IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Rohit Kumar
6th Sep, 2025

Outline

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

Executive Summary

- **Methodologies Implemented:**
  - Data collection through SpaceX REST API and web scraping
  - Data cleaning and wrangling for quality assurance
  - Exploratory Data Analysis (EDA) using statistical visualization and SQL queries
  - Development of interactive visual dashboards using Folium and Plotly Dash
  - Predictive modeling using classification algorithms with hyperparameter tuning and validation
- **Key Results:**
  - Identified critical factors influencing launch success including payload, orbit type, and booster version
  - Created interactive maps and dashboards enabling dynamic data exploration by stakeholders
  - Built and tuned classification models (Logistic Regression, SVM, Decision Tree, KNN)
  - Achieved highest model accuracy with Decision Tree, demonstrating reliable prediction capabilities

Introduction

- **Project Background and Context:**
  SpaceX revolutionizes space travel with reusable rockets and frequent launches. Understanding launch success factors is crucial to improving mission reliability and optimizing operations.
- **Problem Statements:**
  - What are the key factors influencing launch success and failure?
  - How do payload mass, orbit type, launch site, and booster version affect outcomes?
  - Can machine learning models predict launch success with high accuracy?
- **Project Goals:**
  - Collect and organize SpaceX launch data using API and web scraping.
  - Perform exploratory data analysis to uncover launch patterns.
  - Build interactive dashboards for intuitive data exploration.
  - Develop and evaluate classification models to predict launch success.

Section 1

Methodology

Methodology

- Data collection methodology:
  - Collected SpaceX Launch data via REST API calls.
- Perform data wrangling
  - Cleaned and processed raw data for analysis.
  - Handled missing values and transformed relevant features (e.g., date parsing, encoding categorical variables).
- Perform exploratory data analysis (EDA) using visualization and SQL
- Perform interactive visual analytics using Folium and Plotly Dash
- Perform predictive analysis using classification models
  - Trained multiple classification models (Logistic Regression, SVM, Decision Tree, KNN).
  - Tuned hyperparameters using GridSearchCV and validated with 10-fold cross-validation.
  - Selected best model based on accuracy and performance metrics.

Data Collection

- Sources of Data:
  - SpaceX REST API: Used to download comprehensive launch records including flight numbers, launch dates, payload mass, orbit types, launch outcomes, and booster info.
  - Web Scraping: Supplemented additional data attributes relevant to launch sites and booster characteristics.
- Process Overview:
  - Automated REST API calls to retrieve JSON data from SpaceX endpoints.
  - Implemented web scraping scripts to extract details not available via API.
  - Validated and combined data from multiple sources for completeness.

SpaceX REST API Call

- Retrieve launch records JSON data
- Extract key fields: FlightNumber, Date, Payload Mass, Orbit, Booster Info, Launch Site, Outcome etc.

Data Validation & Cleaning

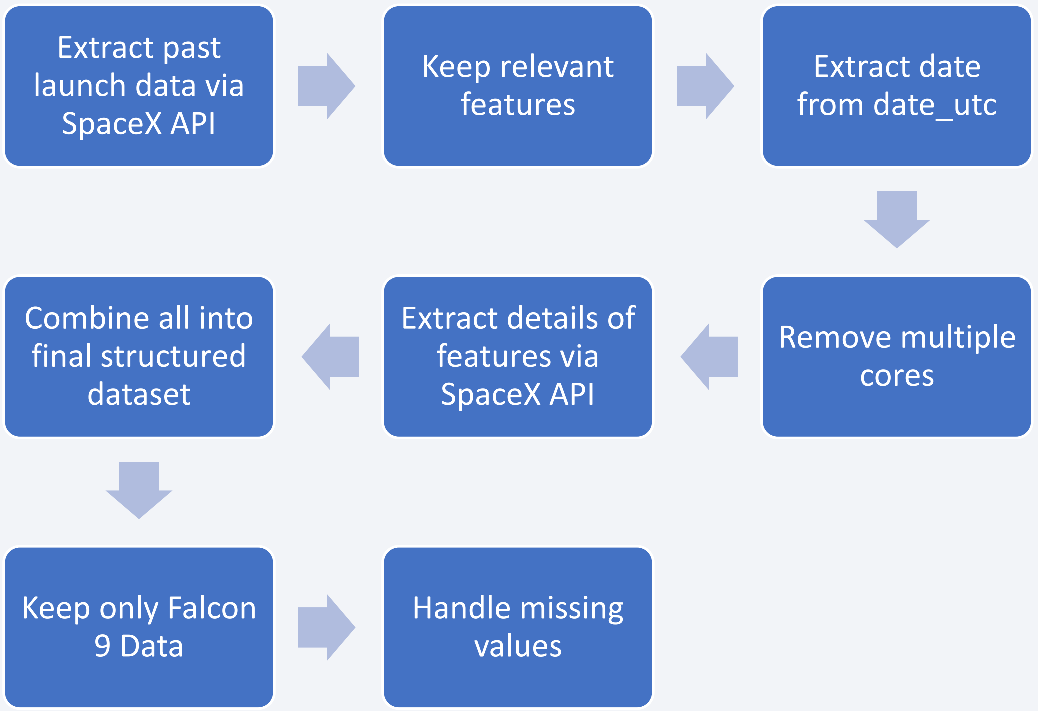
- Check for missing or inconsistent values
- Format dates, convert types

Combining Data Sources

- Merge API data with scraped data based on common keys (e.g., Launch Site, FlightNumber)

Data Collection – SpaceX API

- Overview
  - Data collected from the official SpaceX REST API (https://api.spacexdata.com/v4/launches/past)
  - Data includes detailed records of launches, rockets, payloads, launchpads, and cores
- Data Extraction Process
  - API GET requests used to retrieve raw JSON data for past launches
  - Helper functions implemented for enriching data by querying related endpoints:
  - Booster version: extracted from rocket IDs
  - Launch site: name, longitude, latitude from launchpad IDs
  - Payload: mass (kg) and orbit type from payload IDs
  - Core details: landing success, flight numbers, grid fins, reused status, legs, landing pad, block version, reuse count, serial numbers from cores array

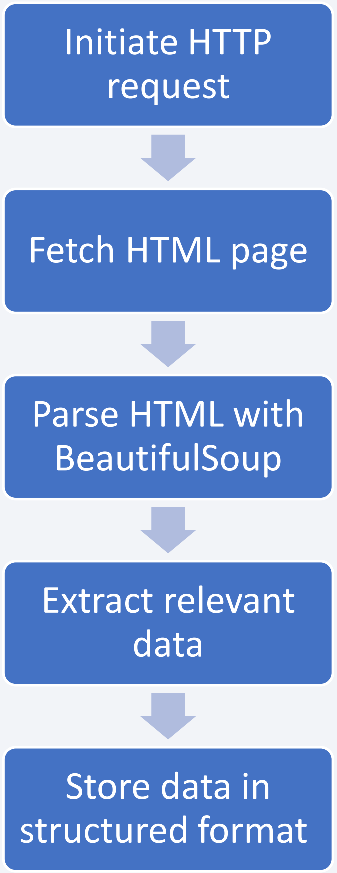


Reference: [Data-Science-Assignment/jupyter-labs-spacex-data-collection-api.ipynb at main · rk-rohitkumar/Data-Science-Assignment]()

Data Collection - Scraping

- Overview
  - Web scraping was performed to gather additional launch data unavailable via API
  - Used Python libraries: requests for HTTP requests, BeautifulSoup for HTML parsing
  - Targeted publicly accessible SpaceX launch information pages

- Methodology
  - Sent HTTP GET requests to SpaceX-related URLs for static and dynamic data
  - Parsed HTML content with BeautifulSoup to extract:
    - Launch dates
    - Mission details
    - Booster and payload information



Reference: *Data-Science-Assignment/jupyter-labs-webscraping.ipynb at main · rk-rohitkumar/Data-Science-Assignment*

Data Wrangling

Identify and calculate the percentage of the missing values in each attribute

Identify which columns are numerical and categorical

Calculate the number of launches on each site

Calculate the number and occurrence of each orbit

Create a landing outcome label from Outcome column

Determine the success rate

Reference: [Data-Science-Assignment/labs-jupyter-spacex-Data wrangling.ipynb at main · rk-rohitkumar/Data-Science-Assignment]

EDA with Data Visualization

- Performed exploratory data analysis (EDA) and feature engineering on SpaceX Falcon 9 launch data using several charts for key insights
  - Scatter plots of Flight Number vs. Payload Mass , Flight Number vs. Launch Site , Payload Mass vs. Launch Site
  - A bar chart of success rates by Orbit type summarized which orbit types yielded the highest booster landing success rates.
  - Additional scatter plots explored relationships between Payload Mass and Orbit type and between Flight Number and Orbit type, providing insights into performance by orbit category.
  - A line chart of yearly average success rate displayed temporal trends in successful booster landings over the years, showing how success rates improved or varied with time.
- These charts were chosen to visualize launch outcomes concerning key variables like flight progression, launch site, payload mass, and orbit type to detect patterns and factors affecting first stage landing success, aiding predictive model development. The use of scatter plots with categorical hues and bar/line charts helped analyze both distribution and aggregated success metrics effectively

Reference: Data-Science-Assignment/edadataviz.ipynb at main · rk-rohitkumar/Data-Science-Assignment

EDA with SQL

- Performed below EDA using SQL
  - Display distinct launch sites
  - Display 5 rerecords 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 succesful landing outcome in ground pad was acheived
  - 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 all the booster_versions that have carried the maximum payload mass, using a subquery with a suitable aggregate function.
  - 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.
  - 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.
  - cords where launch sites begin with the string 'CCA'

Reference: [Data-Science-Assignment/jupyter-labs-eda-sql-coursera_sqllite.ipynb at main · rk-rohitkumar/Data-Science-Assignment]

Build an Interactive Map with Folium

- Markers and Circles:
  - Each SpaceX launch site's geographic coordinates (latitude and longitude) were used to place markers and highlighted circles on the map. Circles were drawn around the launch sites to visually emphasize their locations with a specified radius and popup labels showing the site names.
- DivIcon Markers:
  - Custom text labels were added to the map at the launch site coordinates using Folium's DivIcon to provide visible site names next to the markers.
- Color-coded Markers:
  - Launch success or failed outcomes for each site were represented using color-labeled markers on the map, helping visually differentiate success versus failure by location.
- Distance Lines:
  - Lines or measured distances were calculated and shown between launch sites and nearby reference points such as railways, highways, and coastlines, to understand proximity and possible geographic impact factors on launch success.

Reference: [Data-Science-Assignment/lab_jupyter_launch_site_location.ipynb at main · rk-rohitkumar/Data-Science-Assignment]

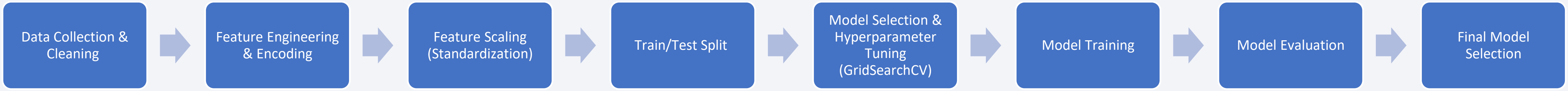
Build a Dashboard with Plotly Dash

- Launch Success Pie Charts:
  - Provide an easy visual summary of the performance distribution among launch sites, highlighting where launches are most successful.
- Scatter Plots with Interactive Filters (payload and flight number):
  - Allow users to dynamically explore how payload weight and launcher experience relate to success, identifying key factors for booster landings.
- Bar and Line Charts capture temporal and categorical trends:
  - Essential for understanding success rates by orbit type and over time, supporting strategic decision-making.
- Folium Maps and Distance Calculations:
  - Bring spatial awareness to analysis, useful for planning optimal launch locations and understanding geographical constraints.
- Interactivity enables users to explore data layers and filter views, enhancing comprehension and engagement with the analytic findings.

Reference: Data-Science-Assignment/spacex-dash-app.py at main · rk-rohitkumar/Data-Science-Assignment

Predictive Analysis (Classification)

- Model Building:
  - Built classification models including Logistic Regression, Support Vector Machine (SVM), and Decision Tree classifiers.
  - Used GridSearchCV with 10-fold cross-validation to tune hyperparameters.
- Model Evaluation:
  - Selected the best hyperparameters based on cross-validation accuracy.
  - Evaluated models on the test data using accuracy scores.
  - Analyzed confusion matrices for detailed performance metrics.
- Best Model Selection:
  - Compared models' performances.
  - The model with the highest test accuracy was selected as the best performing model.



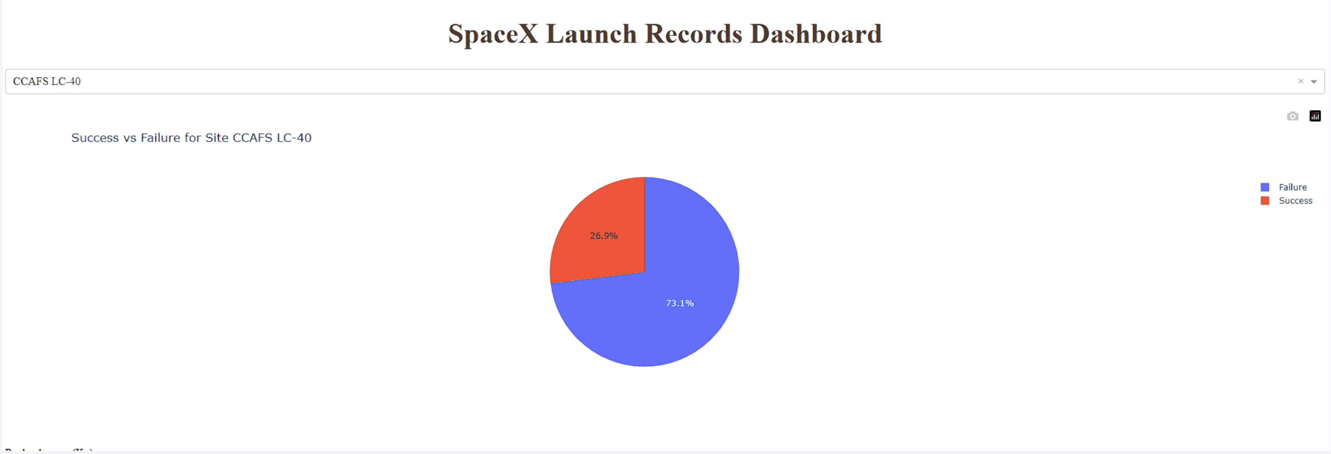
Reference: [Data-Science-Assignment/SpaceX_Machine Learning Prediction_Part_5.ipynb at main · rk-rohitkumar/Data-Science-Assignment]()

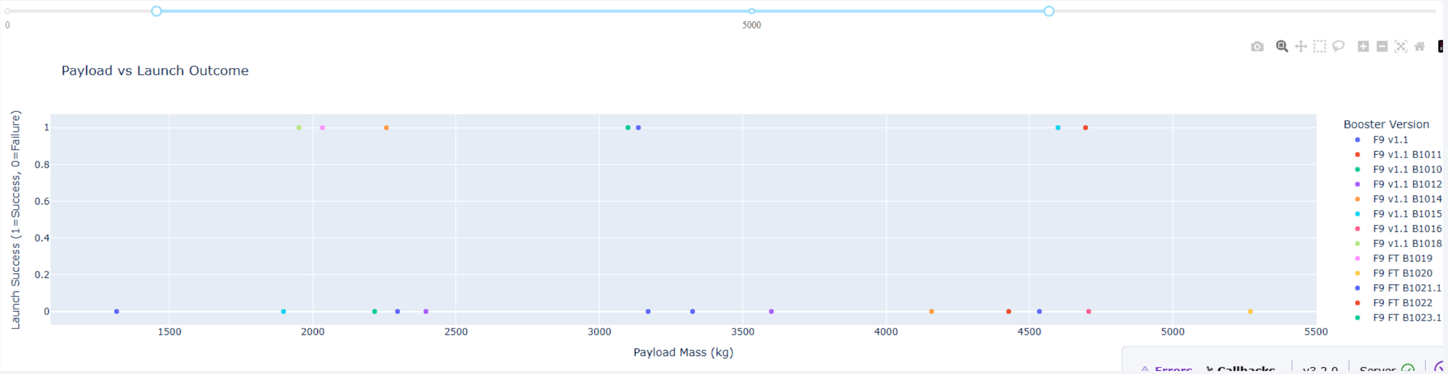
Results

- Exploratory data analysis results
    - As the flight number increases, the first stage is more likely to land successfully.
    - The payload mass also appears to be a factor; even with more massive payloads, the first stage often returns successfully.
    - For the VAFB-SLC Launch Site there are no rockets launched for heavy payload mass.
    - The success rate since 2013 kept increasing till 2020.
    - ES-L1, GEO, HEO and SSO Orbits have the highest success rate.
    - In the LEO orbit, success seems to be related to the number of flights. Conversely, in the GTO orbit, there appears to be no relationship between flight number and success.

Results

- Interactive analytics demo in screenshots





Results

- Predictive analysis results

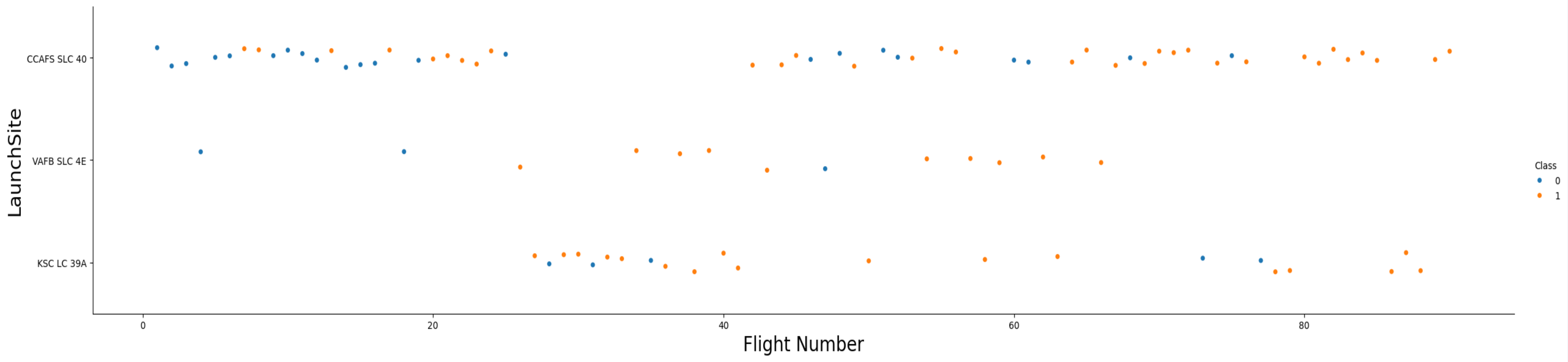
Model	Training Accuracy	Test
LogisticRegression	84.64%	83.33%
SVM	84.21%	83.33%
DecisionTreeClassifier	87.68%	83.33%
KNeighborsClassifier	84.21%	83.33%

Section 2

Insights drawn from EDA

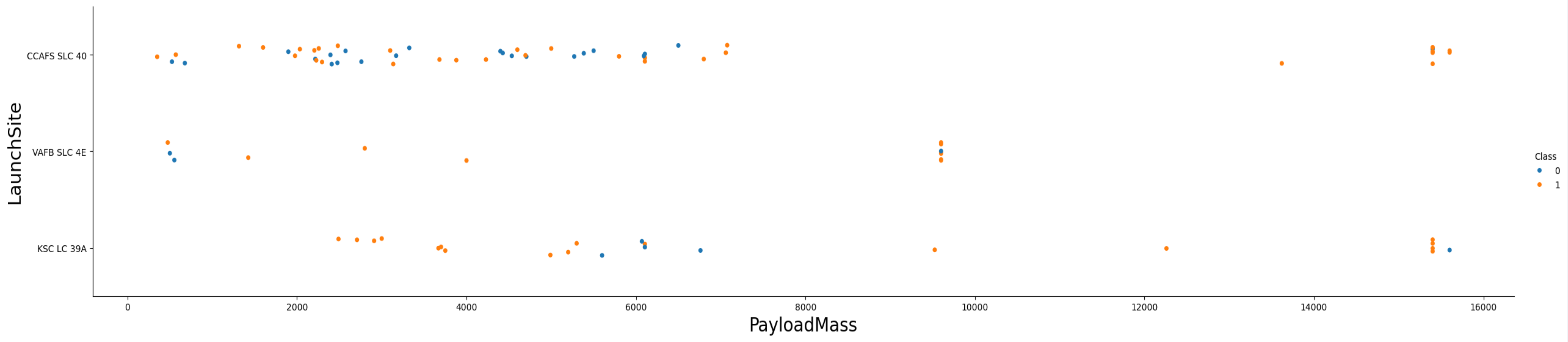
Flight Number vs. Launch Site

- More recent flights show higher success rates, regardless of launch site, indicating significant learning and technological improvement.



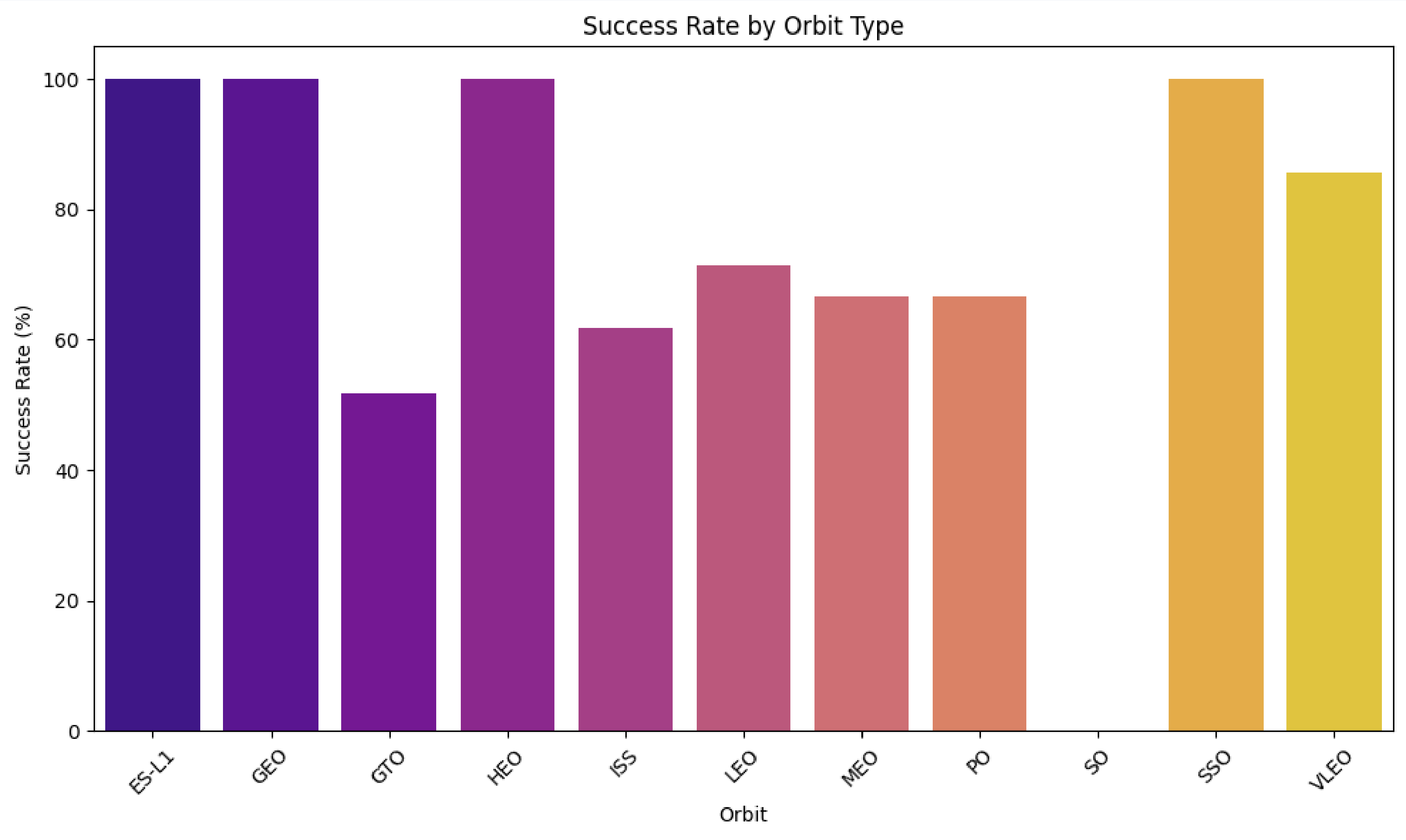
Payload vs. Launch Site

• CCAFS SLC 40 site is the most frequently used, supporting a broad payload spectrum.
• Falcon 9's landing success is not constrained by heavier payloads.
• There is no simple linear relationship between payload mass and landing outcome; both high and low payloads can succeed or fail depending on other factors.
• KSC LC 39A and VAFB SLC 4E sites generally handle heavier payloads more successfully



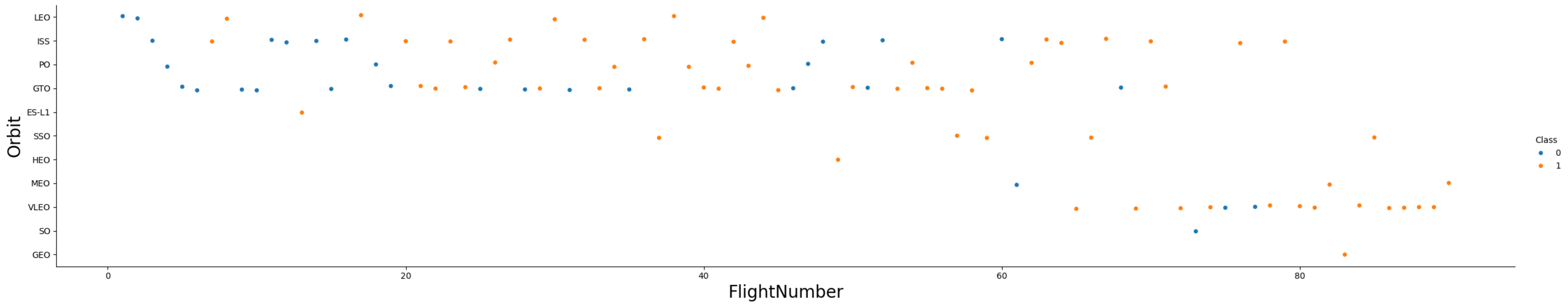
Success Rate vs. Orbit Type

- **ES-L1, GEO, HEO, and SSO all have a 100% success rate**—every attempt for these orbit types resulted in a successful launch.
- **VLEO** also demonstrates a very high success rate.
- **GTO (Geostationary Transfer Orbit)** has the lowest reported success rate in the chart, nearly half of all attempts faced issues.



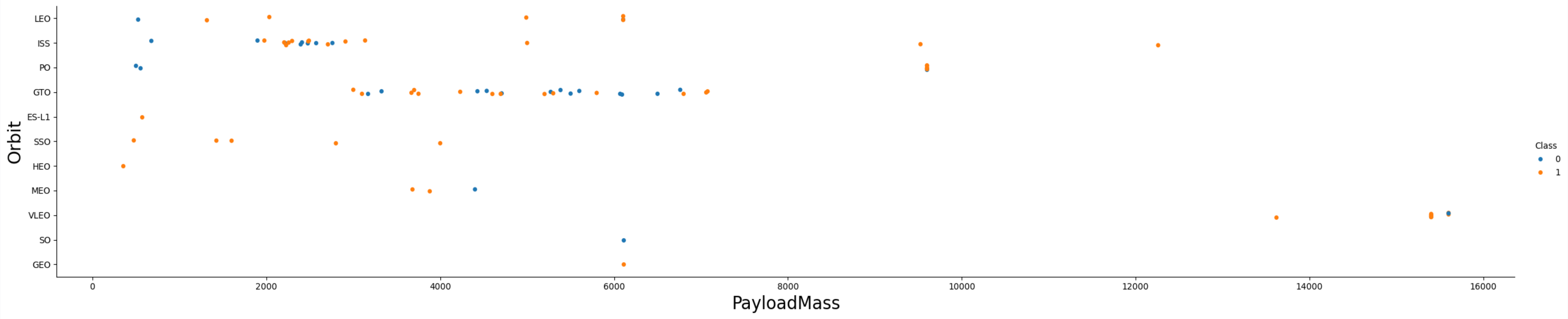
Flight Number vs. Orbit Type

- In the LEO orbit, success seems to be related to the number of flights. Conversely, in the GTO orbit, there appears to be no relationship between flight number and success.



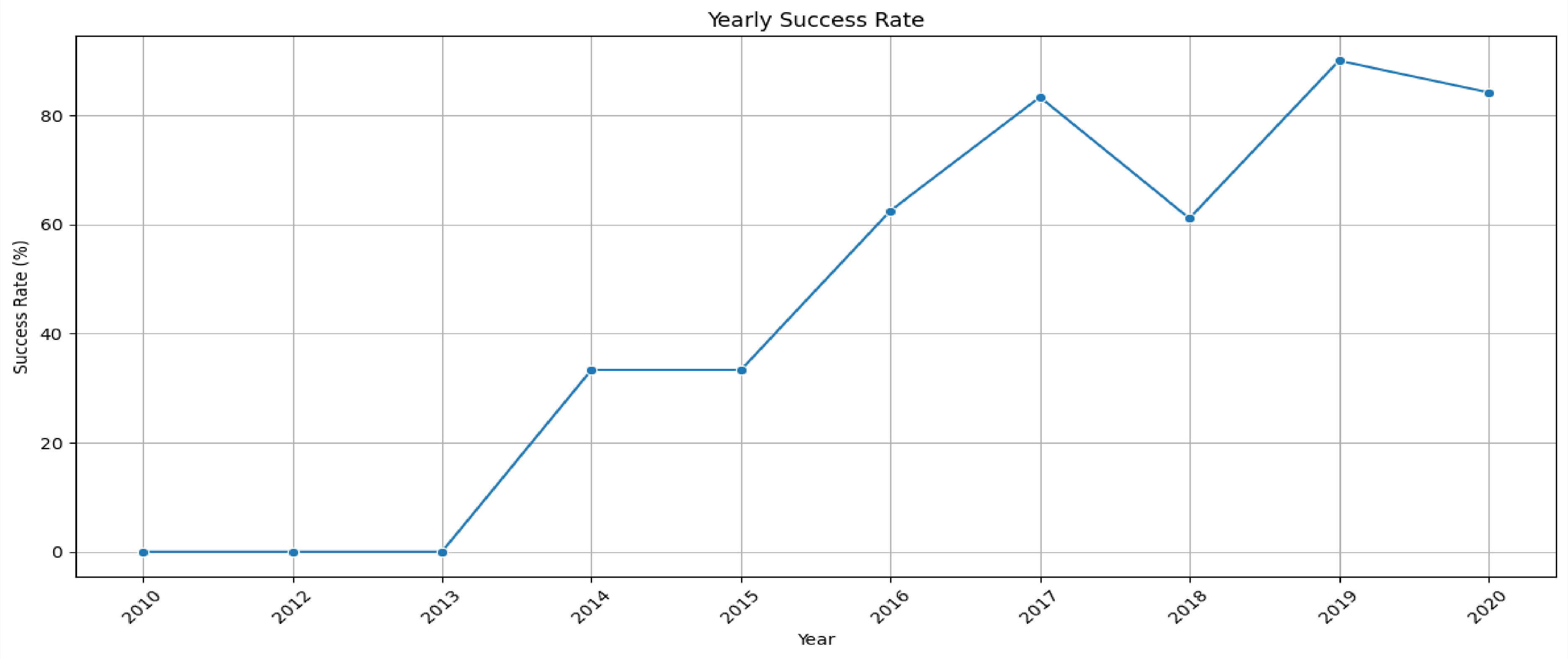
Payload vs. Orbit Type

- The majority of high-mass launches are concentrated in a few orbit types.
- Mission outcomes are clearly influenced by orbit type, not just payload mass



Launch Success Yearly Trend

Success rate since 2013 kept increasing.



All Launch Site Names

Names of the unique launch sites:

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

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

```
 * sqlite:///my_data1.db
Done.
```

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

Explanation:

SELECT DISTINCT is used to select unique (non-duplicate) values from the column Launch_Site.

FROM SPACEXTABLE specifies the table where the launch data is stored.

This query returns a list of all distinct launch site names occurring in the SpaceX missions dataset, helping identify all different launch locations used.

Launch Site Names Begin with 'CCA'

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

[17]:
```
%sql SELECT * FROM SPACEXTABLE WHERE Launch_Site LIKE 'CCA%' LIMIT 5;
```

```
 * sqlite:///my_data1.db
Done.
```

[17]:

Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer	Mission_Outcome	Landing_Outcome
2010-06-04	18:45:00	F9 v1.0 B0003	CCAFS LC-40	Dragon Spacecraft Qualification Unit	0	LEO	SpaceX	Success	Failure (parachute)
2010-12-08	15:43:00	F9 v1.0 B0004	CCAFS LC-40	Dragon demo flight C1, two CubeSats, barrel of Brouere cheese	0	LEO (ISS)	NASA (COTS) NRO	Success	Failure (parachute)
2012-05-22	7:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525	LEO (ISS)	NASA (COTS)	Success	No attempt
2012-10-08	0:35:00	F9 v1.0 B0006	CCAFS LC-40	SpaceX CRS-1	500	LEO (ISS)	NASA (CRS)	Success	No attempt
2013-03-01	15:10:00	F9 v1.0 B0007	CCAFS LC-40	SpaceX CRS-2	677	LEO (ISS)	NASA (CRS)	Success	No attempt

Explanation:

- SELECT *: Selects all columns from the table.
- FROM SPACEXTABLE: Specifies the table named SPACEXTABLE containing SpaceX launch data.
- WHERE Launch_Site LIKE 'CCA%': Filters rows where the Launch_Site column's value starts with the characters "CCA".
  - The LIKE operator enables pattern matching using:
  - % wildcard = zero or more characters.
  - 'CCA%' means any value beginning with "CCA" followed by any characters.
- LIMIT 5: Limits the output to only 5 rows.

Total Payload Mass

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

```
[19]: %sql SELECT SUM(PAYLOAD_MASS__KG_) AS Total_Payload_Mass FROM SPACEXTABLE WHERE Customer = 'NASA (CRS)';
```

```
 * sqlite:///my_data1.db
Done.
```

[19]:

Total_Payload_Mass
45596

Explanation:

SELECT SUM(PAYLOAD_MASS__KG_): The SUM() function calculates the total sum of the numeric values in the column PAYLOAD_MASS__KG_.

AS Total_Payload_Mass: This assigns an alias name Total_Payload_Mass to the resulting sum column for clearer output.

FROM SPACEXTABLE: The data is queried from the table named SPACEXTABLE.

WHERE Customer = 'NASA (CRS)': This filter condition restricts the rows to only those where the Customer column value is exactly 'NASA (CRS)'

Average Payload Mass by F9 v1.1

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

```
[20]: %sql SELECT AVG(PAYLOAD_MASS__KG_) AS Average_Payload_Mass FROM SPACEXTABLE WHERE Booster_Version = 'F9 v1.1';
```

```
 * sqlite:///my_data1.db
Done.
```

[20]:

Average_Payload_Mass
2928.4

Explanation:

SELECT AVG(PAYLOAD_MASS__KG_): The AVG() function calculates the average (mean) of all non-null values in the column PAYLOAD_MASS__KG_.

AS Average_Payload_Mass: Renames the output column to Average_Payload_Mass for clearer display.

FROM SPACEXTABLE: Specifies the table containing the SpaceX launch data.

WHERE Booster_Version = 'F9 v1.1': Filters the rows to include only launches where the booster version matches 'F9 v1.1'.

First Successful Ground Landing Date

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

Hint:Use min function

```
[24]: %sql SELECT Date FROM SPACEXTABLE WHERE Landing_Outcome = 'Success (ground pad)'ORDER BY Date ASC LIMIT 1;
```

```
 * sqlite:///my_data1.db
Done.
```

[24]:

Date
2015-12-22

Explanation:

SELECT Date: Selects the Date column from the table.

FROM SPACEXTABLE: Queries from the table named SPACEXTABLE.

WHERE Landing_Outcome = 'Success (ground pad)': Filters to only include rows where the Landing_Outcome matches exactly 'Success (ground pad)'.

ORDER BY Date ASC: Orders the resulting rows by the Date column in ascending order (earliest date first).

LIMIT 1: Limits the output to only the first row of the ordered results.

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

[25]:
```
%sql SELECT Booster_Version FROM SPACEXTABLE WHERE Landing_Outcome = 'Success (drone ship)' AND PAYLOAD_MASS__KG_ > 4000 AND PAYLOAD_MASS__KG_ < 6000;
```

```
 * sqlite:///my_data1.db
Done.
```

[25]:

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

Explanation:

SELECT Booster_Version: Retrieves the booster version names from the dataset.

FROM SPACEXTABLE: Specifies the SpaceX launch data table.

WHERE Landing_Outcome = 'Success (drone ship)': Filters records where the booster successfully landed on a drone ship.

AND PAYLOAD_MASS__KG_ > 4000 AND PAYLOAD_MASS__KG_ < 6000: Further filters to include only launches with payload masses between 4000 kg and 6000 kg.

Total Number of Successful and Failure Mission Outcomes

List the total number of successful and failure mission outcomes

[28]:
```
%sql SELECT Mission_Outcome, COUNT(*) AS Total FROM SPACEXTABLE GROUP BY Mission_Outcome;
```

```
 * sqlite:///my_data1.db
Done.
```

[28]:

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

Explanation:

SELECT Mission_Outcome: Selects the column Mission_Outcome which contains different mission result statuses.

COUNT(*) AS Total: Counts the number of rows (launches) for each unique mission outcome and assigns the name Total to this count column.

FROM SPACEXTABLE: Specifies the table holding the SpaceX launch data.

GROUP BY Mission_Outcome: Groups the data by the distinct values in the Mission_Outcome column before counting.

Boosters Carried Maximum Payload

List all the booster_versions that have carried the maximum payload mass, using a subquery with a suitable aggregate function.

```
[32]: %sql SELECT DISTINCT Booster_Version FROM SPACEXTABLE WHERE PAYLOAD_MASS__KG_ = (SELECT MAX(PAYLOAD_MASS__KG_) FROM SPACEXTABLE);
```

```
 * sqlite:///my_data1.db
Done.
```

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

Explanation:

SELECT DISTINCT Booster_Version: Selects the unique booster versions from the results.

FROM SPACEXTABLE: The table with SpaceX launch data.

WHERE PAYLOAD_MASS__KG_ = (SELECT MAX(PAYLOAD_MASS__KG_) FROM SPACEXTABLE):

The subquery (SELECT MAX(PAYLOAD_MASS__KG_) FROM SPACEXTABLE) returns the maximum payload mass value present in the entire table.

The outer query filters all rows where the PAYLOAD_MASS__KG_ equals this maximum value.

The DISTINCT clause ensures that if multiple rows have this max payload mass, only unique booster versions are returned.

2015 Launch Records

```
%sql SELECT substr(Date, 6, 2) AS Month, Landing_Outcome, Booster_Version,  Launch_Site FROM SPACEXTABLE WHERE Landing_Outcome = 'Failure (drone ship)
    AND substr(Date, 0, 5) = '2015';
```

```
 * sqlite:///my_data1.db
Done.
```

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

Explanation:

SELECT substr(Date, 6, 2) AS Month: Extracts a substring from the Date column, starting at position 6 with length 2, which corresponds to the month part of a date formatted as YYYY-MM-DD. This is aliased as Month.

Landing_Outcome, Booster_Version, Launch_Site: Selects these additional columns to display.

FROM SPACEXTABLE: Uses the dataset table named SPACEXTABLE.

WHERE Landing_Outcome = 'Failure (drone ship)': Filters rows to include only launch records where the landing outcome was a failure on a drone ship.

AND substr(Date, 0, 5) = '2015': Further filters to rows where the year part of the date equals '2015' by extracting the first 4 characters

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.

[38]:
```
%sql SELECT Landing_Outcome, COUNT(*) AS Outcome_Count FROM SPACEXTABLE WHERE Date BETWEEN '2010-06-04' AND '2017-03-20' GROUP BY Landing_Outcome ORDER BY Outcome_Count DESC;
```

```
 * sqlite:///my_data1.db
Done.
```

Explanation:

SELECT Landing_Outcome: Retrieves distinct values of the Landing_Outcome column that describe the landing results.

COUNT(*) AS Outcome_Count: Counts the number of rows for each unique Landing_Outcome and names the count column Outcome_Count.

FROM SPACEXTABLE: Specifies the table containing the SpaceX launch data.

WHERE Date BETWEEN '2010-06-04' AND '2017-03-20': Filters records to include only those with a Date between (and including) June 4, 2010, and March 20, 2017.

GROUP BY Landing_Outcome: Groups the filtered result rows by different landing outcomes for aggregation.

ORDER BY Outcome_Count DESC: Sorts the grouped results by the count in descending order so the most frequent outcome appears first.

[38]:

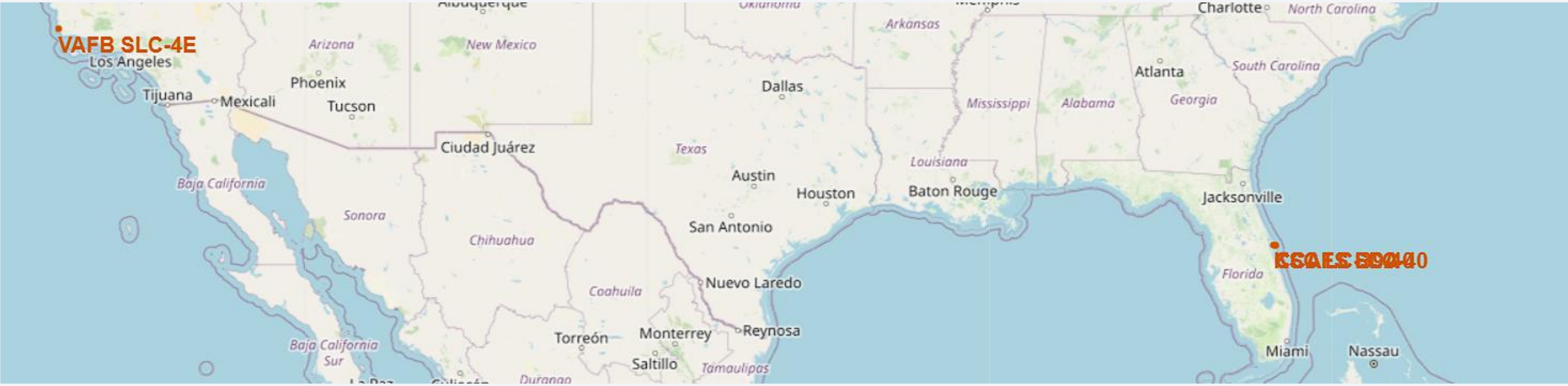
Landing_Outcome	Outcome_Count
No attempt	10
Success (drone ship)	5
Failure (drone ship)	5
Success (ground pad)	3
Controlled (ocean)	3
Uncontrolled (ocean)	2
Failure (parachute)	2
Precluded (drone ship)	1

Section 3

Launch Sites Proximities Analysis

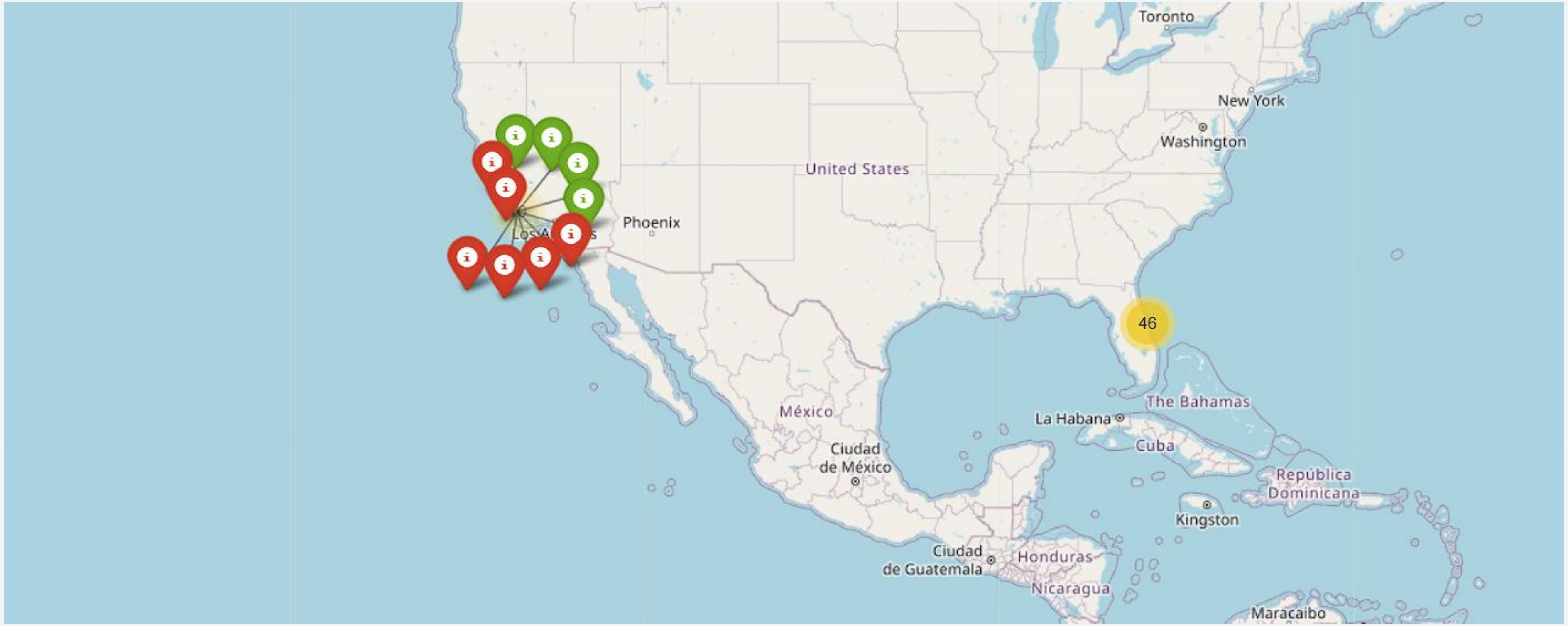
Geographic Distribution of SpaceX Launch Sites

All SpaceX launch sites in the US. They are close to major coastlines but not close to the Equator.
This reflects a balance of safety and operational efficiency



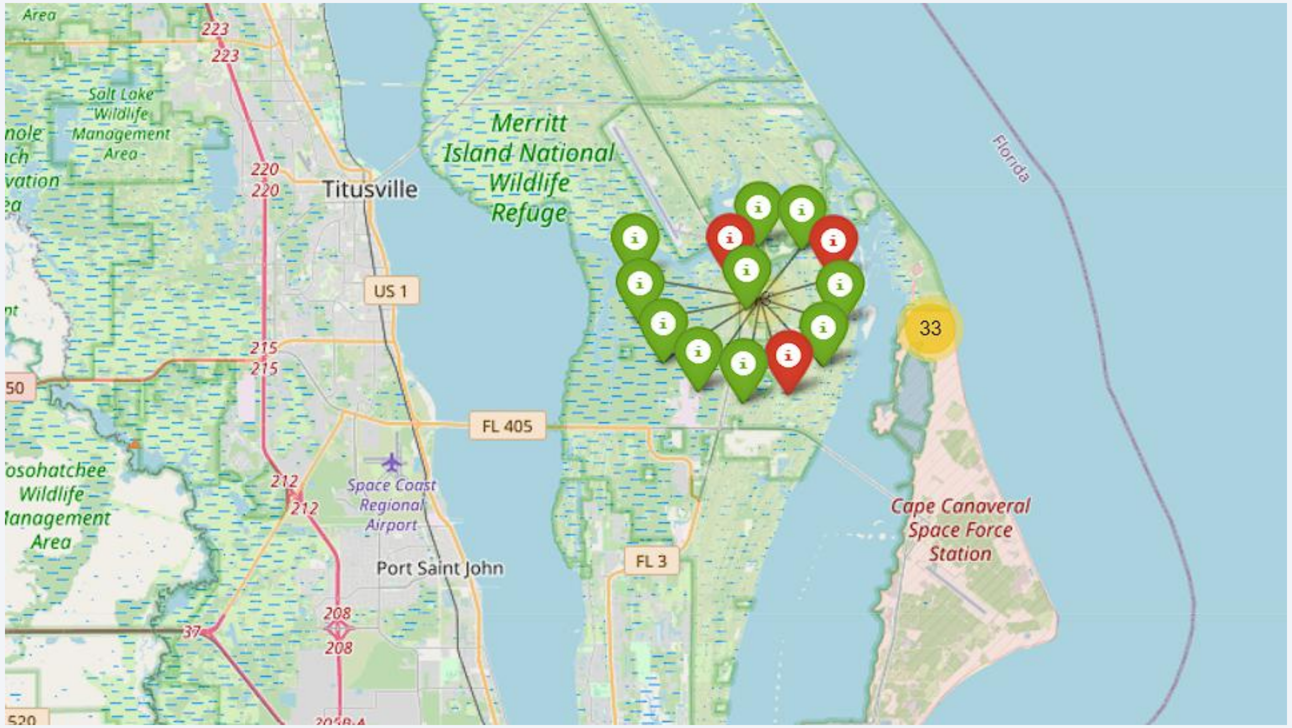
Launch Outcome of Site: VAFB SLC-4E

Fewer launches and low success rate on this Site



Launch Outcome of Site: KSC LC-39A

Very High Success rate and a lot more launches overall.



Launch Outcome of Site: CCAFS SLC-40

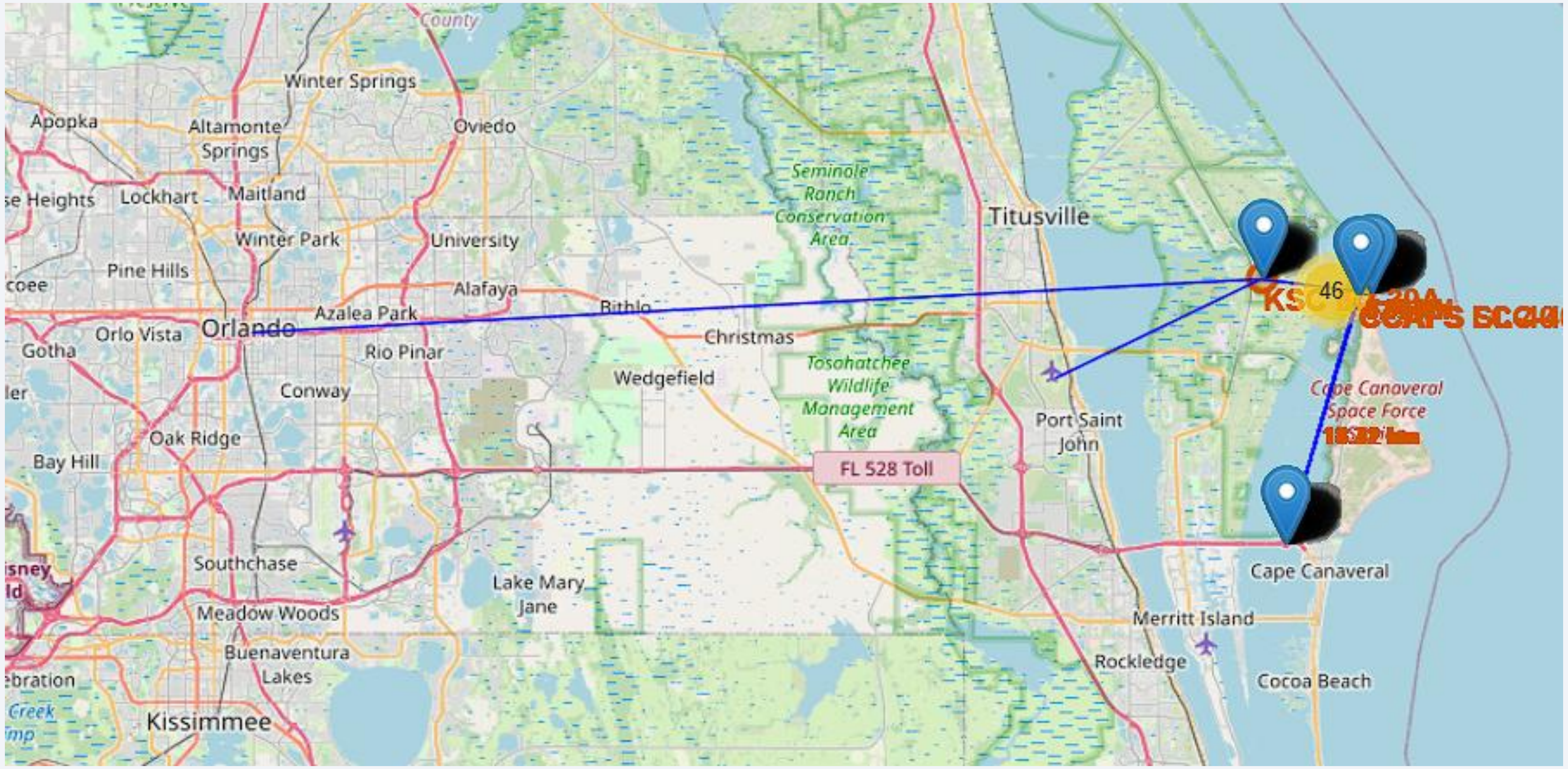
Low Success rate and very few launches overall.

Launch Outcome of Site: CCAFS LC-40

Very Low Success rate and significant launches overall.

Proximity of Launch sites

Launch sites are at least 10 km from highway, airports and cities

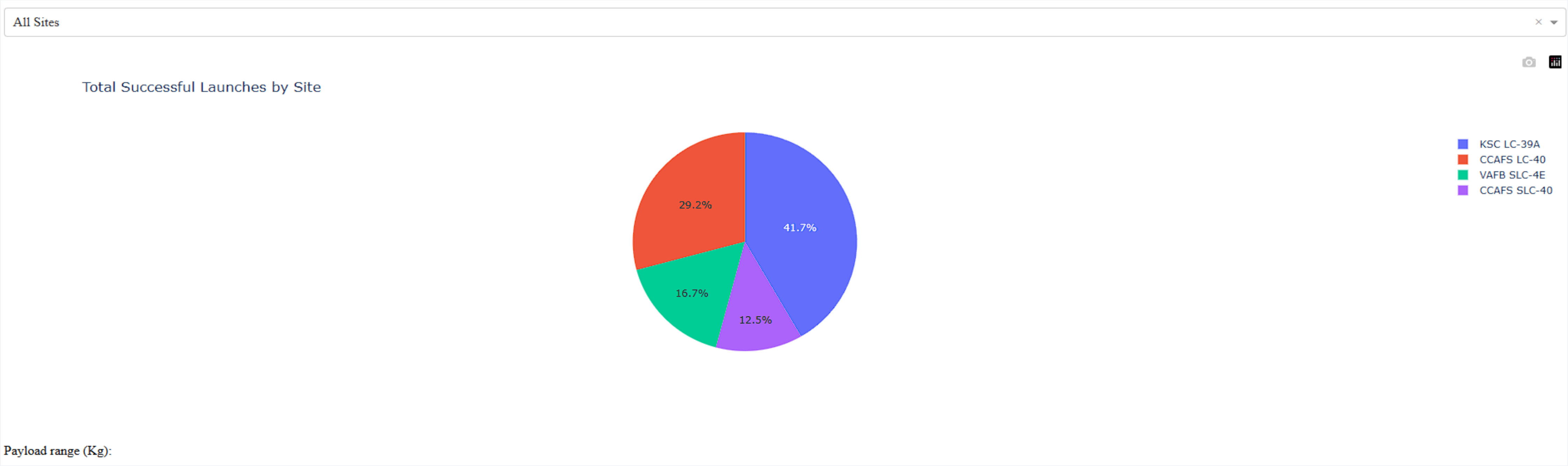


Section 4

Build a Dashboard with Plotly Dash

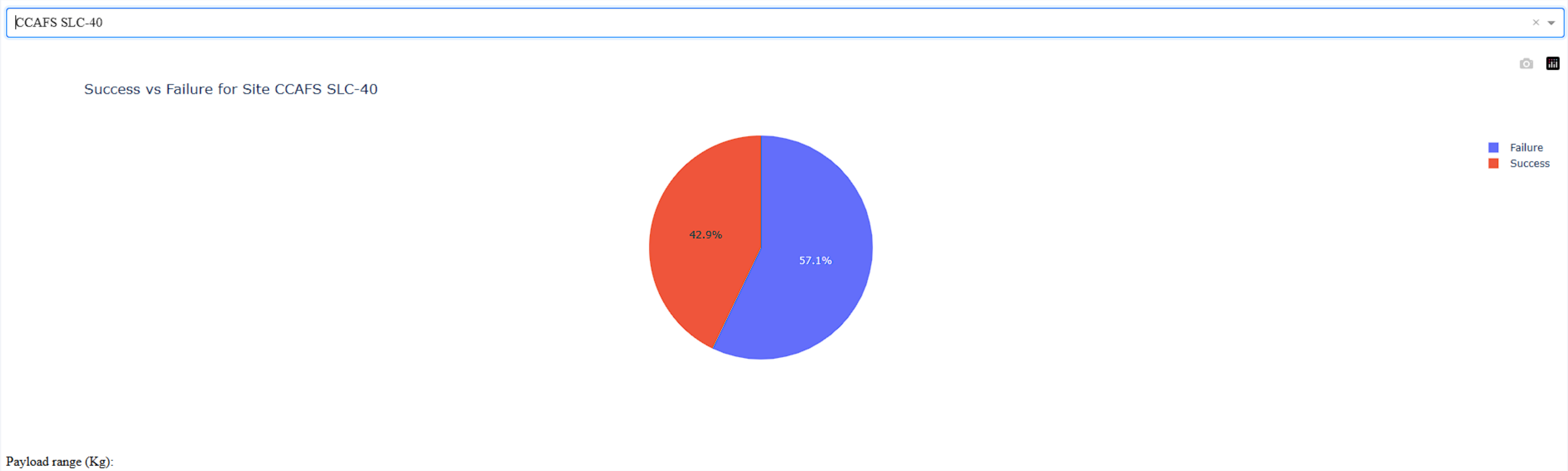
Important Launch Sites

KSC LC-39A accounts for the largest share of successful launches at 41.7% of the total, illustrating its critical role in SpaceX operations.



Highest success rate

CCAFS SLC-40 site has the highest success rate.



Success by payloads

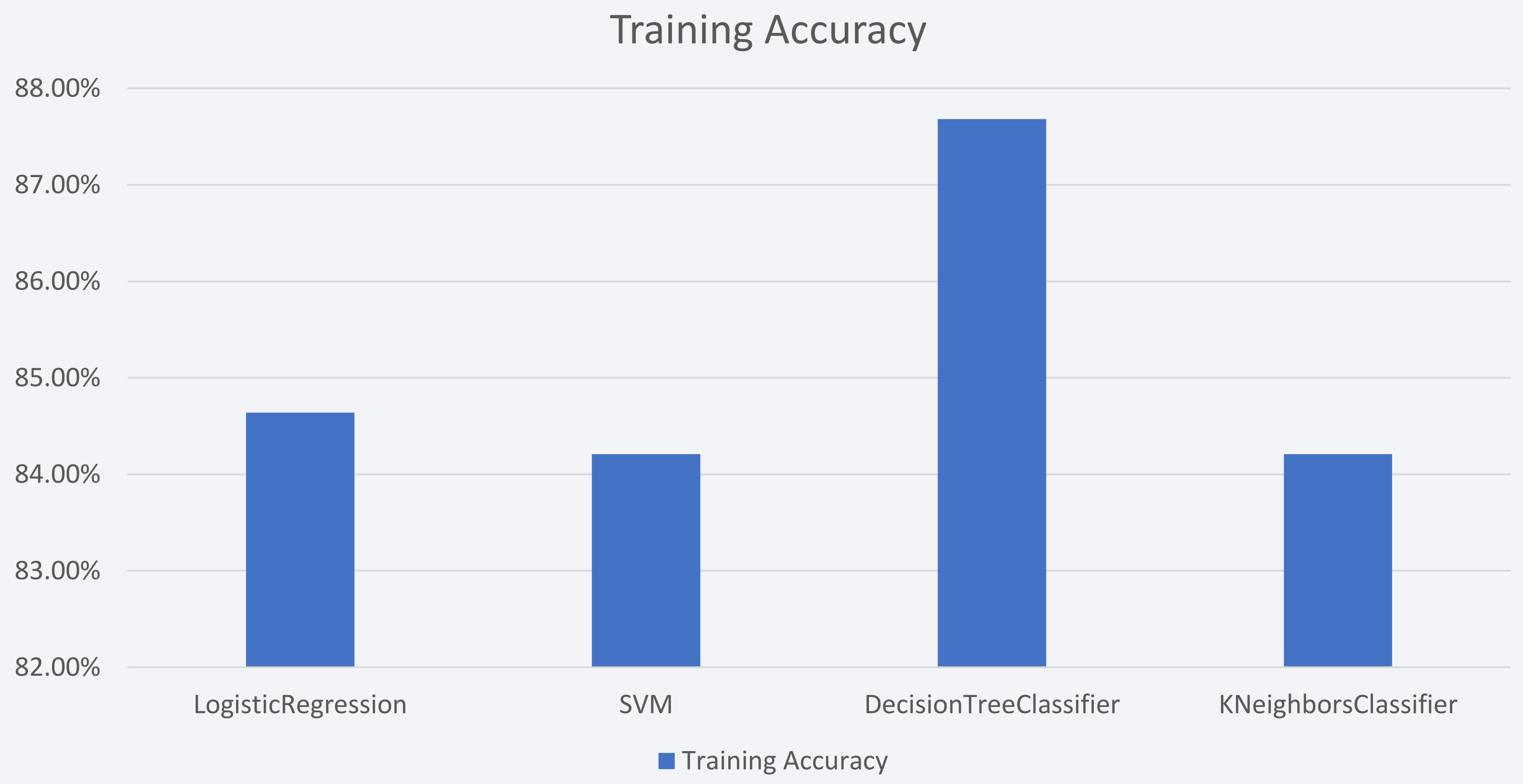
Most successful launches are scattered across the payload range from ~2000 kg up to about 6000 kg, indicating that SpaceX boosters deliver reliable performance for a wide range of payload masses

Section 5

Predictive Analysis (Classification)

Classification Accuracy

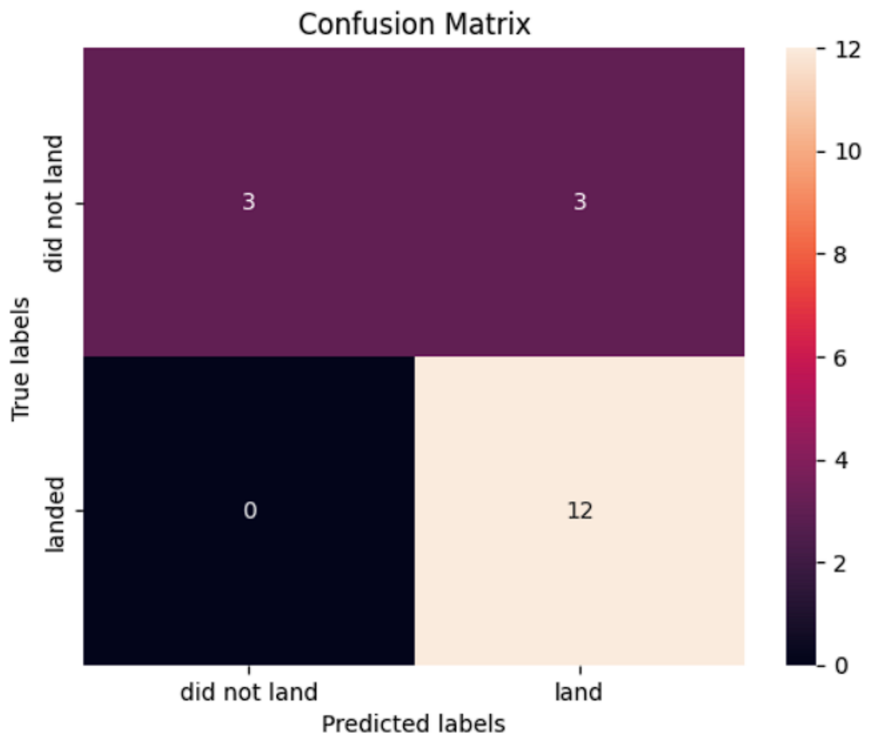
Decision Tree Classifier has the highest accuracy.



Confusion Matrix

The classifier excels at identifying successful landings but is prone to misclassifying some failures as successes.

In [46]:
```
yhat = tree_cv.predict(X_test)
plot_confusion_matrix(Y_test,yhat)
```



Conclusions

- Launch Site Performance:
  - KSC LC-39A contributes the highest share (~42%) of successful launches, underscoring its strategic significance.
  - CCAFS SLC-40 shows the highest success rate despite fewer launches.
  - VAFB SLC-4E and CCAFS LC-40 see lower success rates and fewer launches in comparison.
- Payload and Orbit Influence:
  - Successful launches cover a wide payload mass range (~2000 to 6000+ kg), demonstrating booster versatility.
  - Orbit types ES-L1, GEO, HEO, and SSO achieve near 100% success, highlighting mission-specific reliability.
  - GTO orbit launches exhibit relatively lower success rates, indicating challenges or higher risks.
- Temporal Trends:
  - Overall launch success rates have improved steadily from 2013 to 2020, reflecting technological and operational advancements.
- Flight Experience Impact:
  - Increased flight numbers correlate positively with landing success for orbits like LEO, though not for GTO.

Thank you!