Reusing the Falcon 9 Stage One Booster In the Space Race

Deanna Johansen

Applied Data Science Capstone Project

June 3, 2024

TABLE OF CONTENTS

EXECUTIVE SUMMARY

SpaceX is an American aerospace company that claims that their Falcon 9 Rocket can not only carry payloads to the Earth's Orbit, but that the first stage of this rocket can land and be reused. If true, this is a huge cost savings that they can pass onto their customers. They advertise that their Falcon 9 rocket launches cost 62 million dollars while other providers cost upwards of 165 million dollars each. **The purpose of this research is to determine the factors that improve the success rate of the Falcon 9's stage one landing and to consider if this success rate is high enough for SpaceX to make this claim.**

EXECUTIVE SUMMARY

Summary of Methodologies

- **Collect raw data** using SpaceX REST API and web scraping.
- **Wrangle the data** to create a variable to represent success or failed outcome
- **Explore the data** with data visualization techniques, using the factors of flight number, launch site, payload mass and yearly trend.
- **Analyze the data** with SQL to calculate the total number of successful and failed outcomes related to launch sites and payload.
- **Compare** launch site success rates and their proximity to geographical markers.
- **Visualize** the launch sites with the most success.
- **Build Models to predict** landing outcomes of the stage 1 rocket using logistic regression, support vector machine (SVM), decision tree and K-nearest neighbor (KNN).

Summary of Results

- **Exploratory Data Analysis:**
  - Launch success has improved over time
  - KSC LC-39A has the highest success rate among landing sites
  - Payload mass and booster selection can be a factor in overall success
  - Orbits ES -L1, GEO, HEO, and SSO have a 100% success rate
- **Visualization/Analytics:**
  - Most launch sites are near the equator, and all are close to the coast
- **Predictive Analytics:**
  - All models performed similarly on the test set. The decision tree model slightly outperformed

INTRODUCTION

PROJECT BACKGROUND AND CONTEXT

SpaceX, or Space Exploration Technologies Corporation, is an American aerospace company that designs, manufactures, and launches spacecraft and rockets. This company was founded by Elon Musk in 2002.

SpaceX's headquarters are in Hawthorne, California, and it has launch facilities in Cape Canaveral, Florida, and Vandenberg Space Force Base in California. One of the company’s current products is the Falcon 9. They advertise that their Falcon 9 rocket launches cost 62 million dollars while other providers cost upward of 165 million dollars each. Much of the savings is because SpaceX can reuse the first stage most of the time.

INTRODUCTION

PROJECT BACKGROUND
AND CONTEXT

“Falcon 9 is a reusable, two-stage rocket designed and manufactured by SpaceX for the reliable and safe transport of people and payloads into Earth orbit and beyond. Falcon 9 is the world’s first orbital class reusable rocket. Reusability allows SpaceX to re-fly the most expensive parts of the rocket, which in turn drives down the cost of space access.”

Source: https://www.spacex.com/vehicles/falcon-9/

**See Appendix A for more Falcon 9 information.*

INTRODUCTION

FINDINGS WE ARE INTERESTED IN

The purpose of this study is to identify the factors for a successful Falcon 9 stage one booster landing. This information can help determine the cost of a launch and be used in comparison to the cost of alternate companies that bids against SpaceX for a rocket launch.

Note: For the purposes of this report, ‘launch success’ and “launch success rate” refers to the Falcon 9’s stage one booster landing successfully.

METHODOLOGY

Summary

Data collection methodology:

Describe how data was collected

Perform data wrangling

Describe how data was processed

Perform exploratory data analysis (EDA) using visualization and SQL

Perform interactive visual analytics using Folium and Plotly Dash

Perform predictive analysis using classification models

How to build, tune, evaluate classification models

8

METHODOLOGY

SUMMARY

This section will describe how the data was collected, cleaned, and processed so that it could be studied for analysis with visualizations using various python tools. It will share the results of predictive analysis and how they can be used to predict the success rate of landing stage one of the Falcon 9 rocket.

DATA COLLECTION

With API

Request and parse the SpaceX launch data using the GET request

→ Construct our dataset using the data we have obtained. We will combine the columns into a dictionary.

→ Create a Pandas data frame from the dictionary

→ Filter the dataframe to only include `Falcon 9` launches

Source: https://www.spacex.com/vehicles/falcon-9/

Data collected is from 3/24/2006 – 11/05/2020

	FlightNumber	Date	BoosterVersion	PayloadMass	Orbit	LaunchSite	Outcome	Flights	GridFins	Reused	Legs	La
4	1	2010-06-04	Falcon 9	NaN	LEO	CCSFS SLC 40	None None	1	False	False	False	
5	2	2012-05-22	Falcon 9	525.0	LEO	CCSFS SLC 40	None None	1	False	False	False	
6	3	2013-03-01	Falcon 9	677.0	ISS	CCSFS SLC 40	None None	1	False	False	False	
7	4	2013-09-29	Falcon 9	500.0	PO	VAFB SLC 4E	False Ocean	1	False	False	False	
8	5	2013-12-03	Falcon 9	3170.0	GTO	CCSFS SLC 40	None None	1	False	False	False	
...	...	...	...	...	...	...	...	...	...	...	...	
89	86	2020-09-03	Falcon 9	15600.0	VLEO	KSC LC 39A	True ASDS	2	True	True	True	5e9e3032383ecb6b
90	87	2020-10-06	Falcon 9	15600.0	VLEO	KSC LC 39A	True ASDS	3	True	True	True	5e9e3032383ecb6b

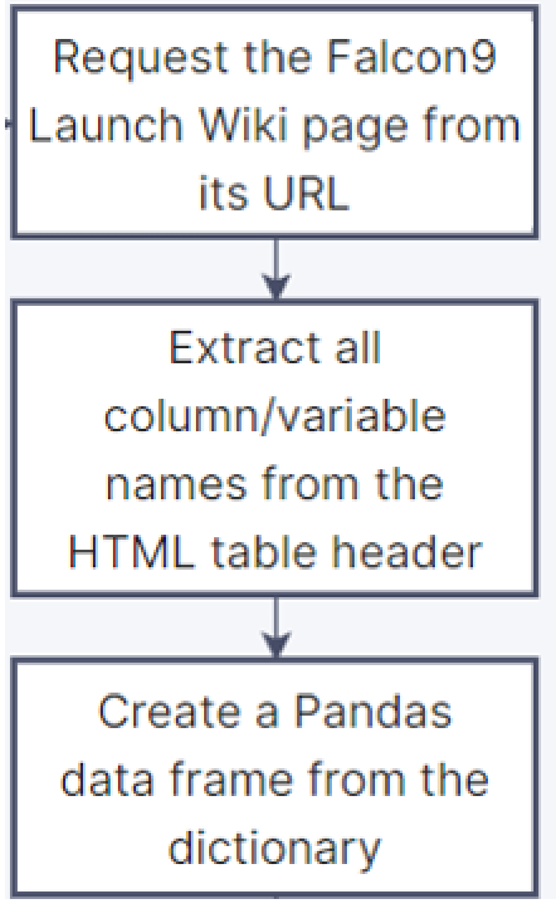
METHODOLOGY

GitHub URL

01_SpaceX_Data_Collection_API.ipynb

DATA COLLECTION

Web Scraping Steps



Dictionary Column Names used in the Pandas Data Frame

```
launch_dict= dict.fromkeys(column_names)

# Remove an irrelvant column
del launch_dict['Date and time ( )']

# Let's initial the launch_dict with each va
launch_dict['Flight No.'] = []
launch_dict['Launch site'] = []
launch_dict['Payload'] = []
launch_dict['Payload mass'] = []
launch_dict['Orbit'] = []
launch_dict['Customer'] = []
launch_dict['Launch outcome'] = []
# Added some new columns
launch_dict['Version Booster']=[]
launch_dict['Booster landing']=[]
launch_dict['Date']=[]
launch_dict['Time']=[]
```

METHODOLOGY

GitHub URL

02_SpaceX_Web_Scraping.ipynb

DATA WRANGLING

- Calculate the number of launches on each site
- Calculate the number and occurrence of each orbit
- Calculate the number of occurence of mission outcome of the orbits
- Create a landing outcome label from the outcome column named "landing_class'

Landing Class refers to whether the Falcon 9 Stage One Booster landed successfully or not.

The Class Label was added to the data frame to clearly display the result that we are after. 0 is an unsuccessful landing and 1 is a successful Landing of the Falcon 9 Stage One Booster.

In [12]:
```
df['Class']=landing_class
df[['Class']].head(8)
```

Out[12]:

	Class
0	0
1	0
2	0
3	0
4	0
5	0
6	1
7	1

METHODOLOGY

GitHub URL

[03_Space_X_Data_Wrangling.ipynb]

EXPLORATORY DATA ANALYSIS (EDA) USING SQL

QUERIES

- Names of unique launch sites
- 5 records where launch site begins with ‘CCA’
- Total payload mass carried by boosters launched by NASA (CRS)
- Average payload mass carried by booster version F9 v1.1.

DETERMINED

- Date of first successful landing for Falcon 9 Stage One on ground pad
- Names of the boosters which had success landing on drone ship and have payload mass greater than 4,000 but less than 6,000
- Total number of successful and failed missions
- Names of booster versions which have carried the max payload
- Failed landing outcomes on drone ship, their booster version and launch site for the months in the year 2015
- Count of landing outcomes between 2010-06-04 and 2017-03-20

METHODOLOGY

GitHub URL

04_SpaceX_EDA_w_SQL.ipynb

EXPLORATORY DATA ANALYSIS (EDA) WITH VISUALIZATIONS

Exploratory Data Analysis (EDA) was conducted to study, explore, and visualize the data to derive important insights.

INCLUDED CHARTS

- Flight Number vs. Payload
- Flight Number vs. Launch Site
- Payload Mass (kg) vs. Launch Site
- Payload Mass (kg) vs. Orbit type

ANALYSIS

- View relationship using scatter plots.
- Show comparisons among discrete categories with bar charts.

METHODOLOGY

GitHub URL

05_SpaceX_EDA_DataViz.ipynb

VISUAL ANALYTICS WITH FOLIUM

MAPS CREATED WITH FOLIUM TO SHOW

- **Markers Indicating Launch Sites**

  Launch sites located using longitude and latitude and marked on a folium map with colored circle and pop-up markers
- **Colored Markers of Launch Outcomes**

  Added colored popup markers of successful (green) and unsuccessful (red) launches at each launch site to visualize which launch sites have high success rates
- **Distances Between a Launch Site to Proximities** such as oceans, highways, cities, and railways

METHODOLOGY

GitHub URL

06_SpaceX_Inter_Visual_Analysis_w_Folium.ipynbv

VISUAL ANALYTICS WITH PLOTLY DASH

DASHBOARD CREATED TO SHOW

- **Dropdown List with Launch Sites** Allow user to select all launch sites or a certain launch site x
- **Pie Chart Showing Successful Launches** Allow user to see successful and unsuccessful launches as a percent of the total
- **Slider of Payload Mass Range** Allow user to select payload mass range
- **Scatter Chart Showing Payload Mass vs. Success Rate by Booster Version** Allow user to see the correlation between Payload and Launch Success

METHODOLOGY

GitHub URL

07_SpaceX_spacex_dash_app.py.txt

PREDICITIVE ANALSYS

DATA COLLECTED, TRAINED, TESTED and DISPLAYED

- Create NumPy array from the ‘Class’ column
- Standardize the data with *StandardScaler*
- Fit and transform the data
- Split the data using the *train_test_split* function.
- Create a *GridSearchCV* object with cv=10 for parameter optimization
- Apply *GridSearchCV* on different algorithms:
    - logistic regression(LogisticRegression())
    - support vector machine (SVC())
    - decision tree (DecisionTreeClassifier())
    - K-Nearest Neighbor (KNeighborsClassifier())
- Calculate accuracy on the test data using *.score()* for all models
- Study the confusion matrix for all four models
- Identify the best model if possible

METHODOLOGY

GitHub URL

[08_SpaceX_Machine Learning Prediction.ipynb]

R E S U L T S

RESULTS SUMMARY

Exploratory Data Analytics Show:

• Show launch success has improved as the number of flights have increased.

• The KSC LC-39A has the highest success rate among landing sites

• The Orbits ES-L1, GEO, HEO and SSO have a 100% success rate

• The launch success rates have increased as the number of flights have increased for all orbits except for the GTO orbit.

•Payload mass and orbit do affect the launch success rate.

•Success rates have climbed nearly all years from 2013 until 2020.

Visual/Interactive Analytics Show:

• Launch sites are in the southern US and all are near the coast.

• Launch sites are chosen such that they are far enough away from anything that a failed launch could not do damage (such as a to a city, highway, railway), while they are still close enough to support launch activities nearby.

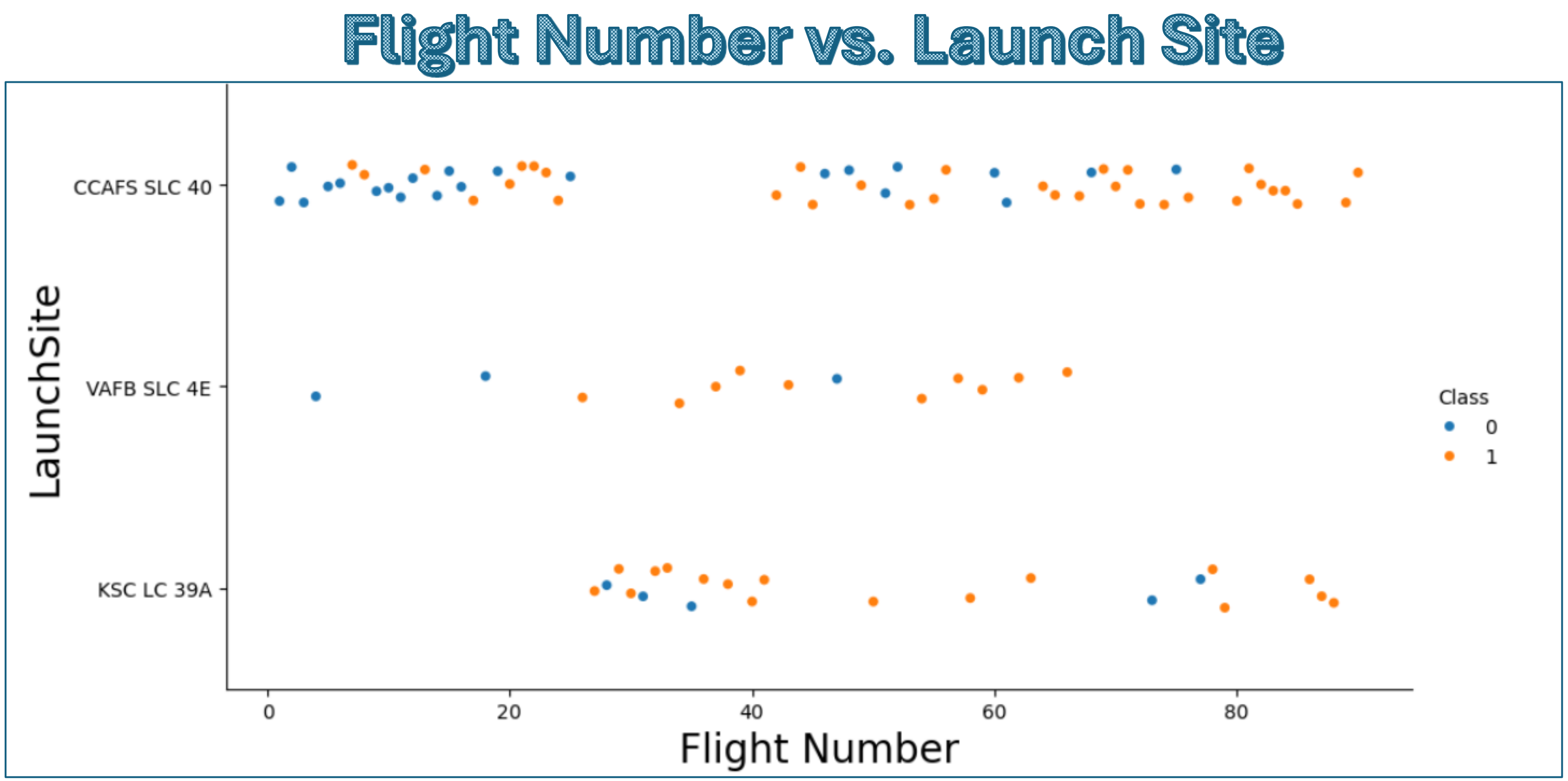
Predictive Analytics Show:

• Of four models studied, the Decision Tree model is the best predictive model for the dataset

FLIGHT NUMBER VS. LAUNCH SITE

Exploratory Data Analysis

- Earlier flights had a lower success rate (blue = fail)
- Later flights had a higher success rate (orange = success)
- Around half of launches were from CCAFS SLC 40 launch site
- VAFB SLC 4E and KSC LC 39A have higher success rates
- This data shows that the success rate of flight launches increased with the number of flights.

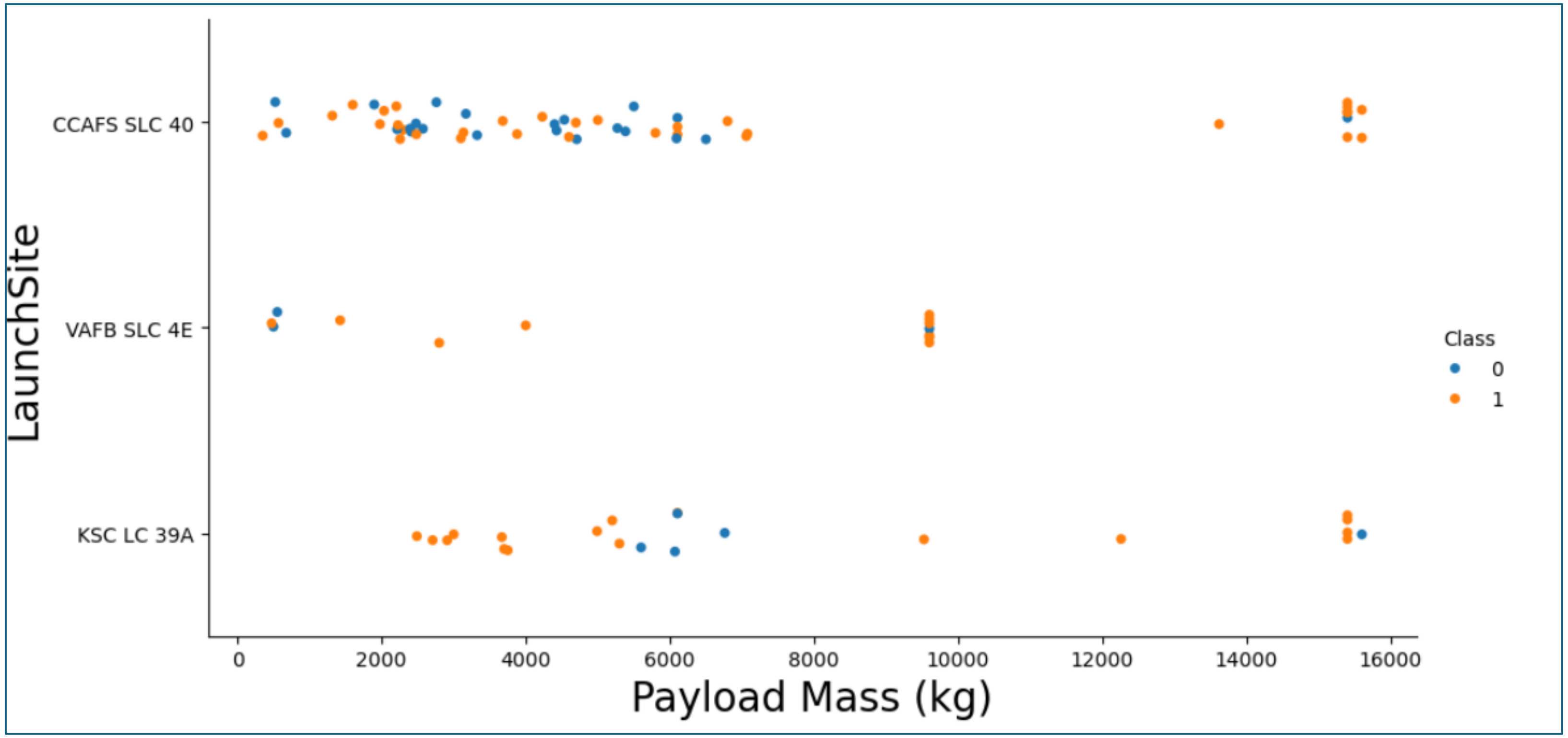


PAYLOAD VS. LAUNCH SITE

Exploratory Data Analysis

- Most launces that had a payload greater than 7,000 kg were successful
- KSC LC 39A has a 100% success rate for launches less than 5,500 kg
- VAFB SKC 4E has not launched anything greater than ~10,000 kg
- The data shows that the higher the payload mass (kg), the higher the success rate

Payload vs. Launch Site

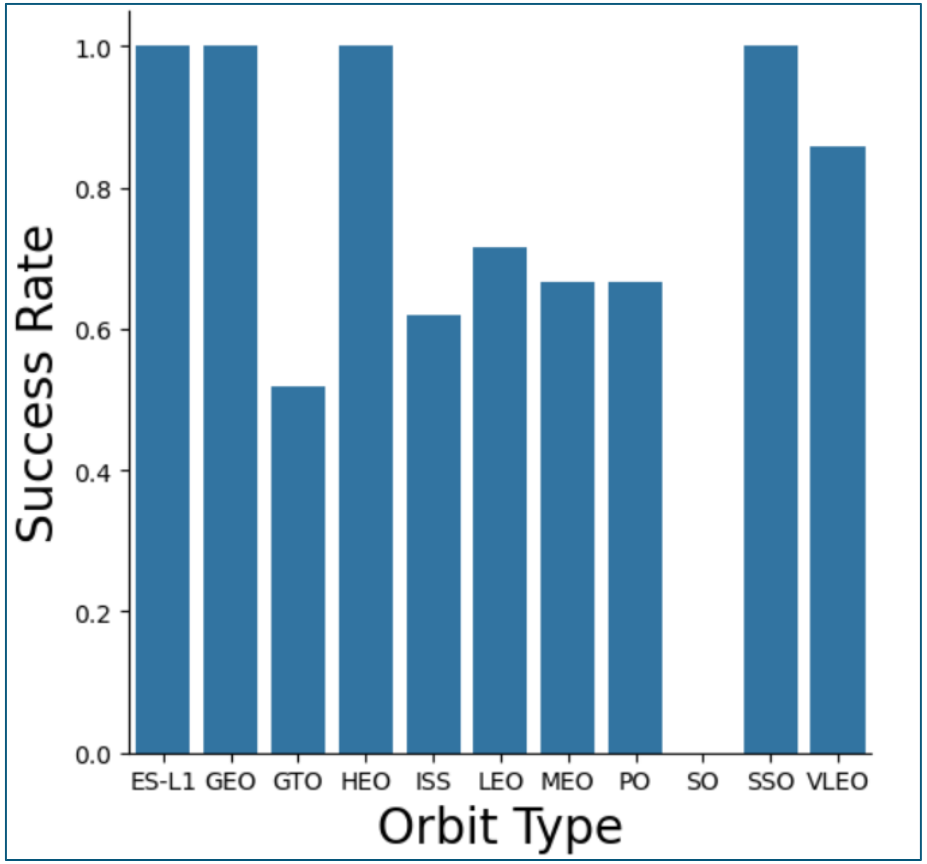


SUCCESS RATE BY ORBIT

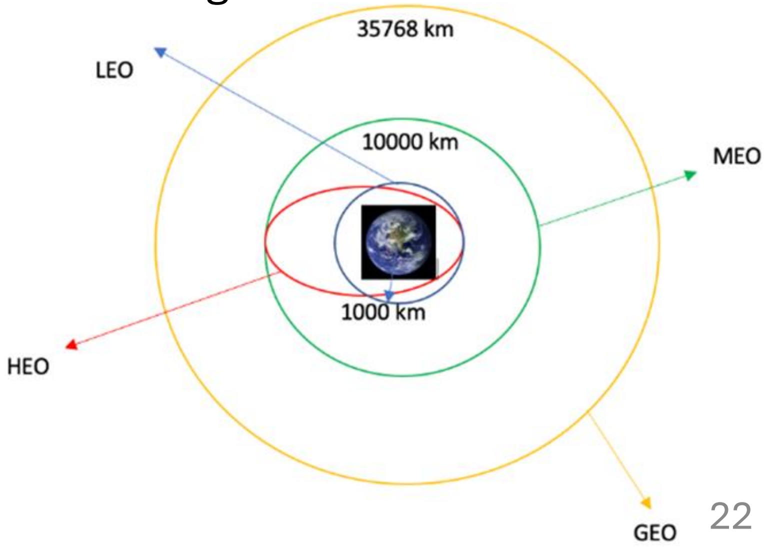
Exploratory Data Analysis

- 100% Success Rate: ES-L1, GEO, HEO and SSO
- 50%-80% Success Rate: GTO, ISS, LEO, MEO, PO
- 0% Success Rate: SO

Success Rate by Orbit



Orbit refers to which orbit the Falcon 9 is delivering space craft to. Its second stage engines are used to move the rocket to the correct orbit and we are interested in how this might affect the stage one landing.

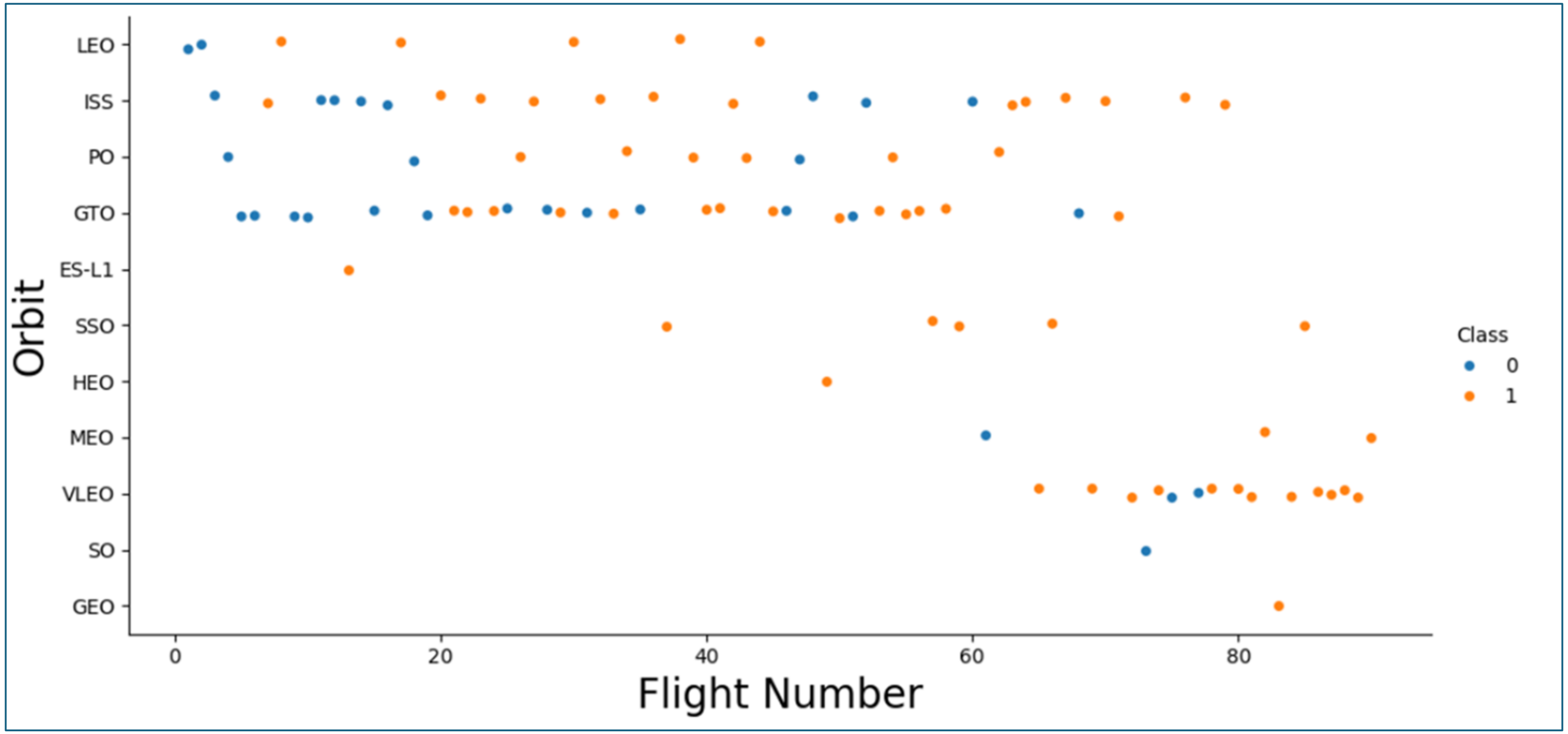


FLIGHT NUMBER VS. ORBIT

Exploratory Data Analysis

- The success rate typically increases with the number of flights for each orbit
- This increase is especially true for flights for the LEO and the ISS orbit
- The only orbit that does not follows this trend is the GTO orbit

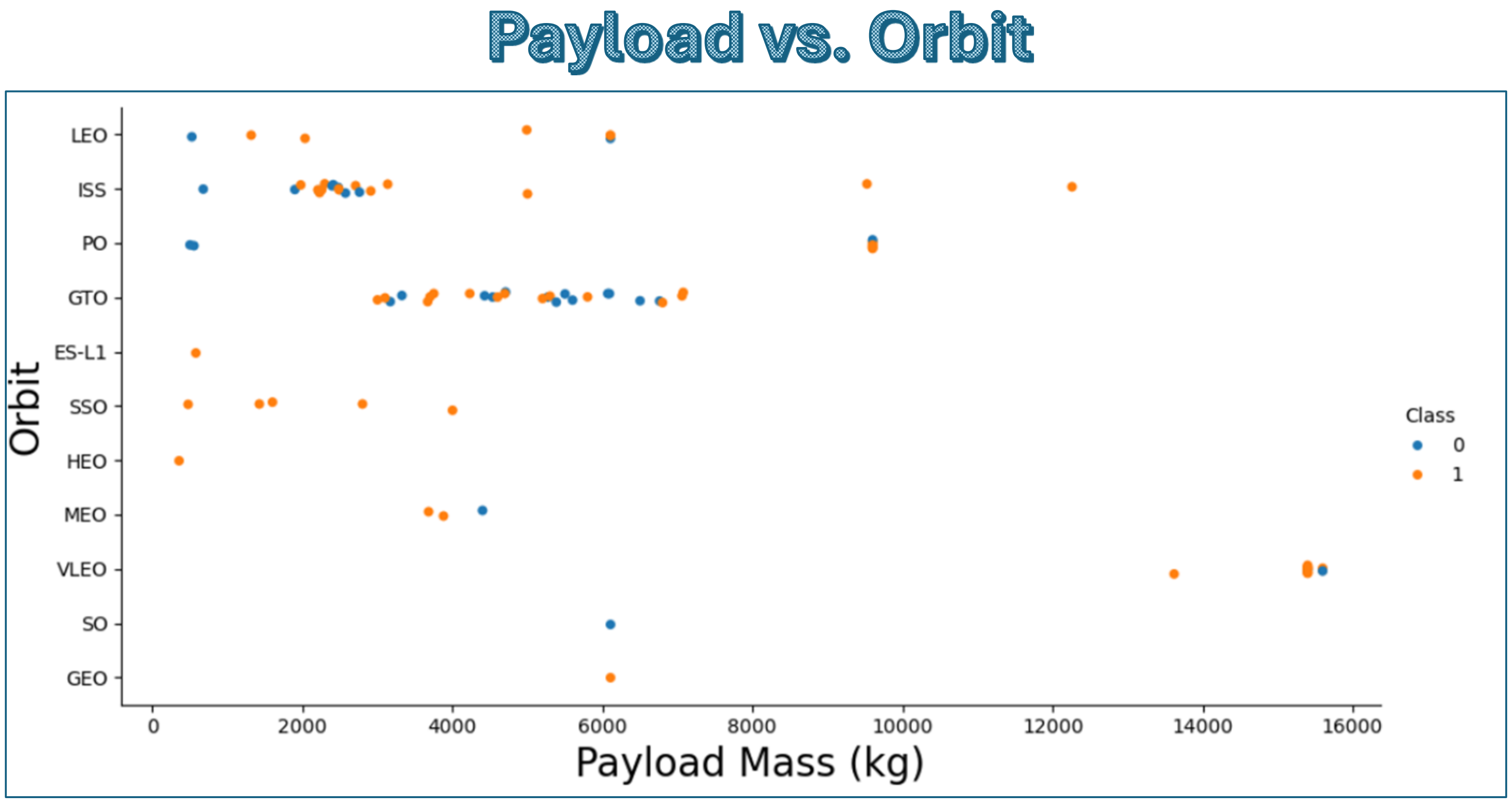
Flight Number vs. Orbit



PAYLOAD VS. ORBIT

Exploratory Data Analysis

- Heavy payloads are more successful with LEO, ISS, PO and VLEO orbits
- The SSO orbit has had the best results for the lightest payloads.
- The GTO orbit has mixed success with heavier payloads
- Lighter payloads are more successful with the remaining orbits.

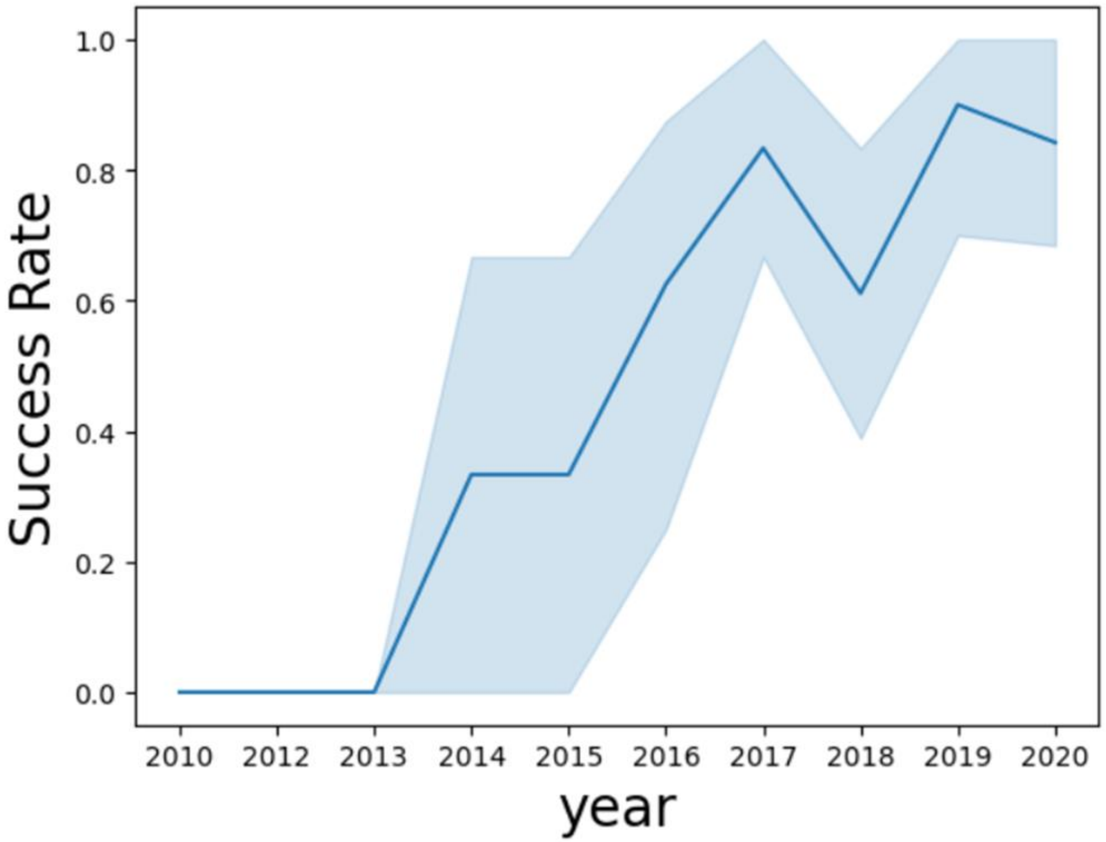


LAUNCH SUCCESS OVER TIME

Exploratory Data Analysis

- The success rate climbed at a high rate from 2013-2017 and from 2018-2019
- The success rate decreased a bit from 2017-2018 and from 2019-2020
- Overall, the success rate has improved since 2013 showing that with time, advancements and improvements in this technology have made a difference.

Launch Success over Time



LAUNCH SITE INFORMATION

EXPLORATORY DATA ANALYSIS (EDA) USING SQL

All Launch Site Names

```
%sql SELECT DISTINCT(Launch_Site) \
     FROM SPACEXTBL;
```

```
 * sqlite:///my_data1.db
Done.
```

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

Launch Site	Abbreviation
Cape Canaveral Space Launch Complex - Florida	CCAFS LC-40
Cape Canaveral Space Launch Complex - Florida	CCAFS SLC-40
Kennedy Space Center Launch Complex - Florida	KSC LC-39A
Vandenberg Space Launch Complex - California	VAFB SLC-4E

- SpaceX has launched the Falcon 9 rocket from these 4 launch sites. The Vandenberg Space Launch Complex is in California while the other three launch sites are near each other in Florida.
- The data collected included the launch site for each Falcon 9 launch.

Launch Site Names that Begin with 'CCA'

```
%sql SELECT * \
    FROM SPACEXTBL \
    WHERE LAUNCH_SITE LIKE'CCA%';
```

```
 * sqlite:///my_data1.db
Done.
```

Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer	Mission_Outcome	Landing_Outco
2010-06-04	18:45:00	F9 v1.0 B0003	CCAFS LC-40	Dragon Spacecraft Qualification Unit	0	LEO	SpaceX	Success	Failure (parachu
2010-12-08	15:43:00	F9 v1.0 B0004	CCAFS LC-40	Dragon demo flight C1, two CubeSats, barrel of Brouere cheese	0	LEO (ISS)	NASA (COTS) NRO	Success	Failure (parachu
2012-05-22	7:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525	LEO (ISS)	NASA (COTS)	Success	No atten
2012-10-08	0:35:00	F9 v1.0 B0006	CCAFS LC-40	SpaceX CRS-1	500	LEO (ISS)	NASA (CRS)	Success	No atten
2013-03-01	15:10:00	F9 v1.0 B0007	CCAFS LC-40	SpaceX CRS-2	677	LEO (ISS)	NASA (CRS)	Success	No atten
2013-			CCAFS LC-						

This is an example of a query that is pulling launches for all launch sites that began with “CCA” which will therefore pull all launches from Cape Canaveral only. This is important in comparing success rates at each launch site against each other.

Payload Mass

TOTAL PAYLOAD MASS

```
%sql SELECT SUM(PAYLOAD_MASS__KG_) \
    FROM SPACEXTBL \
    WHERE CUSTOMER = 'NASA (CRS)';
```

```
 * sqlite:///my_data1.db
Done.
```

SUM(PAYLOAD_MASS__KG_)
45596

The total payload mass carried by SpaceX boosters is 45,596 kg (total)

Average Payload Mass by F9 v1.1

```
%sql SELECT AVG(PAYLOAD_MASS__KG_) \
    FROM SPACEXTBL \
    WHERE BOOSTER_VERSION = 'F9 v1.1';
```

```
 * sqlite:///my_data1.db
Done.
```

AVG(PAYLOAD_MASS__KG_)
2928.4

The average payload carried by the booster version F9 v1.1 was 2,928 kg.

Mission Outcomes

Total Number of Successful and Failed Mission Outcomes from 3/24/2006 and 11/05/2020

- 1 Failure in Flight
- 99 Successful Flights
- 1 Success (where payload status unclear)

```
%sql SELECT MISSION_OUTCOME, COUNT(*) as total_number \
FROM SPACEXTBL \
GROUP BY MISSION_OUTCOME;
```

```
 * sqlite:///my_data1.db
Done.
```

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

1st Successful Landing of Stage One Booster

```
%sql SELECT MIN(DATE) \
FROM SPACEXTBL \
WHERE LANDING_OUTCOME = 'Success (ground pad)'
```

```
 * sqlite:///my_data1.db
Done.
```

MIN(DATE)
2015-12-22

The first successful ground landing of a Falcon 9 Stage One was on December 12, 2015.

Booster Information

```
%sql SELECT BOOSTER_VERSION \
FROM SPACEXTBL \
WHERE PAYLOAD_MASS__KG_ = (SELECT MAX(PAYLOAD_MASS__KG_) FROM SPACEXTBL);
```

```
 * sqlite:///my_data1.db
Done.
```

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

This queried table shows the max payload per booster.

Booster Drone Ship Landing

Successful landing for ships with a booster mass greater than 4,000 but less than 6,000 included the JSCAT-14, JSCAT-16, SES-10, and the SES-11 / EchoStar 105

```
%sql SELECT PAYLOAD \
FROM SPACEXTBL \
WHERE LANDING_OUTCOME = 'Success (drone ship)' \
AND PAYLOAD_MASS__KG_ BETWEEN 4000 AND 6000;
```

```
 * sqlite:///my_data1.db
Done.
```

Payload
JCSAT-14
JCSAT-16
SES-10
SES-11 / EchoStar 105

Failed Landing Outcomes in 2015

We can also query failed booster landings. Here wee can see that in 2015, there were two failed (drone ship) booster landings. These occurred on January 10 and April 14.

```
%sql SELECT substr(Date,6,2) as month, DATE, BOOSTER_VERSION, LAUNCH_SITE, [Landing_Outcome] \
FROM SPACEXTBL \
where Landing_Outcome = 'Failure (drone ship)' and substr(Date,1,4)='2015';
```

```
 * sqlite:///my_data1.db
Done.
```

month	Date	Booster_Version	Launch_Site	Landing_Outcome
01	2015-01-10	F9 v1.1 B1012	CCAFS LC-40	Failure (drone ship)
04	2015-04-14	F9 v1.1 B1015	CCAFS LC-40	Failure (drone ship)

We are More Interested in Success

Here we have queried the count of landing outcomes between 2010-06-04 and 2017-03-20 in descending order.

Of attempted landings during this time period:

- 13 were successful
- 3 were planned failures
- 4 were unsuccessful
- And 1 was precluded (prevented)

```
%sql SELECT [Landing_Outcome], count(*) as count
     FROM SPACEXTBL \
     WHERE DATE between '2010-06-04' and '2017-0
```

```
 * sqlite:///my_data1.db
Done.
```

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

Visual/Interactive Analytics

Launch Site
Proximity Analysis

LAUNCH SITES

	Launch Site	Lat	Long
0	CCAFS LC-40	28.562302	-80.577356
1	CCAFS SLC-40	28.563197	-80.576820
2	KSC LC-39A	28.573255	-80.646895
3	VAFB SLC-4E	34.632834	-120.610745

Visual/Interactive Analysis

- The markers show the location of the 4 SpaceX launch sites. 1 in California and 3 in Florida.

All are as near the equator as possible

- The closer the launch site is to the equator, the easier it is to launch a rocket into orbit .
- Rockets launched from sites near the equator get an additional natural boost - due to the rotational speed of earth. This helps save on the cost of putting in extra fuel and boosters .

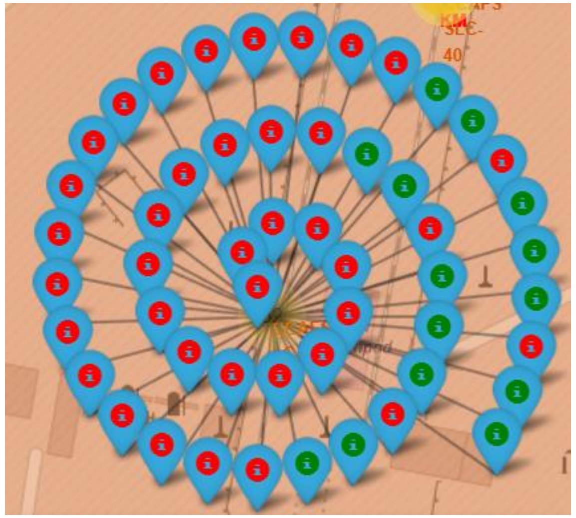
LAUNCH OUTCOMES

Visual/Interactive Analysis

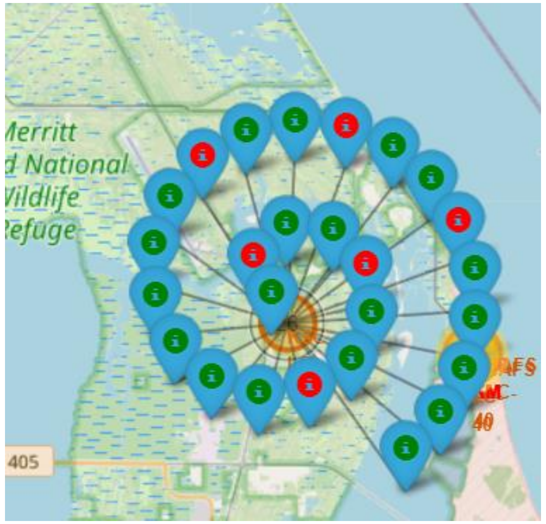
- This visual shows the number of successful and unsuccessful launch outcomes (stage one landings) at each launch site. Green markers show successful launches and red markers show unsuccessful launches.
- Launch site KSC LC-39 has the most successful launches at 20/27 launches or a 76.9% success rate.

VAFB SLC-4E
(California)

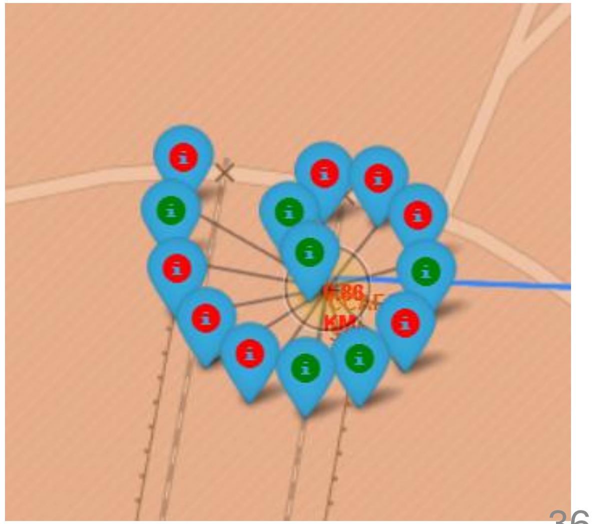
CCAFS SLC-40
(Florida)



KSC LC-39A
(Florida)

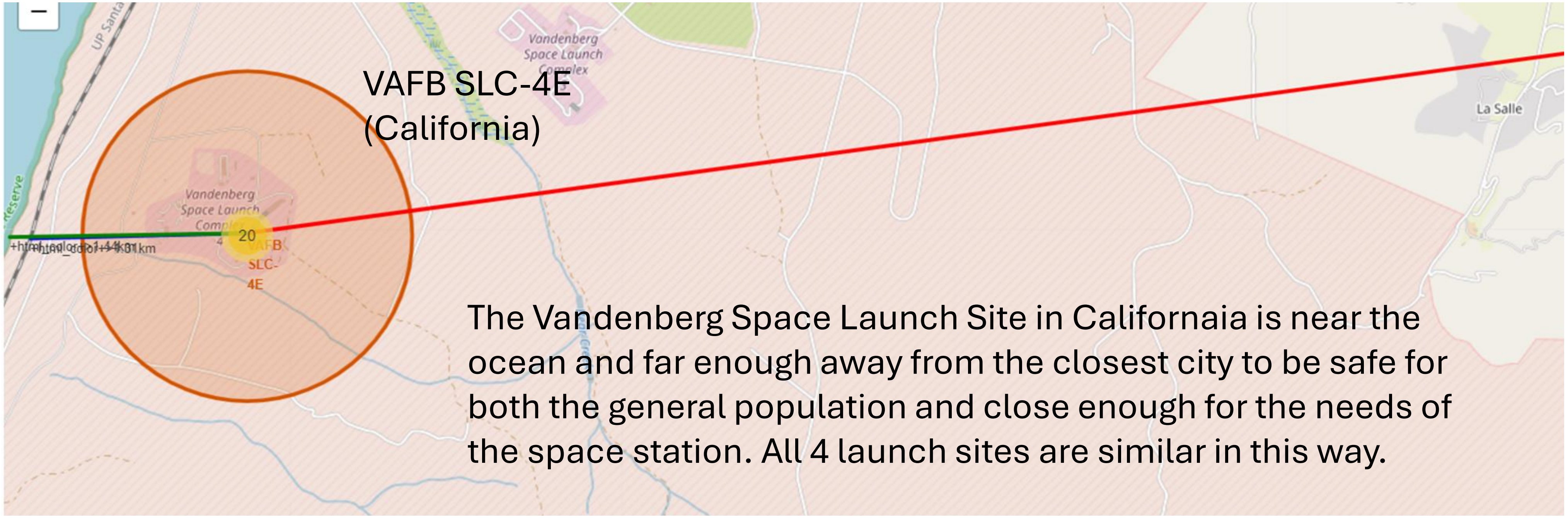


CCAFS SLC-40
(Florida)



DISTANCE TO PROXIMITIES

Visual/Interactive Analysis
It is important to note the distance a launch pad is from proximities for safety and security.
Safety: to ensure the rocket launches are far enough away from the general population as possible.
Security: an exclusion zone around a launch pad to keep unauthorized people out and keep everyone as safe as possible.

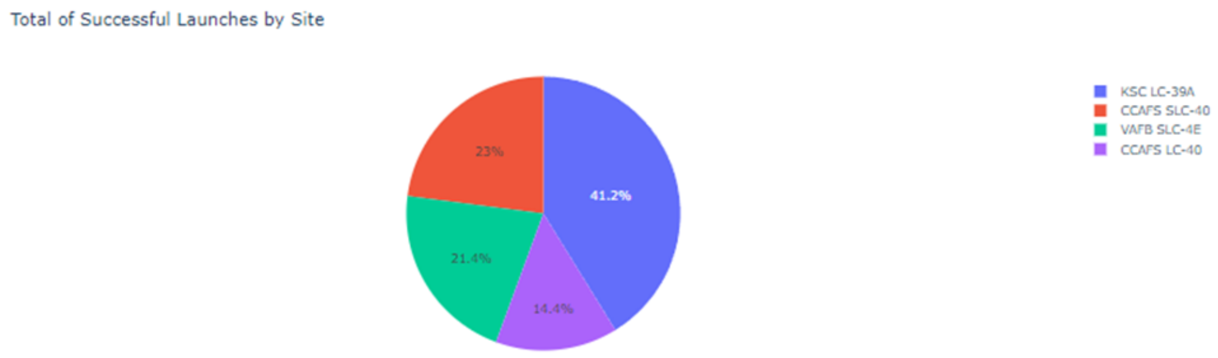


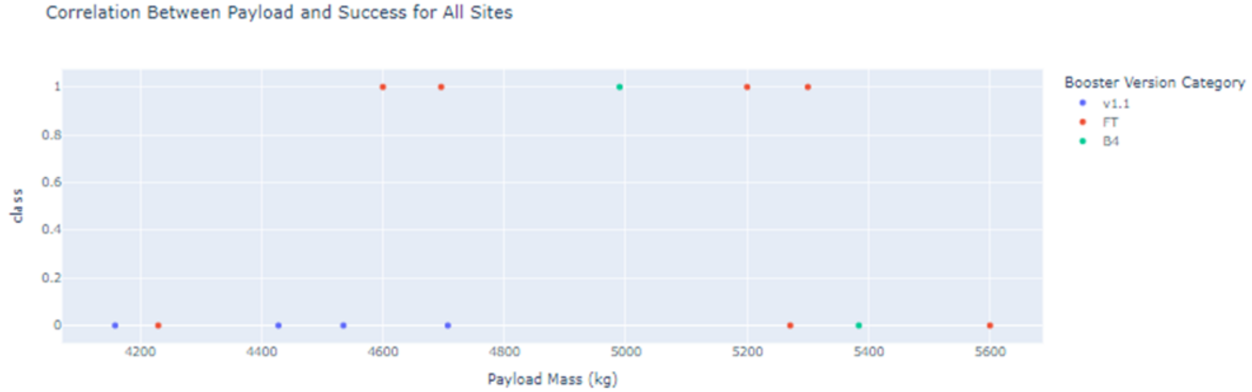
VAFB SLC-4E (California)

The Vandenberg Space Launch Site in Californaia is near the ocean and far enough away from the closest city to be safe for both the general population and close enough for the needs of the space station. All 4 launch sites are similar in this way.

INTERACTIVE ANAYLTYICS USING DASHBBOARD WITH PLOTLY

This interactive dashboard shares data about the success rate of each launch site using a drop-down menu. It can also be used to show the success rate by payload size by adjusting the range on the payload range bar.

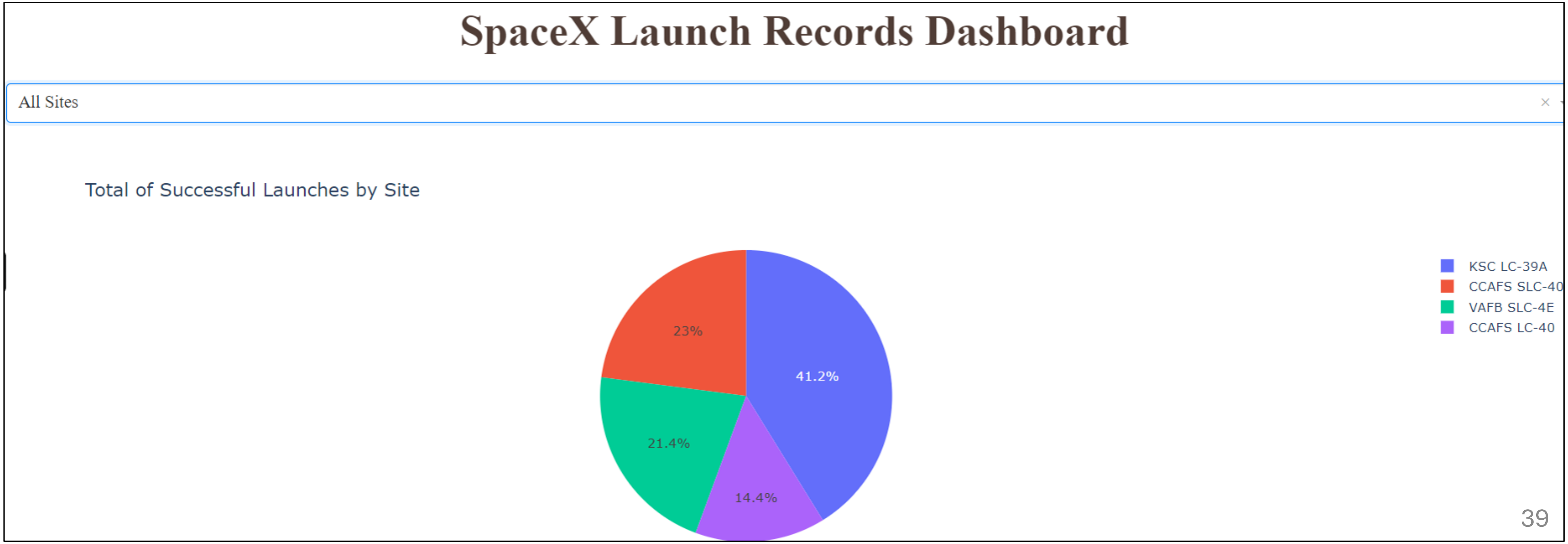




Launch Success by Site

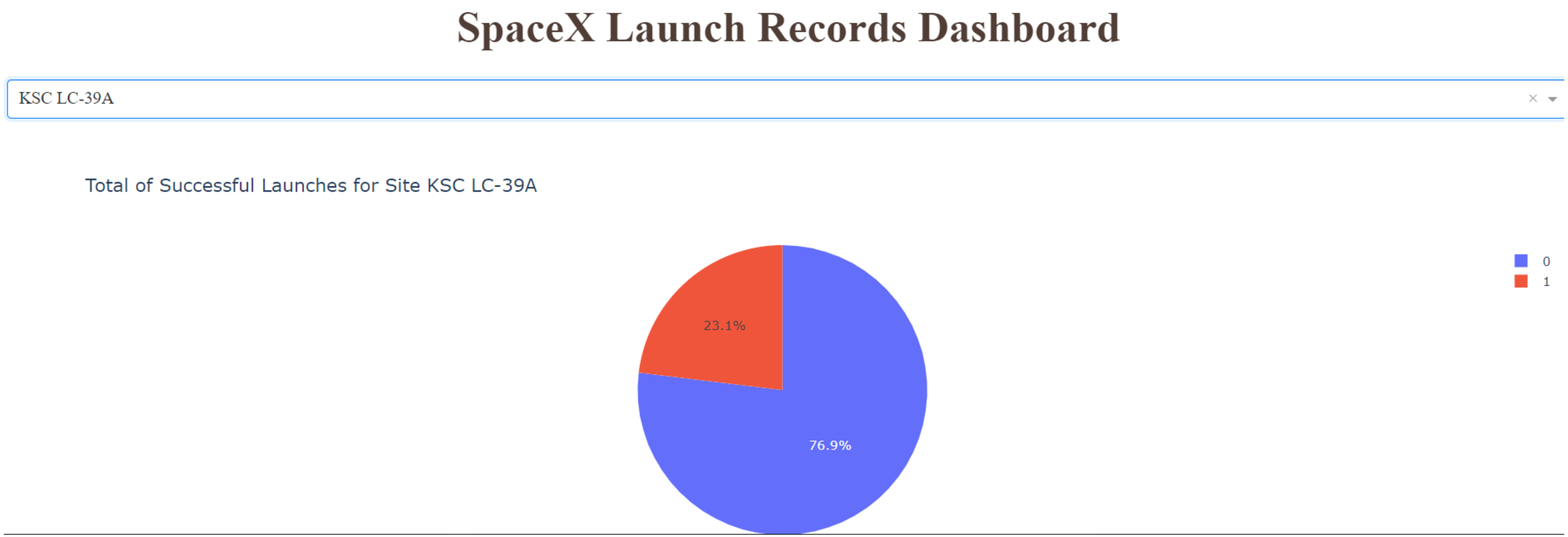
Success as Percent of Total

- KSC LC-39A has the most successful launches amongst launch sites at 41.2%



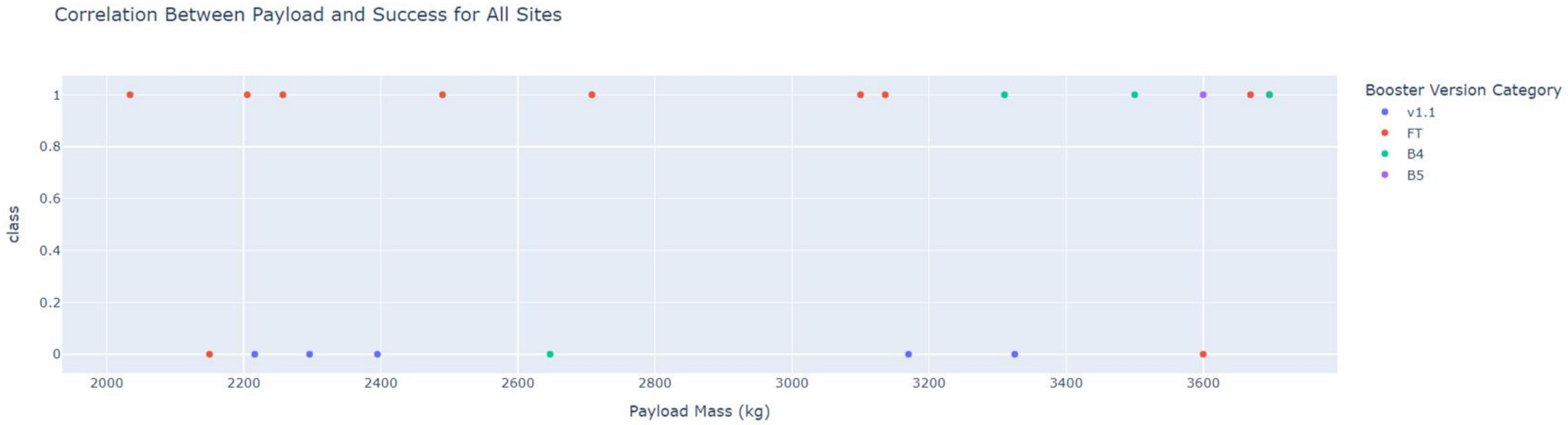
LAUNCH SITE WITH THE HIGHEST LAUNCH SUCCESS RATIO

KSC LC-39A has the highest success rate of stage one landings among the four launch sites (76.9%)



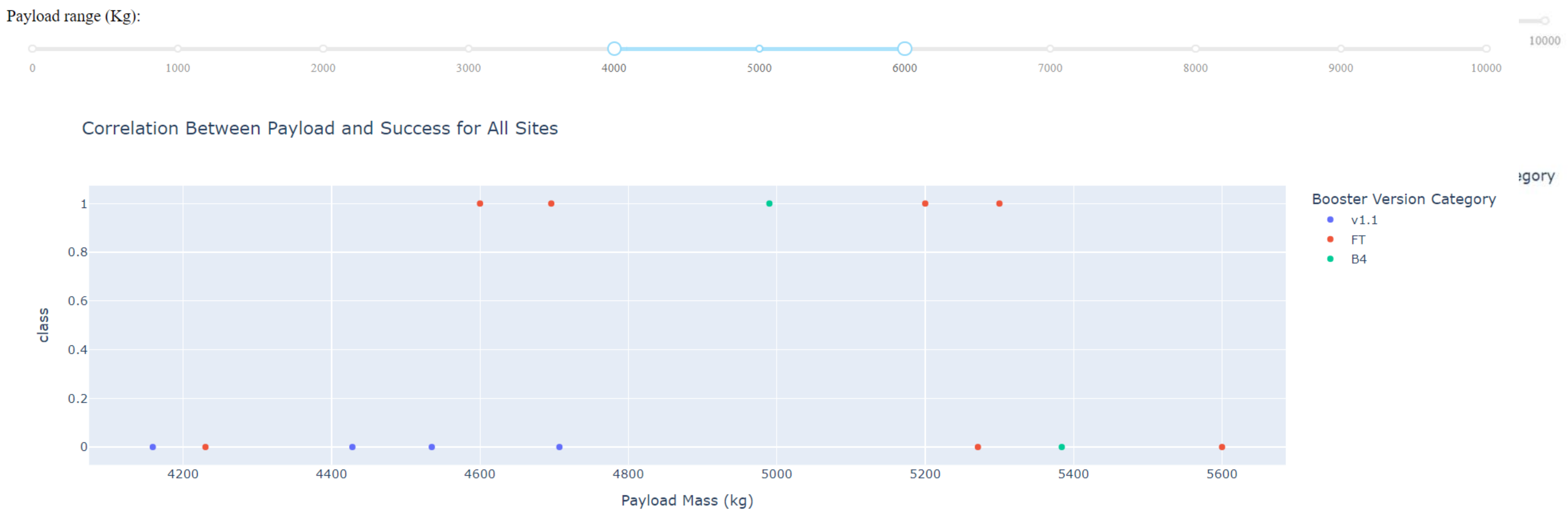
PAYLOAD VS LAUNCH OUTCOME

- Payloads between 2,000 kg and 4,000 kg have the highest success rate
- 1 indicating successful outcome and 0 indicating an unsuccessful outcome
- Most of the successful outcomes were from launches using the FT Booster Version



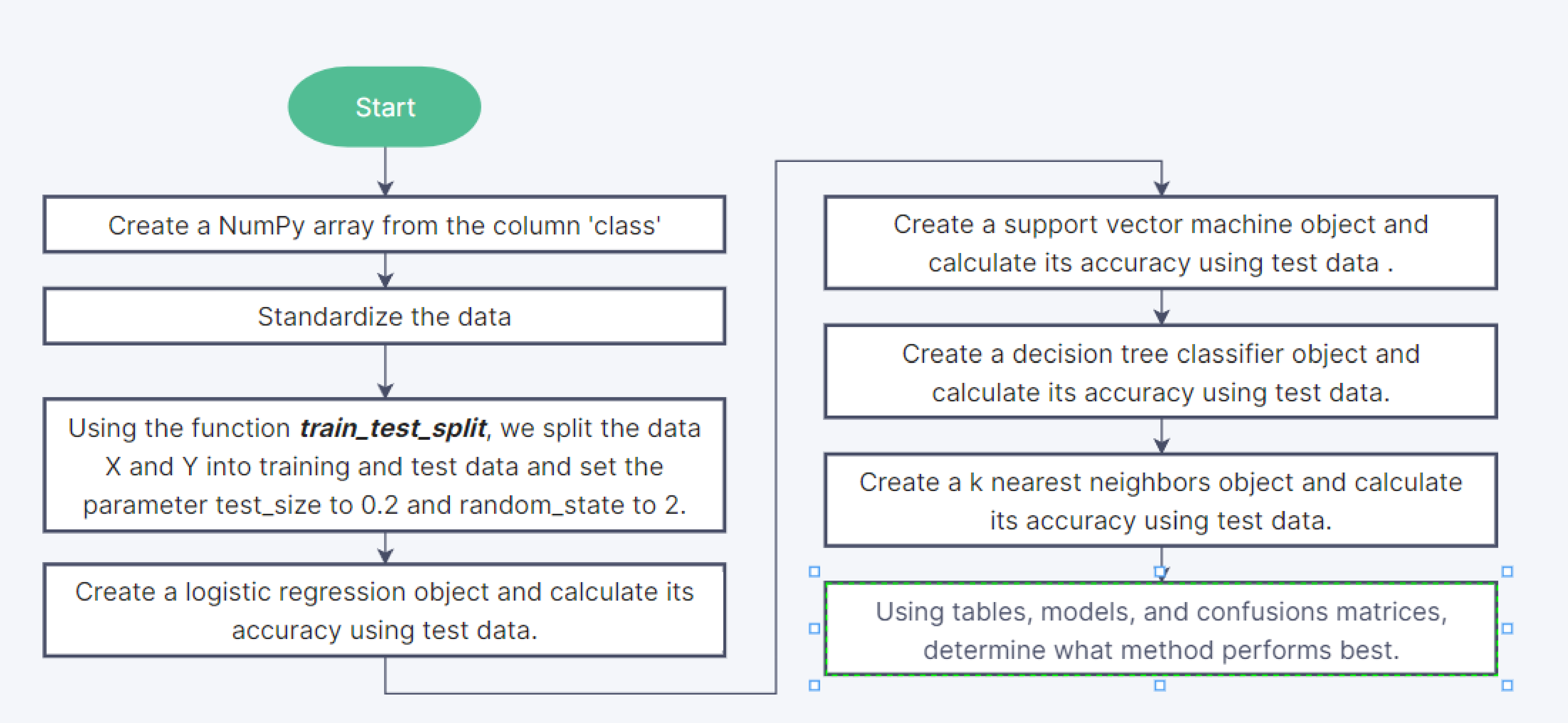
PAYLOAD VS LAUNCH OUTCOME

- Payloads between 4,000 kg and 6,000 kg have the lowest success rate
- 1 indicating successful outcome and 0 indicating an unsuccessful outcome
- There is no difference in booster version at this payload.



Predictive Analysis (Classification)

METHOD OF PREDICTIVE ANALYSIS



PREDICTIVE ANALYSIS (CLASSIFICATION)

The class values were put into an array using numpy.

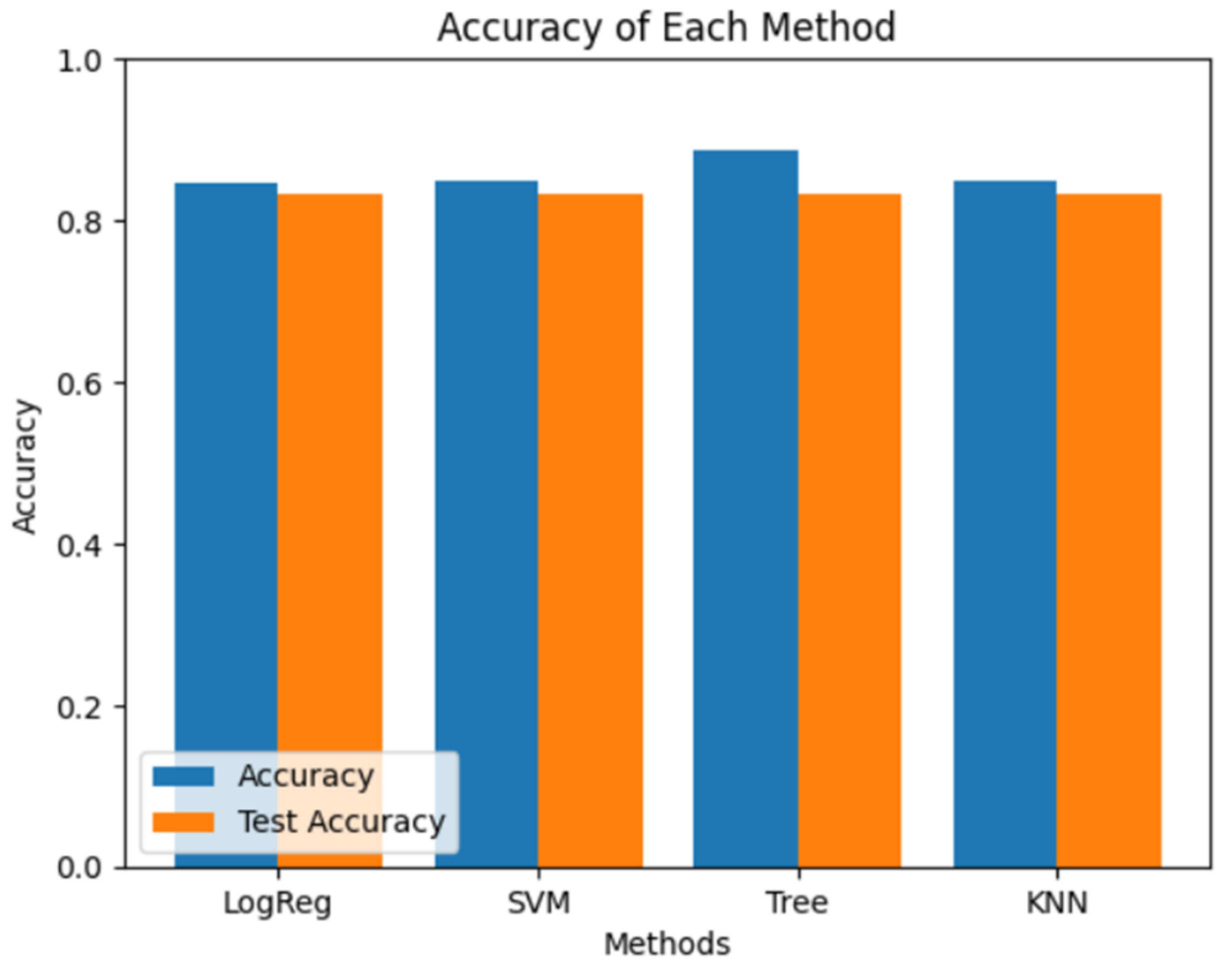
```
array([0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 1, 1, 1,
       1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1,
       1, 0, 0, 0, 1, 1, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1, 1, 1, 1, 1,
       1, 0, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
       1, 1], dtype=int64)
```

- The data was then standardized before being split in to training data and testing data using the train_test_split function. 80% of the data was used for training and 20% was used for testing.
- Then four models were created with the training and testing data.
    - logistic regression(LogisticRegression())
    - support vector machine (SVC())
    - decision tree (DecisionTreeClassifier())
    - K-Nearest Neighbor (KNeighborsClassifier())

CLASSIFICATION ACCURACY

- All the models performed at about the same level and had the same scores and accuracy. This is likely due to the small dataset.
- The Decision Tree model slightly outperformed the rest when looking at .best_score_
- .best_score_ is the average of all cv folds for a single combination of the parameters

Model	Accuracy	Test Accuracy
LogReg	0.84643	0.83333
SVM	0.84821	0.83333
Tree	0.8875	0.83333
KNN	0.84821	0.83333



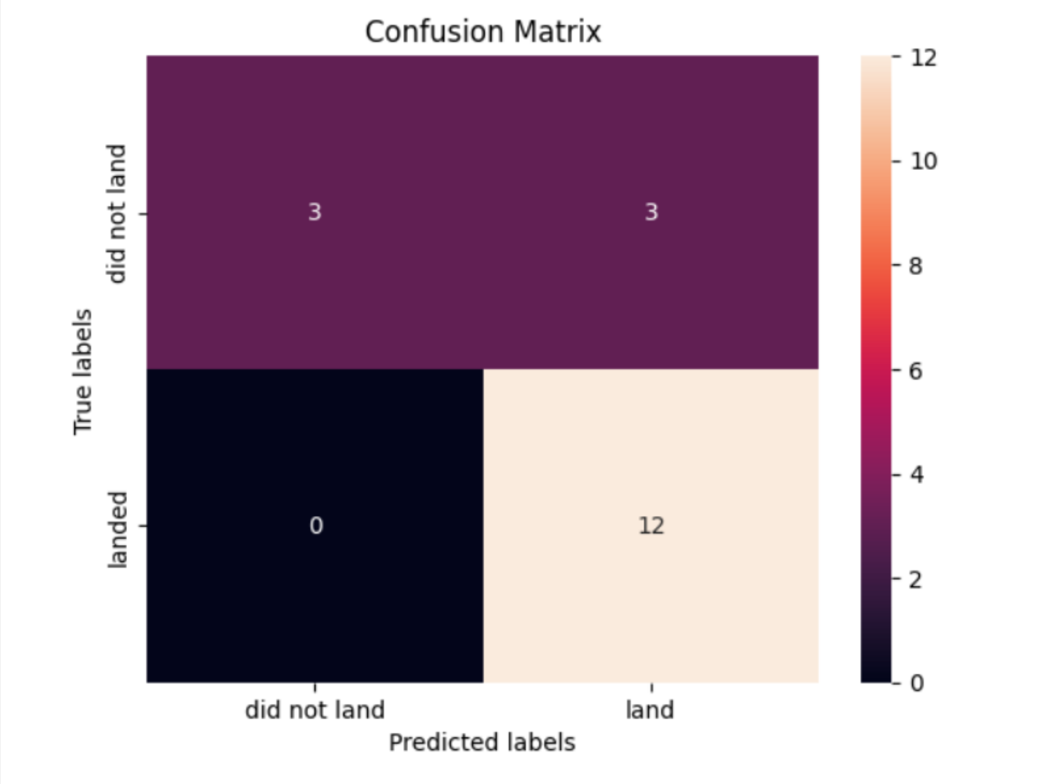
PERFORMANCE SUMMARY WITH THE CONFUSION MATRIX

- A confusion matrix summarizes the performance of a classification algorithm

- All confusion matrices were identical for all four models.

<u>Confusion Matrix Outputs:</u>

- 12 True positive
- 3 True negative
- 3 False positive
- 0 False Negative



Accuracy = (TP + TN) / (TP + TN + FP + FN) = .833

CONCLUSIONS

CONCLUSIONS

FROM THE RESEARCH

Note: Launch success refers to Falcon 9's stage one booster landing successfully.

- Launch success has improved as the number of flights have increased over time. This shows success in improvements made.
- The KSC LC-39A launch site has highest success rate among launch sites. In fact, it has a 100% success rate for launches less than 5,500 kg
- Launches to orbits ES-L1, GEO, HEO, and SSO have a 100% success rate in landing their stage one booster.
- On the other hand, launches to LEO and the ISS orbits have shown the greatest increase over time.
- Heavy payloads are more successful with LEO, ISS, PO and VLEO orbits while the lighter payloads do best with the SSS orbit.
- Across all launch sites, the higher the payload mass (kg), the higher the success rate
- However, the FT booster has had the highest success rate with payloads between 2000 kg and 4000 kg.

CONCLUSIONS

FROM THE RESEARCH (Continued)

• Predictive analysis shows that all four models perform similarly on the test set with the decision tree model slightly outperforming when using the .best_score_ function.

•All four models showed an 83.3% success rate with the test data.

• Most of the launch sites are set as close to the equator as possible. This is because rockets receive an additional natural boost from the rotational speed of earth . This helps save the costs of putting in extra fuel and boosters.

• All the launch sites are close to the coast for security and safety reasons.

The factors that affect the success of landing a Falcon 9 stage one booster include the following:

- **Improvements made over time in technology and booster selection**
- **Launch Site Location**
- **Destination Orbit**
- **Payload Mass**

CONCLUSIONS

Things to Consider

- A larger data set may help with the predictive analytics, however as the data set grows with time, the factors also continue to change and improve.
- It would be interesting to compare the early data to more current data. What is the success rate over the last year alone?
- What new technology will help improve success rate on landing first stage boosters?
- An 83.3% predictive success rate is considered good. How does this compare to how the Falcon 9 is doing today?
- What other improvements in technology could decrease the cost of a rocket launch in the future?
- How many times on average has a Falcon 9 stage one booster been reused?
- I am curious about the number of objects we have put into space, which I know is substantial, and how will this affect us long-term.

Appendices

Appendix A

What is the Falcon 9?

FALCON 9

OVERVIEW

HEIGHT	**70 m** / 229.6 ft
DIAMETER	**3.7 m** / 12 ft
MASS	**549,054 kg** / 1,207,920 lb
PAYLOAD TO LEO	**22,800 kg** / 50,265 lb
PAYLOAD TO GTO	**8,300 kg** / 18,300 lb
PAYLOAD TO MARS	**4,020 kg** / 8,860 lb

Source: https://www.spacex.com/vehicles/falcon-9/

Appendix B

Current Statistics on the Falcon 9 First Stage Landing

341	298	272
TOTAL LAUNCHES	TOTAL LANDINGS	TOTAL REFLIGHTS

Total Landings Success Rate: 87.3%
Total Reflights Success Rate: 79.8%

Source: https://www.spacex.com/vehicles/falcon-9/

Thank You