IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

HAZMAN NAIM BIN AHSAN

24 AUGUST 2023

Outline

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

Executive Summary

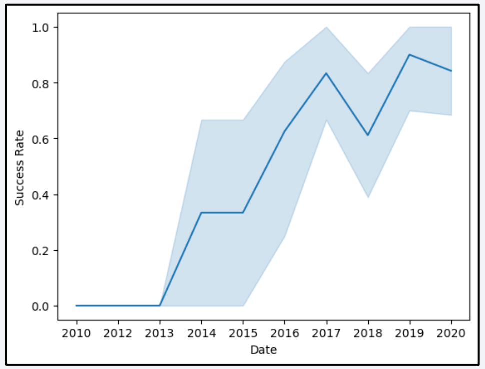
Summary of methodologies

- Data Collection
- Data Wrangling
- Exploratory Data Analysis
- Interactive Visual Analytics
- Predictive Analysis (Classification)

Summary of all results

1. Exploratory Data Analysis (EDA
2. Geospatial Analytics
3. Interactive Dashboard
4. Predictive Analysis of Classification Models

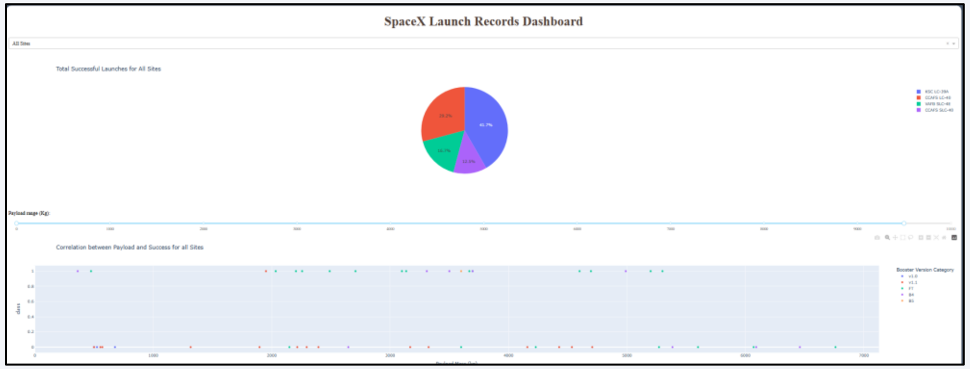
1



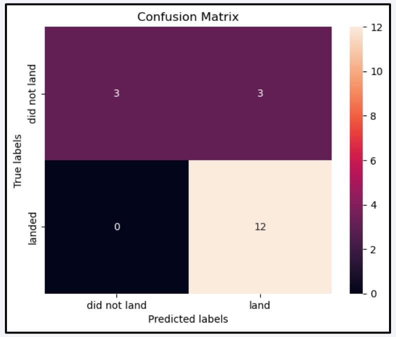
2

VAFB
SLC
4E
CCAFS
SLC-
40A
Sacramento
San Jose
Fresno
Los Angeles
Las Vegas
Salt Lake City
Denver
Albuquerque
Phoenix
Tucson
Dallas
Austin
Houston
San Antonio
Kansas City
Memphis
Atlanta
Chicago
Miami
Monterrey
México

3


SpaceX Launch Records Dashboard
All Sites
Total Successful Launches for All Sites
Payload range (Kg):
Correlation between Payload and Success for all Sites
Booster Version Category
v1.0
v1.1
FT
B4
B5
class
0
1000
2000
3000
4000
5000
6000
7000

4



Introduction

- This project aims to predict the success of Falcon 9 first stage landings, crucial due to SpaceX's cost-saving reusability, enabling accurate launch cost estimation and informed competition.

- Can we predict Falcon 9 landing success accurately considering complex variables, ensuring reliable bid strategy for alternate launch companies?

Section 1

Methodology

Methodology

Executive Summary

1. Data collection methodology

- GET Request to SpaceX API
- Web Scraping

2. Perform data wrangling

- Handle missing values
- Handle duplicates
- Calculate the number of launches on each site
- Calculate the number and occurrence of each orbit
- Calculate the number and occurence of mission outcome per orbit type
- Labelling the landing outcome

3. Perform exploratory data analysis (EDA) using visualization and SQL

- Using SQL queries to examine the dataset
- Using Pandas and Matplotlib to perform EDA and feature engineering

4. Perform interactive visual analytics using Folium and Plotly Dash

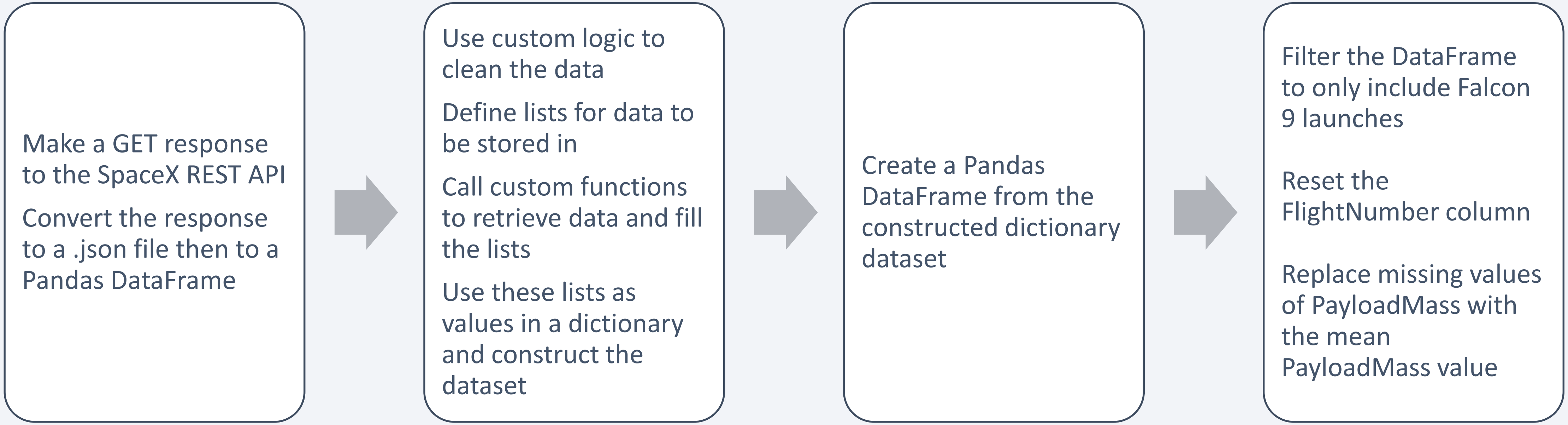
- Geospatial analytics using Folium
- Building dashboard using Plotly Dash

5. Perform predictive analysis using classification model

- Building classification models: Logistic regression, Support vector classifier, Decision tree classifier, KNN
- Calculating accuracy score for each models and visualized with confusion matrix

Data Collection – SpaceX API

Obtaining launch data through the SpaceX API, which includes details about the rocket utilized, the payload dispatched, launch and landing specifications, as well as the outcome of the landing.



Make a GET response to the SpaceX REST API

Convert the response to a .json file then to a Pandas DataFrame

Use custom logic to clean the data

Define lists for data to be stored in

Call custom functions to retrieve data and fill the lists

Use these lists as values in a dictionary and construct the dataset

Create a Pandas DataFrame from the constructed dictionary dataset

Filter the DataFrame to only include Falcon 9 launches

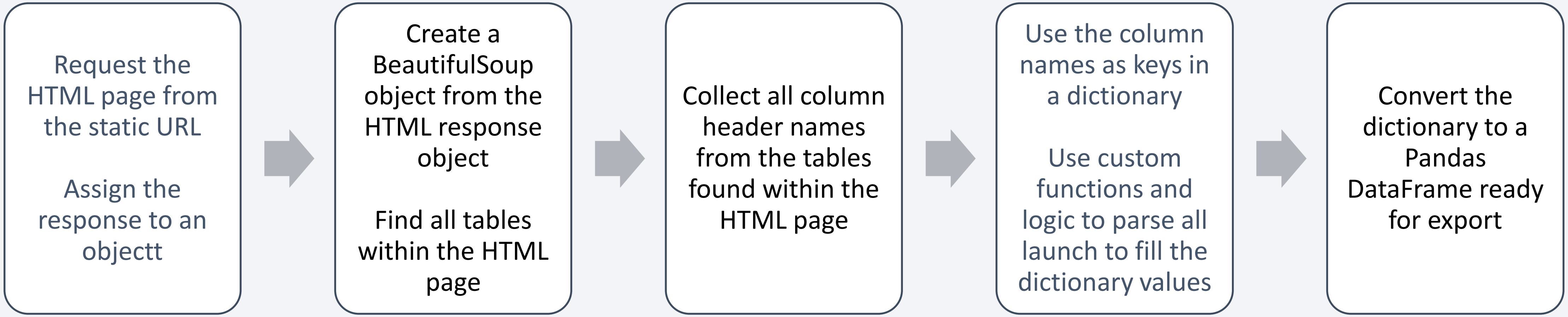
Reset the FlightNumber column

Replace missing values of PayloadMass with the mean PayloadMass value

Github Link

Data Collection – Web Scraping

Gathering historical launch records of Falcon 9 from a Wikipedia page named "List of Falcon 9 and Falcon Heavy launches" through web scraping.



Request the HTML page from the static URL

Assign the response to an objectt

Create a BeautifulSoup object from the HTML response object

Find all tables within the HTML page

Collect all column header names from the tables found within the HTML page

Use the column names as keys in a dictionary

Use custom functions and logic to parse all launch to fill the dictionary values

Convert the dictionary to a Pandas DataFrame ready for export

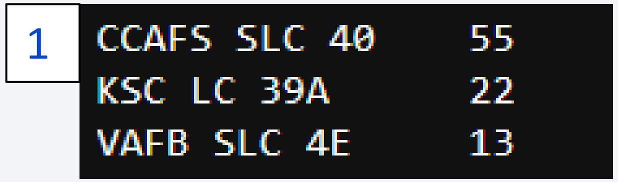
Github Link

Data Wrangling

In this section, the initial exploration and analysis of the dataset aim to provide insights into SpaceX launches. The dataset includes various columns, each offering valuable information about the launches. Let's examine the key insights extracted:

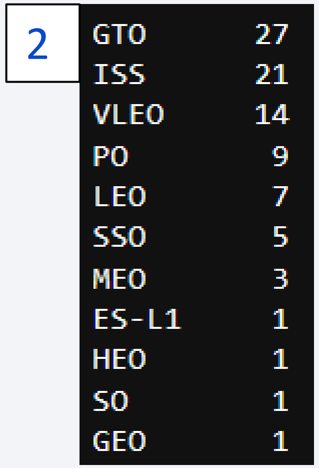
1. **Number of Launches per Site**
2. **Distribution of Launch Orbits**
3. **Mission Outcomes**
4. **Creating Landing Outcome Labels (1=Success, 0=Failed)**

1



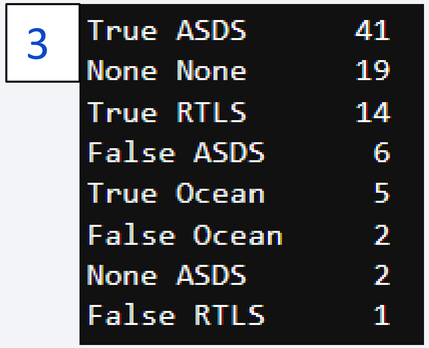
```
CCAFS SLC 40     55
KSC LC 39A       22
VAFB SLC 4E      13
```

2



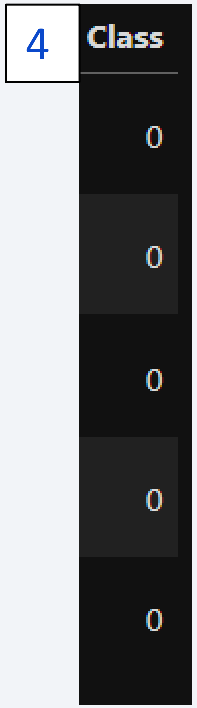
```
GTO        27
ISS        21
VLEO       14
PO          9
LEO         7
SSO         5
MEO         3
ES-L1       1
HEO         1
SO          1
GEO         1
```

3



```
True ASDS      41
None None      19
True RTLS      14
False ASDS      6
True Ocean      5
False Ocean     2
None ASDS       2
False RTLS      1
```

4



```
Class
    0
    0
    0
    0
    0
```

Github Link

EDA with Data Visualization

In the EDA - Visualizations section, several types of charts were utilized to analyze and present the data.

Scatter Chart

- Flight Number and Launch Site
- Payload and Launch Site
- Orbit Type and Flight Number
- Payload and Orbit Type

Scatter charts are useful to observe relationships, or correlations, between two numeric variables.

Bar Chart

- Success Rate and Orbit Type

Bar charts are used to compare a numerical value to a categorical variable. Horizontal or vertical bar charts can be used, depending on the size of the data.

Line Chart

- Success Rate and Year (i.e. the launch success yearly trend)

Line charts contain numerical values on both axes, and are generally used to show the change of a variable over time.

Github Link

EDA with SQL

In the EDA - SQL section, the following analyses were conducted:

1. The names of the unique launch sites in the space mission were displayed.
2. Five records where launch sites began with the string 'CCA' were displayed.
3. The total payload mass carried by boosters launched by NASA (CRS) was calculated and displayed.
4. The average payload mass carried by the booster version F9 v1.1 was computed and displayed.
5. The date when the first successful landing outcome on a ground pad was achieved was listed.
6. The names of the boosters which had success on a drone ship and a payload mass between 4000 and 6000 kg were listed.
7. The total number of successful and failed mission outcomes were calculated and listed.
8. The names of the booster versions which have carried the maximum payload mass were listed.
9. The failed landing outcomes on drone ships, their booster versions, and launch site names for 2015 were listed.
10. The count of landing outcomes (such as Failure (drone ship) or Success (ground pad)) between the date 2010-06-04 and 2017-03-20, were ranked in descending order.

Github Link

Build an Interactive Map with Folium

In the Interactive Map with Folium section, the following steps were executed:

1. **Marking Launch Sites:**
    1. All launch sites were marked on a map.
    2. This was achieved by initializing a map using a Folium Map object.
    3. A *folium.Circle* and *folium.Marker* were added for each launch site on the launch map.
2. **Marking Launch Outcomes:**
    1. The success and failure of launches for each site were marked on a map.
    2. Given that many launches share the same coordinates, they were clustered together for better visualization.
    3. The color of the markers was set to green for successful launches and red for failed ones.
    4. Each launch was added to a *MarkerCluster()* object with an icon as a text label, and the *icon_color* was set based on the marker color determined previously.
3. **Calculating Distances:**
    1. The distances between a launch site and its proximities were calculated using the latitude and longitude values.
    2. A point was marked using these values, and a *folium.Marker* object was created to display the distance.
    3. A *folium.PolyLine* was drawn to display the distance line between two points and added to the map.

Github Link

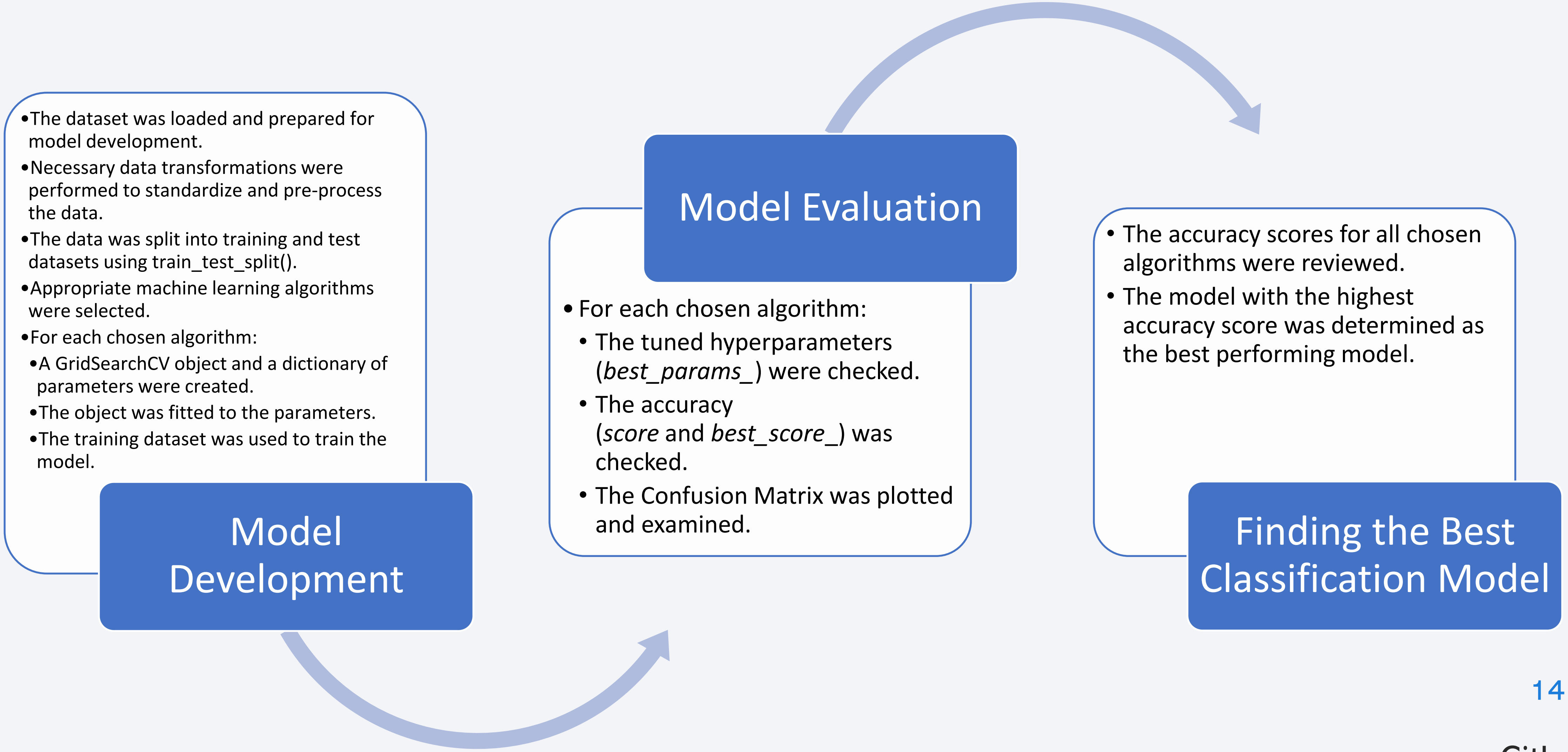
Build a Dashboard with Plotly Dash

Interactive visualizations of the data were created using a Plotly Dash dashboard. The following plots were included:

1. **Pie Chart:**
    1. A pie chart was created using *px.pie()* to display the total successful launches per site.
    2. This visualization made it clear which sites had the most successful launches.
    3. A filter was added using a *dcc.Dropdown()* object to view the success/failure ratio for individual sites.
2. **Scatter Graph:**
    1. A scatter graph was created using *px.scatter()* to show the correlation between outcome (success or not) and payload mass (kg).
    2. A filter was added using a *RangeSlider()* object to view data within specific ranges of payload masses.

Github Link

Predictive Analysis - Classification



Model Development

- The dataset was loaded and prepared for model development.
- Necessary data transformations were performed to standardize and pre-process the data.
- The data was split into training and test datasets using train_test_split().
- Appropriate machine learning algorithms were selected.
- For each chosen algorithm:
  - A GridSearchCV object and a dictionary of parameters were created.
  - The object was fitted to the parameters.
  - The training dataset was used to train the model.

Model Evaluation

- For each chosen algorithm:
  - The tuned hyperparameters (*best_params_*) were checked.
  - The accuracy (*score* and *best_score_*) was checked.
  - The Confusion Matrix was plotted and examined.

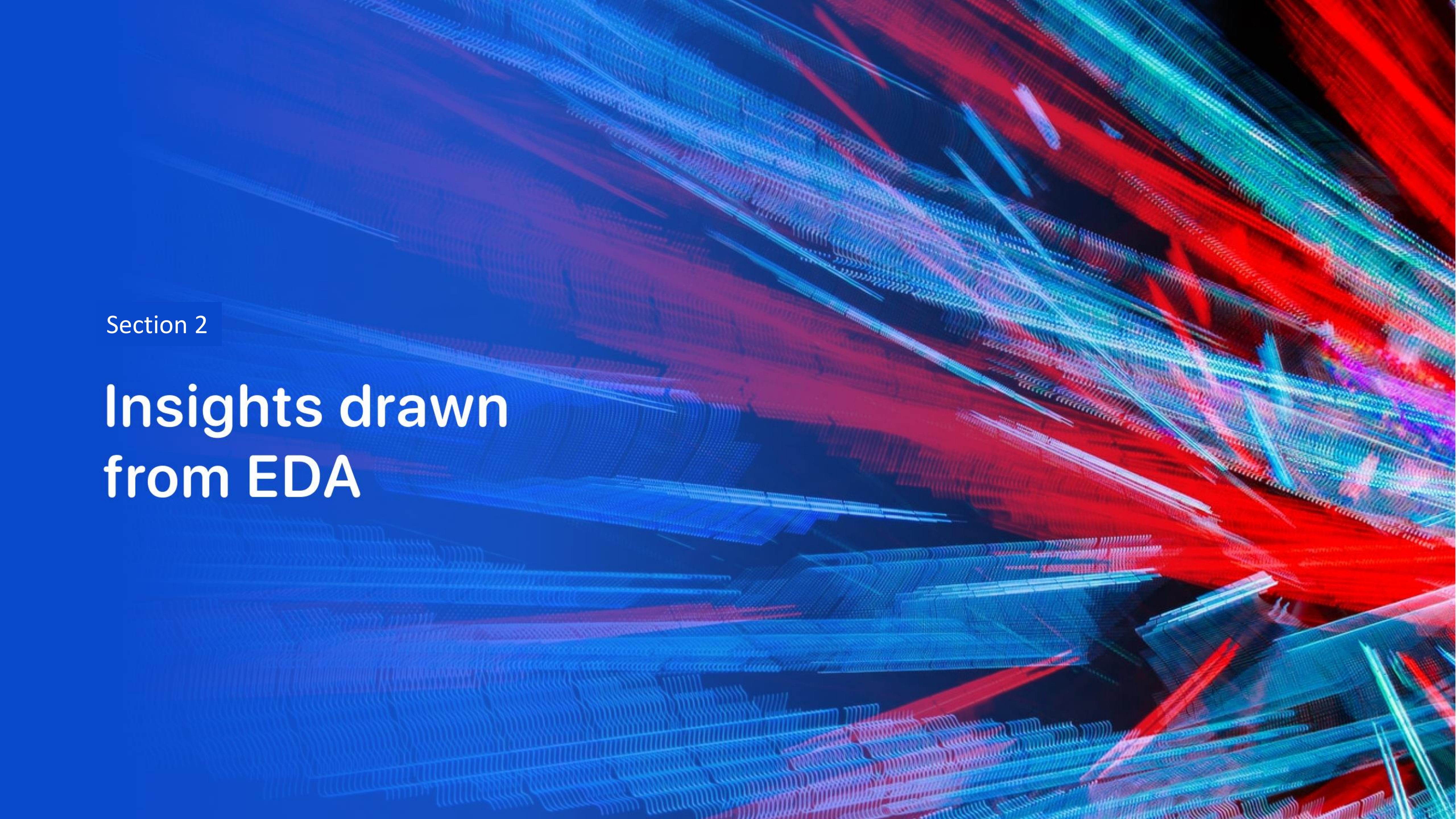
Finding the Best Classification Model

- The accuracy scores for all chosen algorithms were reviewed.
- The model with the highest accuracy score was determined as the best performing model.

Github Link

Section 2

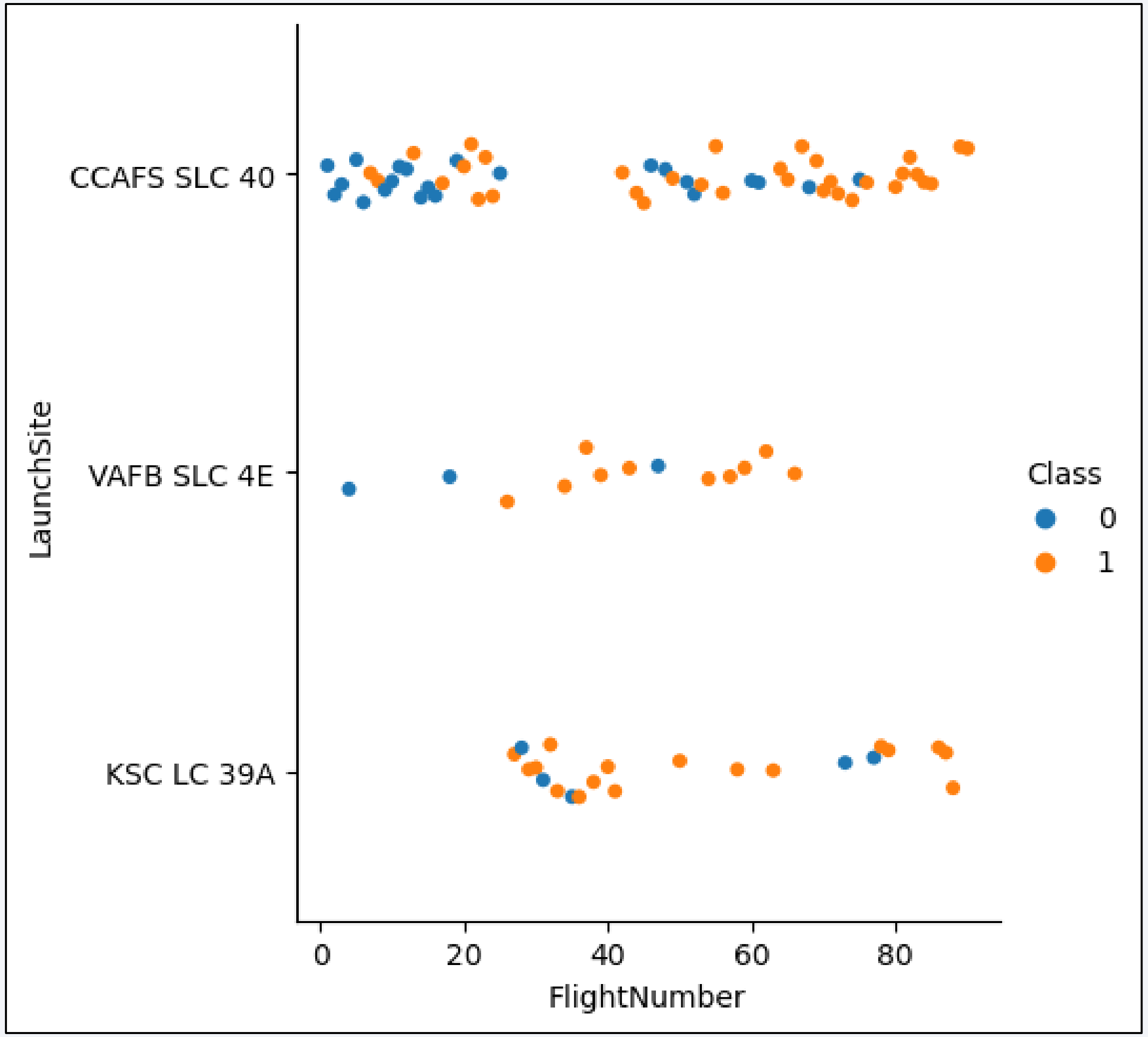
Insights drawn from EDA



Flight Number vs. Launch Site

The scatter plot illustrating the relationship between Launch Site and Flight Number reveals several insights:

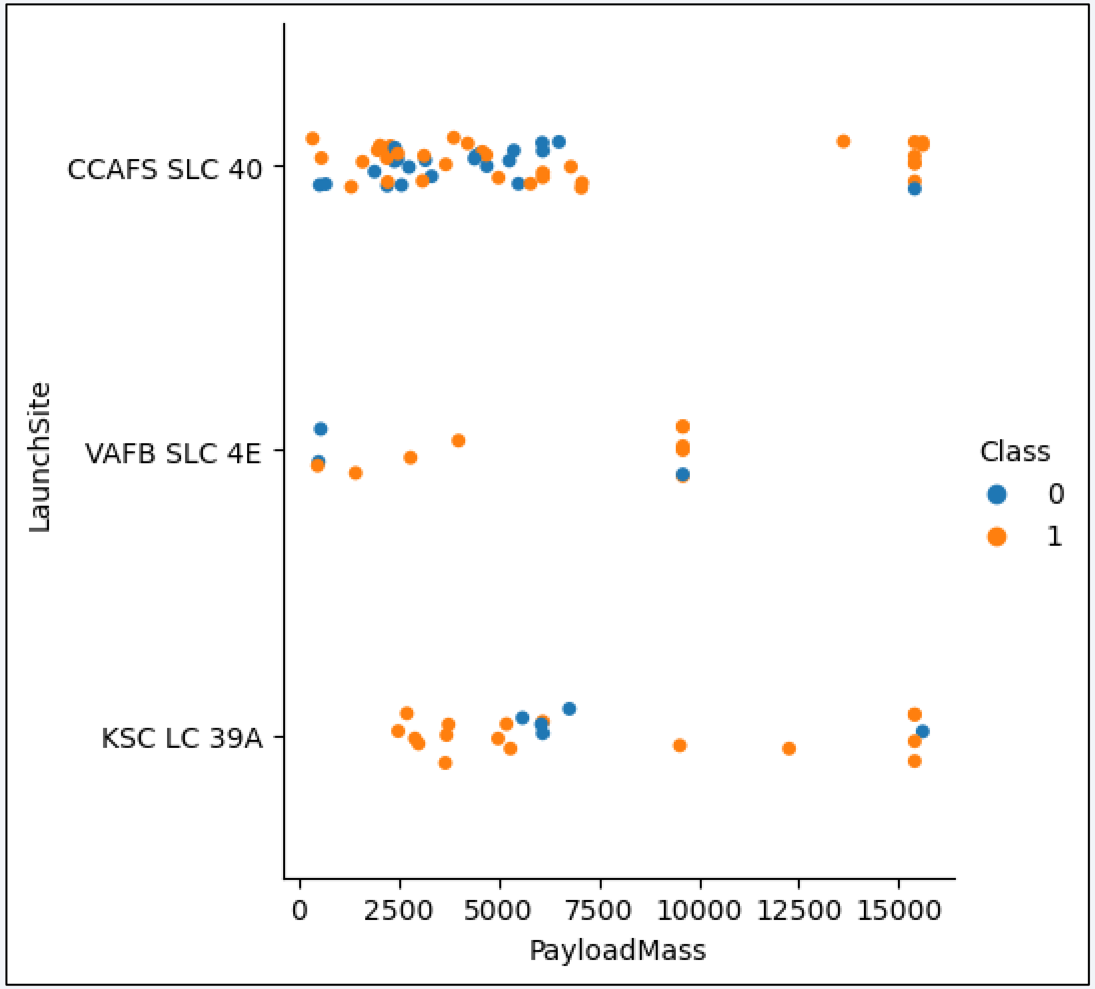
- There is a positive correlation between the number of flights and the success rate at a launch site. As the flight count increases, so does the success rate.
- The majority of initial flights (those with flight numbers less than 30) were launched from CCAFS SLC 40, and these generally did not succeed.
- A similar trend is observed with flights from VAFB SLC 4E, where earlier flights experienced less success.
- KSC LC 39A did not host any early flights, resulting in a higher success rate for launches from this site.
- For flights with a number greater than approximately 30, there is a noticeable increase in successful landings (Class = 1).



Payload vs. Launch Site

The scatter plot depicting the relationship between Launch Site and Payload Mass provides several key observations:

- For payload masses exceeding approximately 7000 kg, unsuccessful landings are quite rare. However, it's important to note that there is less data available for these heavier launches.
- There doesn't appear to be a distinct correlation between payload mass and success rate for a specific launch site.
- All sites have launched a range of payload masses. Notably, most of the launches from CCAFS SLC 40 carried comparatively lighter payloads, although there were some exceptions.



Success Rate vs. Orbit Type

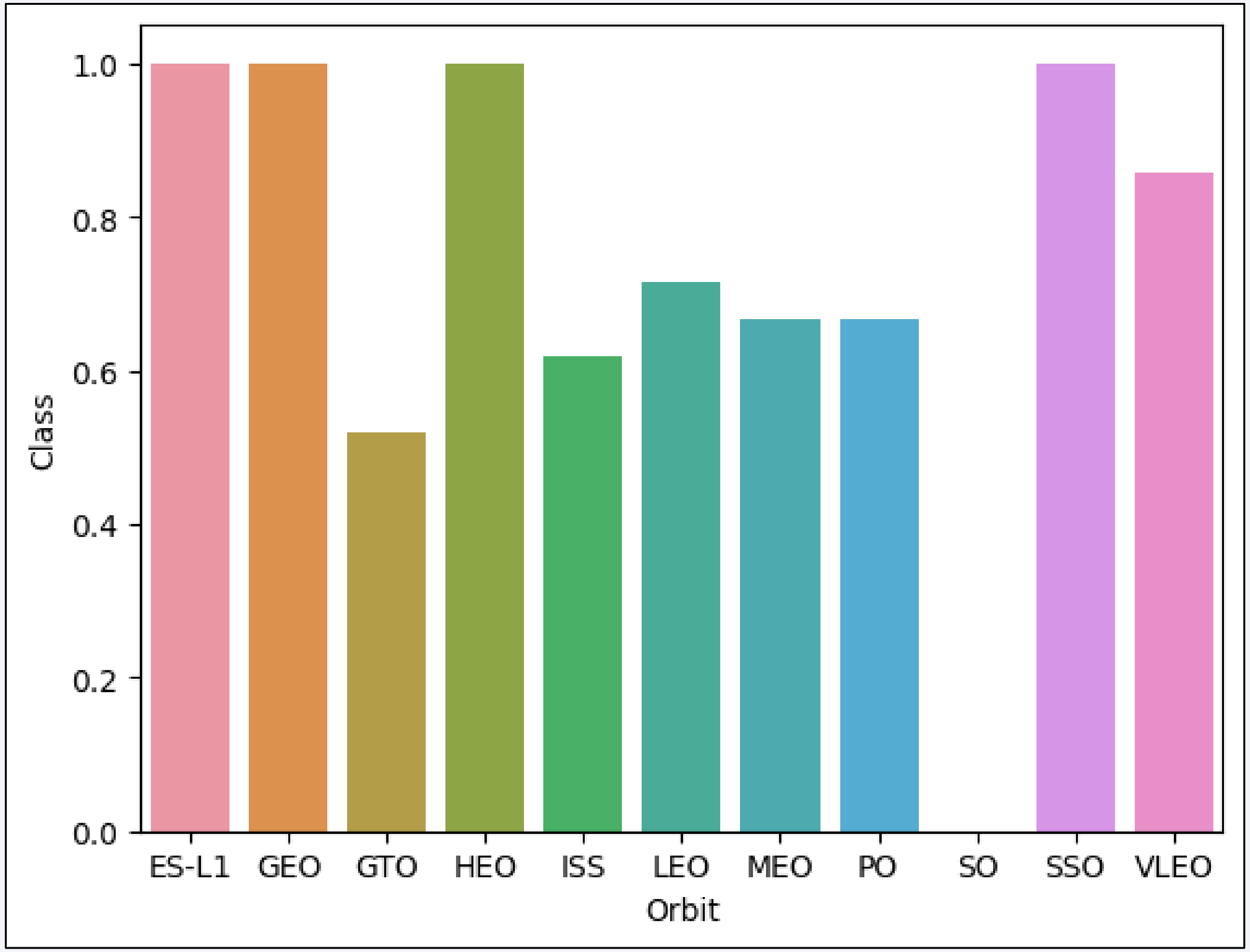
The bar chart illustrating Success Rate versus Orbit Type reveals that the following orbits have achieved a perfect success rate of 100%:

- ES-L1 (Earth-Sun First Lagrangian Point)
- GEO (Geostationary Orbit)
- HEO (High Earth Orbit)
- SSO (Sun-synchronous Orbit)

On the other hand, the orbit with the lowest success rate, specifically 0%, is:

- SO (Heliocentric Orbit)

This analysis provides valuable insights into the success rates associated with different orbit types.



Flight Number vs. Orbit Type

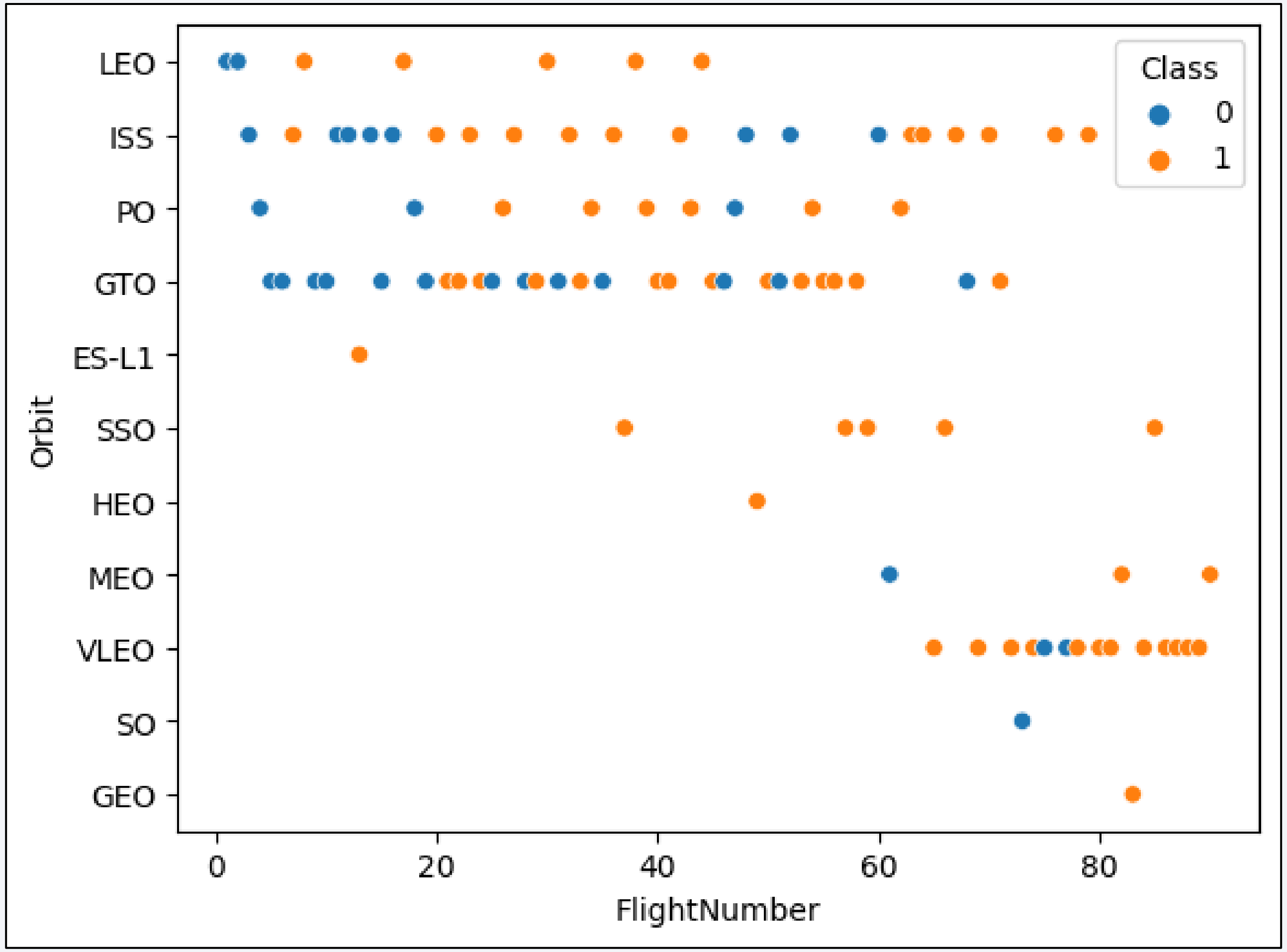
The scatter plot comparing Orbit Type and Flight Number offers several unique insights:

- The 100% success rate of GEO, HEO, and ES-L1 orbits can be attributed to the fact that there has only been one flight into each of these respective orbits.

- The 100% success rate in SSO is particularly noteworthy, given that it encompasses five successful flights.

- There appears to be little correlation between Flight Number and Success Rate for GTO.

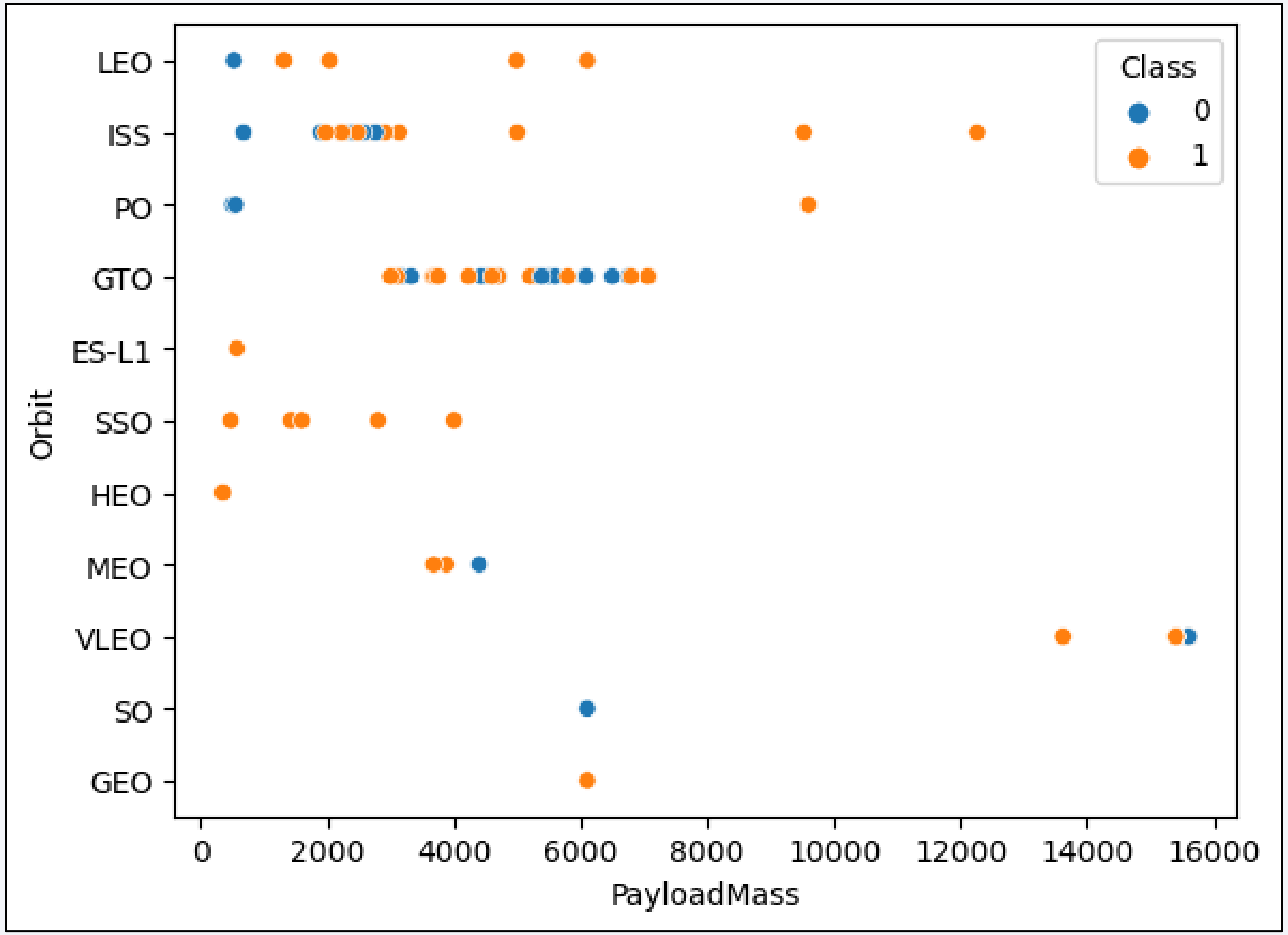
- Generally, an increase in Flight Number corresponds to an increase in the success rate. This trend is most pronounced for LEO, where unsuccessful landings predominantly occurred during the early launches with lower flight numbers.



Payload vs. Orbit Type

The scatter plot comparing Orbit Type and Payload Mass reveals several key observations:

- Certain orbit types, such as PO (Polar Orbit), ISS (International Space Station), and LEO (Low Earth Orbit), tend to have more successful launches with heavier payloads. However, it’s worth noting that the number of data points for PO is relatively small.
- For GTO (Geostationary Transfer Orbit), the relationship between payload mass and success rate is not clearly defined.
- Launches into VLEO (Very Low Earth Orbit) are typically associated with heavier payloads, which aligns with intuitive expectations.

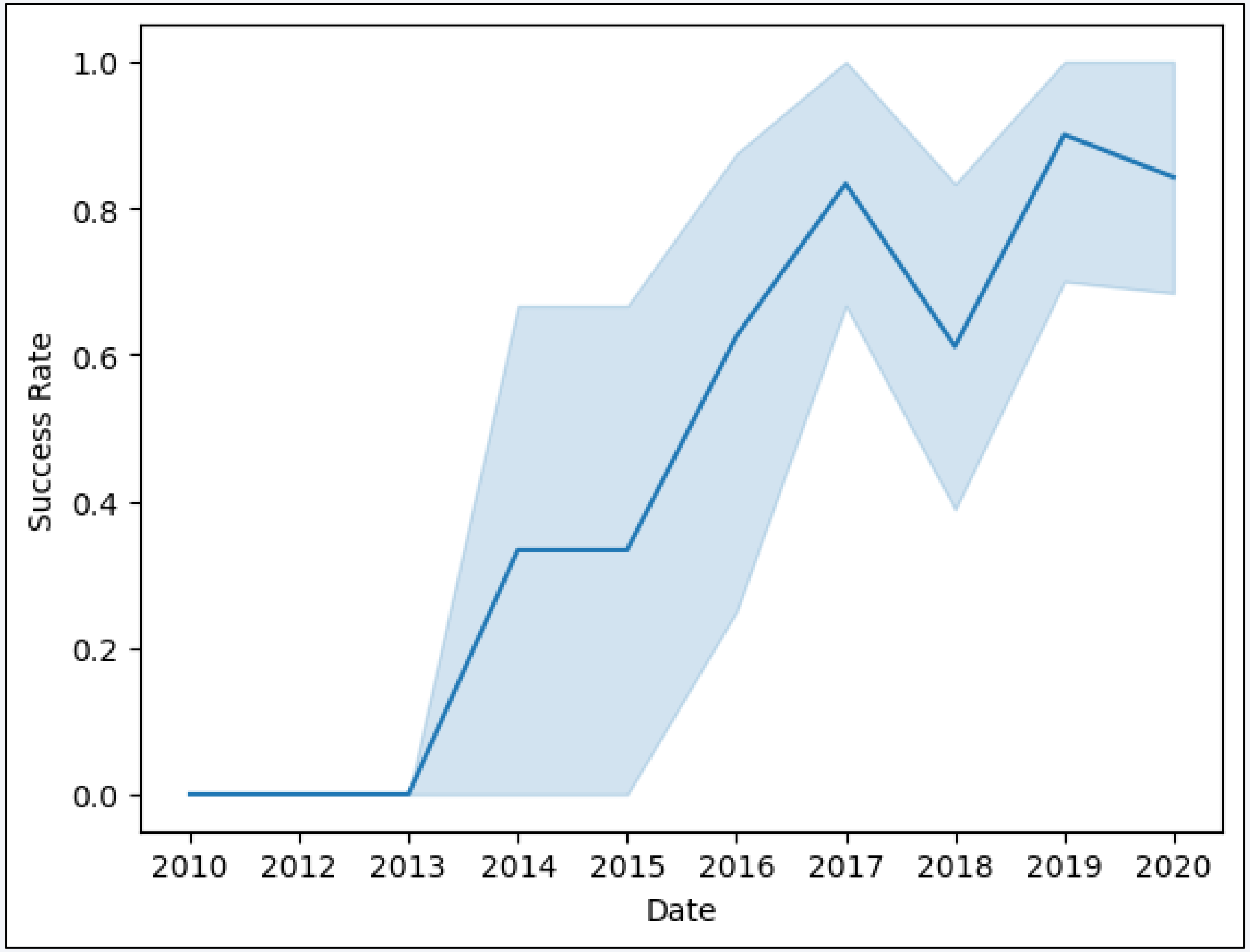


Launch Success Yearly Trend

The line chart depicting the yearly average success rate provides several key insights:

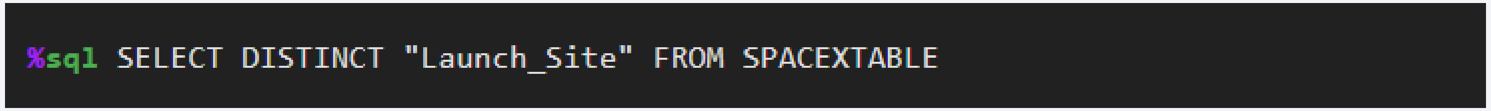
- From 2010 to 2013, all landings were unsuccessful, as indicated by a success rate of 0%.
- Following 2013, there was a general upward trend in the success rate, despite minor dips observed in 2018 and 2020.
- Post-2016, the chance of a successful landing was consistently above 50%.

This analysis highlights the significant improvements in landing success rates over time.



All Launch Site Names

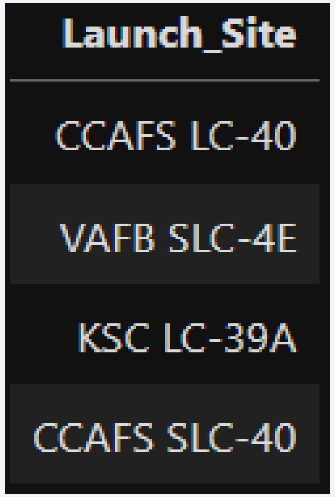
Display the names of the unique launch sites in the space mission:



```
%sql SELECT DISTINCT "Launch_Site" FROM SPACEXTABLE
```

Use DISTINCT to select only unique values from the table.

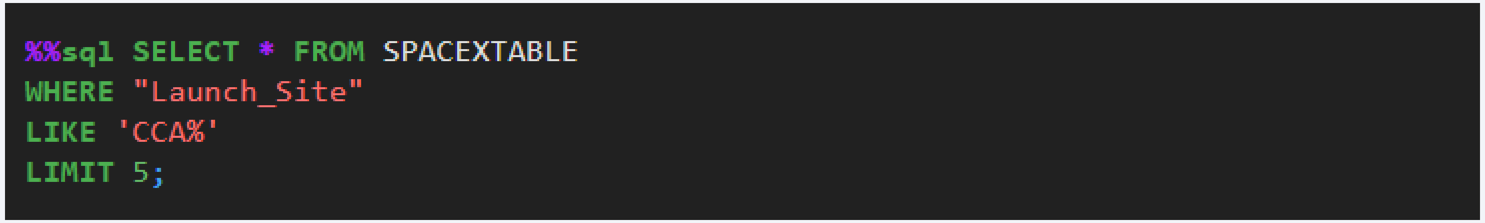
Result:



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

Launch Site Names Begin with 'CCA'

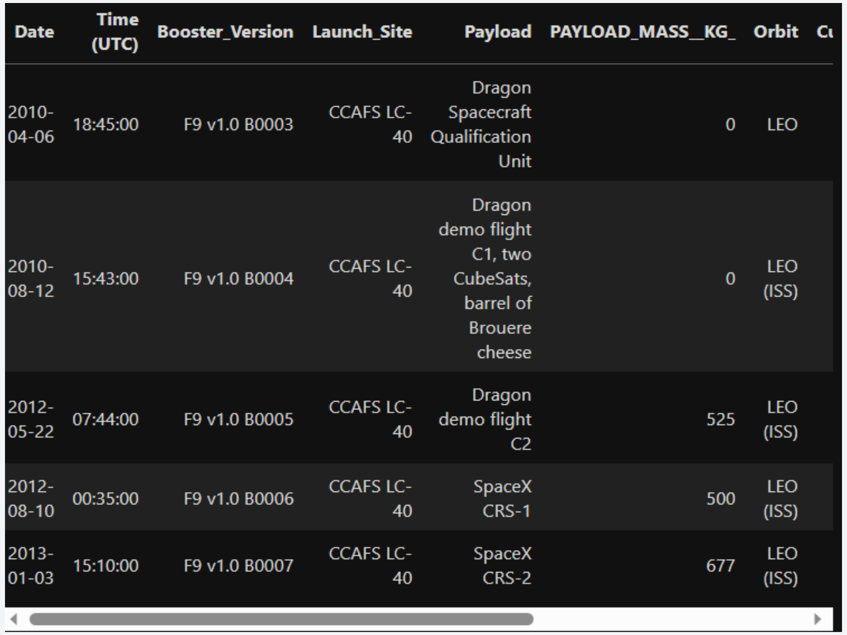
Display 5 records where launch sites begin with the string 'CCA'



```
%%sql SELECT * FROM SPACEXTABLE
WHERE "Launch_Site"
LIKE 'CCA%'
LIMIT 5;
```

LIMIT 5 fetches only 5 records, and the LIKE keyword is used with the wild card ‘CCA%’ to retrieve string values beginning with ‘CCA’.

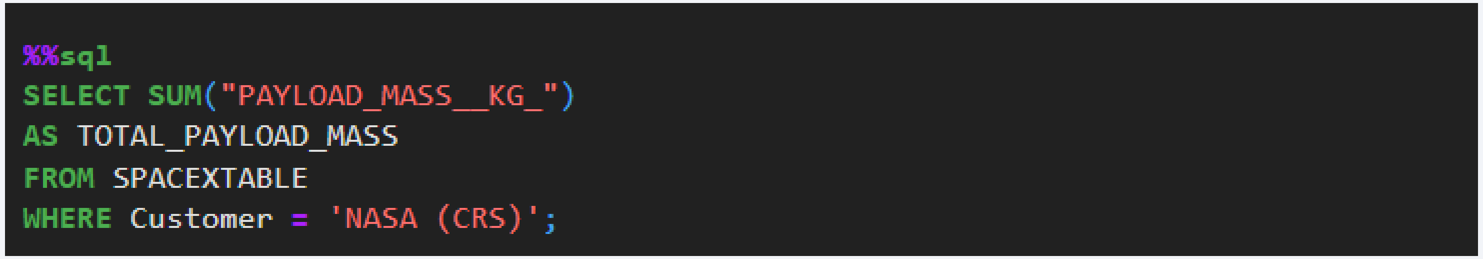
Result:



Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Cu
2010-04-06	18:45:00	F9 v1.0 B0003	CCAFS LC-40	Dragon Spacecraft Qualification Unit	0	LEO	
2010-08-12	15:43:00	F9 v1.0 B0004	CCAFS LC-40	Dragon demo flight C1, two CubeSats, barrel of Brouere cheese	0	LEO (ISS)	
2012-05-22	07:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525	LEO (ISS)	
2012-08-10	00:35:00	F9 v1.0 B0006	CCAFS LC-40	SpaceX CRS-1	500	LEO (ISS)	
2013-01-03	15:10:00	F9 v1.0 B0007	CCAFS LC-40	SpaceX CRS-2	677	LEO (ISS)	

Total Payload Mass

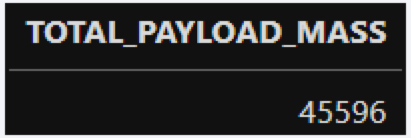
Display the total payload mass carried by boosters launched by NASA (CRS)



```
%%sql
SELECT SUM("PAYLOAD_MASS__KG_")
AS TOTAL_PAYLOAD_MASS
FROM SPACEXTABLE
WHERE Customer = 'NASA (CRS)';
```

The SUM keyword is used to calculate the total of the LAUNCH column, and the SUM keyword (and the associated condition) filters the results to only boosters from NASA (CRS).

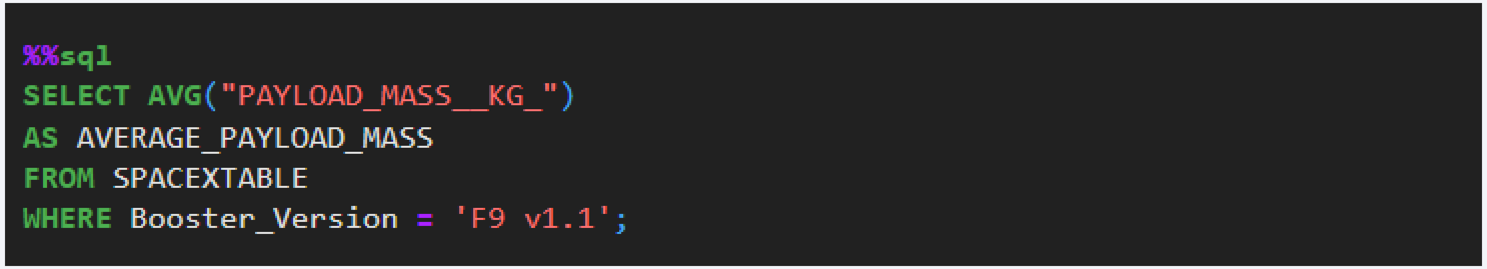
Result:



TOTAL_PAYLOAD_MASS
45596

Average Payload Mass by F9 v1.1

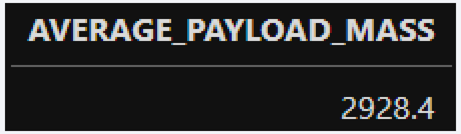
Display average payload mass carried by booster version F9 v1.1



```
%%sql
SELECT AVG("PAYLOAD_MASS__KG_")
AS AVERAGE_PAYLOAD_MASS
FROM SPACEXTABLE
WHERE Booster_Version = 'F9 v1.1';
```

The AVG keyword is used to calculate the average of the PAYLOAD_MASS__KG_ column, and the WHERE keyword (and the associated condition) filters the results to only the F9 v1.1 booster version.

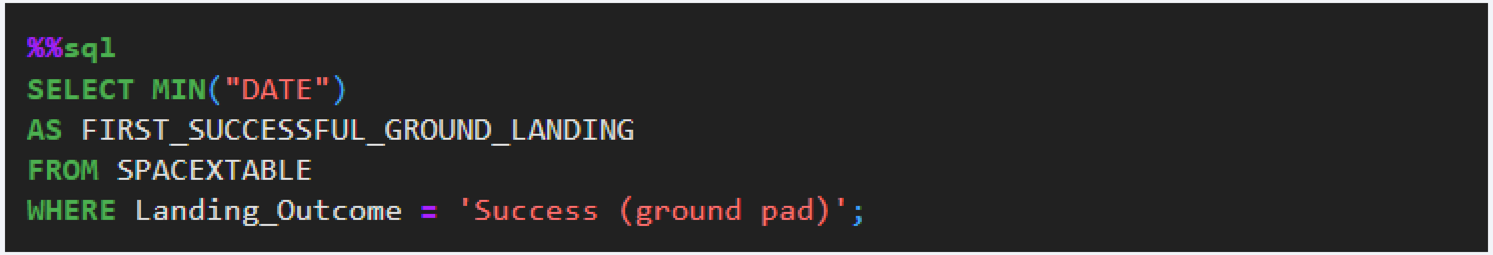
Result:



AVERAGE_PAYLOAD_MASS
2928.4

First Successful Ground Landing Date

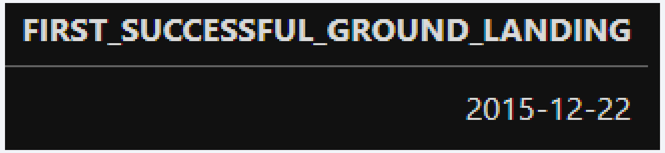
List the date when the first succesful landing outcome in ground pad was acheived.



```
%%sql
SELECT MIN("DATE")
AS FIRST_SUCCESSFUL_GROUND_LANDING
FROM SPACEXTABLE
WHERE Landing_Outcome = 'Success (ground pad)';
```

The MIN keyword is used to calculate the minimum of the DATE column, i.e. the first date, and the WHERE keyword (and the associated condition) filters the results to only the successful ground pad landings.

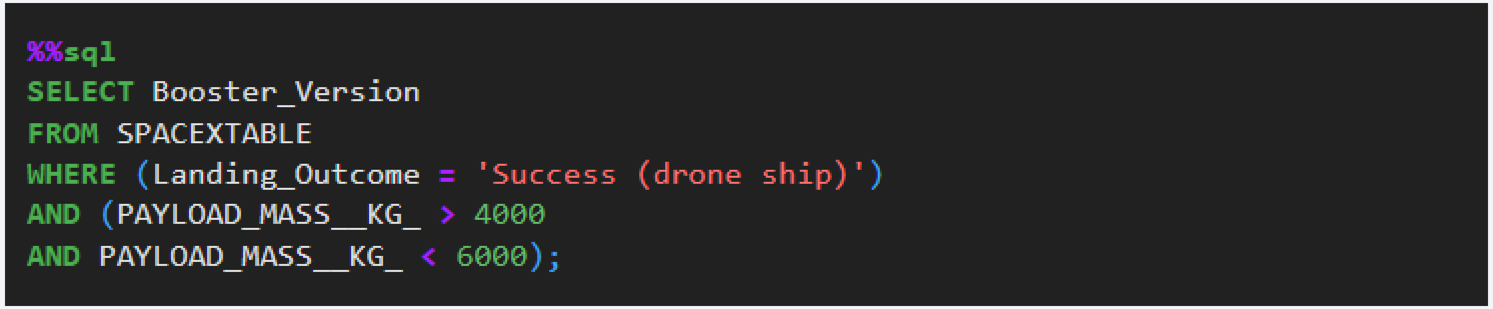
Result:



FIRST_SUCCESSFUL_GROUND_LANDING
2015-12-22

Successful Drone Ship Landing with Payload between 4000 and 6000

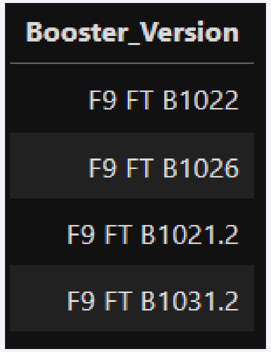
List the names of the boosters which have success in drone ship and have payload mass greater than 4000 but less than 6000



```
%%sql
SELECT Booster_Version
FROM SPACEXTABLE
WHERE (Landing_Outcome = 'Success (drone ship)')
AND (PAYLOAD_MASS__KG_ > 4000
AND PAYLOAD_MASS__KG_ < 6000);
```

The WHERE keyword is used to filter the results to include only those that satisfy both conditions in the brackets (as the AND keyword is also used). The BETWEEN keyword allows for 4000 < x < 6000 values to be selected.

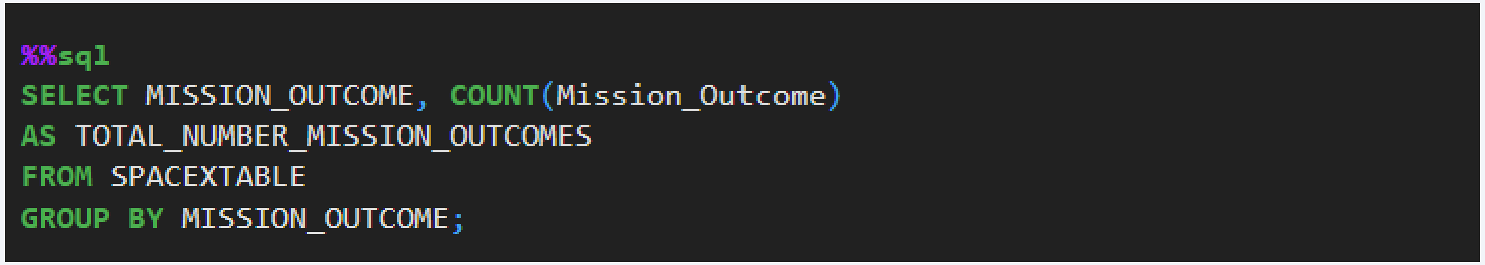
Result:



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

Total Number of Successful and Failure Mission Outcomes

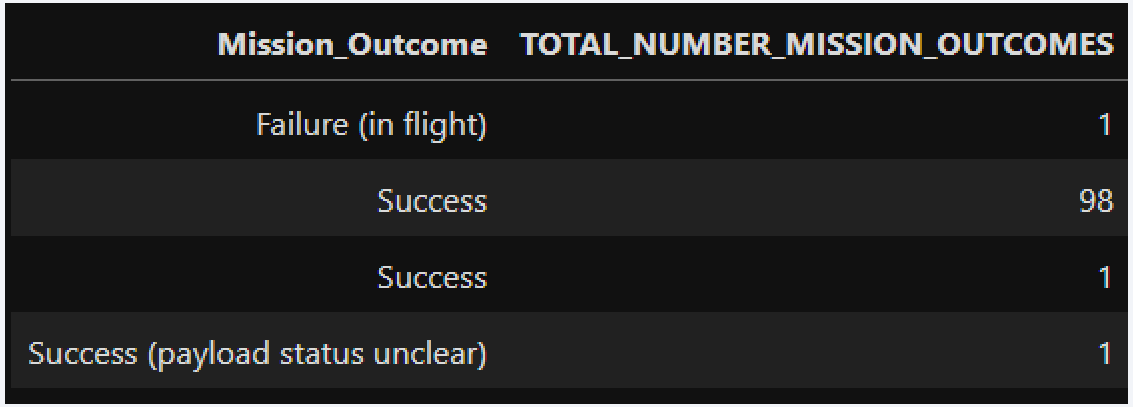
List the total number of successful and failure mission outcomes



```
%%sql
SELECT MISSION_OUTCOME, COUNT(Mission_Outcome)
AS TOTAL_NUMBER_MISSION_OUTCOMES
FROM SPACEXTABLE
GROUP BY MISSION_OUTCOME;
```

The COUNT keyword is used to calculate the total number of mission outcomes, and the GROUPBY keyword is also used to group these results by the type of mission outcome.

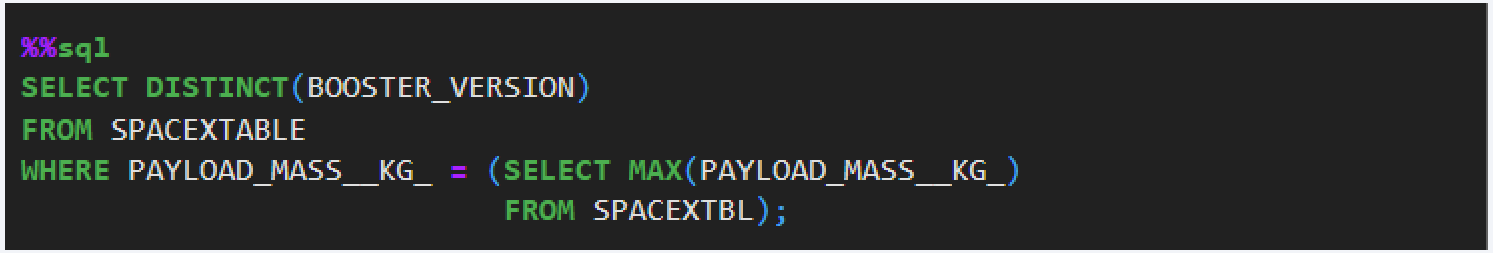
Result:



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

Boosters Carried Maximum Payload

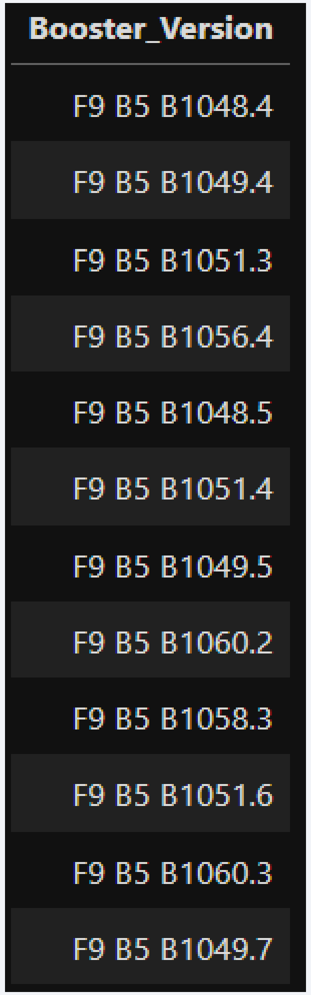
List the names of the booster_versions which have carried the maximum payload mass.



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

A subquery is used here. The SELECT statement within the brackets finds the maximum payload, and this value is used in the WHERE condition. The DISTINCT keyword is then used to retrieve only distinct /unique booster versions.

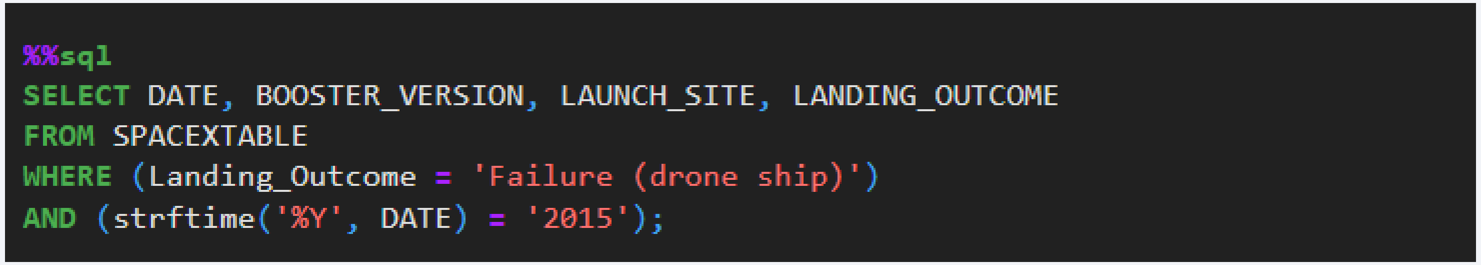
Result:



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

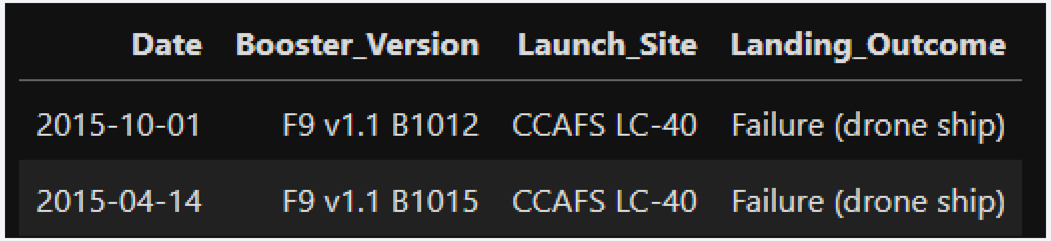
List the records which will display the month names, failure landing_outcomes in drone ship, booster versions, launch_site for the months in year 2015.



```
%%sql
SELECT DATE, BOOSTER_VERSION, LAUNCH_SITE, LANDING_OUTCOME
FROM SPACEXTABLE
WHERE (Landing_Outcome = 'Failure (drone ship)')
AND (strftime('%Y', DATE) = '2015');
```

The WHERE keyword is used to filter the results for only failed landing outcomes, AND only for the year of 2015.

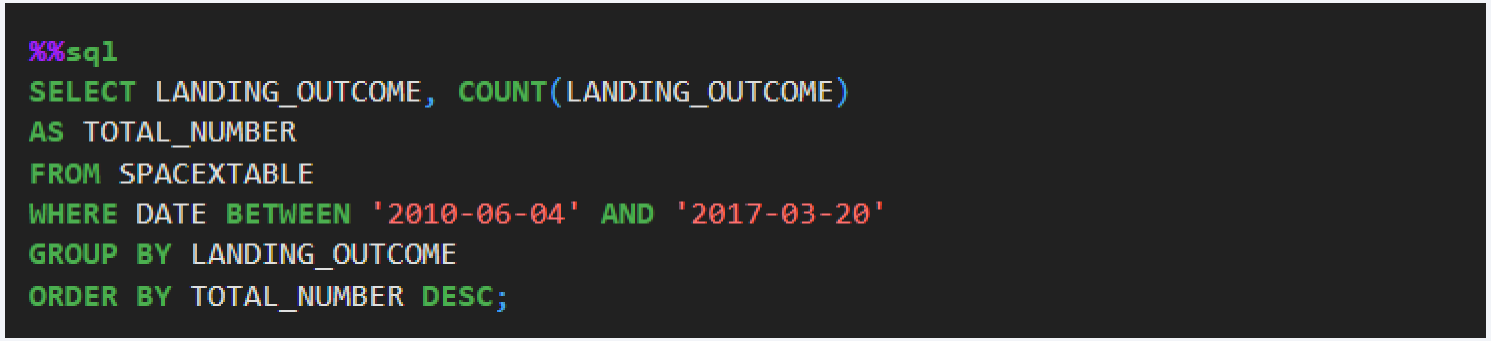
Result:



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

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

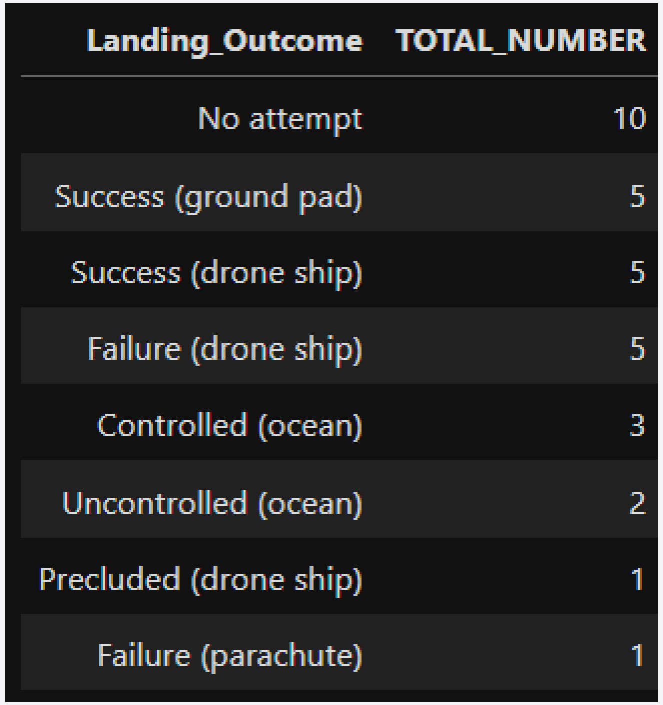
Rank the count of landing outcomes (such as Failure (drone ship) or Success (ground pad)) between the date 2010-06-04 and 2017-03-20, in descending order.



```
%%sql
SELECT LANDING_OUTCOME, COUNT(LANDING_OUTCOME)
AS TOTAL_NUMBER
FROM SPACEXTABLE
WHERE DATE BETWEEN '2010-06-04' AND '2017-03-20'
GROUP BY LANDING_OUTCOME
ORDER BY TOTAL_NUMBER DESC;
```

The WHERE keyword is used with the BETWEEN keyword to filter the results to dates only within those specified. The results are then grouped and ordered, using the keywords GROUP BY and ORDER BY, respectively, where DESC is used to specify the descending order.

Result:

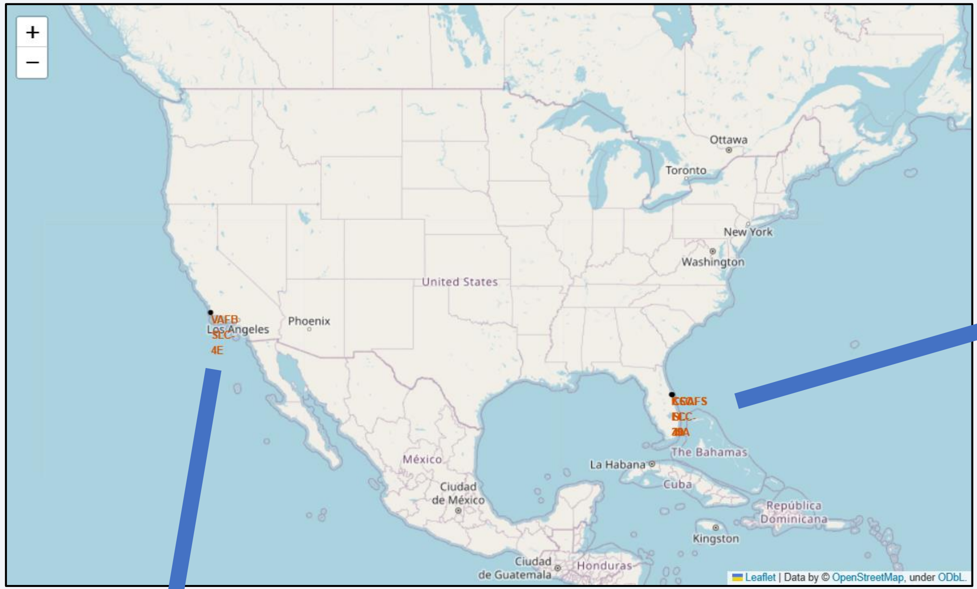


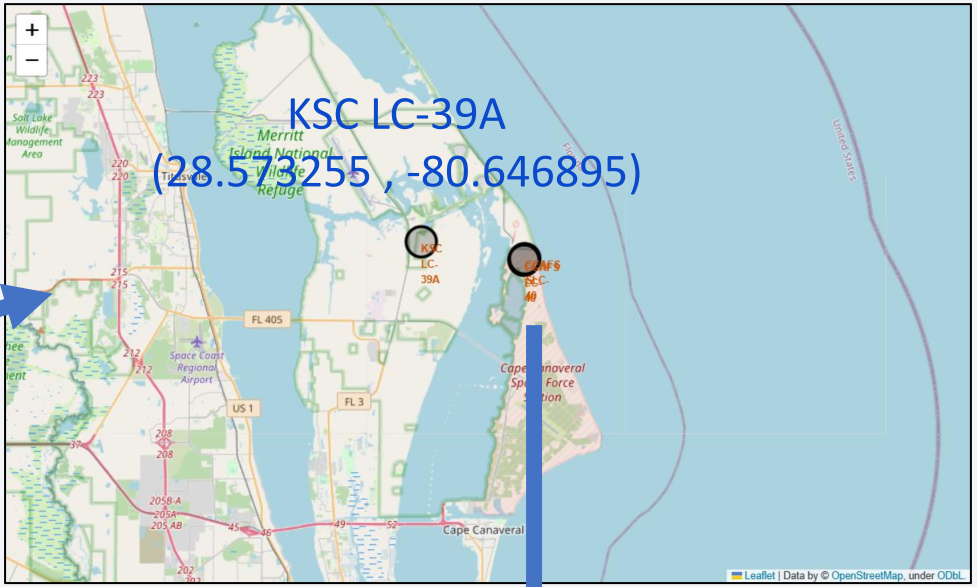
Landing_Outcome	TOTAL_NUMBER
No attempt	10
Success (ground pad)	5
Success (drone ship)	5
Failure (drone ship)	5
Controlled (ocean)	3
Uncontrolled (ocean)	2
Precluded (drone ship)	1
Failure (parachute)	1

Section 3

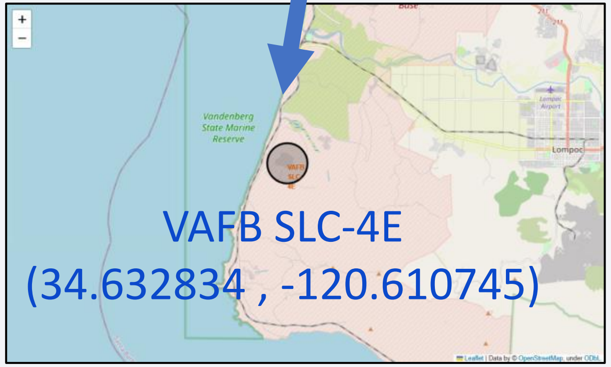
Launch Sites Proximities Analysis

All Launch Sites

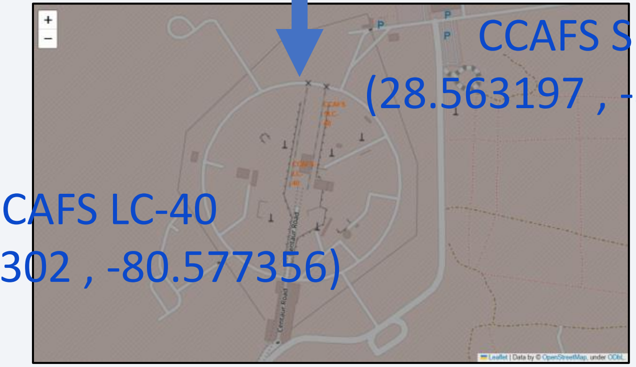



+
−
Merritt Island National Wildlife Refuge
Titusville
KSC LC-39A
CCAFS SLC 40
FL 405
Space Coast Regional Airport
US 1
FL 3
Cape Canaveral Space Force Station
Cape Canaveral
Leaflet | Data by © OpenStreetMap, under ODbL.

KSC LC-39A
(28.573255 , -80.646895)


+
−
Vandenberg State Marine Reserve
Lompoc
VAFB SLC 4E
Leaflet | Data by © OpenStreetMap, under ODbL.

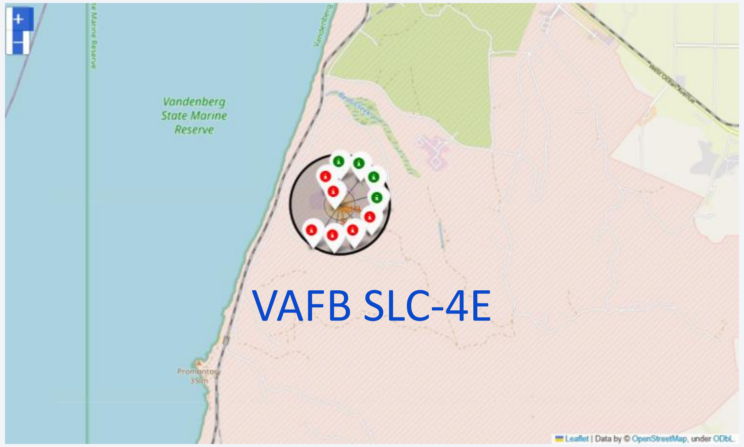
VAFB SLC-4E
(34.632834 , -120.610745)



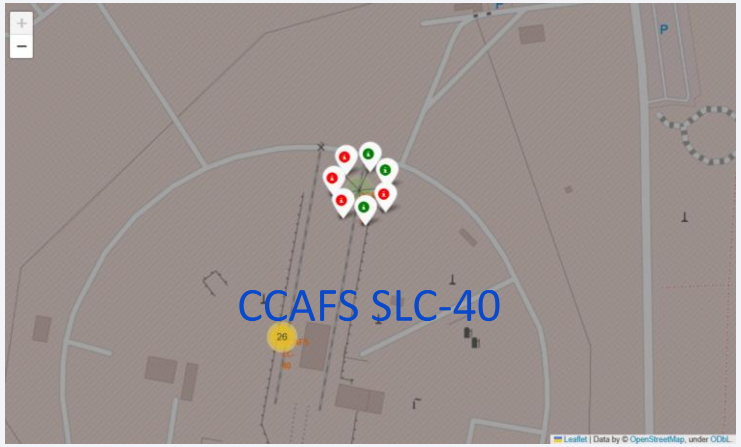
CCAFS SLC-40
(28.563197 , -80.576820)

CCAFS LC-40
(28.562302 , -80.577356)

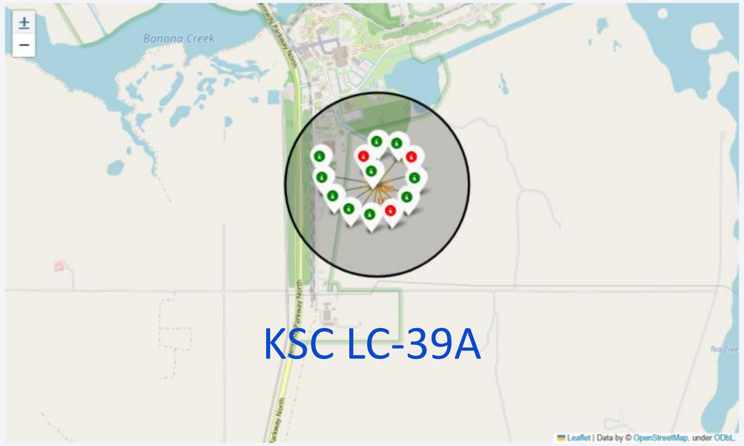
Success Status For Each Launch Site



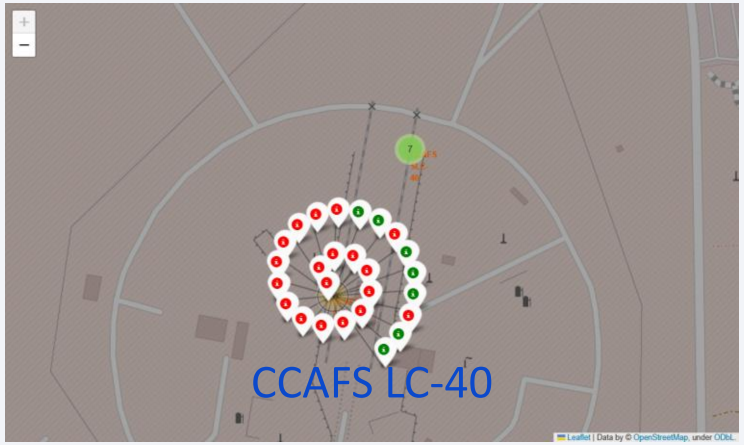
VAFB SLC-4E



CCAFS SLC-40

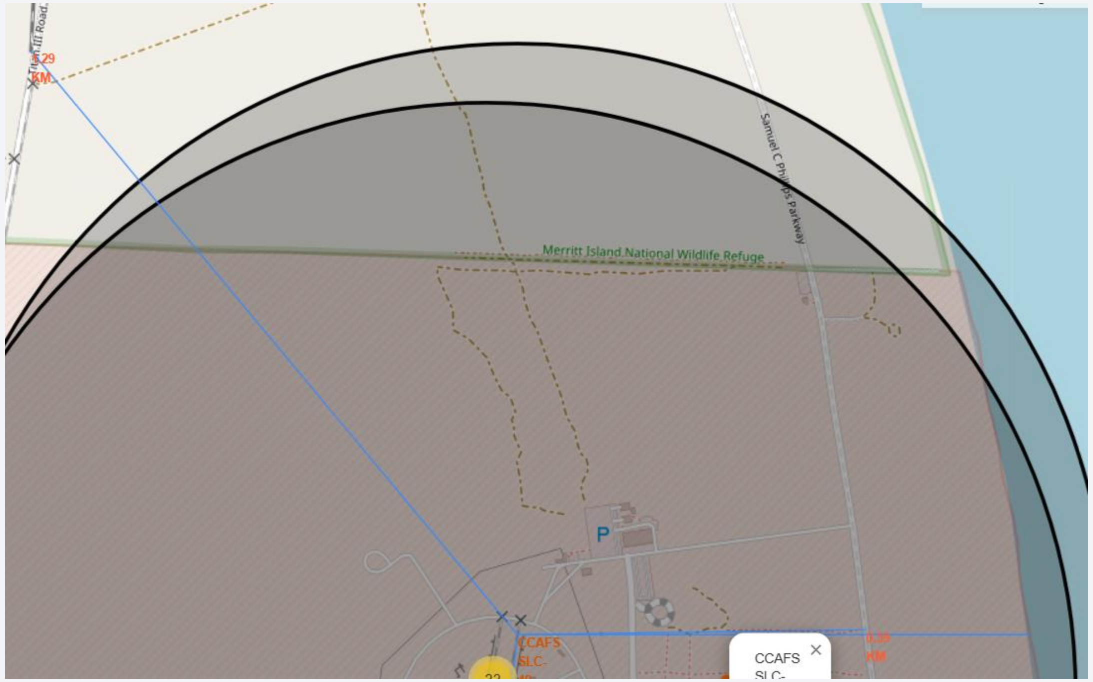


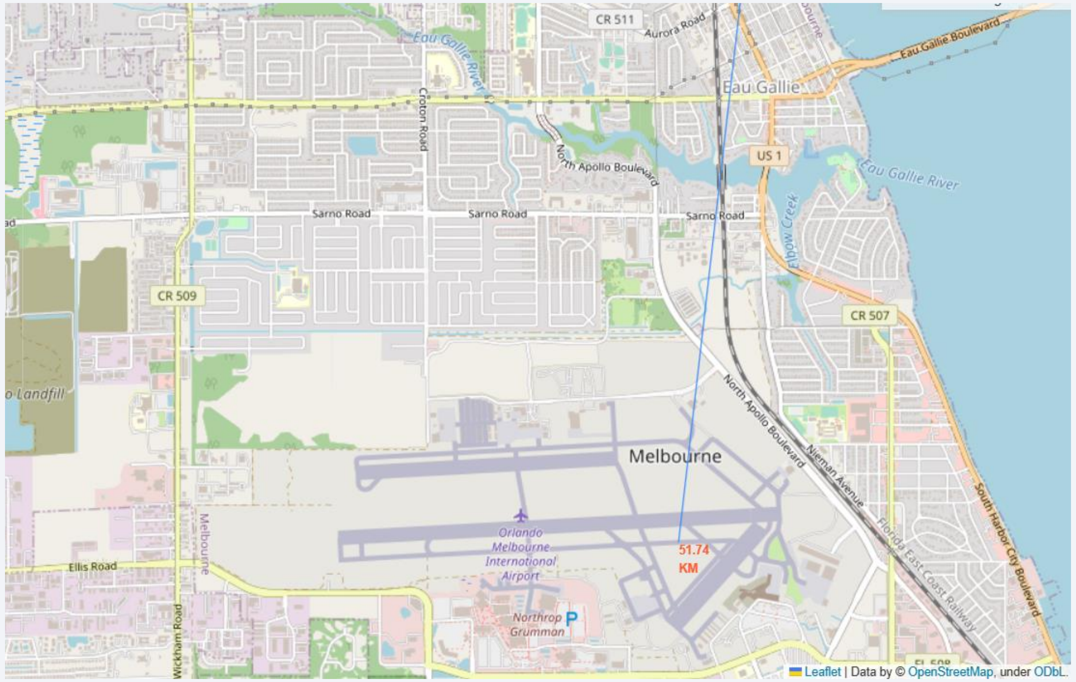
KSC LC-39A



CCAFS LC-40

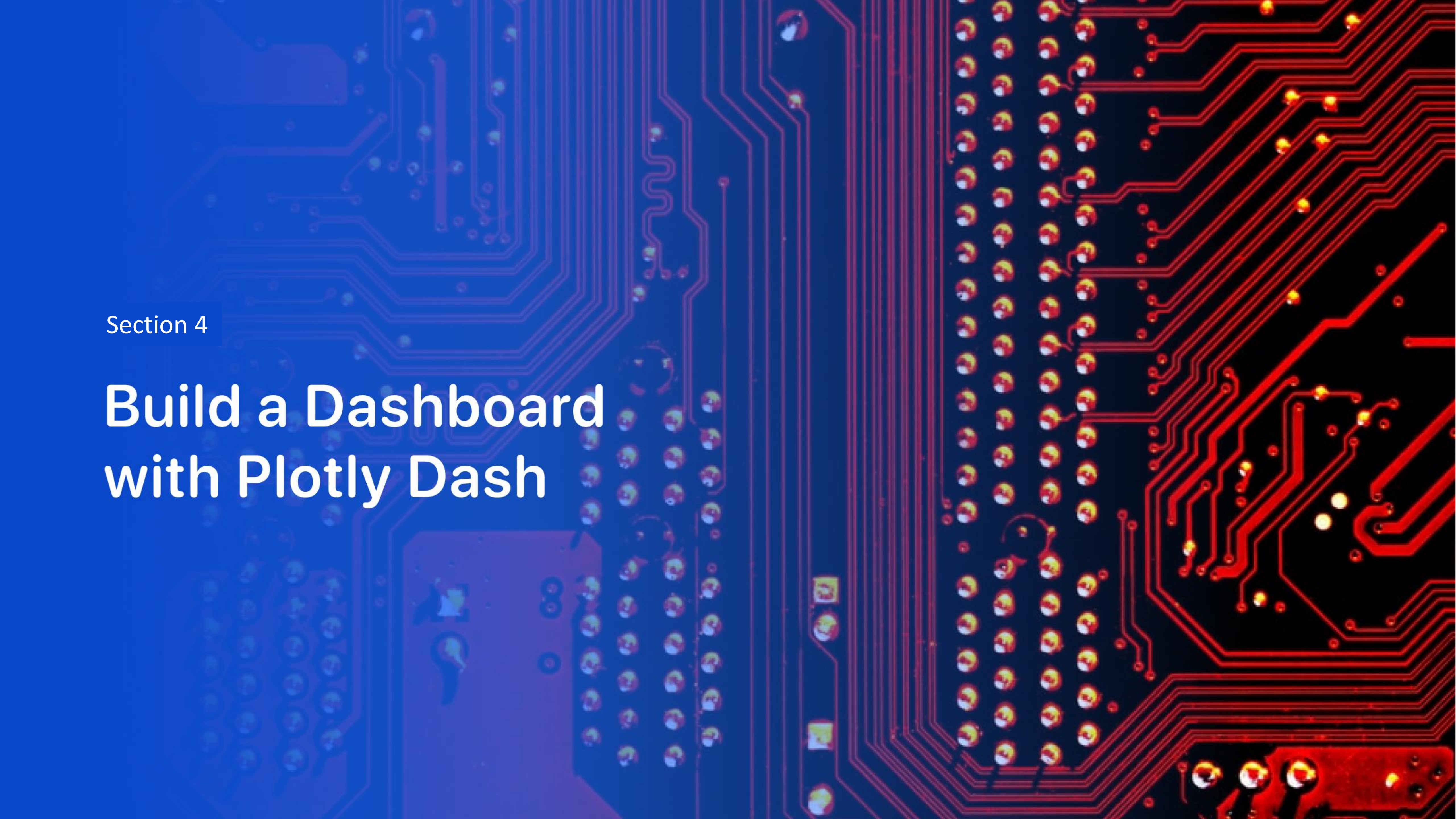
Points Of Interest Proximity


1.29
KM
Samuel C Phillips Parkway
Merritt Island National Wildlife Refuge
P
CCAFS
SLC-
CCAFS
SLC-



For CCAFS SLC-40 :

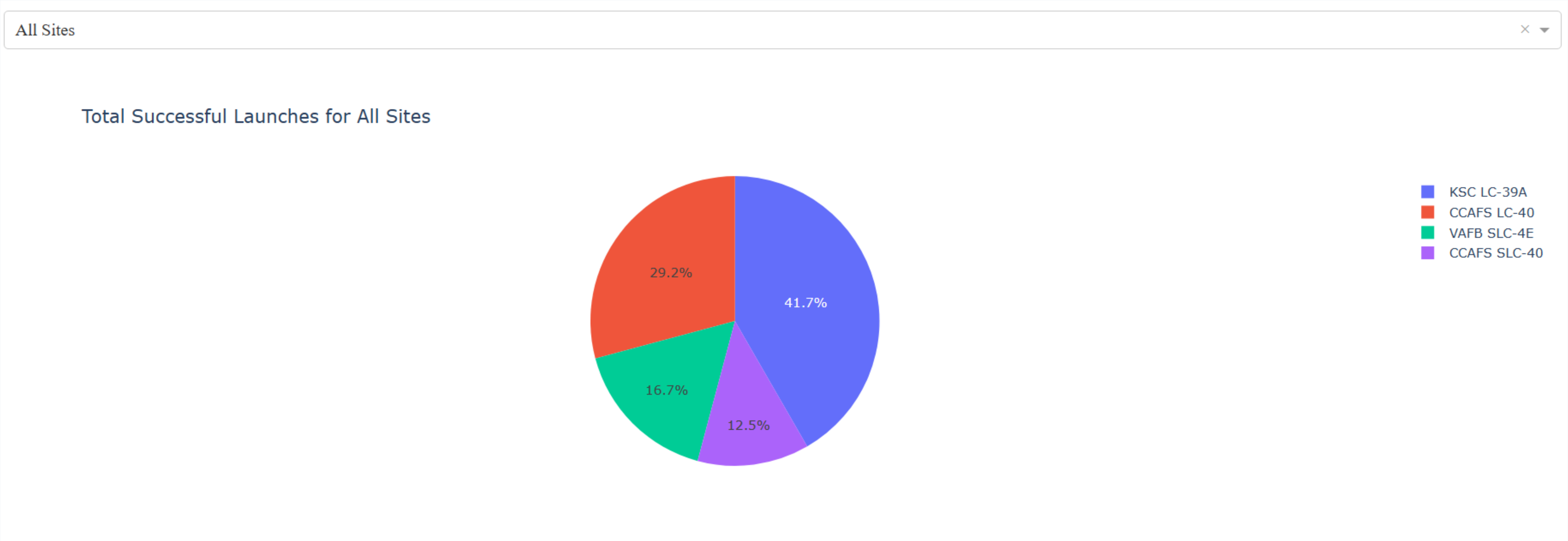
1. Nearest Railway : 1.29 km
2. Nearest Highway: 0.59 km
3. Nearest Coastline: 0.87 km
4. Nearest Airport: 51.74 km



Section 4

Build a Dashboard with Plotly Dash

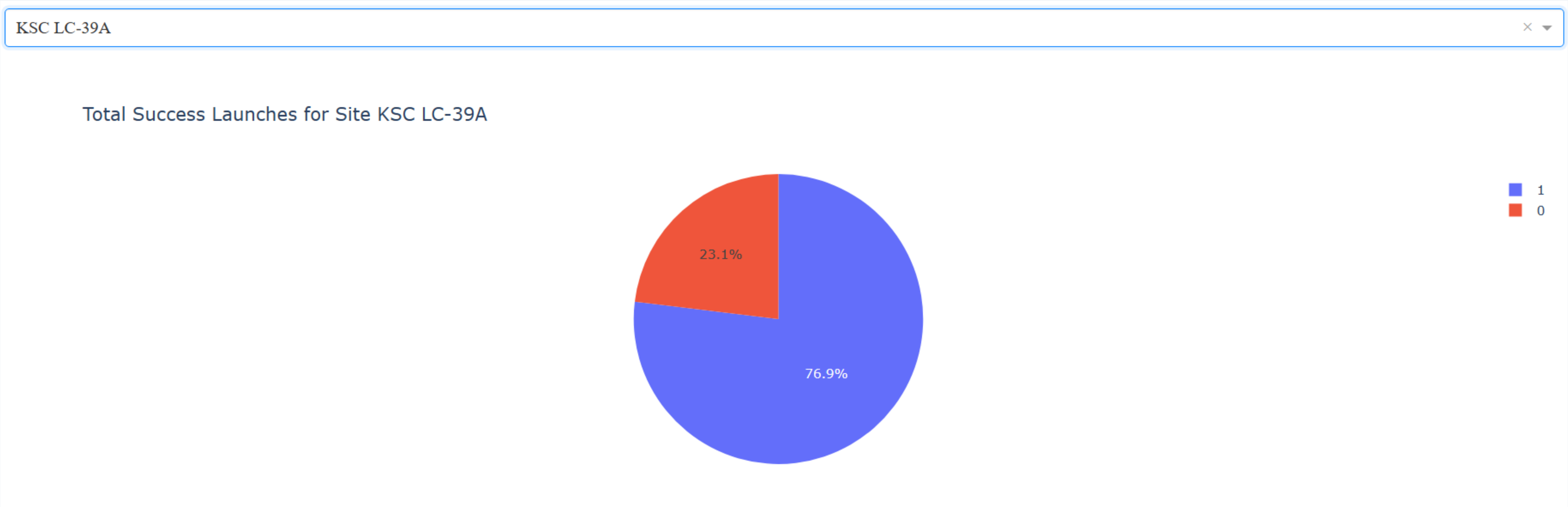
Successful Launcher for All Sites



All Sites

The launch site KSC LC-39A had the most successful launches, with 41.7% of the total successful launches while the launch site CCAFS SLC-40 had the least successful launcher with 12.5% of the total successful launches.

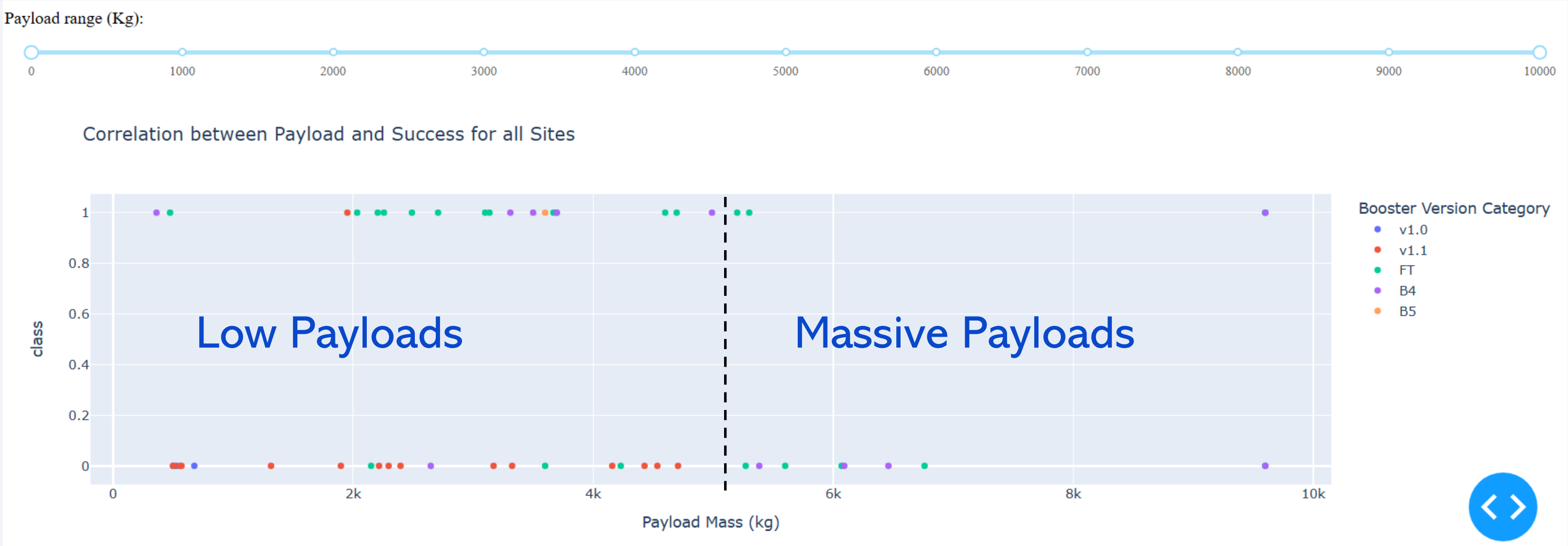
Launch Site with Highest Success Ratio



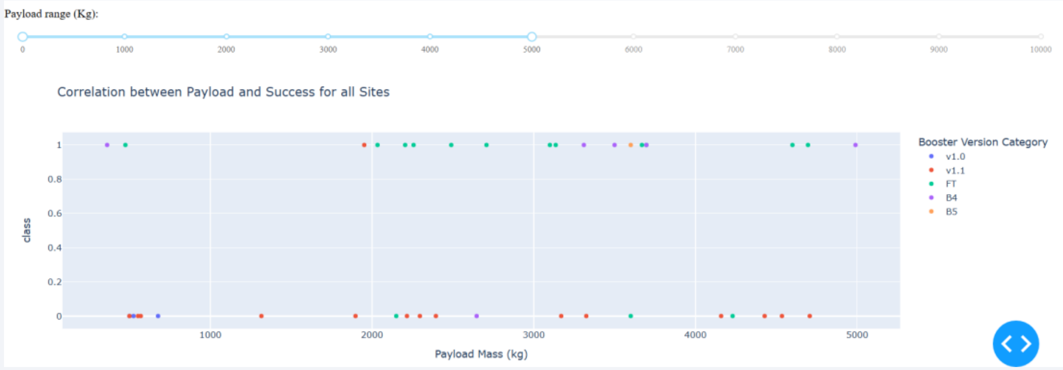
KSC LC-39A

The launch site KSC LC-39A had the highest success launch ratio, with 76.9% of the total launches are successful.

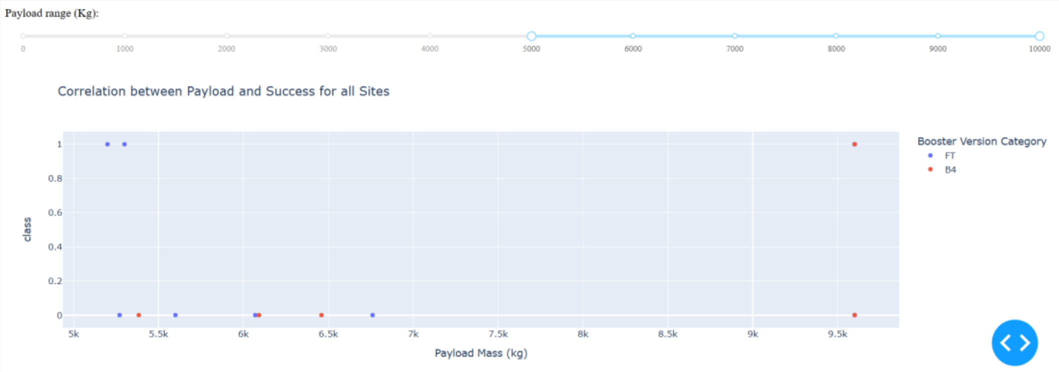
Launch Outcome VS. Payload Scatter Plot



1. There is a gap around 4000 kg when plotting the launch outcome against the payload for all sites.
2. The data can be divided into two ranges: low payloads (0-4000 kg) and massive payloads (4000-10000 kg).
3. The success rate for massive payloads is lower than that for low payloads.
4. Some booster types, such as v1.0, v1.1 and B5, have not been used to launch massive payloads.



Low Payloads



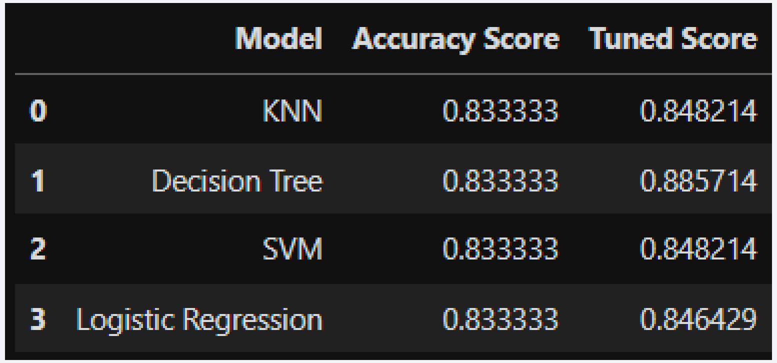
Massive Payloads

Section 5

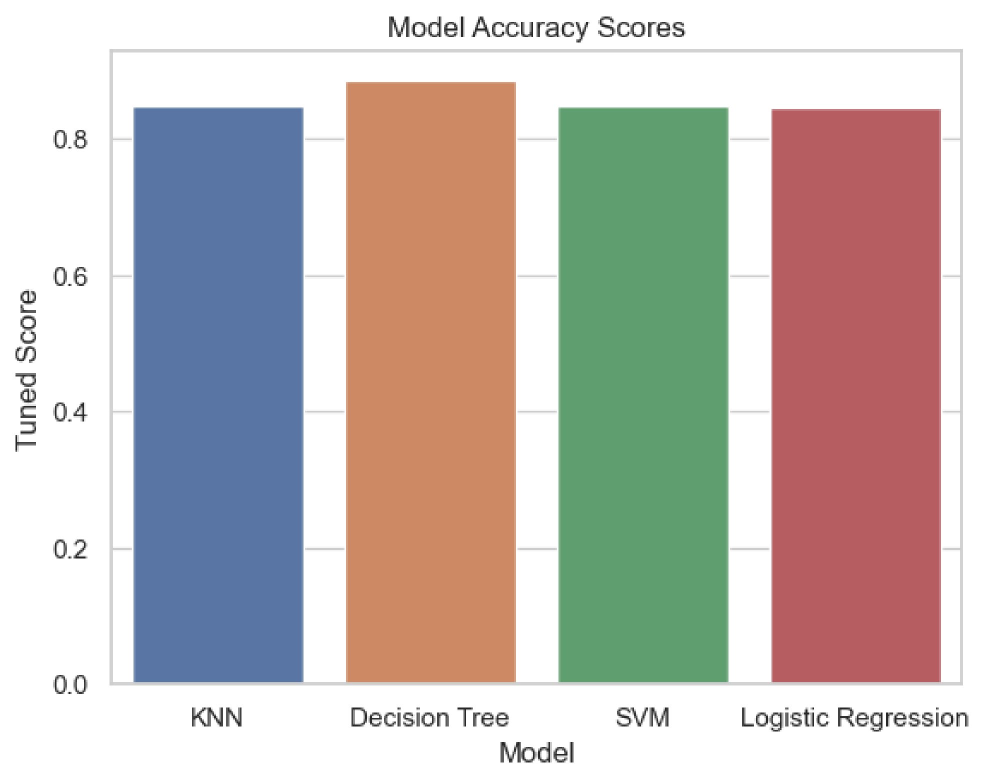
Predictive Analysis (Classification)

Classification Accuracy

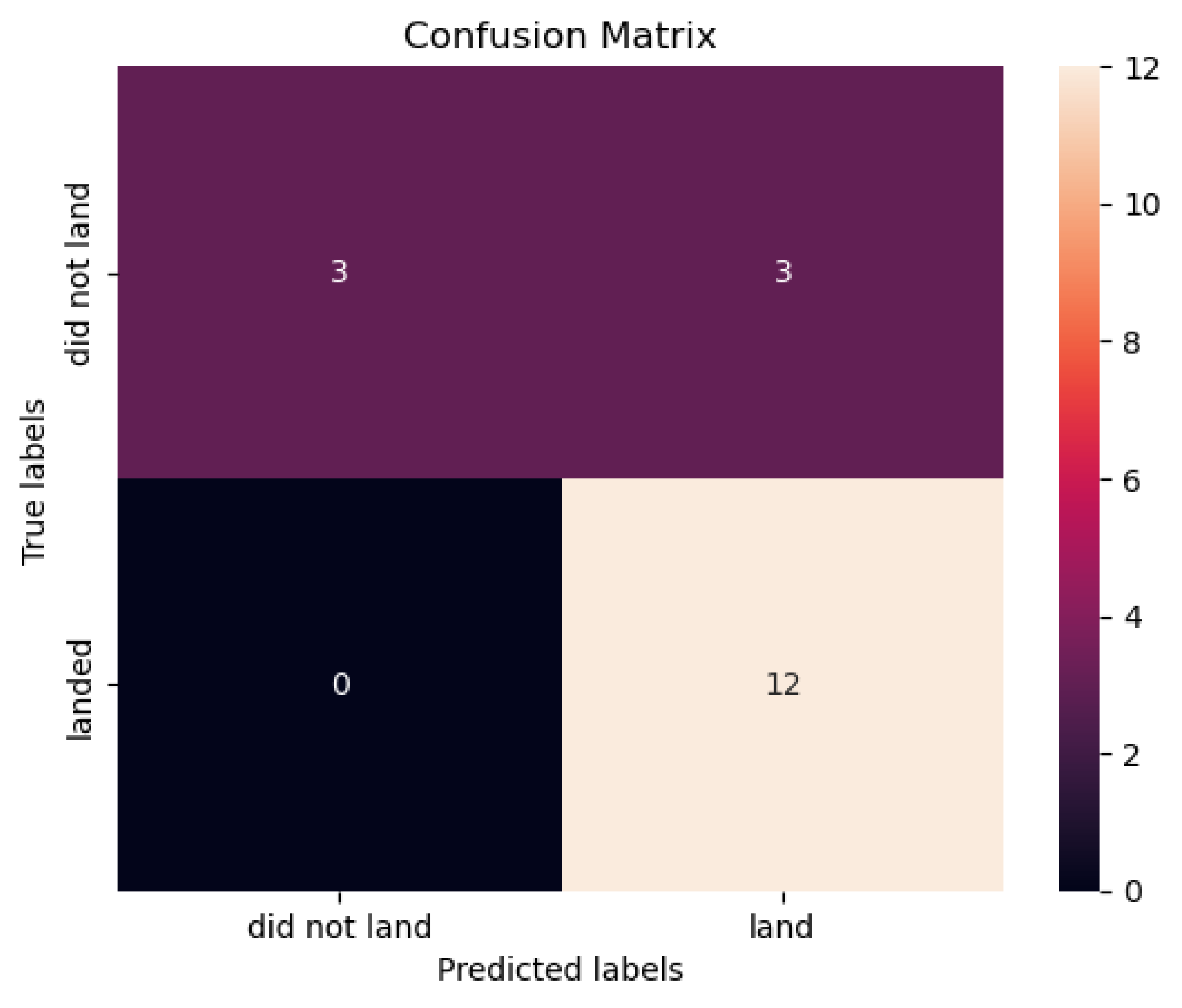
- Four classification models have been created: KNN, Decision Tree, SVM, and Logistic Regression.
- The Decision Tree model is the most accurate, with a tuned accuracy score of 88.6%.



	Model	Accuracy Score	Tuned Score
0	KNN	0.833333	0.848214
1	Decision Tree	0.833333	0.885714
2	SVM	0.833333	0.848214
3	Logistic Regression	0.833333	0.846429



Confusion Matrix



- The Decision Tree model is the most accurate, with a tuned accuracy score of 88.6%.
- The confusion matrix on the left shows that out of a total of 18 launches, only 3 were misclassified while the remaining 15 were correctly classified.

Conclusions

- The success rate at a launch site increases as the number of flights increases, with most early flights being unsuccessful. This suggests that with more experience, the success rate increases.
  - Between 2010 and 2013, all landings were unsuccessful, as indicated by a success rate of 0%.
  - After 2013, the success rate generally increased, despite experiencing small dips in 2018 and 2020.
  - After 2016, there was always a greater than 50% chance of success.
- Orbit types ES-L1, GEO, HEO, and SSO have the highest success rate at 100%. The 100% success rate of GEO, HEO, and ES-L1 orbits can be attributed to the fact that there was only one flight into each of these respective orbits. The 100% success rate in SSO is particularly impressive, with 5 successful flights.
  - The orbit types PO, ISS, and LEO have more success with heavy payloads. VLEO (Very Low Earth Orbit) launches are associated with heavier payloads, which is intuitively understandable.
- The launch site KSC LC-39A had the most successful launches, accounting for 41.7% of the total successful launches. It also had the highest rate of successful launches at 76.9%.
  - The success rate for massive payloads (over 4000kg) is lower than that for low payloads.
- The best performing classification model is the Decision Tree model, with an accuracy of 88.6%.

Appendix

1. Relevant Python Libraries:
   - Matplotlib
   - Numpy
   - Seaborn
   - Plotly
   - Dash
   - Scikit-Learn
   - Folium
2. Do check the Github links provided in each section for the source code.

Thank you!