IBM Developer
SKILLS NETWORK

Winning the Space Race with Data Science

Dan Barstow B.S. Mathematics: Operations Research
12/15/21

Outline

- Executive Summary
- Introduction
- Methodology
- Results
- Conclusion
- Appendix

Executive Summary – “What’s the Upshot?”

- Summary of methodologies
- Summary of all results

Introduction

According to SpaceX's website, the total cost to launch a Falcon 9 rocket is $62 million, which is roughly half of what other space agencies can offer.

SpaceX can launch at a cheaper cost because it can **reuse** the first stage.

<u>What pertinent factors affect whether the first stage returns?</u>

<u>Given that it's so expensive, can we predict the success of a future launch?</u>

The first stage of a Falcon 9 rocket being recovered and lifted via crane.

<image source> = http://kenkremer.blogspot.com/2018/11/spacex-tilts-recovered-falcon-9-1st.html

Section 1

Methodology

Methodology

- Data collection methodology
  - Launch data was collected using the SpaceX REST API and by web scraping a Wikipedia page using Beautiful Soup.
- Perform data wrangling
  - Data was processed using the API and Pandas, then stored on an SQL database for preliminary analysis.
- Perform exploratory data analysis (EDA) using visualization and SQL
  - Explanatory data analysis was performed on the data using SQL queries and Seaborn to plot graphs and charts.
- Perform interactive visual analytics using Folium and Plotly Dash
  - An interactive map marking the launch sites and their success rates was created using Folium and Plotly Dash.
- Perform predictive analysis using classification models
  - 4 Different Machine Learning models were built, trained, and tested suggesting an 83.3% predicted success rate.

Data Collection

- Data was collected using the SpaceX REST API.
- Additional data was acquired through web scraping launch data from Wikipedia using the Beautiful Soup Python library.
- The data was then loaded into an SQL database and a Pandas data frame to simplify the analysis process.


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

response =requests.get(url)

response.json()
```

Data Collection – SpaceX API

```python
spacex_url="https://api.spacexdata.com/v4/launches/past"
```

```python
response = requests.get(spacex_url)
```

- Using the API, we obtain a nice table of data including flight number, booster version, type of orbit, etc.
- Extraneous data not related to the Falcon 9 is not included in the working dataset.

	FlightNumber	Date	BoosterVersion	PayloadMass	Orbit	LaunchSite	Outcome	Flights	GridFins	Reused	Legs	LandingPad
4	1	2010-06-04	Falcon 9	NaN	LEO	CCSFS SLC 40	None None	1	False	False	False	None
5	2	2012-05-22	Falcon 9	525.0	LEO	CCSFS SLC 40	None None	1	False	False	False	None
6	3	2013-03-01	Falcon 9	677.0	ISS	CCSFS SLC 40	None None	1	False	False	False	None
7	4	2013-09-29	Falcon 9	500.0	PO	VAFB SLC 4E	False Ocean	1	False	False	False	None
8	5	2013-12-03	Falcon 9	3170.0	GTO	CCSFS SLC 40	None None	1	False	False	False	None
...	...	...	...	...	...	...	...	...	...	...	...	...
89	86	2020-09-03	Falcon 9	15600.0	VLEO	KSC LC 39A	True ASDS	2	True	True	True	5e9e3032383ecb6bb234e7ca
90	87	2020-10-06	Falcon 9	15600.0	VLEO	KSC LC 39A	True ASDS	3	True	True	True	5e9e3032383ecb6bb234e7ca
91	88	2020-10-18	Falcon 9	15600.0	VLEO	KSC LC 39A	True ASDS	6	True	True	True	5e9e3032383ecb6bb234e7ca
92	89	2020-10-24	Falcon 9	15600.0	VLEO	CCSFS SLC 40	True ASDS	3	True	True	True	5e9e3033383ecbb9e534e7cc
93	90	2020-11-05	Falcon 9	3681.0	MEO	CCSFS SLC 40	True ASDS	1	True	False	True	5e9e3032383ecb6bb234e7ca

Github URL:
https://github.com/barstow2/IBM-Applied-Data-Science-Capstone/blob/master/Data%20Collection%20API%20Lab.ipynb

Data Collection - Scraping

- Data was also Scraped from this Wikipedia page using Beautiful Soup:

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

- This text file was then processed to extract the rows and columns present on the website.

```
<tr>
<th scope="col">Flight No.
</th>
<th scope="col">Date and<br/>time (<a href="/wiki/Coordinated_Universal_
Time" title="Coordinated Universal Time">UTC</a>)
</th>
<th scope="col"><a href="/wiki/List_of_Falcon_9_first-stage_boosters" ti
tle="List of Falcon 9 first-stage boosters">Version,<br/>Booster</a> <su
p class="reference" id="cite_ref-booster_11-0"><a href="#cite_note-boost
er-11">[b]</a></sup>
</th>
<th scope="col">Launch site
</th>
<th scope="col">Payload<sup class="reference" id="cite_ref-Dragon_12-0">
<a href="#cite_note-Dragon-12">[c]</a></sup>
</th>
<th scope="col">Payload mass
</th>
<th scope="col">Orbit
</th>
<th scope="col">Customer
</th>
<th scope="col">Launch<br/>outcome
</th>
<th scope="col"><a href="/wiki/Falcon_9_first-stage_landing_tests" title
="Falcon 9 first-stage landing tests">Booster<br/>landing</a>
</th></tr>
```

Github URL: https://github.com/barstow2/IBM-Applied-Data-Science-Capstone/blob/master/Data%20Collection%20with%20Web%20Scraping%20Lab.ipynb

Data Wrangling – Preventing Garbage in, Garbage Out

- Issues resolved:
  - Disambiguated whether a launch was a success or not.
  - Identified what data was missing and converted NaN or Null values.
  - Used One-Hot Encoding to turn categorical variables into numerical ones.

We can use the following line of code to determine the success rate:

```
df["Class"].mean()
```

```
0.6666666666666666
```

```
landing_outcomes
```

```
True ASDS      41
None None      19
True RTLS      14
False ASDS      6
True Ocean      5
None ASDS       2
False Ocean     2
False RTLS      1
Name: Outcome, dtype: int64
```

Github URL: https://github.com/barstow2/IBM-Applied-Data-Science-Capstone/blob/master/Data%20Wrangling%20Lab.ipynb

EDA with Data Visualization

- Graphed the relationship between the following variables to determine any correlation:
    - Flight Number and Launch Site.
    - Payload and Launch Site.
    - Success rate of each different orbit.
    - Flight Number and Orbit Type.
    - Payload and Orbit Type.
- Graphed the success rate over time.

Github URL: https://github.com/barstow2/IBM-Applied-Data-Science-Capstone/blob/master/Explanatory%20Data%20Analysis%20with%20Visualization.ipynb

EDA with SQL

- SQL Queries performed:
    - Calculated total and average payload mass.
    - Found the first successful launch date.
    - Got the total number of mission successes and failures.
    - Found the booster version which carries the minimum payload mass.
    - Ranked the counts of frequency of different unsuccessful landing outcomes

Github URL: https://github.com/barstow2/IBM-Applied-Data-Science-Capstone/blob/master/Exploratory%20Data%20Analysis%20with%20SQL.ipynb

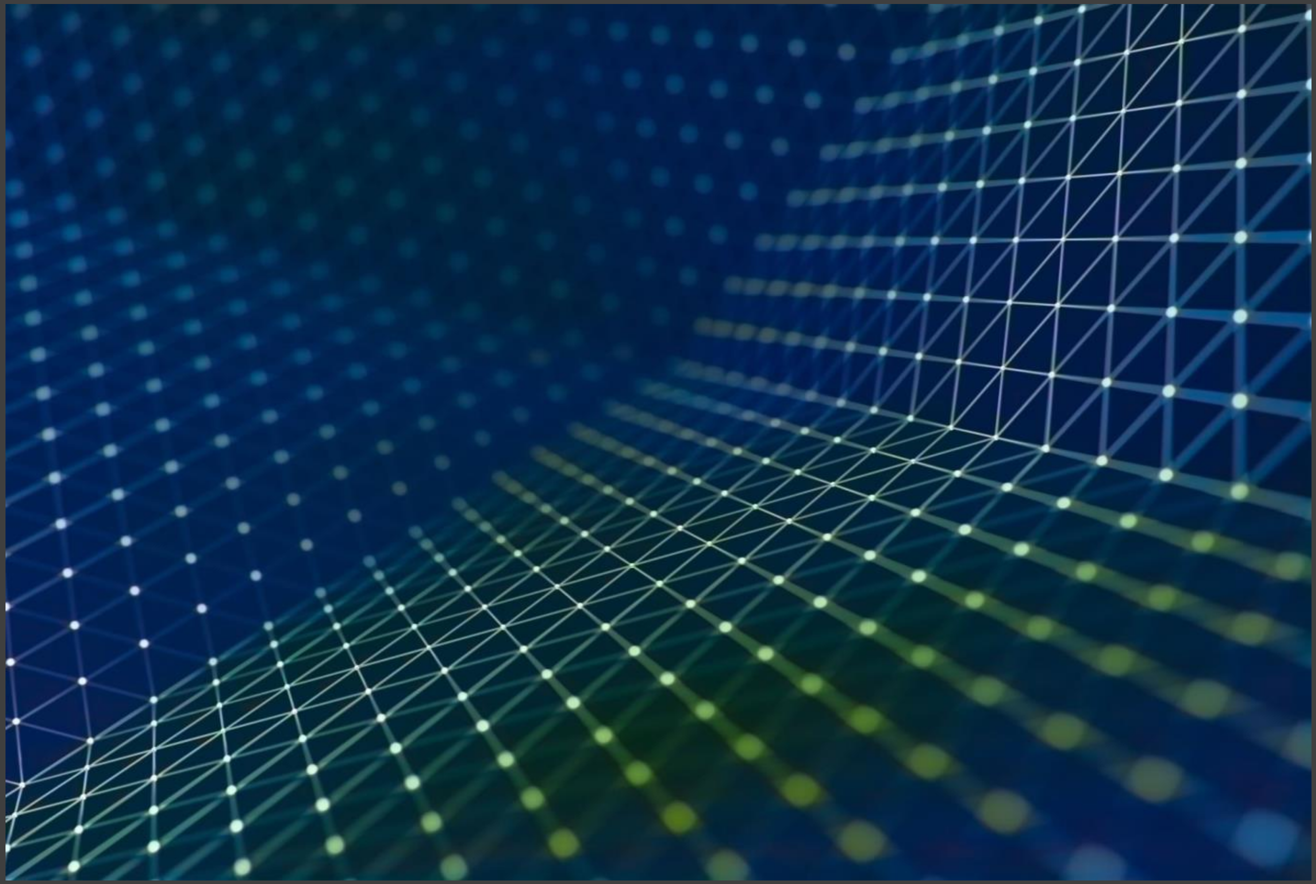
Build an Interactive Map with Folium

- Marked all launch sites on interactive Folium map with labels.
- Mark the successful/unsuccessful launches for each site on the map.
- Calculated the distances between launch sites and coastlines, railways, and cities to determine if they play a role.

Github URL:

https://github.com/barstow2/IBM-Applied-Data-Science-Capstone/blob/master/Interactive%20Visual%20Analytics%20with%20Folium.ipynb

Build a Dashboard with Plotly Dash



The web app allows users to:

- View the launch success rates by site in the form of an interactive pie chart.
- Vary the payload range on the Falcon 9 to see how the success rate changes.
- View the correlation between payload and launch site with an interactive scatter plot.

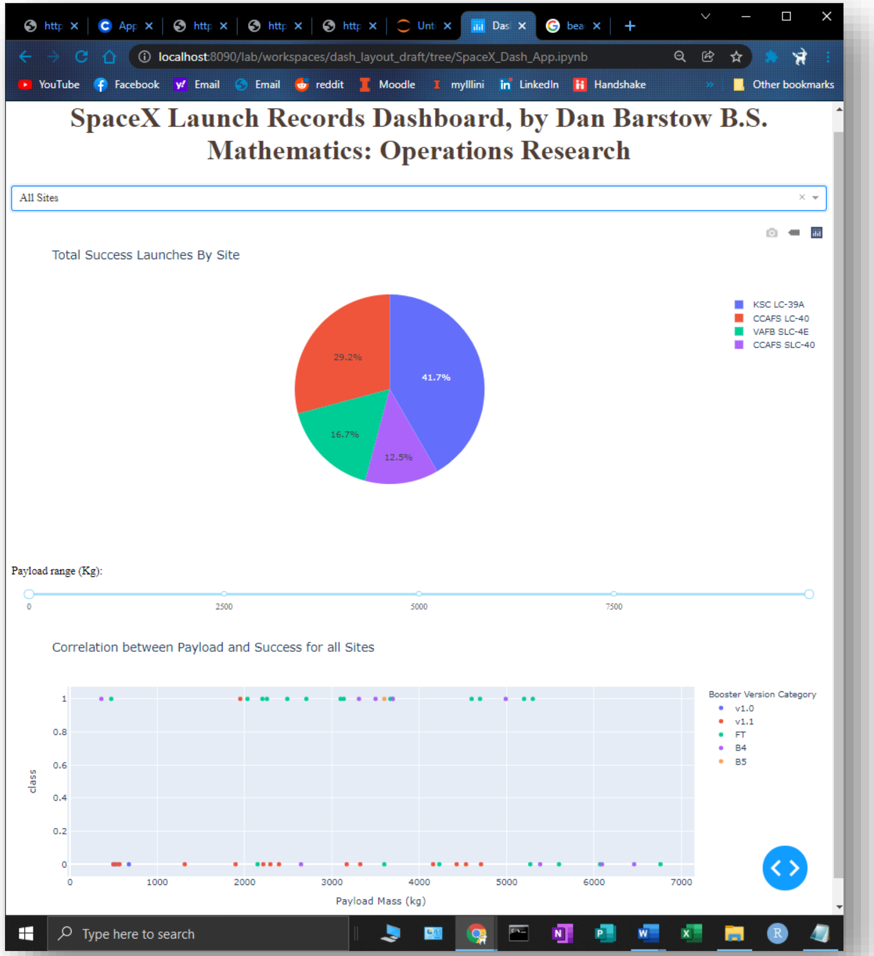
Github URL:

https://github.com/barstow2/IBM-Applied-Data-Science-Capstone/blob/master/Interactive%20SpaceX%20Dashboard.ipynb

Predictive Analysis (Classification)

- Built and tested 4 different machine learning models to predict the success rate of the first stage of the Falcon 9.
  - Logistic Regression (prediction = 83.3% success)
  - Support Vector Machine (prediction = 83.3% success)
  - Decision Tree (prediction = 83.3% success)
  - K-Nearest Neighbors (prediction = 83.3% success)
- We see that all 4 models agree with a success rate of 83.3% given our data.
- Github URL: https://github.com/barstow2/IBM-Applied-Data-Science-Capstone/blob/master/Interactive%20Visual%20Analytics%20with%20Folium.ipynb

Results



- We can see that all 4 models agree on the predicted success rate of 83.3%.
- This is likely due to the small size of the dataset (18 samples).
- Since we have been thorough in our Data Wrangling stage, we can be confident in our prediction.

Find the method performs best:

```
scores = [lr_score,svm_score,tree_score,knn_score]
print(scores)
print(scores.index(max(scores)))
```

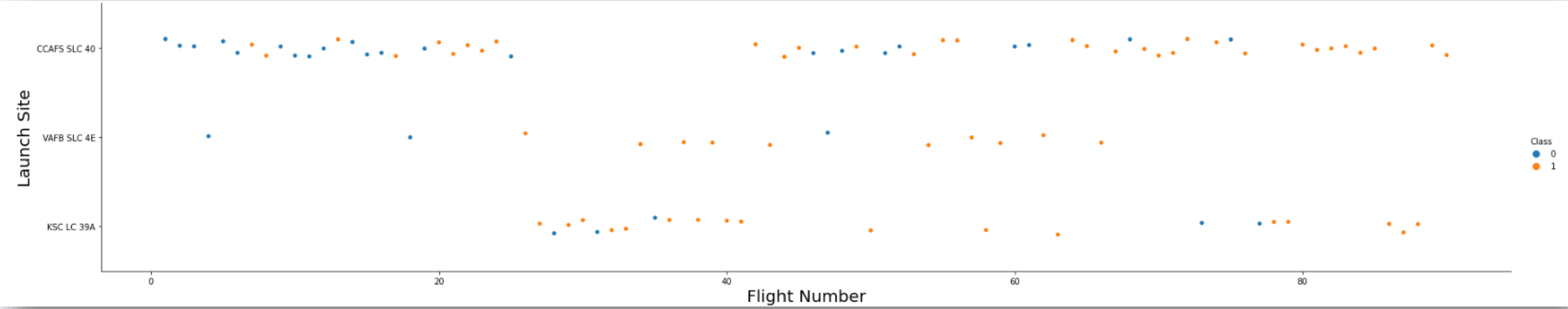
```
[0.8333333333333334, 0.8333333333333334, 0.8333333333333334, 0.8333333333333334]
0
```

Section 2

Insights drawn from EDA

Flight Number vs. Launch Site

Scatter plot of Flight Number vs. Launch Site

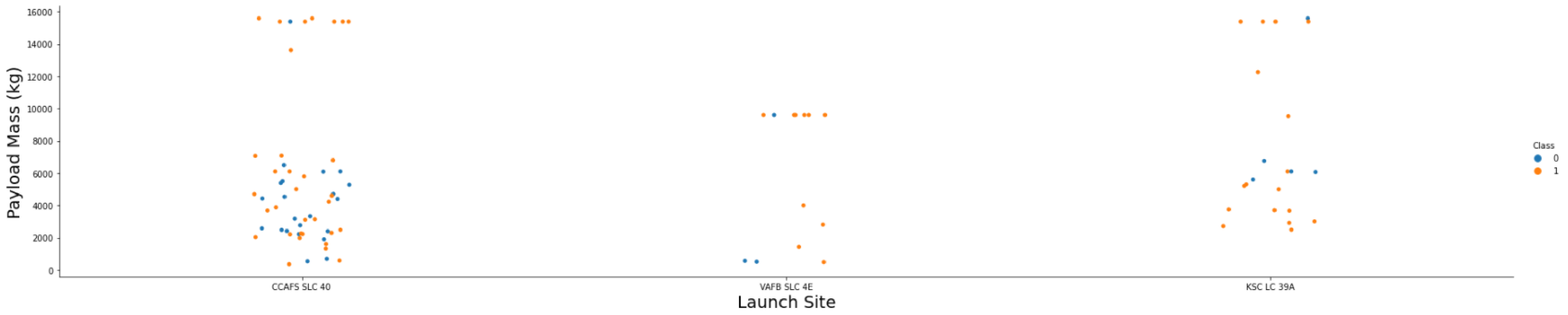


What does the plot tell us?

- We see that CCAFS SLC 40 doesn't do any launches for flight numbers 25-40, VFAB SLC 4E doesn't do any launches for flight numbers above 66, and KSC LC 39A doesn't do any launches for flight numbers less than 27.
    - Perhaps the flight numbers also have meaning with regard to where they launch from.
- VFAB SLC 4E appears to launch less frequently than the others.
    - Maybe it is more specialized than the other two launch sites.
- CCAFS SLC 40 has the most unsuccessful launches.
    - This could simply be since it is the oldest launch site, so there is more data inherently.

Payload vs. Launch Site

Scatter plot of Payload vs. Launch Site



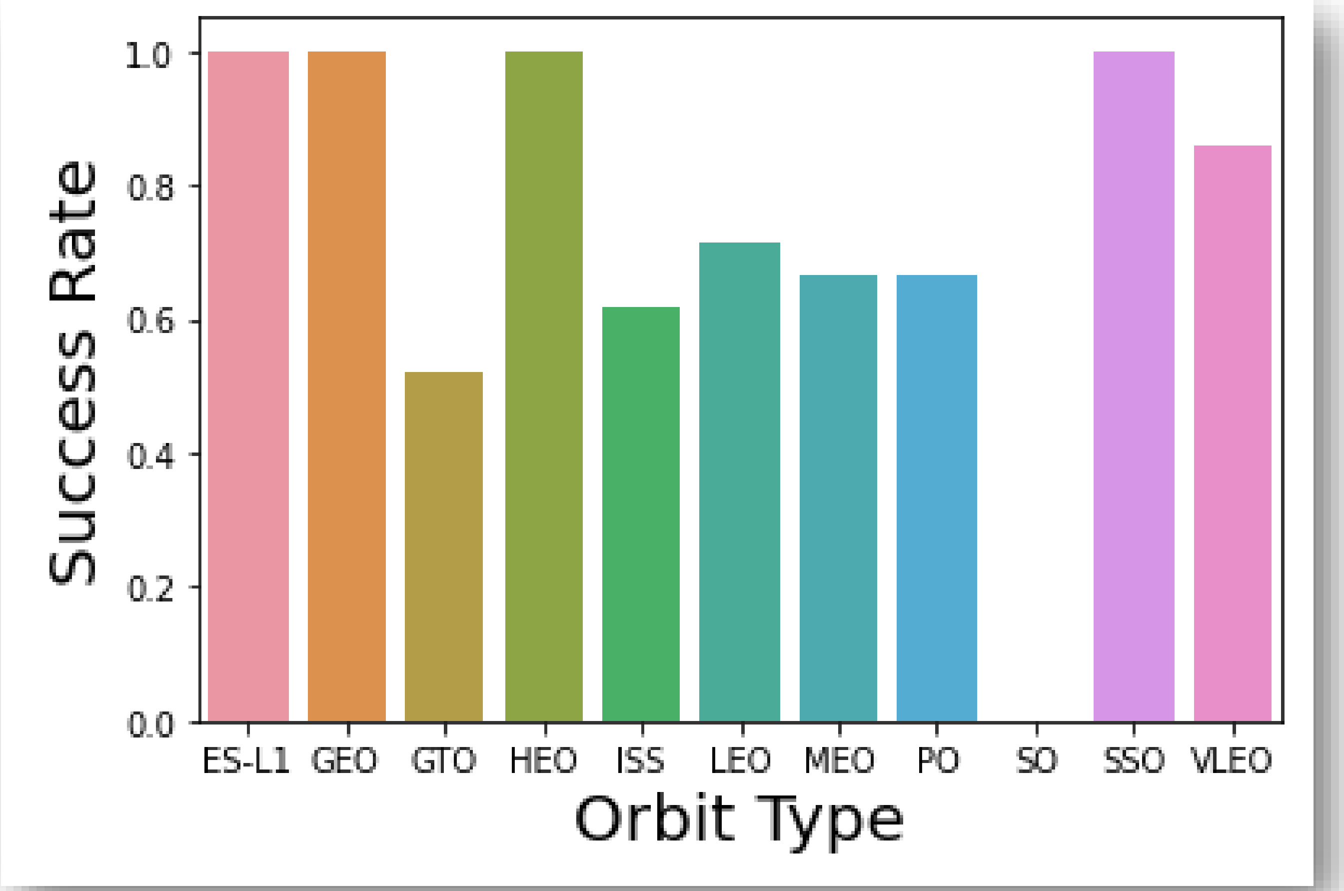
What does this plot tell us?

- We can see that VFAB SLC 4E never launches a payload with a mass greater than 10,000 kg.
  - This could be due to lack of infrastructure or that the site's purpose is to launch smaller payloads.
- KSC LC 39A seems to be the most "well-rounded" because it has launched payloads of almost all masses.
  - It never launches a payload smaller than 4,000 kg, maybe it's designed for heavier ones.
- We also see that CCAFS SLC 40 has launched the most massive payloads.
  - This site seems to be doing most of the launches and could be a "workhorse" for SpaceX.

Success Rate vs. Orbit Type

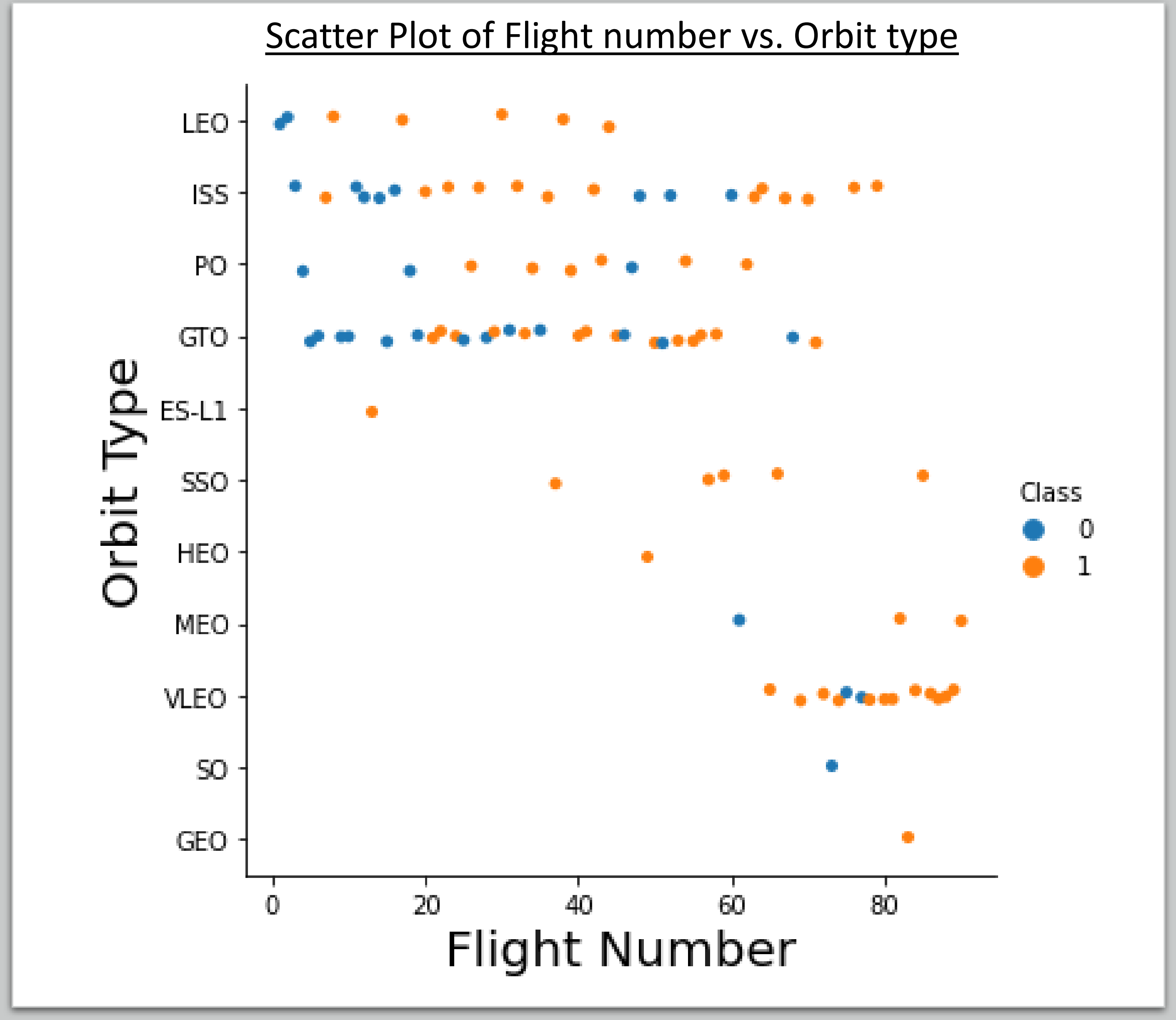
- We can see that the ES-L1, GEO, HEO, and SSO orbits have a success rate of 100%!
- We also note that the SO orbit has a success rate of 0%.
- The chart suggests that the most challenging orbits are those in the middle: GTO, ISS, LEO, MEO, PO, SO.
  - This is likely because this is the point at which maximum dynamic pressure is exerted on the rocket.

Bar Chart for the Success Rate of each Orbit Type



Flight Number vs. Orbit Type

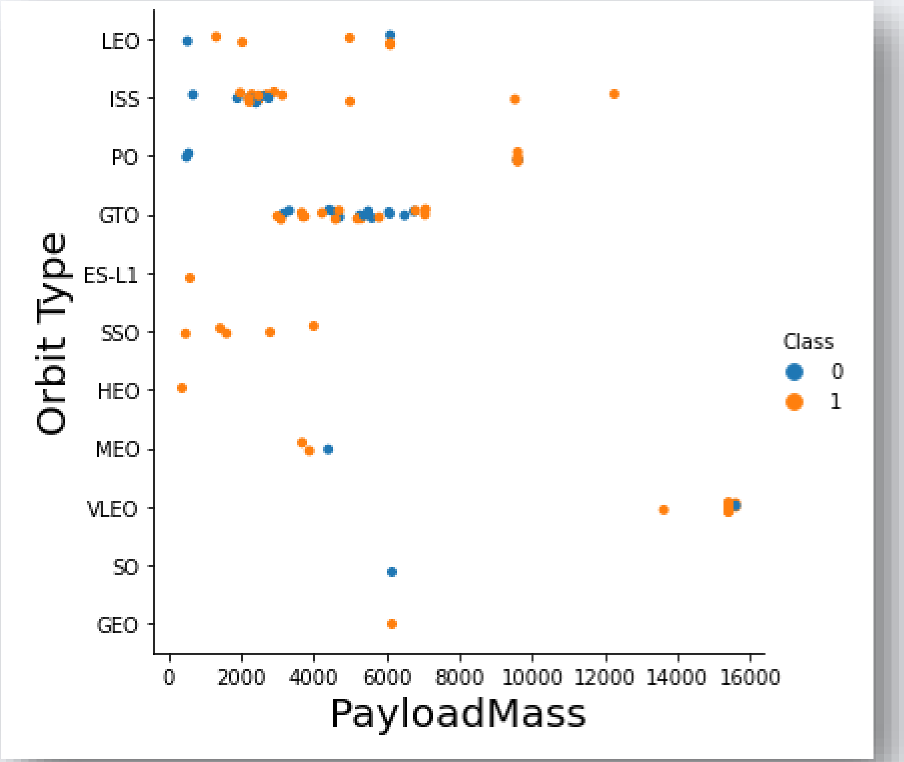
- We see that as the flight number increases for the LEO orbit, the probability of success increases.
  - This is not the case for the GTO orbit, suggesting that this is a particularly difficult orbit to succeed in.
- We also see that there is only one launch for the ES-L1, HEO, SO, and GEO orbits, which would explain why we saw 100% and 0% on the bar graph.
- The graph suggests that the best orbits for success are VLEO, SSO, and LEO due to their per-capita success rates.



Payload vs. Orbit Type

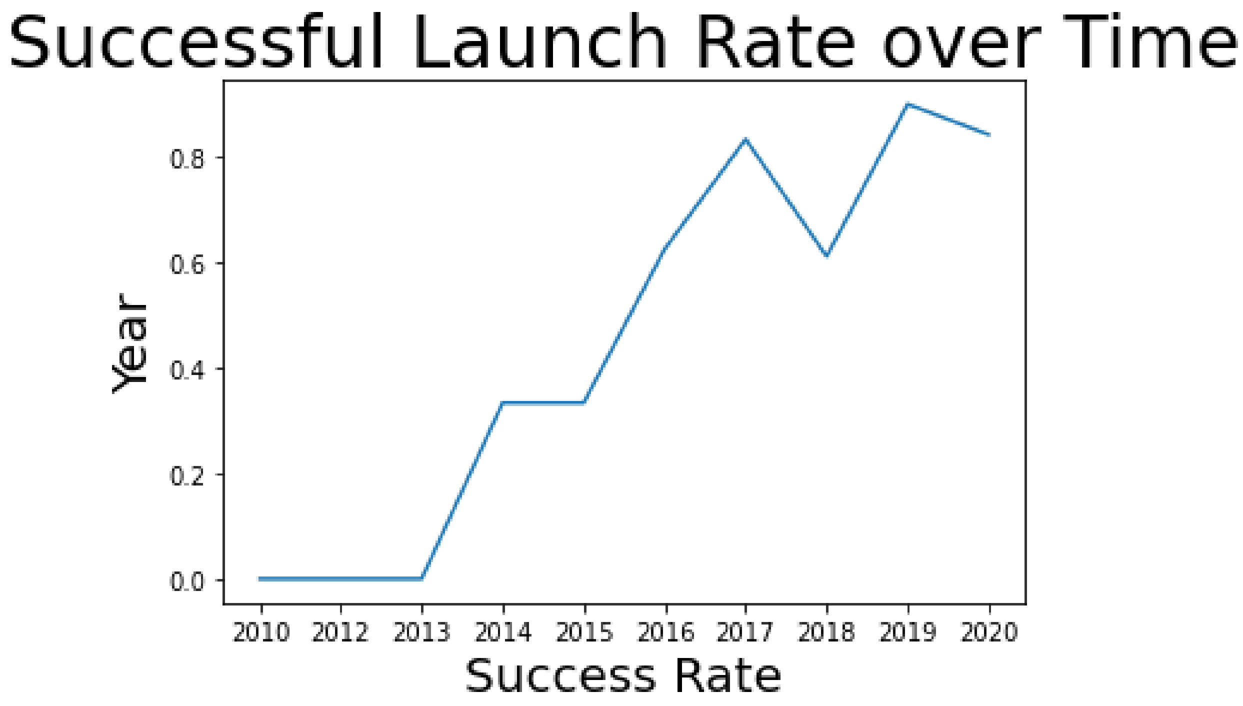
- The plot suggests that the success rates for the PO, LEO, and ISS orbits are higher with heavier payloads.
  - Maybe this has to do with the extra mass adding some stability/inertia?
- Lower masses for the ISS orbit, and virtually all masses for the GTO orbit have precarious success rates.
  - This further supports our idea that the GTO orbit is very difficult to succeed in.
- We also notice that the ES-L1, SSO, and HEO orbits tend to be successful at lower payload masses.
  - The issue is that there's only 7 launches to go off, so we can't necessarily assume that.

Scatter plot of Payload Mass vs. Orbit type



Launch Success Yearly Trend

- For the first three years, we see that SpaceX couldn’t land the first stage at all!
- Then in 2014 the success rate jumps up to 33% and continues climbing above 80% until a dip in 2018 to 61%.
- Finally, the success rate rises to 90% in 2019 but falls to 84.2% in 2020.
- *One question about this plot is: why is there a dip in 2018?*
- *One possible answer could be that in 2018, SpaceX began to branch out from what they had succeeded doing in the past to try a different orbit or booster version, for example.*



Section 3

Explanatory Data Analysis using SQL

All Launch Site Names

- The SQL query shows the names of the launch sites.
- The different launch sites are:
  - CCAFS LC-40
  - CCAFS SLC-40
  - KSC LC-39A
  - VAFB SLV-4E

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

```
%sql select unique(LAUNCH_SITE) from SPACEXDATASET
```

```
 * ibm_db_sa://bpp91703:***@2f3279a5-73d1-4859-88f0-a6c3e
Done.
```

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

Launch Site Names Begin with 'CCA'

- The query displays the launch sites which start with ‘CCA’.
- There are 5 records in the dataset where a launch occurred at the CCAFS LC-40 site.

Display 5 records where launch sites begin with the string 'CCA'

```
%sql select LAUNCH_SITE from SPACEXDATASET where (LAUNCH_SITE) like 'CCA%' limit 5
```

```
 * ibm_db_sa://bpp91703:***@2f3279a5-73d1-4859-88f0-a6c3e6b4b907.c3n41cmd0nqnrk39u9
Done.
```

launch_site
CCAFS LC-40
CCAFS LC-40
CCAFS LC-40
CCAFS LC-40
CCAFS LC-40

Total Payload Mass

- The query calculates the sum of all the payload masses in kilograms.
- The total payload mass carried by all boosters launched is 619,967 kg.

Display the total payload mass carried by boosters launched by NASA (CRS)

```
%sql select sum(PAYLOAD_MASS__KG_) as totalPayloadMass from SPACEXDATASET
```

```
 * ibm_db_sa://bpp91703:***@2f3279a5-73d1-4859-88f0-a6c3e6b4b907.c3n41cmd0
Done.
```

totalpayloadmass
619967

Average Payload Mass by F9 v1.1

- The query calculates the average payload mass of all Falcon 9 v1.1 boosters.
- The average payload mass carried by all the F9 v1.1 boosters is 2,534 kg.

Display average payload mass carried by booster version F9 v1.1

```
%%sql select avg(PAYLOAD_MASS__KG_) as averagePayloadMass from SPACEXDATASET
where BOOSTER_VERSION like 'F9 v1.1%'
```

```
 * ibm_db_sa://bpp91703:***@2f3279a5-73d1-4859-88f0-a6c3e6b4b907.c3n41cmd
Done.
```

1]:

averagepayloadmass
2534

First Successful Ground Landing Date

- This query calculates the first successful ground landing date.
- The first successful ground landing date is December 22nd, 2015.

List the date when the first successful landing outcome in ground pad was acheived.

```
%sql select min(DATE) from SPACEXDATASET where LANDING__OUTCOME like 'Success%'
```

```
 * ibm_db_sa://bpp91703:***@2f3279a5-73d1-4859-88f0-a6c3e6b4b907.c3n41cmd0n
Done.
```

2]:

1
2015-12-22

List the names of the boosters which have success in drone ship and have payload mass greater than 4000 but less than 6000

```
%%sql select BOOSTER_VERSION from SPACEXDATASET where LANDING__OUTCOME like 'Success (drone ship)'
and PAYLOAD_MASS__KG_ between 4000 and 6000
```

```
 * ibm_db_sa://bpp91703:***@2f3279a5-73d1-4859-88f0-a6c3e6b4b907.c3n41cmd0nqnrk39u98g.databases.appdomain.c
Done.
```

5]:

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

Successful Drone Ship Landing with Payload between 4000 and 6000

- This query returns the booster versions which successfully landed on a drone ship and had payload masses between 4,000 kg and 6,000 kg.
- The names of boosters which satisfied this criteria are F9 FT B1022, F9 FT B1026, F9 FT B1021.2, and F9 FT B1031.2.

Total Number of Successful and Failure Mission Outcomes

List the total number of successful and failure mission outcomes

```
%%sql select MISSION_OUTCOME, count(*) from SPACEXDATASET
group by MISSION_OUTCOME
```

```
 * ibm_db_sa://bpp91703:***@2f3279a5-73d1-4859-88f0-a6
Done.
```

1]:

mission_outcome	2
Failure (in flight)	1
Success	99
Success (payload status unclear)	1

- This query calculates the total number of successful and unsuccessful mission outcome (not based on a successful ground landing).
- There was 1 failure in flight, 1 success where the payload status was undetermined, and 99 successful missions.

Boosters that Carried the Maximum Payload

- This query lists the names of boosters which have carried the maximum payload mass.
- It appears that the F9 B5 B10XX.X models are the ones which can carry the maximum payload mass.

List the names of the booster_versions which have carried the maximum payload mass. Use a subquery

```
%%sql select BOOSTER_VERSION from SPACEXDATASET
where PAYLOAD_MASS__KG_ = (select max(PAYLOAD_MASS__KG_) from SPACEXDATASET)
```

```
 * ibm_db_sa://bpp91703:***@2f3279a5-73d1-4859-88f0-a6c3e6b4b907.c3n41cmd0nqnrk39u98g.d
Done.
```

5]:

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

- This query displays the unsuccessful drone ship landings, their booster versions, and the launch site names for the year 2015
- The two booster versions which failed in 2015 were the F9 v1.1 B1012 and F9 v1.1 B1015. They were launched from the same site and failed 3 months apart.

List the failed landing_outcomes in drone ship, their booster versions, and launch site names for in year 2015

```
%%sql select DATE, BOOSTER_VERSION, LANDING__OUTCOME, LAUNCH_SITE from SPACEXDATASET
where extract(YEAR FROM DATE)='2015' and LANDING__OUTCOME = 'Failure (drone ship)'
```

```
 * ibm_db_sa://bpp91703:***@2f3279a5-73d1-4859-88f0-a6c3e6b4b907.c3n41cmd0nqnrk39u98g.databa
Done.
```

4]:

DATE	booster_version	landing__outcome	launch_site
2015-01-10	F9 v1.1 B1012	Failure (drone ship)	CCAFS LC-40
2015-04-14	F9 v1.1 B1015	Failure (drone ship)	CCAFS LC-40

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

- This query ranks the number of times each landing outcome occurred from June 4th, 2010, to March 20th, 2017, in descending order
- The most common landing outcome was no attempt at a landing at all, followed by an equal number of successful and unsuccessful drone ship landings.

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 count(LANDING__OUTCOME), LANDING__OUTCOME from SPACEXDATASET
where DATE between '2010-06-04' and '2017-03-20'
group by LANDING__OUTCOME
order by COUNT(LANDING__OUTCOME) desc
```

```
 * ibm_db_sa://bpp91703:***@2f3279a5-73d1-4859-88f0-a6c3e6b4b907.c3n41cmd0nqnrk39u98g.databases.appdomain.cloud:30756/bludb
Done.
```

8]:

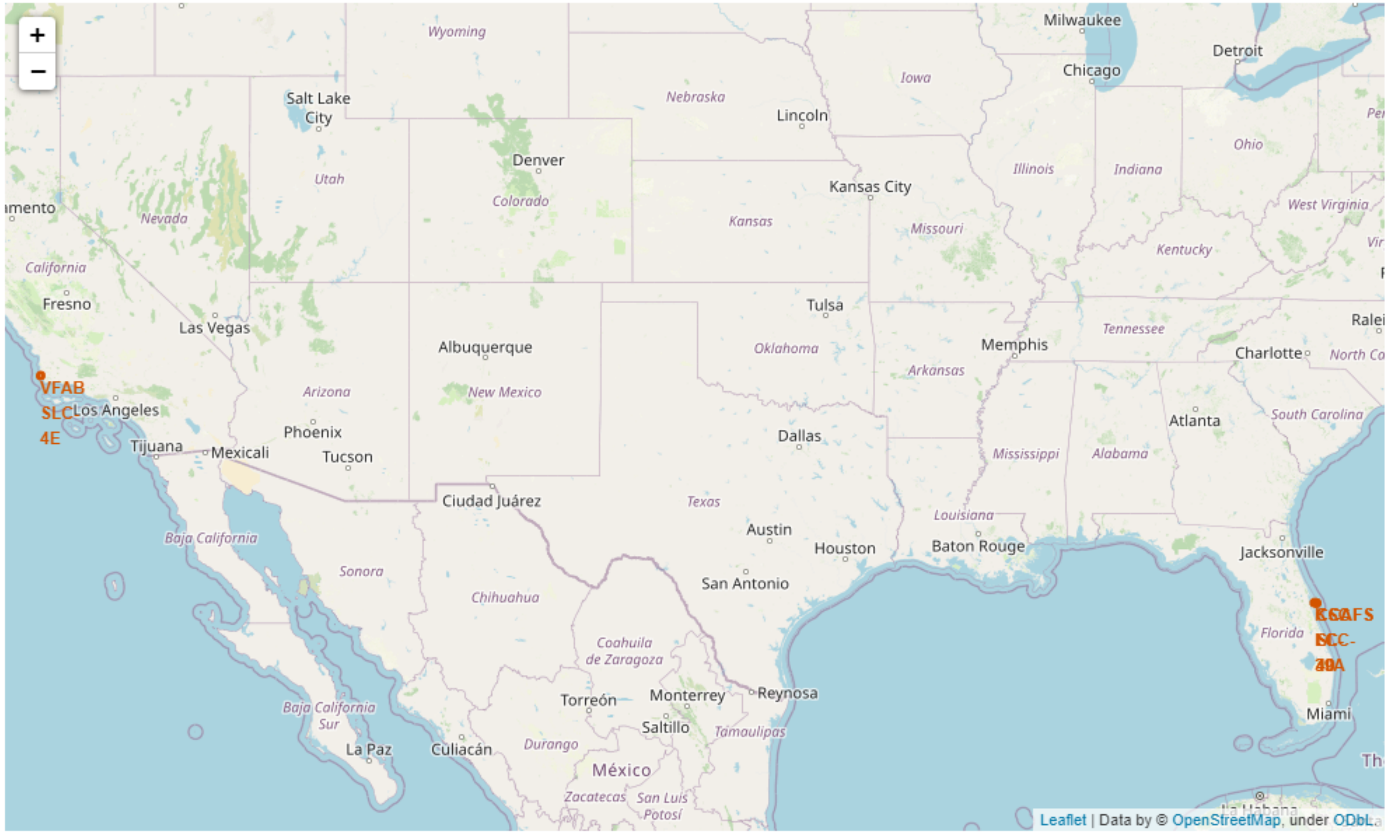
1	landing__outcome
10	No attempt
5	Failure (drone ship)
5	Success (drone ship)
3	Controlled (ocean)
3	Success (ground pad)
2	Failure (parachute)
2	Uncontrolled (ocean)
1	Precluded (drone ship)

Section 4

Launch Sites Proximities Analysis

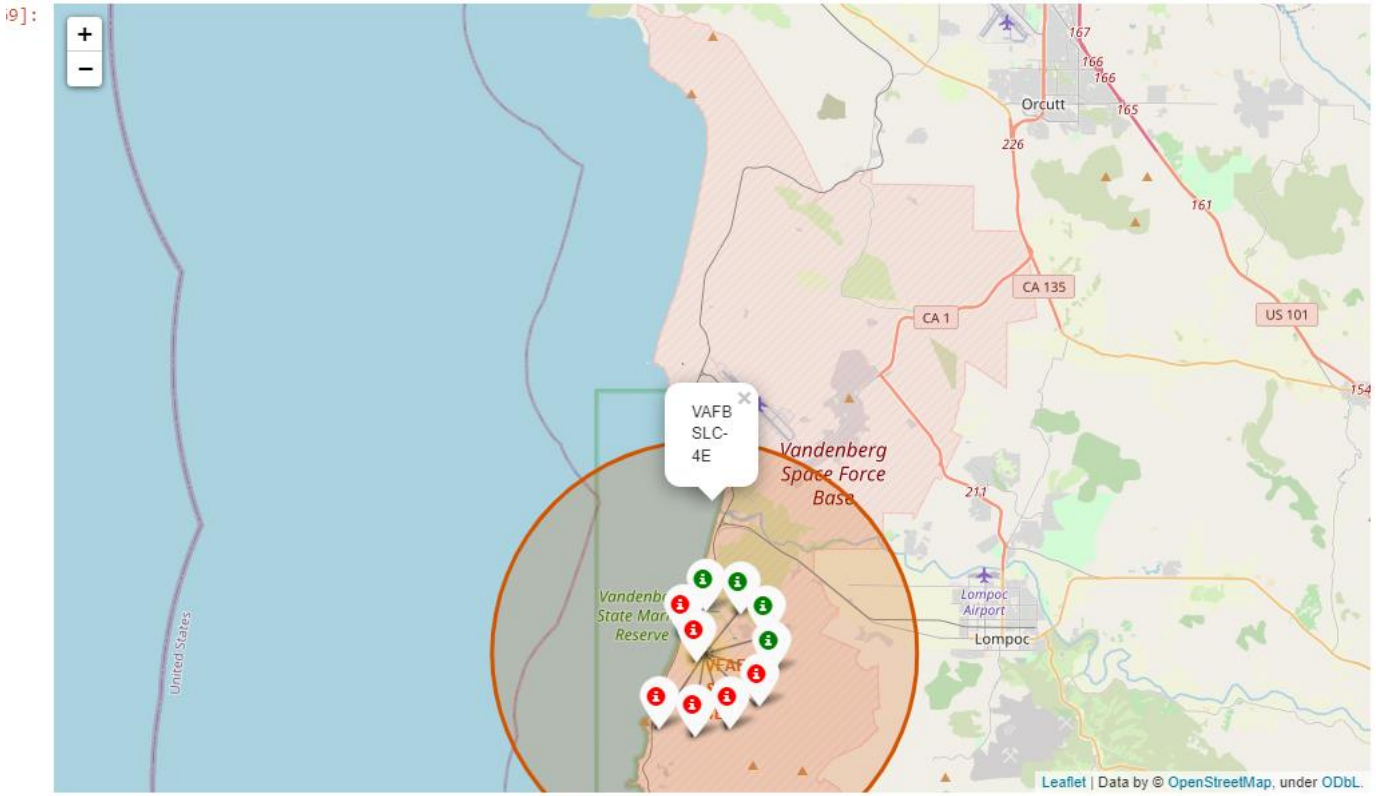
Interactive Folium Map marking Launch Sites

Out[49]:

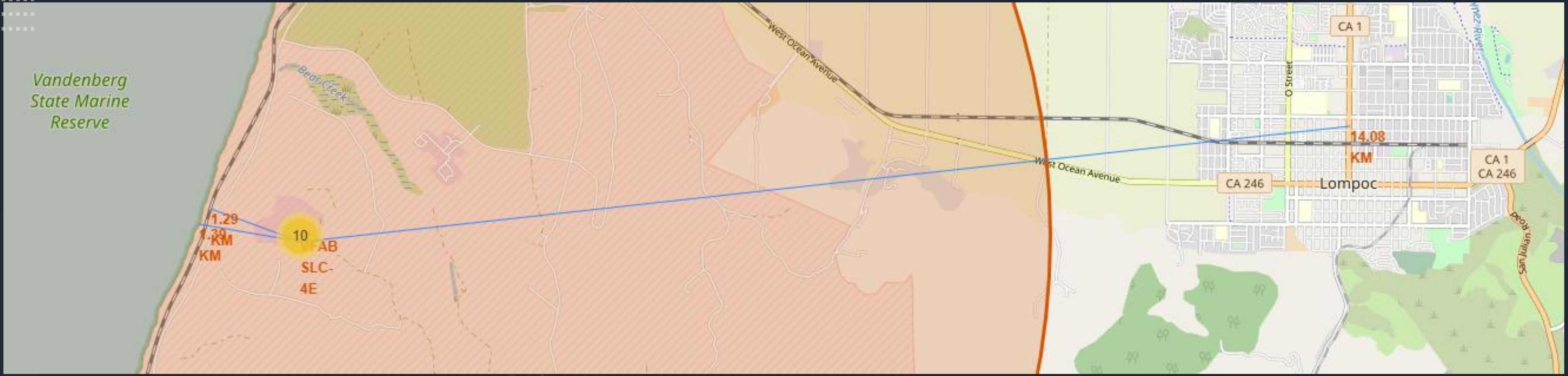


- We can see the launch sites clustered in orange toward southwest California and southeast Florida.
- One very important thing to note is that all the launch sites are close to the equator!
- This is because Earth is a rigid body, so points on the hemispheres must travel at the same speed as those on the equator.
- In other words, if a rocket takes off near the equator, it gets a natural speed boost from the Earth!

Interactive Folium Map with Color-Coded Launch Outcomes



- When we click on the launch site, we can view a history of its launches and outcomes.
- In this case, we see that the VAFB SLC-4E site had 4/10 successful launch outcomes.
- The KSC LC-39A site had 10/13 successful launch outcomes.
- CCAFS SLC-40 had 3/7 successful launch outcomes.
- And CCAFS LC-40 had 7/26 successful launches.
- Perhaps there is a connection between the KSC LC-39A launch site and its high success rate?

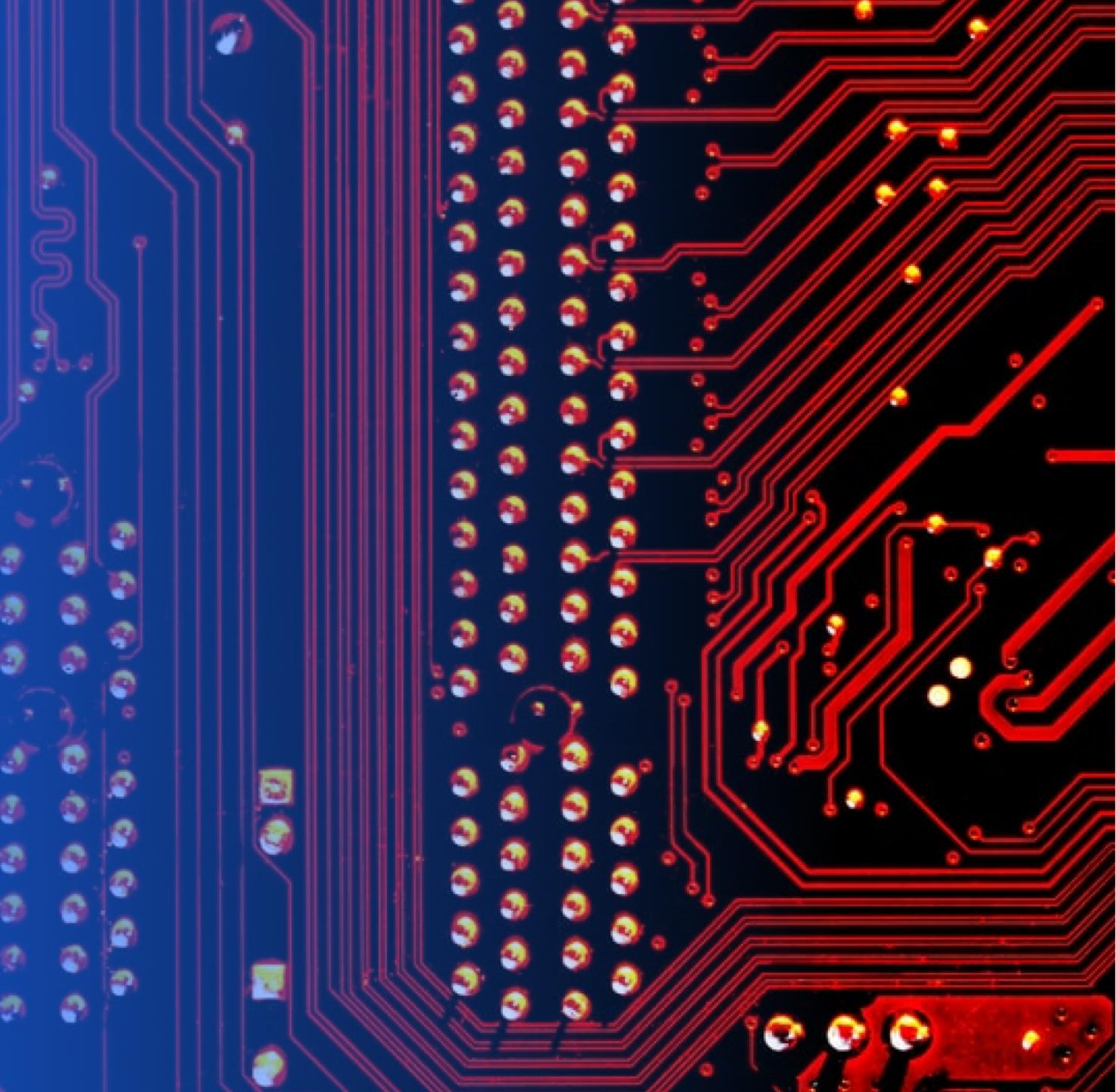


Interactive Folium Map with Distances to Infrastructure

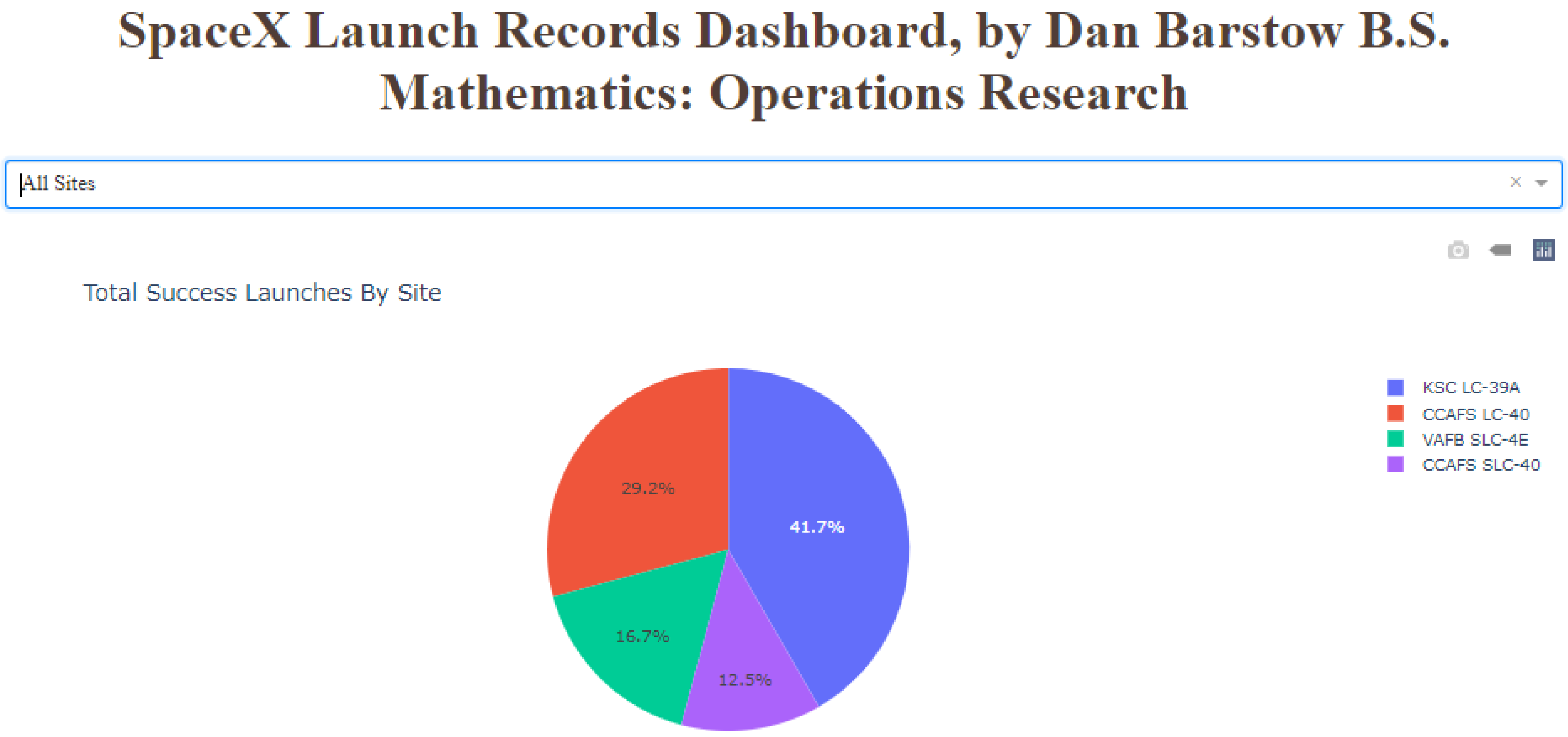
- The map shows 3 distances to the VAFB SLC-4E launch site:
  - The distance from the coastline to the launch site (1.39 km).
  - The distance from the nearest railway to the launch site (1.29 km).
  - The distance from the nearest highway to the launch site (14.08 km).
- We can discern that a launch should be as close to this infrastructure as possible to minimize the time wasted while components are in transport.
- As it relates to this specific site, it’s nearly 3 times further from the nearest highway than the KSC LC-39A site—and the KSC site has a higher success rate.

Section 5

Build a Dashboard with Plotly Dash

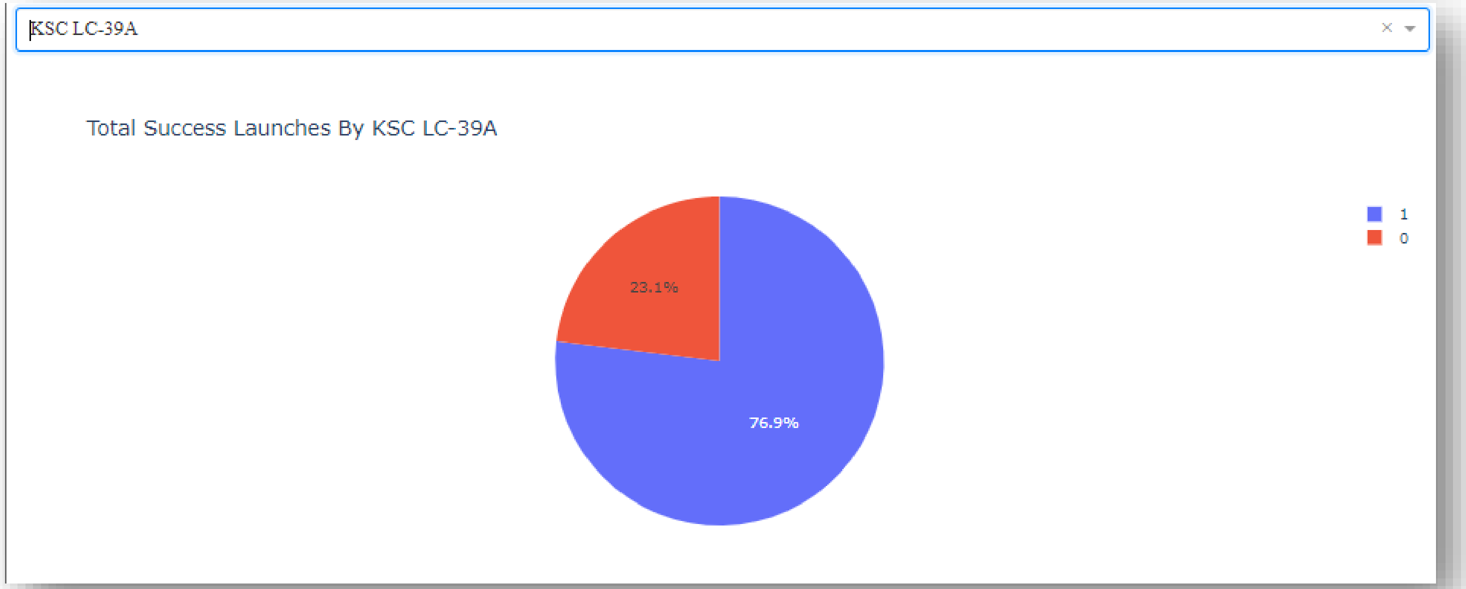


SpaceX Launch Dashboard – All Sites Pie Chart



- Here we see the breakdown of successful launch rates by site.
- We note that the KSC LC-39A site has the highest success rate of 41.7% of successful launches.
- CCAFS SLC-40 has the lowest success rate of 12.5% of the successful launches.

SpaceX Launch Dashboard – KSC LC-39A Pie Chart

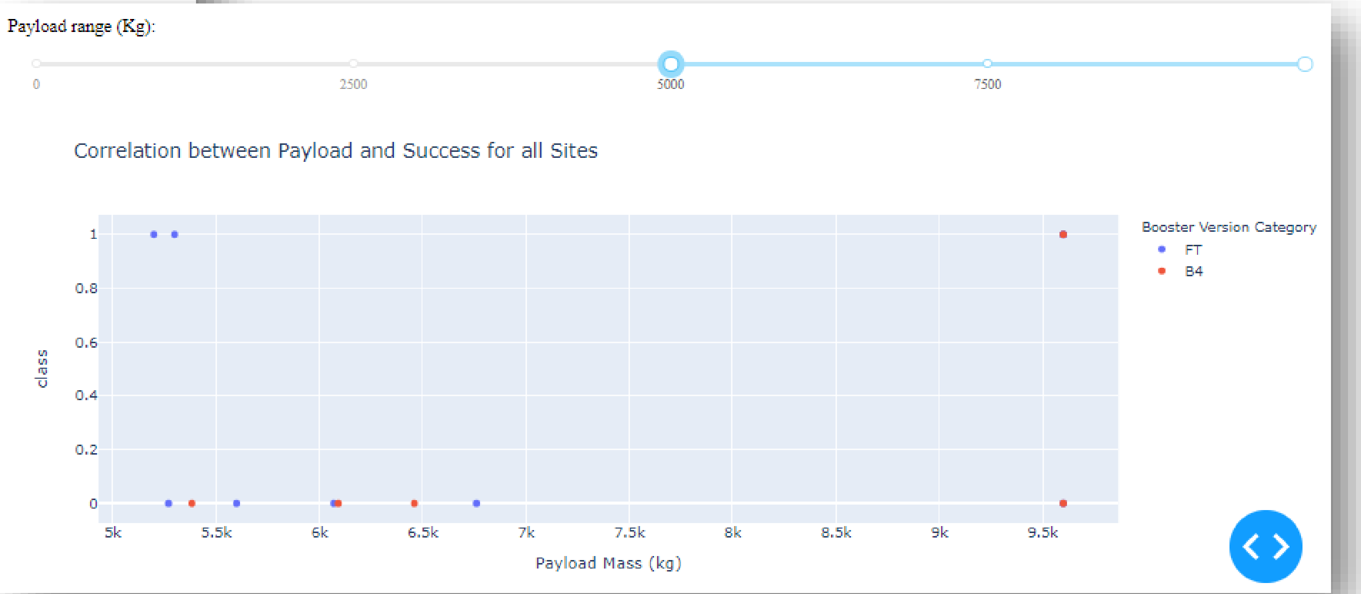


- Here we see that KSC LLC-39A has a launch success rate of 76.9% and a launch failure rate of 22.1%.

SpaceX Launch Dashboard – Scatter Plot of Payload vs. Launch Outcome for all Sites

- For the 0 kg – 10,000 kg payload range, the FT booster seems to have the highest overall success rate.
- It would also seem that most successful launches carry between 2,000 kg and 6,000 kg.

- For the 5,000 kg – 10,000 kg payload range, the FT booster seems to faire better at lower masses, while the B4 booster does better at the higher masses.

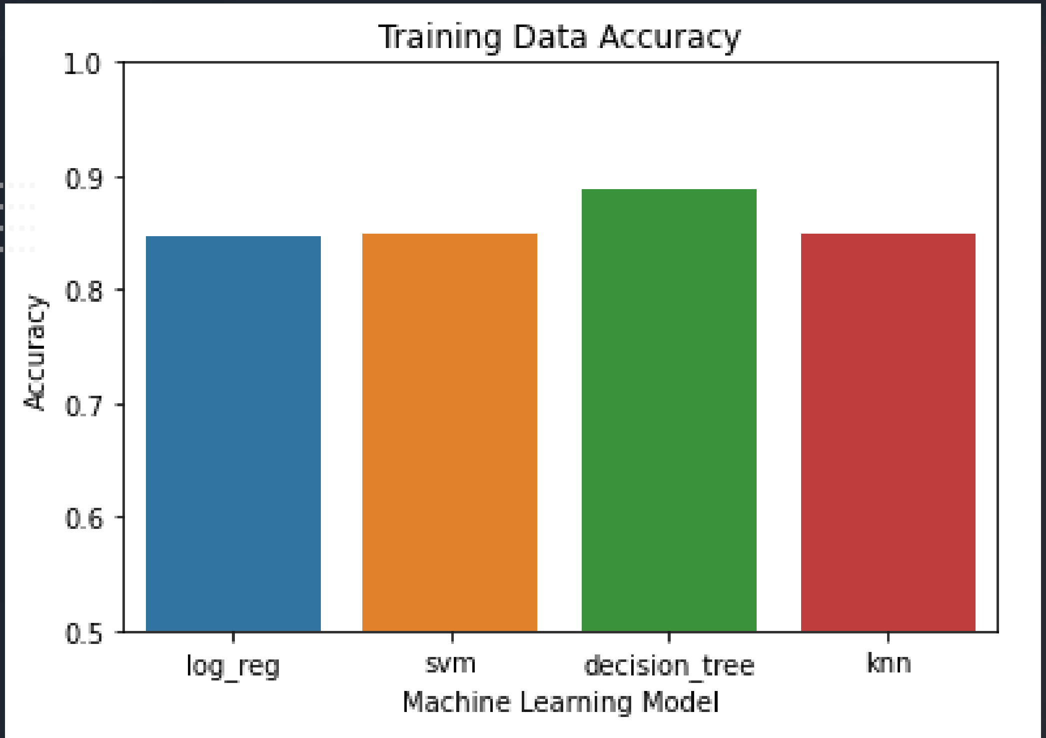


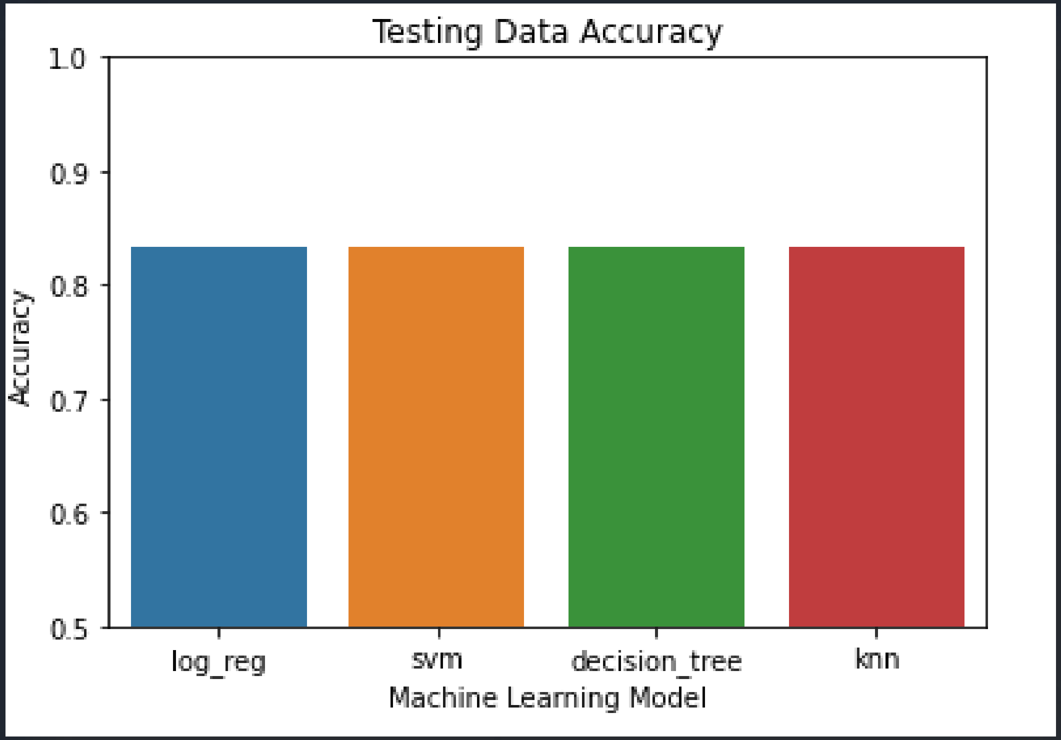
Section 6

Predictive Analysis (Classification)

Classification Accuracy

- With the training set, the decision tree model performed the best with 88.93% accuracy.
- With the testing set, all the models agreed on a predicted launch success rate of 83.33%





In [74]:

```
train_accuracy_3 = tree_cv.best_score_
print("tuned hpyerparameters :(best parameters) ",tree_cv.best_params_)
print("accuracy :",tree_cv.best_score_)
```

```
tuned hpyerparameters :(best parameters)  {'criterion': 'gini', 'max_depth': 12, 'max_features': 'auto', 'min_samples_leaf': 4, 'min_samples_split': 5, 'splitter': 'random'}
accuracy : 0.8892857142857142
```

Confusion Matrix

- Here we see the confusion matrix the for the decision tree model.
- The model correctly classified 15/18 successful launches from the test set.
- The bottom left corner of the confusion matrix tells us that the model made no false negative predictions.
- However, there were 3 false positives or type-I errors in the upper right corner of the confusion matrix.

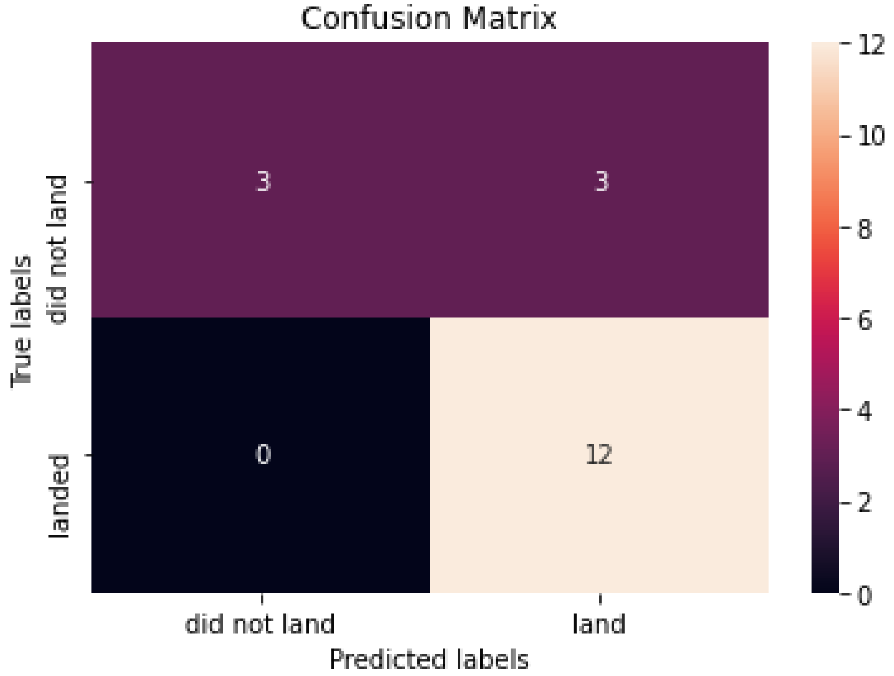
Calculate the accuracy of tree_cv on the test data using the method `score` :

In [60]:
```
tree_score = tree_cv.score(X_test,Y_test)
tree_score
```

Out[60]: 0.8333333333333334

We can plot the confusion matrix

In [61]:
```
yhat = svm_cv.predict(X_test)
plot_confusion_matrix(Y_test,yhat)
```



Conclusions

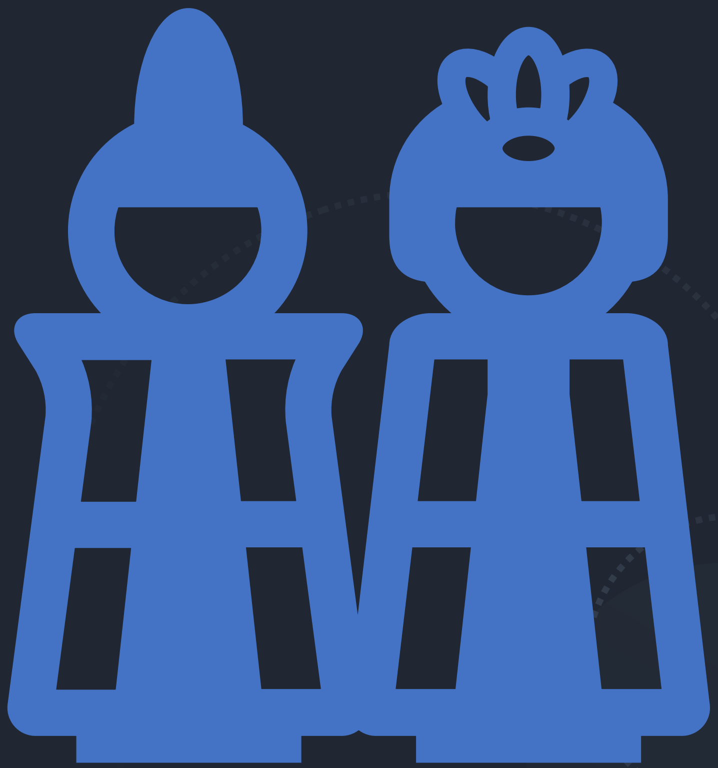
- We have learned that the KSC LC-39A launch site has been the most successful per capita, while the CCAFS SLC-40 launch site seems to have been the workhorse for much of SpaceX's tenure.
- We also learned that the most successful booster has been the Falcon 9 FT booster and had most successful missions tend to occur with payloads between 2,000 kg and 6,000 kg.
- Further, we learned that part of the reason the CCAFS SLC-40 site has been so pivotal to SpaceX's success is because of its proximity to railways, highways, and coastlines.
- From our EDA stage, we found that the most successful orbits are the ES-L1, GEO, HEO, and SSO orbits, while the other orbits in between them are more precarious.

Conclusions (cont.)

- Since all 4 of our machine learning models yielded a predicted success rate of 83.3%, we can be somewhat confident in our prediction.
- However, because the test dataset only consists of 18 samples, we cannot be too certain of our findings.
- We conclude that there is at least some evidence to suggest that the next launch will be successful.

Appendix

- There is certainly more work to be done in this area.
- The limiting factor in this case is the number of launches, given that SpaceX has only completed 134 launches over its 12 years (as of 2021).
- However, one possible remedy to this problem is to add more data to the existing dataset like weather conditions, for example.
- A special thanks to Joseph Santarcangelo, Ph.D. and Lakshmi Holla from IBM for their exceptional tutelage and support—thank you!



Thank you!