Winning Space Race with Data Science

Mateo Cherasco
January 3, 2023

Table of contents

Executive Summary

- All over this document, Space Y is going to determine, using machine learning algorithms, wether the Falcon 9 first stage will land succesfully.
- Estimating and taking into consideration variables like Launch Site, Payload Mass, Booster Version, time, orbit and many others.
- Space X statements declarate services of rocket launching starting at 62$ million, while others providers escalate to over 165 million. By determining if the Falcon 9 first stage will land, we can determine the cost of launch.
- The methodology is based on data collection, wrangling prepocessing and further data analysis and data science approaches.
- As result, we were able to predict the first stage rocket booster will land successfully with an accuracy percentage of 83.3%.

Introduction

- When rockets are reused, Space X estimated that the average cost for a launch is 62 million dollars.
- However, the first stage not always has succesfully landings and other times is sacrifieced by Space X due to certain misión parameters like payload, orbit and customer.
- By predicting the likelihood of succes of the first stage landing, we can save millions of dollars in costs of reconstruction, which escalates up to 15 million dolars for each Falcon 9.

Falcon 9 landing in a ocean plataform.

Section 1

Methodology

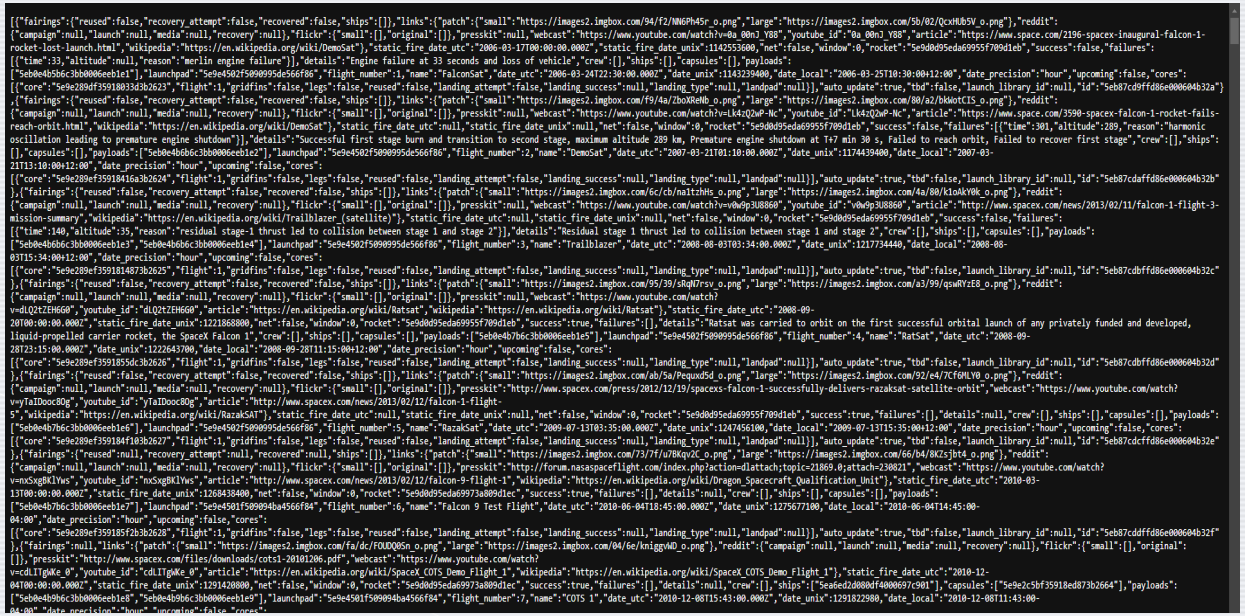
Methodology

The data science methodology applied goes allong with the following:

- Data collection
- Data wrangling
- Exploratory data analysis (EDA) with SQL
- Data visualization (visual analytics using Folium and Plotly Dash)
- Model development
- Share and report outcomes to stakeholders

Data Collection – SpaceX API

1. Request and parse the launch data using the “get” request from the Open Source REST API that Space X provides.
2. Normalize JSON response into a dataframe.
3. Filter the columns in the dataframe to include only the Falcon 9 launches.
4. Deal with missing values (replace the missing payload mass values with the mean)
5. Export to .CSV file.

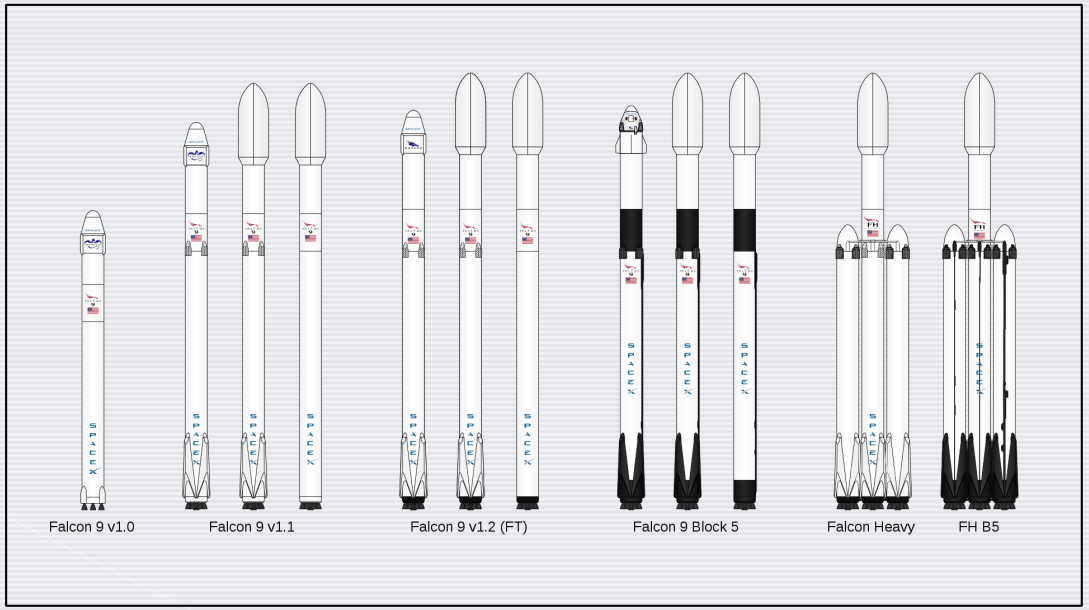

[{"fairings":{"reused":false,"recovery_attempt":false,"recovered":false,"ships":[]},"links":{"patch":{"small":"https://images2.imgbox.com/94/f2/NN6Ph45r_o.png","large":"https://images2.imgbox.com/5b/02/QcxHUb5V_o.png"},"reddit":
{"campaign":null,"launch":null,"media":null,"recovery":null},"flickr":{"small":[],"original":[]},"presskit":null,"webcast":"https://www.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/f9/4a/ZboXReNb_o.png","large":"https://images2.imgbox.com/80/a2/bkWotCIS_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/6c/cb/na1tzhHs_o.png","large":"https://images2.imgbox.com/4a/80/k1oAkY0k_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":
[{"core":"5e9e289ef3591814873b2625","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":"5eb87cdbffd86e000604b32c"
},{"fairings":{"reused":false,"recovery_attempt":false,"recovered":false,"ships":[]},"links":{"patch":{"small":"https://images2.imgbox.com/95/39/sRqN7rsv_o.png","large":"https://images2.imgbox.com/a3/99/qswNYzE8_o.png"},"reddit":
{"campaign":null,"launch":null,"media":null,"recovery":null},"flickr":{"small":[],"original":[]},"presskit":null,"webcast":"https://www.youtube.com/watch?
v=dLQ2tZEH6G0","youtube_id":"dLQ2tZEH6G0","article":"https://en.wikipedia.org/wiki/Ratsat","wikipedia":"https://en.wikipedia.org/wiki/Ratsat"},"static_fire_date_utc":"2008-09-
20T00:00:00.000Z","static_fire_date_unix":1221868800,"net":false,"window":0,"rocket":"5e9d0d95eda69955f709d1eb","success":true,"failures":[],"details":"Ratsat was carried to orbit on the first successful orbital launch of any privately funded and developed,
liquid-propelled carrier rocket, the SpaceX Falcon 1","crew":[],"ships":[],"capsules":[],"payloads":["5eb0e4b7b6c3bb0006eeb1e5"],"launchpad":"5e9e4502f5090995de566f86","flight_number":4,"name":"RatSat","date_utc":"2008-09-
28T23:15:00.000Z","date_unix":1222643700,"date_local":"2008-09-28T11:15:00+12:00","date_precision":"hour","upcoming":false,"cores":
[{"core":"5e9e289ef3591855dc3b2626","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":"5eb87cdbffd86e000604b32d"
},{"fairings":{"reused":false,"recovery_attempt":false,"recovered":false,"ships":[]},"links":{"patch":{"small":"https://images2.imgbox.com/ab/5a/Pequxd5d_o.png","large":"https://images2.imgbox.com/92/e4/7CF6MLY0_o.png"},"reddit":
{"campaign":null,"launch":null,"media":null,"recovery":null},"flickr":{"small":[],"original":[]},"presskit":"http://www.spacex.com/press/2012/12/19/spacexs-falcon-1-successfully-delivers-razaksat-satellite-orbit","webcast":"https://www.youtube.com/watch?
v=yTaIDooc80g","youtube_id":"yTaIDooc80g","article":"http://www.spacex.com/news/2013/02/12/falcon-1-flight-
5","wikipedia":"https://en.wikipedia.org/wiki/RazakSAT"},"static_fire_date_utc":null,"static_fire_date_unix":null,"net":false,"window":0,"rocket":"5e9d0d95eda69955f709d1eb","success":true,"failures":[],"details":null,"crew":[],"ships":[],"capsules":[],"payloads":
["5eb0e4b7b6c3bb0006eeb1e6"],"launchpad":"5e9e4502f5090995de566f86","flight_number":5,"name":"RazakSat","date_utc":"2009-07-13T03:35:00.000Z","date_unix":1247456100,"date_local":"2009-07-13T15:35:00+12:00","date_precision":"hour","upcoming":false,"cores":
[{"core":"5e9e289ef359184f103b2627","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":"5eb87cdcffd86e000604b32e"
},{"fairings":{"reused":null,"recovery_attempt":null,"recovered":null,"ships":[]},"links":{"patch":{"small":"https://images2.imgbox.com/73/7f/u78KQv2C_o.png","large":"https://images2.imgbox.com/66/b4/8KZsjbt4_o.png"},"reddit":
{"campaign":null,"launch":null,"media":null,"recovery":null},"flickr":{"small":[],"original":[]},"presskit":"http://forum.nasaspaceflight.com/index.php?action=dlattach;topic=21869.0;attach=230821","webcast":"https://www.youtube.com/watch?
v=nxSxgBKlYws","youtube_id":"nxSxgBKlYws","article":"http://www.spacex.com/news/2013/02/12/falcon-9-flight-1","wikipedia":"https://en.wikipedia.org/wiki/Dragon_Spacecraft_Qualification_Unit"},"static_fire_date_utc":"2010-03-
13T00:00:00.000Z","static_fire_date_unix":1268438400,"net":false,"window":0,"rocket":"5e9d0d95eda69973a809d1ec","success":true,"failures":[],"details":null,"crew":[],"ships":[],"capsules":[],"payloads":
["5eb0e4b7b6c3bb0006eeb1e7"],"launchpad":"5e9e4501f509094ba4566f84","flight_number":6,"name":"Falcon 9 Test Flight","date_utc":"2010-06-04T18:45:00.000Z","date_unix":1275677100,"date_local":"2010-06-04T14:45:00-
04:00","date_precision":"hour","upcoming":false,"cores":
[{"core":"5e9e289ef35918f2b3b2628","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":"5eb87cddffd86e000604b32f"
},{"fairings":null,"links":{"patch":{"small":"https://images2.imgbox.com/fa/dc/FOUDQ0Sn_o.png","large":"https://images2.imgbox.com/04/6e/knigqvWD_o.png"},"reddit":{"campaign":null,"launch":null,"media":null,"recovery":null},"flickr":{"small":[],"original":
[]},"presskit":"http://www.spacex.com/files/downloads/cots1-20101206.pdf","webcast":"https://www.youtube.com/watch?
v=cdLITgpKe_Q","youtube_id":"cdLITgpKe_Q","article":"https://en.wikipedia.org/wiki/SpaceX_COTS_Demo_Flight_1","wikipedia":"https://en.wikipedia.org/wiki/SpaceX_COTS_Demo_Flight_1"},"static_fire_date_utc":"2010-12-
04T00:00:00.000Z","static_fire_date_unix":1291420800,"net":false,"window":0,"rocket":"5e9d0d95eda6973a809d1ec","success":true,"failures":[],"details":null,"crew":[],"ships":["5ea6ed2d00df4000697c901"],"capsules":["5e9e2c5bf35918ed873b2664"],"payloads":
["5eb0e4b9b6c3bb0006eeb1e8","5eb0e4b9b6c3bb0006eeb1e9"],"launchpad":"5e9e4501f509094ba4566f04","flight_number":7,"name":"COTS 1","date_utc":"2010-12-08T15:43:00.000Z","date_unix":1291822980,"date_local":"2010-12-08T11:43:00-
04:00","date_precision":"hour","upcoming":false,"cores":

REST API shown raw.

Github URL: Data Collection

Data Collection - Scraping

1. Request the Falcon 9 launch information from a Wikipedia table. Transfrom it using BeautifulSoup.
2. Extract all column/variable names from the HTML table header.
3. Create a dataframe with the scrapped information.
4. Export to .CSV file.



Falcon 9 and other heavy launches extracted from the page used in webscraping

Github URL: Data Collection - Web Scrapping

* Class 0: first stage booster did not land successfully
Class 1: first stage booster landed successfully

Data Wrangling

- First, we initialize by calculating the number of launches on each site.
- Calculate the number of occurrence by orbit.
- Calculate the number and occurence of mission outcome per orbit type
- Create a columns called landing outcome

Class 0	Class 1 *
• None None: not attempted • None ASDS: unable to attempt due to launch failure • False ASDS: drone ship landing failed • False Ocean: ocean landing failed • False RTLS: ground pad landing failed	• True ASDS: drone ship landing succed. • True RTLS: ground pad landing succed. • True Ocean: ocean landing succed.

```
CCAFS SLC 40     55
KSC LC 39A       22
VAFB SLC 4E      13
Name: LaunchSite, dtype: int64
```

Launches per site

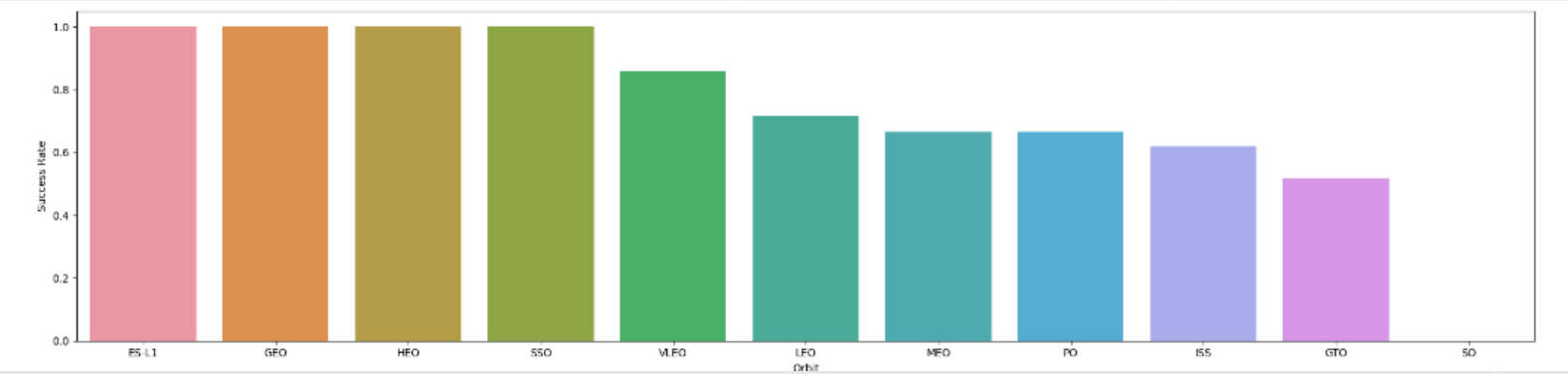
```
True ASDS       41
None None       19
True RTLS       14
False ASDS       6
True Ocean       5
False Ocean      2
None ASDS        2
False RTLS       1
Name: Outcome, dtype: int64
```

Count of outcomes

Github URL: Data Wrangling

EDA with Data Visualization

- Using Matplotlib and Seaborn
- Most of the graffics were used to visualize relationships between features:
  - Flight Number vs. Payload Mass
  - Flight Number vs. Launch Site
  - Payload Mass vs. Launch Site
  - Flight Number vs. Orbit type
  - Payload Mass vs Orbit type
  - Success rate for each orbit type



Success rate for each orbit type

Github URL: Data Visualization

EDA with SQL

- Load into a DB2 instance
- The SQL queries were based on displaying and listing information in regard to the following features:
  - Launch sites
  - Payload mass
  - Customers
  - Booster Version
  - Date

```
%sql select min(Date) as "Date" from SPACEXTBL where "Landing _Outcome" = "Success (ground pad)"
```

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

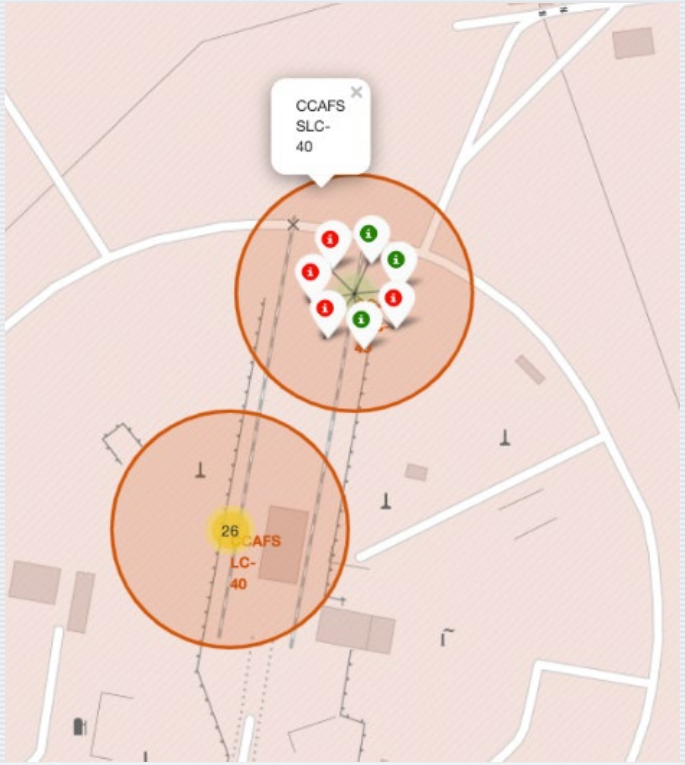
```
%sql select Mission_Outcome, count(Mission_Outcome) as "Mission_Result" from SPACEXTBL group by Mission_Outcome
```

List the total number of successful and failure mission outcomes

Gihub URL: EDA with SQL

Build an Interactive Map with Folium

- Using Folium, the following marks were added to a map:
  - Launch Sites (4)
  - Succesfull (class 1, green) and failed (class 0, red) launches on each site.
- In addition, we added lines representing the distance to its proximities:
  - City
  - Railways
  - Highways
  - Coastlines
- This was calculated with the purpose of determining why are the lauch sites far away or near certain entities.



CCAFS SLC-40 launch site showing "0 class" and "1 class" launches

Github URL: Folium Map Lab

Build a Dashboard with Plotly Dash

- Using Dash Core and HTML Components we got the following interactive elements.
  - A dropdown list to enable the user to select each launch site, also containing the option to review all of them.
  - A pie chart that shows the total successful launches ratio given the dropdown output.
  - A scatter chart that allows the user to visualize the relationship between payload mass and launch success given a certain payload range.
  - A slider to select a payload mass range.

```
#Question 1: The site with the most succesfull launches is KSC LC-39A
#Question 2: The site with the highest success rate is CCAFS SLC-40
#Question 3: The payload range with the most success rate is 3100kg-3700kg (7)
#Question 4: The payload range with the most success rate is 5500kg-9500kg (0)
#Question 5: The F9 Booster Version with the highest succes rate is the FT (14)
```

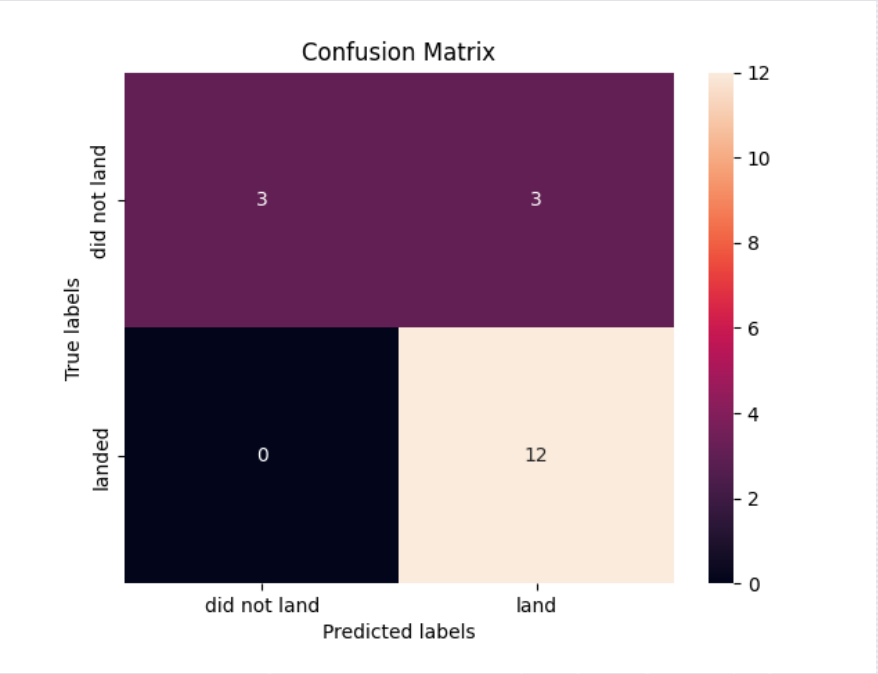
Questions asked in the lab (answers)

Github URL: Dashboard with plotly dash

Github URL: Predictive Analysis

Predictive Analysis (Classification)

- Multiple machine learning algorithms were used to find the best accuracy, but we can conclude in the following steps:
  1. Create and standardize to arrays containing the data that will be used.
  2. Apply the “train_test_split” function.
  3. Create a machine learning model given the algorithm required.
  4. Fit the model.
  5. Get the accuracy with the “score” function.
- Algorithms used:
  - Logistic Regression
  - Support Vector Machine
  - Decision Tree
  - K nearest neighbors



Confussion Matrix of the SVM.

Results

Some EDA outcomes are:

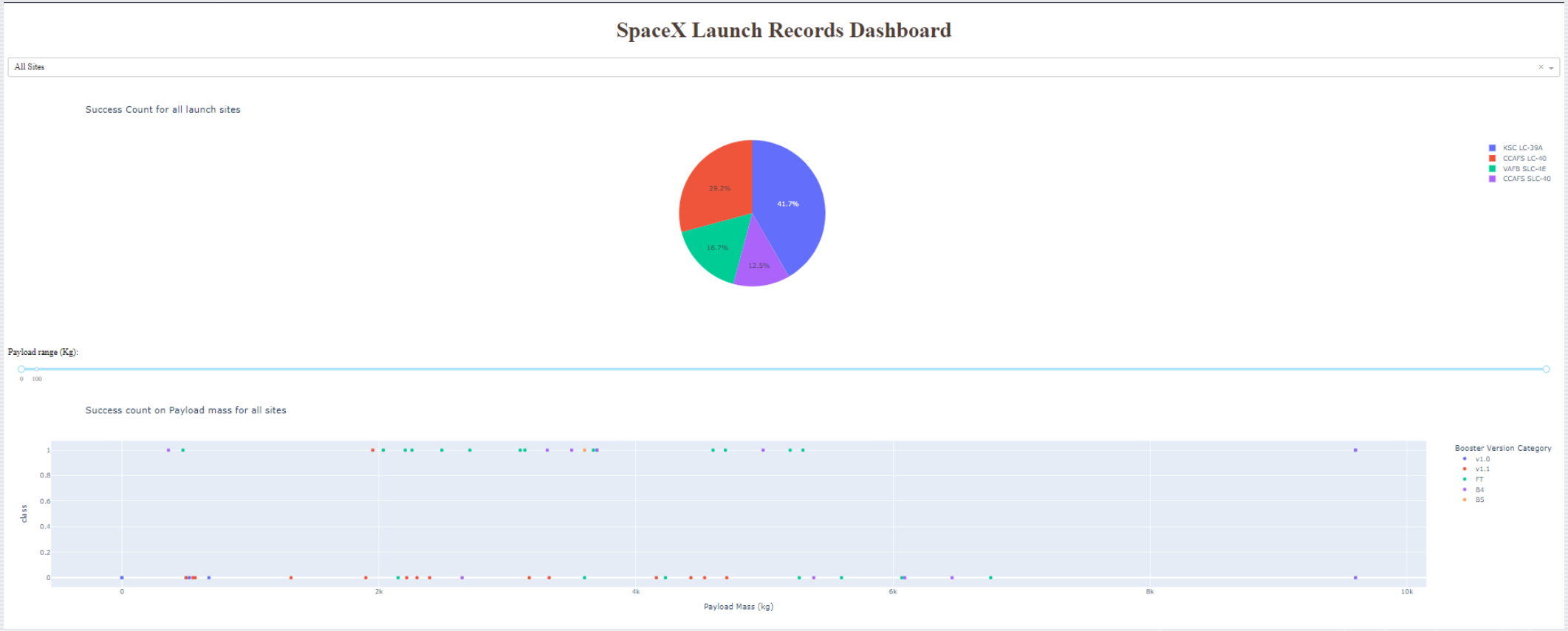
- The site with the most succesfull launches is KSC LC-39A.
- The F9 Booster Version with the highest succes rate is the FT (14)
- The payload range with the most success rate is 3100kg-3700kg (7)
- The launch sites tend to be near the coastline and far away from the cities.
- The sites are located as near as possible to the Equator, because the Earth's rotation is fastest at the equator, providing an extra boost to the rocket.

Some predictive analysis outcomes are:

- The accuracy of every model is 83.33%
- The confussion matrix does a very good job, but it has problems handling false positives.
- The decision tree had a slightly better accuracy when using the “best_score_” function that others algorithms.

Results

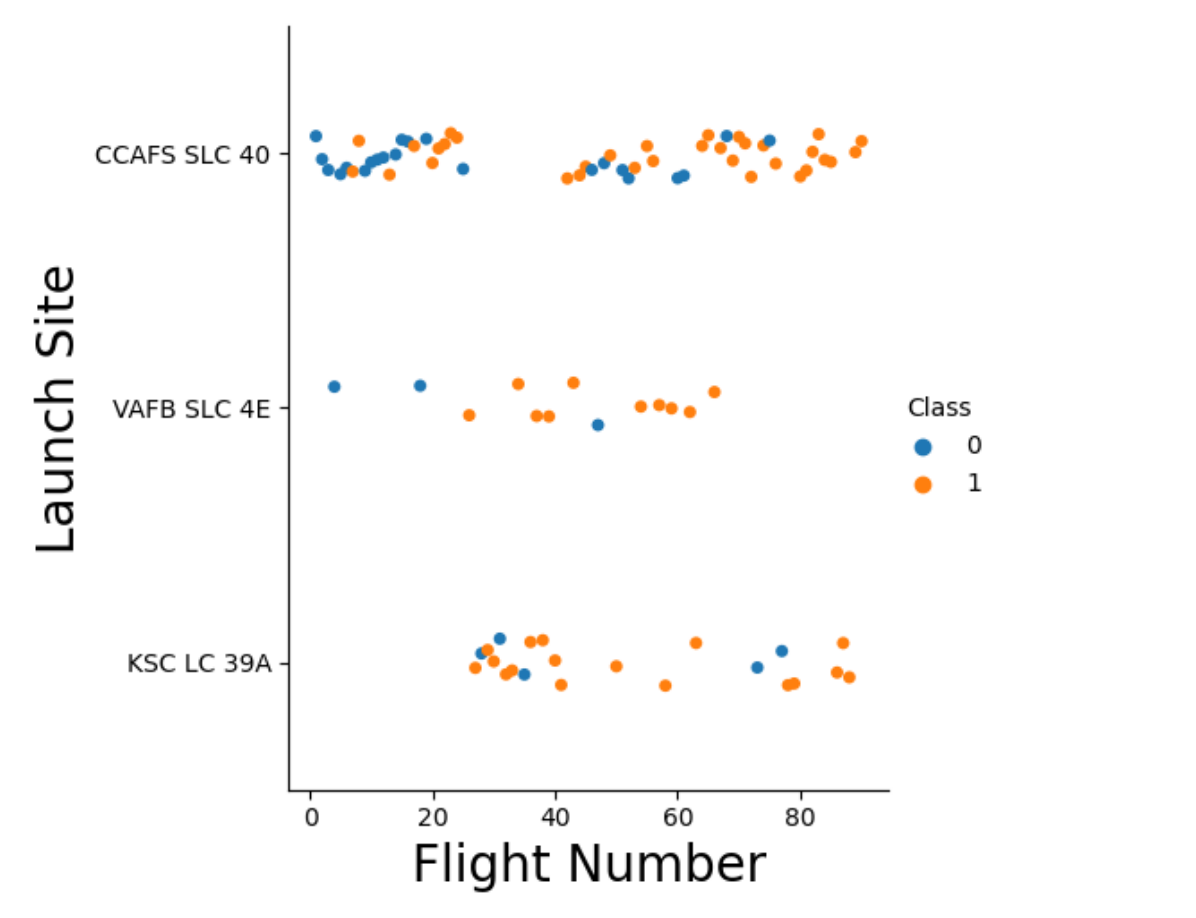
- Plotly dash demo:



Section 2

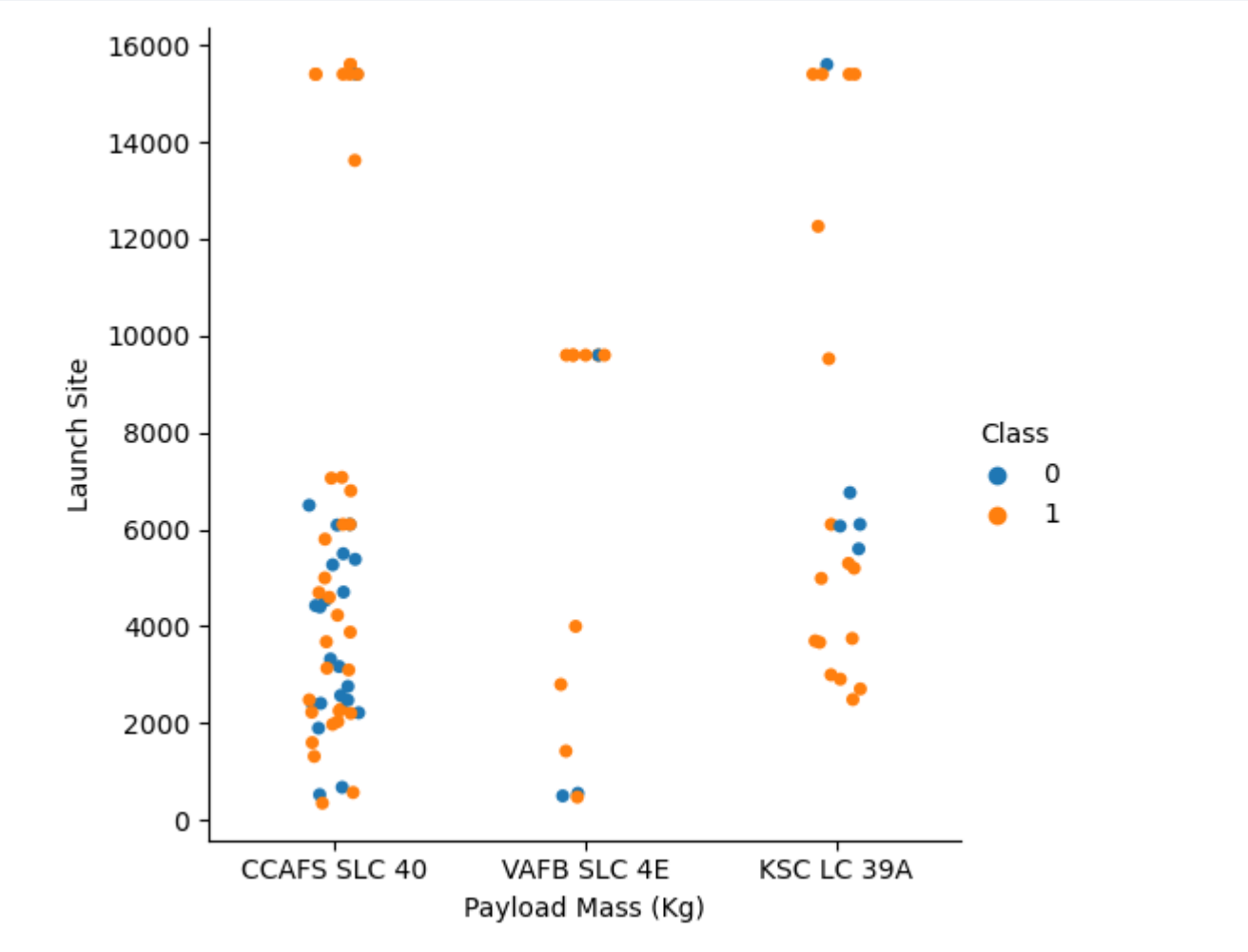
Insights drawn from EDA

Flight Number vs. Launch Site



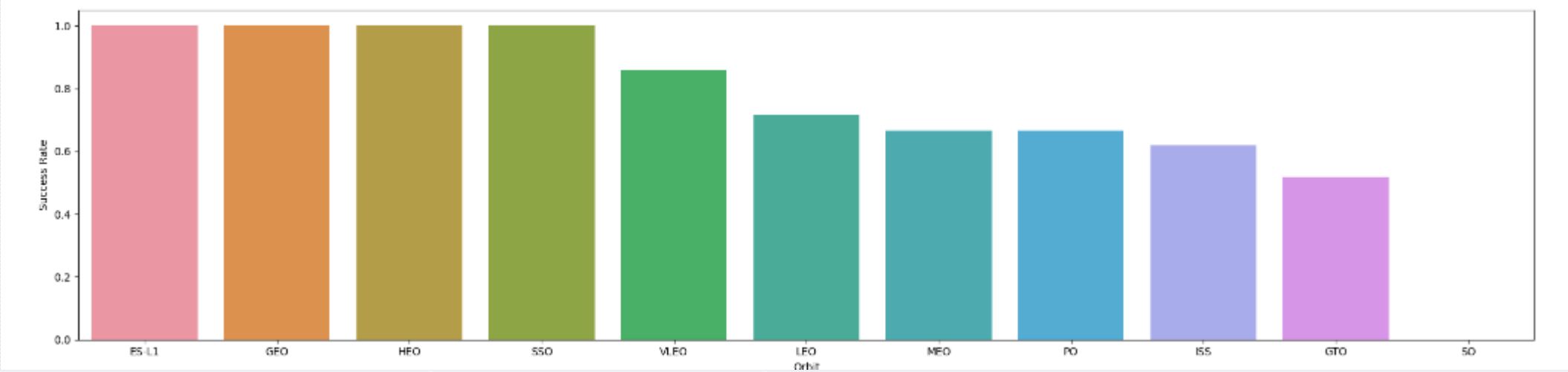
- The graph shows the orange dots as succesful launches and blue dots as failed launches
- We can observe that as the flight number increases, it tends to be more succesful launches
- We can also see that most of the first launches were located in the CCAFS SLC 40 site, this range involves most of the failed launches.

Payload vs. Launch Site



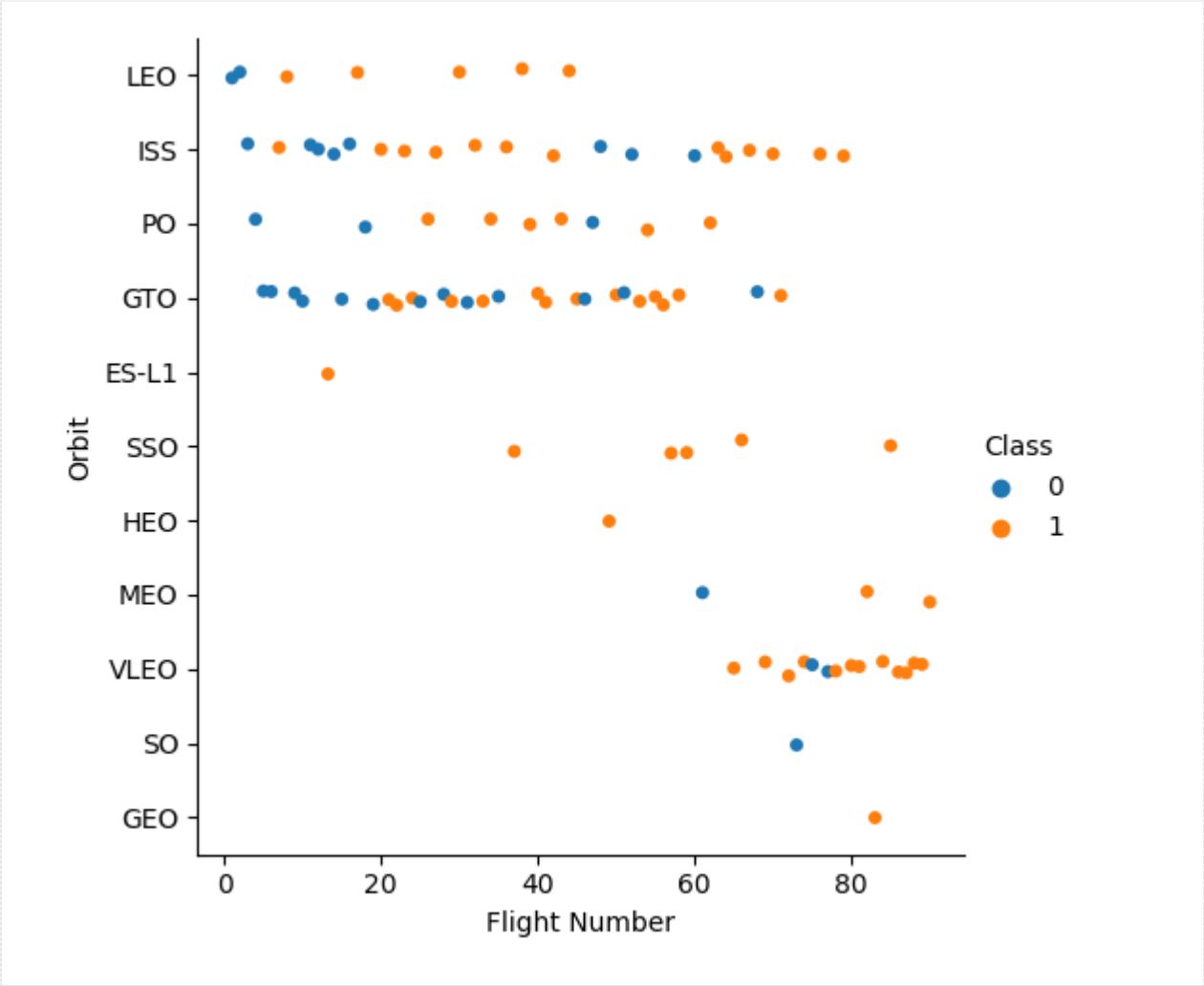
- We can see that most of the launches with low payload mass tend to be deploy in CCADS SLC 40.
- Also, we can visualize that heavier payloads result more effective.

Success Rate vs. Orbit Type



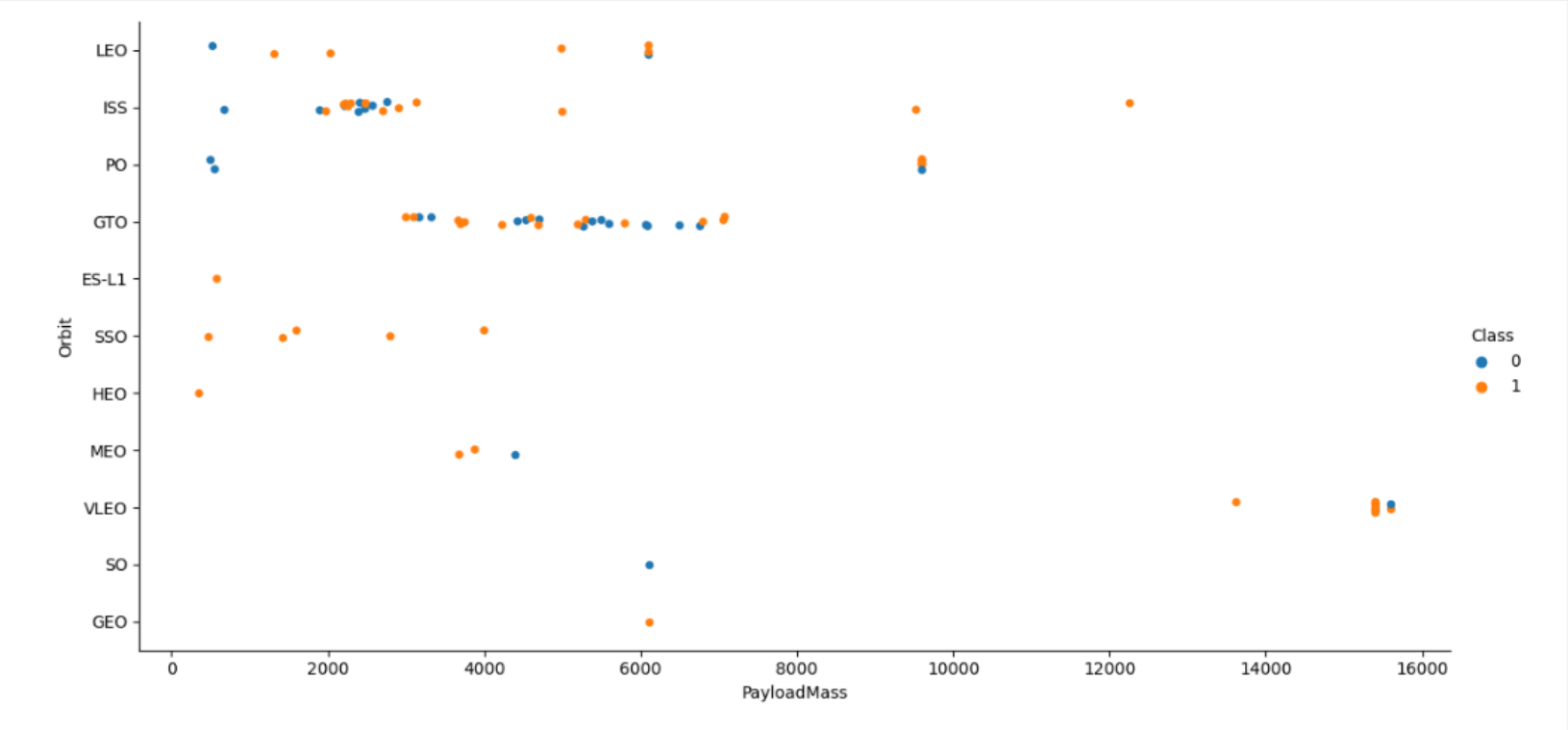
- The orbits ES-L1, GEO, HEO and SSO are the most effective ones.
- All orbits have at least one succesful first stage landing except for SO.
- The orbit with the lower success ratio is GTO. (Except for SO)

Flight Number vs. Orbit Type



- As we can visualize, the most used orbit in the last flight numbers is “VLEO”.
- Most of the failed landings are in the range of 0-40 flight number
- “SSO” has a 100% succes ratio, however, has fewer launches than other orbits.
- “LEO” appears to have a possitive correlation with flight number.

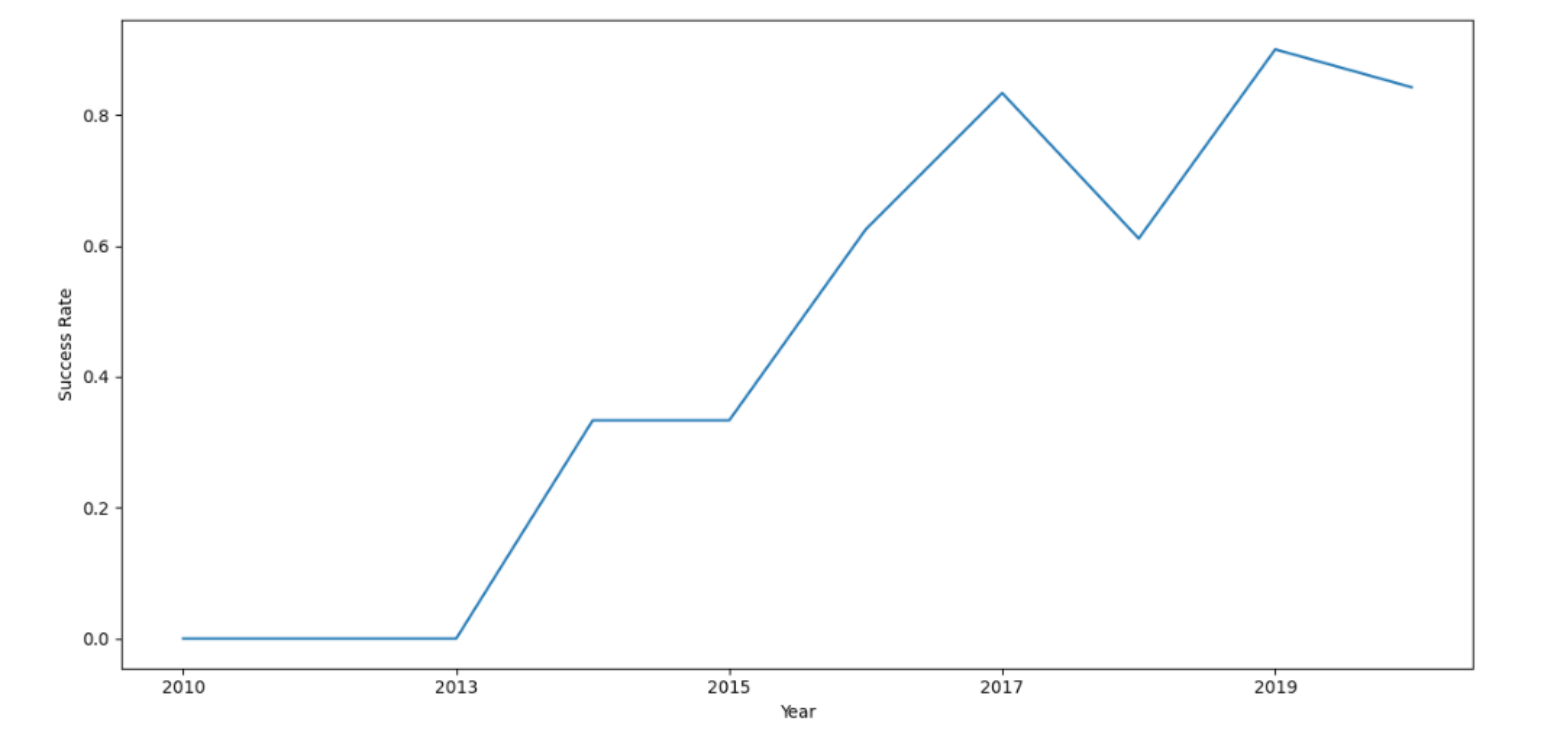
Payload vs. Orbit Type



- Heavier launches are usually deployed in “VLEO”
- Most of the light and medium launches are distributed in “GTO” and “ISS”

- “LEO” and “ISS” appear have a possitive correlation with Payload Mass.

Launch Success Yearly Trend



- We can see an enormous increase in the success rate over the years since 2013.
- However, there is a decrease in 2018.

All Launch Site Names

- Select “DISTINCT” is used for selecting the unique Launch Site.

```
%sql select DISTINCT "Launch_Site" from SPACEXTBL
```

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

Launch Site Names Begin with 'CCA'

- Find 5 records where launch sites begin with `CCA`.

```
%sql select * from SPACEXTBL where "Launch_Site" like "CCA%" limit 5
```

Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer	Mission_Outcome	Landing _Outcome
04-06-2010	18:45:00	F9 v1.0 B0003	CCAFS LC-40	Dragon Spacecraft Qualification Unit	0	LEO	SpaceX	Success	Failure (parachute)
08-12-2010	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)
22-05-2012	07:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525	LEO (ISS)	NASA (COTS)	Success	No attempt
08-10-2012	00:35:00	F9 v1.0 B0006	CCAFS LC-40	SpaceX CRS-1	500	LEO (ISS)	NASA (CRS)	Success	No attempt
01-03-2013	15:10:00	F9 v1.0 B0007	CCAFS LC-40	SpaceX CRS-2	677	LEO (ISS)	NASA (CRS)	Success	No attempt

- We can observe that all the launches come from “CCAFS LC-40”

Total Payload Mass

- Calculate the total payload carried by boosters from NASA

```
%sql select sum(PAYLOAD_MASS__KG_) as "Total_Payload_Kgs", "Customer" from SPACEXTBL where "Customer" = "NASA (CRS)"
```

Total_Payload_Kgs	Customer
45596	NASA (CRS)

- The “sum” function was used to accumulate all the payloads weights that matches with the required customer.

Average Payload Mass by F9 v1.1

- Calculate the average payload mass carried by booster version F9 v1.1

```
%sql select avg(PAYLOAD_MASS__KG_) as "Average_Payload_Mass_Kgs","F9 v1.1" as "Booster_Version" from SPACEXTBL where "Booster_Version" like "F9 v1.1%"
```

Average_Payload_Mass_Kgs	Booster_Version
2534.6666666666665	F9 v1.1

- Booster Versions taken into account:
  - F9 v1.1
  - F9 v1.1 B1011
  - F9 v1.1 B1010
  - F9 v1.1 B1012
  - F9 v1.1 B1013
  - F9 v1.1 B1014
  - F9 v1.1 B1015
  - F9 v1.1 B1016
  - F9 v1.1 B1017
  - F9 v1.1 B1018

First Successful Ground Landing Date

- Find the dates of the first successful landing outcome on ground pad

```
%sql select min(Date) as "Date" from SPACEXTBL where "Landing _Outcome" = "Success (ground pad)"
```

Date
01-05-2017

- “Min” function is used to obtain the first recorded date that fits with “Successful landing on ground pad.”

Successful Drone Ship Landing with Payload between 4000 and 6000

- List the names of boosters which have successfully landed on drone ship and had payload mass greater than 4000 but less than 6000

```
%sql select distinct Booster_Version from SPACEXTBL where "Landing _Outcome" = "Success (drone ship)" and PAYLOAD_MASS__KG_ > 4000 and PAYLOAD_MASS__KG_ < 6000
```

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

- Here we do an arithmetic comparison to find the launches with a payload between the required ranges.

Total Number of Successful and Failure Mission Outcomes

- Calculate the total number of successful and failure mission outcomes

```
%sql select Mission_Outcome, count(Mission_Outcome) as "Mission_Result" from SPACEXTBL group by Mission_Outcome
```

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

- Here we can visualize 99 successful outcomes.

Boosters Carried Maximum Payload

- List the names of the booster which have carried the maximum payload mass

```
%sql select distinct "Booster_Version" from SPACEXTBL where PAYLOAD_MASS__KG_ = (select max(PAYLOAD_MASS__KG_) from SPACEXTBL)
```

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

- We can distinguish the use of a subquery easily by observing both of the select methods applied.
- “Max” function was use with the purpose of selecting the máximum payload mass carried in the registers.

2015 Launch Records

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

```
%sql select *, substr(Date, 4, 2) as "Month", substr(Date,7,4) as "Year" from SPACEXTBL WHERE substr(Date,7,4)='2015' AND "Landing _Outcome" = 'Failure (drone ship)'
```

Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer	Mission_Outcome	Landing _Outcome	Month	Year
10-01-2015	09:47:00	F9 v1.1 B1012	CCAFS LC-40	SpaceX CRS-5	2395	LEO (ISS)	NASA (CRS)	Success	Failure (drone ship)	01	2015
14-04-2015	20:10:00	F9 v1.1 B1015	CCAFS LC-40	SpaceX CRS-6	1898	LEO (ISS)	NASA (CRS)	Success	Failure (drone ship)	04	2015

- Here we can see how we use the “substr” fuction to get the month of the year.
- Both of the launches were directed to the LEO(ISS) Orbit.

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

```
%sql select "Landing _Outcome", count("Landing _Outcome") as "Count" from SPACEXTBL WHERE Date BETWEEN '04-06-2010' AND '20-03-2017' GROUP BY "Landing _Outcome" ORDER BY "Count" DESC
```

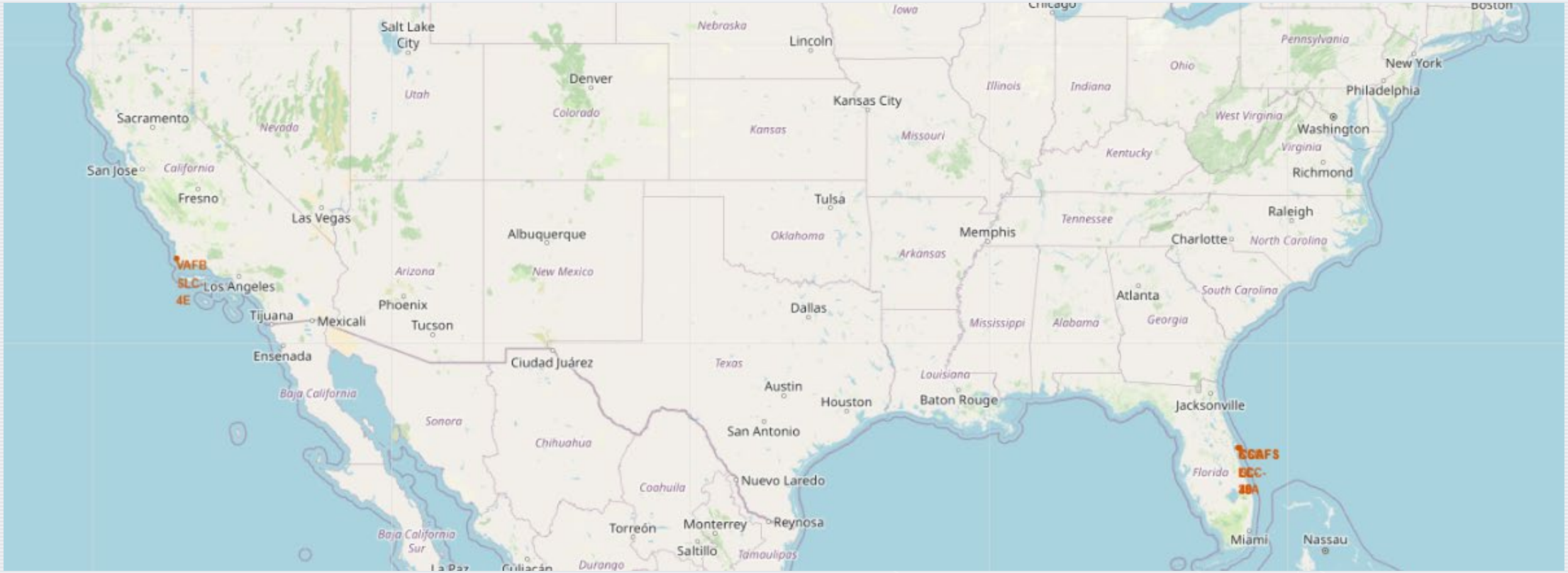
Landing _Outcome	Count
Success	20
No attempt	10
Success (drone ship)	8
Success (ground pad)	6
Failure (drone ship)	4
Failure	3
Controlled (ocean)	3
Failure (parachute)	2
No attempt	1

- Most of the landings ended up in success.
- The most common failure type is the drone ship failure.
- There are more tan 10 not attempted landings
- The coun column is ordered in descending order
- We use the keyword “between” to select a certain range between dates.

Section 3

Launch Sites Proximities Analysis

Location analysis of SPACE X launch sites

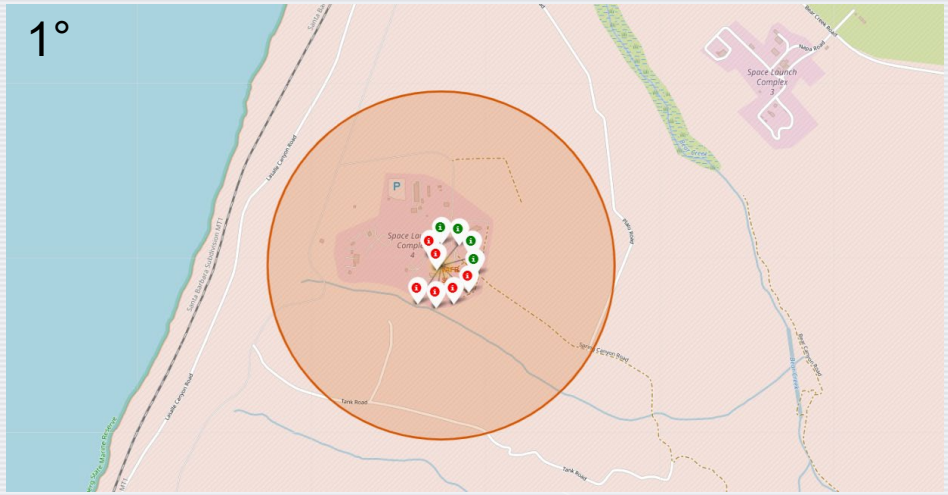


- We can visualize that SPACE X has launch sites both on the east and west coast
- Sites are as near as posible to the Equator

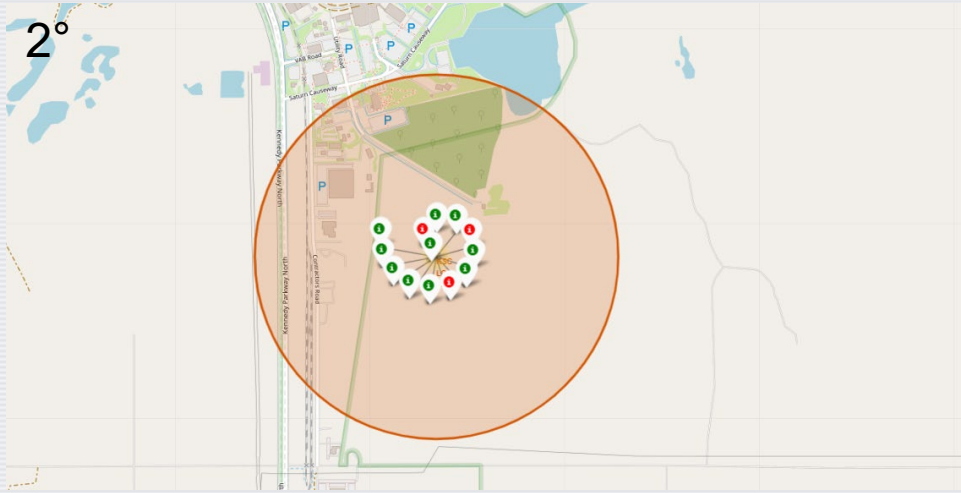
Landing outcomes with color markers

1°- VAFB SLC – 4E
2°- KSC LC - 39A
3°- CCAFS LC - 40
4°- CCAFS SLC - 40

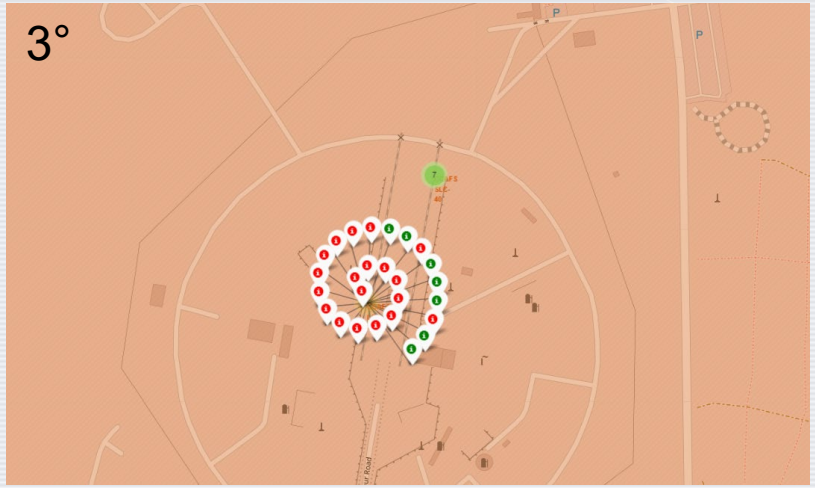
1°



2°



3°



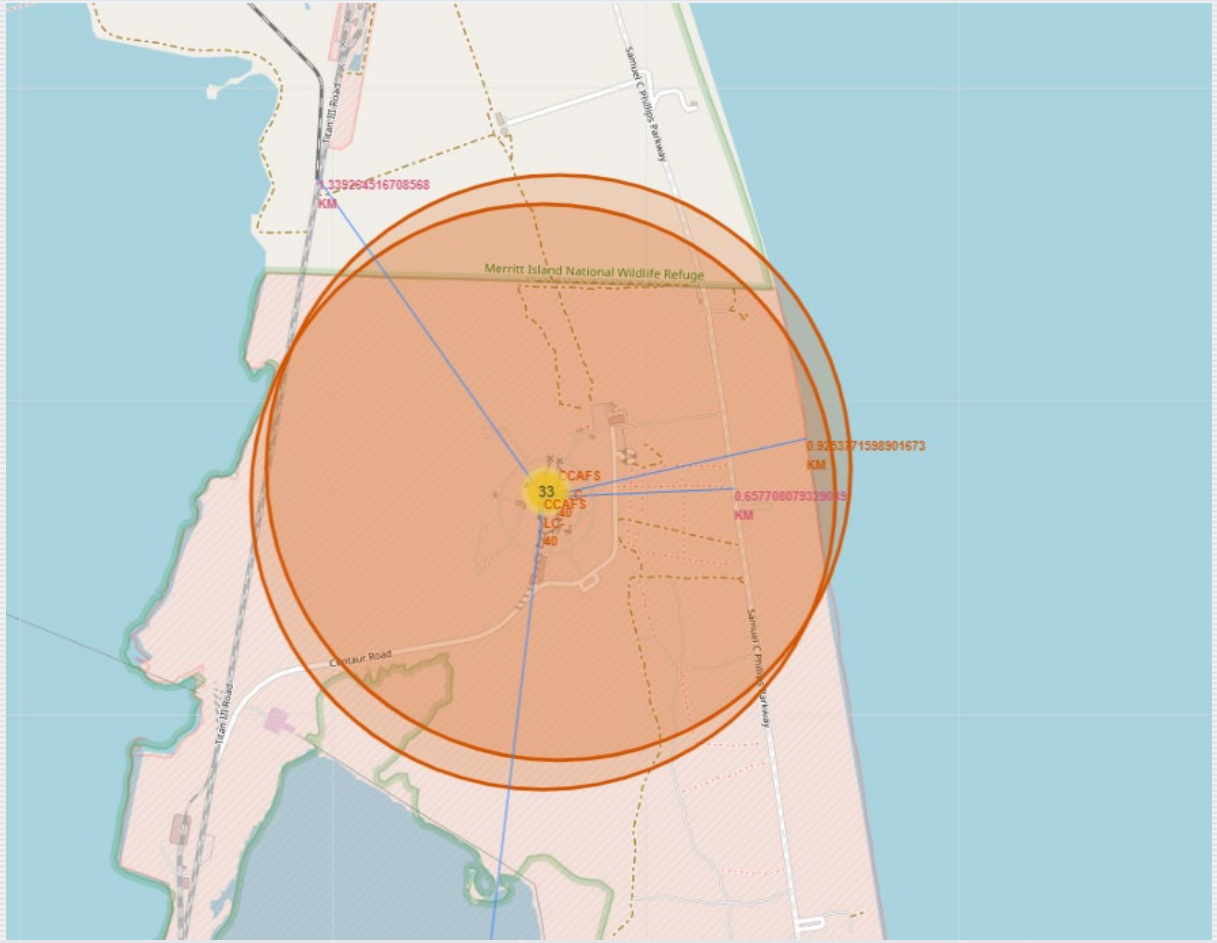
4°

Landing outcomes with color markers

- “Class 0” landings (failed) are marked in red, “class 1” landings (successful) are marked in green.
- We can see that very few landings were sited in CCAFS SLC – 40.
- The launch site with the most activity is CCAFS LC – 40.
- Anyways, CCAFS LC – 40 has not a really good succesful rate.
- The site with the best successful rate is located in Florida, KSC LC – 39A.

Launch sites proximities

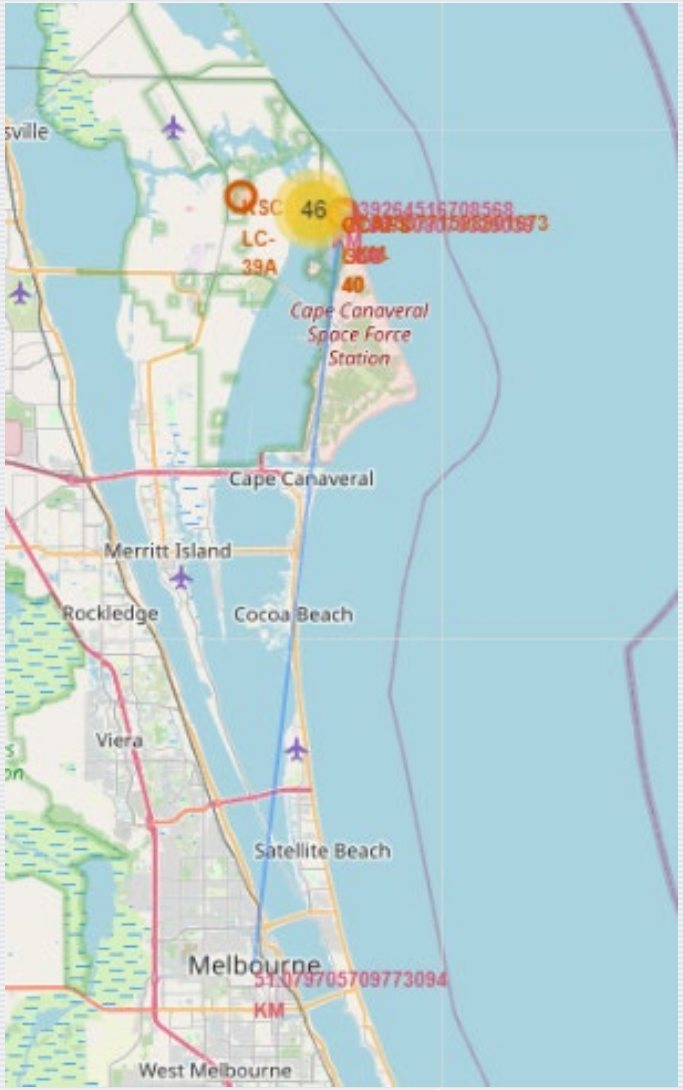
- Here, we can visualize how close is the launch site CCAFS LC – 40 to the coast line. This may be for safety reasons
- Also, the launches sites tend to be close to railways and highways due to the necesity to transport different materials and parts for the mission. Also this helps to transport personal and equipment,
- Distances:
  - Coastline: 0.92 kilometers
  - Highway: 0.65 kilometers
  - Railways: 1.34 kilometers

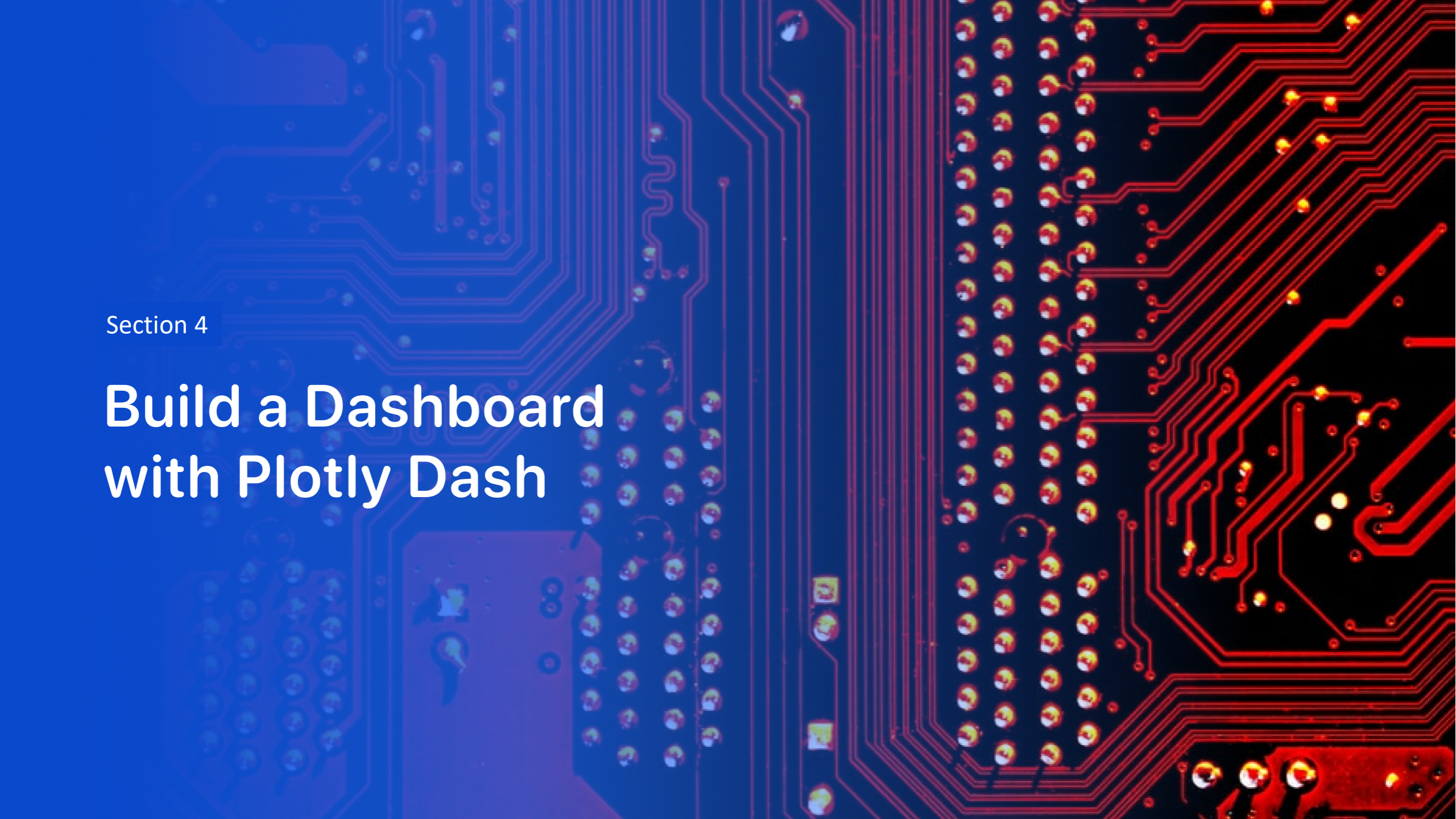


Launch sites proximities

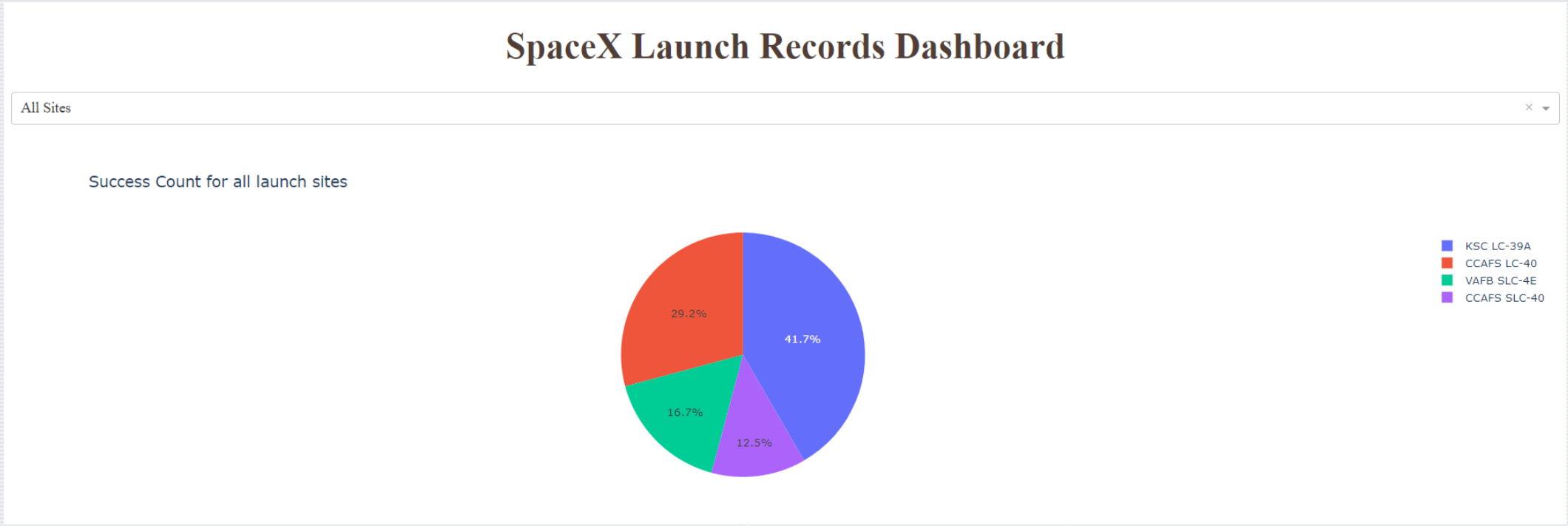
- However, the launch sites are ofter far away from cities, the purpose is to reduce the risk in dense populated areas.

- Distance:
  - City (Melbourne, Florida): 51.08 kilometers.



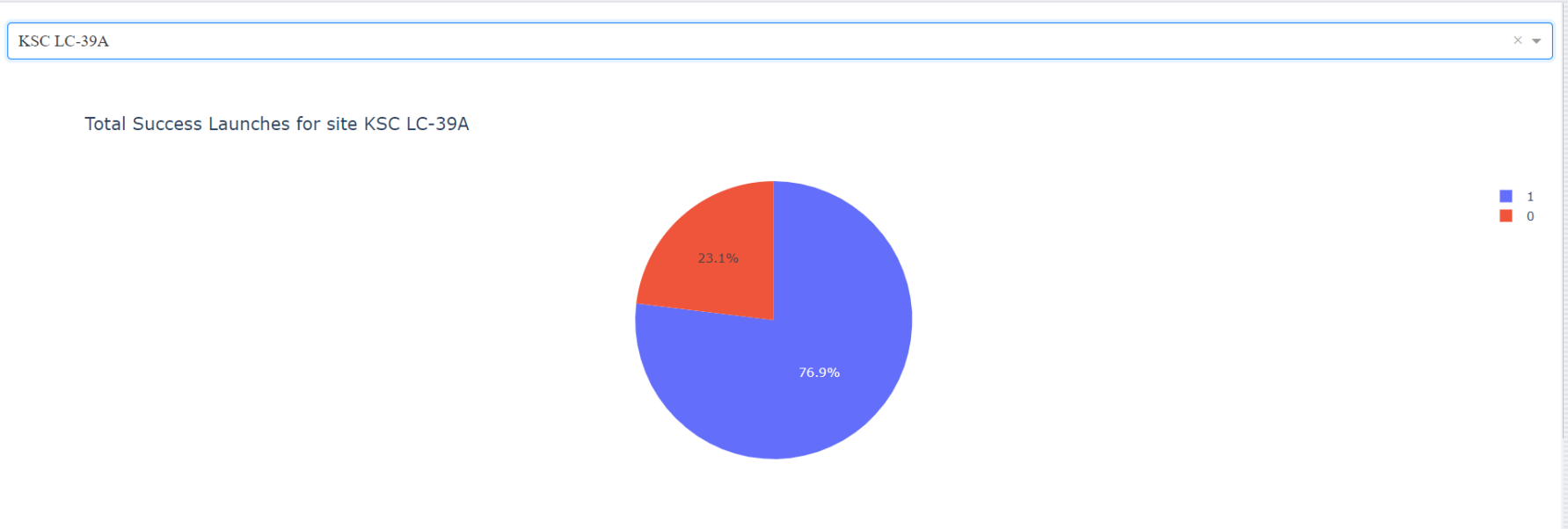

Section 4
Build a Dashboard with Plotly Dash

Pie chart of the count of successful launches



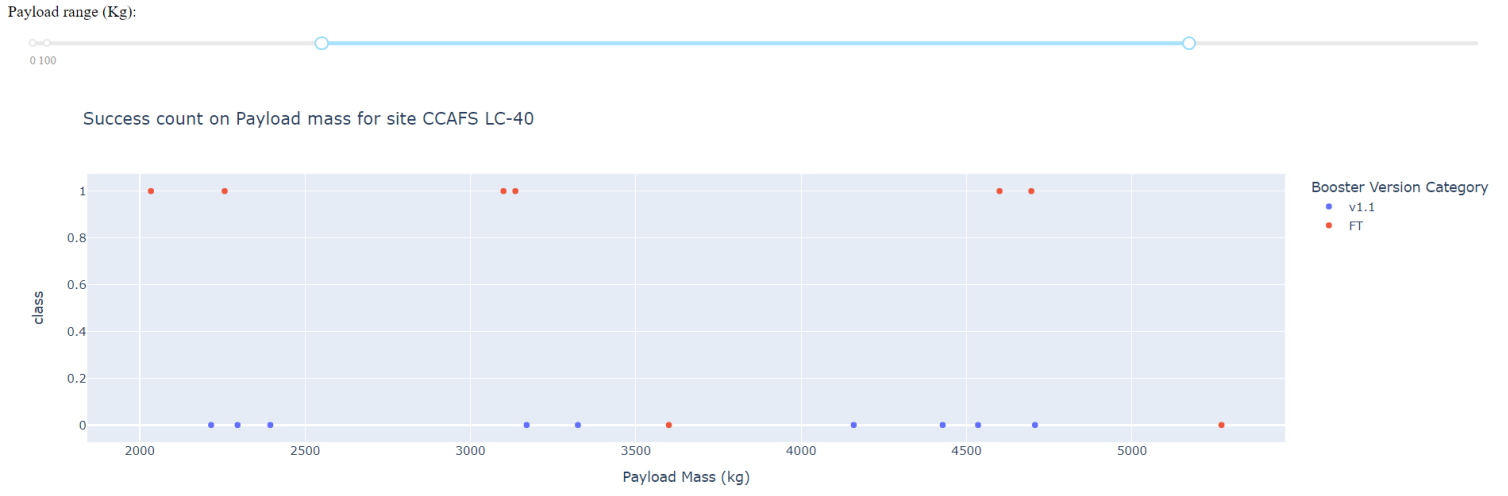
- Here we can visualize the count of successful launches for all sites.
- The site with the most successful launches is KSC LC-39A

Site with the highest succes rate

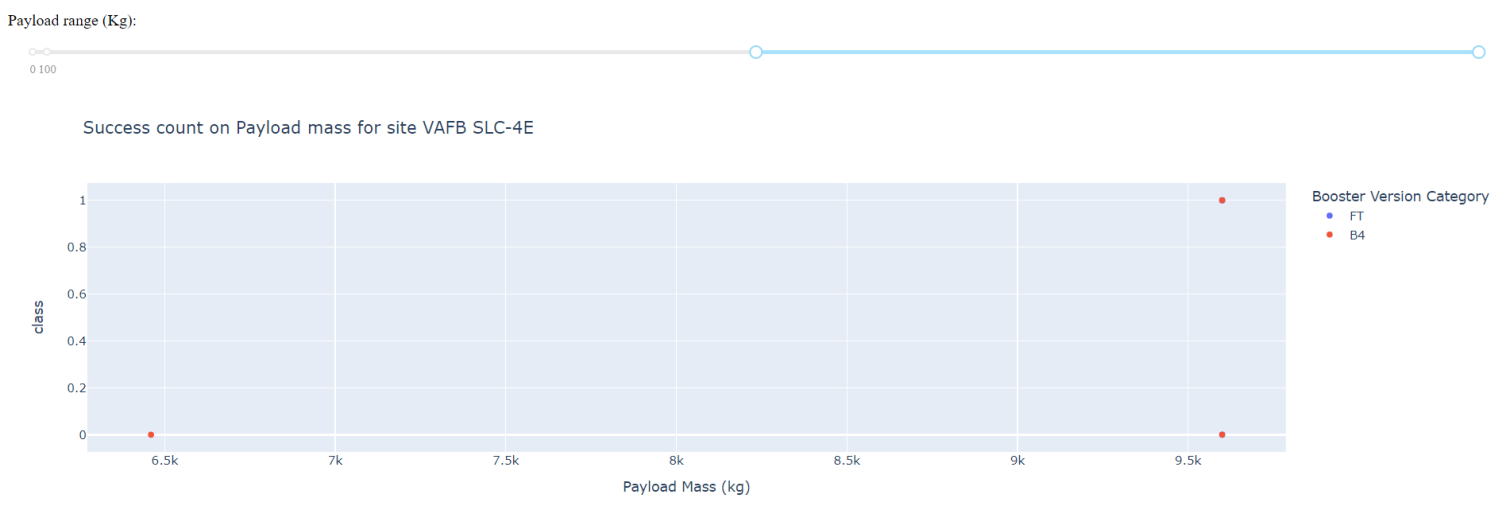


- The site with the highest success ratio is also KSC LC-39A located in Florida.
- This site has 10 successful landings, and 3 failed ones.
- 76.9% of succes rate.

Payload vs. Launch Outcome scatter plot

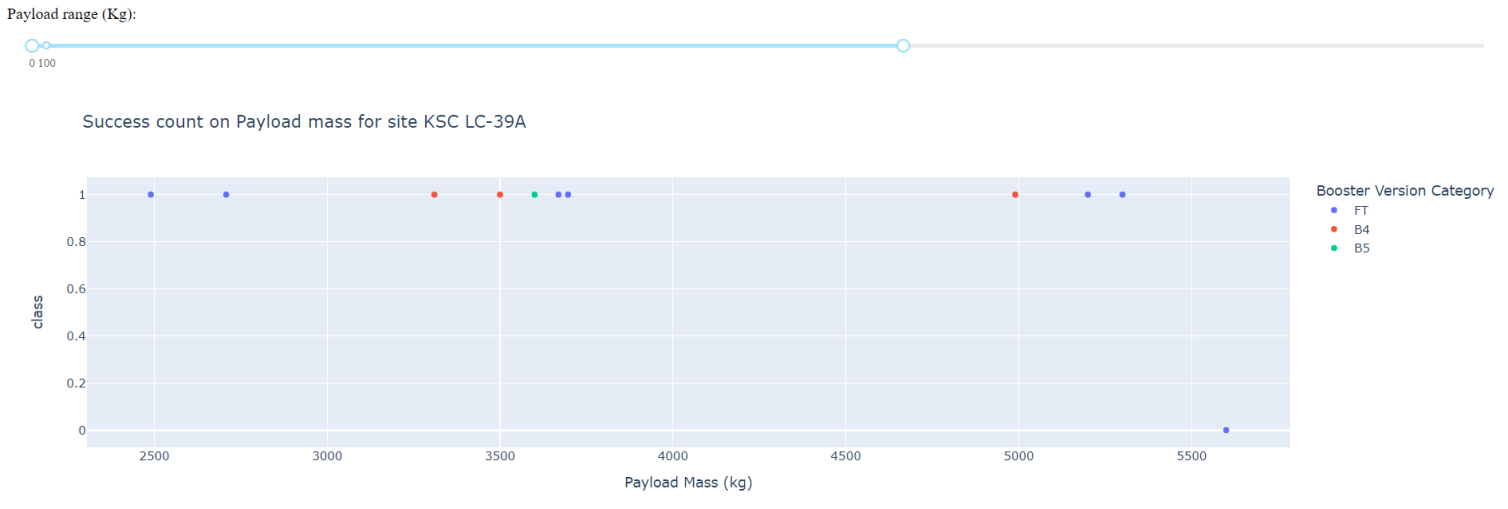


- CCAFS LC-40 is the most used site, we can observe that the most used Boster Versions are v1.1 and FT, also it does not have a really good succes rate

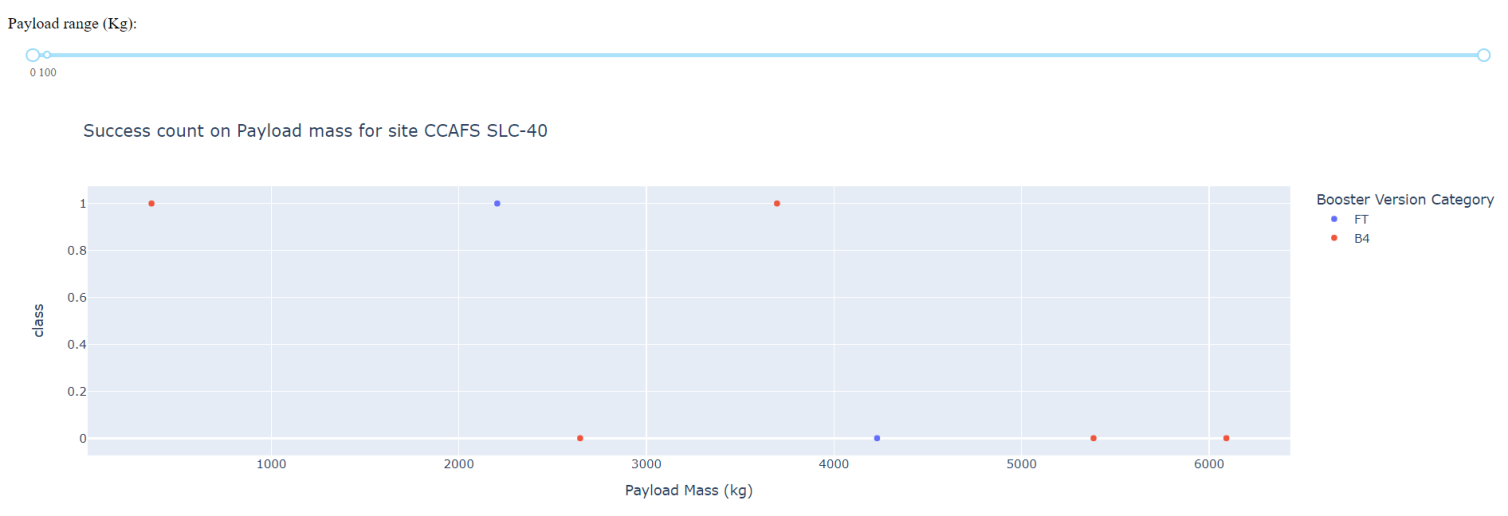


- The VAFB SLC-4E is the only site located in the wast cost, is also where the heaviest payloads are used. We can visualize two “B4” launches that weighted more tan 9.5 thousand kilograms.

Payload vs. Launch Outcome scatter plot



- The KSC LC-39 A is the site with the highest number of succesful launches. Is the only site that uses “B5” Booster Versions.

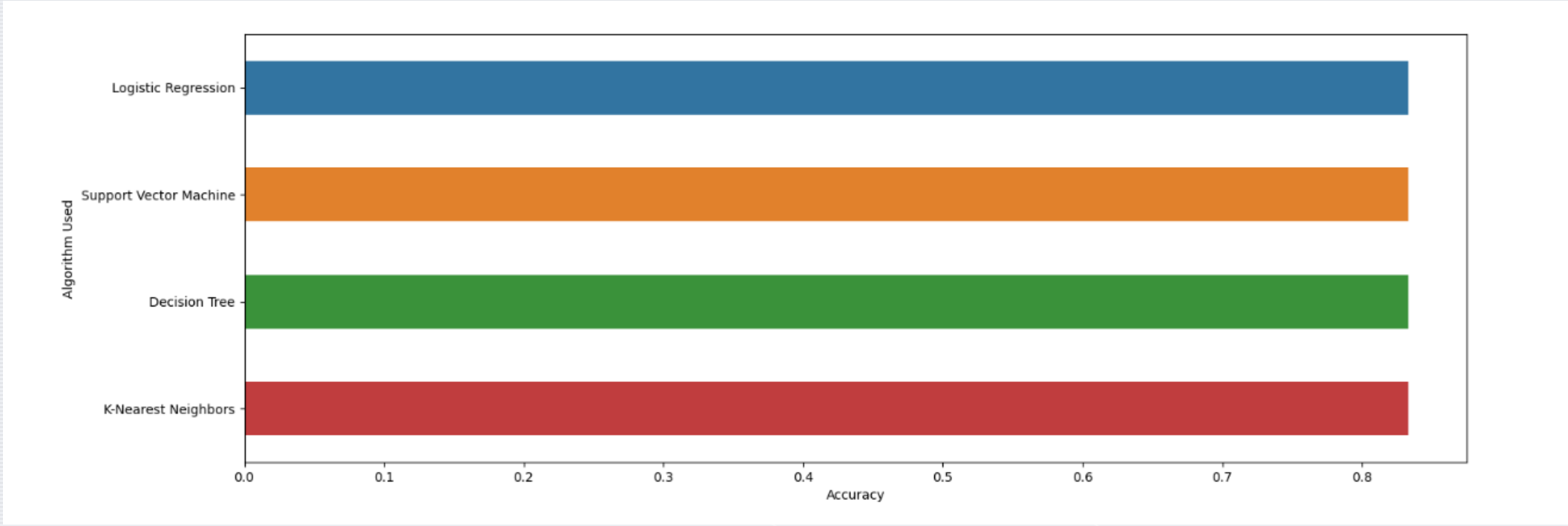


- The CCAFS SLC-40 contains the fewest number of landings of all sites. We can observe only 7 launches in all the payload mass range. It also contains the lightest launch of all, less than a thousand kilograms.

Section 5

Predictive Analysis (Classification)

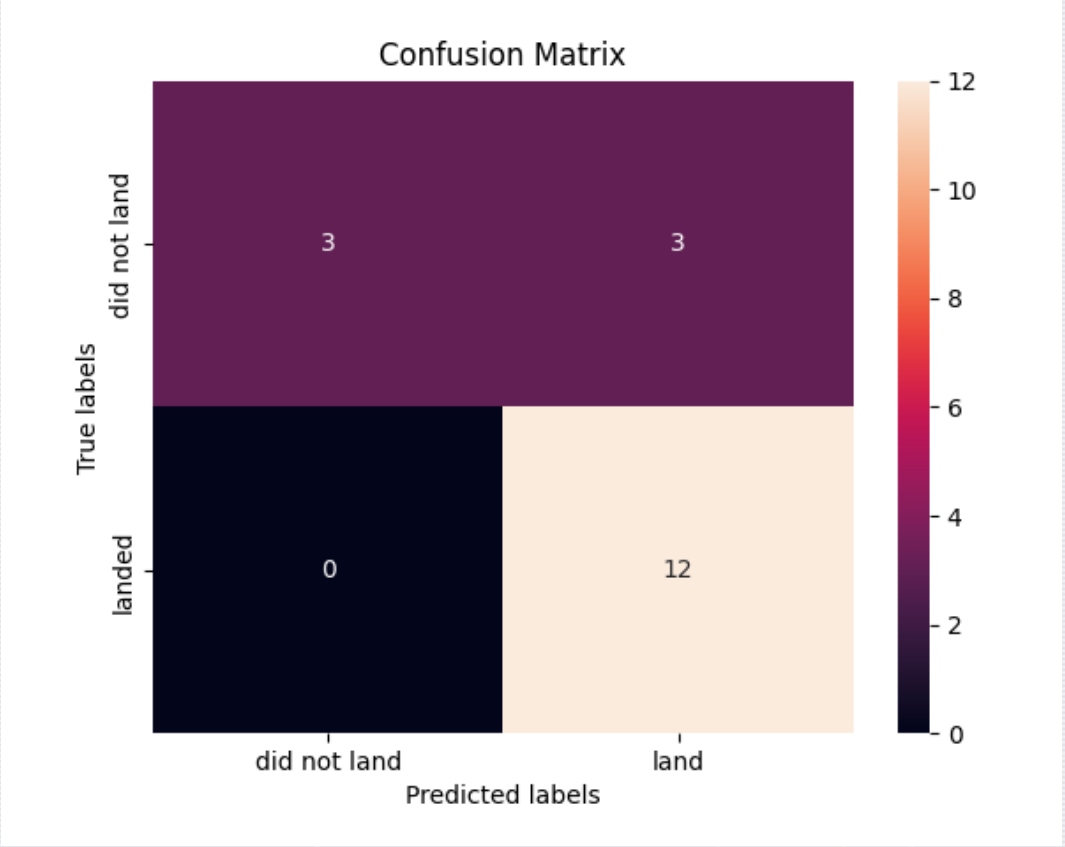
Classification Accuracy



- As we can see, all 4 models had the same accuracy calculated using the “score” function.
- Decision Tree had a slightly difference when we use the “best_score_” function. (0.9 comapred to 0.848 that the other models had)

Confusion Matrix

- The model did a good work predicting 12 true positives (successful landings) and 3 true negatives (failed landings).
- All 4 models have the same confusion matrix, this is due to their tie in accuracy.
- The model predicted 3 successful landings that were a fail. Ending with 3 false positives.



Conclusions

- The biggest problem for the predictive models are the false positives.
- Using any of the 4 algorithms to predict the landing of the Falcon 9 First Stage we will get a accuracy of 83.33% in our result.
- The success rate has improved over the years, so we can say that there is a positive correlation between number of flights and success rate.
- Launches from Florida tend to be more successful, especifically in KSC LC-39A site.
- The most used orbit in the present is VLEO.
- Even adding the cost of the Falcon 9 First Stage (approximately 15 million dollars) to the 62 million that costs a launch, it will still be relatively inexpesive compared to the other competitors. (77 millions vs. 165 millions)

Appendix

- My Github repository: https://github.com/teocherasco/Capstone_Unit
- Course link: https://www.coursera.org/learn/applied-data-science-capstone
- Space X Falcon 9 page: https://www.spacex.com/vehicles/falcon-9/

Thank you!