - value
Criminal	10%	174	1244	0.001%
HR	52%	324	577	1.751%
Intelligence	63%	192	287	8.558%
IT	35%	253	637	1.006%
National Security	72%	442	587	2.144%
Science & Tech	18%	237	1084	0.040%

Fig 7: Table depicting the p-values for various departments within the FBI

Step 4: Draw conclusions

From the above table (Fig 7), we compare the p-values to significance level (α) based on the conditions in Fig 5.

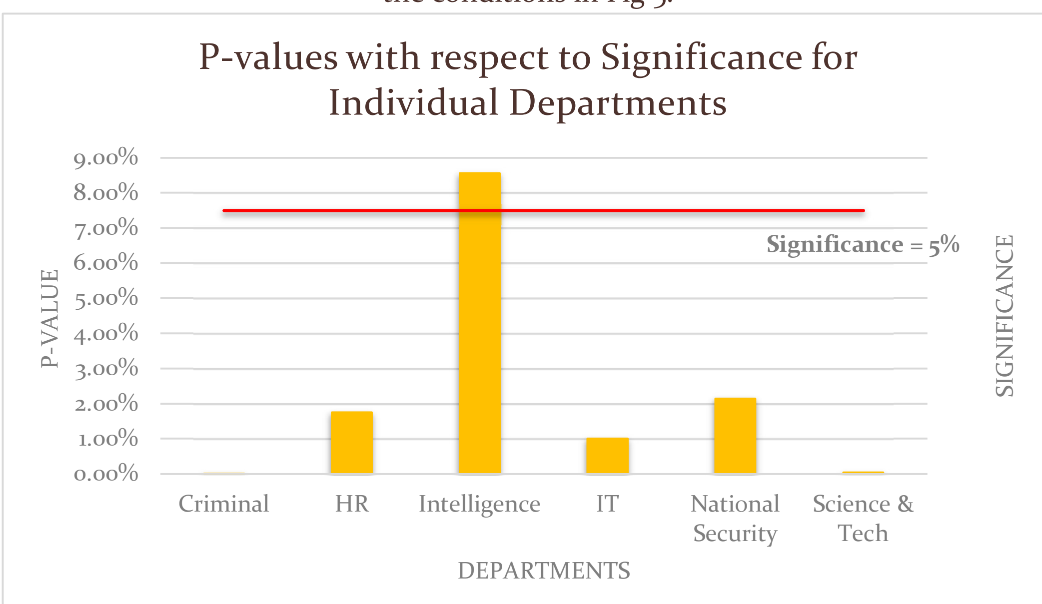

Fig 8: Chart depicting the conclusion from the hypothesis test of individual departments

From Fig 8, we can see that the p-values are less than 5% (Significance) for all departments other than Intelligence. We can reject null hypothesis and conclude that the New system (T) is better for all departments other than Intelligence.

However, we know that the sample size directly effects p-values. A sample size that is too small can result in a failure to identify a difference when one truly exists.

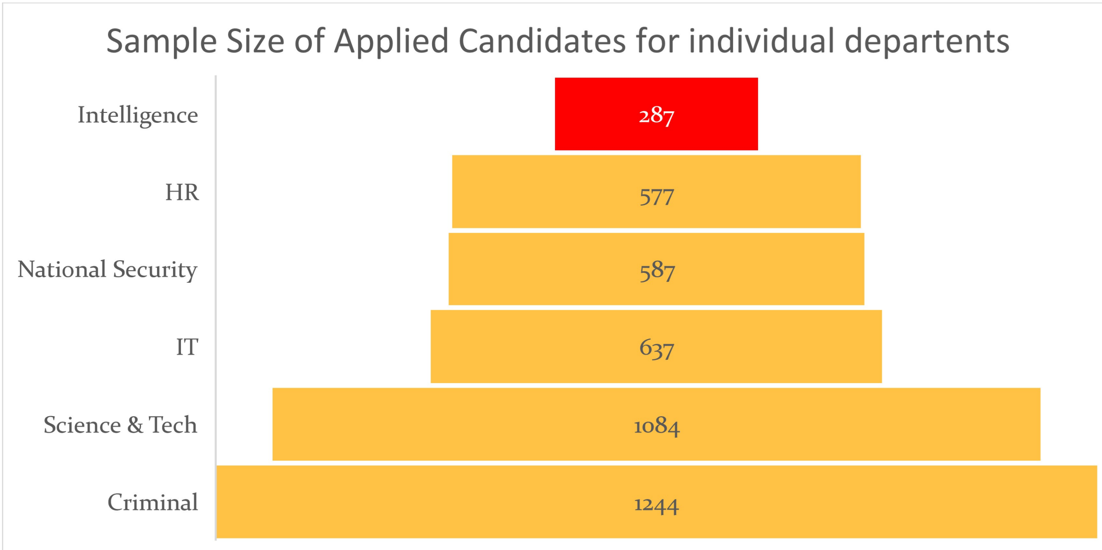

Fig 9: Chart depicting the sample sizes of the candidates applied in various departments

From Fig 9, we can see that Intelligence has the least number of applied candidates that could have resulted in the higher p-value.

Thus, we can conclude that **New system (T)** is better for each individual department

Conclusion

The New system (T) is better for the FBI Recruitment process.

We performed data analysis of the dataset at a comprehensive as well as granular level for every individual department. The results of the granular level were however contradictory to that of the comprehensive level. But, since every applicant lands only in a single department in the FBI after selection, we prefer the granular method of data analysis over the comprehensive level.

Furthermore, the hypothesis testing done for each individual department proved that the New system (T) is better for all departments other than Intelligence. But, since Intelligence had the least sample size compared to the other departments, which effected its test statistic, we can conclude that the New system (T) works for Intelligence department as well and it is the best system to use in the FBI recruitment process.