IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Karen Stagg

June 06, 2022

Outline

Executive Summary

- SpaceX is very competitive in the rocket launch business because it is able to reuse the first stage of its Falcon 9 rockets, cutting its costs (estimated at 62 million dollars) down to that of less than one third of its competitors price (estimated at upwards of 165 million dollars).
- If the success of landing of the first stage can be determined, then launch cost can also be determined.
- This project uses data collection and analysis measures using data collected from API import requests from HTTP web requests along with collecting web-scraped data on SpaceX Falcon 9 rocket launches, using the booster version F9.
- Data wrangling allows us to classify the successes and failures, and database SQL inquiries allow for further insight.
- Data visualization techniques using maps, graphs, charts and a range slider further allow for finding relationships and further data exploration and analysis.
- Lastly, Machine learning algorithms based of this data were utilized and determined with 83.34% accuracy the probability of the first stage landing success to help answer the question of whether landing success is a viable predictor for determining launch cost.

Introduction

- Can SpaceY compete with SpaceX when pricing its' launches?
- We know that SpaceX is able to charge less than one-third the price of its' competitors for a launch. This is mainly due to the fact that SpaceX reuses its' first stage, which significantly cuts the cost of a launch.
- Because SpaceX has been launching rockets for more than ten years, we have an abundance of data that can be analyzed for patterns and then use machine learning modeling to help predict the likelihood of the first stage landing successfully so that it can be re-used.
- We can use this knowledge to help SpaceY competitively price a launch.

Section 1

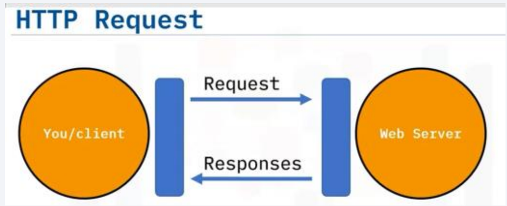
Methodology

Methodology

Executive Summary

- Data collection methodology:
  - API import requests from HTTP requests of the SpaceX API and creating a dataframe of this information
  - Webscraping using Beautiful Soup the Falcon 9 launch records from a HTML table from WIkipedia
- Perform data wrangling
  - NaN values were identified, bad_outcomes were identified and made into a set and used in comparison to create a 'Class' column where 0 meant a bad outcome and a 1 meant a successful outcome. This helped to calculate the mean of successful outcome of 66%
- Perform exploratory data analysis (EDA) using visualization and SQL
- Perform interactive visual analytics using Folium and Plotly Dash
- Perform predictive analysis using classification models
  - Build multiple classification models, split data to train and test, evaluate accuracy scores, and confusion matrix to help determine best model

Data Collection- SpaceX API



Spacex_url = "https://api.spacexdata.com/v4/launches/past"

Make a 'Get' Request requests.get(spacex_url)

Get a Response back as a JSON

Normalize JSON response and convert result into a pandas dataframe

Take a subset of the dataframe keeping only relevant information

Use the API again to get specific data on launches given ID's for each launch and store inside global vaiable lists

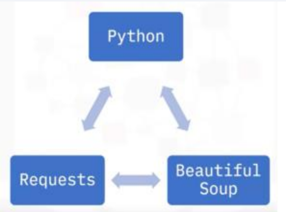
Fill the global variable lists by running the pre-defined methods provided

Construct the dataset by combining columns into a dictionary of the filled global variables

Create a pandas dataframe from the dictionary and save as a csv

[Capstone/jupyter-labs-spacex-data-collection-api (1).ipynb at master · karenstagg/Capstone (github.com)]

Data Collection - Scraping



static_url = "https://en.wikipedia.org/w/index.php?title=List_of_Falcon_9_and_Falcon_Heavy_launches&oldid=1027686922"

Make a 'Get' Request requests.get(static_url).text

Create a Beautiful Soup object from the HTML response

Extract all the column/variable names from the HTML table header

From the column names <th> elements, create a dictionary

Rid irrelevant data and initialize the dictionay with each value as an empty list.

Loop through and parse the beautiful soup object populating the empty lists inside the dictionary with the relevant row values

Create a pandas dataframe from the dictionary and save as a csv

Capstone/jupyter-labs-webscraping.ipynb at master · karenstagg/Capstone (github.com)

Data Wrangling

Load the csv file from the API exercise

Identify the attributes which have missing values

Perform Calculations: - the number of launches on each site, -number & occurrence of each of 11 orbits, and of mission outcome per orbit type

Loop through the mission outcomes and setup 0-7 numerical keys for the 8 types

Create a set of bad outcomes using the keys for bad badoutcomes

Create a list and loop through the outcomes placing bad outcomes in the list as a '0' (failure) otherwise as a '1'(success)

Set the attribute 'Class' to be equal to the contents of the list(thereby making contents only 0 or 1)

Since the Class is now only numeric, you can cetermine the success rate mean of 66%

Save the file as a csv

Capstone/labs-jupyter-spacex-Data wrangling (1).ipynb at master · karenstagg/Capstone (github.com)

EDA with Data Visualization

- In this section, we used the output csv file from the data wrangling section to visualize the results in different ways in order to visualize how each of the variables impacts the success rate.
- Scatterplots were used to show relationships between flight number and launch site, between payload and launch site, flight number and orbit type, and payload and orbit type.
- A bar chart showed the relationship between success rate and orbit type.
- A line chart was used to show relationships between date and class in order visualize the launch success yearly trend. It was easy to observe that the success rate since 2013 kept increasing until 2020.
- A features dataframe was created with selected variables that will be used in success prediction. Dummy variables were assigned to some of the variables using one hot encoding to be able to have the features df completely numeric for easy manipulation.
- URL: Capstone/jupyter-labs-eda-dataviz (1).ipynb at master · karenstagg/Capstone (github.com)

EDA with SQL

DB2 table queries were done on the converted spacex.csv file.

- Names of unique launch sites in the space mission
- Records of launch sites beginning with 'CCA'
- The total calculated payload mass carried by boosters launched by NASA (CRS)
- The average payload mass carred by booster version F9 v1.1
- The date of the first successful landing outcome in ground pad
- The names of the boosters having success in drone ship landing and having a payload between 4000-6000 kg
- The total number of success and failure mission outcomes
- The names of the booster versions which have carried the maximum payload mass of 15600 kg
- The failed landing outcomes in drone ship landing , their booster versions, and launch site names for the year 2015
- Descending rankings by count of landing outcomes between 06/04/10 and 03/20/20
- URL:Capstone/jupyter-labs-eda-sql-coursera.ipynb at master · karenstagg/Capstone (github.com)

Build an Interactive Map with Folium

- All the launch sites, as well as the initial center to be NASA Johnson Space Center in Houston, Texas, were added to the map by their lat and long coordinates, with circles and markers for each.
- Each of the success/failed launches were added to the map as color-coded marker clusters. Red for failed and green for success missions.
- Distances between launch sites and proximities were calculated. I used the CCAFS SLC-40 site coordinates and identified the closest coastline, with a folium marker was created showing the distance, and a polyline was drawn between the launch site and coastline marker.
- CCAFS SLC-40 site was again used to find the closest city of Titusville, the closest railway of NASA Rail, and the closest highway of Samuel C. Phillips Pkwy. Distance markers and poly lines were drawn for all 3.
- URL: Capstone/lab_jupyter_launch_site_location (2).ipynb at master · karenstagg/Capstone (github.com)

Build a Dashboard with Plotly Dash

The Dashboard application contains:

- an input component of a drop-down list of individual launch sites as well as the category 'all' if visualizations want to be seen collectively for all launch sites.
- a payload range slider, where the uses can choose specif payload ranges from the minimum to maximum kg.
- pie chart indicating landing success for a site. Color-coding for '0' is failure, and '1' is success. If 'all' is chosen, then the color-coded percentage of successes at each site is shown.
- and a scatter plot to display all values for Payload Mass (kg) and the variable class and using a point color chart to show the booster version categories, where class of failure = '0', and class of success = '1'. This helps to observe mission outcomes with different boosters.

URL: Capstone/spacex_dash.py at master · karenstagg/Capstone (github.com)

Predictive Analysis (Classification)

2 csv files were read in to obtain our X and Y data.

Y data was created from a numpy array from the variable 'Class' in the csv X data was read in from a second csv and transformed

A test-train split of 80/20 was made from the X and Y data of 90 records. The split resulted in 72 records for training and 18 records for testing.

4 Classification models were used:
- Logistic Regression (LR),
- Support Vector Machine (SVM),
- Decision Tree (DT), and
- K Nearest Neighbors (KNN)

For each model object, a set of parameters was given, and a GridSearchCV object was created and trained

The best parameters for each model 's accuracy were computed using the data attribute best_params

The accuracy of the validation test data is calculated using the method 'score'

A confusion matrix was prepared for each model, and further accuracy metrics were computed with a classification report, jaccard and f1 scores, and logloss for LR

A report was prepared reflecting a matrix of all scores

URL: [Capstone/Spacex Machine Learning Prediction(1).ipynb at master · karenstagg/Capstone (github.com)]

Results

- <u>Exploratory data analysis results (Helpful SQL and Visualization Examples)</u>

Display the names of the unique launch sites in the space mission

```
%sql SELECT DISTINCT(Launch_Site) from SPACEXTBL
```

```
 * ibm_db_sa://zbb82269:***@54a2f15b-5c0f-46df-8954-7e38e612c2
Done.
```

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

Identify Unique Launch Sites

List the total number of successful and failure mission outcomes

```
%%sql SELECT COUNT(mission_outcome) AS success_mission_counts FROM SPACEXTBL
WHERE mission_outcome LIKE 'S%'
```

```
 * ibm_db_sa://zbb82269:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0
Done.
```

success_mission_counts
100

Count the total Number of Successful Missions

Rank the count of landing outcomes (such as Failure (drone ship) or Success (ground pad)) between the date 2010-06-04 and 2017-03-20, in descending order

```
%%sql SELECT landing__outcome, COUNT(landing__outcome) AS count FROM SPACEXTBL
WHERE date BETWEEN '2010-06-04' AND '2017-03-20'
GROUP BY landing__outcome
ORDER BY count DESC
```

```
 * ibm_db_sa://zbb82269:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32733/BLUDB
Done.
```

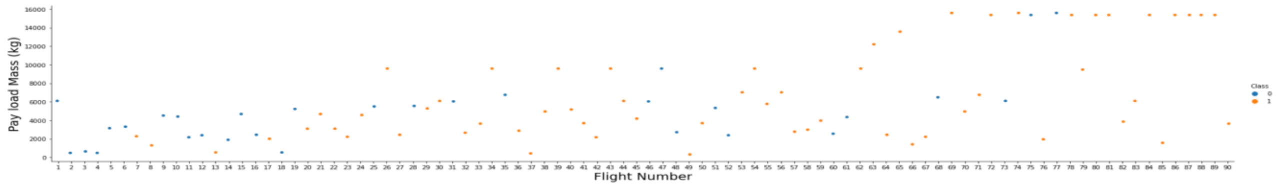
landing__outcome	COUNT
No attempt	10
Failure (drone ship)	5
Success (drone ship)	5
Controlled (ocean)	3
Success (ground pad)	3
Failure (parachute)	2
Uncontrolled (ocean)	2
Precluded (drone ship)	1

Descending Count of Landing Outcome Scenarios

First, let's try to see how the `FlightNumber` (indicating the continuous launch attempts.) and `Payload` variables would affect the launch outcome.

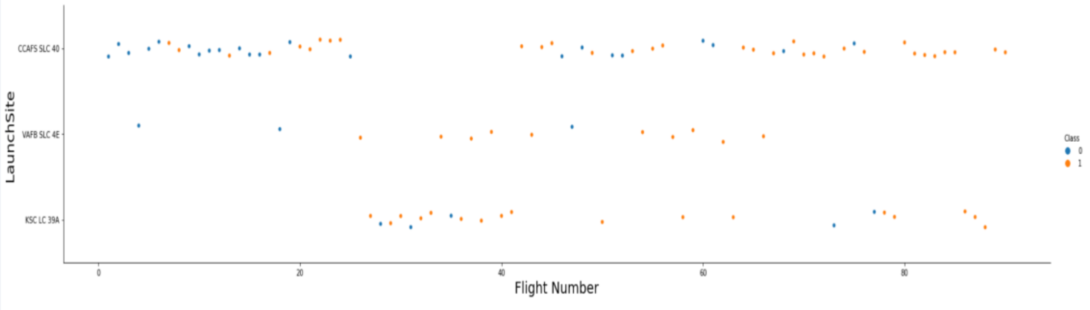
We can plot out the `FlightNumber` vs. `PayloadMass` and overlay the outcome of the launch. We see that as the flight number increases, the first stage is more likely to land successfully. The payload mass is also important; it seems the more massive the payload, the less likely the first stage will return.

```
sns.catplot(y="PayloadMass", x="FlightNumber", hue="Class", data=df, aspect=5)
plt.xlabel("Flight Number",fontsize=20)
plt.ylabel("Pay load Mass (kg)",fontsize=20)
plt.show()
```



We see that different launch sites have different success rates. `CCAFS LC-40`, has a success rate of 60 %, while `KSC LC-39A` and `VAFB SLC 4E` has a success rate of 77%.

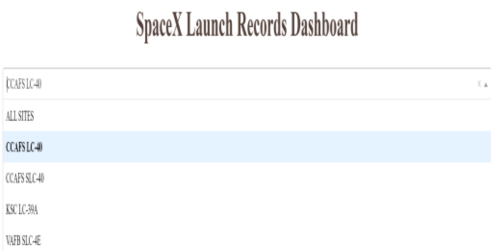
Flight Number Vs. Payload Mass



Drill Down on Flight Number Vs. Launch Site

Results Continued:

- <u>Interactive analytics demo in screenshots</u>



Dashboard Choices

Correlation between Payload and Success for all sites

<u>Predictive analysis results</u>

********************************** EVALUATING MODEL PREFORMANCE ***

	Best_Params	Accuracy_Score	Jaccard	F1-Score	Logloss
KNN	0.8482142857142858	0.8333333333333334	0.8333333333333334	0.8148148148148149	N/A
Decision Tree	0.8750000000000000	0.8333333333333334	0.8333333333333334	0.8148148148148149	N/A
SVM	0.8482142857142856	0.8333333333333334	0.8333333333333334	0.8148148148148149	N/A
Logistic Regression	0.8464285714285713	0.8333333333333334	0.8333333333333334	0.8148148148148149	0.4786666968559153

****

The best model is the model with the best params score of: 0.875
The best model is the model with the highest accuracy score of: 0.8333333333333334
The best model is the model with the jaccard score of: 0.8333333333333334

Scores Matrix for all Classification Models

Section 2

Insights drawn from EDA

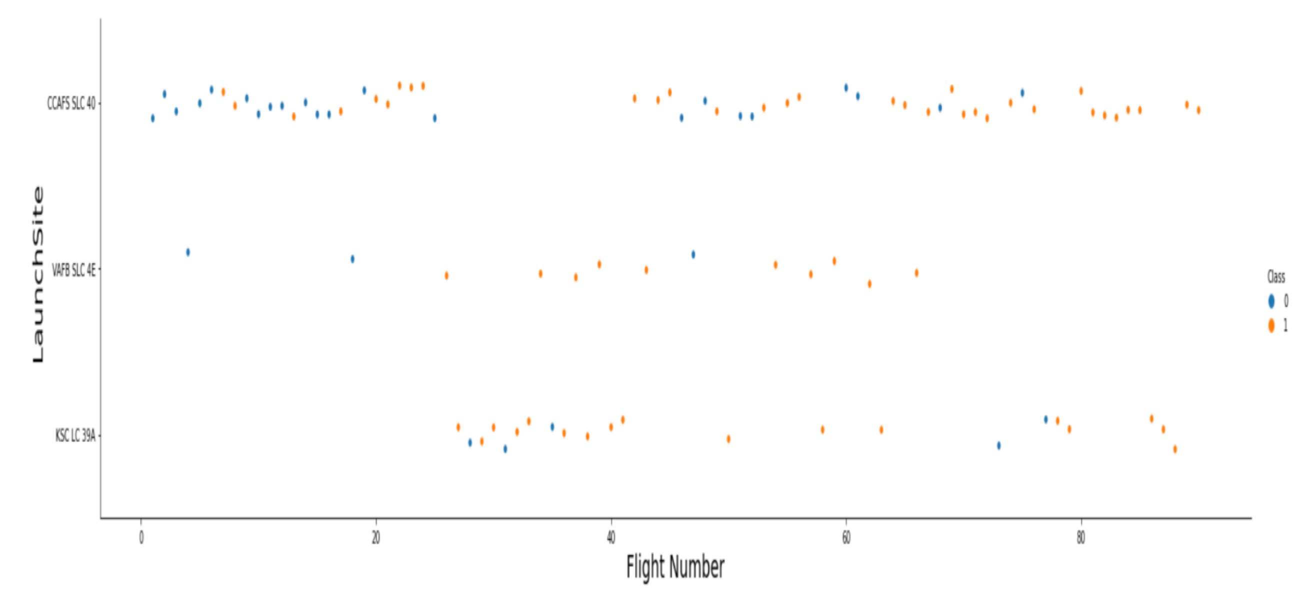
Flight Number vs. Launch Site

- The scatter plot clearly shows that the launch site CCAFS SLC-40 had the most flights, and VAFB SLC-4E had the least flights.

- Site CCAFS SLC-40 didn't have flights between @25-40, but rather KSC LC-39A did have many in that period.

- Site VAFB SLC-4E stopped having launches after @ flight 65.

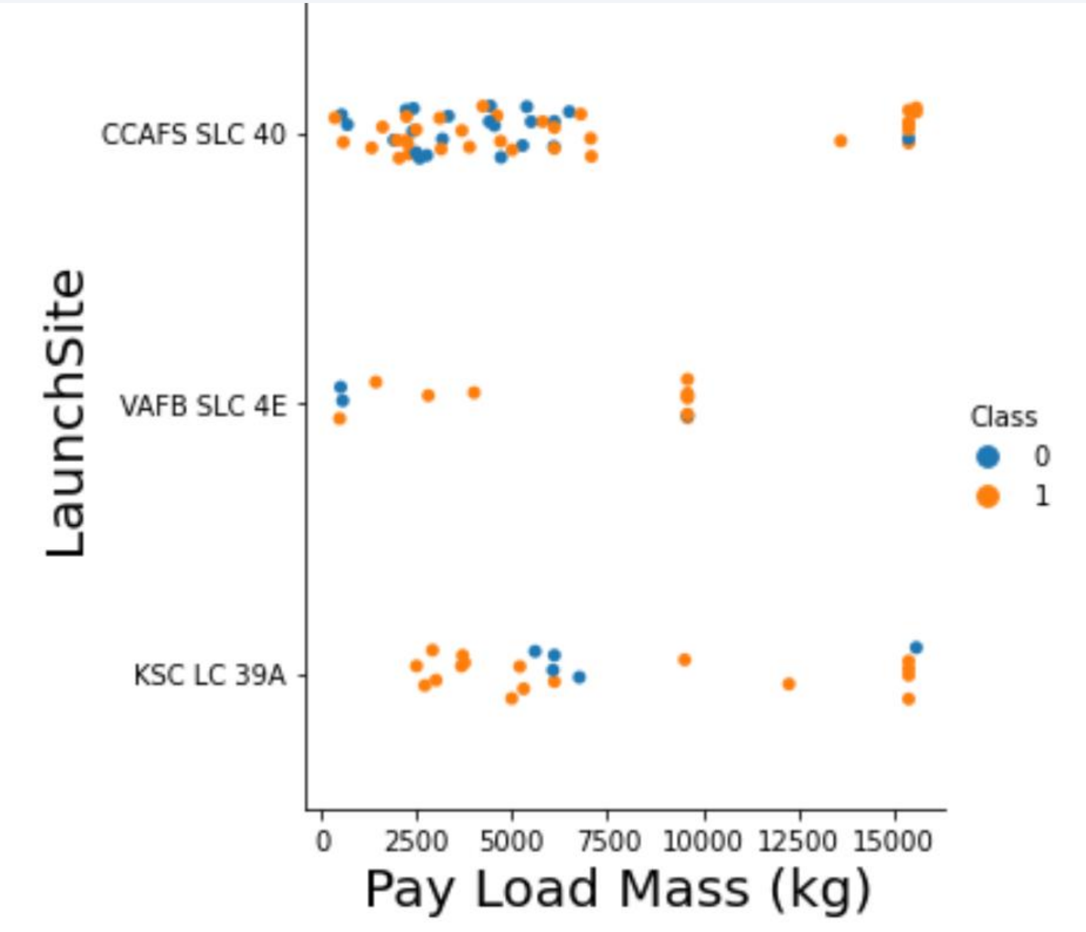
- As the flight numbers increase, the success rate also increases.



Payload vs. Launch Site

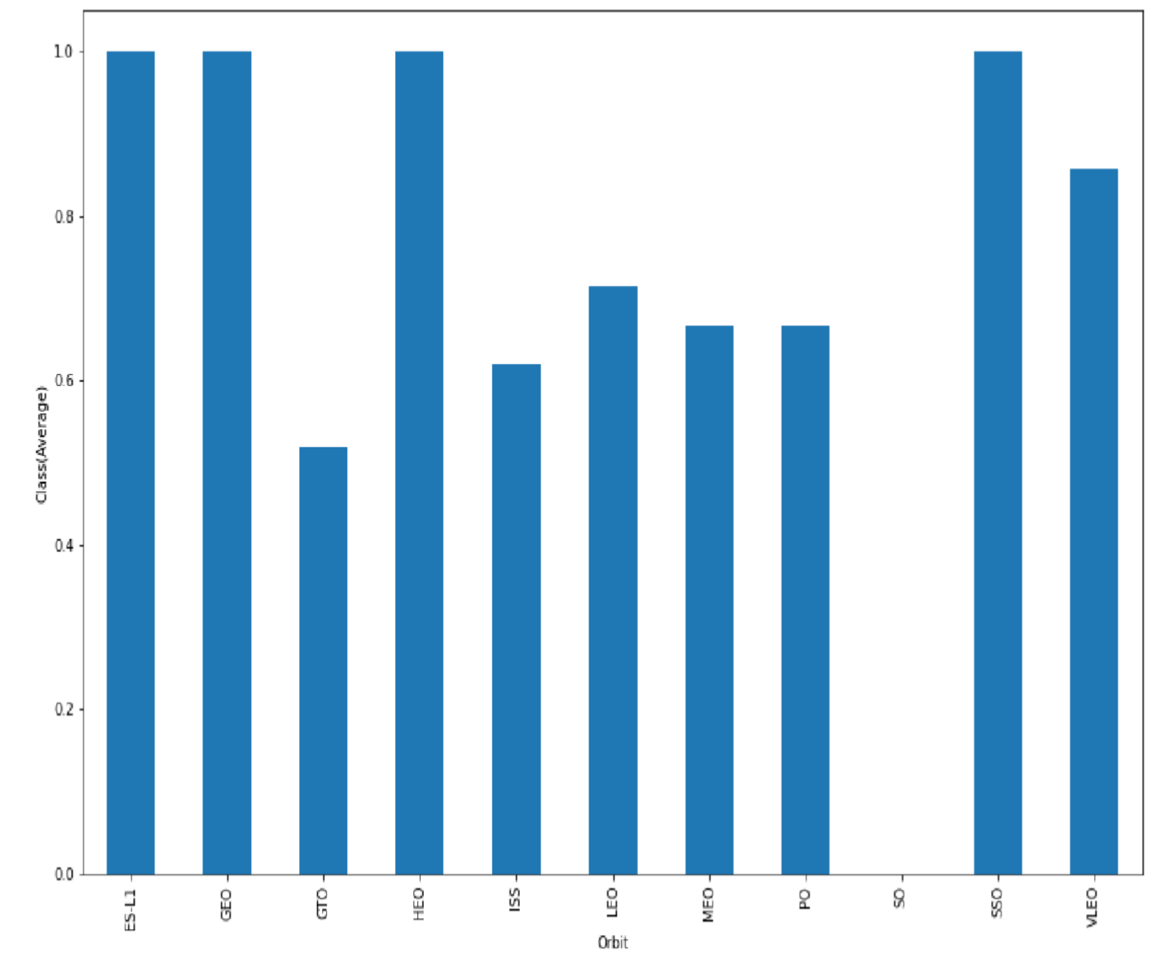
- The scatter plot clearly shows that site VAFB SLC-4E has no launches for a heavy payload mass over 10000 kg, and was very successful with launches containing payload.

- KSC LC-39A had quite a few successful launches when the payload was either light or heavy, not in-between.



Success Rate vs. Orbit Type

- The bar chart reflects the mean success rate for each of the orbits.
- The highest mean success rate orbits are ES-L1, GEO, HEO and SSO
- The lowest mean success rate orbits are SO with no successes, and GTO.



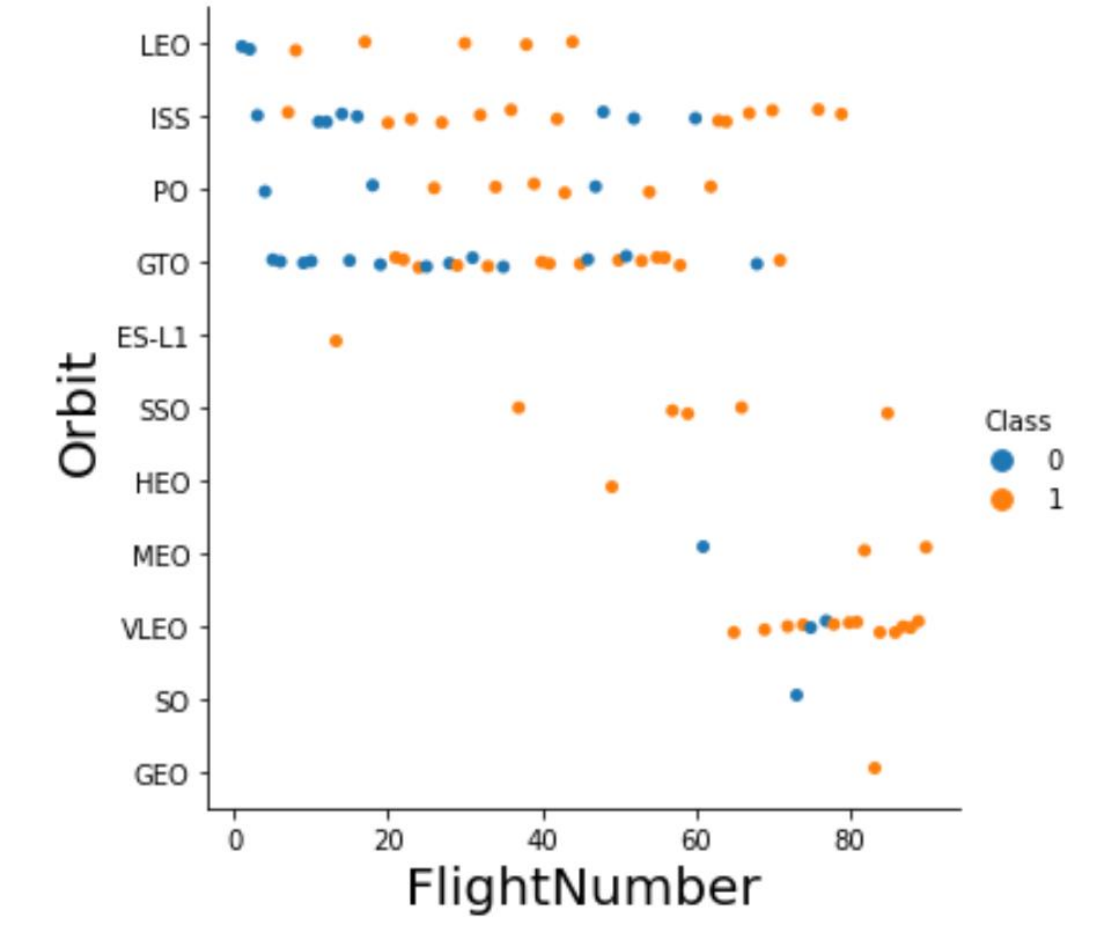
Flight Number vs. Orbit Type

- In the LEO orbit the success appears related to the number of flights; on the other hand, there seems to be no relationship between flight number when in GTO orbit.

- ISS and GTO have the most flights.

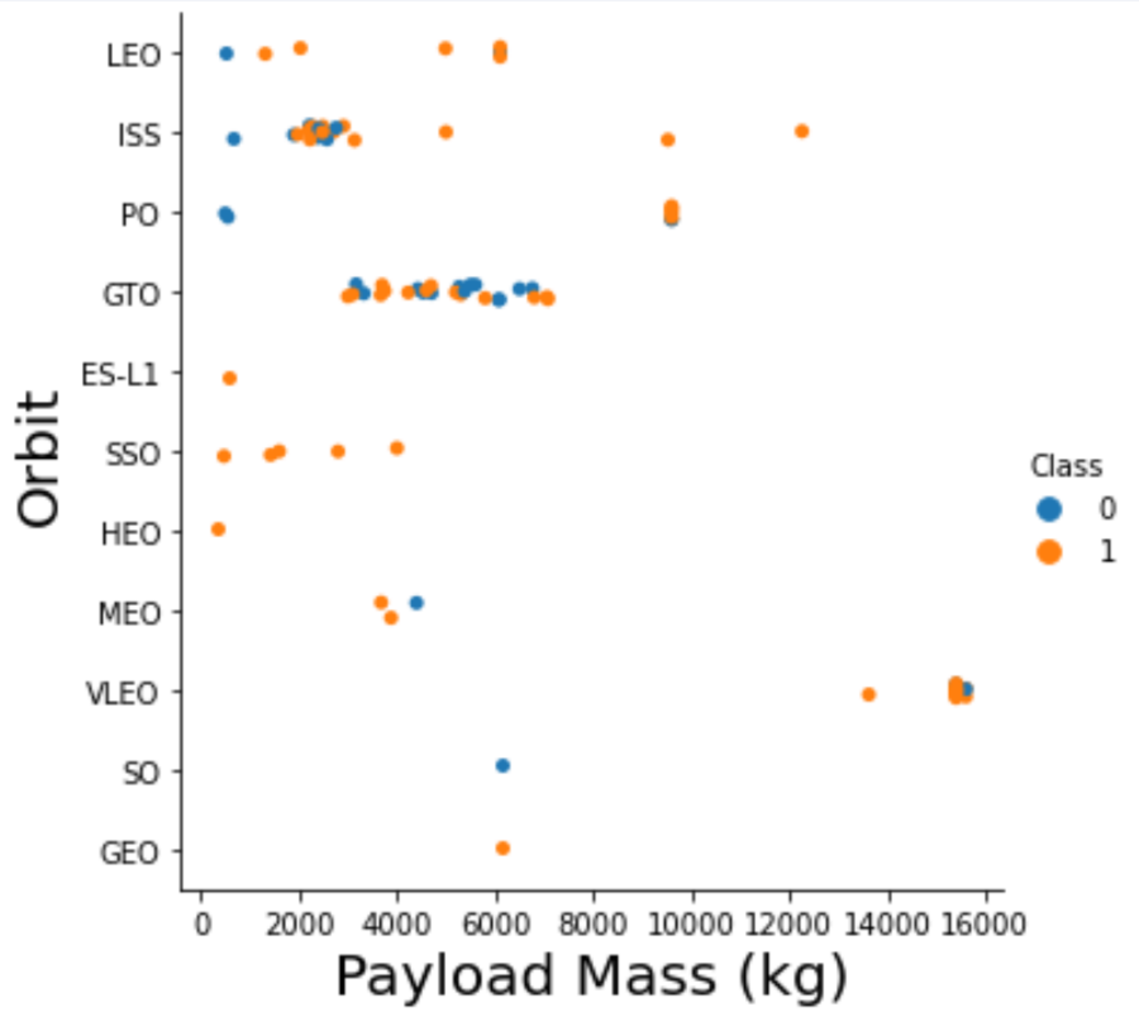
- LEO, ES-L1, SSO, HEO, MEO, SO and GEO orbits all have a limited number of flights.

- SSO orbit only had complete success over multiple flights.



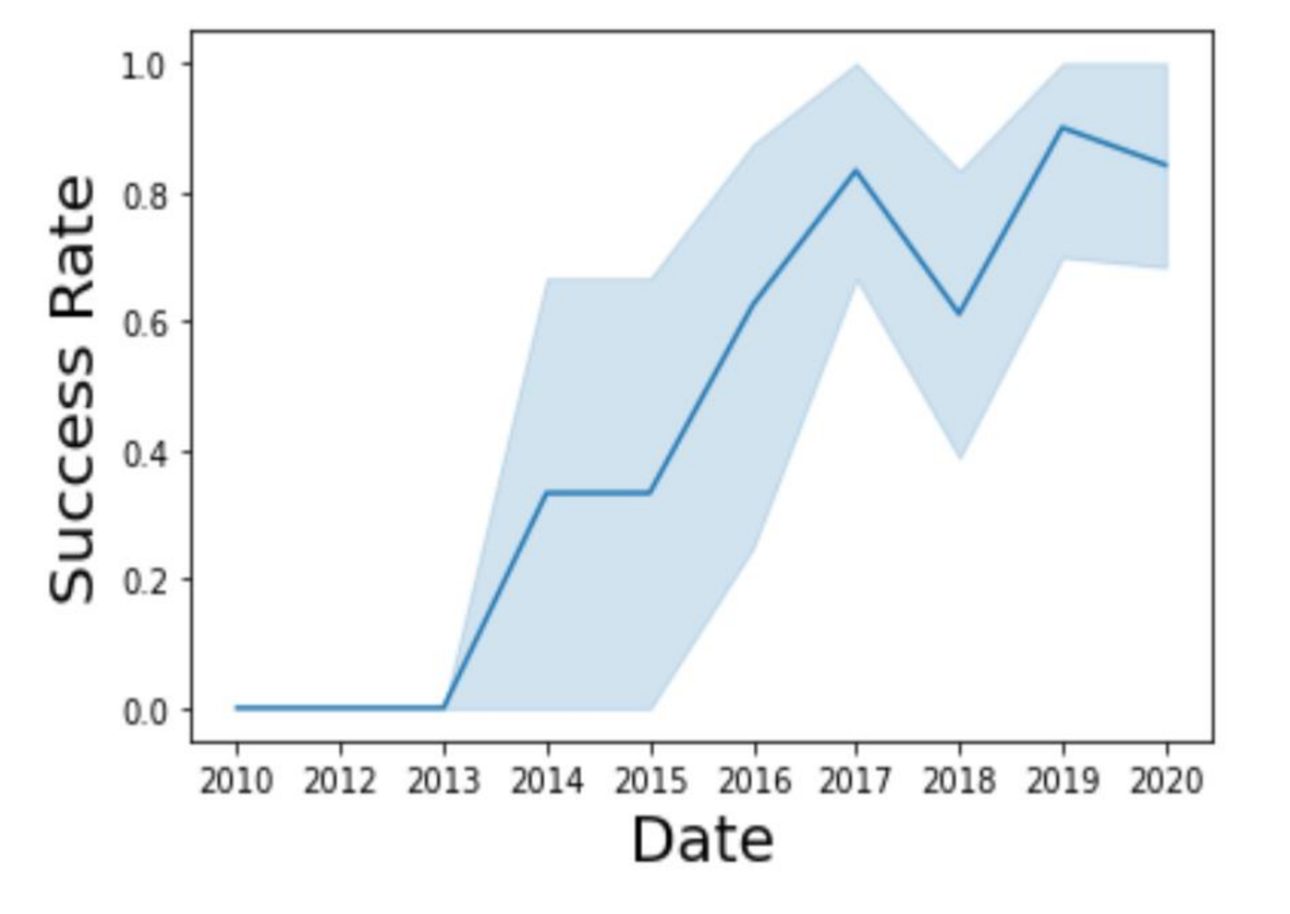
Payload vs. Orbit Type

- With heavy payloads, the successful landing or positive landing rate are more for PO,LEO and ISS.
- For GTO it's hard to distinguish results well because of the both positive land negative landing rates.
- Multiple successful launches of light payloads(<6000kg) are found with LEO, ISS, and SSO orbits.
- LEO, ISS and PO orbits have failures at low payloads, but then successes at higher payloads.
- VLEO orbit has the heaviest payload success.



Launch Success Yearly Trend

- It’s easy to observe that the average success rate since 2013 kept increasing till 2020.



All Launch Site Names

- The SQL query returns four unique launch sites.

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

Launch Site Names Begin with 'CCA'

- The SQL query returns the 5 records where launch sites begin with `CCA`.

DATE	time_utc_	booster_version	launch_site	payload	payload_mass__kg_	orbit	customer	mission_outcome	landing__outcome
2010-06-04	18:45:00	F9 v1.0 B0003	CCAFS LC-40	Dragon Spacecraft Qualification Unit	0	LEO	SpaceX	Success	Failure (parachute)
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 (parachute)
2012-05-22	07:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525	LEO (ISS)	NASA (COTS)	Success	No attempt
2012-10-08	00:35:00	F9 v1.0 B0006	CCAFS LC-40	SpaceX CRS-1	500	LEO (ISS)	NASA (CRS)	Success	No attempt
2013-03-01	15:10:00	F9 v1.0 B0007	CCAFS LC-40	SpaceX CRS-2	677	LEO (ISS)	NASA (CRS)	Success	No attempt

Total Payload Mass

- The SQL query for the total payload carried by boosters from NASA using the phrase

  “WHERE customer LIKE ‘NASA(CRS)%’ ”

  finds 2 matches:

1) (CRS), Kacific 1 = 2,617 kg
2) (CRS) = 45,596 kg

- For a total payload mass of 48,213 kg.

payload_mass
48213

Average Payload Mass by F9 v1.1

- Average payload mass carried by booster version F9 v1.1 = 2,534 kg.

payload_mass
2534

First Successful Ground Landing Date

- The date of the first successful landing outcome on ground pad is December 22, 2015.

1
2015-12-22

Successful Drone Ship Landing with Payload between 4000 and 6000

- Here is the list of the 4 names of boosters which have successfully landed on drone ship and had payload mass greater than 4000 but less than 6000.

booster_version	landing__outcome	payload_mass__kg_
F9 FT B1022	Success (drone ship)	4696
F9 FT B1026	Success (drone ship)	4600
F9 FT B1021.2	Success (drone ship)	5300
F9 FT B1031.2	Success (drone ship)	5200

WHERE mission_outcome LIKE 'S%'

Total Number of Successful and Failure Mission Outcomes

- Calculating the total number of successful and failure mission outcomes, we find that when using the SQL search criteria:

“WHERE mission_outcome LIKE ‘S%’ “ gives us a count of 100.

Using the SQL query of:

“WHERE mission_outcome LIKE ‘F%’ “ gives us a count of 1.

success_mission_counts
100

failure_mission_counts
1

Boosters Carried Maximum Payload

- There are 12 booster versions which have carried the maximum payload mass of 15,600 kg.

booster_version	payload_mass__kg_
F9 B5 B1048.4	15600
F9 B5 B1048.5	15600
F9 B5 B1049.4	15600
F9 B5 B1049.5	15600
F9 B5 B1049.7	15600
F9 B5 B1051.3	15600
F9 B5 B1051.4	15600
F9 B5 B1051.6	15600
F9 B5 B1056.4	15600
F9 B5 B1058.3	15600
F9 B5 B1060.2	15600
F9 B5 B1060.3	15600

2015 Launch Records

- There are 2 records which have failed landing_outcomes in drone ship for 2015. Both are from the same launch site.

YEAR	booster_version	launch_site	landing__outcome
2015	F9 v1.1 B1012	CCAFS LC-40	Failure (drone ship)
2015	F9 v1.1 B1015	CCAFS LC-40	Failure (drone ship)

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

- Here is a ranking in descending order on the count of landing outcomes between the date 2010-06-04 and 2017-03-20.
- The ‘No attempt’ landing outcome had the most, and Precluded (drone ship) had the least.

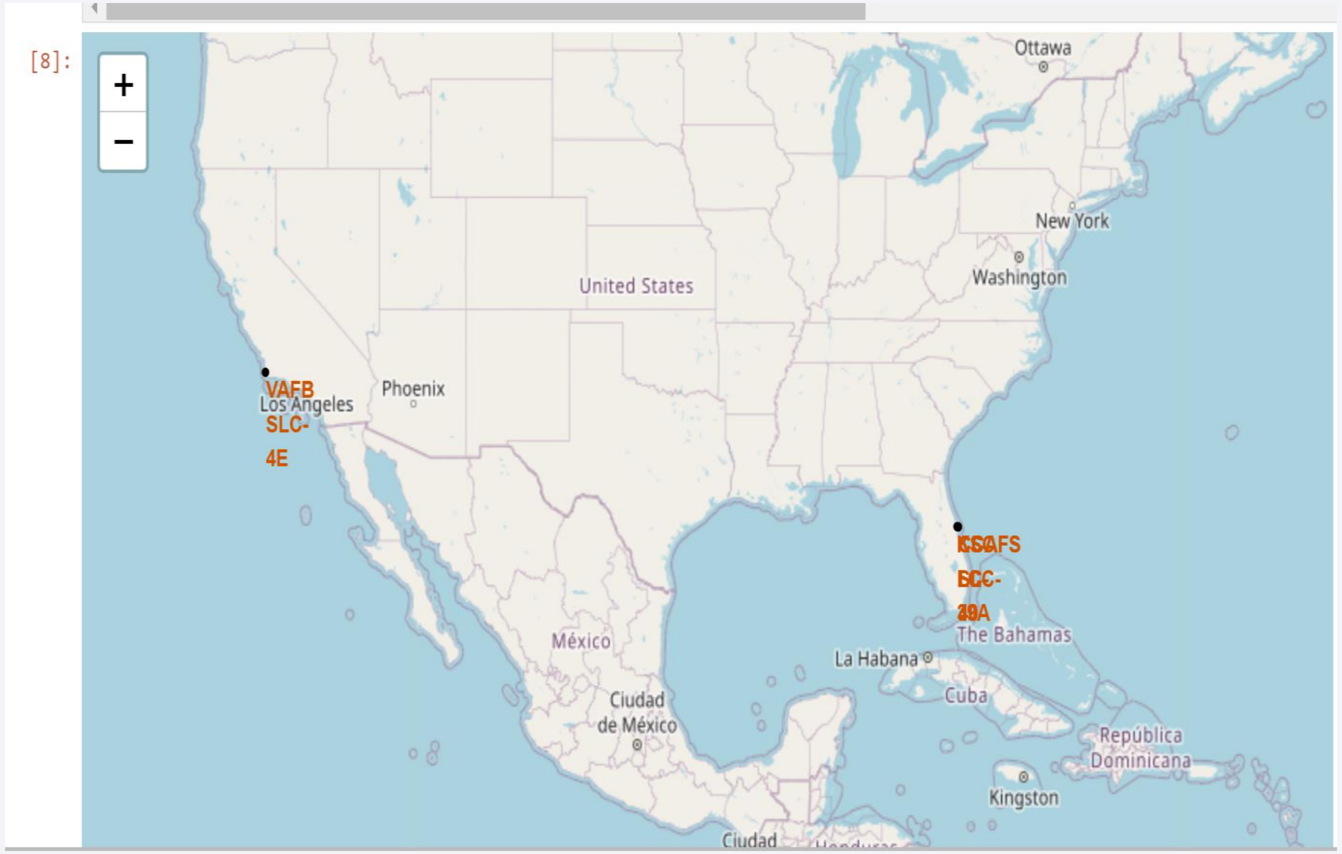
landing__outcome	COUNT
No attempt	10
Failure (drone ship)	5
Success (drone ship)	5
Controlled (ocean)	3
Success (ground pad)	3
Failure (parachute)	2
Uncontrolled (ocean)	2
Precluded (drone ship)	1

Section 3

Launch Sites Proximities Analysis

Falcon 9 Launch Site Locations

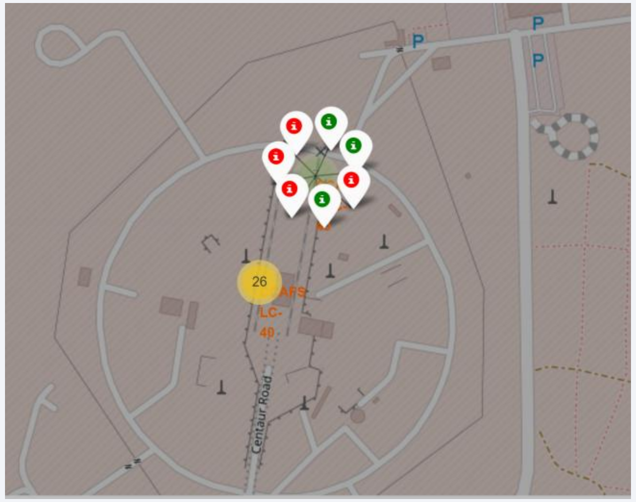
- There are multiple launch sites, mainly located in either Florida or California.
- It is easy to see that all the launch sites are located near the ocean, and are in the southernmost part of the US.

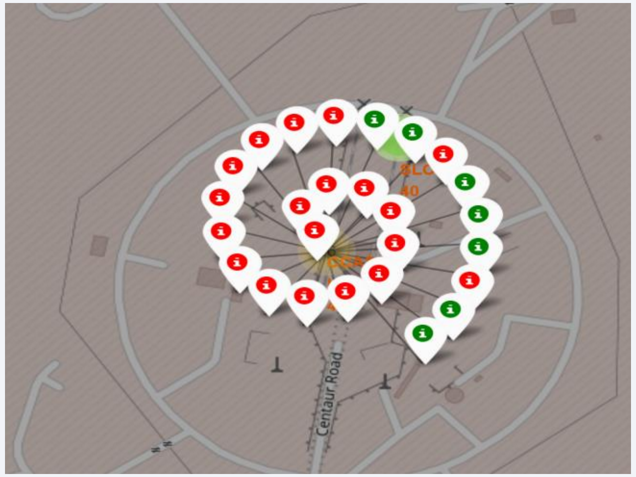


Launch Outcomes at Sites CCAFS LC-40 & CCAFS SLC-40

- Marker clusters are an easy way to denote launch sites successes (in green) and failures (in red).

- Since the two sites CCAFS LC-40 & CCAFS SLC-40 have similar coordinates, it’s easy to see that CCAFS LC-40 had more launches and a number of failures than site CCAFS SLC-40.



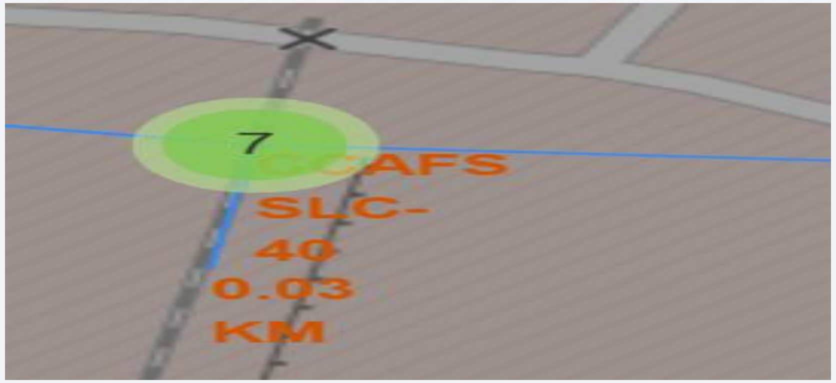


Surrounding Landmarks for site CCAFS SLC-40

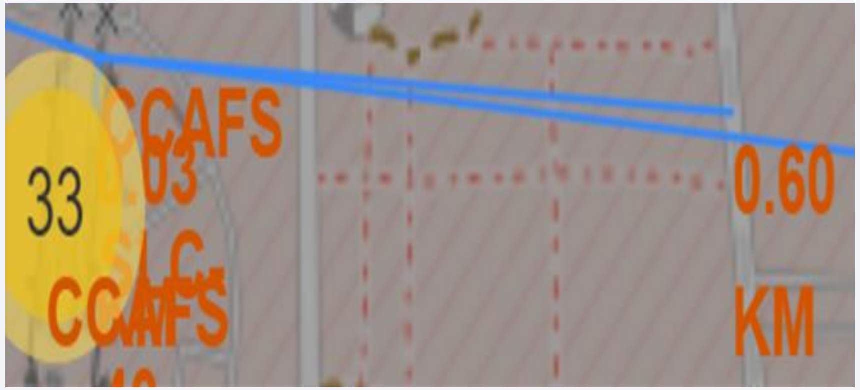
- Site CCAFS SLC-40 is located near the Florida Coast.
- The coastline is very close (.88KM) allowing for launches/landings to be safe and nearby.
- The fail and highway systems are close too (.03& .60KM) making travel to/from the site easy.
- The farthest is a nearby city (Titusville) at 23.30KM away. This allows the public to be at a safe distance from the launch site.

Nearest Coastline for site CCAFS SLC-40 is .88 KM

Closest city to site CCAFS SLC-40 is Titusville, at 23.30 KM



Closest rail to site CCAFS SLC-40 is NASA Rail, at .03 KM



Closest highway for site CCAFS SLC-40 is Samuel C. Phillips, at .60 KM

Section 4

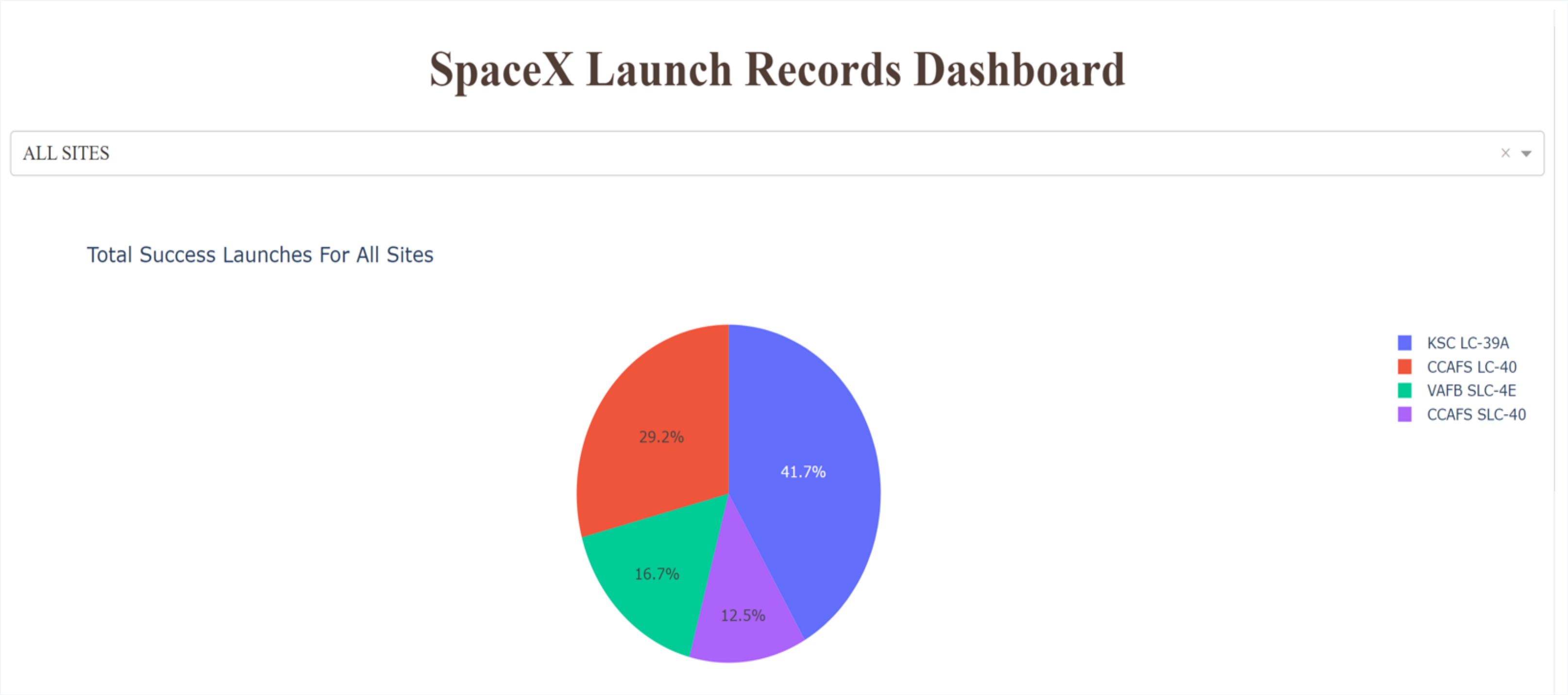
Build a Dashboard with Plotly Dash

Total Success Launches for All Sites

- Site KSC LC-39A had the most successes (41.7%)

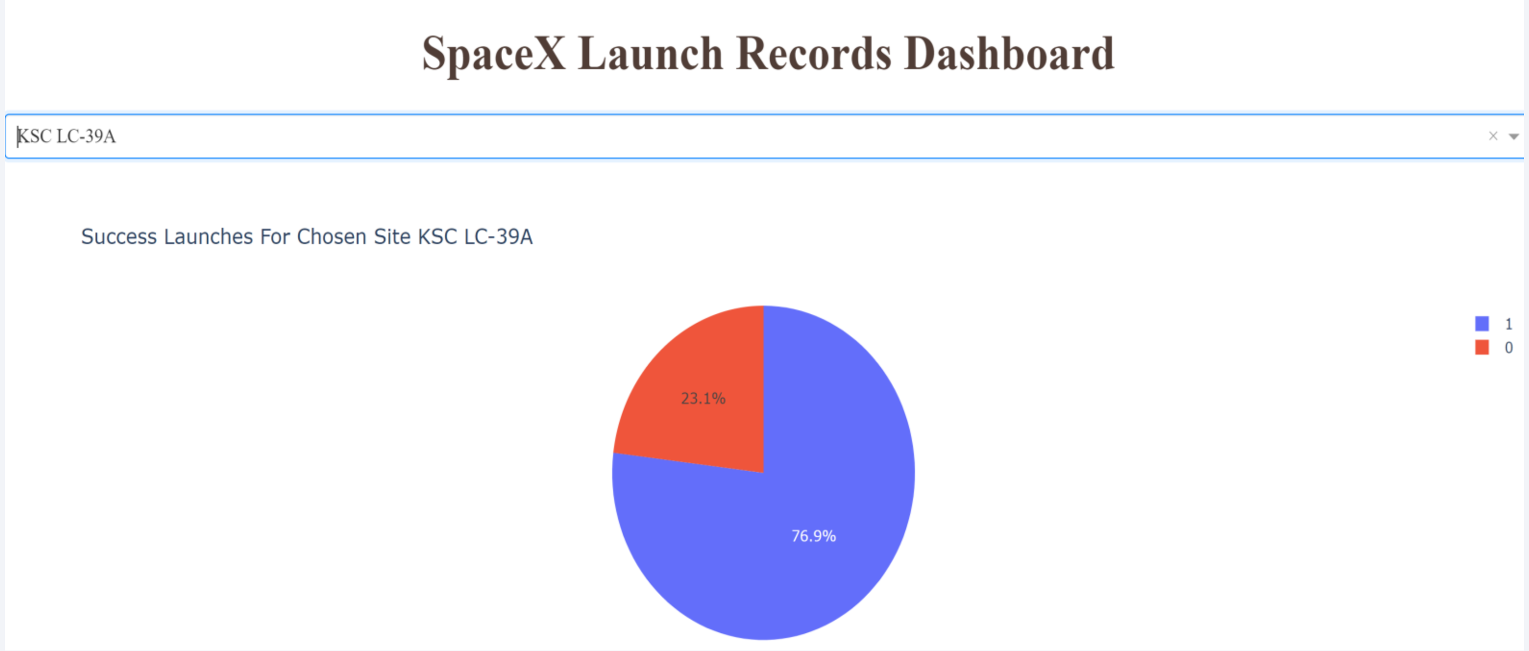
- Site CCAFS LC-40 had the second-most successes (29.2%)

- Site CCAFS SLC-40 had the least successes (12.5%)



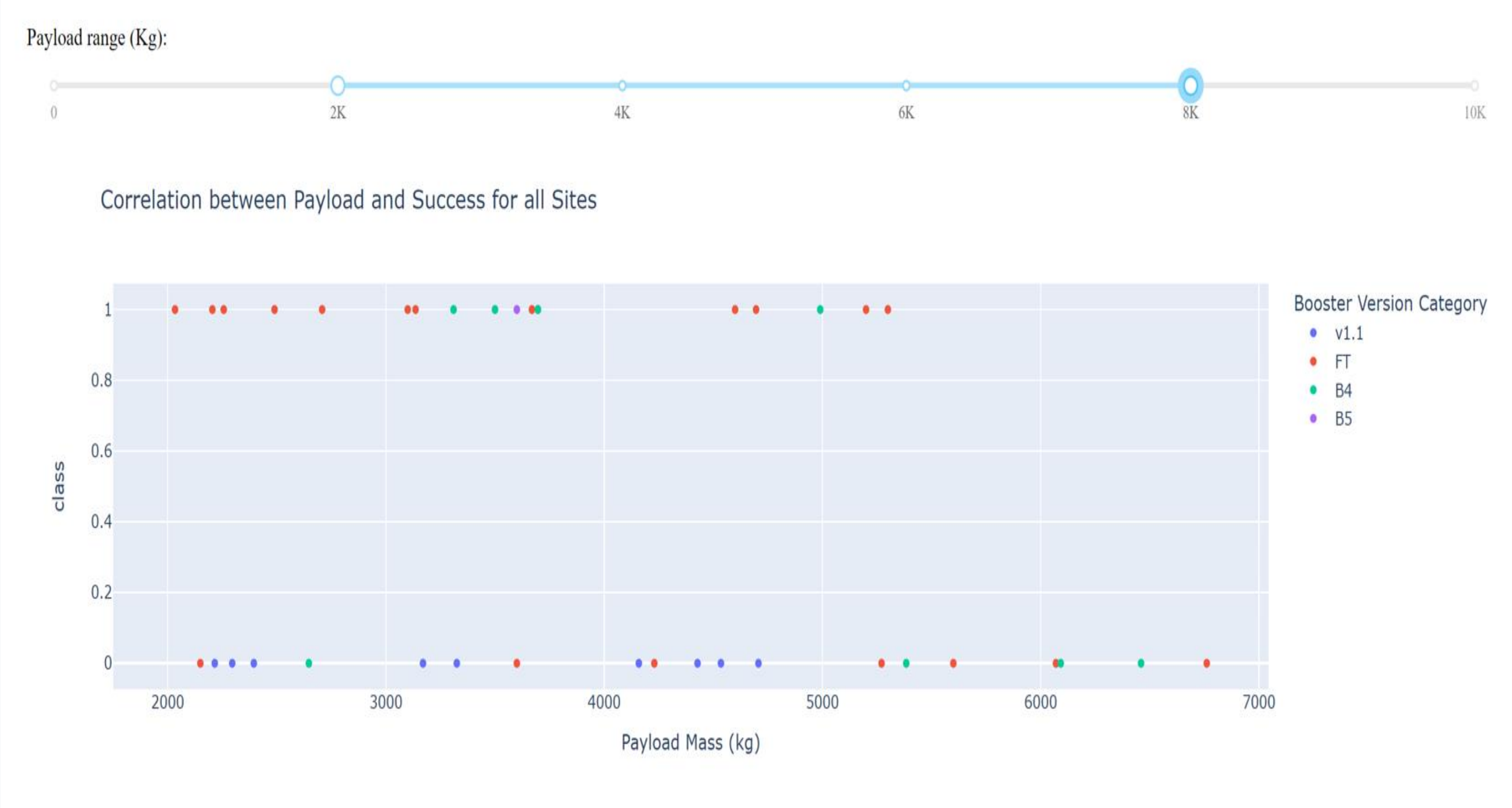
KSC LC-39A –Highest Launch Success Ratio

- Site KSC LC-39A had the highest success ratio of all sites at 76.9% success and only 23.1% failure.



Payload Mass (kg) Vs. Launch Outcome for All Sites (Range)

- The Payload Mass (kg) slider was set to show payloads between 2K and 8K.
- In this range, we can see that there were more failures (class = 0) than successes (class=1).
- In this range we can also see that booster version FT had the most successes, while v1.1 had the most failures.

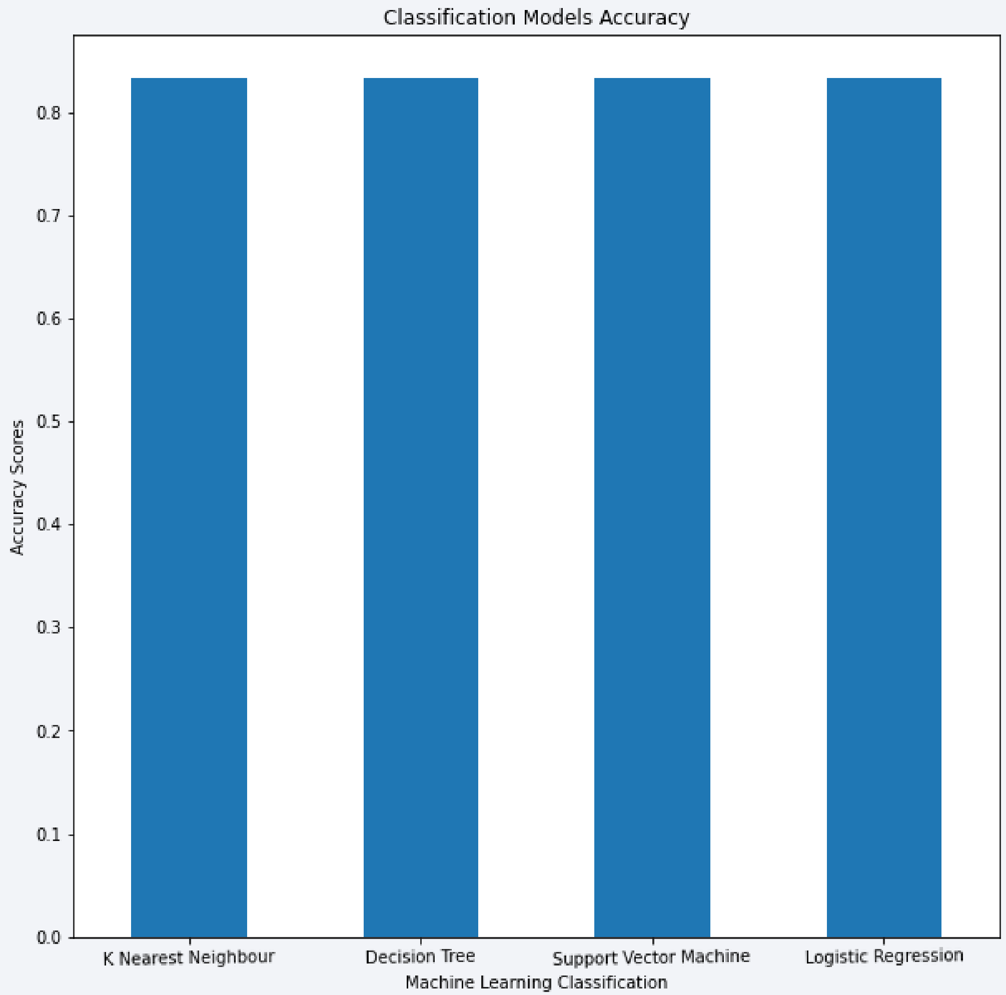


Section 5

Predictive Analysis (Classification)

Classification Accuracy

- All 4 of the machine learning classification models had the same accuracy score of 83.34%., so we can use any of them to use for predicting first stage launch success.
- Since the best_params score is the highest for decision tree (87.%% vs. @84% for each of the other three models), I will highlight detail using the Decision Tree model.



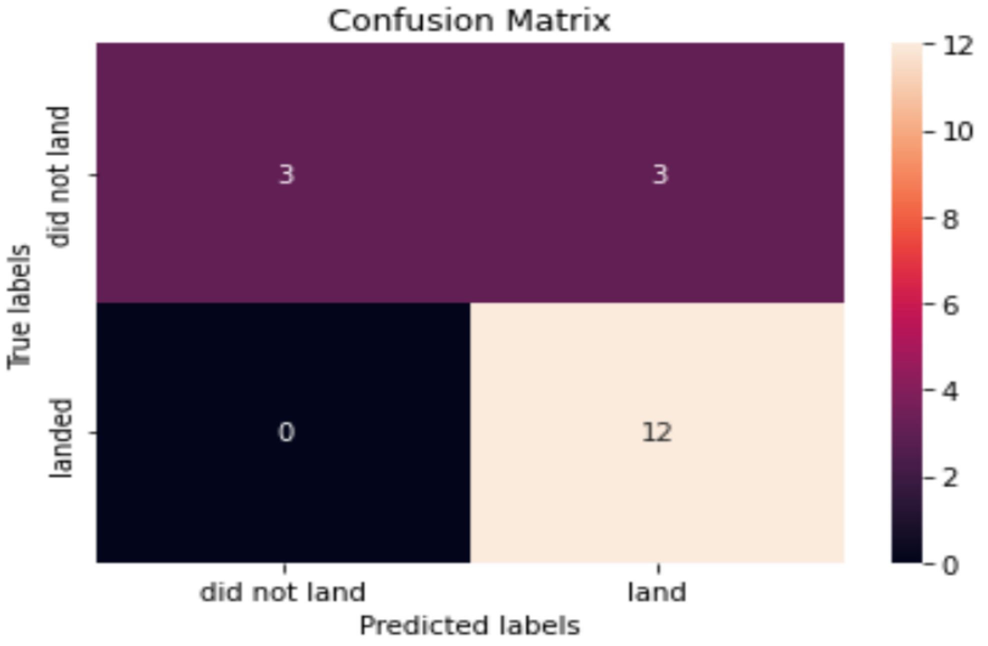
	Best_Params
KNN	0.8482142857142858
Decision Tree	0.8750000000000000
SVM	0.8482142857142856
Logistic Regression	0.8464285714285713

Confusion Matrix

- The Confusion Matrix for Decision Tree Classification shows correctly that all 12 launches that landed successfully were correctly reflected as such.
- However, it also shows that out of the 6 unsuccessful launches, 3 were marked wrongly as false positive.

	precision	recall	f1-score	support
0	1.00	0.50	0.67	6
1	0.80	1.00	0.89	12
accuracy			0.83	18
macro avg	0.90	0.75	0.78	18
weighted avg	0.87	0.83	0.81	18

```
Test set Tree Jaccard Accuracy:  0.8333333333333334
Test set Tree F1 Accuracy: 0.8148148148148149
```



Conclusions

- Folium maps show that launch sites are concentrated near the coasts of Florida and California. Nearby available necessary resources to sites include highway, rail, and coastline, while keeping general population of surrounding cities distant.
- Visualization through charts and plots enabled us to easily see many details, including the highest successful launch site is KSC LC-39A in Florida., and the success rate for launches has steadily increased from 2015-2020, site CCAFS SLC-40 had the most launches, and that launch success increased with flight number., and that payload mas s seems to has an effect on launch success, and booster version FT has the most successes for payload masses between 2-8k.
- Machine Learning using 4 different classification models all had an accuracy of launch success prediction of 83.4%. This is a fairly high degree of accuracy using the given historical data.
- All of these findings in combination help to identify the specific areas to concentrate on if/when SpaceY decides to compete with SpaceX.

Appendix

	precision	recall	f1-score	support
0	1.00	0.50	0.67	6
1	0.80	1.00	0.89	12
accuracy			0.83	18
macro avg	0.90	0.75	0.78	18
weighted avg	0.87	0.83	0.81	18

```
Test set LR Jaccard Accuracy:  0.8333333333333334
Test set LR F1 Accuracy: 0.8148148148148149
Test set LR Log_Loss: 0.4786666968559153
```

Linear Regression

	precision	recall	f1-score	support
0	1.00	0.50	0.67	6
1	0.80	1.00	0.89	12
accuracy			0.83	18
macro avg	0.90	0.75	0.78	18
weighted avg	0.87	0.83	0.81	18

```
Test set SVM Jaccard Accuracy:  0.8333333333333334
Test set SVM F1 Accuracy: 0.8148148148148149
```

Support Vector Machine

	precision	recall	f1-score	support
0	1.00	0.50	0.67	6
1	0.80	1.00	0.89	12
accuracy			0.83	18
macro avg	0.90	0.75	0.78	18
weighted avg	0.87	0.83	0.81	18

```
Test set Tree Jaccard Accuracy:  0.8333333333333334
Test set Tree F1 Accuracy: 0.8148148148148149
```

Decision Tree

	precision	recall	f1-score	support
0	1.00	0.50	0.67	6
1	0.80	1.00	0.89	12
accuracy			0.83	18
macro avg	0.90	0.75	0.78	18
weighted avg	0.87	0.83	0.81	18

```
Test set KNN Jaccard Accuracy:  0.8333333333333334
Test set KNN F1 Accuracy: 0.8148148148148149
```

K Nearest Neighbour

I prepared classification reports for each of the 4 models, and was able to collect a matrix of scores to evaluate model performance.

As it turns out, all 4 models had similar scores.

********************************** EVALUATING MODEL PREFORMANCE **

	Best_Params	Accuracy_Score	Jaccard	F1-Score	Logloss
KNN	0.8482142857142858	0.8333333333333334	0.8333333333333334	0.8148148148148149	N/A
Decision Tree	0.8750000000000000	0.8333333333333334	0.8333333333333334	0.8148148148148149	N/A
SVM	0.8482142857142856	0.8333333333333334	0.8333333333333334	0.8148148148148149	N/A
Logistic Regression	0.8464285714285713	0.8333333333333334	0.8333333333333334	0.8148148148148149	0.4786666968559153

****

```
The best model is the model with the best params score of:  0.875
The best model is the model with the highest accuracy score of:  0.8333333333333334
The best model is the model with the jaccard score of:  0.8333333333333334
```

Thank you!