Data Science Capstone: Final Report

Rishikesh Kakde

17th Oct 2022

OUTLINE

- Executive Summary
- Introduction
- Methodology
- Results
  - Visualization – Charts
  - Dashboard
- Discussion
  - Findings & Implications
- Conclusion
- Appendix

IBM Developer

SKILLS NETWORK

EXECUTIVE SUMMARY

- Data Collection via API, SQL, and Web Scraping

- Data Wrangling and Analysis

- Interactive Maps with Folium

- Predictive Analysis for each Classification Model

- Data Analysis along with the Interactive Visualizations

- Best Model for Predictive Analysis

IBM Developer

SKILLS NETWORK

INTRODUCTION

SpaceX takes pride in its cost efficient rockets. The reason their rockets are so cost efficient is because unlike other companies, SpaceX can reuse the first stage of the rocket.

In this report, we analyze the data collected by SpaceX on its Falcon 9 rocket landings. We use this data to make a predictive model, to predict if the first stage of a given rocket will land successfully or not.

We will explore the factors on which successful landing of a rocket depends on and the effect of each feature on the landing.
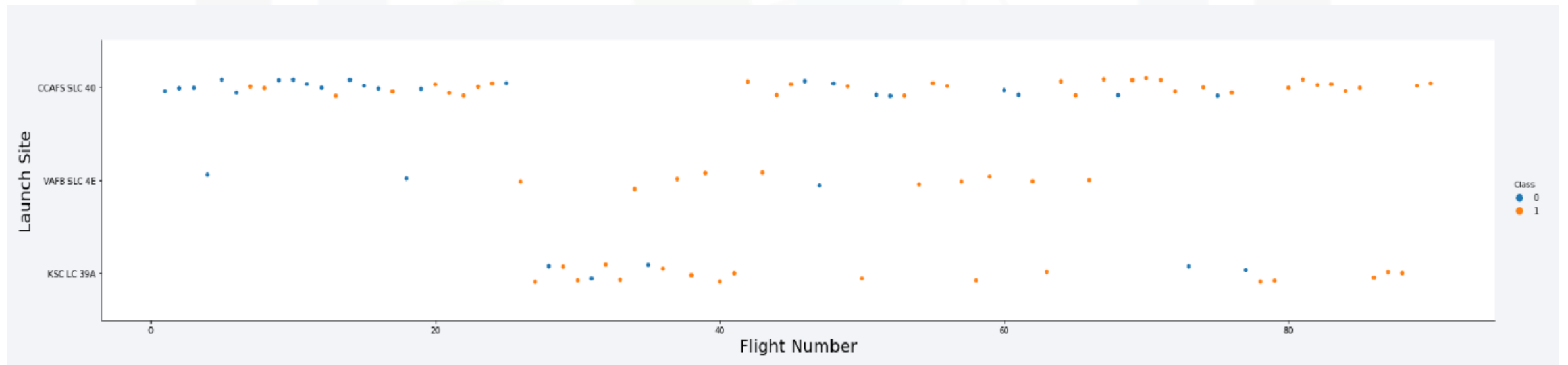
METHODOLOGY

- Data Collection using API (SpaceX REST) and Web Scraping (Wikipedia)
- Data Preprocessing and cleaning
- EDA using SQL and visualization
- Interactive Visualizations using Folium library
- Building Classification Models
- Model Evaluation
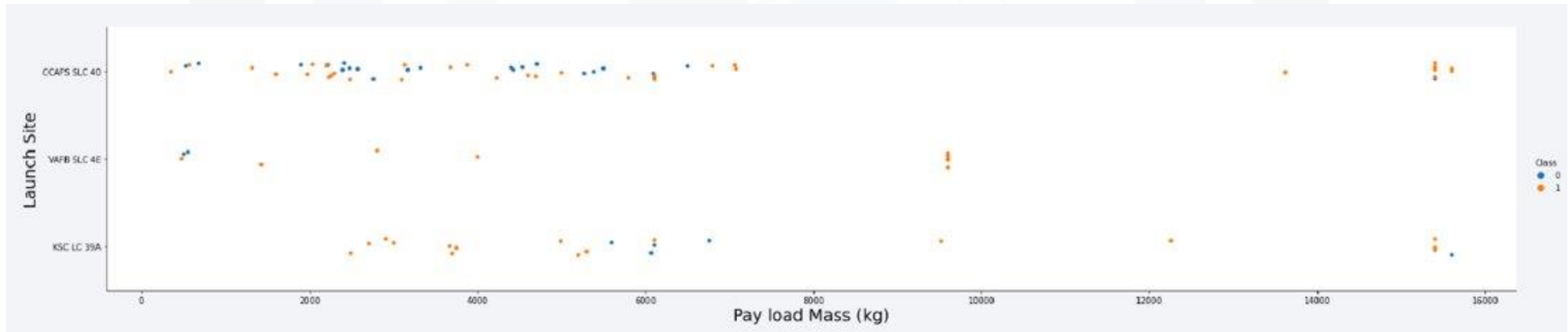
IBM Developer

SKILLS NETWORK

Results from EDA

Flight Number vs Launch Site

- The success rate shows a up-trend for each of the sites

Results from EDA

Payload vs Launch Site

- Heavy payload seems to be positively correlated with successful landing but too heavy payload has a negative impact on the landing.

Results from EDA

Success Rate vs Orbit Type

- This bar chart shows the different success rates associated with different orbit types

- The following orbits have the highest success rates:
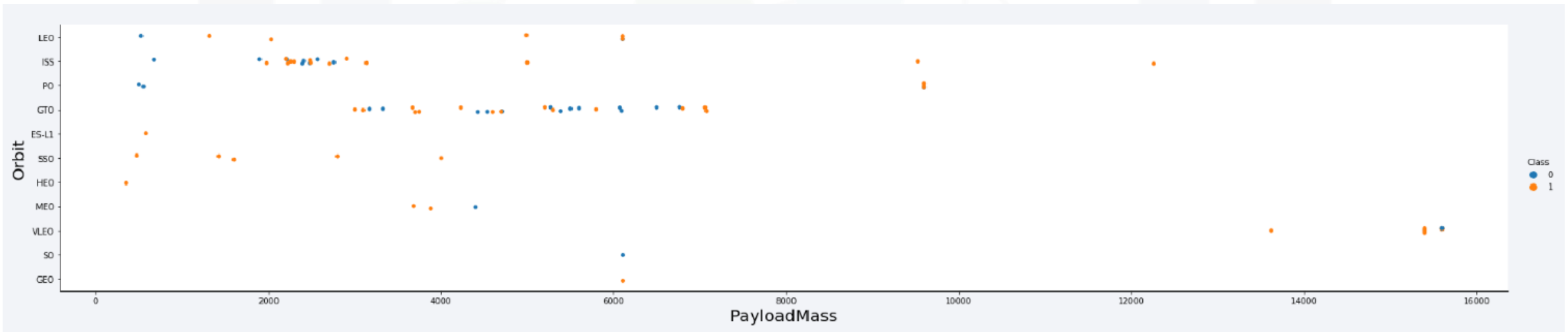 - ES-L1
  - GEO
  - HEO
  - SSO

IBM **Dev** loper

SKILLS NETWORK
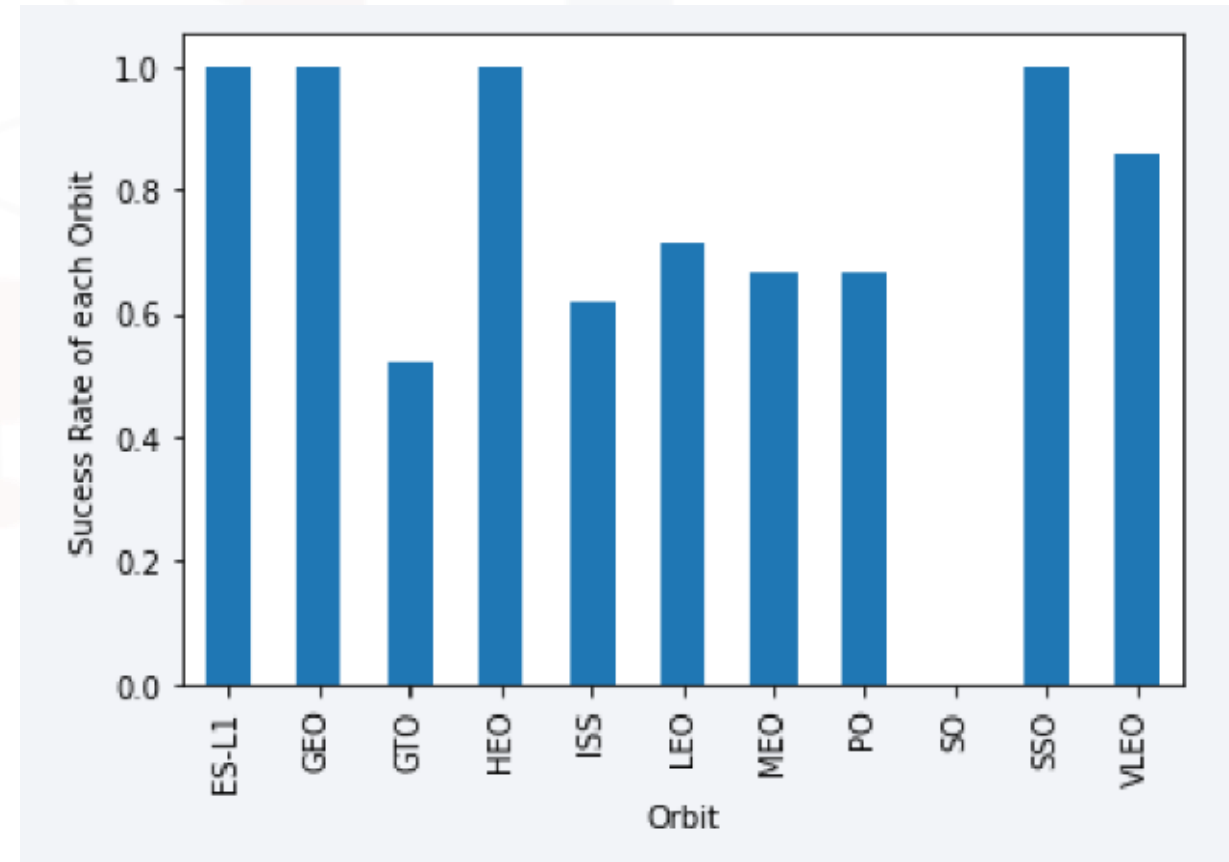
Results from EDA

Flight Number vs Orbit Type

- GTO seems to have no correlation between Success rate and number of flights.

- For LEO, the success rate increased with the number of flights
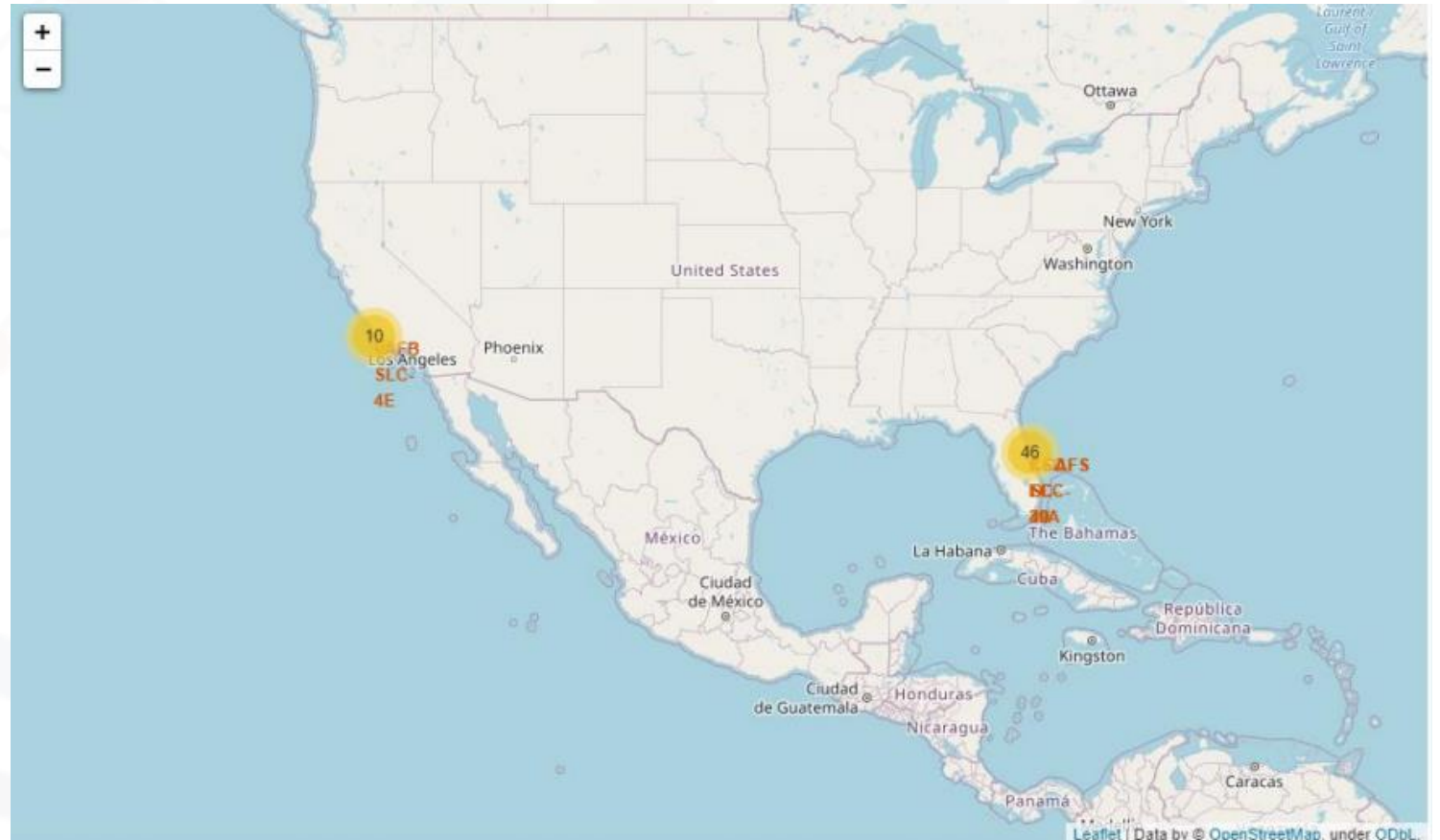
Results from EDA

Payload vs Orbit Type

- Payload in general is positively correlated with the LEO orbit
- But for GTO orbit, less payload helps in successful landing
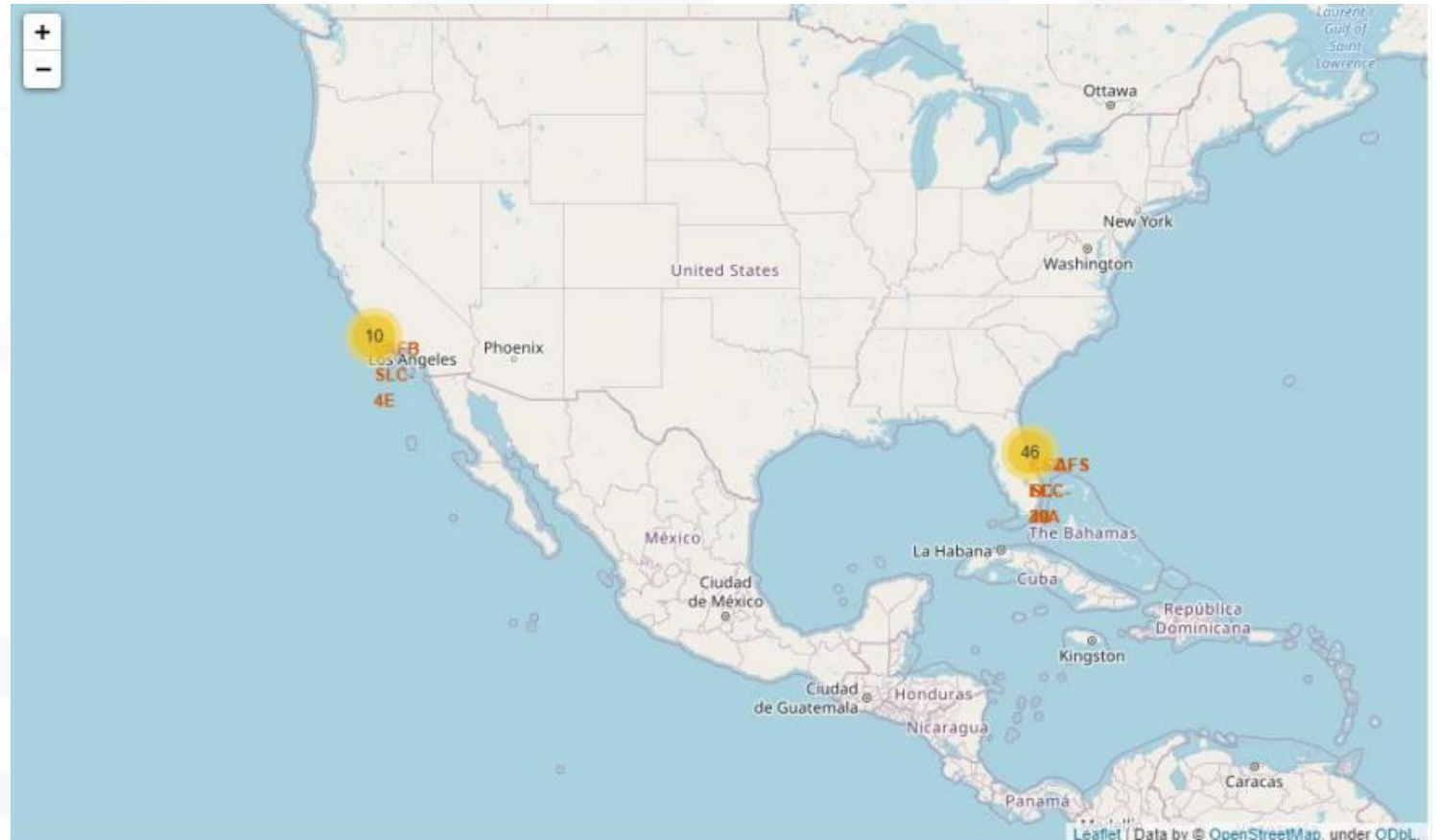
IBM Developer

SKILLS NETWORK

Folium Map

- The adjacent map shows the SpaceX launch sites

Folium Map

- The adjacent map shows the SpaceX launch sites

Dashboard

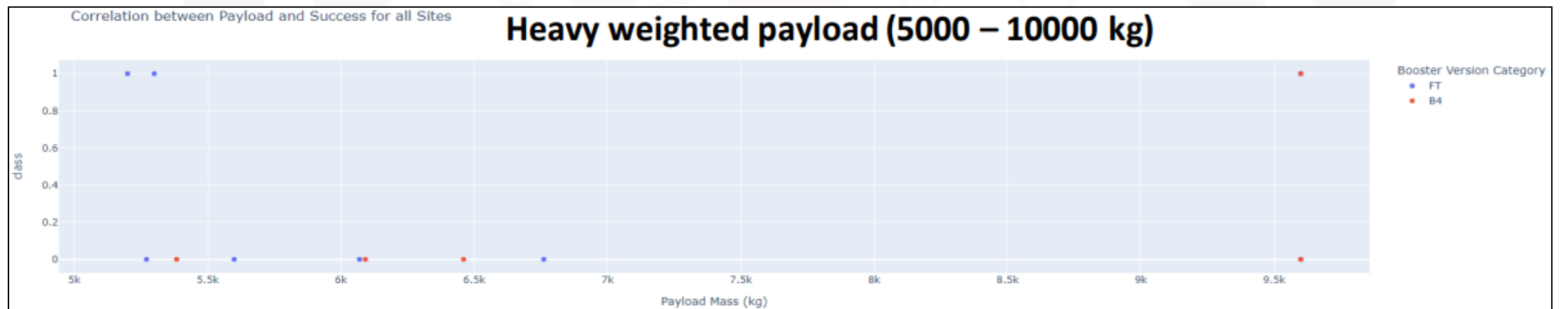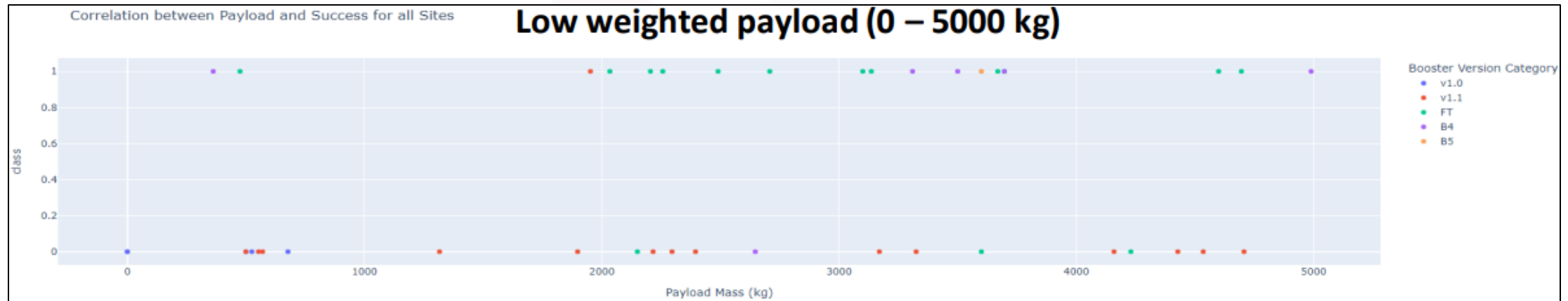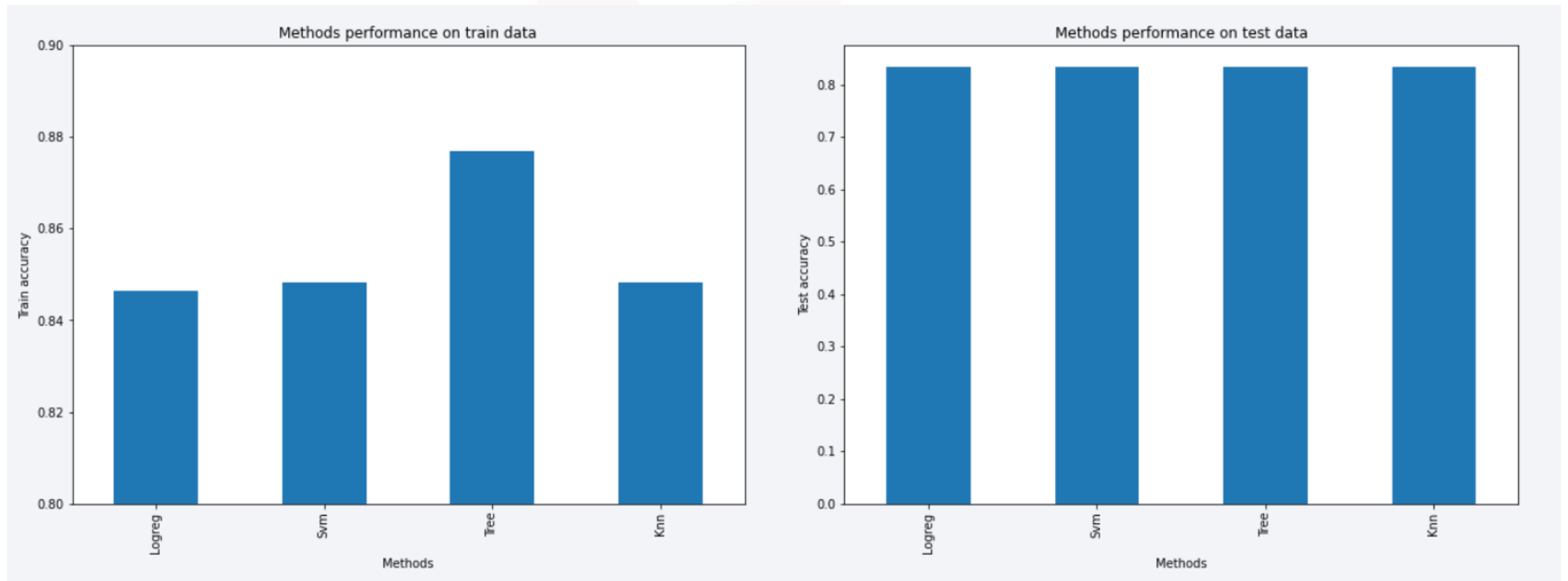Total Success Launches by Site

- KSC LC-39A
- CCAFS LC-40
- VAFB SLC-4E
- CCAFS SLC-40

41.7%
29.2%
16.7%
12.5%

Total Success Launches for Site KSC LC-39A

- 1
- 0

76.9%
23.1%

Dashboard

Low weighted payload (0 – 5000 kg)

Correlation between Payload and Success for all Sites

Heavy weighted payload (5000 – 10000 kg)

Correlation between Payload and Success for all Sites

Predictive Analysis: Classification Accuracy

- We select Decision Tree model as it performs equally well on test data and the best among all on the training data.

CONCLUSION

- The orbits with the best success rates are GEO, HEO, SSO, and ES-L1.

- Based on the orbit, the payload might positively or negatively affect the landing.

- Decision Tree model performed the best among all the classification models. It has the same accuracy as the other models on the test set but it was better than the other on the training set.

Thank You!