

The background is a deep purple space scene. In the top left is a large planet with horizontal purple and white stripes. To its right is a smaller planet with a ring. In the bottom left, an astronaut in a white suit floats with a coiled tether. In the bottom right is a large, cratered moon. The sky is filled with white stars and soft, wavy nebulae in shades of purple and blue.

APPLIED DATA SCIENCE CAPSTONE

Cody Oneil Harry

2023-08-24



TABLE OF CONTENTS

• Executive Summary	2
• Introduction	3
• Methodology	4
• Results	5-9
• Charts	
• Dashboard	
• Discussion	10
• Findings	
• Conclusion	11
• Appendix	12

EXECUTIVE SUMMARY

Applying data science skills as a Data Scientist for a private space launch company which includes recognizing the significance of gathering as much pertinent data as feasible, collecting data from diverse sources enhancing its quality through data wrangling becomes a focal point and delving into the processed data.

Guided by expertise, the exploration of captivating real-world datasets occurs collaboratively with the team. The opportunity arises to practice SQL skills for data querying and gaining valuable insights. Further comprehension of the data unfolds while employing rudimentary statistical analysis and data visualization methods. These tools allow for direct observation of relationships among variables. Data segmentation into groups characterized by categorical variables or factors reveals more intricate insights within the dataset, unlocking potential for even more riveting discoveries.

INTRODUCTION

In this capstone project, the goal is to predict the successful landing of the Falcon 9 first stage. SpaceX prominently features Falcon 9 rocket launches on its website, pricing them at 62 million dollars. In contrast, other providers charge upwards of 165 million dollars per launch. SpaceX's cost efficiency is largely due to its ability to reuse the first stage of the rocket. Therefore, by accurately predicting the first stage landing outcome, we can effectively estimate the launch cost. This predictive insight holds particular value when considering competitive bids against SpaceX for rocket launches.

Within this lab, the primary task involves gathering data from an API and ensuring its conformity to the correct format. The following excerpt illustrates both a successful launch and its respective outcome.

METHODOLOGY

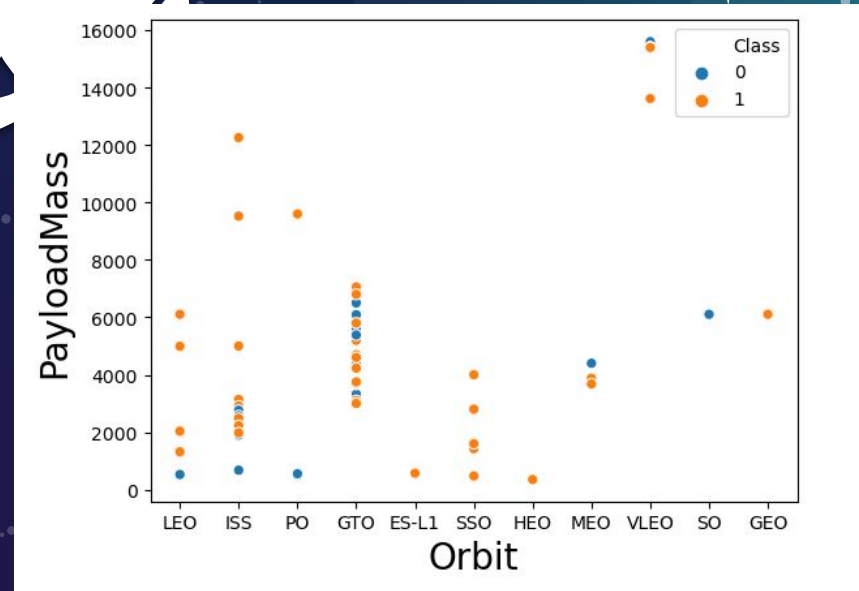
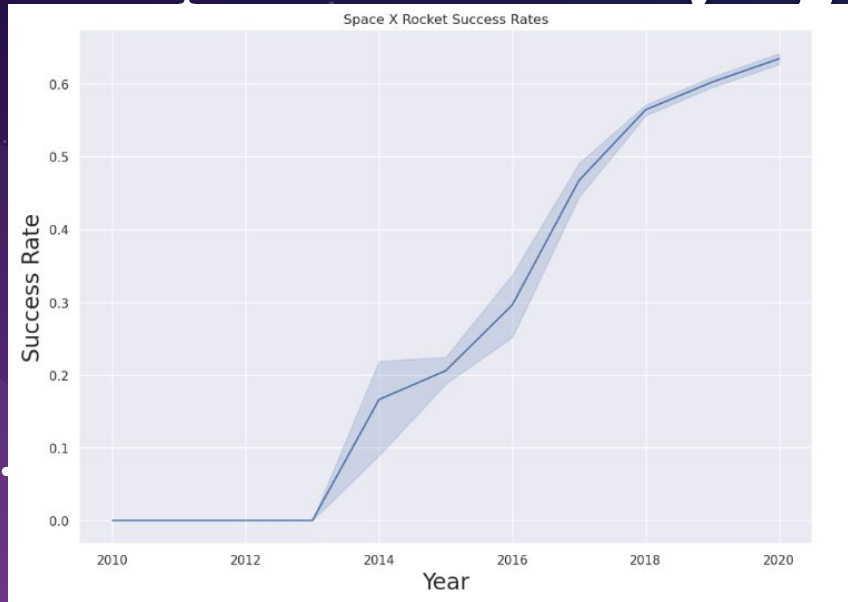
Firstly, our focus will be on conducting Exploratory Data Analysis (EDA) to uncover discernible patterns within the data. Additionally, we will ascertain the most appropriate label for training supervised models.

We worked with SpaceX launch data that is gathered from the SpaceX REST API. This API will give us data about launches, including information about the rocket used, payload delivered, launch specifications, landing specifications, and landing outcome.

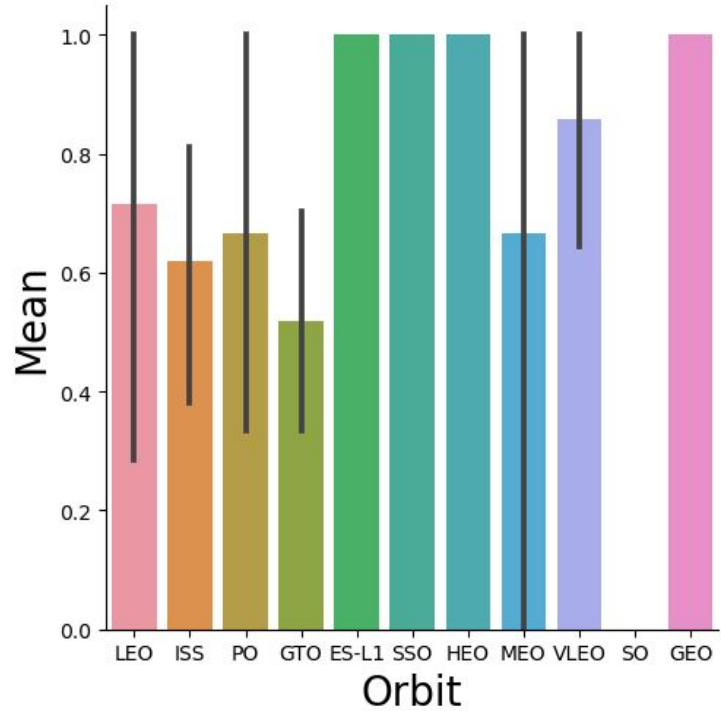
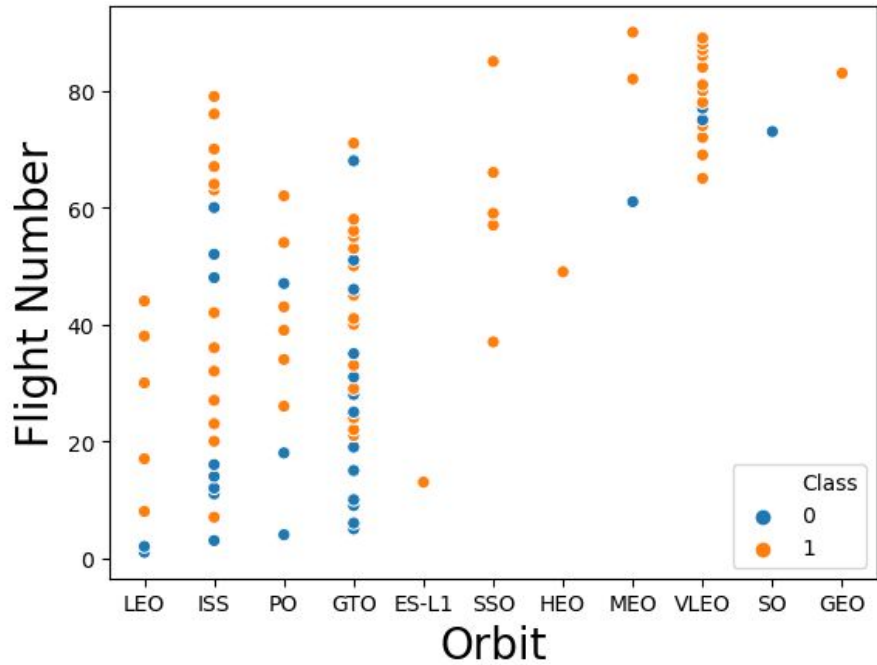
Contained within the dataset are multiple instances where the booster's landing was not successful. These cases encompass a range of scenarios. Sometimes, an attempted landing resulted in failure due to unforeseen events. For instance, when the outcome is labeled as "True Ocean," it signifies that the mission concluded with a successful landing within a specific region of the ocean. Conversely, when labeled as "False Ocean," it indicates an unsuccessful landing within a designated ocean region.

RESULTS

CHARTS



CHARTS - 2



DASHBOARDS

SpaceX Launch Records Dashboard

All Sites

Total Success Launches by Site

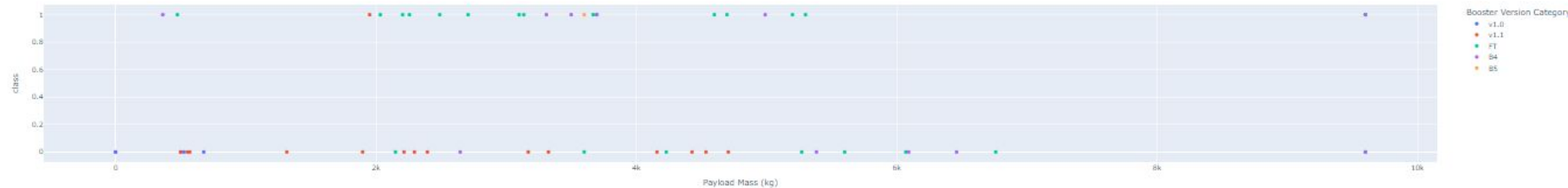


KSC LC-39A
CCAPS LC-40
VAFB SLC-4E
CCAPS SLC-4E

Payload range (Kg):



Correlation between Payload and Success for all Sites



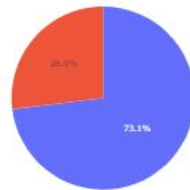
Booster Version Category
v1.0
v1.1
FT
B4
B5

DASHBOARDS +2

SpaceX Launch Records Dashboard

CCAFS LC-40

Total Success Launches for CCAFS LC-40



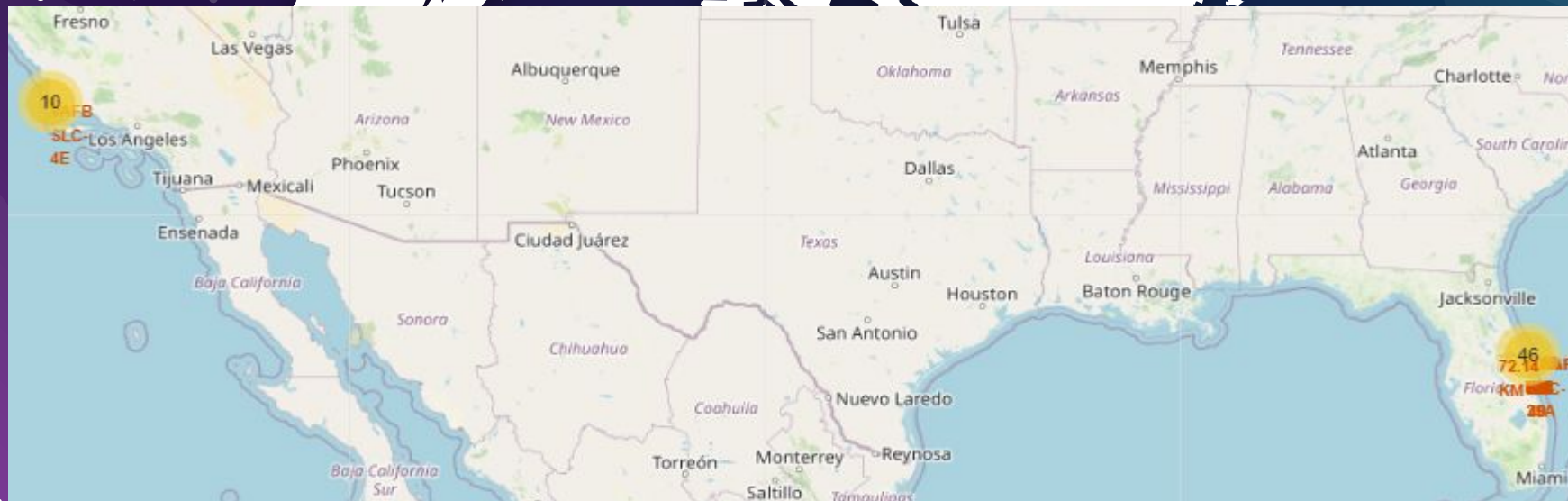
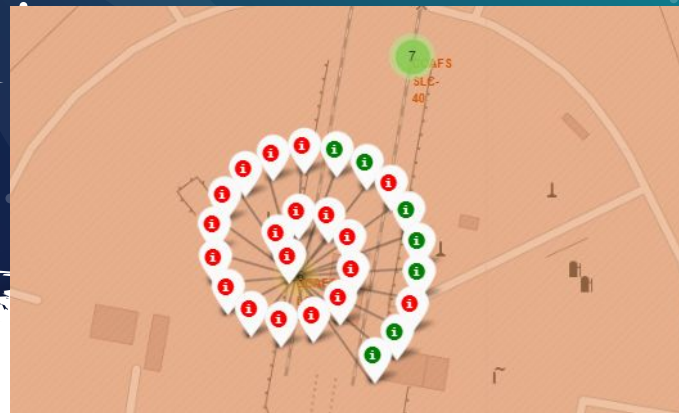
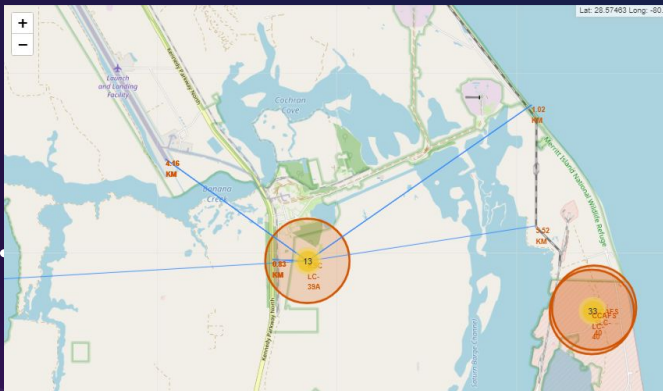
Payload range (Kg):



Correlation between Payload and Success for CCAFS LC-40



MAPS



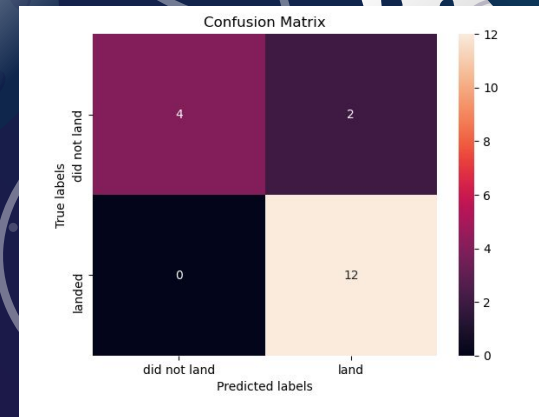
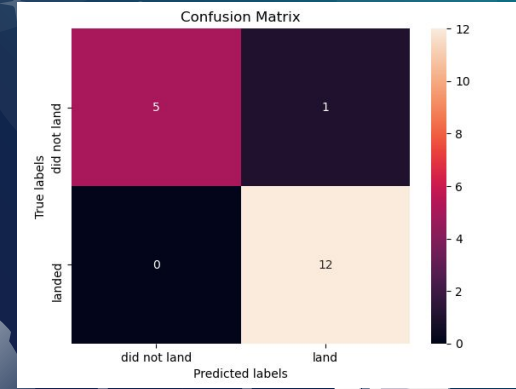
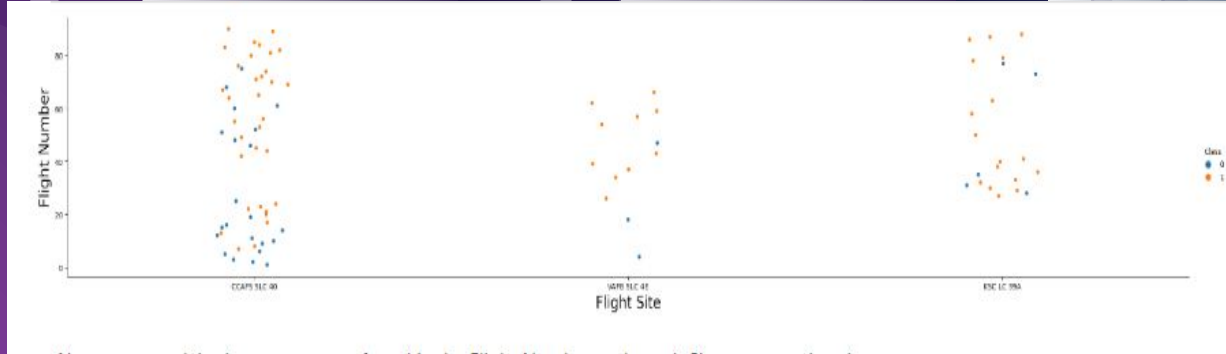
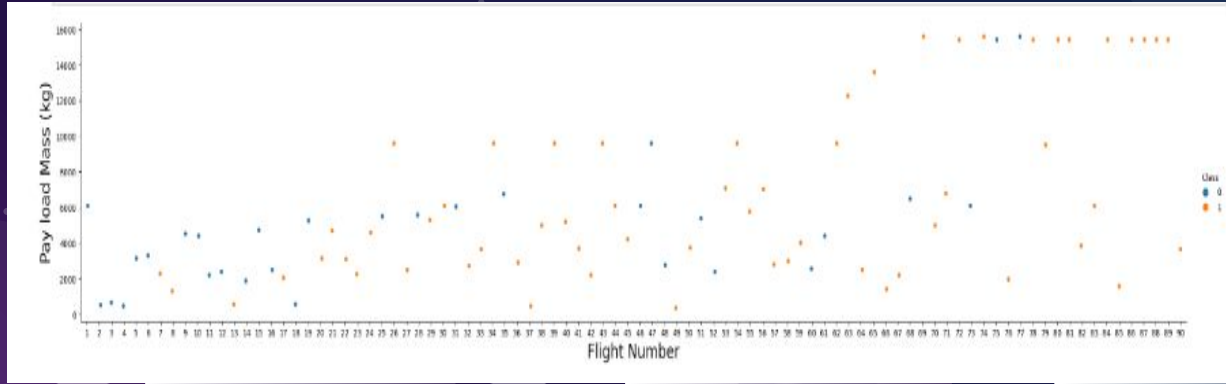
DISCUSSION

- The success rate has increased significantly over the years.
- The first successful landing outcome on ground pad occurred on December 22, 2015.
- CCAFS SLC-40 was the launch site that had the highest Falcon 9 first stage landing success rate (42.9%).
- The payload range from about 2,000 kg to 5,000 kg has the largest success rate.
- All models performed equally well except for the Decision Tree model which performed poorly relative to the other models.
- ES-L1, SSO, HEO and GEO orbits have no failed first stage landings.

CONCLUSION

- > The Tree Classifier Algorithm is the best for Machine Learning for this dataset
- > Low weighted payloads perform better than the heavier payloads
- > The success rates for SpaceX launches is directly proportional time in years they will eventually perfect the launches
- > We can see that KSC LC-39A had the most successful launches from all the sites
- > Orbit GEO,HEO,SSO,ES-L1 has the best Success Rate

APPENDIX





THANKS!

