IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Saad Alhuthali
14/03/2025

Outline

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

Executive Summary

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

Introduction

- Project background and context:

  This test can be improved by focusing on predicting the cost of a SpaceX launch based on reusability. We can leverage public data on past launches and machine learning to identify patterns that indicate a reusable first stage (like launch location or payload weight). If the model predicts reusability, the launch cost is likely around $62 million. Conversely, a predicted non-reusable launch suggests a higher cost exceeding $165 million. This approach directly addresses the cost prediction goal, using SpaceX's reusability as the key factor.

- Problems you want to find answers:
  - What are the key mission parameters and environmental factors that influence a successful rocket landing?
  - How do different aspects of the mission, like launch trajectory, booster performance, and weather conditions, interact to influence the success rate of a rocket landing?
  - What operational considerations, from pre-launch checks to flight control strategies and landing site preparation, are crucial for a successful rocket landing program?

Section 1

Methodology

Methodology

Executive Summary

- Data collection methodology:
  - Data was collected using SpaceX API and web scraping from Wikipedia.
- Perform data wrangling
  - Data wrangling prepares data for machine learning. We'll analyze SpaceX launch data to identify successful booster.
- Perform exploratory data analysis (EDA) using visualization and SQL
- Perform interactive visual analytics using Folium and Plotly Dash
- Perform predictive analysis using classification models
  - By analyzing launch data, we'll build a model to predict if the first stage lands (saving millions) compared to competitor costs. We'll explore data, define success labels, and train models (SVM, Trees, Logistic Regression) to find the best predictor.

Data Collection

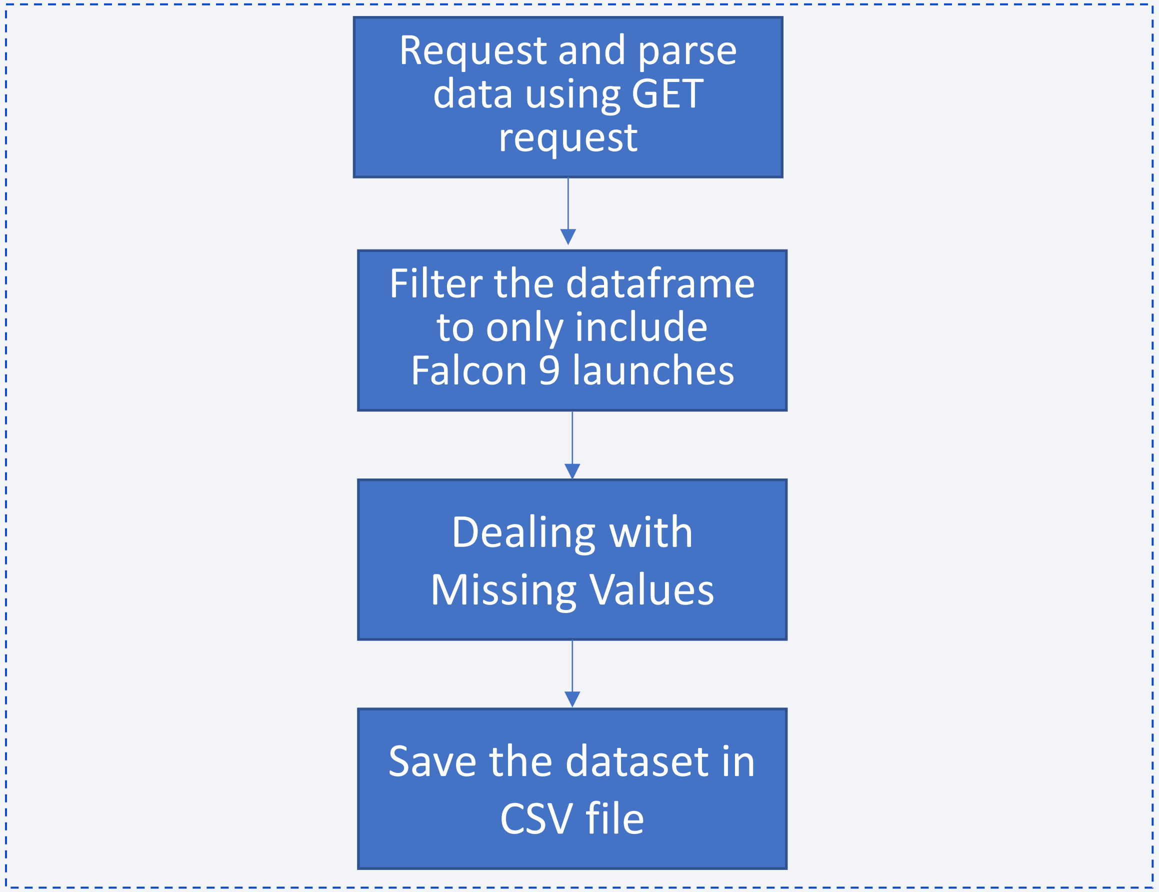
The data sets are collected using 2 methods:

1. **Request to the SpaceX API:**
    - Gathered SpaceX's past launch data via their open-source API.
    - Retrieved and processed this data with GET request.
    - Ensured the data included only Falcon 9 launches.
    - Filled in missing payload weights from secret missions with average values.
2. **Web Scraping:**
    - Requested past Falcon 9 and Falcon Heavy launch data from Wikipedia's relevant page.
    - Accessed the Falcon 9 Launch page via its direct Wikipedia link.
    - Extracted all the column names from the HTML table.
    - Parsed and transformed the table into a Pandas data frame suitable for analysis.

Data Collection – SpaceX API

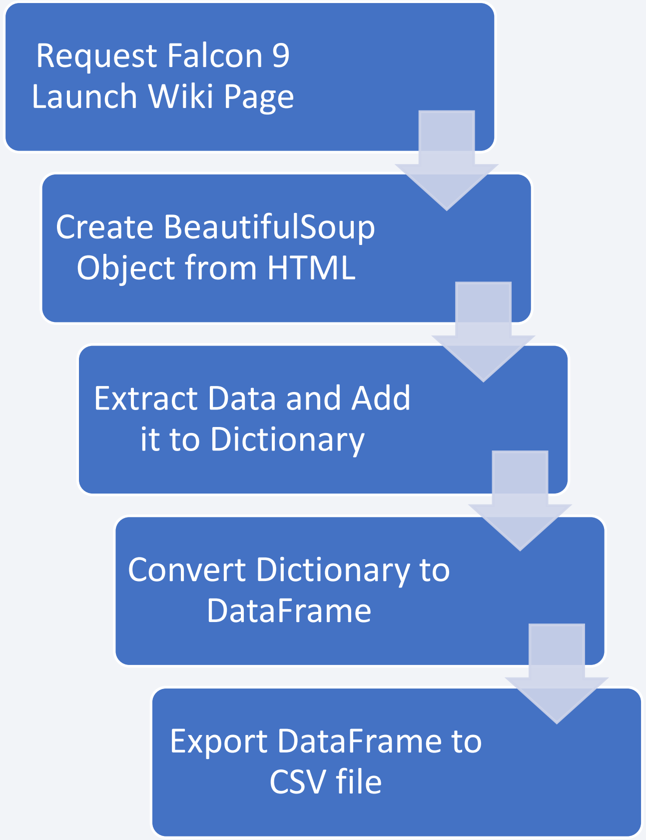
This task involves retrieving data from the SpaceX API using a GET request. The fetched data will then undergo a cleaning process to organize and standardize it for subsequent analysis or use. This cleaning process might involve handling missing values, converting data types, and ensuring consistency in formatting, ultimately making the data more suitable for further exploration and interpretation.

- GitHub URL of the completed SpaceX API calls.



Data Collection - Scraping

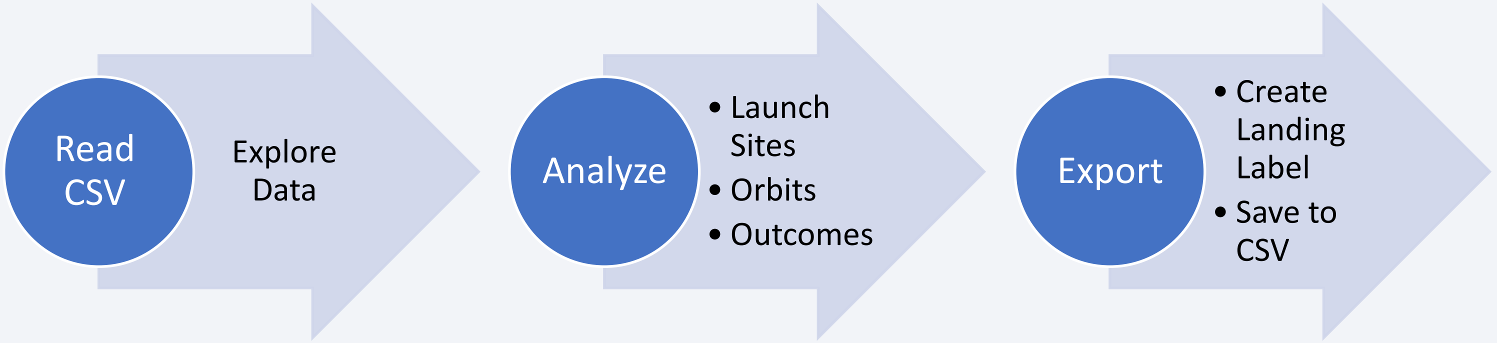
- This project extracts historical Falcon 9 launch data from a Wikipedia page. It utilizes BeautifulSoup, a Python library, to request the webpage's HTML content, parse the relevant table, and extract column names. The data from each table row is then stored in a dictionary, which is subsequently transformed into a Pandas DataFrame. Finally, the structured DataFrame containing the Falcon 9 launch records is saved as a CSV file.
- GitHub URL of the completed web scraping notebook.



Data Wrangling

This project aims to analyze SpaceX launch data, focusing on landing success prediction. By exploring patterns in launch sites, orbit types, and specific landing outcomes (ocean, ground pad, drone ship), the goal is to create a single binary label (landing_class) representing successful or unsuccessful landings. This label will be used to train machine learning models for predicting future launch outcomes. The analysis will involve identifying relevant features and transforming the data to ensure it's suitable for model training and evaluation.

- GitHub URL of your completed data wrangling related notebooks.



EDA with Data Visualization - 1

Summary of the charts plotted and their rationale:

1. Flight Number vs. Payload Mass (with Outcome Overlay): This scatter plot aimed to visualize the relationship between the number of launch attempts, payload mass, and launch outcome (success or failure). The goal was to see if more launches or heavier payloads correlated with higher or lower success rates.
2. Launch Site vs. Success Rate (Bar Chart): This chart compared the success rates of launches from different SpaceX launch sites. The purpose was to identify if certain sites were more conducive to successful landings than others.
3. Flight Number vs. Launch Site (Scatter Plot): This plot aimed to show the distribution of launches over time at each site. The objective was to see if the success rate improved with experience at a particular launch site.
4. Payload vs. Launch Site (Scatter Plot): This chart explored the relationship between payload mass and launch site. The goal was to determine if certain launch sites handled heavier payloads more successfully.
5. Success Rate by Orbit Type (Bar Chart): This chart illustrated the success rates for different orbit types. It aimed to reveal which orbits were more challenging or easier for successful landings.

EDA with Data Visualization - 2

6. Flight Number vs. Orbit Type (Scatter Plot): This plot investigated if the success rate within a specific orbit type changed with more launch attempts. It helped to understand if experience improved outcomes for particular orbits.
7. Payload vs. Orbit Type (Scatter Plot): This chart examined the relationship between payload mass and orbit type. It aimed to see if heavier payloads were more or less successful in certain orbits.
8. Year vs. Average Success Rate (Line Chart): This chart tracked the trend of average launch success rates over time. Its purpose was to observe if there was a general improvement in success rates as SpaceX gained more experience.

Overall, the use of these various charts provided a comprehensive visual exploration of the dataset, revealing potential relationships between launch outcomes and factors like flight number, payload mass, launch site, orbit type, and time. These insights are valuable for understanding the key drivers of successful SpaceX landings.

GitHub URL of your completed EDA with data visualization notebook.

EDA with SQL

- Query unique launch sites: Retrieved distinct launch sites from the space_mission table.
- Filter launch sites by prefix: Selected records where launch sites start with 'CCA'.
- Calculate total payload mass: Summed payload mass for boosters launched by NASA (CRS).
- Calculate average payload mass: Computed average payload mass for booster version F9 v1.1.
- Find first successful ground pad landing date: Queried the earliest date of successful landing on a ground pad.
- Filter boosters by landing outcome and payload mass: Selected boosters that landed successfully on drone ships and had payload mass between 4000 and 6000.
- Count successful and failed mission outcomes: Grouped mission outcomes by type (success/failure) and counted their occurrences.
- Find booster versions with max payload mass: Used a subquery to identify booster versions that carried the maximum payload mass.
- Filter by year and landing outcome: Selected records from 2015 with failed drone ship landings, displaying month, outcome, booster version, and launch site.
- Rank landing outcomes by count: Ranked the count of landing outcomes (e.g., 'Failure (drone ship)', 'Success (ground pad)') within a specific date range, sorted in descending order.

- GitHub URL of your completed EDA with SQL notebook.

Build a Dashboard with Plotly Dash

In summary, the chosen plots and interactions were carefully selected to offer a multi-faceted understanding of SpaceX launch data, balancing high-level overview with detailed, customizable analysis, and empowering users to extract meaningful insights That dashboard leverages two key visualizations: a pie chart and a scatter plot, coupled with a dropdown menu and a slider for interactive exploration.

- **Pie Chart:** This chart primarily visualizes the success rate of SpaceX launches. When 'All Sites' is selected, it provides a clear breakdown of successful launches across all locations, immediately highlighting overall performance. Selecting a specific site shifts the focus to that location's success/failure ratio, enabling comparisons and drawing attention to potential performance variations across different launch sites.
- **Scatter Plot:** This visualization dives into the relationship between payload mass and launch success. It allows users to visually assess if heavier payloads correlate with lower success rates. The coloring by booster version further enhances this analysis, possibly revealing trends where certain booster types might be more suitable for different payload ranges.
- **Dropdown and Slider:** These interactive elements empower users to personalize their exploration of the data. The dropdown enables a targeted view of specific launch sites, while the slider offers dynamic control over the payload range considered in the scatter plot. This flexibility allows for on-the-fly hypothesis testing and in-depth exploration of the dataset.
- GitHub URL of your completed Plotly Dash lab.

Build an Interactive Map with Folium

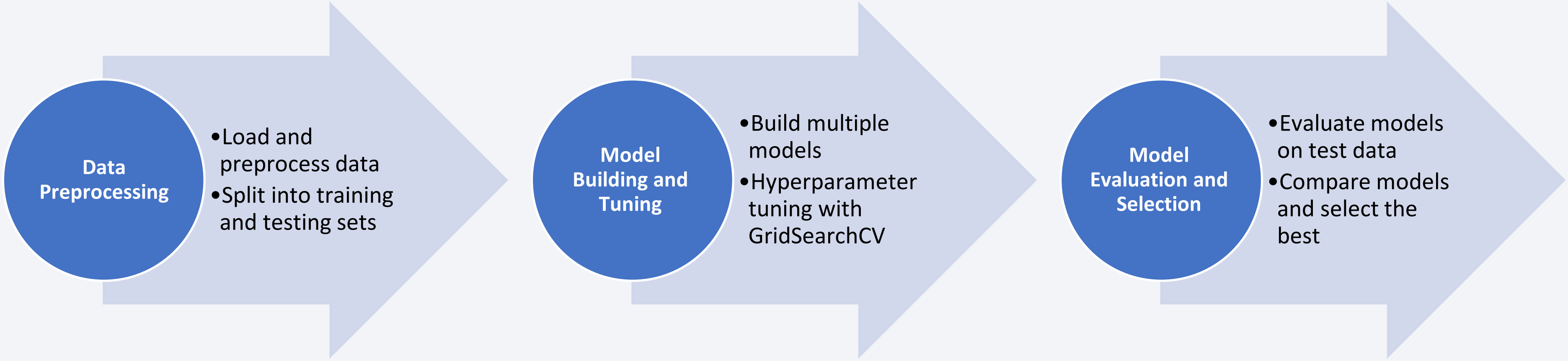
In essence, these objects transform a basic map into an interactive and informative tool. They enable easy visualization of launch site distribution, provide insights into launch success rates, and facilitate exploration of the spatial relationships between launch sites and their surroundings. that code constructs a Folium map with a few key objects that enhance its visualization and functionality:

- **Circles and Markers for Launch Sites:** These serve as the foundational elements of the map, visually pinpointing the exact locations of the various launch sites. Circles provide a larger, highlighted area for easy identification, while markers pinpoint the precise coordinates and offer the site name as a popup label.
- **Color-Coded Markers for Launch Outcomes:** For every recorded launch, a marker is added to the map, colored green for success and red for failure. This visual cue immediately conveys the launch success rate associated with each site, allowing for quick comparisons and pattern recognition.
- **Marker Clusters:** These clusters group together multiple markers at the same location, preventing map overcrowding and improving visual clarity, especially for sites with frequent launches.
- **Lines and Distance Markers:** By connecting launch sites to nearby points of interest like coastlines, cities, railways, etc., these lines and accompanying distance markers visually represent the proximity of launch sites to key infrastructure and geographical features. This facilitates analysis of potential locational factors influencing launch success.
- GitHub URL of your completed interactive map with Folium map.

Predictive Analysis (Classification)

The machine learning process involved several crucial steps, culminating in the selection of the best-performing model.

- **Data Preprocessing**: We started by preparing the data, which included converting categorical variables into numerical ones using one-hot encoding and casting all features to a consistent float64 data type.
- **Model Building and Tuning**: We then constructed multiple classification models: Logistic Regression, Support Vector Machine (SVM), Decision Tree, and K-Nearest Neighbors (KNN). To optimize their performance, we employed GridSearchCV for hyperparameter tuning, systematically searching for the best parameter combinations.
- **Model Evaluation and Selection**: We evaluated each model's accuracy on the test data and used confusion matrices to gain insights into their strengths and weaknesses. Finally, we compared the models based on their accuracy scores and selected the one with the highest performance, which turned out to be the Support Vector Machine.
- GitHub URL of your completed predictive analysis lab.



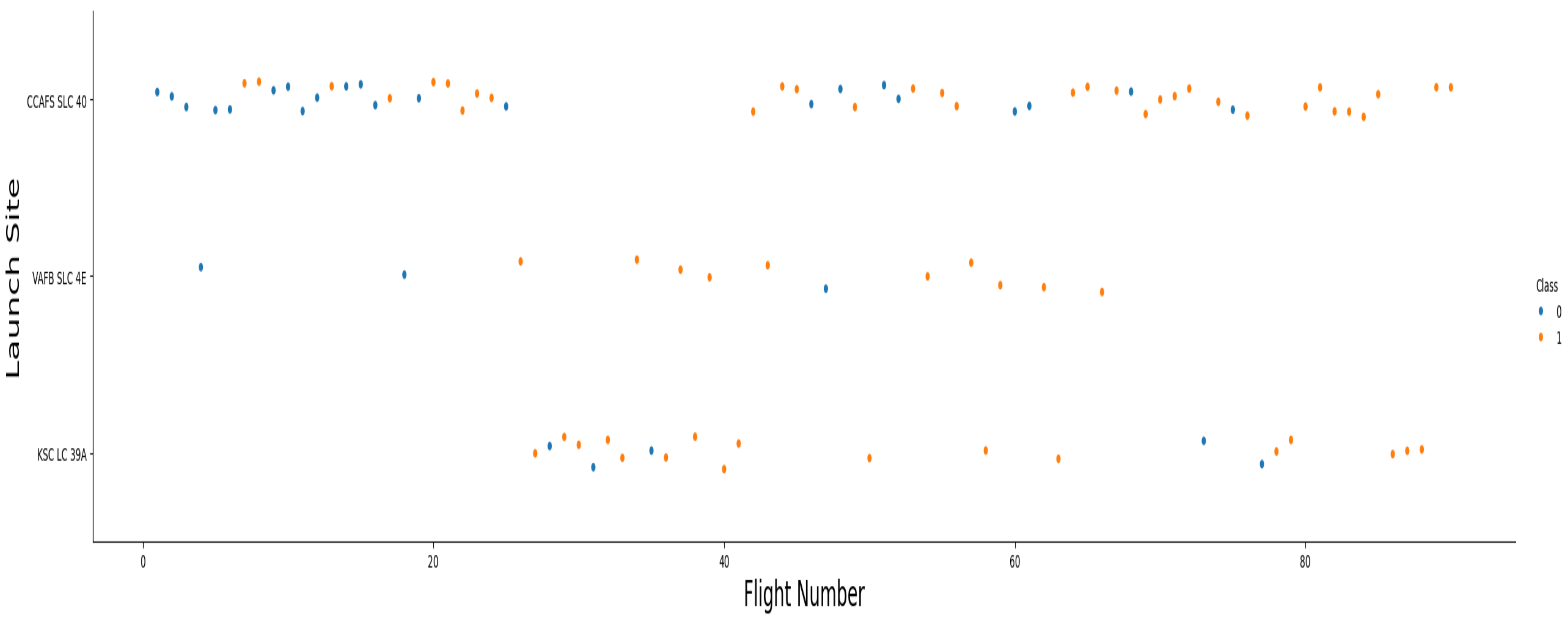
Results

- **Exploratory data analysis results:**
  - The analysis highlights several key factors potentially influencing SpaceX launch success: flight experience, payload mass, launch site, orbit type, and technological progress over time.
- **Interactive analytics:**
  - Interactive analytics empowers users to explore and analyze data dynamically. It involves real-time data manipulation, visualization, and filtering capabilities, enabling users to uncover insights, identify trends, and make data-driven decisions quickly. It is often facilitated through user-friendly dashboards, drag-and-drop interfaces, and customizable reports, fostering a deeper understanding of data for a wider range of users.
- **Predictive analysis results:**
  - SVM is the Champion: Support Vector Machine clinched the top spot with the highest accuracy, proving its prowess in handling this dataset's complexities.
  - All Models Show Promise: Logistic Regression, Decision Tree, and KNN also demonstrated good accuracy, highlighting their potential for this type of prediction.
  - Room for Growth: Despite these positive results, there's room for improvement through larger datasets and exploring other advanced machine learning techniques.

Section 2

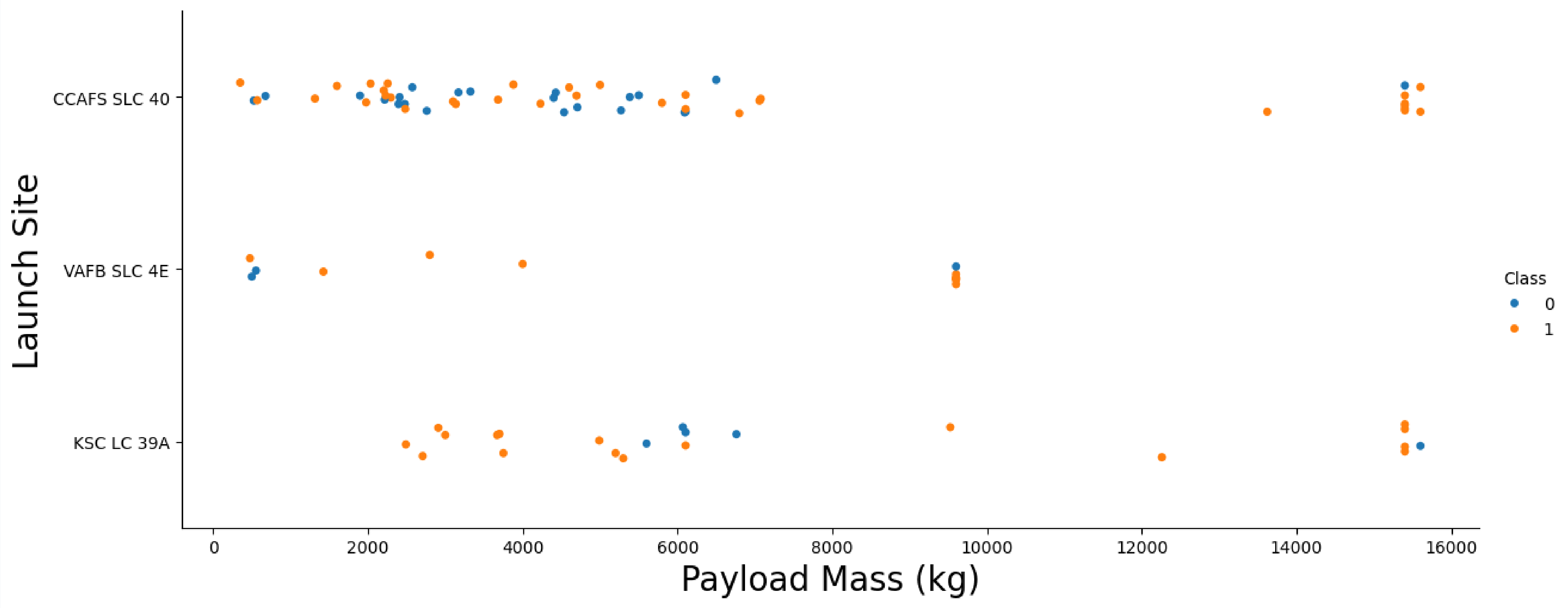
Insights drawn from EDA

Flight Number vs. Launch Site



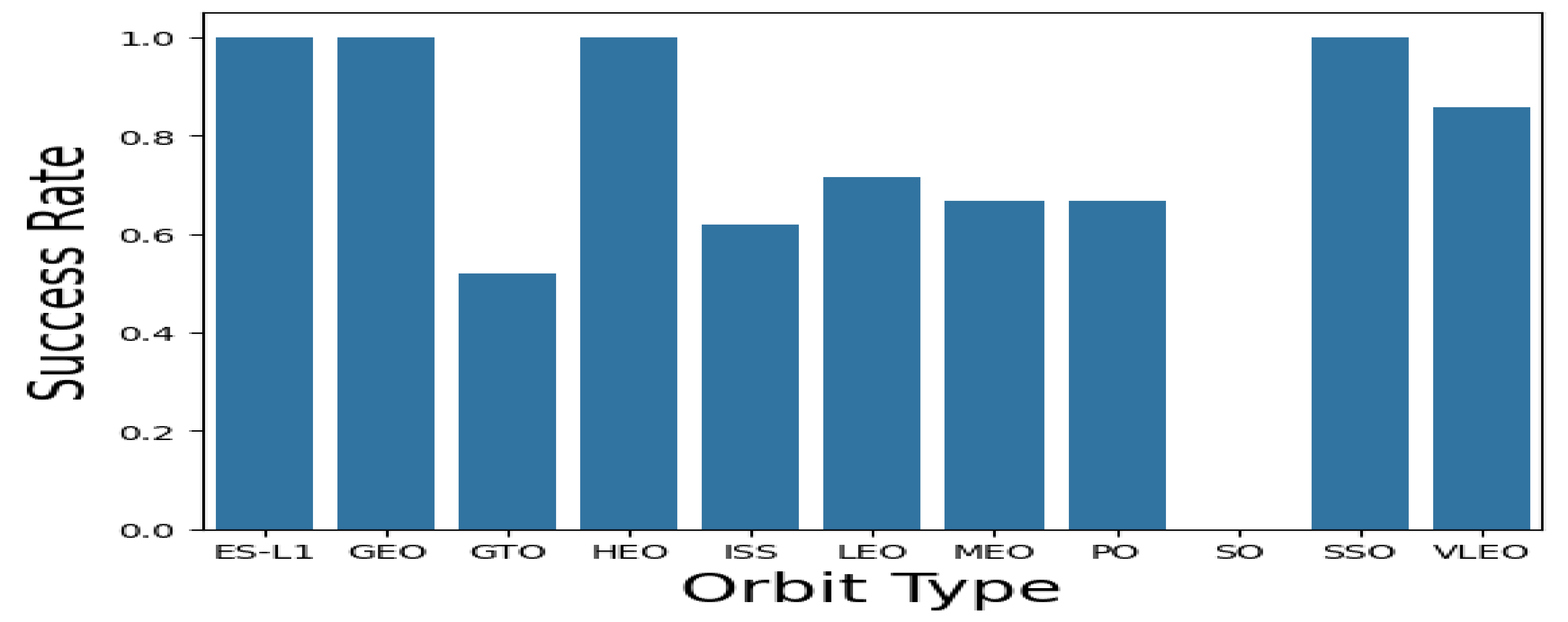
This scatterplot visualizes the relationship between Flight Number and Launch Site for SpaceX launches, with color-coded points indicating success (orange) or failure (blue). It seems that launch frequency increases over time at all sites, and while CCAFS SLC 40 shows a mix of successes and failures, VAFB SLC 4E has notably fewer launches overall.

Payload vs. Launch Site



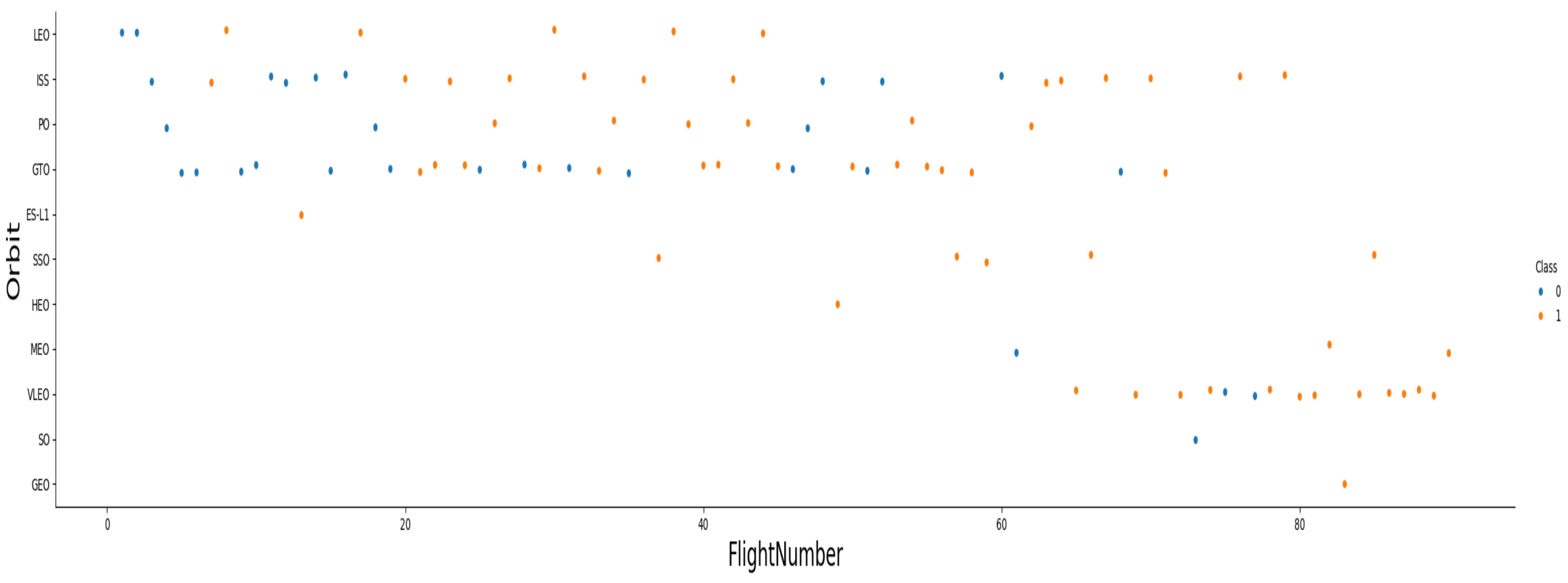
This scatter plot examines the relationship between Payload Mass and Launch Site for SpaceX launches, with the color distinguishing between successful (Class 1) and unsuccessful (Class 0) missions. It appears that KSC LC 39A handles the heaviest payloads, and all three sites have a mix of success and failures across different payload ranges.

Success Rate vs. Orbit Type



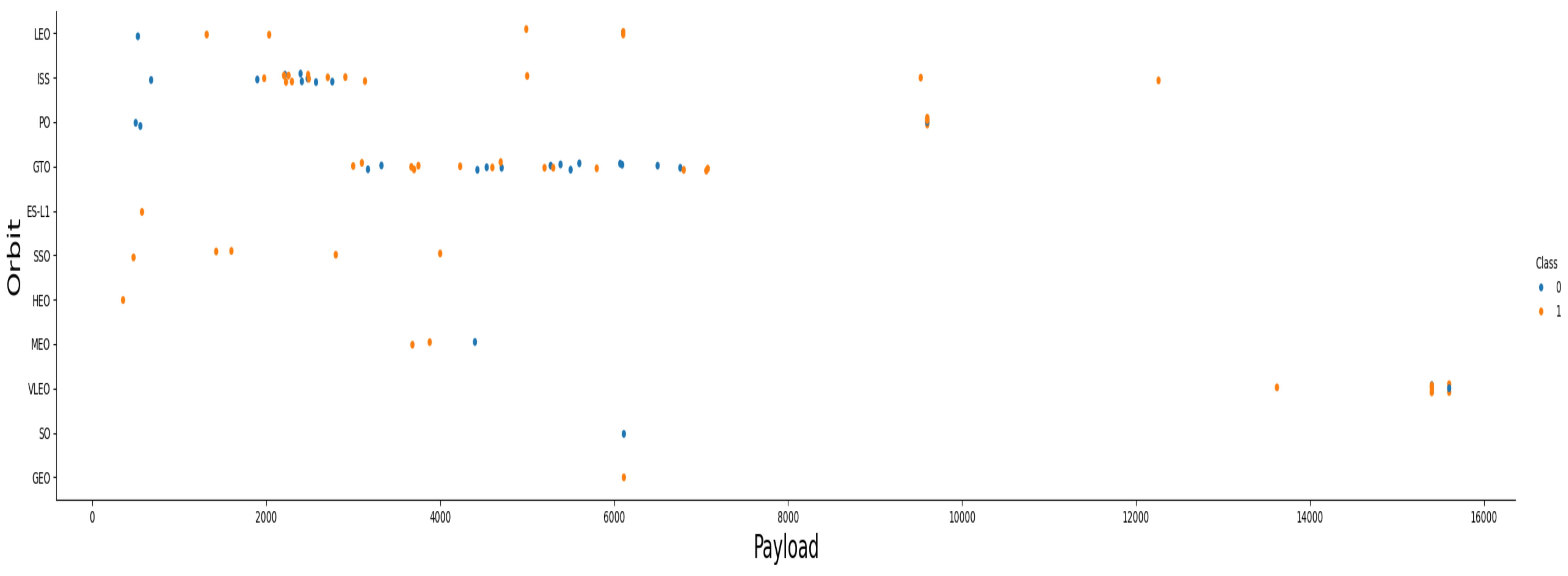
This bar chart depicts the success rate of SpaceX launches across various orbit types. We can observe that orbits like ES-L1, GEO, HEO, and SSO have the highest success rates, while GTO and others exhibit lower or more variable success.

Flight Number vs. Orbit Type



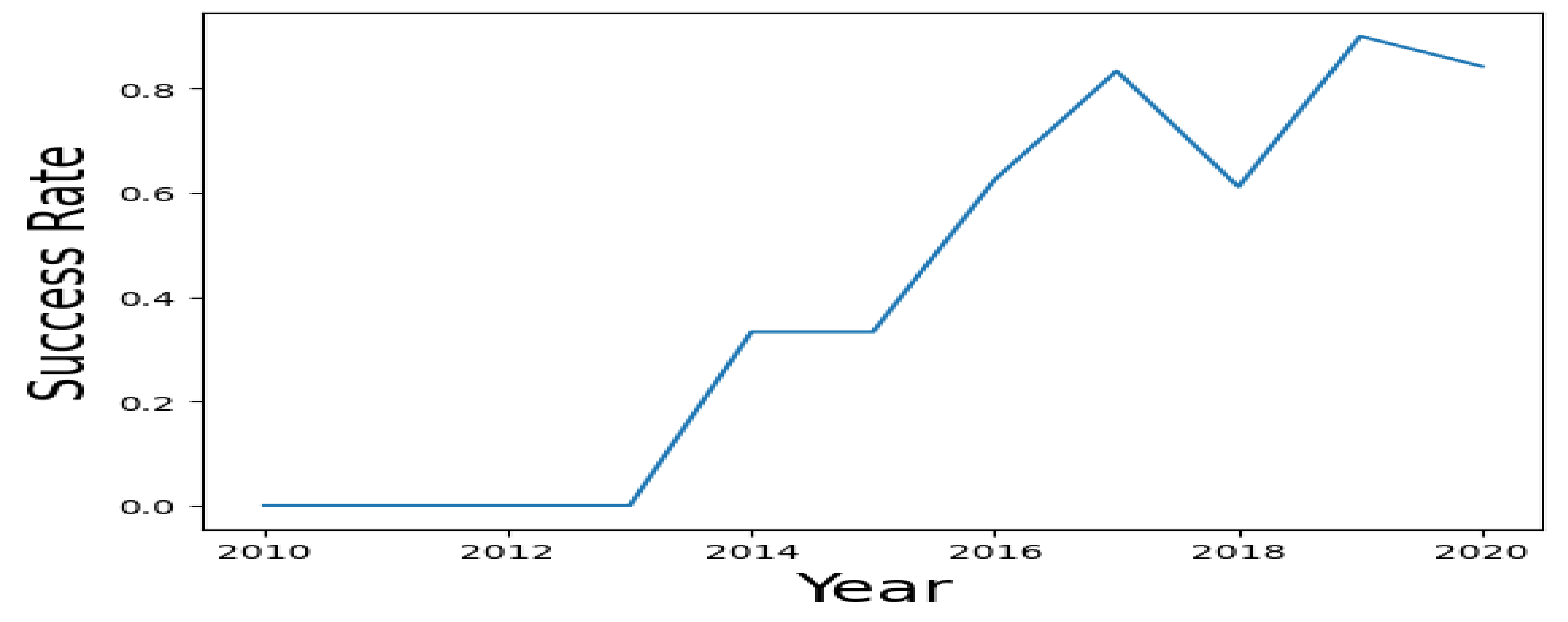
This scatter plot illustrates the relationship between Flight Number and Orbit type for SpaceX launches, with the color distinguishing between successful (orange) and unsuccessful (blue) missions. It appears that certain orbits like ES-L1, SSO, and VLEO have high success rates regardless of flight number, while others like GTO and LEO show a mix of outcomes, potentially suggesting a learning curve or greater complexity in those orbits.

Payload vs. Orbit Type



This scatter plot shows the relationship between Payload Mass (kg) and Orbit type for SpaceX launches, with the color distinguishing success (orange) or failure (blue). It suggests that heavier payloads tend to be associated with specific orbits like GTO, while lighter payloads are distributed across various orbits.

Launch Success Yearly Trend



This line chart illustrates the trend of SpaceX launch success rates over time, from 2010 to 2020. It shows a significant increase in success rate starting from 2013 with a plateau in 2014, followed by a steeper rise from 2015 onwards, peaking in 2019 before a slight dip in 2020.

All Launch Site Names

- The names of the unique launch sites in the space mission:
  - CCAFS LC-40
  - VAFB SLC-4E
  - KSC LC-39A
  - CCAFS SLC-40
- The SQL query retrieves the unique launch site names from the 'SPACEXTABLE' database.

```
[8]: %sql SELECT DISTINCT LAUNCH_SITE FROM SPACEXTABLE;

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

[8]:

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

Launch Site Names Begin with 'CCA'

- The query aimed to fetch 5 records from a table named "SPACEXTABLE" where the "LAUNCH_SITE" column starts with the string 'CCA'.

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

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

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

Total Payload Mass

- The SQL query that calculates and displays the total payload mass (in kilograms) carried by boosters launched by NASA (CRS). The result indicates that the total payload mass is 45596 kg.

```
%sql SELECT SUM(PAYLOAD_MASS__KG_) AS total_payload_mass FROM SPACEXTABLE WHERE CUSTOMER = 'NASA (CRS)';
```

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

total_payload_mass
45596

Average Payload Mass by F9 v1.1

- The SQL query being used to calculate the average payload mass carried by SpaceX's Falcon 9 v1.1 booster. The result of the query shows that the average payload mass is 2928.4 kg.

```
%sql SELECT AVG(PAYLOAD_MASS__KG_) AS average_payload_mass FROM SPACEXTABLE WHERE BOOSTER_VERSION = 'F9 v1.1';

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

average_payload_mass
2928.4

First Successful Ground Landing Date

- This SQL query aims to find the earliest date when a SpaceX landing was successful on a ground pad. The output shows that the first successful ground pad landing occurred on December 22, 2015.

```
%sql SELECT MIN(DATE) AS first_successful_ground_pad_date,LANDING_OUTCOME FROM SPACEXTABLE WHERE LANDING_OUTCOME = 'Success (ground pad)';
```

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

first_successful_ground_pad_date	Landing_Outcome
2015-12-22	Success (ground pad)

Successful Drone Ship Landing with Payload between 4000 and 6000

- The query selects the unique booster versions from the SPACEXTABLE that successfully landed on a drone ship and carried a payload mass between 4000 kg and 6000 kg. The output lists these specific booster versions along with their landing outcome and payload mass.

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 DISTINCT BOOSTER_VERSION,LANDING_OUTCOME,PAYLOAD_MASS__KG_ FROM SPACEXTABLE WHERE \
LANDING_OUTCOME = 'Success (drone ship)' AND PAYLOAD_MASS__KG_ > 4000 AND PAYLOAD_MASS__KG_ < 6000;
```

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

]:

Booster_Version	Landing_Outcome	PAYLOAD_MASS__KG_
F9 FT B1022	Success (drone ship)	4696
F9 FT B1026	Success (drone ship)	4600
F9 FT B1021.2	Success (drone ship)	5300
F9 FT B1031.2	Success (drone ship)	5200

Total Number of Successful and Failure Mission Outcomes

- This SQL query counts the occurrences of each MISSION_OUTCOME in the SPACEXTABLE. The results show there was 1 in-flight failure, 99 successes, and 1 success with unclear payload status.

List the total number of successful and failure mission outcomes

```
%sql SELECT MISSION_OUTCOME, COUNT(*) AS total_number FROM SPACEXTABLE GROUP BY MISSION_OUTCOME;
```

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

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

Boosters Carried Maximum Payload

- The query aims to identify the booster versions in the SPACEXTABLE that carried the maximum payload mass. It first finds the maximum PAYLOAD_MASS_KG and then selects the booster versions associated with that maximum payload.

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

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

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

2015 Launch Records

- The SQL query retrieves the month, date, mission outcome, booster version, and launch site from the SPACEXTBL database for missions that resulted in a failed drone ship landing and occurred in the year 2015. The output displays two such missions, both occurring at CCAFS LC-40 in January and April 2015, despite having successful mission outcomes.

```
%sql SELECT substr(Date, 6, 2) as month, Date, MISSION_OUTCOME, BOOSTER_VERSION, LAUNCH_SITE FROM SPACEXTBL \
WHERE LANDING_OUTCOME = 'Failure (drone ship)' AND substr(Date, 0, 5)='2015';
```

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

month	Date	Mission_Outcome	Booster_Version	Launch_Site
01	2015-01-10	Success	F9 v1.1 B1012	CCAFS LC-40
04	2015-04-14	Success	F9 v1.1 B1015	CCAFS LC-40

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

- The query fetches the count of different landing outcomes (success or failure types) from the SPACEXTABLE for launches between June 4th, 2010 and March 20th, 2017. The results are then sorted in descending order based on the frequency of each outcome.

```
%sql SELECT LANDING_OUTCOME, COUNT(*) AS total_outcomes FROM SPACEXTABLE \
WHERE "Landing _Outcome" LIKE "%Success%" OR "F DATE BETWEEN '2010-06-04' AND '2017-03-20' GROUP BY LANDING_OUTCOME ORDER BY total_outcomes DESC;
```

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

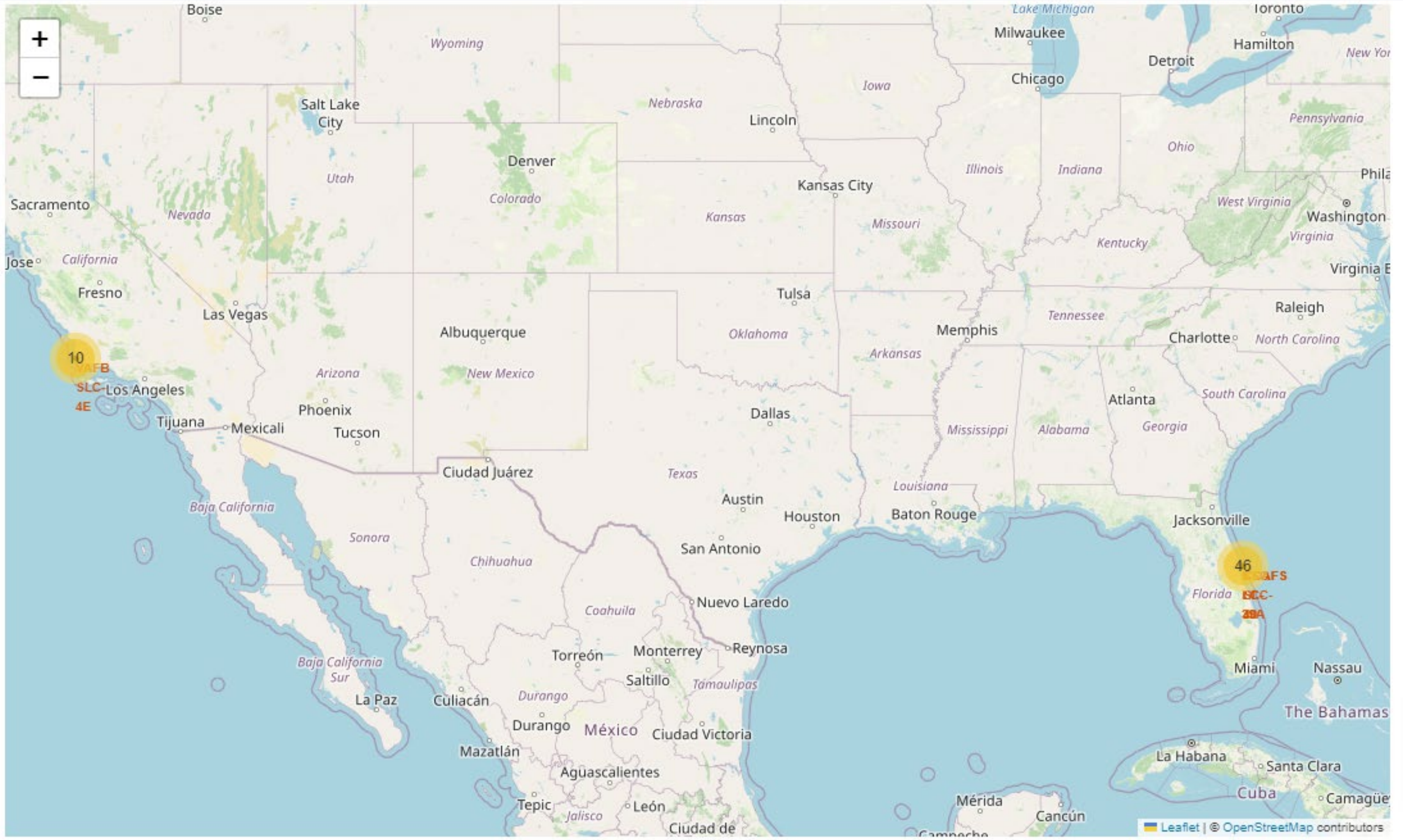
Landing_Outcome	total_outcomes
No attempt	10
Success (drone ship)	5
Failure (drone ship)	5
Success (ground pad)	3
Controlled (ocean)	3
Uncontrolled (ocean)	2
Failure (parachute)	2
Precluded (drone ship)	1

Section 3

Launch Sites Proximities Analysis

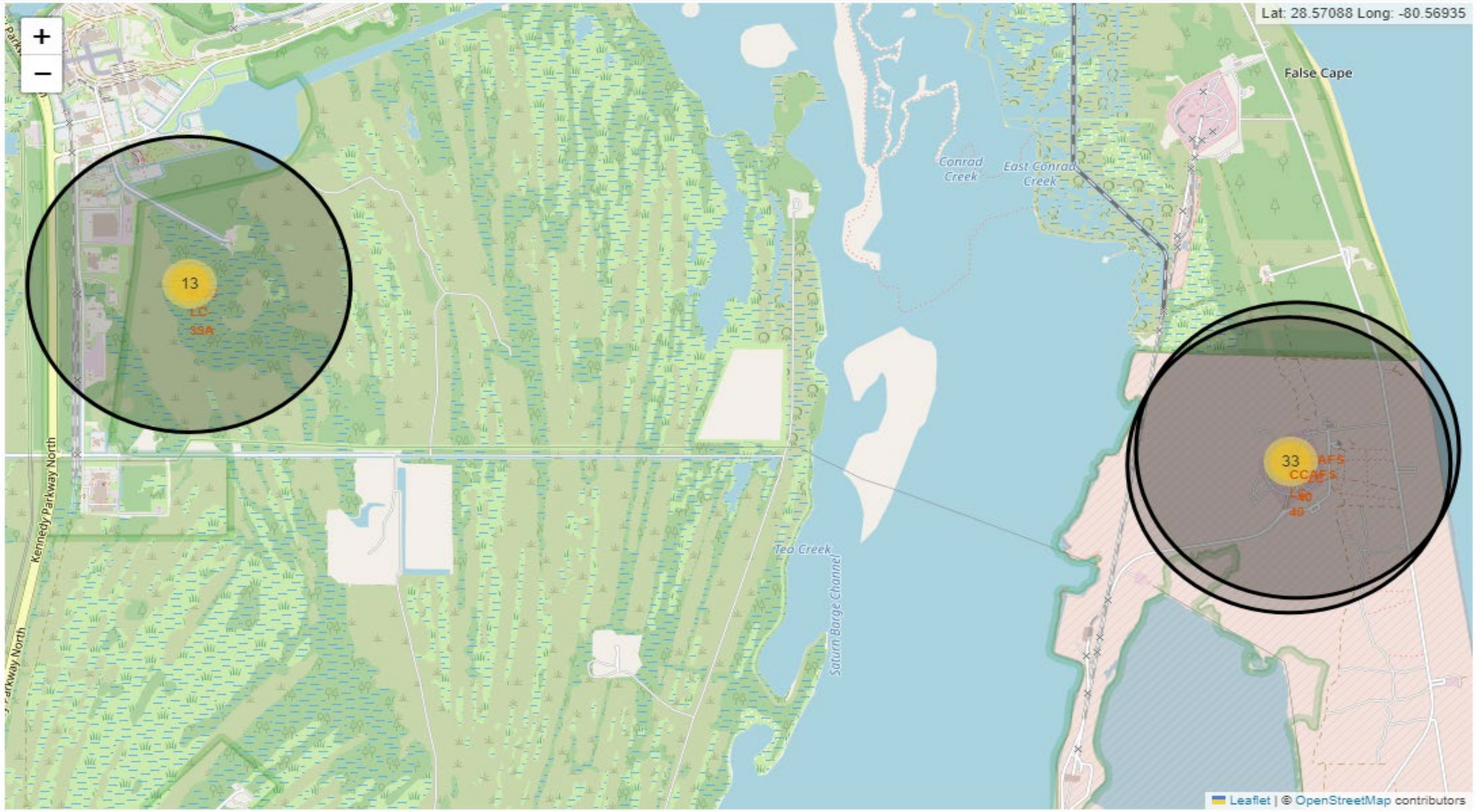
All active SpaceX launch sites

This map displays all active SpaceX launch sites, marked with clustered circles and numbers indicating the quantity of launches from each site.



<Folium Map Screenshot 2>

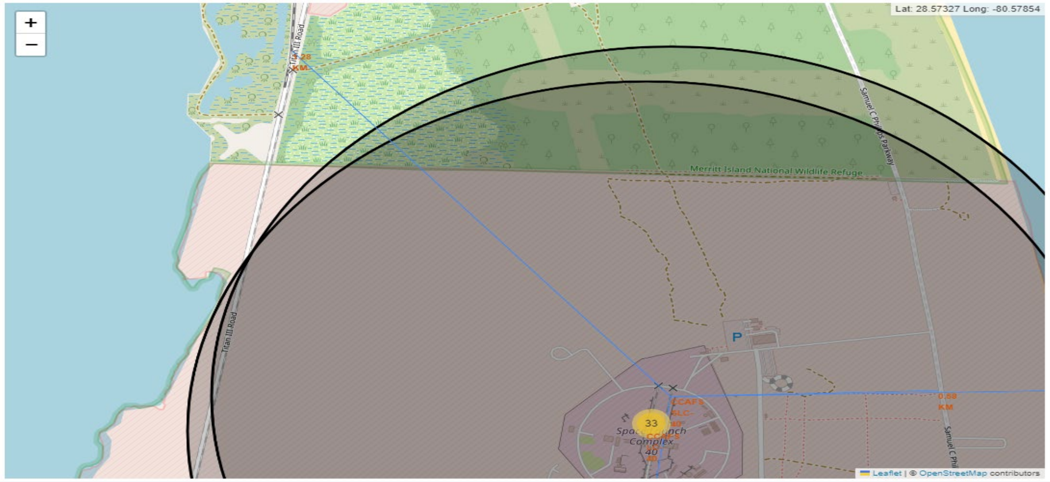
This map displays two SpaceX launch sites in Florida: Cape Canaveral Space Force Station (CCAFS) and Kennedy Space Center (KSC). CCAFS, highlighted by the top circle, has hosted 13 launches, while KSC, shown in the bottom circle, has had significantly more launches (33). The color variations within the clusters likely represent different launch outcomes (successes and failures).

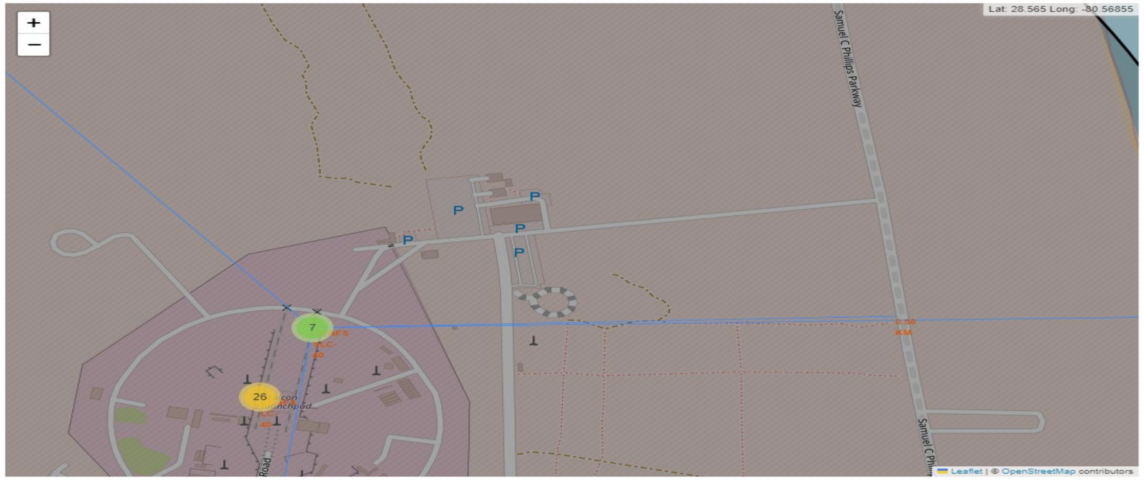


The nearest coastline, the closest railway and highway

The map displays a detailed view of the Cape Canaveral Space Force Station (CCAFS) launch site and its surrounding area. It highlights the proximity of the launch site to critical infrastructure, including:

- **Coastline:** The nearest point on the coastline is a mere 0.83 km away.
- **Railway:** A railway line runs 1.42 km from the launch site.
- **Highway:** A highway is located 1.15 km away.



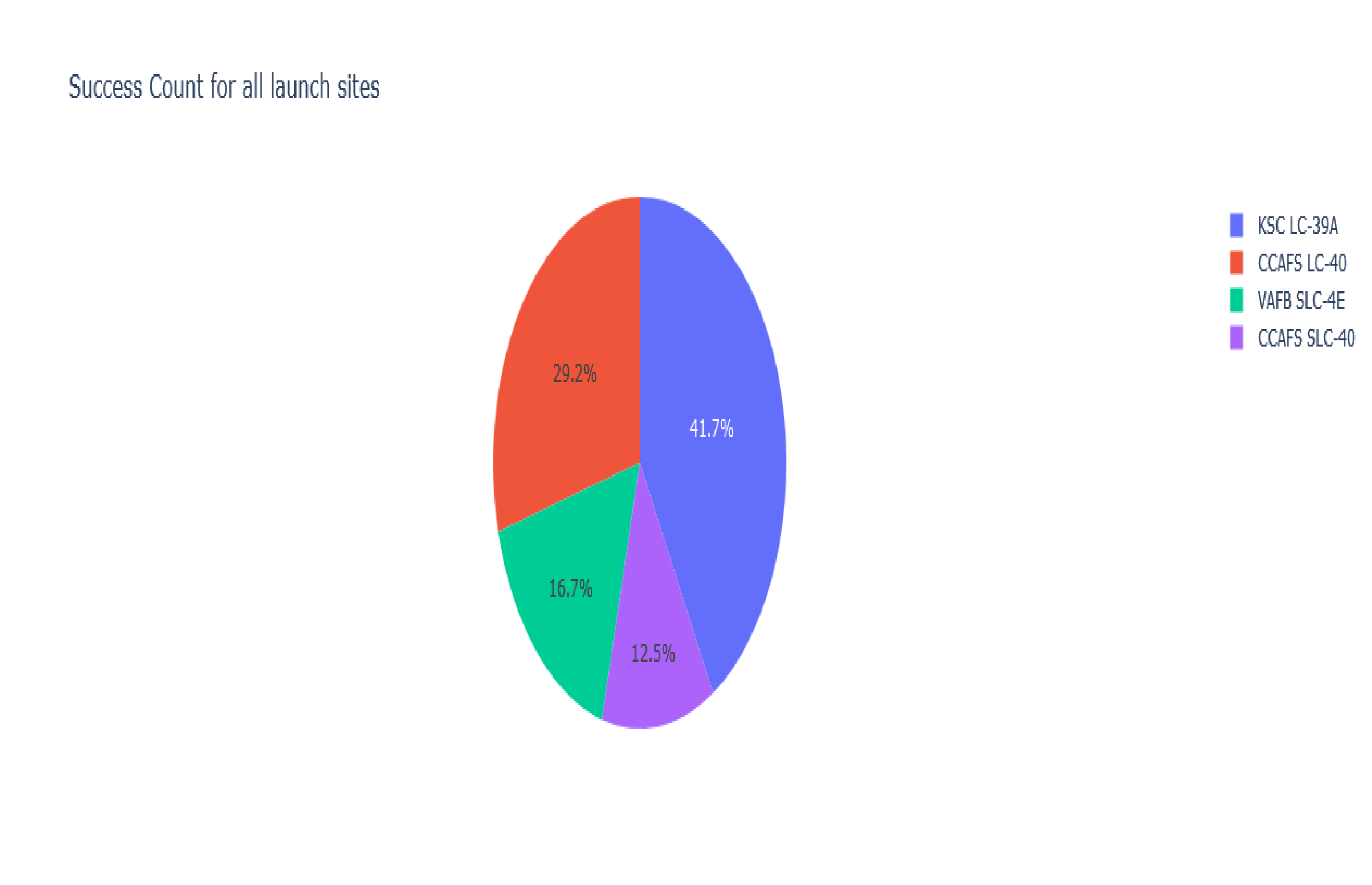


Section 4

Build a Dashboard with Plotly Dash

Successful SpaceX launches by launch site

- This pie chart presents a breakdown of successful SpaceX launches by launch site.
- CCAFS LC-40 leads with 41.7% of all successful launches.
- KSC LC-39A follows closely with 29.2%.
- VAFB SLC-4E and CCAFS SLC-40 account for 16.7% and 12.5% respectively.

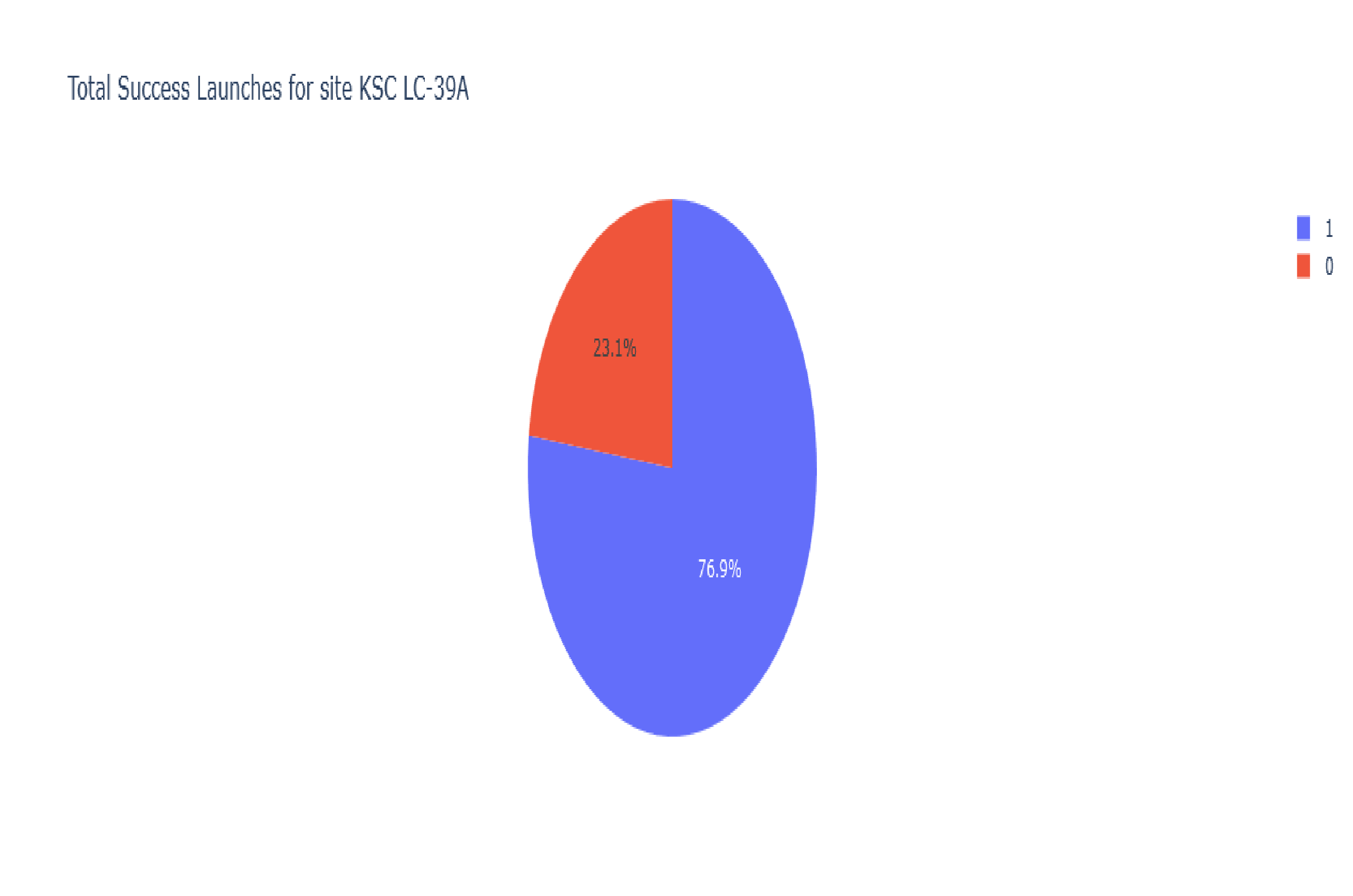


Success rate of launches from the KSC LC-39A launch site

This pie chart displays the success rate of SpaceX launches specifically from the KSC LC-39A launch site.

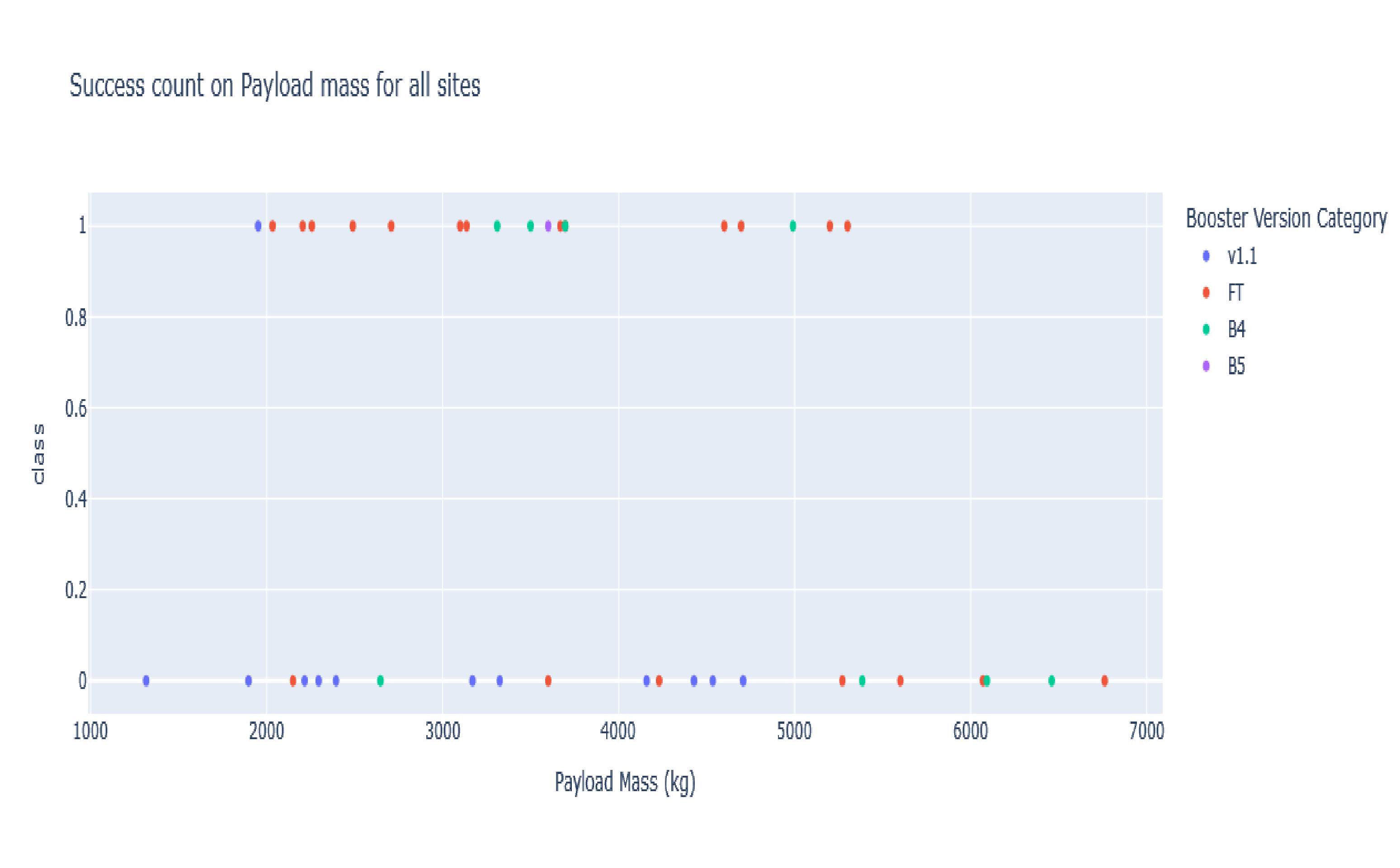
It shows that:

- 76.9% of launches from this site were successful.
- 23.1% of launches from this site were unsuccessful.



The relationship between payload mass and launch success for all launch sites

This scatter plot shows the relationship between payload mass and launch success for all SpaceX launch sites. It seems that heavier payloads don't necessarily reduce success rates, as successful launches (class 1) are distributed across the entire payload range. Different booster versions are also represented, with v1.1 being the most frequently used.

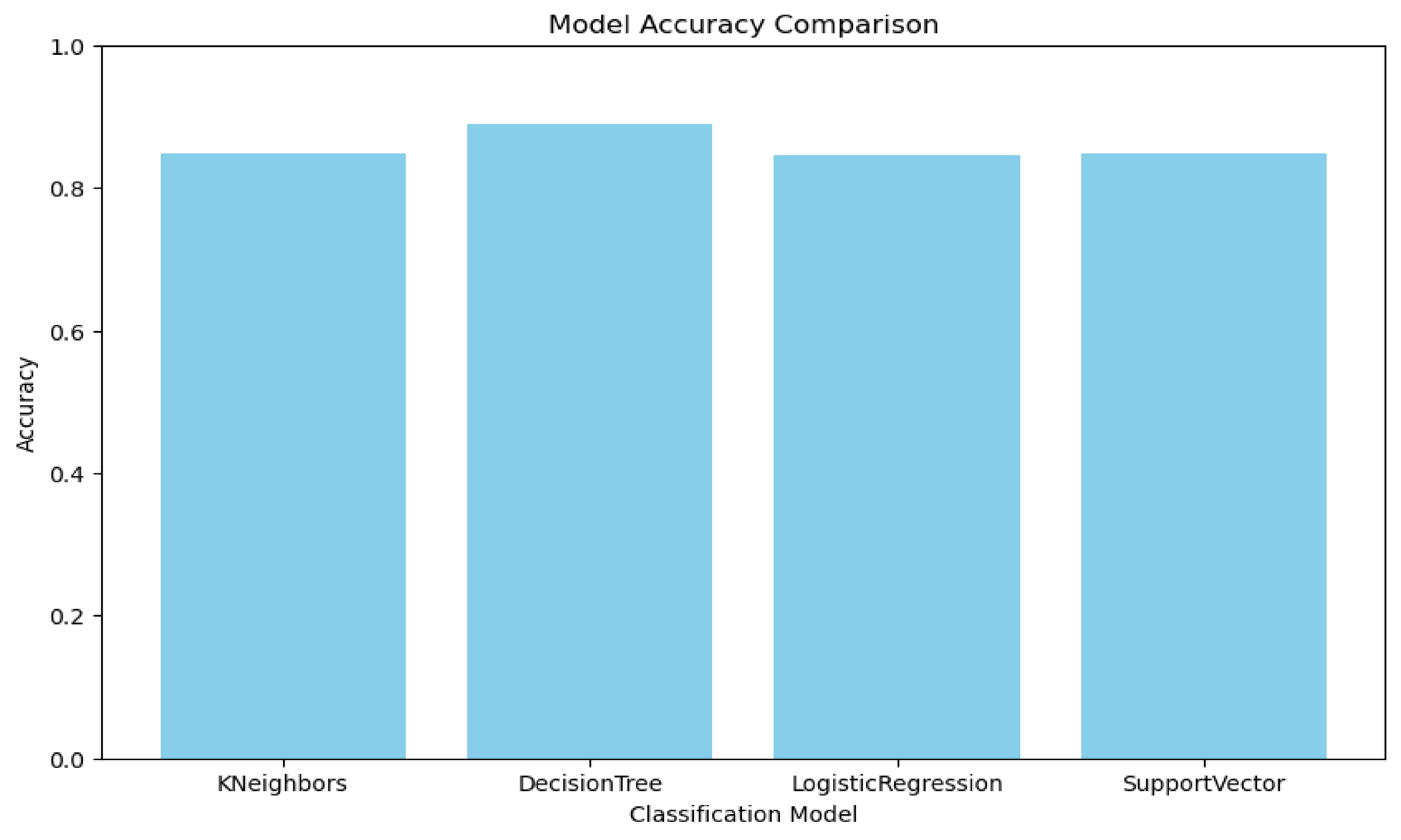


Section 5

Predictive Analysis (Classification)

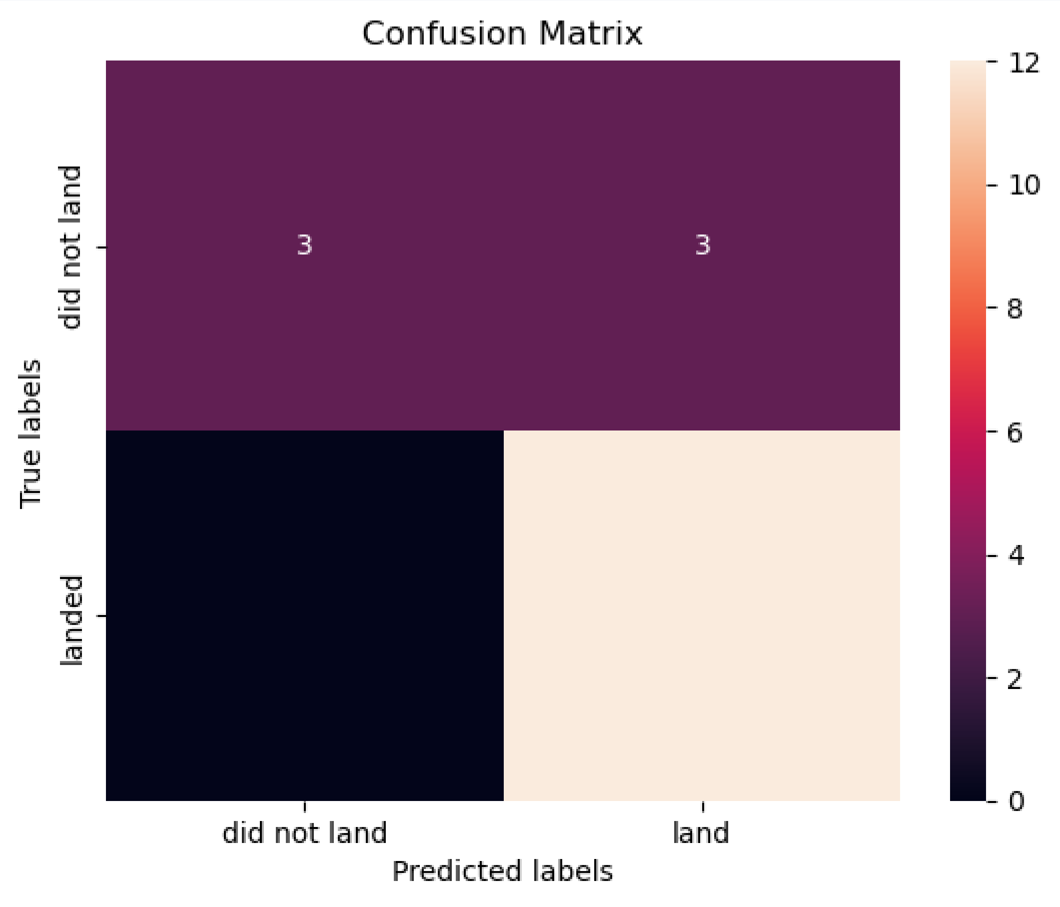
Classification Accuracy

It looks like we have a tie! The Decision Tree, Logistic Regression, and Support Vector Machine models all share the highest accuracy score of 0.833. This means they correctly predicted the outcome of approximately 83% of the SpaceX launches in the test dataset.



Confusion Matrix

We built a machine-learning model to predict the success of SpaceX Falcon 9 first-stage landings. We trained and evaluated four different models: Logistic Regression, Support Vector Machine, Decision Tree, and K-Nearest Neighbors. All models showed good accuracy, but the Decision Tree, Logistic Regression, and SVM tied for the best performance. We also visualized the results using a confusion matrix to understand the model's predictions in more detail.



Conclusions

- The success rate of SpaceX launches generally increased from 2013 to 2020, indicating improvements in technology and processes.
- Heavier payloads tend to have a slightly lower success rate, especially for certain orbit types.
- Different launch sites exhibit varying success rates, suggesting that site-specific factors could play a role.
- Certain orbits (ES-L1, GEO, HEO, SSO) show a 100% success rate, while others (GTO, SO) have more variability.
- The choice of booster version might also influence the success rate, with newer versions potentially performing better.
- SVM, Logistic Regression, and Decision Tree models emerged as the top performers in predicting launch success.
- Further analysis is needed to definitively select the best model among the top three, considering factors like interpretability, complexity, and training time.
- Potential for improvement exists by further tuning hyperparameters or exploring ensemble methods.
- Confusion matrices provide valuable insights into model predictions, highlighting areas where the model excels or struggles.
- The analysis successfully identified key factors that may influence the success of SpaceX launches, including payload mass, launch site, orbit type, and booster version.
- The machine learning models developed can be valuable tools for predicting launch success and optimizing mission parameters.

Appendix

- Include any relevant assets like Python code snippets, SQL queries, charts, Notebook outputs, or data sets that you may have created during this project

Thank you!