IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Lee Det***
6/1/2022

Outline

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

Executive Summary

- Summary of methodologies
  - Data Collection through API
  - Data Collection with Web Scraping
  - Data Wrangling
  - Exploratory Data Analysis with SQL
  - Exploratory Data Analysis with Data Visualization
  - Interactive Visual Analytics with Folium
  - Machine Learning Prediction
- Summary of all results
  - Exploratory Data Analysis result
  - Interactive analytics in screenshots
  - Predictive Analytics result

Introduction

- The mission of this project is to be able to predict whether a Falcon 9 launch will land successfully. Space X has undercut the market because they are able to reuse their rockets by landing them, so their costs are reduced. The company Space Y has hired data scientists to create a machine learning model that will be able to predict whether a rocket will land successfully. With this insight, Space Y will be able to undercut Space X, which is currently leading this space race.
- Problems we wish to solve:
  - Which attributes of a rocket launch contribute most to whether a rocket will successfully land?
  - What, if any, is the interaction between these attributes? For example, is there a correlation between payload mass of the rocket and the orbit type?

Section 1

Methodology

Methodology

Executive Summary

- Data collection methodology:
  - Data was collected from public sources
    - SpaceX Rest API
    - Webscraping Falcon 9 Wikipedia tables
- Perform data wrangling
  - Landing outcomes were simplified into training labels 1 and 0, meaning successful landing and unsuccessful landing respectively.
  - Aids in simplifying data analysis
- Perform exploratory data analysis (EDA) using visualization and SQL
- Perform interactive visual analytics using Folium and Plotly Dash
- Perform predictive analysis using classification models
  - Testing efficacy of various machine learning classification models to determine most accurate model

Data Collection

- The data used in this study was taken from the following public sources
  - SpaceX Rest API with endpoint api.spacexdata.com/v4/launches/past.


```
[{"fairings":{"reused":false,"recovery_attempt":false,"recovered":false,"ships":[]},"links":{"patch":{"small":"https://images2.imgbox.com/3c/0e/T8iJcSN3_o.png","large":"https://images2.imgbox.com/40/e3/GypSkayF_o.png"},"reddit":
{"campaign":null,"launch":null,"media":null,"recovery":null},"flickr":{"small":[],"original":[]},"presskit":null,"webcast":"https://www.youtube.com/watch?v=0a_00nJ_Y88","youtube_id":"0a_00nJ_Y88","article":"https://www.space.com/2196-spacex-inaugural-falcon-1-
rocket-lost-launch.html","wikipedia":"https://en.wikipedia.org/wiki/DemoSat"},"static_fire_date_utc":"2006-03-17T00:00:00.000Z","static_fire_date_unix":1142553600,"net":false,"window":0,"rocket":"5e9d0d95eda69955f709d1eb","success":false,"failures":
[{"time":33,"altitude":null,"reason":"merlin engine failure"}],"details":"Engine failure at 33 seconds and loss of vehicle","crew":[],"ships":[],"capsules":[],"payloads":
["5eb0e4b5b6c3bb0006eeb1e1"],"launchpad":"5e9e4502f5090995de566f86","flight_number":1,"name":"FalconSat","date_utc":"2006-03-24T22:30:00.000Z","date_unix":1143239400,"date_local":"2006-03-25T10:30:00+12:00","date_precision":"hour","upcoming":false,"cores":
[{"core":"5e9e289df35918033d3b2623","flight":1,"gridfins":false,"legs":false,"reused":false,"landing_attempt":false,"landing_success":null,"landing_type":null,"landpad":null}],"auto_update":true,"tbd":false,"launch_library_id":null,"id":"5eb87cd9ffd86e000604b32a"}
,{"fairings":{"reused":false,"recovery_attempt":false,"recovered":false,"ships":[]},"links":{"patch":{"small":"https://images2.imgbox.com/4f/e3/I0lkuJ2e_o.png","large":"https://images2.imgbox.com/be/e7/iNqsqVYM_o.png"},"reddit":
{"campaign":null,"launch":null,"media":null,"recovery":null},"flickr":{"small":[],"original":[]},"presskit":null,"webcast":"https://www.youtube.com/watch?v=Lk4zQ2wP-Nc","youtube_id":"Lk4zQ2wP-Nc","article":"https://www.space.com/3590-spacex-falcon-1-rocket-fails-
reach-orbit.html","wikipedia":"https://en.wikipedia.org/wiki/DemoSat"},"static_fire_date_utc":null,"static_fire_date_unix":null,"net":false,"window":0,"rocket":"5e9d0d95eda69955f709d1eb","success":false,"failures":[{"time":301,"altitude":289,"reason":"harmonic
oscillation leading to premature engine shutdown"}],"details":"Successful first stage burn and transition to second stage, maximum altitude 289 km, Premature engine shutdown at T+7 min 30 s, Failed to reach orbit, Failed to recover first stage","crew":[],"ships":
[],"capsules":[],"payloads":["5eb0e4b6b6c3bb0006eeb1e2"],"launchpad":"5e9e4502f5090995de566f86","flight_number":2,"name":"DemoSat","date_utc":"2007-03-21T01:10:00.000Z","date_unix":1174439400,"date_local":"2007-03-
21T13:10:00+12:00","date_precision":"hour","upcoming":false,"cores":
[{"core":"5e9e289ef35918416a3b2624","flight":1,"gridfins":false,"legs":false,"reused":false,"landing_attempt":false,"landing_success":null,"landing_type":null,"landpad":null}],"auto_update":true,"tbd":false,"launch_library_id":null,"id":"5eb87cdaffd86e000604b32b"
},{"fairings":{"reused":false,"recovery_attempt":false,"recovered":false,"ships":[]},"links":{"patch":{"small":"https://images2.imgbox.com/3d/86/cnu0pan8_o.png","large":"https://images2.imgbox.com/4b/bd/d8UxLh4q_o.png"},"reddit":
{"campaign":null,"launch":null,"media":null,"recovery":null},"flickr":{"small":[],"original":[]},"presskit":null,"webcast":"https://www.youtube.com/watch?v=v0w9p3U8860","youtube_id":"v0w9p3U8860","article":"http://www.spacex.com/news/2013/02/11/falcon-1-flight-3-
mission-summary","wikipedia":"https://en.wikipedia.org/wiki/Trailblazer_(satellite)"},"static_fire_date_utc":null,"static_fire_date_unix":null,"net":false,"window":0,"rocket":"5e9d0d95eda69955f709d1eb","success":false,"failures":
[{"time":140,"altitude":35,"reason":"residual stage-1 thrust led to collision between stage 1 and stage 2"}],"details":"Residual stage 1 thrust led to collision between stage 1 and stage 2","crew":[],"ships":[],"capsules":[],"payloads":
["5eb0e4b6b6c3bb0006eeb1e3","5eb0e4b6b6c3bb0006eeb1e4"],"launchpad":"5e9e4502f5090995de566f86","flight_number":3,"name":"Trailblazer","date_utc":"2008-08-03T03:34:00.000Z","date_unix":1217734440,"date_local":"2008-08-
03T15:34:00+12:00","date_precision":"hour","upcoming":false,"cores":
```

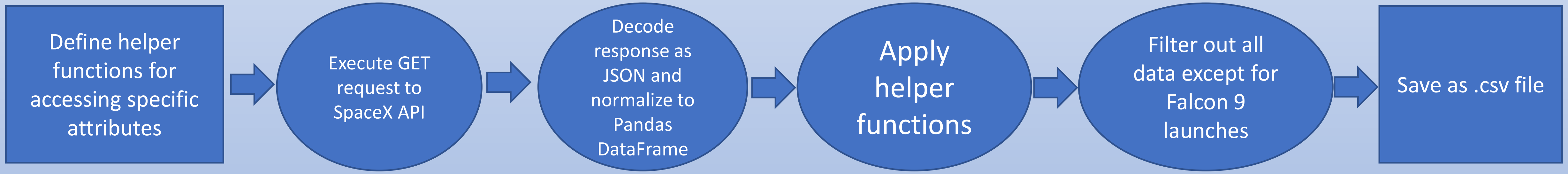

(The contents of the API before import and processing)

  - Data table from Wikipedia titled List of Falcon 9 and Falcon Heavy launches

[hide] Flight No.	Date and time (UTC)	Version, Booster [b]	Launch site	Payload[c]	Payload mass	Orbit	Customer	Launch outcome	Booster landing
1	4 June 2010, 18:45	F9 v1.0[7] B0003.1[8]	CCAFS, SLC-40	Dragon Spacecraft Qualification Unit		LEO	SpaceX	Success	Failure[9][10] (parachute)
	First flight of Falcon 9 v1.0.[11] Used a boilerplate version of Dragon capsule which was not designed to separate from the second stage.(more details below) Attempted to recover the first stage by parachuting it into the ocean, but it burned up on reentry, before the parachutes even deployed.[12]								
2	8 December 2010, 15:43[13]	F9 v1.0[7] B0004.1[8]	CCAFS, SLC-40	Dragon demo flight C1 (Dragon C101)		LEO (ISS)	NASA (COTS) NRO	Success[9]	Failure[9][14] (parachute)
	Maiden flight of Dragon capsule, consisting of over 3 hours of testing thruster maneuvering and reentry.[15] Attempted to recover the first stage by parachuting it into the ocean, but it disintegrated upon reentry, before the parachutes were deployed.[12] (more details below) It also included two CubeSats,[16] and a wheel of Brouère cheese.								
3	22 May 2012, 07:44[17]	F9 v1.0[7] B0005.1[8]	CCAFS, SLC-40	Dragon demo flight C2+[18] (Dragon C102)	525 kg (1,157 lb)[19]	LEO (ISS)	NASA (COTS)	Success[20]	No attempt
	Dragon spacecraft demonstrated a series of tests before it was allowed to approach the International Space Station. Two days later, it became the first commercial spacecraft to board the ISS.[17] (more details below)								
4	8 October 2012, 00:35[21]	F9 v1.0[7] B0006.1[8]	CCAFS, SLC-40	SpaceX CRS-1[22] (Dragon C103)	4,700 kg (10,400 lb)	LEO (ISS)	NASA (CRS)	Success	No attempt
				Orbcomm-OG2[23]	172 kg (379 lb)[24]	LEO	Orbcomm	Partial failure[25]	
	CRS-1 was successful, but the secondary payload was inserted into an abnormally low orbit and subsequently lost. This was due to one of the nine Merlin engines shutting down during the launch, and NASA declining a second reignition, as per ISS visiting vehicle safety rules, the primary payload owner is contractually allowed to decline a second reignition. NASA stated that this was because SpaceX could not guarantee a high enough likelihood of the second stage completing the second burn successfully which was required to avoid any risk of secondary payload's collision with the ISS.[26][27][28]								

(A sample of the data from the Wikipedia page before import and processing)

Data Collection - SpaceX API



[Link to Notebook on GitHub]

Data Collection - Scraping

Define helper functions for processing data → Perform HTTP GET request to Wikipedia page → Create BeautifulSoup object and extract title and column names → Create dictionary with column names as keys → Parse through tables to fill dictionary values with column information. Aided with helper functions → Convert dictionary to data frame then convert to .csv

[Link to Notebook on Github]

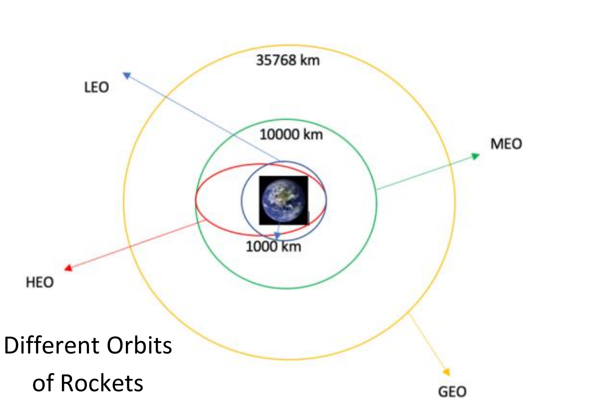
Data Wrangling

Explore how many values are missing for each attribute and which attributes are categorical vs. numerical

Calculate number of launches per site, per orbit. Calculate outcomes for each orbit type

Simplify the different types of good and bad outcomes to a binary system of either good or bad

Add column of binary training labels to dataframe and save as .csv

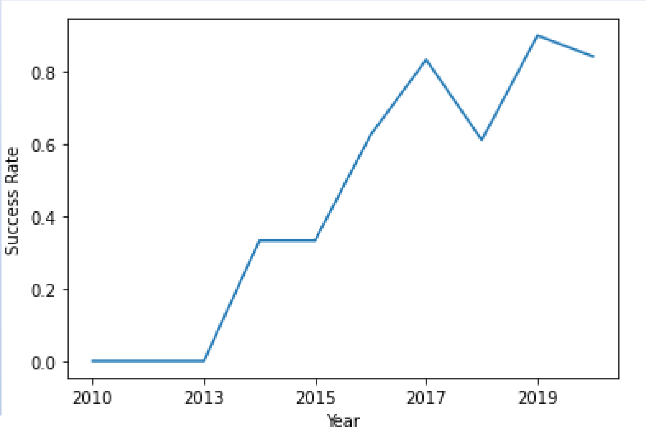


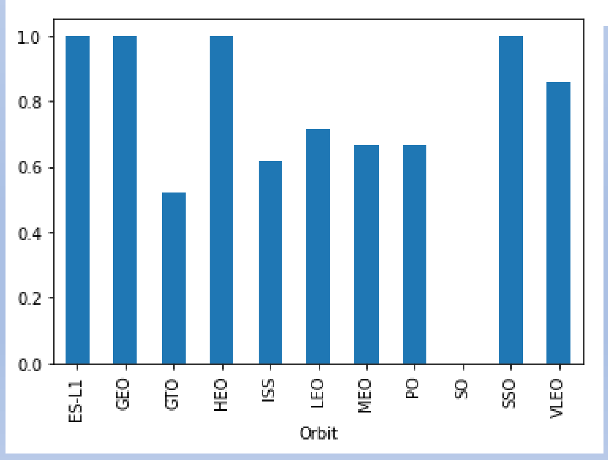
Different Orbits of Rockets

[Link to Notebook on Github]

EDA with Data Visualization

- Preliminary insights can be gained by seeing how the relationship between two attributes affects success rate
- Relationships that were plotted include:
  - Pay load vs. Flight number
  - Flight number vs. Launch Site
  - Payload vs. Launch Site
  - Flight number vs. Orbit Type
  - Payload vs. Orbit Type
- The charts on this page explore the relationships between orbit type vs. success rate and year vs. success rate





[Link to Notebook on Github]

EDA with SQL

- Further Exploratory Data Analysis was performed using SQL
- The queries performed include:
  - Display the names of the unique launch sites in the space mission
  - Display 5 records where launch sites begin with the string 'CCA‘
  - Display the total payload mass carried by boosters launched by NASA (CRS)
  - Display average payload mass carried by booster version F9 v1.1
  - List the date when the first successful landing outcome in ground pad was achieved
  - List the names of the boosters which have success in drone ship and have payload mass greater than 4000 but less than 6000
  - List the total number of successful and failure mission outcomes
  - List the names of the booster versions which have carried the maximum payload mass. Use a subquery
  - List the failed landing outcomes in drone ship, their booster versions, and launch site names for in year 2015
  - Ranking the count of landing outcomes (such as Failure (drone ship) or Success (ground pad)) between the date 2010-06-04 and 2017-03-20

[Link to Notebook on Github]()

Build an Interactive Map with Folium

- A folium map was created that displayed circles for the coordinates of the different launch sites
- The launch cites are added as circles
- Markers for the launches added to clusters and are color coded to indicate if the launch was a success of failure
- Lines were drawn between the launch sites and landmarks, such as the ocean coast and the nearest airport.
  - The lines measure distance and can be used to answer questions like, are launch sites a certain distance away from major cities?

Link to Notebook on Github

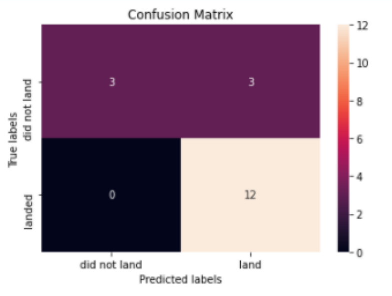
Build a Dashboard with Plotly Dash

- An interactive dashboard was built using Plotly Dash
- The dashboard allows a user to choose a launch site from a drop down menu and payload mass from a range slider
- From these inputs a pie chart showing the total launches by a certain site (or from all sites if no site was selected) and a scatter plot showing the relationship between Payload Mass and Outcome for different booster versions
- This app is a great way to quickly gain insights based on the different launch sites and payload masses

Link to App Source Code on Github

Predictive Analysis (Classification)

- Four different machine learning models were used to see which would be the best models. The models are logistic regression, support vector machine, decision tree, and k nearest neighbors.
- A confusion matrix was made for each model



Standardize Data → Break data up into test and training sets → Build different models and train on the training set → For each model use GridSearchCV to determine the optimal parameters → Test each model on the test set and calculate accuracy → Determine best model based on accuracy score

[Link to Notebook on Github]

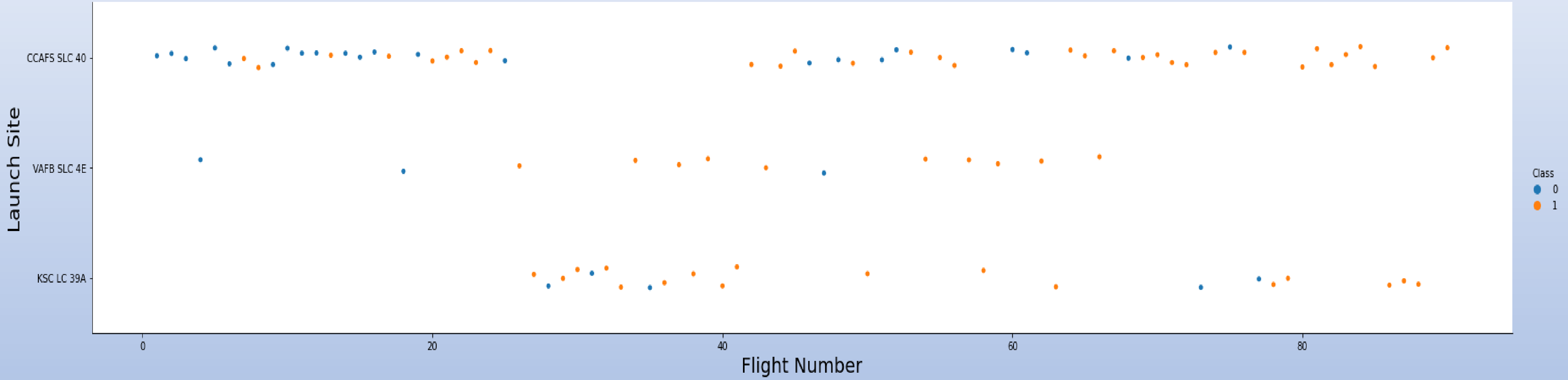
Results

- Exploratory data analysis results
- Interactive analytics demo in screenshots
- Predictive analysis results

Section 2

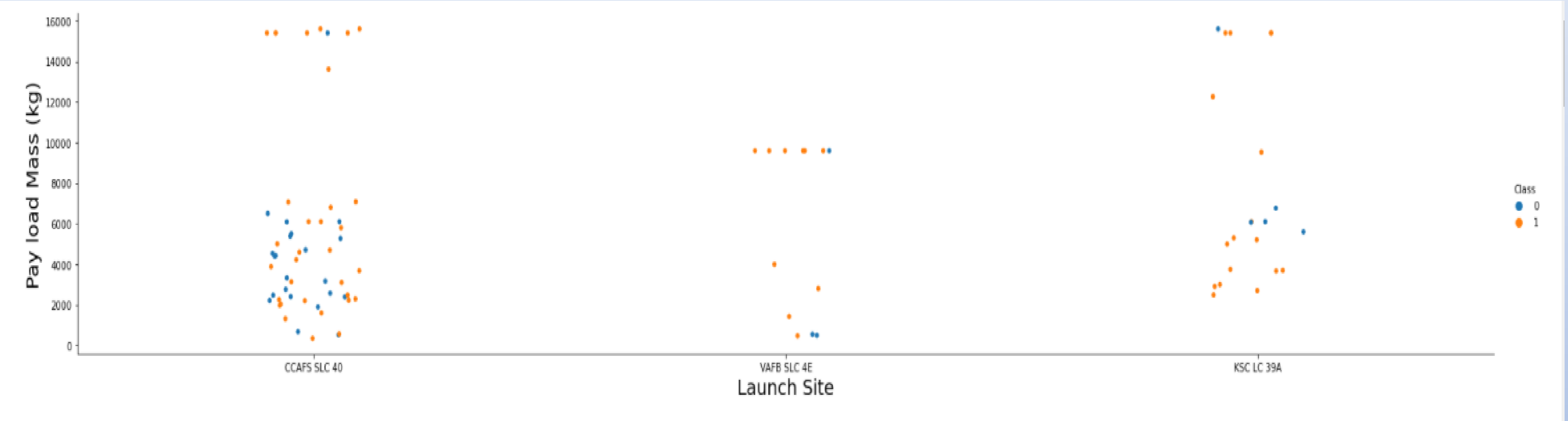
Insights drawn from EDA

Flight Number vs. Launch Site



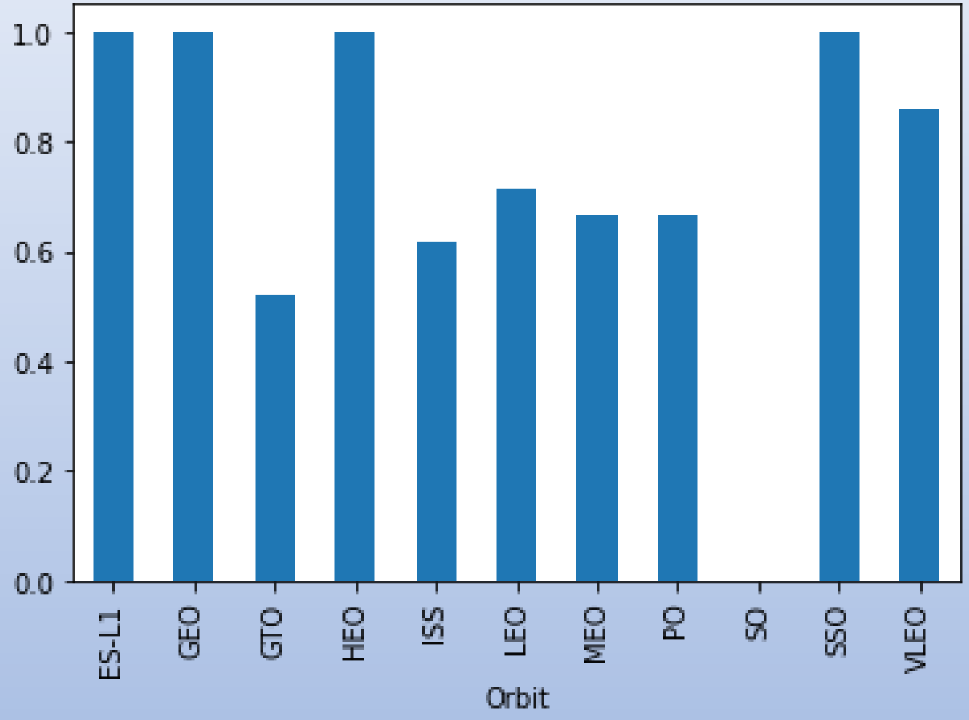
From the plot, it can be concluded that as the flight number increases the chance of a successful launch for the three sites increases

Payload vs. Launch Site



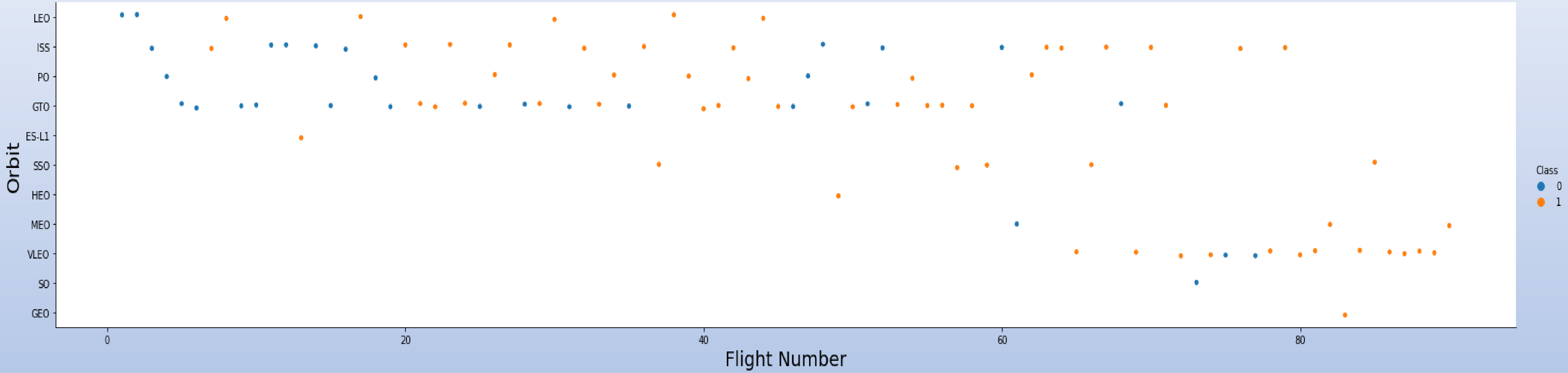
- For the site CCAFS SLC 40, nearly all rockets with a pay load mass above 10,000 kg were successfully landed
- For the site VAFB SLC 4E, no rockets above 10,000 kg were launched
- For site KSC LC 39A, rockets around 6,000 kg had a low rate of being successfully landed

Success Rate vs. Orbit Type



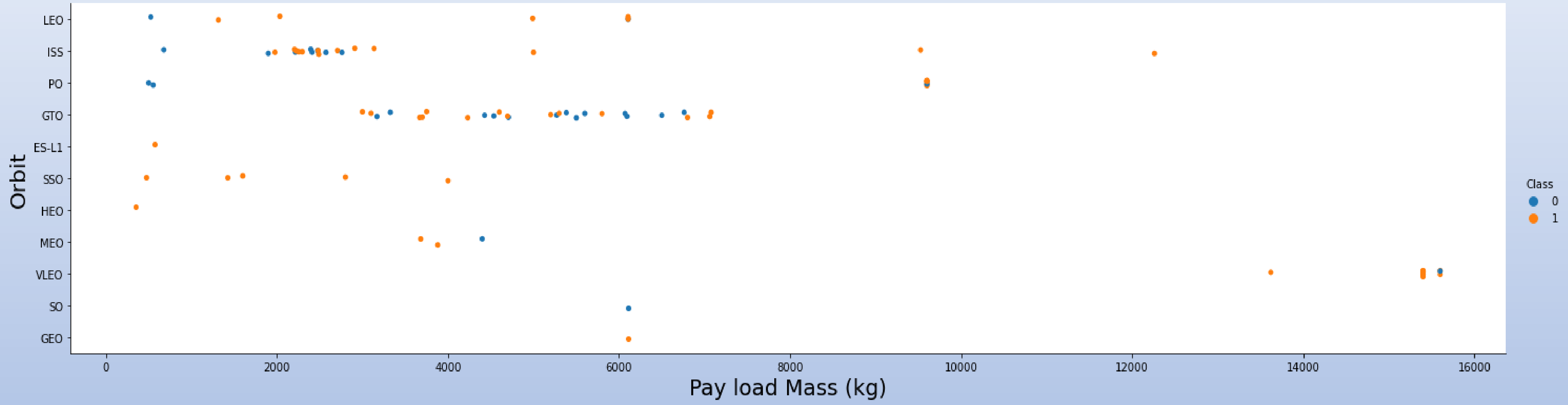
Orbits ES-LL1, GEO, HEO, SSO, and VLEO had that highest success rates

Flight Number vs. Orbit Type



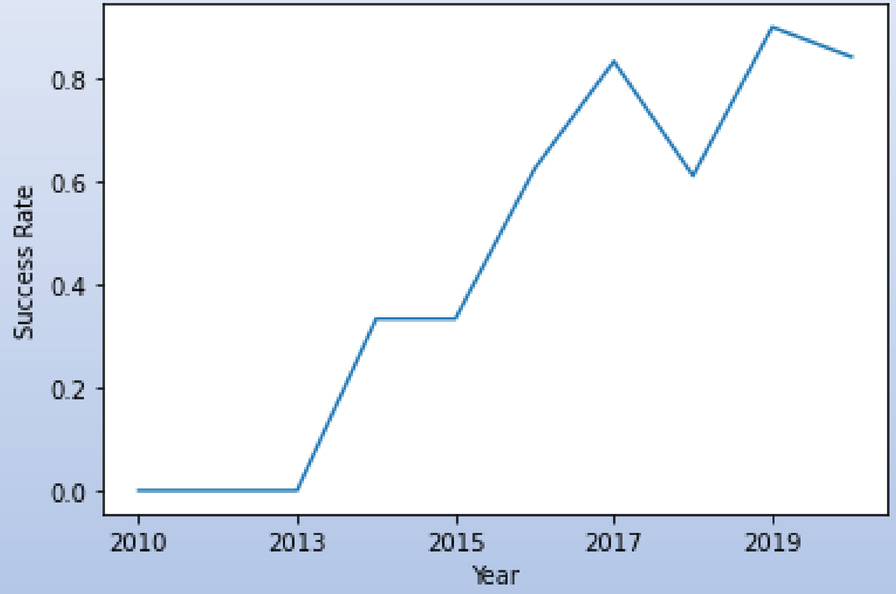
- It is evident that in the LEO orbit the success appears related to the number of flights, but there seems to be no relationship between flight number when in GTO orbit
- This scatter plot gives further insight to the bar chart on the previous slide. For example, the bar chart displayed the GEO orbit as having a 100% success rate. Although, looking at this graph it is seen that only one launch one attempted in this orbit and it was successful

Payload vs. Orbit Type



For some orbits, e.g. Polar,LEO and ISS, a positive relationship is shown between the increase of Pay Load Mass and success rate. Whereas, for other orbits, such as GTO, there is no distinguishable relationship.

Launch Success Yearly Trend



From 2013 to 2020 that success rate has steadily increased.

All Launch Site Names

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

In [37]:
```
%sql select distinct launch_site from SPACEXDATASET
```

```
 * ibm_db_sa://zjv16128:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32733/bludb
Done.
```

Out[37]:

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

Using the keyword Distinct, we are able to query for the unique launch sites in the data set

Launch Site Names Begin with 'CCA'

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

In [13]:

```
%sql select launch_site from SPACEXDATASET where launch_site like 'CCA%' limit 5
```

```
 * ibm_db_sa://zjv16128:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32733/bludb
Done.
```

Out[13]:

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

Using the keyword 'like', we are able to query for launch sites beginning with 'CCA'

Total Payload Mass

```
In [28]: %sql select sum(payload_mass__kg_) as total_payload_mass_for_nasa from SPACEXDATASET where customer='NASA (CRS)'

 * ibm_db_sa://zjv16128:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32733/bludb
Done.
```

Out[28]:

total_payload_mass_for_nasa
45596

Using Sum and specifying the customer as NASA, we are able to calculate the total payload mass for boosters launched by NASA

Average Payload Mass by F9 v1.1

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

In [38]:
```
%sql select avg(payload_mass__kg_) as avg_payload_mass_F9_one_one from SPACEXDATASET where booster_version = 'F9 v1.1'
```

```
 * ibm_db_sa://zjv16128:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32733/bludb
Done.
```

Out[38]:

avg_payload_mass_f9_one_one
2928

Using avg (function for average) and specifying the booster version as ‘F9 v1.1’, we can find the average payload mass for that booster.

First Successful Ground Landing Date

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

Hint:Use min function

In [52]:
```
%sql select min(DATE) from SPACeXDATASET where mission_outcome='Success'
```

```
 * ibm_db_sa://zjv16128:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32733/bludb
Done.
```

Out[52]:

1
2010-06-04

The min (minimum) function allows us to find the earliest date where the outcome of the mission, i.e. launch, was a success.

Successful Drone Ship Landing with Payload between 4000 and 6000

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

In [55]:
```
%sql select booster_version,payload_mass__kg_ from SPACEXDATASET where (payload_mass__kg_ >4000 and payload_mass__kg_ <6000) order by payload_mass__kg
```

```
 * ibm_db_sa://zjv16128:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32733/bludb
Done.
```

Out[55]:

booster_version	payload_mass__kg_
F9 v1.1 B1014	4159
F9 B5 B1051.2	4200
F9 FT B1032.2	4230
F9 B5B1060.1	4311
F9 B5B1062.1	4311
F9 B5B1054	4400
F9 v1.1 B1011	4428
F9 v1.1	4535
F9 FT B1026	4600
F9 FT B1022	4696
F9 v1.1 B1016	4707
F9 B5 B1048.3	4850
F9 B4 B1040.1	4990
F9 B4 B1043.1	5000
F9 FT B1031.2	5200
F9 FT B1020	5271
F9 FT B1021.2	5300
F9 FT B1032.1	5300
F9 B5 B1047.2	5300
F9 B4 B1040.2	5384
F9 B5 B1058.2	5500
F9 FT B1030	5600
F9 B5 B1046.2	5800

We were able to query for a list of booster names with a payload mass between 4000 kg and 6000 kg that successfully landed on a drone ship. We presented the list in ascending order of payload mass

Total Number of Successful and Failure Mission Outcomes

List the total number of successful and failure mission outcomes

In [68]:
```
%sql select  mission_outcome , count(mission_outcome) from SPACEXDATASET group by mission_outcome
```

```
 * ibm_db_sa://zjv16128:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32733/bludb
Done.
```

Out[68]:

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

We found the distinct mission outcomes and how instances of each outcome there were. Effectively 100 successes and 1 failure.

Boosters Carried Maximum Payload

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

In [69]:

```
%sql select booster_version, payload_mass__kg_ from SPACEXDATASET where payload_mass__kg_ = (select max(payload_mass__kg_) from SPACEXDATASET)
```

```
 * ibm_db_sa://zjv16128:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32733/bludb
Done.
```

Out[69]:

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

By using a subquery, we’re able to query for the names of booster versions that carried the maximum payload mass

2015 Launch Records

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

In [77]:

```
%sql select DATE, booster_version, launch_site from SPACEXDATASET where (landing__outcome='Failure (drone ship)' and year(DATE)=2015)
```

```
 * ibm_db_sa://zjv16128:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32733/bludb
Done.
```

Out[77]:

DATE	booster_version	launch_site
2015-01-10	F9 v1.1 B1012	CCAFS LC-40
2015-04-14	F9 v1.1 B1015	CCAFS LC-40

In this query, we find the date, booster version, and launch site of failed landings in the year 2015. Notice the use of year() to specifically extract the year from a date so it can be used in the query.

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

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

In [110…

```
%sql select landing__outcome , count(landing__outcome) as "count"  from SPACEXDATASET where (DATE > '2010-06-04' and DATE<'2017-03-20') group by landi
```

```
 * ibm_db_sa://zjv16128:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32733/bludb
Done.
```

Out[110…

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

In this final query, we find the number of unique landing outcomes between the two specified dates. We include the landing outcome and numbers of instances in the resulting query.

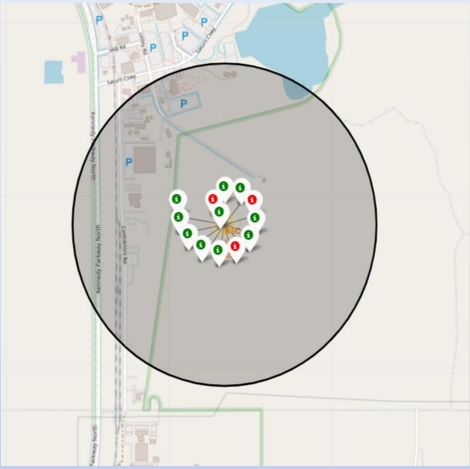
Section 3

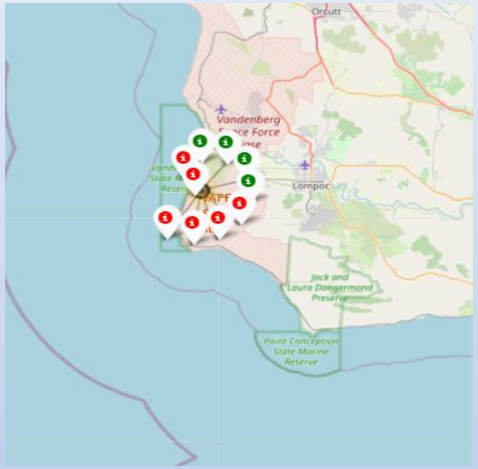
Launch Sites Proximities Analysis

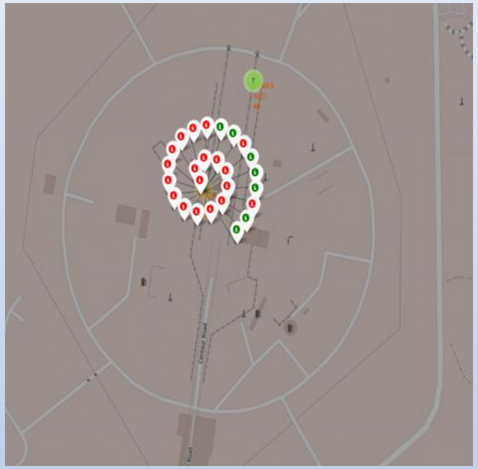
All Launch Sites displayed on Global Map

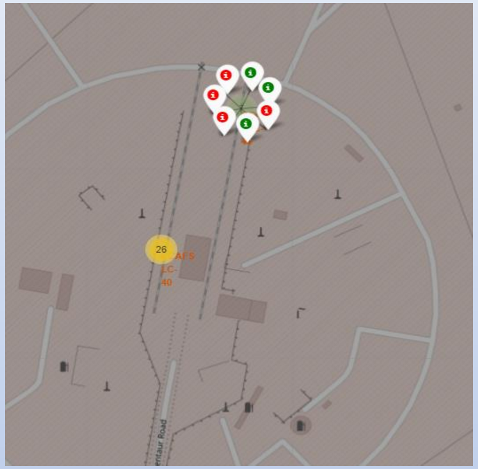
This map of the entire globe displays the locations of the launch sites of this data set as well as NASA headquarters. The Launch sites are on the East and West Coast of the USA

Color Coded Launch Sites





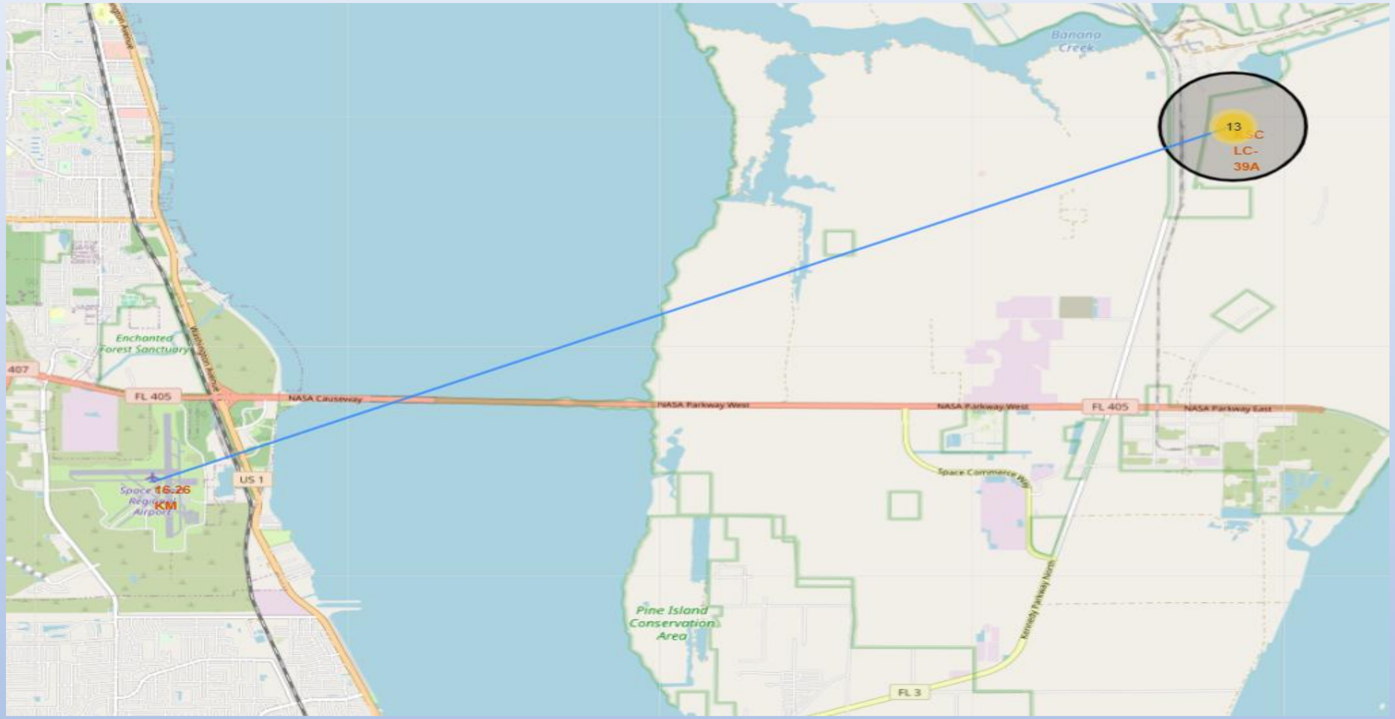




Green Marker = Success Red Marker = Failure

The above images display enhanced views of the four launch sites. They display the launch and the color of the marker indicates if the launch was successful or unsuccessful.

Launch Site Distance to Landmarks

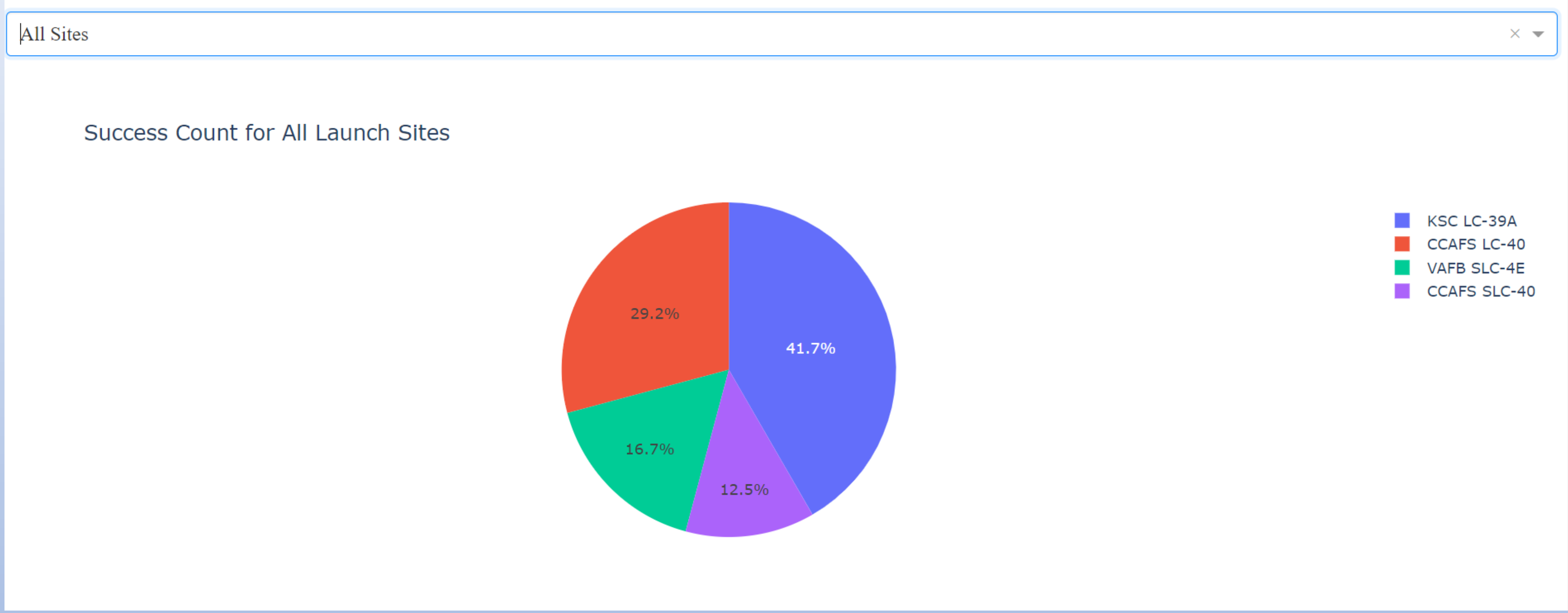


- This image shows the distance between a launch site and the closest regional airport. The distance displayed is 16.26 km
- Distances to other landmarks can also be measured to answer questions such as:
  - Do launch sites keep certain distance away from cities?
  - Are launch sites in close proximity to railways, highways, and/or coastlines?

Section 4

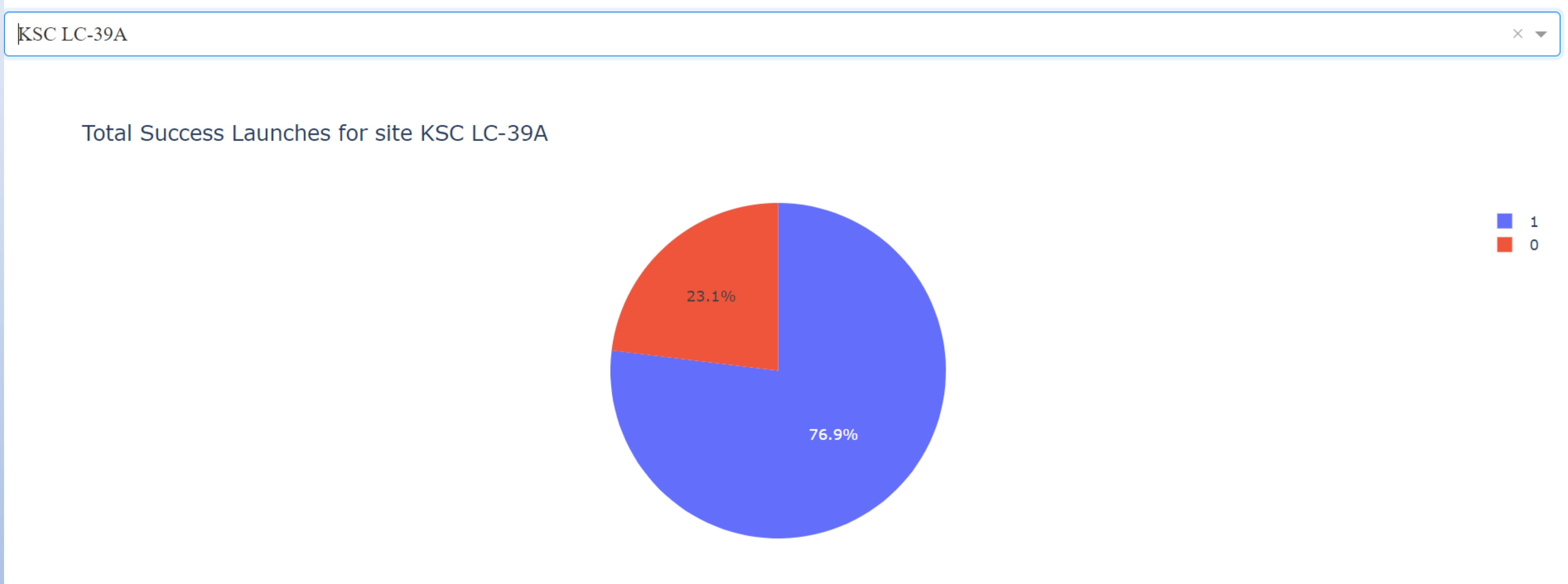
Build a Dashboard with Plotly Dash

Success Count for All Launch Sites Pie chart



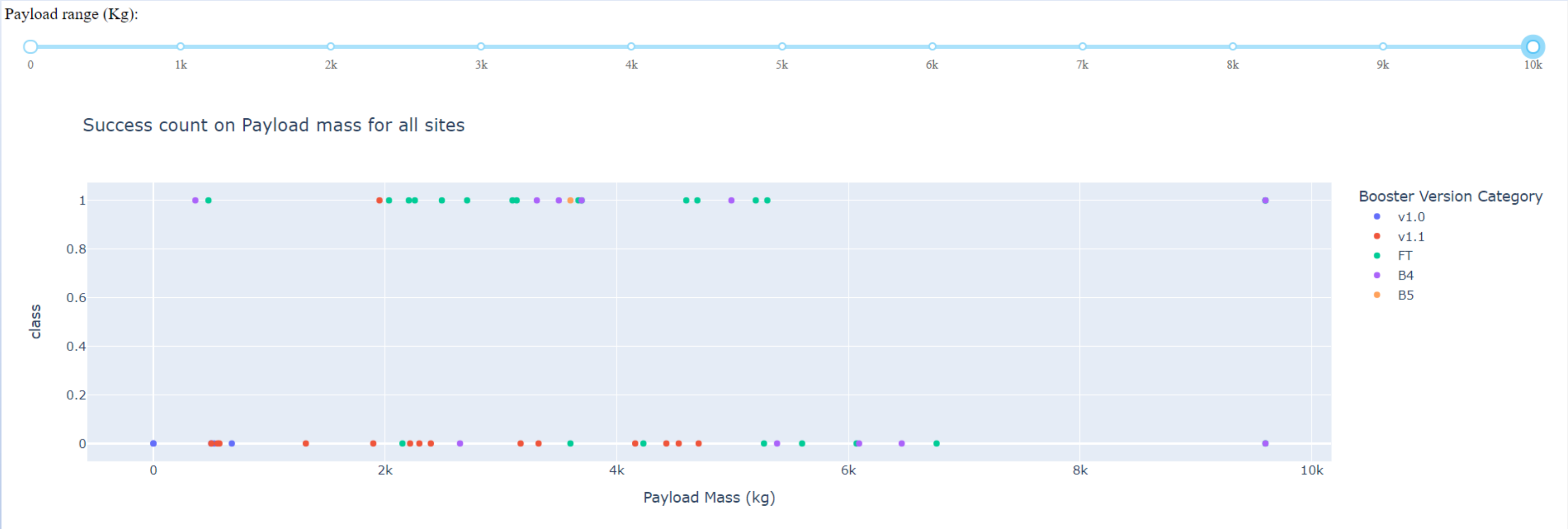
Just a quick glance at this gives a wealth of information. For instance, we can quickly learn that the launch site KSC LC-39A was the site with the best success rate, followed by CCAFS LC-40.

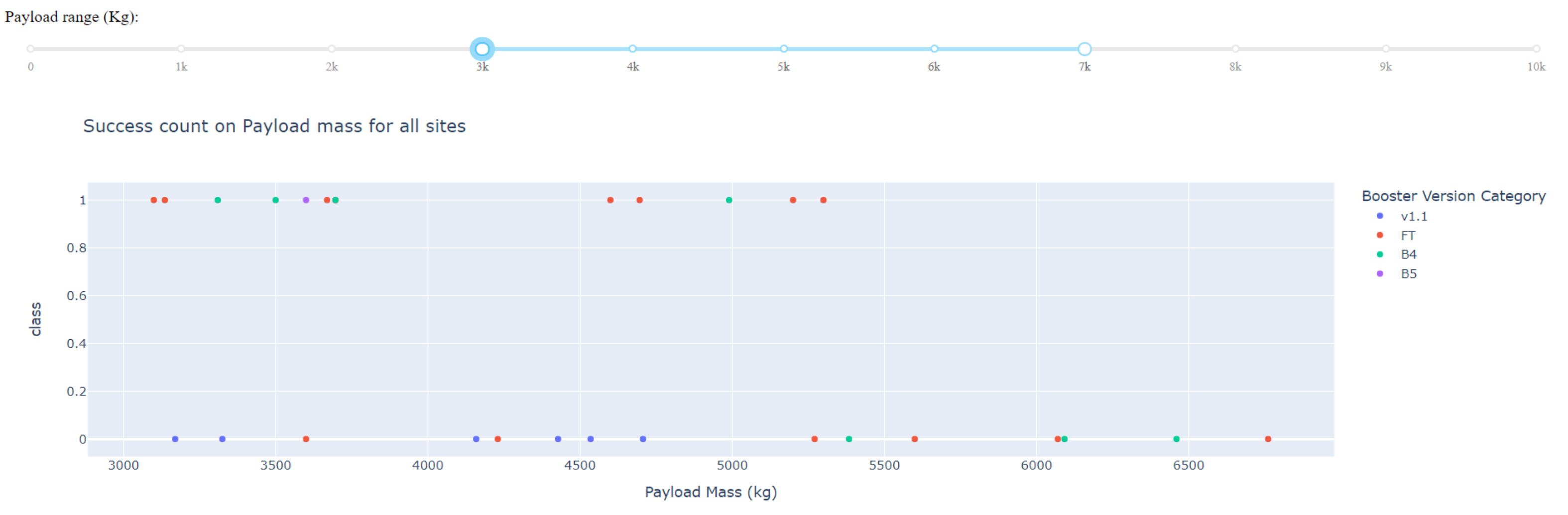
Pie chart for Most Successful Site



Choosing the site with the best success rate, KSC LC-39A, from the drop down menu gives an enhanced view of that site. The piechart shows the percent success vs the percent failure for that site

Scatter Plots for Different Payload Mass Ranges vs. Outcome for All Sites



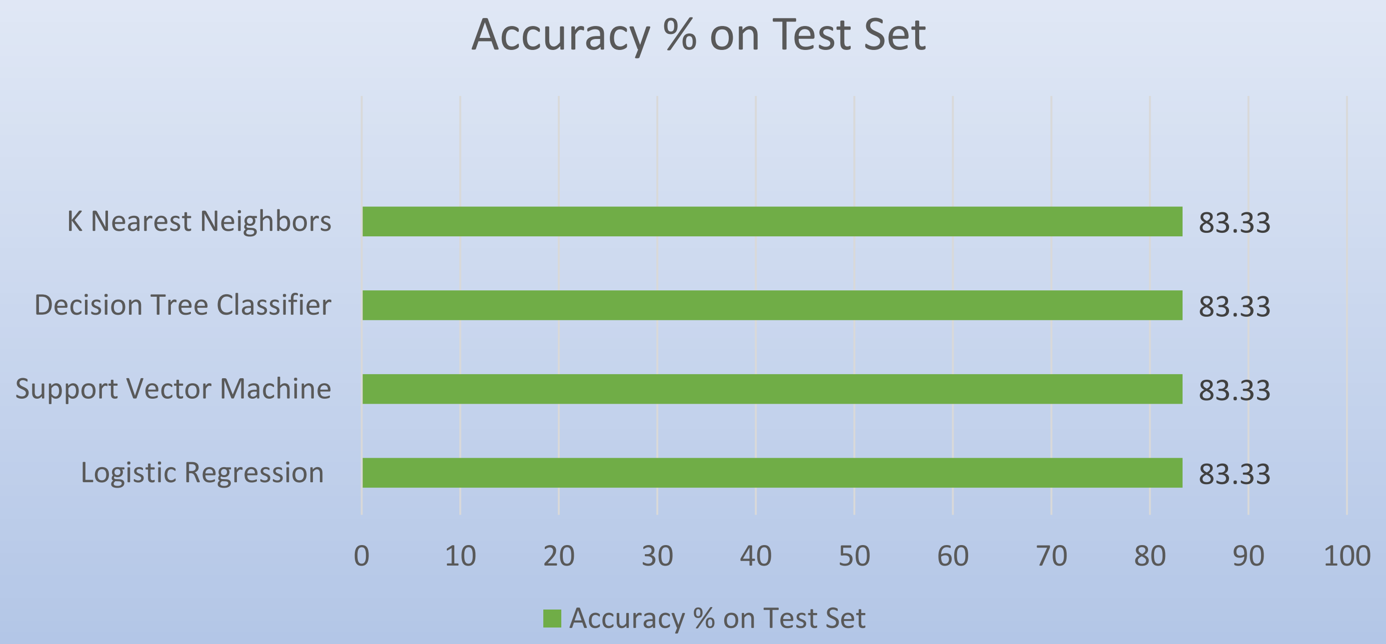


The two charts, which show Payload Mass vs Success for different Booster versions for two different ranges of payload mass, show that there are more successful launches at lower payload masses than at higher payload masses

Section 5

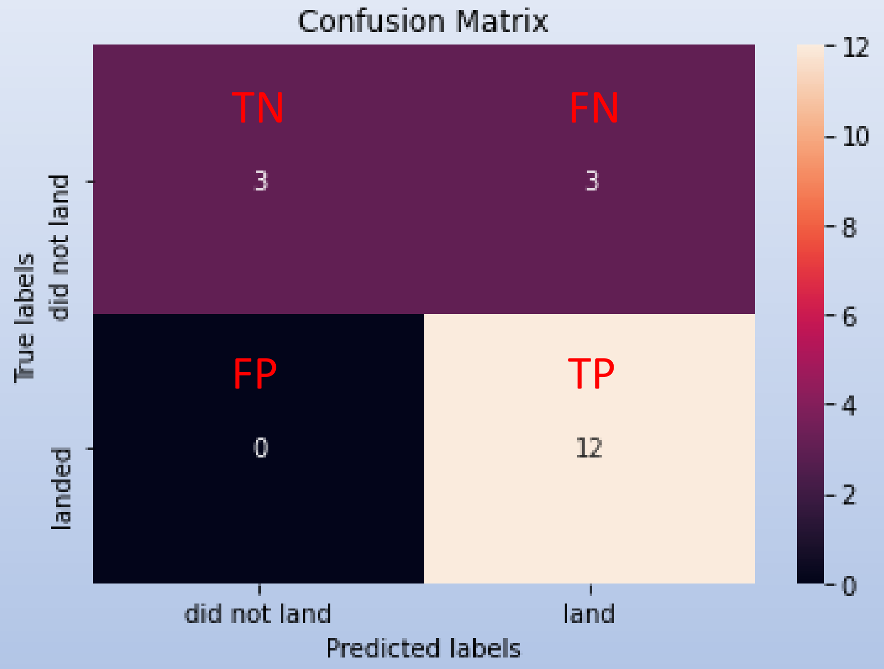
Predictive Analysis (Classification)

Classification Accuracy



The analysis of the success for the four different machine learning models shows that all four had the same level of success with their respective optimal parameters. The accuracy score is based on the model's accuracy when tested with the test data (not training data), which best represents new data that would be tested with the model. Any of these models could be chosen for the final model.

Confusion Matrix



All four models had the same confusion matrix. Out of 18 samples there were 12 true positive results, 3 true negative results, 3 false negative results, and 0 false positive results

Conclusions

- We have four machine learning models to choose from because they all produce the same results
  - Logistic Regression
  - Support Vector Machine
  - Decision Tree Classifier
  - K Nearest Neighbor
- There is room for improvement to refine one of these models or explore other models to reduce false negative results and increase the model accuracy to above 83.33%
- As the amount of flights at a launch site increases, the chance for success also increases
- The orbits with the highest success rates are ES-L1, GEO, HEO, SSO, and VLEO
- From 2013 to 2020 that success rate has steadily increased
- KSC LC-39A is the most successful launch site

Appendix

[Link to GitHub Repository which includes all relevant files for this project]

Thank you!