Oreoluwa Olatunji

19th October, 2023

First Landing: A Competitive Edge In The Space Market

EXECUTIVE SUMMARY

To be able to predict prices of launches and compete with Space X, Space Y gave me the task of collecting, wrangling and analysing the data of various Space X launches from different sites over the years.

The aim of this project was to see if the first landing stage would land successfully after a first launch in order to be reused which would be able to help determine the cost to launch.

TABLE OF CONTENT

INTRODUCTION

- We simply want to predict the successful landing of the first landing stage.

- Things to consider:
- What factors affecting the landing?
- Can it be used to predict the prices of a launch?

METHODOLOGY

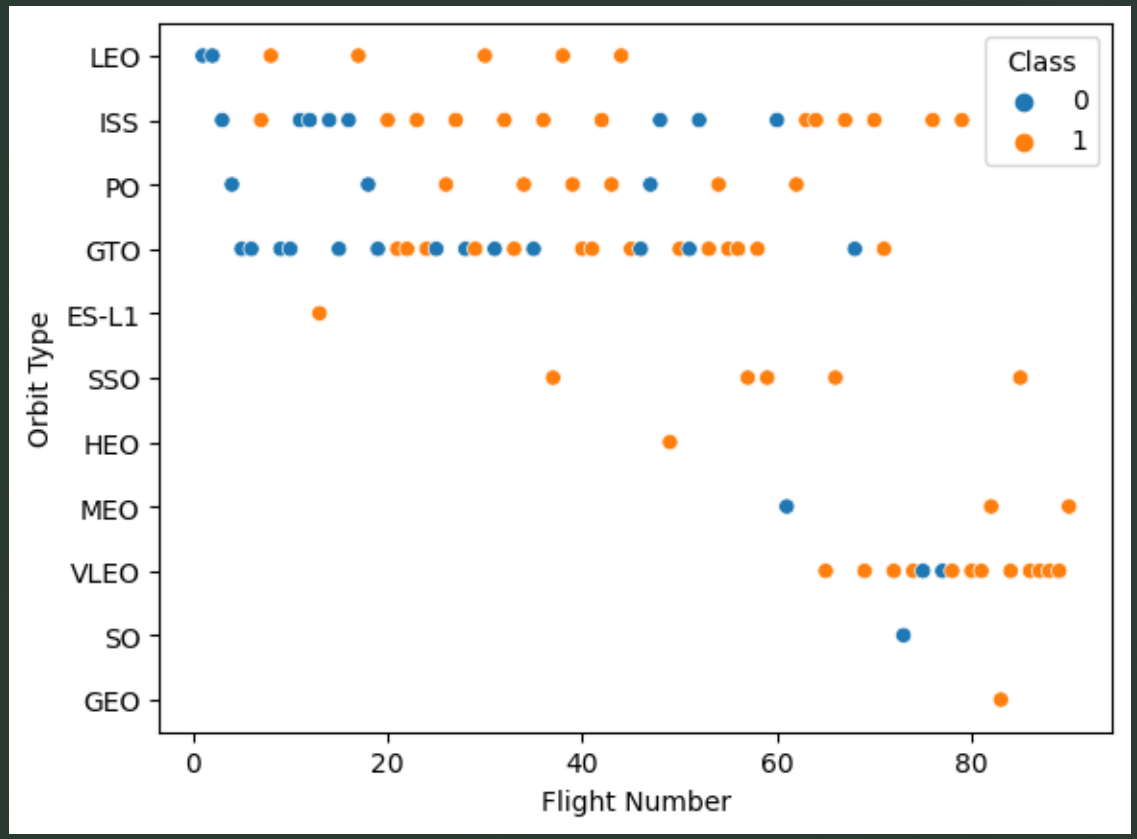
- Data Sources: SpaceX API and Wikipedia.
- Data Collection: The data used for this project was gotten from the SpaceX API and web scraped from Wikipedia using Python and converted to a csv file.
- Data Wrangling: Using python, the raw data in the csv format was converted into meaningful data which was used to carry out the project which was also saved as a csv file. An EDA was carried out to find patterns in the data using Pandas and numPy.
- Further EDA was done with visuals – Scatterplots, Bar Charts, Line Charts.
- Predictive Analysis: The Machine Learning pipeline included KNN, Logistic Regression, SVM and Decision Tree and the best of the lot was used to predict the success of the first landing.

RESULTS

EDA WITH VISUALISATIONS

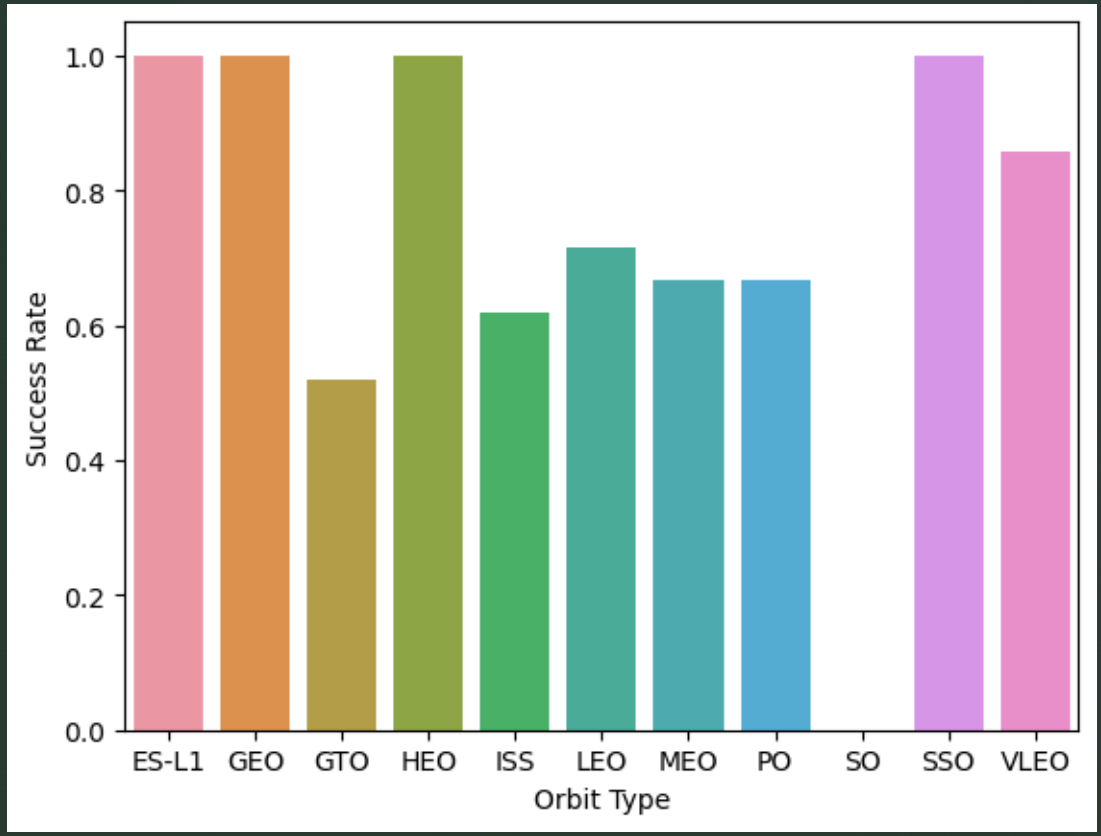
Flight Number v Orbit Type

The LEO orbit success rate appears related to the number of flights while there seems to be no relationship when in GTO orbit.



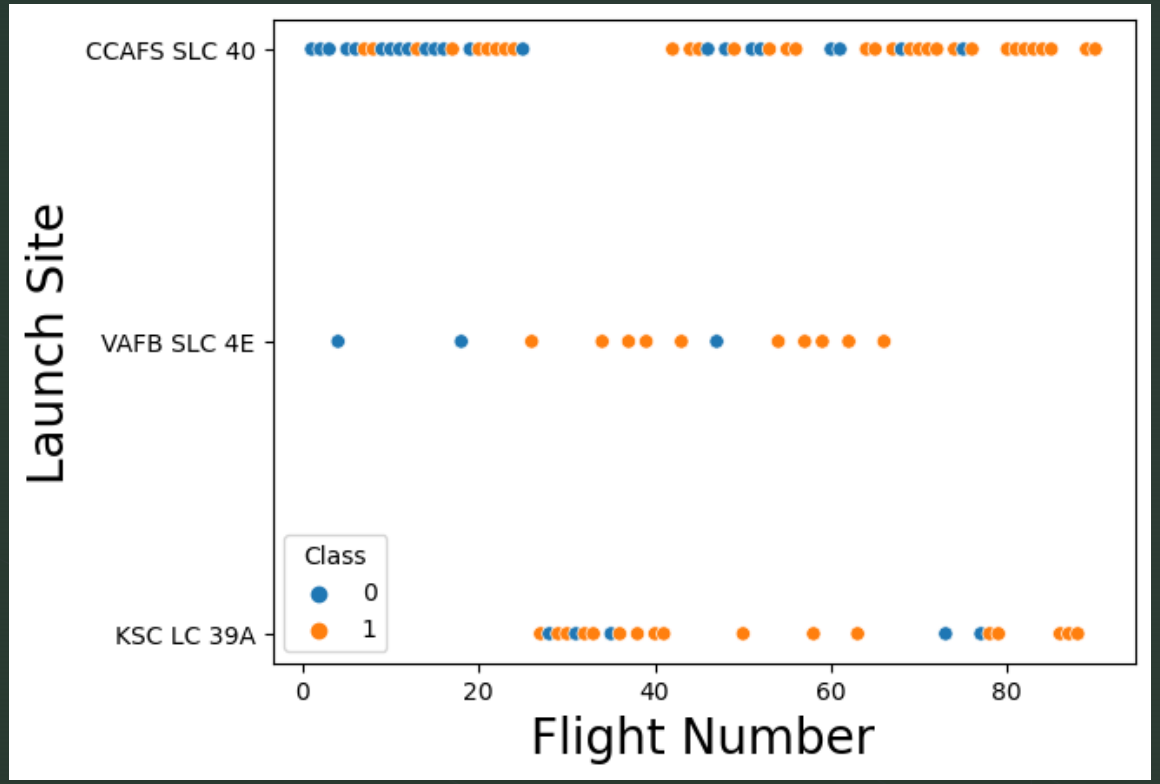
Orbit Type v Success Rate

ES-L1, GEO, HEO and SSO had the highest success rates: 1 (Only SSO had more than 1 launch attempt)



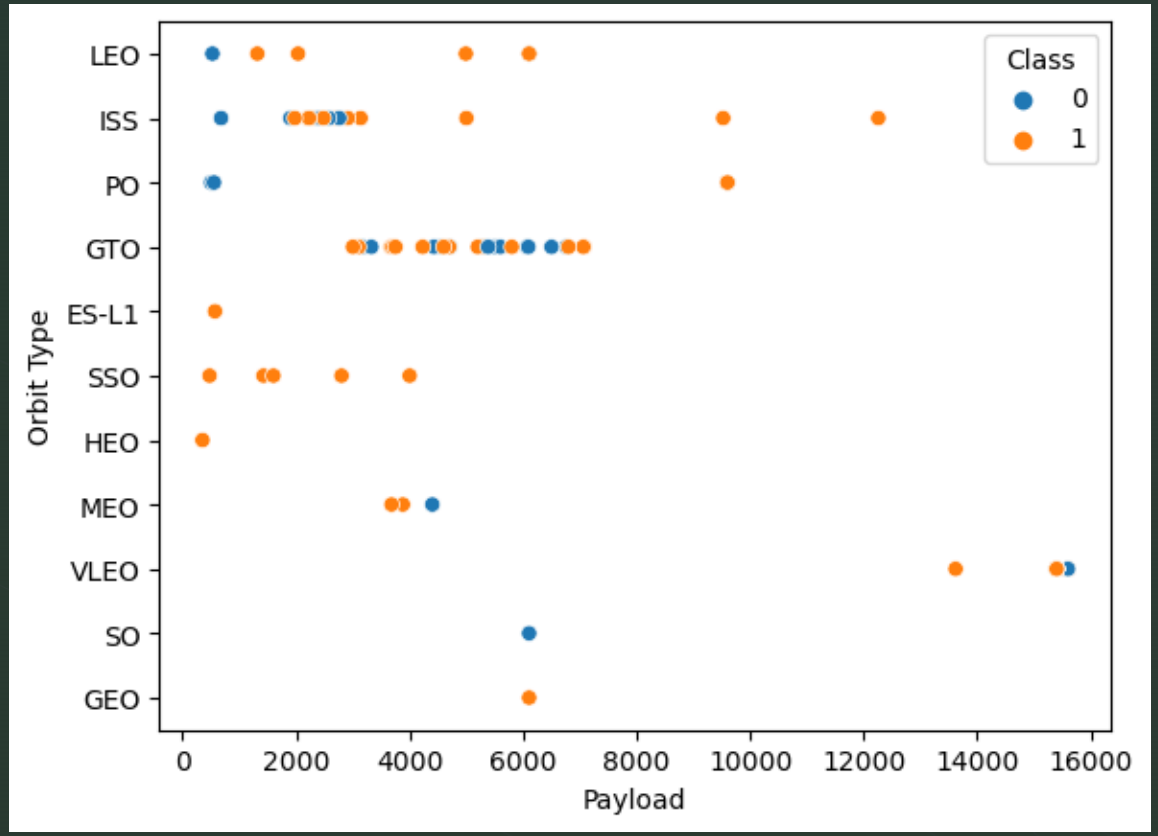
Flight Number v Launch Site

From 40, the number of successful outcomes increased.



Orbit Type V Payload

The heavier the payload, the more successful landing rate for LEO and ISS.

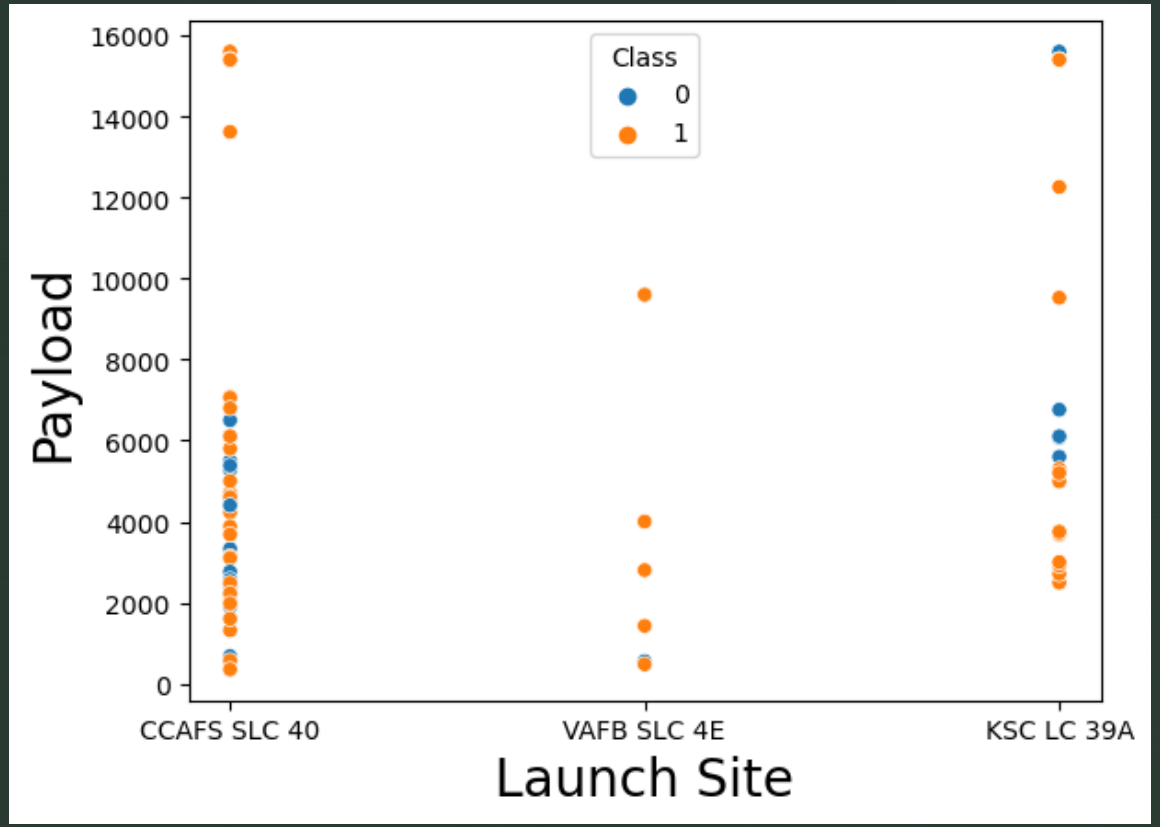


Payload v Launch Site

CCAFS SLC 40: All payloads greater than 10,000kg successful.

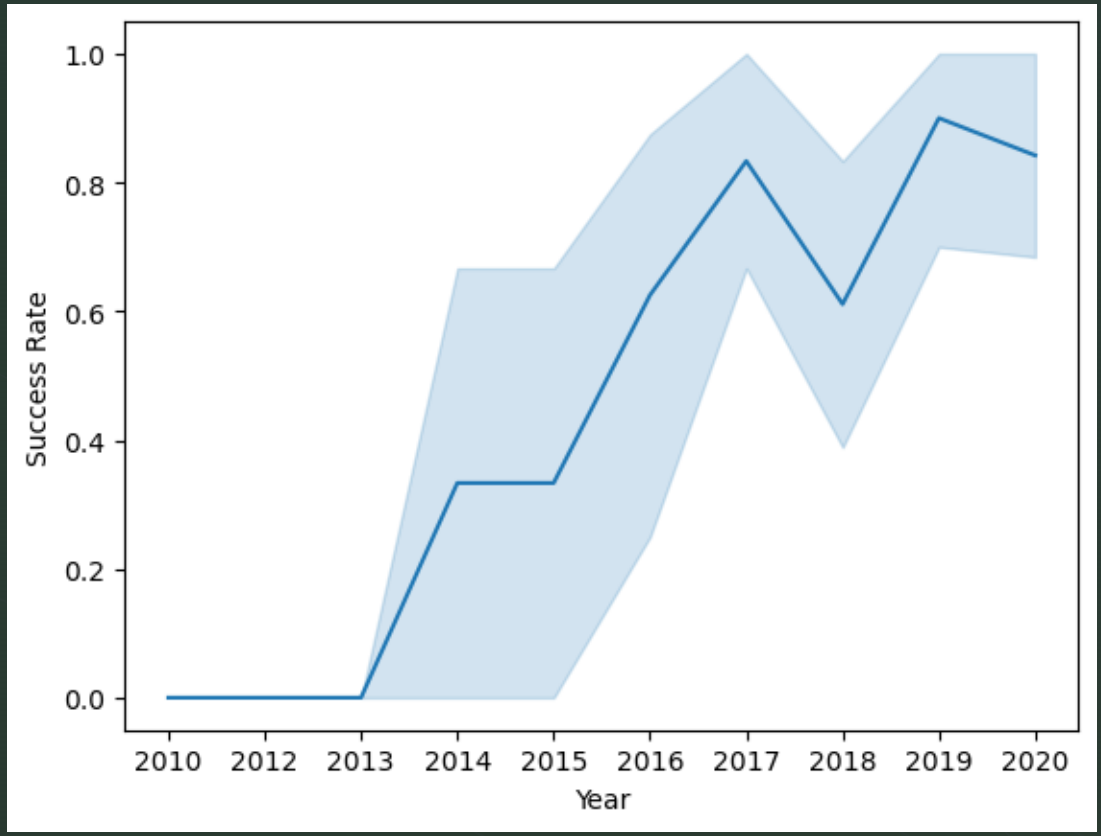
No payload greater than 10,000kg launched at VAFBSLC 4E.

In KSC LC 39A, the success rate is greater for payloads above 8,000kg than it is for payloads lesser than 8,000kg.



Yearly Trend

Steady increase in successful launches from 2013 till 2018 which had a sharp decrease then another increase the following year in 2019.

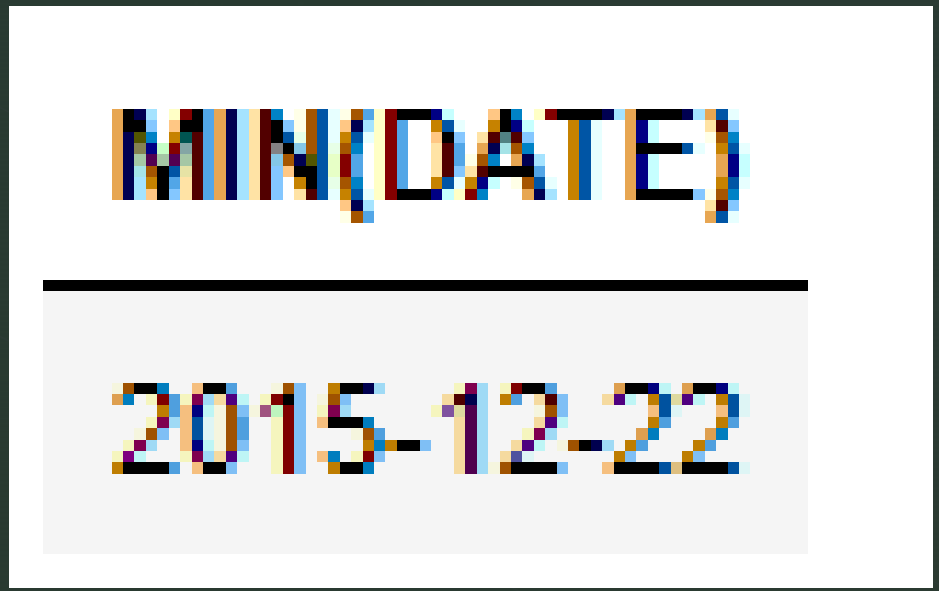


EDA With SQL

The names of the launch sites in the space mission:

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

The date the first successful landing outcome on the pad was achieved:



- List of boosters with a successful landing on a drone ship which have a payload mass greater than 4,000kg and lesser than 6,000kg:

Booster_Version
F9 FT B1022
F9 FT B1026
F9 FT B1021.2
F9 FT B1031.2

Number of successes and failures:

Mission_Outcome	COUNT(*)
Failure (in flight)	1
Success	98
Success	1
Success (payload status unclear)	1

List of boosters which have carried the maximum payload:

Booster_Version
F9 B5 B1048.4
F9 B5 B1049.4
F9 B5 B1051.3
F9 B5 B1056.4
F9 B5 B1048.5
F9 B5 B1051.4
F9 B5 B1049.5
F9 B5 B1060.2
F9 B5 B1058.3
F9 B5 B1051.6
F9 B5 B1060.3
F9 B5 B1049.7

List displaying records from 2015:

Year	Month	Landing_Outcome	Booster_Version	Launch_Site
2015	10	Failure (drone ship)	F9 v1.1 B1012	CCAFS LC-40
2015	04	Failure (drone ship)	F9 v1.1 B1015	CCAFS LC-40

List of landing outcomes between 2016-06-04 and 2017-03-20 in descending order:

Landing_Outcome	COUNT(*)
No attempt	10
Success (ground pad)	5
Success (drone ship)	5
Failure (drone ship)	5
Controlled (ocean)	3
Uncontrolled (ocean)	2
Precluded (drone ship)	1
Failure (parachute)	1

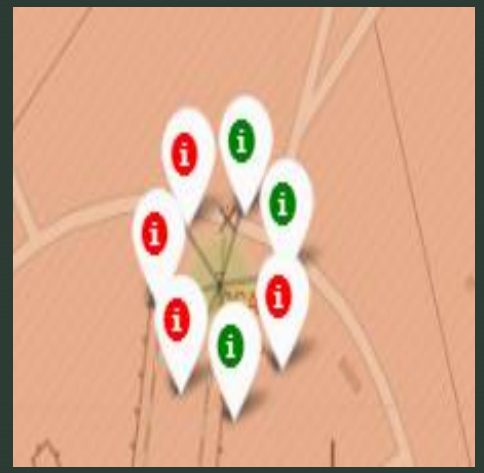
Folium

A map showing the locations of the launch sites:

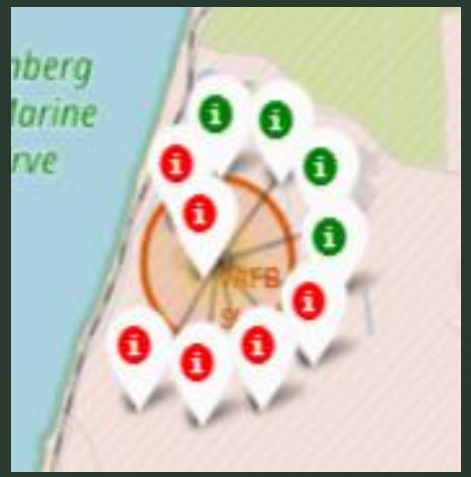
Clusters showing the outcomes in the launch sites:

KSC LC-39A

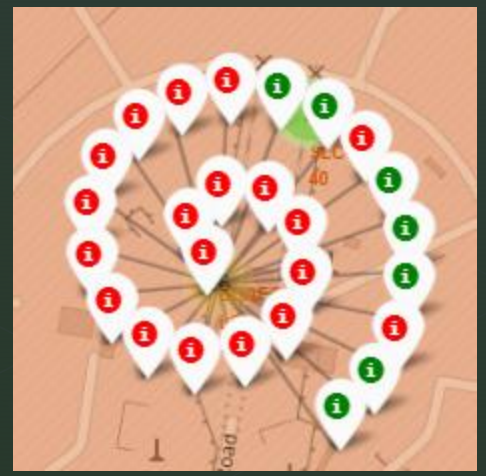
CCAFS SLC-40



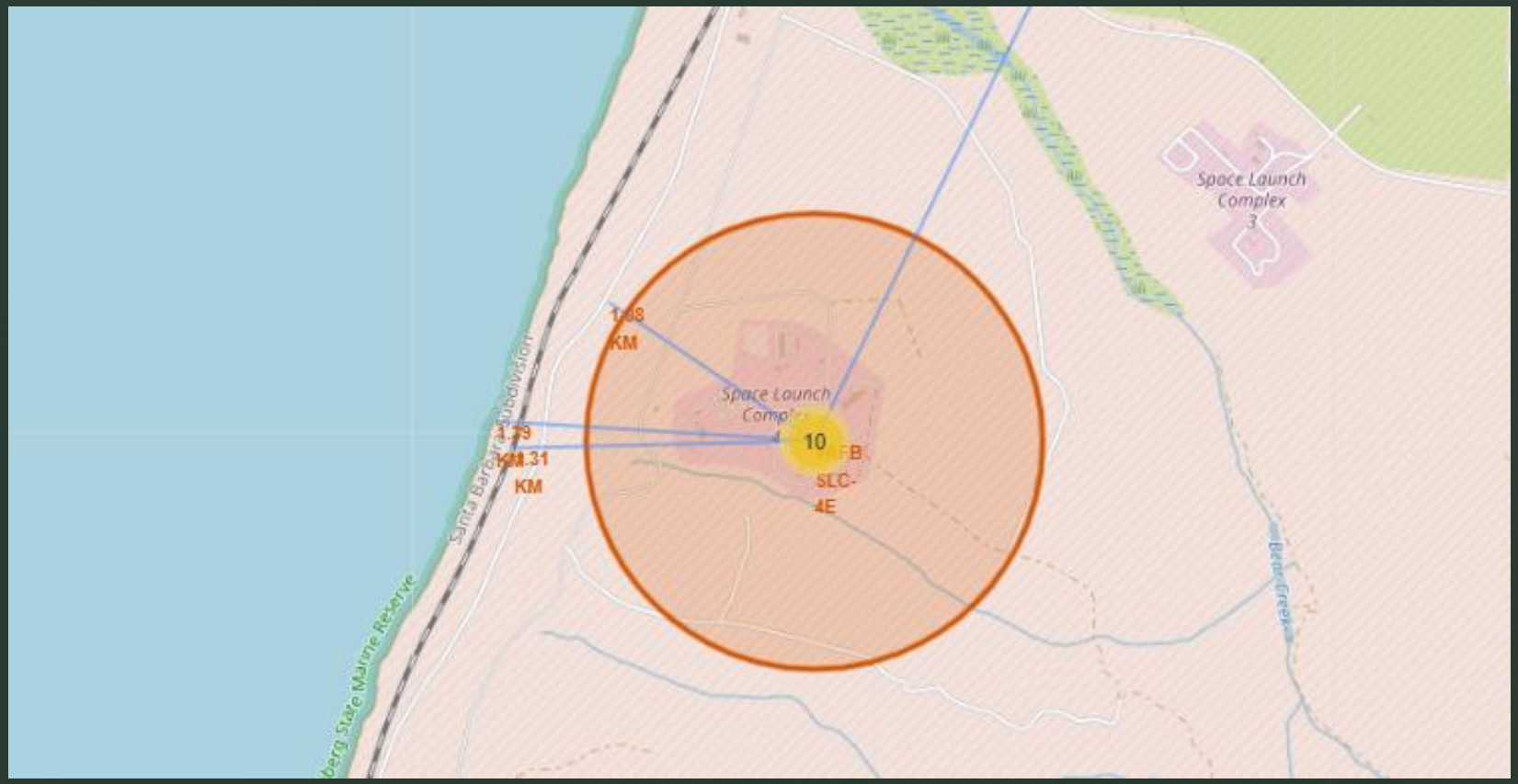
VAFB SLC-4E



CCAFS LC-40

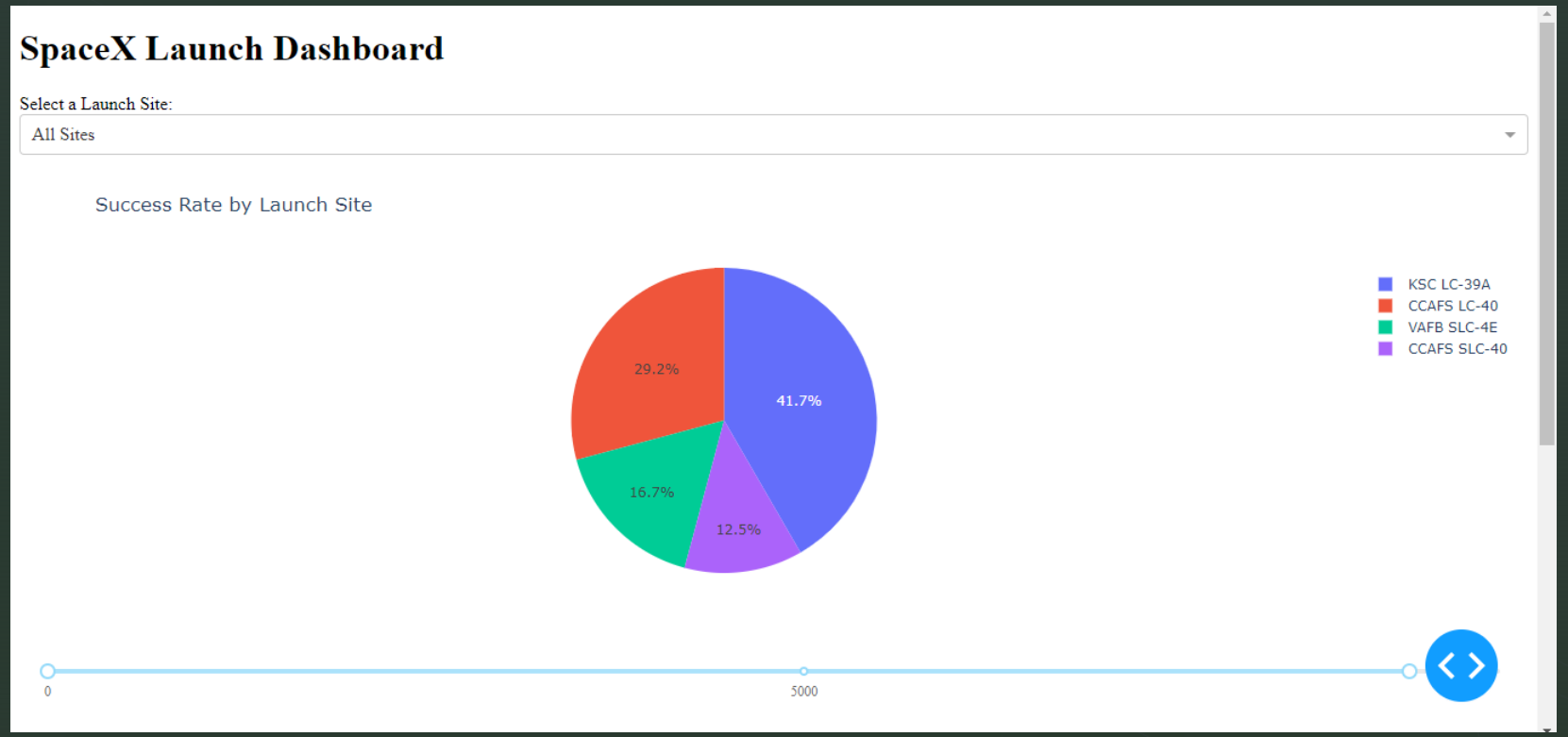


A map showing the launch site's proximity to specific places:

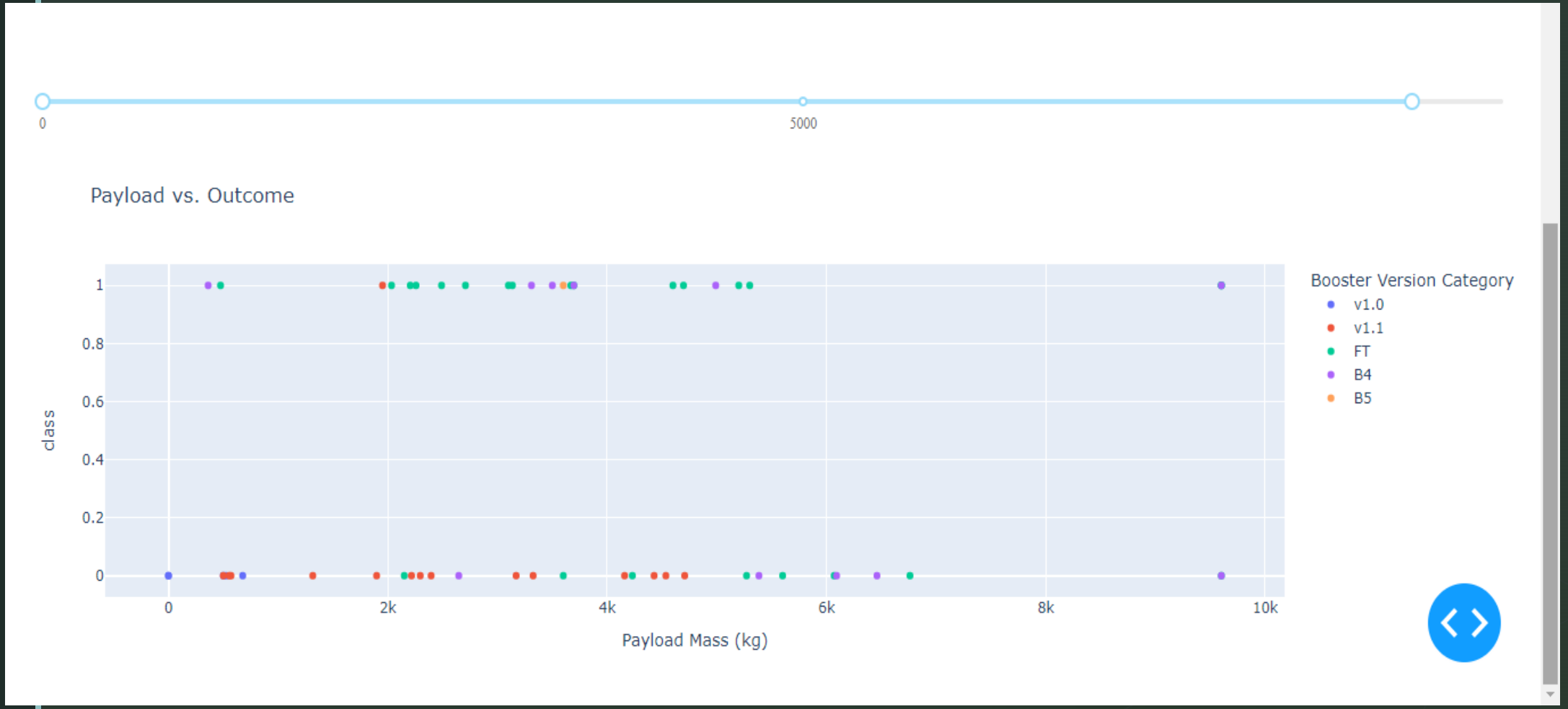


Plotly Dash

The Dashboard showing all successful launches and launches by site:

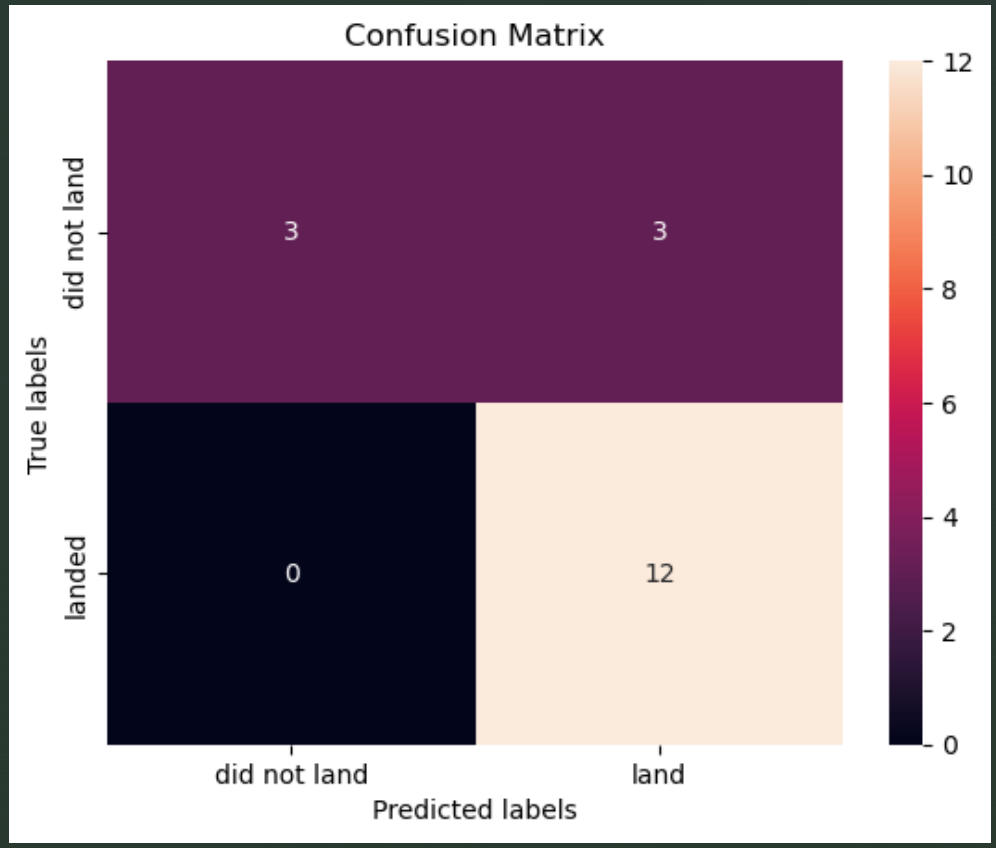


The booster version and payload:



Predictive Analysis

The confusion matrix for KNN, Logistic Regression, SVM and Decision Tree:



CONCLUSION

- The best method for the prediction was Logistic Regression with 0.83.
- The successful launches got better as the years went by.
- A bigger payload mass could possibly impact the successful outcome of the first stage landing.

APPENDIX

- All the files, notebooks and reports can be found here - Github.