ANALYZING SPACEX'S SUCCESS:

PREDICTING THE FUTURE OF A SIMILAR SPACE AGENCY

ANMOL BISWAS

SPACEX

CONTENT

EXECUTIVE SUMMARY

The subsequent techniques were employed to examine the data:

1. Gathering data through web scraping and SpaceX API.
2. Conducting Exploratory Data Analysis (EDA), which involved data preprocessing, data visualization, and interactive visual analytics.
3. Employing Machine Learning for predictive purposes.

Summary:
Valuable data was collected from public sources, and by using Exploratory Data Analysis (EDA), the most significant features that predict the success of launchings were identified. Additionally, Machine Learning Prediction revealed the best model for accurately determining the relevant characteristics.



INTRODUCTION

The main objective is to evaluate the viability of Space Y, a new company, in its ability to compete with Space X. The assessment will focus on determining the total cost for launches by predicting the success rate of first-stage rocket landings. Additionally, identifying the best launch location will be a key aspect of the evaluation process.

METHODOLGY

- DATA COLLECTION
- DATA WRANGLING
- EDA (DATA VISUALIZATION)
- EDA (SQL)
- INTERACTIVE MAPPING WITH FOLIUM
- DYNAMIC DASHBOARD WITH PLOTLY
- PREDICTIVE ANALYSIS (CLASSIFICATION)

DATA COLLECTION

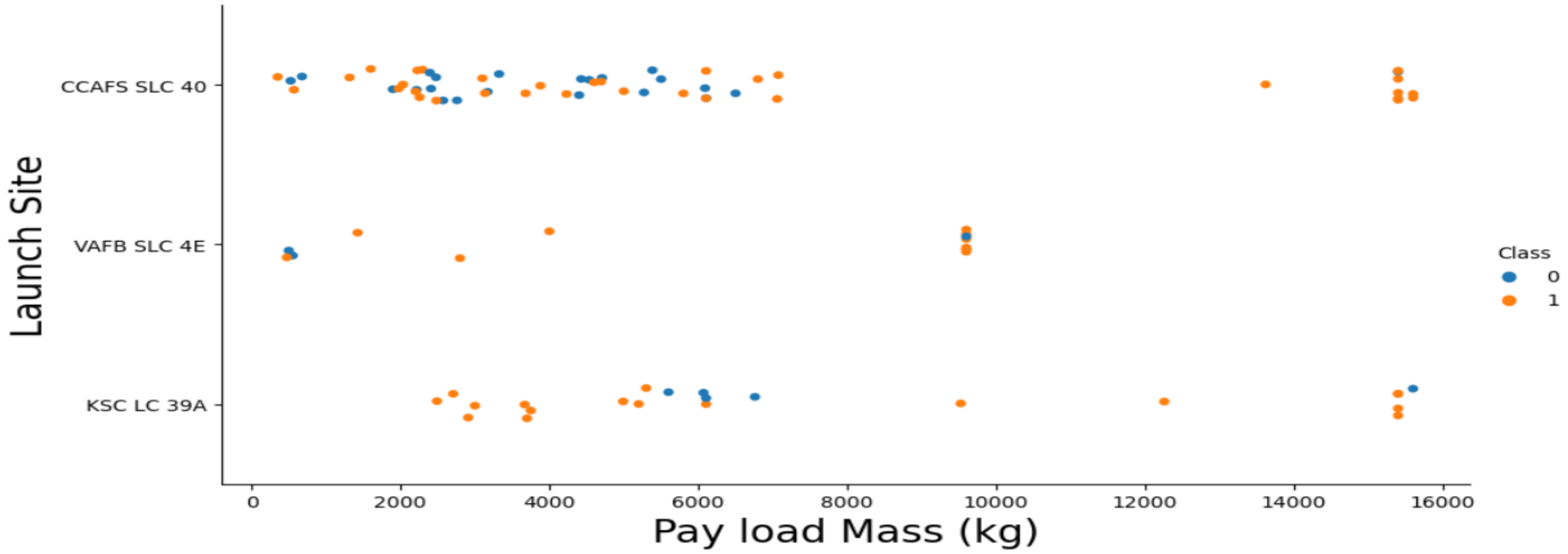
- SpaceX offers a public API for users to access and utilize data. The API was employed as depicted in the accompanying flowchart, enabling the retrieval of data, which was then stored for future use.
- Data was gathered from two sources: the Space X API (https://api.spacexdata.com/v4/rockets/) and Wikipedia (https://en.wikipedia.org/wiki/List_of_Falcon/_9/_and_Falcon_Heavy_launches), utilizing web scraping methods.

DATA WRANGLING

- The dataset initially underwent exploratory data analysis (EDA). Following this, summaries of launches were computed based on site, orbit occurrences, and mission outcomes by orbit type. Ultimately, the "Outcome" column was utilized to create the label for the landing outcome.

EDA (DATA VISUALIZATION)

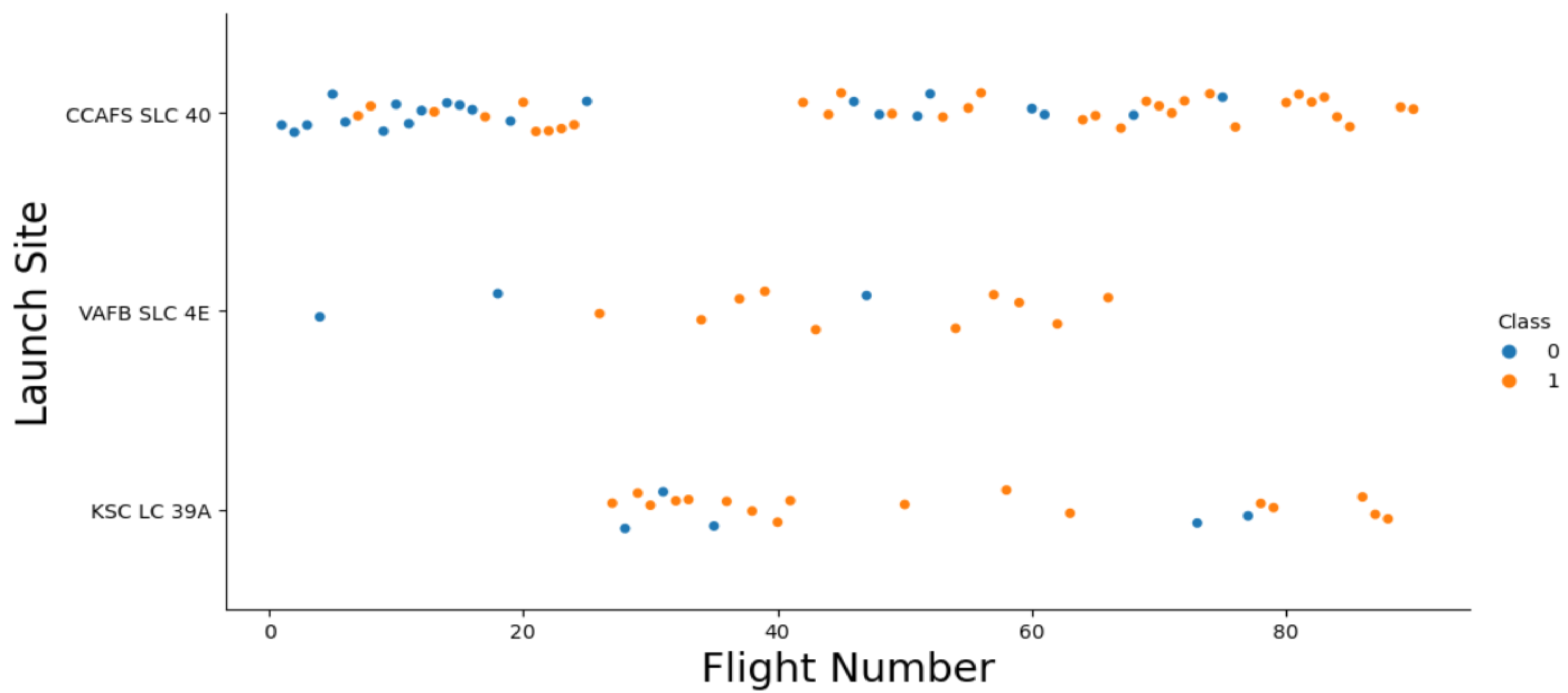
- The data investigation involved using scatterplots and barplots to illustrate the relationships between various features. These included Orbit and Flight Number, Payload and Orbit, Payload Mass multiplied by Flight Number, as well as Launch Site and Flight Number.



Why The map above demonstrates that CCAF5 SLC 40, where the majority of recent launches were successful, is currently the best launch point.

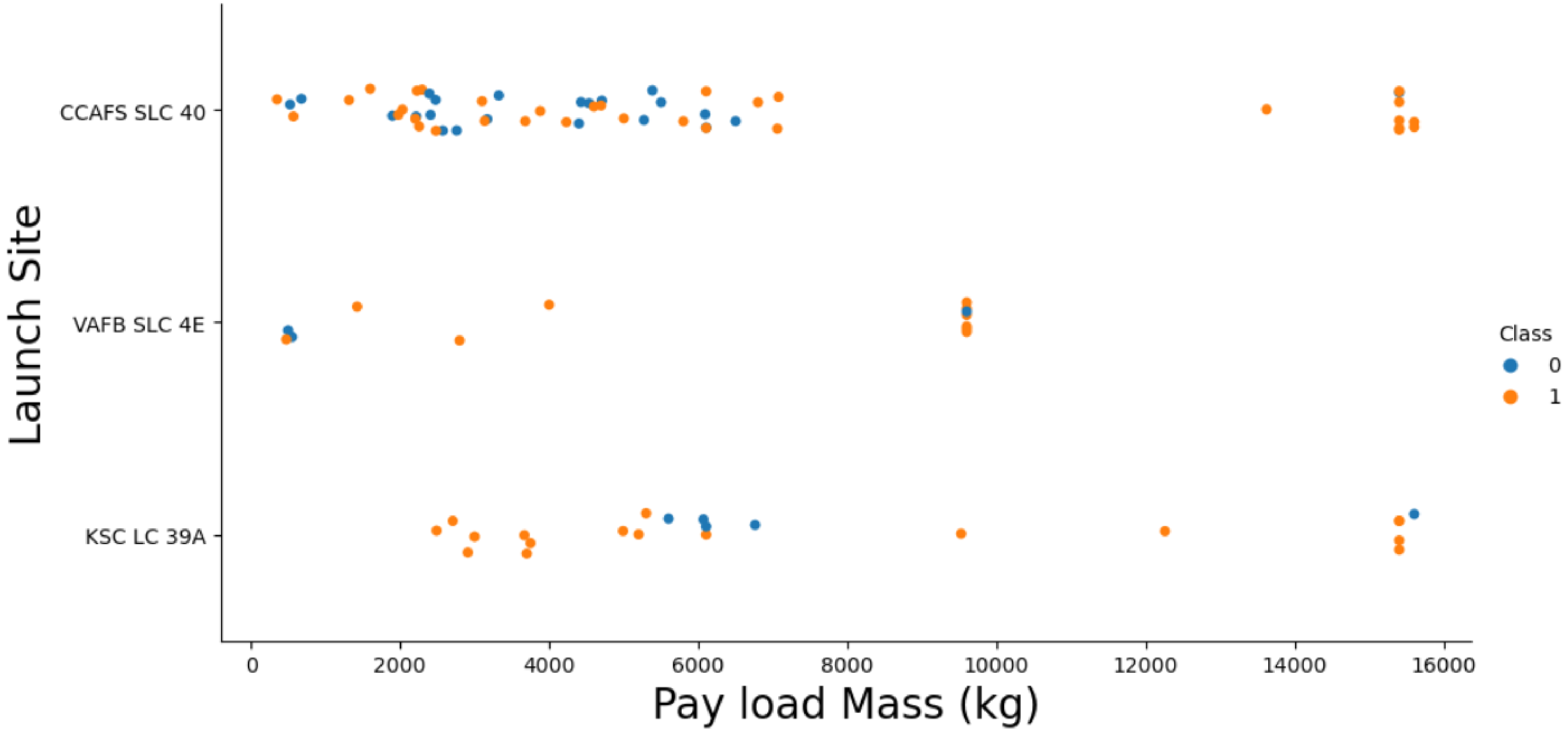
VAFB SLC 4E was in second position, while KSC LC 39A was in third

Additionally, the general success rate can be seen to have increased over time.



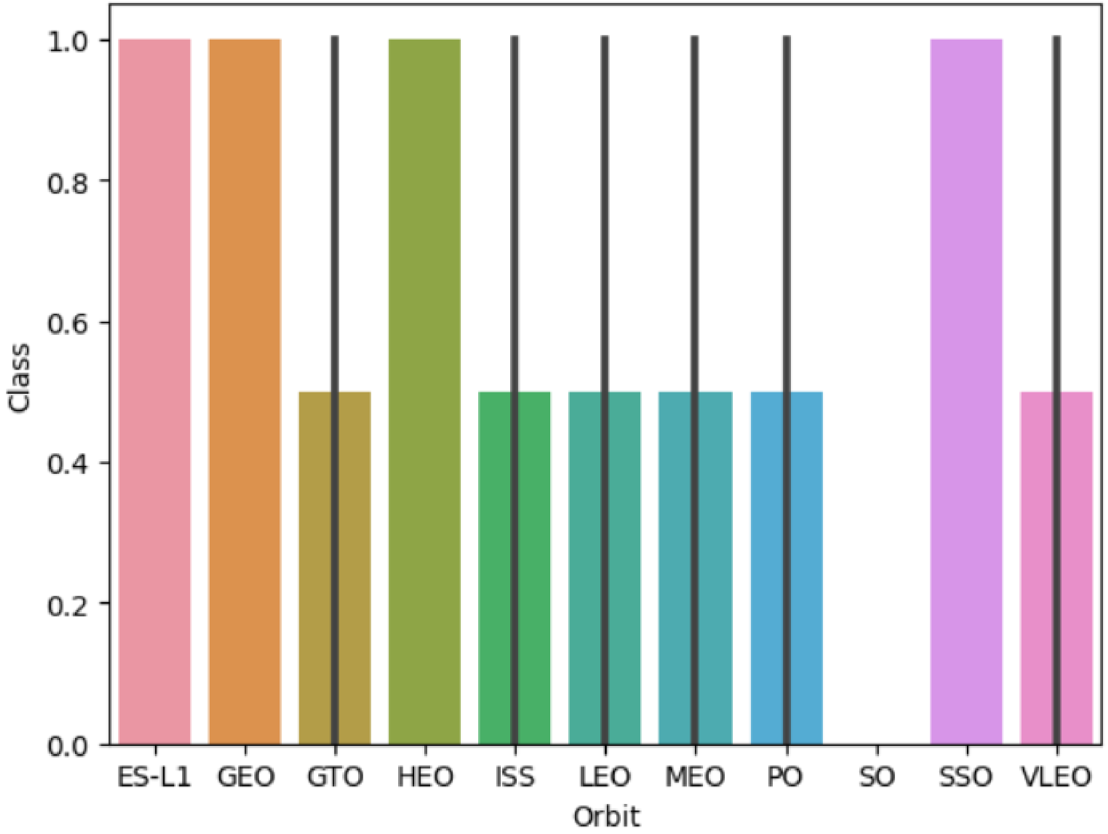
Payloads above 12,000kg appear to be only possible on CCAFS SLC 40 and KSC LC 39A launch sites

Payloads over 9,000kg (about the weight of a school bus) have high success rates.



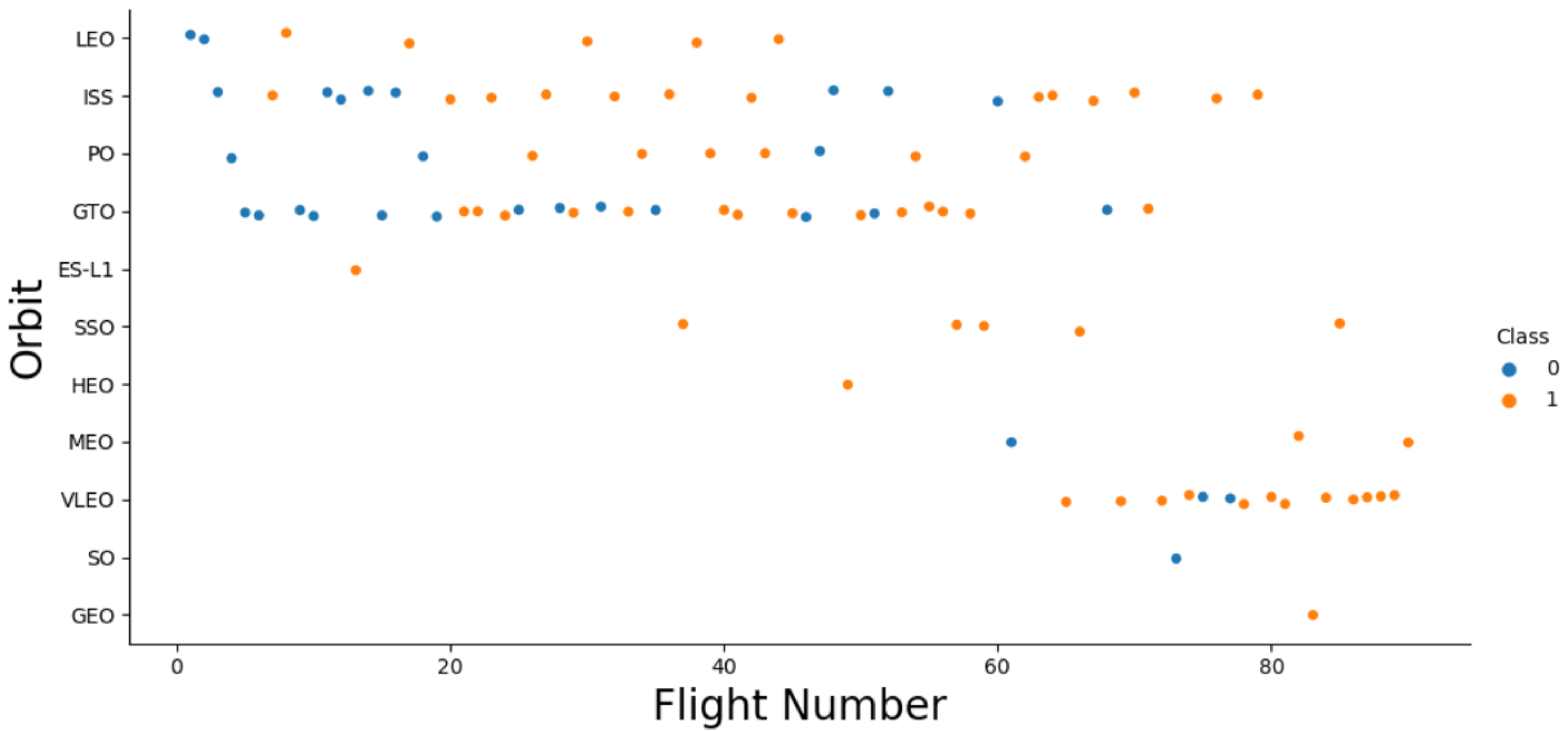
The orbits with the highest success rates are: ES L1, GEO, HEO, and SSO.

VLEO (above 80%) and LFO (above 70%) are the next two variables.

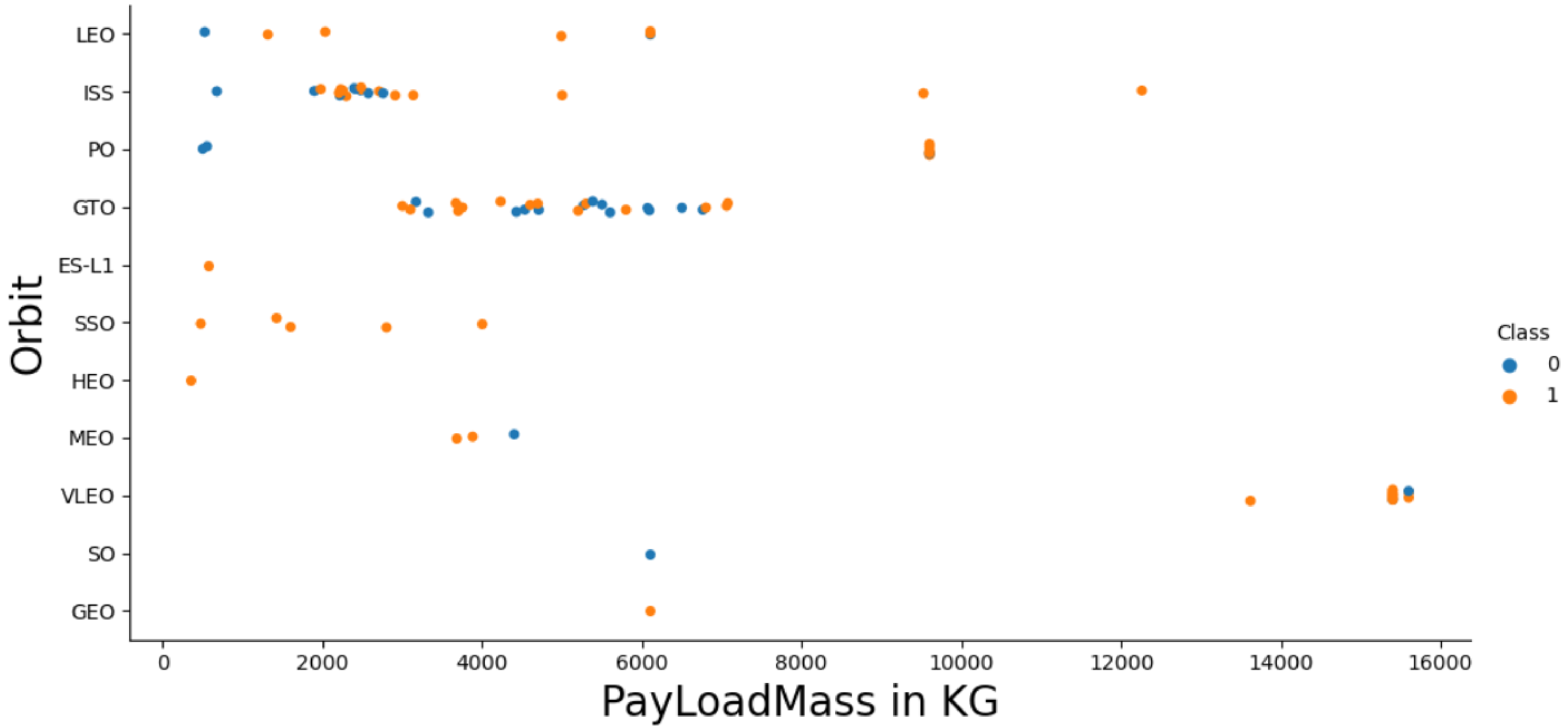


It appears that all orbits' success rates have increased with time

Due to the recent increase in the market size for VLEO orbit frequency.

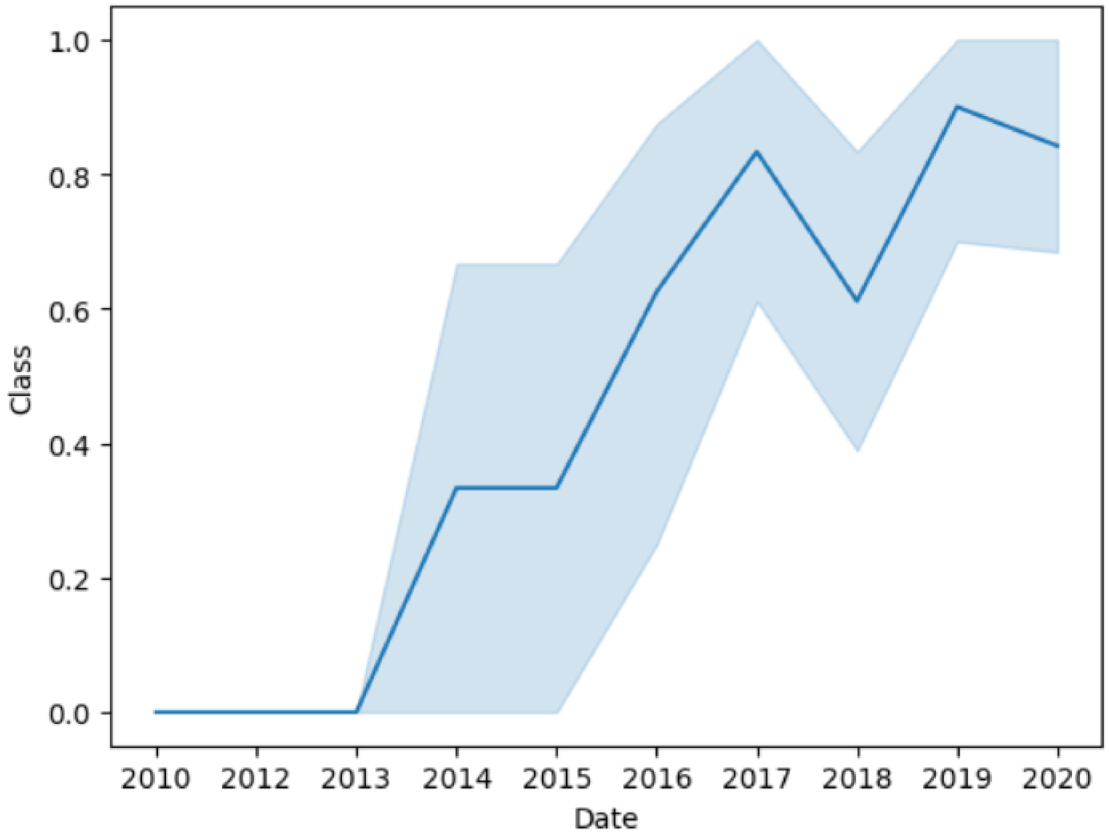


There appears to be no correlation between payload and success rate for launches to the GTO orbit; The ISS orbit has the broadest payload range and a high success rate and there are few launches to the SO and GEO orbits.



The success rate began to rise in 2013 and continued to do so until 2020.

It appears that the first three years were a time of adjustments and technological advancement.



EDA WITH SQL

RESULTS:

All Launch Site Names :

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

Launch Site Names Begin with 'CCA'

Launch_Site
CCAFS LC-40
CCAFS LC-40
CCAFS LC-40
CCAFS LC-40
CCAFS LC-40

Total Payload Mass

sum(PAYLOAD_MASS__KG_)
45596.0

Average Payload Mass by F9 v1.1

AVG(PAYLOAD_MASS__KG_)
2928.4

First Successful Ground Landing Date

min(Date)
01/07/2020

Successful Drone Ship Landing with Payload between 4000 and 6000

Booster_Version
F9 v1.1
F9 v1.1 B1011
F9 v1.1 B1014
F9 v1.1 B1016
F9 FT B1020
F9 FT B1022
F9 FT B1026
F9 FT B1030
F9 FT B1021.2
F9 FT B1032.1
F9 B4 B1040.1
F9 FT B1031.2
F9 B4 B1043.1
F9 FT B1032.2
F9 B4 B1040.2
F9 B5 B1046.2
F9 B5 B1047.2
F9 B5 B1046.3
F9 B5B1054
F9 B5 B1048.3
F9 B5 B1051.2
F9 B5B1060.1
F9 B5 B1058.2

Total Number of Successful and FailureMission Outcomes

Count(*)
98

Boosters Carried 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

2015 Launch Records

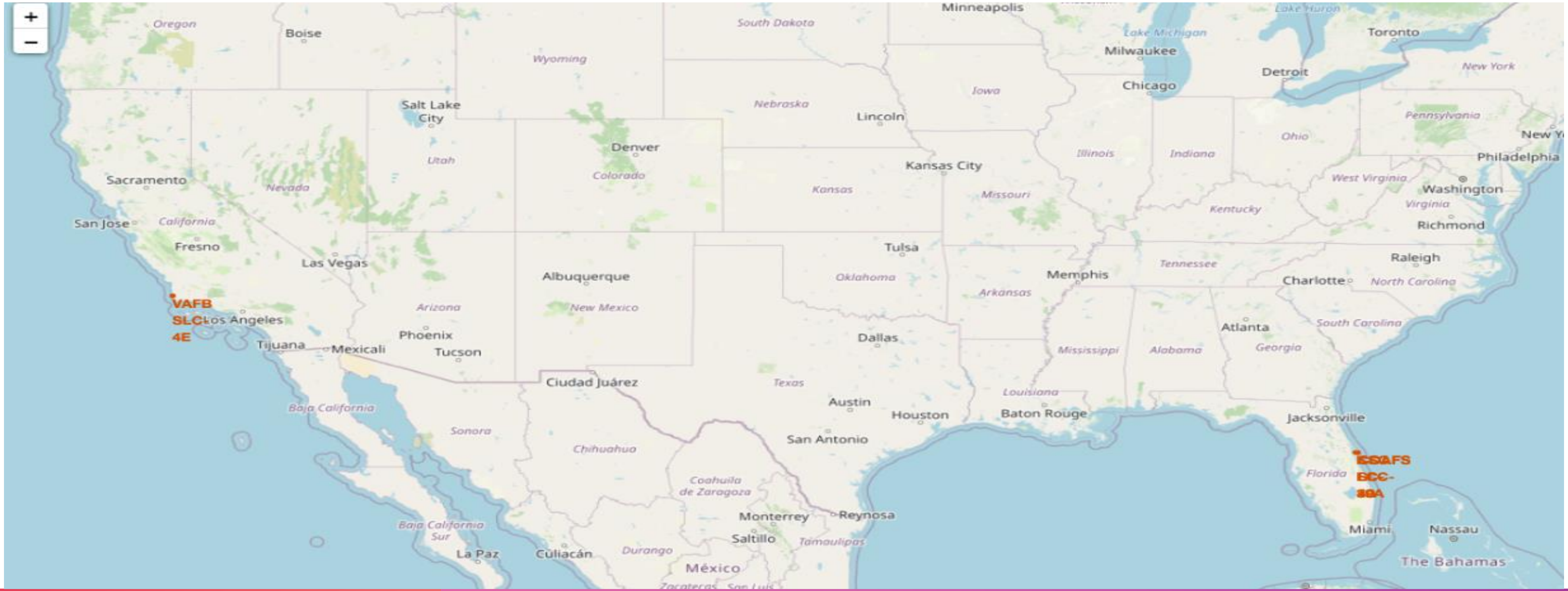
substr(Date, 4, 2)	substr(Date,7,4)	Landing_Outcome	Booster_Version	Launch_Site
10	2015	Failure (drone ship)	F9 v1.1 B1012	CCAFS LC-40
04	2015	Failure (drone ship)	F9 v1.1 B1015	CCAFS LC-40

Rank Landing Outcomes Between 2010-06-04 and 2017-03-20

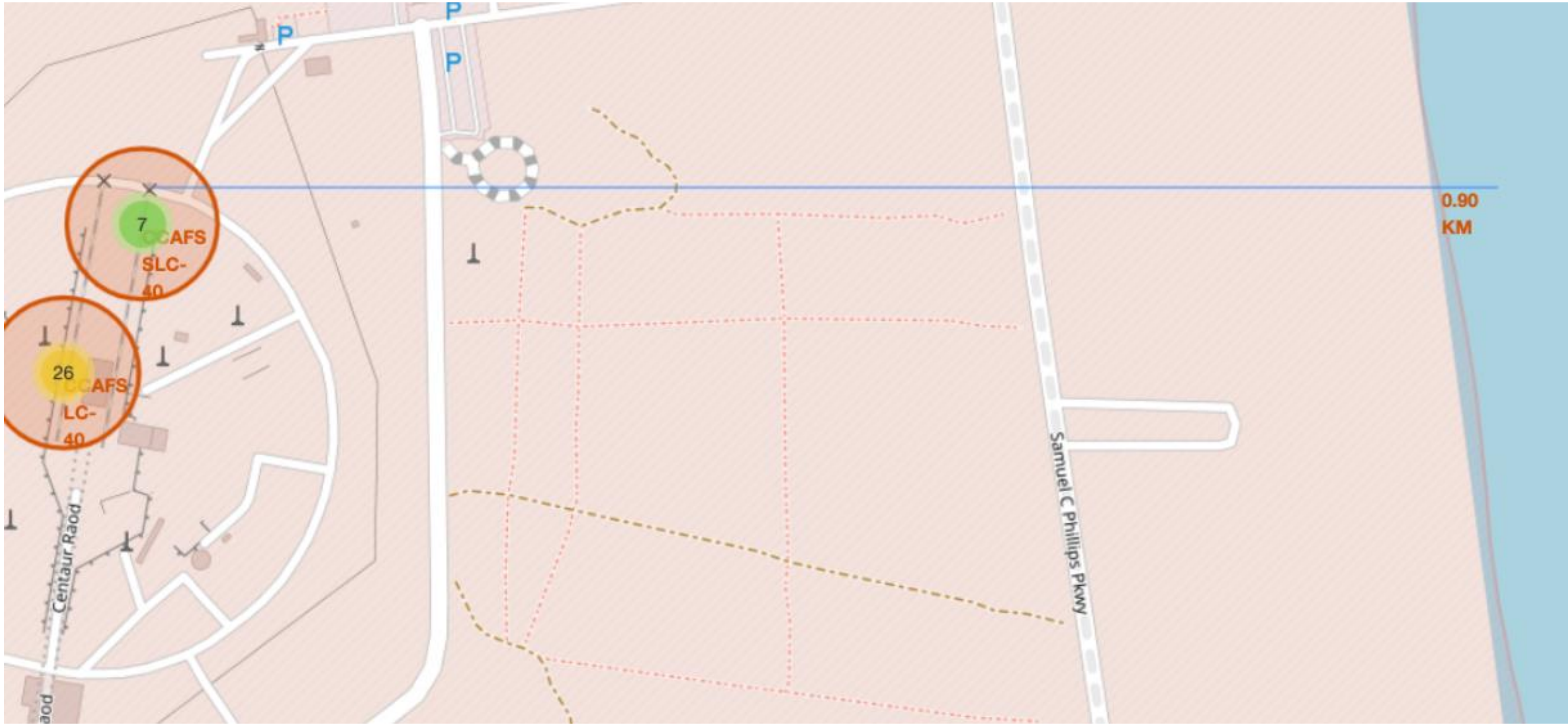
Landing_Outcome	count(Landing_Outcome)
Success	20
Success (drone ship)	8
Success (ground pad)	7

MAPPING USING FOLIUM

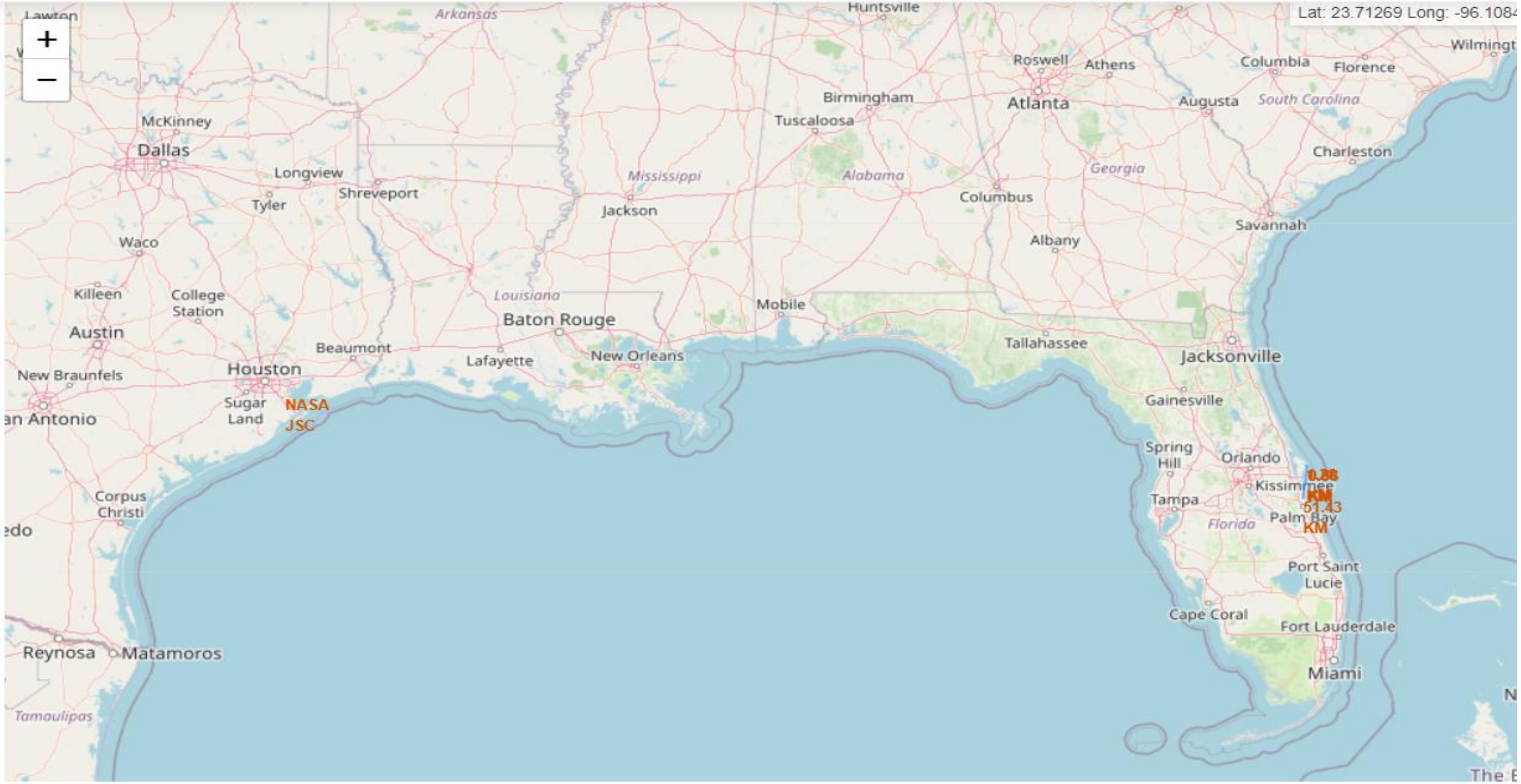
RESULTS:



Launch Outcomes by Site

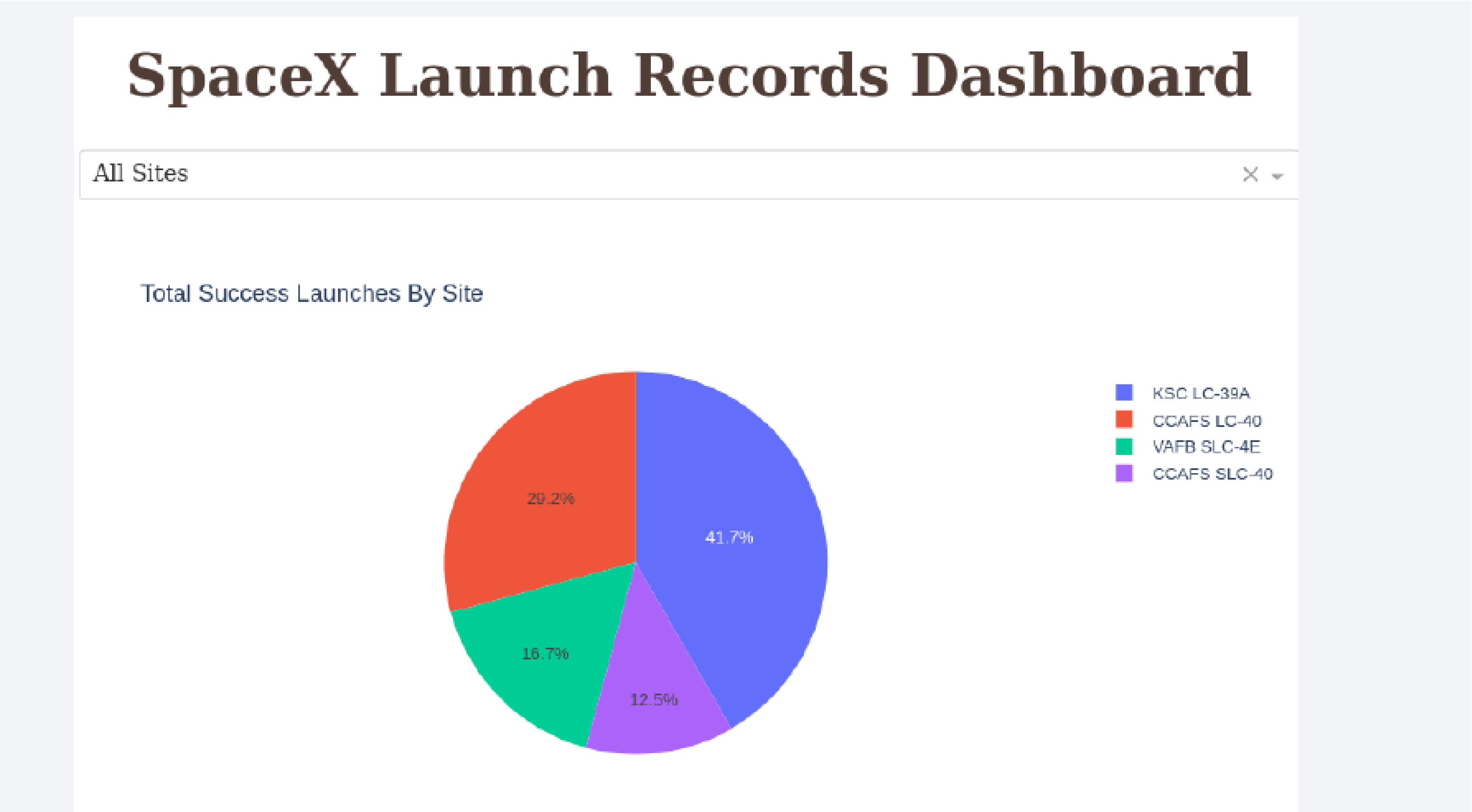


Logistics and Safety

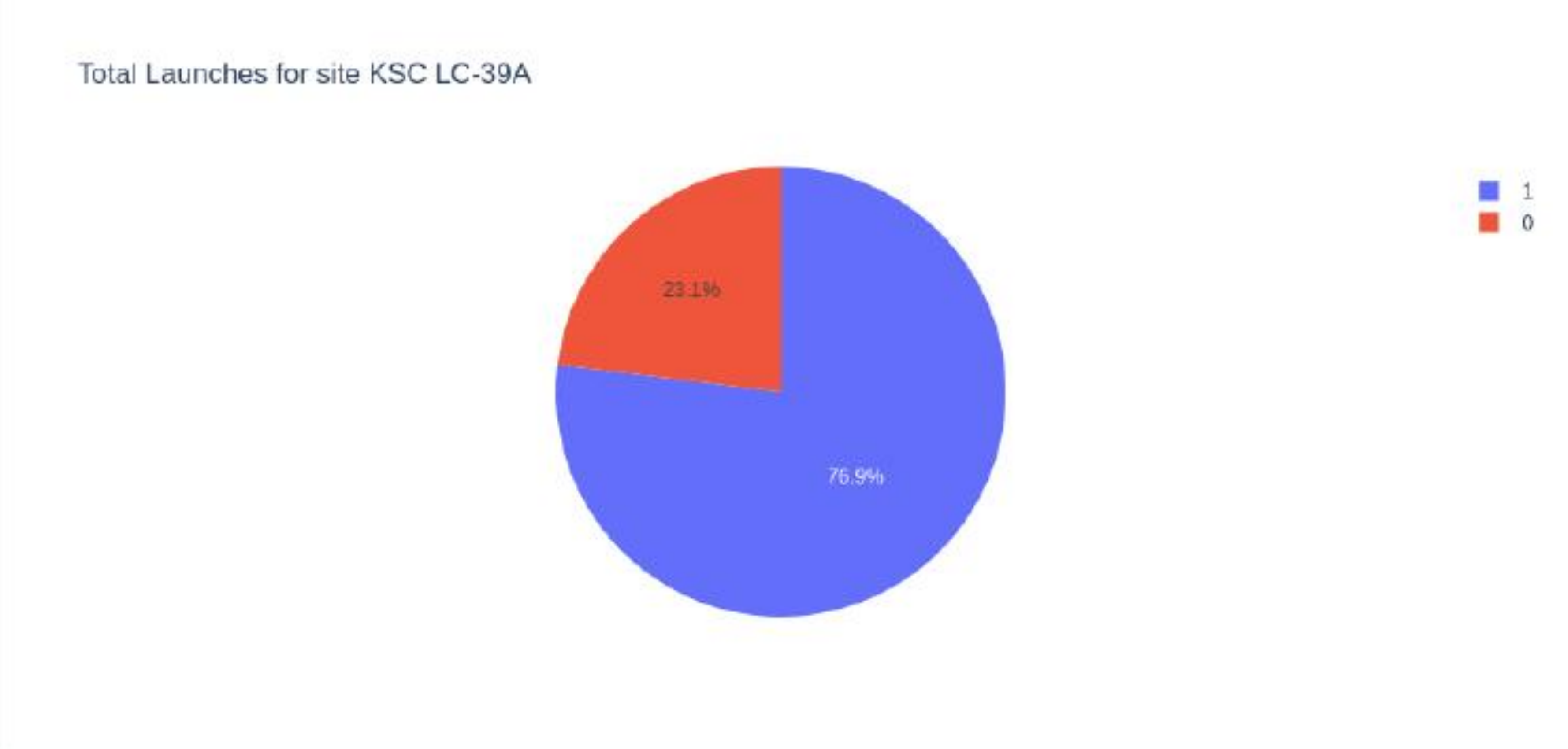


DASHBOARD WITH PLOTLY DASH

Successful Launches by Site



Launch Success Ratio for KSC LC-39A



Payload vs. Launch Outcome

Payload vs. Launch Outcome

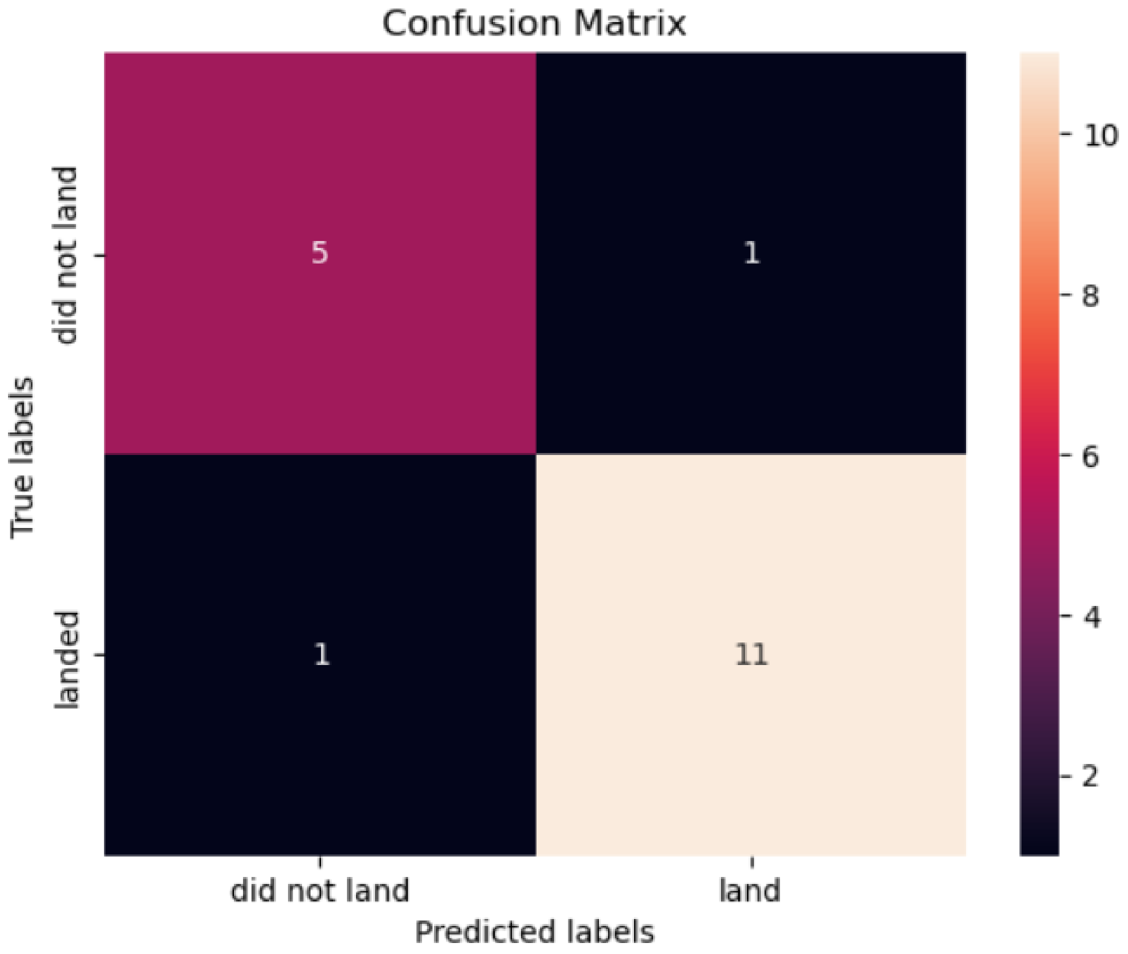
PREDICTIVE ANALYSIS

Classification Accuracy

	Model	Accuracy	Prediction score
0	LogisticRegression()	0.8464285714285713	0.8333333333333334
1	SVC()	0.8482142857142856	0.8333333333333334
2	DecisionTreeClassifier()	0.8732142857142856	0.8888888888888888
3	KNeighborsClassifier()	0.8482142857142858	0.8333333333333334

The highest classification accuracy was achieved by Decision Tree Classifier, which had accuracies of over 87%, out of the four classification models that were tested.

Confusion Matrix of Decision Tree Classifier



CONCLUSIONS

- The best launch site is KSC LC 39A, with launches over 7,000 kg are less risky although most missions are successful.
- Successful landing outcomes appear to improve over time, according to the evolution of processes and rockets; and Decision Tree Classifier can be used to predict successful landings and increase profits.

THANK YOU