IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Matias Brandt Møller
October 9, 2024

Outline

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

Executive Summary of Methodologies

Exploratory Data Analysis (EDA):

- Used Seaborn and Matplotlib to investigate patterns in launch data, focusing on factors such as launch sites, payload mass, and orbit types.
- Key tools included scatter plots and bar charts to visualize relationships between variables and identify trends over time.

2. **Data Wrangling**:
    - Cleaned and processed the dataset by handling missing values, normalizing data, and applying one-hot encoding for categorical variables.
    - Filtered and aggregated the data for analysis, ensuring the dataset was ready for machine learning modeling.
3. **Predictive Modeling**:
    - Trained multiple classification models (Logistic Regression, K-Nearest Neighbors, Decision Tree, and Support Vector Machine) to predict launch success.
    - Evaluated models based on accuracy, precision, recall, and F1 score, and improved model performance using GridSearchCV for hyperparameter tuning.
4. **Interactive Visualizations**:
    - Built an interactive dashboard with Plotly Dash to display visual insights, including a scatter plot of payload mass vs. launch success, a pie chart of success rates, and dynamic interactions with dropdowns and sliders.
    - Created an interactive map with Folium to plot launch sites and visualize geographical factors affecting launch outcomes.

Executive Summary of Results

1. **Launch Success Rates**:
   - Launch sites like **CCAFS SLC-40** and **KSC LC-39A** have higher success rates compared to others, indicating that launch location plays a significant role in the outcome.
2. **Payload Mass vs. Success**:
   - Higher payload mass correlates with a greater chance of success, especially at sites with frequent launches. However, success is still possible with lower payloads, depending on other conditions.
3. **Best Performing Model**:
   - **Logistic Regression** emerged as the best-performing model, achieving an accuracy of **83.33%** after applying **GridSearchCV** for hyperparameter tuning. Other models like KNN and SVM performed equally well, while Decision Tree had a lower accuracy.
4. **Yearly Trends**:
   - The success rate of SpaceX launches has steadily improved over the years, especially from 2013 onwards, reflecting the company’s learning curve and technological advancements.

Introduction

- The goal of this project was to analyze SpaceX's Falcon 9 launch data to predict the likelihood of a successful first-stage landing, a key factor in SpaceX's cost-saving strategy. Reusability of the first stage plays a vital role in reducing launch costs, making predictive analysis critical for operational planning.

- Using a combination of Exploratory Data Analysis (EDA), data wrangling, and predictive modeling, we sought to uncover key insights into what factors influence the success or failure of a launch. These factors include the launch site, payload mass, orbit type, and booster version, among others.

- The methodology involved analyzing historical launch data, building classification models to predict launch success, and creating interactive dashboards for visualizing these insights. Through these methods, we aimed to identify patterns and trends, and ultimately recommend the best models for predicting future outcomes with high accuracy.

- This paper will show which parameters are instrumental for a succesful SpaceX launch and landing of the first stage.

Section 1

Methodology

Methodology

Executive Summary

- Data collection methodology:
  - Data was collected from the SpaceX API
  - Historical launch data was collected using webscraping on https://en.wikipedia.org/wiki/List_of_Falcon_9_and_Falcon_Heavy_launches
- Perform data wrangling
  - The data was cleaned and one-hot encoding was applied to categorical values.
- Perform exploratory data analysis (EDA) using visualization and SQL
- Perform interactive visual analytics using Folium and Plotly Dash
- Perform predictive analysis using classification models
  - Logistic Regression, Support Vector Machine, Decision Tree and KNN models where trained tested and compared using train/test splits and GridSerchCV to get the optimal model and model parameters.

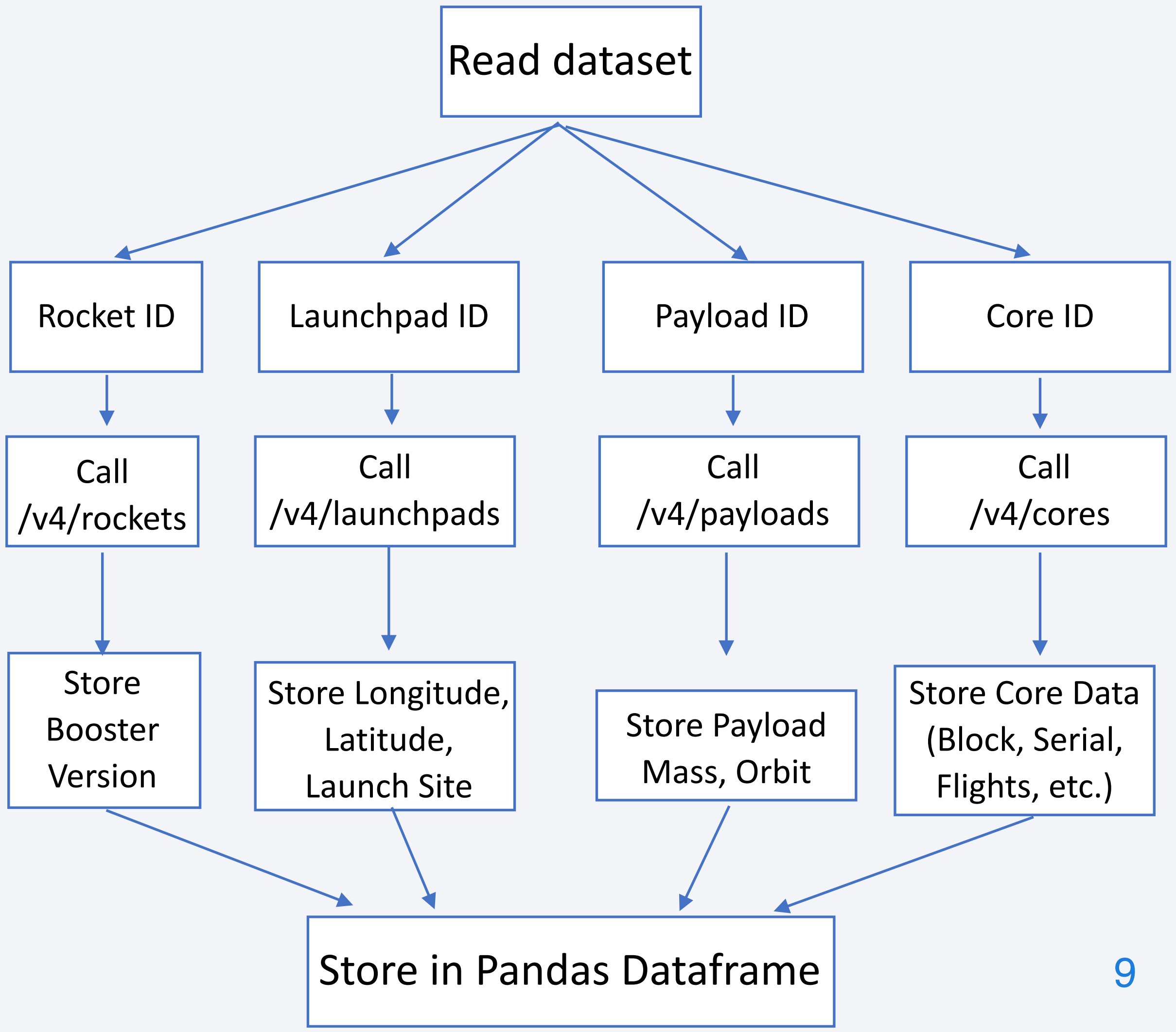
Data Collection

1. **SpaceX API**:
    - We made API calls to retrieve up-to-date launch records from SpaceX's public API. The data included details about each launch such as payload mass, orbit type, launch site, and whether the first-stage landing was successful or not.
    - The API provided real-time data access, allowing for continuous updates and the inclusion of the latest launch information.
2. **Web Scraping from Wikipedia**:
    - We complemented the API data with historical launch records from the Wikipedia page titled "List of Falcon 9 and Falcon Heavy launches."
    - Using BeautifulSoup, we scraped detailed tables listing launch dates, booster versions, payloads, and outcomes to ensure our dataset was comprehensive.

Data Collection – SpaceX API

- Data collection with SpaceX REST calls:
  - API Setup:
    - Use requests library to access SpaceX API endpoints like /v4/rockets and /v4/launchpads.
  - Booster Version:
    - Retrieve booster version names using rocket IDs from /v4/rockets/{id}.
  - Launch Site Data:
    - Gather longitude, latitude, and launch site names from /v4/launchpads/{id}.
  - Payload Data:
    - Collect payload mass and orbit data from /v4/payloads/{id}.
  - Core Data:
    - Extract core block, flight, and landing outcome from /v4/cores/{id}.

- GitHub URL: https://github.com/BrandtBrandtBrandt/winningthespacerace/blob/d3c0e77e39596ce5b7ec49a7129633c2ef0f952b/jupyter-labs-spacex-data-collection-api%20(1).ipynb



Data Collection - Scraping

- Send HTTP GET Request:
    - Use the requests.get() method to retrieve the HTML content of the Wikipedia page on Falcon 9 launches.
- Create BeautifulSoup Object:
    - Parse the HTML content using BeautifulSoup to create a structured object for easier extraction.
- Locate Target Table:
    - Find the relevant launch records table using soup.find_all('table'), specifically targeting the third table containing the data.
- Extract Table Headers:
    - Identify column names by iterating over the <th> elements and extracting their contents.
- Parse Table Rows:
    - Loop through the table rows (<tr>), extracting relevant data points such as Flight Number, Date, Launch Site, Payload, Orbit, and Landing Outcome.
- Store Data in Dictionary:
    - Populate a dictionary (launch_dict) with extracted data from the table rows.
- Convert to DataFrame:
    - Convert the populated dictionary into a Pandas DataFrame for further analysis.

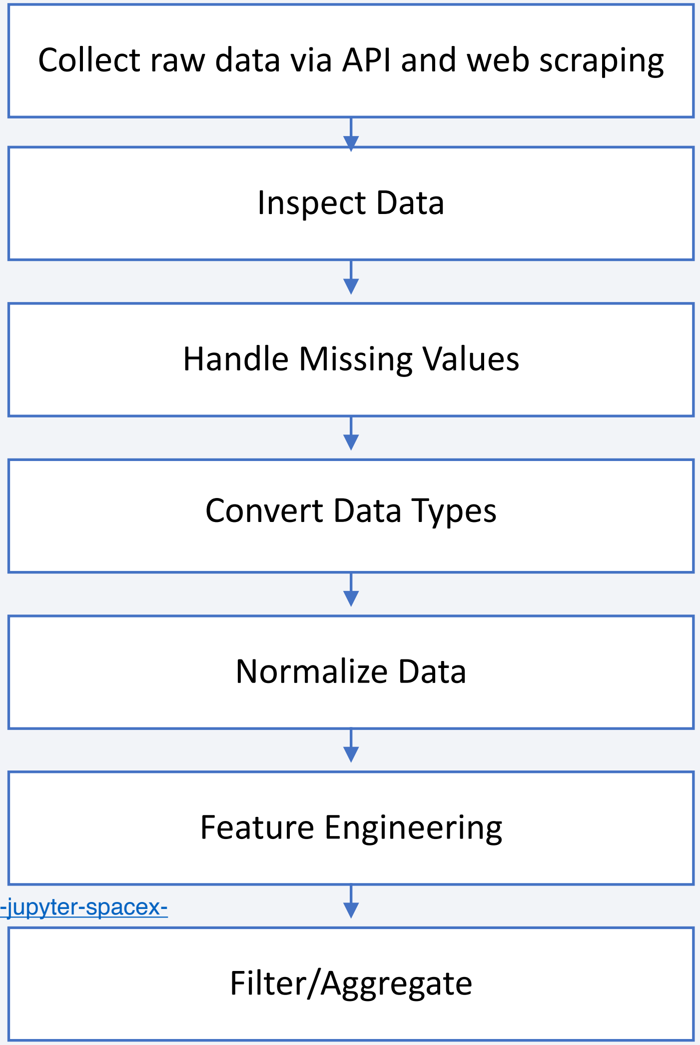
- GitHub URL: https://github.com/BrandtBrandtBrandt/winningthespacerace/blob/642a9cf3473744a35c079b2b158329fb40031271/jupyter-labs-webscraping.ipynb

Send HTTP GET request to Wikipedia page
↓
Create BeautifulSoup Object
↓
Find Table
↓
Extract Headers
↓
Extract Data
↓
Store Data
↓
Convert to DataFrame

Data Wrangling

1. **Data Collection**:
    - Use API calls or web scraping to gather raw data (e.g., SpaceX launch records).
2. **Inspecting the Data**:
    - Use Pandas to load data into a DataFrame and inspect the structure (e.g., missing values, inconsistent formatting).
3. **Handling Missing Values**:
    - Identify missing values using .isnull() and handle them by removing or imputing based on context.
4. **Converting Data Types**:
    - Convert columns to appropriate data types (e.g., datetime, int, float) to ensure consistency.
5. **Data Normalization**:
    - Scale numeric data, clean up string values, and ensure consistency across categorical data.
6. **Feature Engineering**:
    - Create new features based on existing ones (e.g., combining date and time columns or adding binary indicators).
7. **Filtering and Aggregation**:
    - Remove irrelevant data, filter based on conditions, and aggregate as needed to focus on relevant insights.

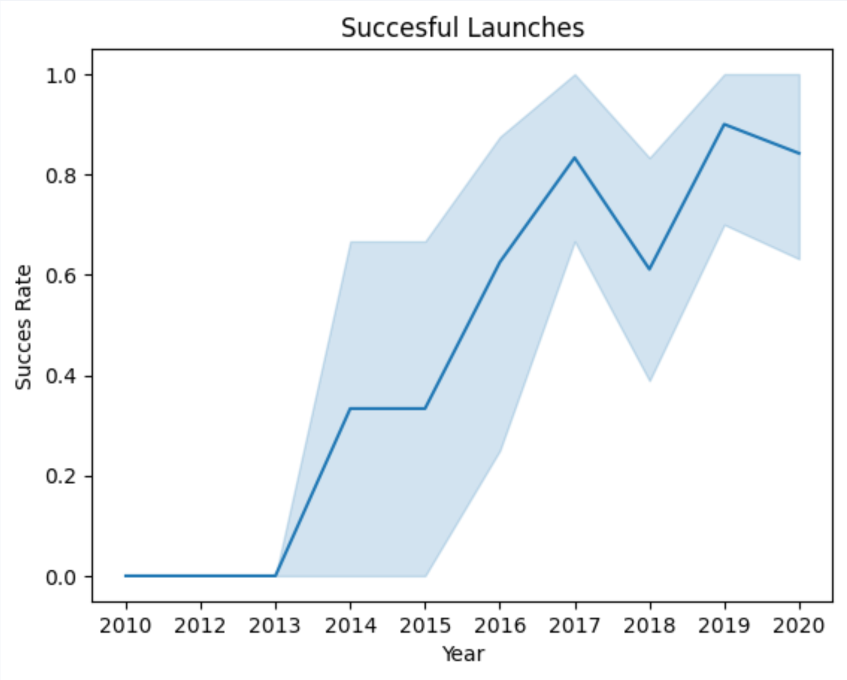
- **GitHub URL**: https://github.com/BrandtBrandtBrandt/winningthespacerace/blob/ae8465da428f2fc4a550f237dc1329bd116062c6/labs-jupyter-spacex-Data%20wrangling%20(1).ipynb



EDA with Data visualisation

Summary of Data Visualization Charts:

1. **Flight Number vs. Payload Mass (with Launch Outcome)**:
   - **Chart**: Scatter Plot
   - **Reason**: To show the relationship between the number of launches and the likelihood of successful landings, while also considering the payload mass.
2. **Flight Number vs. Launch Site**:
   - **Chart**: Scatter Plot
   - **Reason**: To observe if different launch sites show any patterns in the success rate of launches as the flight numbers increase.
3. **Payload Mass vs. Launch Site**:
   - **Chart**: Scatter Plot
   - **Reason**: To explore the distribution of payload mass across various launch sites and its correlation with launch outcomes.
4. **Success Rate vs. Orbit Type**:
   - **Chart**: Bar Chart
   - **Reason**: To visualize the average success rate for each orbit type and identify which orbits have the highest success rates.
5. **Flight Number vs. Orbit Type**:
   - **Chart**: Scatter Plot
   - **Reason**: To explore if there's a relationship between the number of flights and orbit types, and how it affects launch success.
6. **Payload Mass vs. Orbit Type**:
   - **Chart**: Scatter Plot
   - **Reason**: To see how payload mass influences the success rate across different orbit types.
7. **Yearly Launch Success Trend**:
   - **Chart**: Line Chart
   - **Reason**: To visualize how the average success rate of launches has changed over the years, revealing trends over time.



These charts provide insights into the variables that affect the success of Falcon 9 landings and help inform future prediction models.

- **GitHub URL:** https://github.com/BrandtBrandtBrandt/winningthespacerace/blob/73855ffc586c99e2110c11dfdc8d390d373f1f5d/edadataviz%20(2).ipynb

EDA with SQL

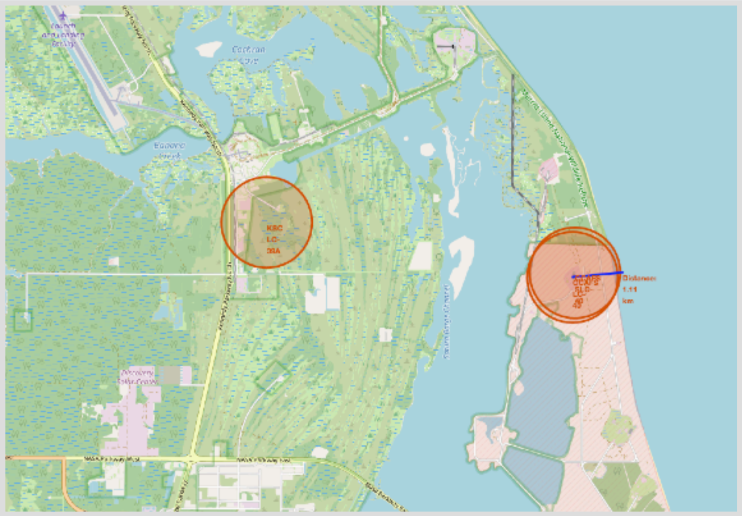
- **Loaded Dataset into SQL**:
  - Imported the dataset into a SQLite database for structured querying and analysis.
- **Basic Descriptive Statistics**:
  - Queried the dataset to get counts, means, and ranges for key variables like PayloadMass and FlightNumber.
  - Example: SELECT AVG(PayloadMass), MAX(FlightNumber), MIN(FlightNumber) FROM launches;
- **Launch Site Distribution**:
  - Identified the number of launches per launch site.
  - Example: SELECT LaunchSite, COUNT(*) AS launch_count FROM launches GROUP BY LaunchSite;
- **Success Rate by Orbit Type**:
  - Calculated the success rate for different orbit types.
  - Example: SELECT Orbit, AVG(Class) AS success_rate FROM launches GROUP BY Orbit;
- **Payload vs. Success Analysis**:
  - Investigated the relationship between payload mass and launch success.
  - Example: SELECT PayloadMass, Class FROM launches WHERE PayloadMass IS NOT NULL;
- **Yearly Trends**:
  - Analyzed yearly trends in launch outcomes and frequency.
  - Example: SELECT strftime('%Y', Date) AS year, COUNT(*) FROM launches GROUP BY year;

These queries helped explore key relationships in the dataset, uncovering trends and patterns in launch success rates.

- **GitHub URL:** https://github.com/BrandtBrandtBrandt/winningthespacerace/blob/4e7d637af780a4fa651839239c4f29c64bf7bf5b/jupyter-labs-eda-sql-coursera_sqllite%20(1).ipynb

Build an Interactive Map with Folium

- Launch sites were plotted to an interactive map as well as information on succes rate and lines visualising the distance to areas of interest.
- These where added to visualise launches and what geographical factors play a role in a succesful launch.
- GitHub URL: https://github.com/BrandtBrandtBrandt/winningthespacerace/blob/c73a49fa2274a9ac182dfeb4530278faf12ee43c/lab_jupyter_launch_site_location.ipynb



Build a Dashboard with Plotly Dash

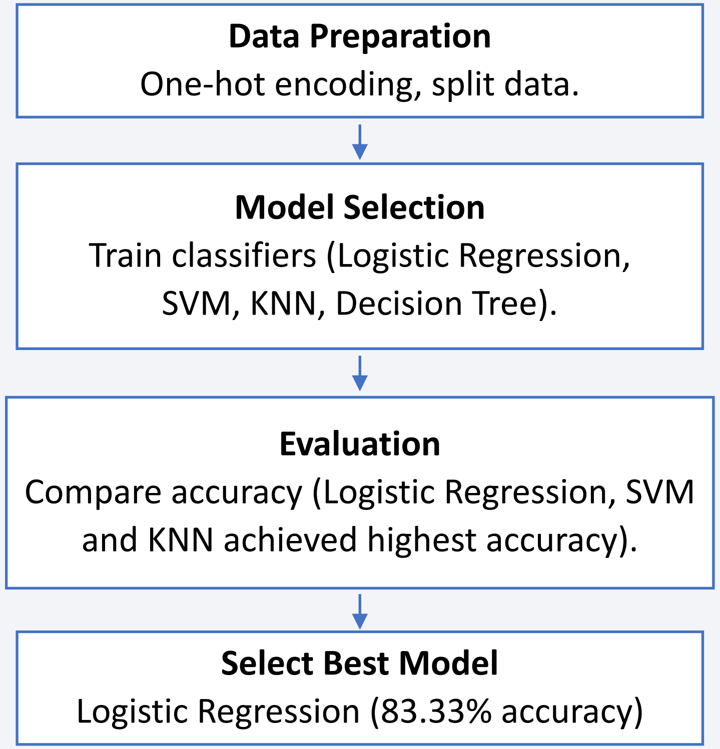
- **Pie Chart: Success Rates by Launch Site**:
  - **Why**: To show the distribution of successful launches across different SpaceX launch sites.
  - **Interaction**: Users can filter the pie chart by selecting a specific launch site from the dropdown to view success vs. failure for that site.
- **Payload Range Slider**:
  - **Why**: Allows users to dynamically adjust the payload mass range and filter the data in the scatter plot accordingly.
  - **Interaction**: Users can select a payload range to update the scatter plot and analyze how payload mass affects launch outcomes.
- **Scatter Plot: Payload Mass vs. Launch Success**:
  - **Why**: Visualize the correlation between payload mass and the success of launches. Color-coded by booster version to identify patterns based on rocket type.
  - **Interaction**: Users can filter the plot by launch site and payload range to explore how different variables impact success rates.

These features were added to make the dashboard interactive, providing users with the ability to explore SpaceX launch data and discover relationships between payload mass, launch sites, and success outcomes.

- **GitHub URL**: https://github.com/BrandtBrandtBrandt/winningthespacerace/blob/fae6f985cbf9332710ddef0954bd100d8e564277/spacex_dash_app.py

Predictive Analysis (Classification)

- **Data Preparation**:
  - **Why**: Cleaned and normalized the dataset, applied one-hot encoding for categorical variables, and split the data into training and test sets.
  - **Key Task**: Ensured all features were in numerical format for compatibility with classification models.
- **Model Selection**:
  - **Why**: Trained and tested multiple classifiers: Logistic Regression (LR), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Decision Tree.
  - **Key Task**: Used accuracy as the primary evaluation metric for comparing models.
- **Evaluation**:
  - **Why**: Evaluated models using test accuracy. Logistic Regression, SVM, and KNN achieved the highest accuracy at 83.33%, while the Decision Tree had a lower accuracy of 66.67%.
- **Best Performing Model**:
  - **Why**: Logistic Regression was selected as the best performing model due to its high accuracy.
  - **Key Task**: Focused on Logistic Regression for further improvements, using cross-validation and hyperparameter tuning to confirm the model's performance.
- **GitHub URL**: https://github.com/BrandtBrandtBrandt/winningthespacerace/blob/0df073150df1963ff33c32d83a1df7faad0ef5fc/SpaceX_Machine%20Learning%20Prediction_Part_5%20(2).ipynb



Results

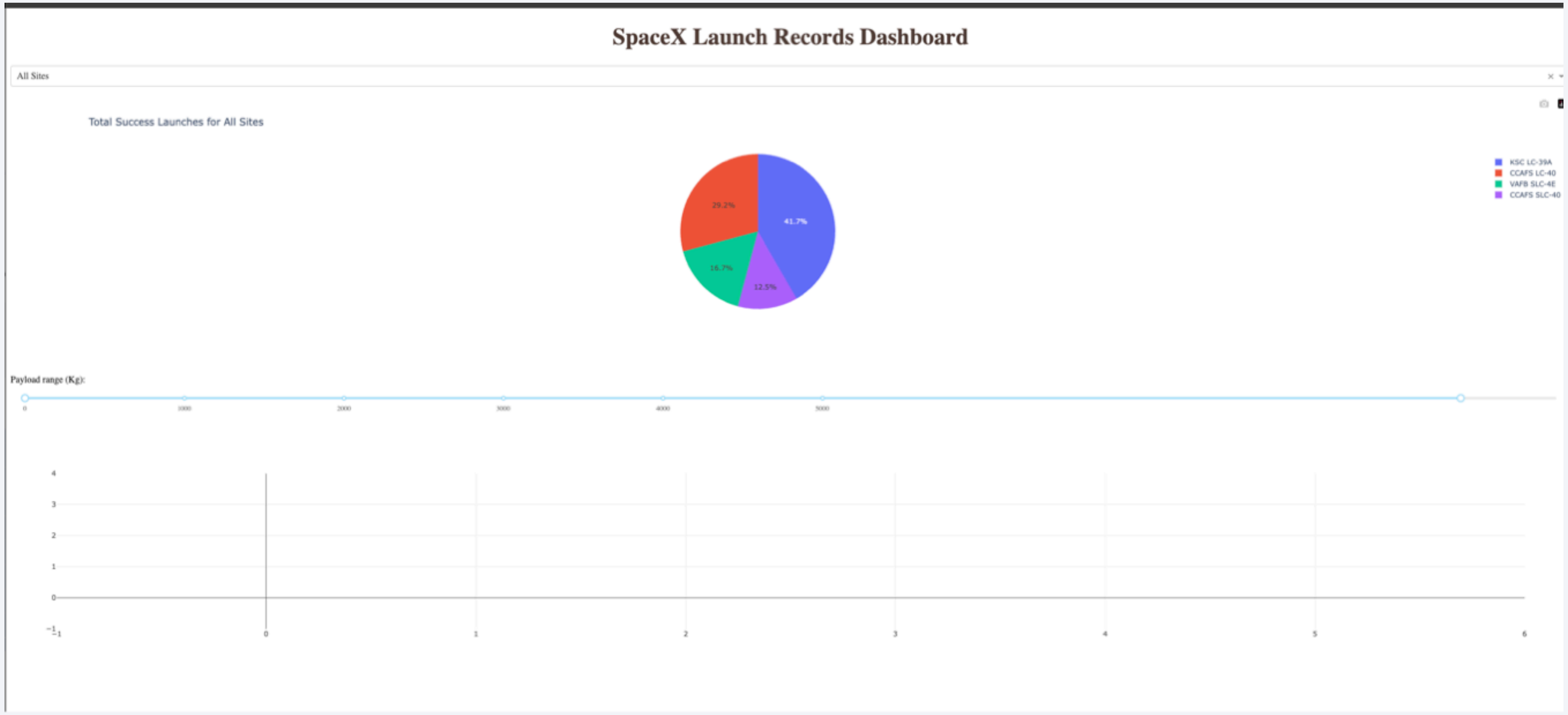
Results: Exploratory Data Analysis (EDA)

- **Launch Success Rates by Site**:
  - **Insight**: Certain launch sites (e.g., KSC LC-39A and CCAFS SLC-40) have higher success rates compared to others, indicating site-specific factors may influence launch outcomes.
- **Payload Mass vs. Success**:
  - **Insight**: Higher payload masses generally show a positive correlation with successful landings, though success is still possible with lower payloads depending on other conditions.
- **Orbit Type vs. Success**:
  - **Insight**: Orbits such as LEO (Low Earth Orbit) exhibit higher success rates, while orbits like GTO (Geostationary Transfer Orbit) show more mixed results, suggesting certain orbits are more challenging for successful landings.
- **Flight Number vs. Success**:
  - **Insight**: Success rates tend to increase as the number of flights increases, reflecting SpaceX's improvements and learning curve over time.
- **Yearly Trends in Success**:
  - **Insight**: The success rate has consistently improved over the years, especially since 2013, showing the technological advancements and operational optimizations by SpaceX.

These insights from EDA provided a strong foundation for building the classification model and understanding key factors influencing launch success.

Results

Interactive Analytics



Results

Results: Predictive Analysis

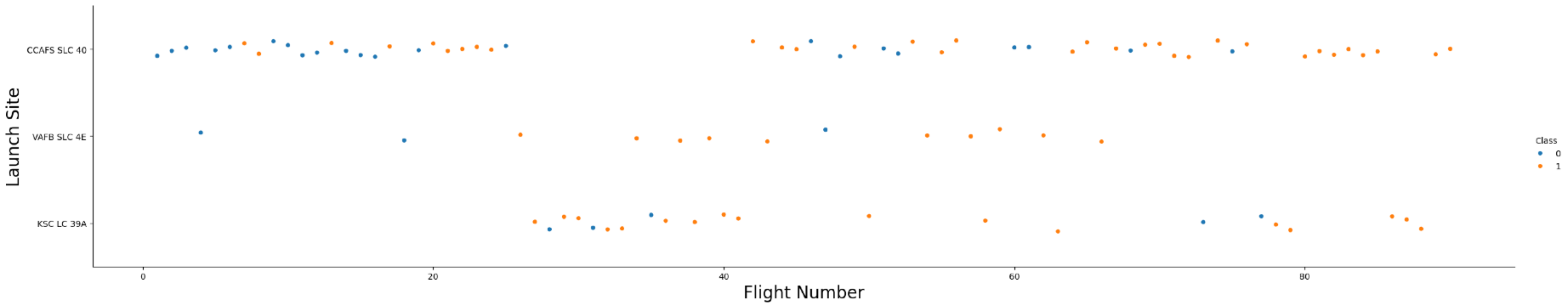
- **Best Performing Model**:
  - **Logistic Regression** achieved the highest accuracy at **83.33%**, outperforming other models like Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Decision Tree.
- **Model Evaluation**:
  - Logistic Regression, SVM, and KNN all achieved **83.33% accuracy**, while the Decision Tree performed worse with **66.67% accuracy**.
  - Evaluation metrics such as precision, recall, and F1-score also supported the selection of Logistic Regression as the best model.
- **Hyperparameter Tuning**:
  - **GridSearchCV** was used to optimize Logistic Regression parameters, specifically tuning the regularization strength (C), resulting in the most effective configuration.
- **Key Findings**:
  - The classification model successfully predicted the likelihood of a successful launch based on factors such as payload mass, launch site, and booster version.
  - The model highlights the importance of payload mass and the launch site in determining the success of a Falcon 9 first-stage landing.
- **Further Improvements**:
  - With additional data or more sophisticated algorithms, there may be potential to further improve the accuracy beyond the current 83.33%.

These predictive results provide a strong foundation for forecasting future SpaceX launch outcomes.

Section 2

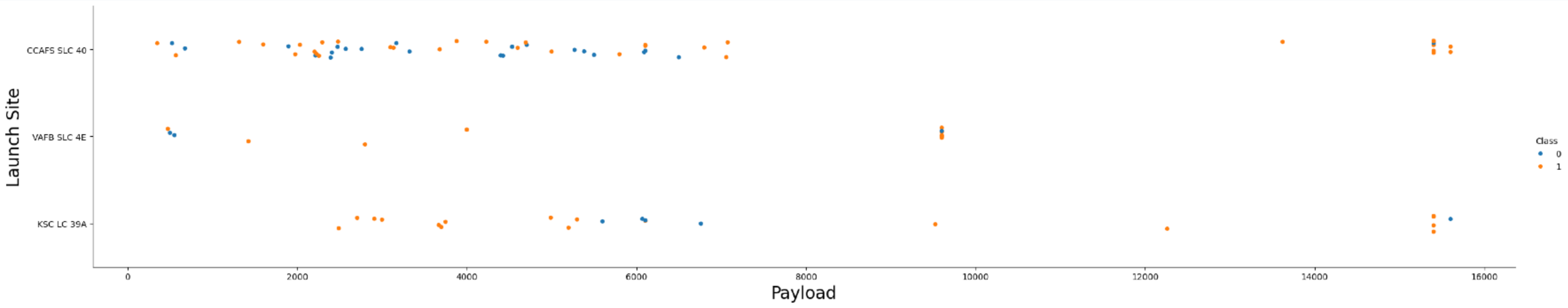
Insights drawn from EDA

Flight Number vs. Launch Site



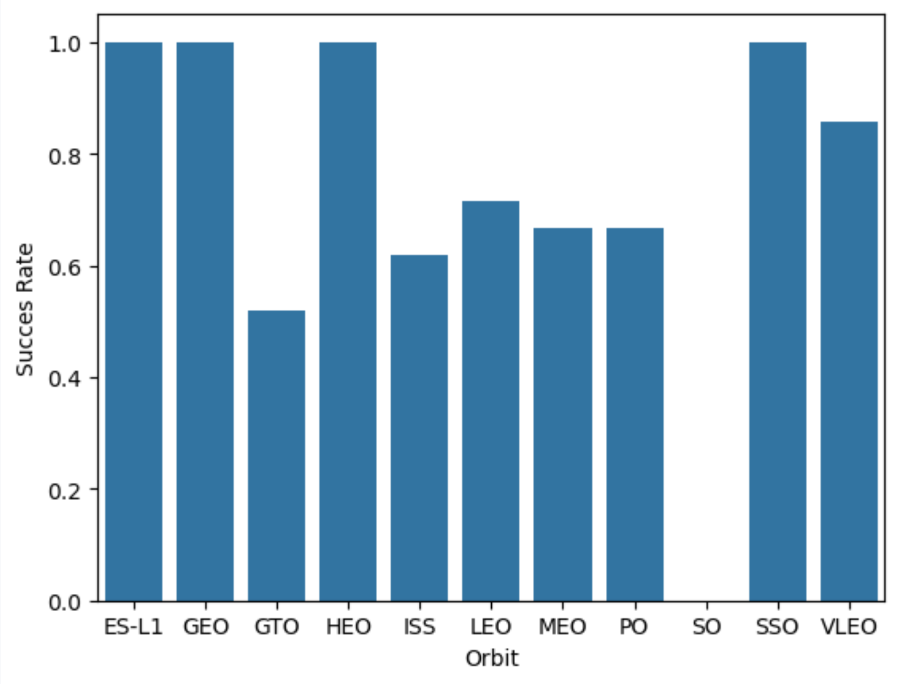
- As the **Flight Number** increases, the likelihood of a successful launch tends to improve across most launch sites, suggesting that experience and technological improvements play a role in the success rates.
- **CCAFS SLC-40** and **KSC LC-39A** exhibit higher success rates overall, while **VAFB SLC-4E** shows more variability in outcomes.

Payload vs. Launch Site



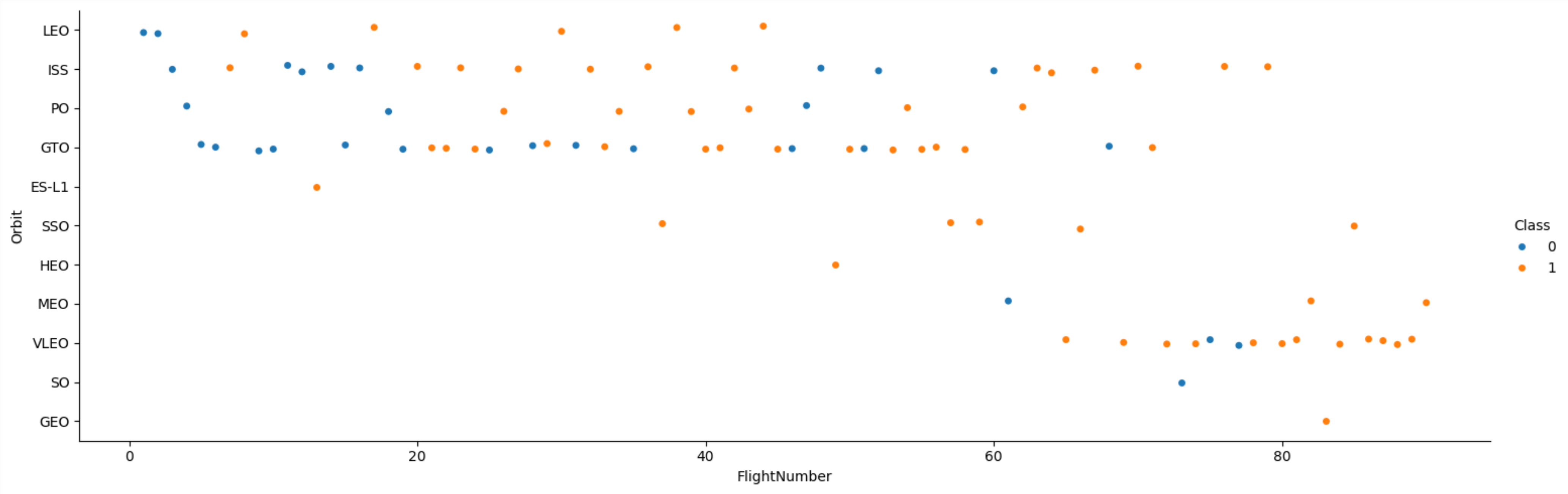
A high payload (above 7000) seems to indicate a higher succes rate, independent of launch site.

Success Rate vs. Orbit Type



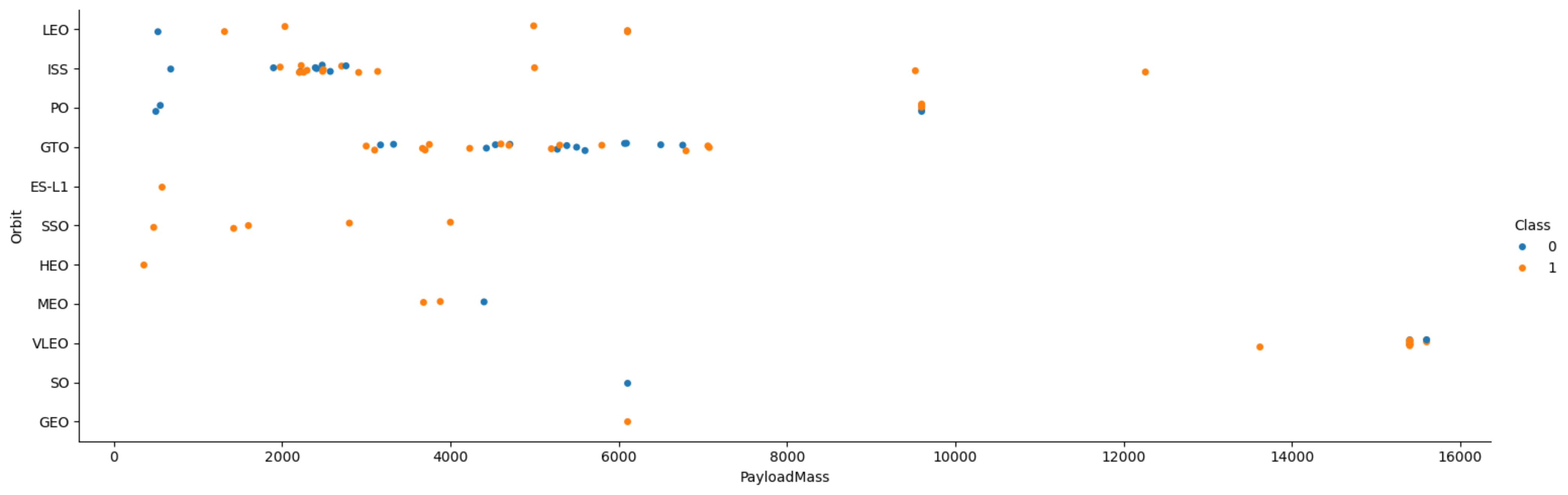
ES-L1, GEO and HEO has the highest success rate, with SSO following closely.

Flight Number vs. Orbit Type



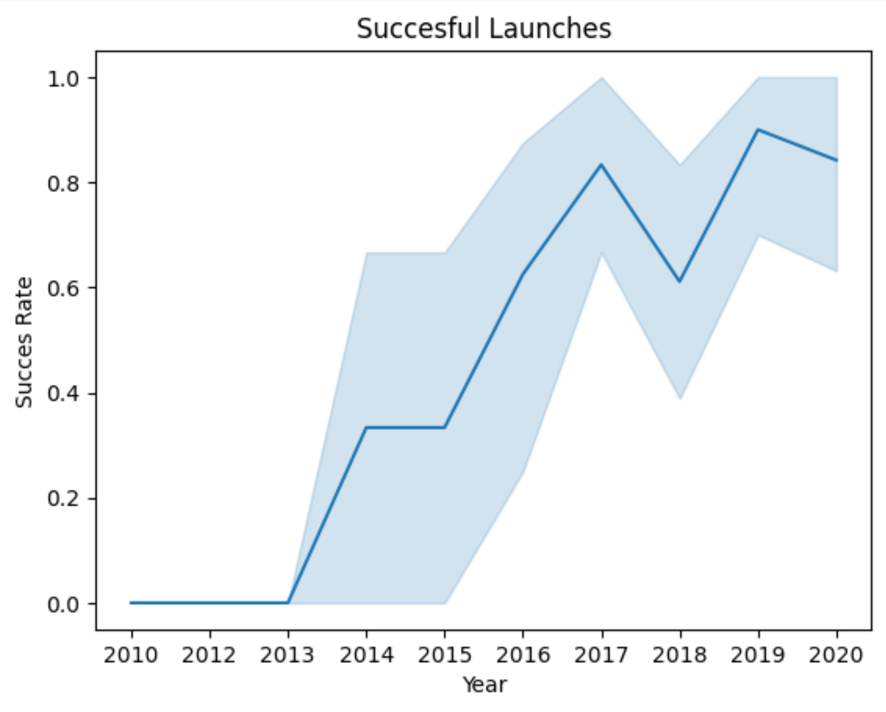
The succesrate has increased for all orbits, as the flight number (experience) increases. The more succesfull orbits can be explained by the fact that they mostly show up in later flights.

Payload vs. Orbit Type



With heavy payloads the successful landing or positive landing rate are more for Polar,LEO and ISS. However, for GTO, it's difficult to distinguish between successful and unsuccessful landings as both outcomes are present.

Launch Success Yearly Trend



you can observe that the success rate since 2013 kept increasing till 2020

All Launch Site Names

- Running %sql SELECT DISTINCT(Launch_Site) FROM SPACEXTBL; showed the following four distinct launch sites

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

Launch Site Names Begin with 'CCA'

- Running %sql SELECT * FROM SPACEXTBL WHERE Launch_Site LIKE 'CCA%' LIMIT 5; showed the following.

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

Running %sql SELECT SUM(PAYLOAD_MASS__KG_) FROM SPACEXTBL WHERE Customer LIKE ‘%NASA%'; showed the total mass carried by boosters for NASA.

SUM(PAYLOAD_MASS__KG_)
107010

Average Payload Mass by F9 v1.1

Running %sql SELECT AVG(PAYLOAD_MASS__KG_) FROM SPACEXTBL WHERE Booster_Version LIKE '%F9 v1.1%’; showed the average payload mass carried by booster F9 v1.1 to be 2534.67 KG

AVG(PAYLOAD_MASS__KG_)
2534.6666666666665

First Successful Ground Landing Date

- %sql SELECT MIN(Date) FROM SPACEXTBL WHERE Landing_Outcome IS "Success (ground pad)";
- The date when the first succesful landing outcome in ground pad was achieved was 2015-12-22

MIN(Date)
2015-12-22

Successful Drone Ship Landing with Payload between 4000 and 6000

- Boosters which have successfully landed on drone ship and had payload mass greater than 4000 but less than 6000:
  - F9 FT B1022
  - F9 FT B1026
  - F9 FT B1021.2
  - F9 FT B1031.2

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

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

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

Total Number of Successful and Failure Mission Outcomes

- Out of 101,only one mission was a complete failure.

```
%%sql SELECT
    CASE
        WHEN Mission_Outcome LIKE '%Success%' THEN 'Success'
        WHEN Mission_Outcome LIKE '%Failure%' THEN 'Failure'
        ELSE 'Other'
    END AS Outcome_Category,
    COUNT(Mission_Outcome) AS Outcome_Count
FROM SPACEXTBL
GROUP BY Outcome_Category;
```

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

Outcome_Category	Outcome_Count
Failure	1
Success	100

Boosters Carried Maximum Payload

- Names of the booster which have carried the maximum payload mass

```
%%sql SELECT Booster_Version
FROM SPACEXTBL
WHERE PAYLOAD_MASS__KG_ = (SELECT MAX(PAYLOAD_MASS__KG_) FROM SPACEXTBL);
```

```
* 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

2015 Launch Records

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

```
%%sql SELECT SUBSTR(Date, 6, 2) AS Month, Booster_Version, Launch_Site
FROM SPACEXTBL
WHERE Landing_Outcome = 'Failure (drone ship)'
AND SUBSTR(Date, 1, 4) = '2015';
```

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

Month	Booster_Version	Launch_Site
01	F9 v1.1 B1012	CCAFS LC-40
04	F9 v1.1 B1015	CCAFS LC-40

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

- 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 Outcome_Count
FROM SPACEXTBL
WHERE Date BETWEEN '2010-06-04' AND '2017-03-20'
GROUP BY Landing_Outcome
ORDER BY Outcome_Count DESC;
```

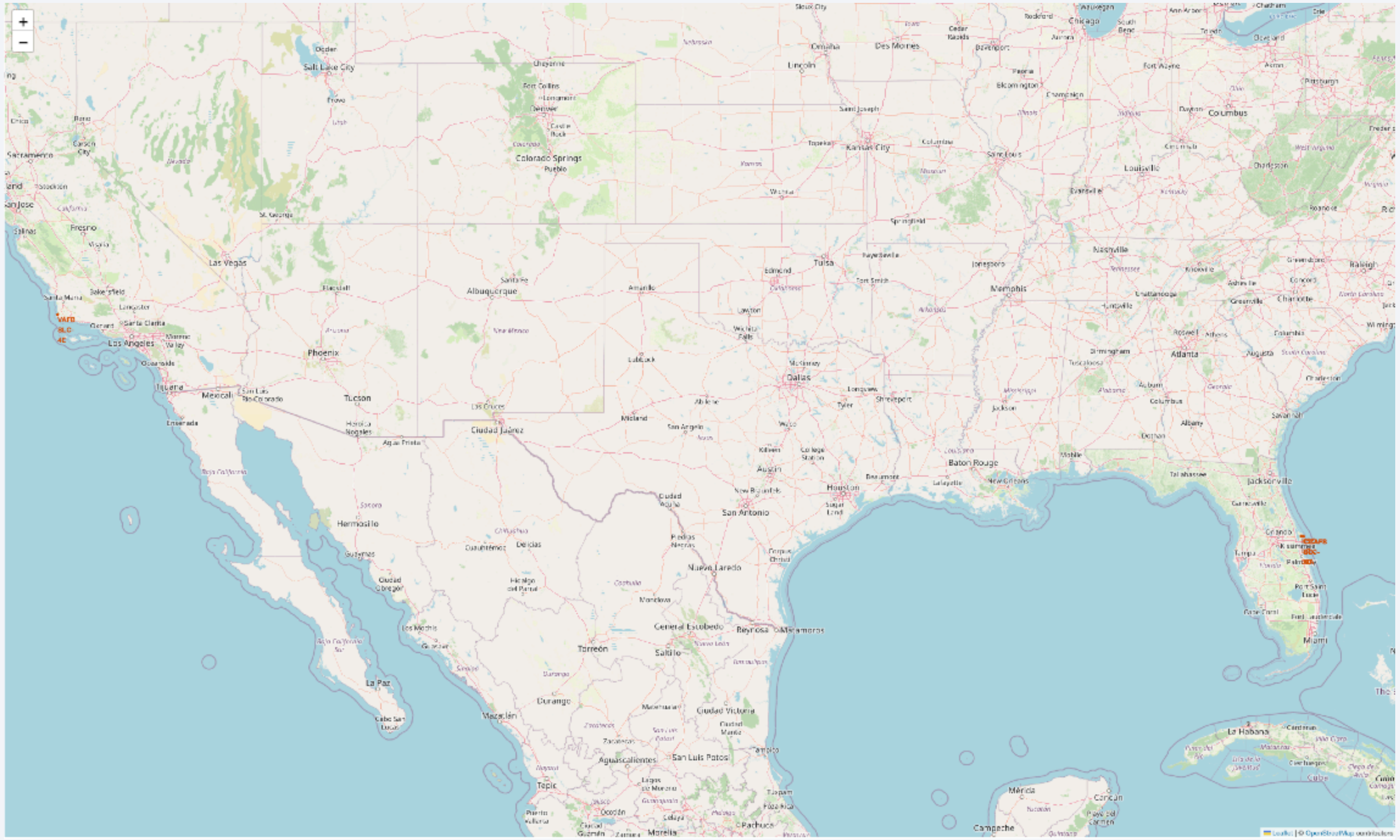
```
* sqlite:///my_data1.db
Done.
```

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

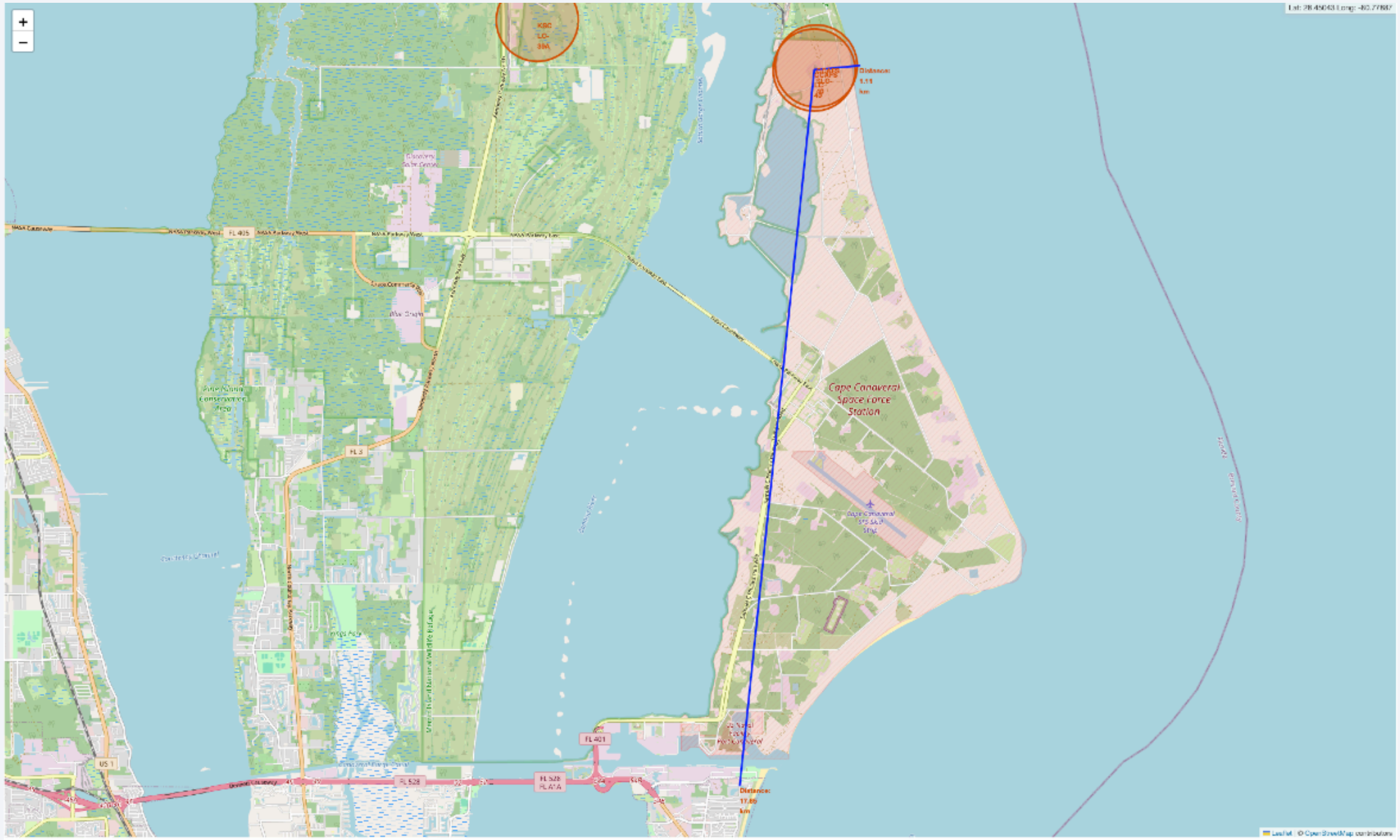
Launch Sites Proximities Analysis

SpaceX Launch Sites



There are sights one the east and west coast of America. Both close to the water.

CAFFS Launch Site



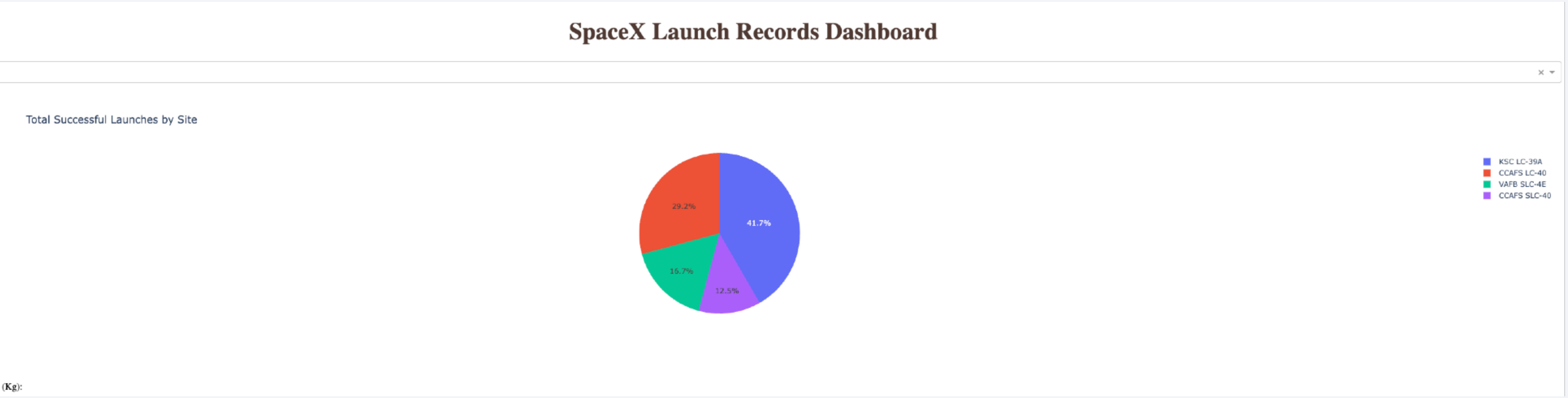
- Launch site CCAFS with distances to the ocean as well as nearest city

Section 4

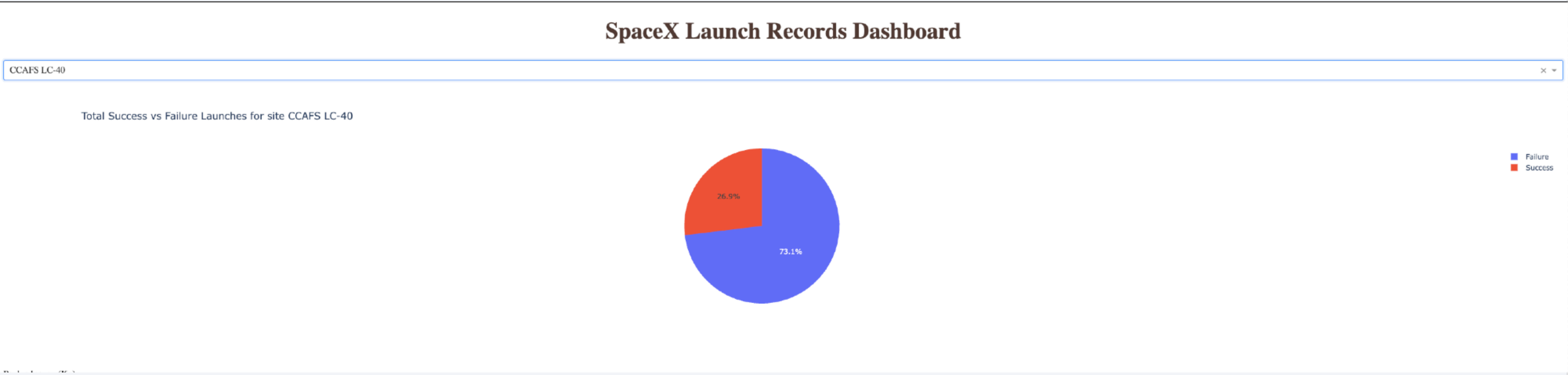
Build a Dashboard with Plotly Dash

Total Succesful Launches by Site

- KSK LC-39A has the highest succes rate

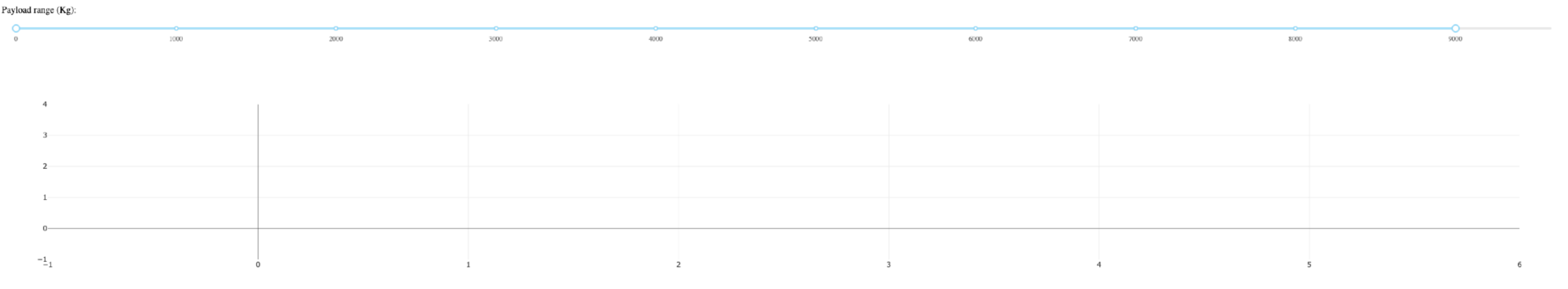


Succes rate for CCAFS LC-40



Payload vs. Launch Outcome

- Something is not working in my code, it would seem :’(

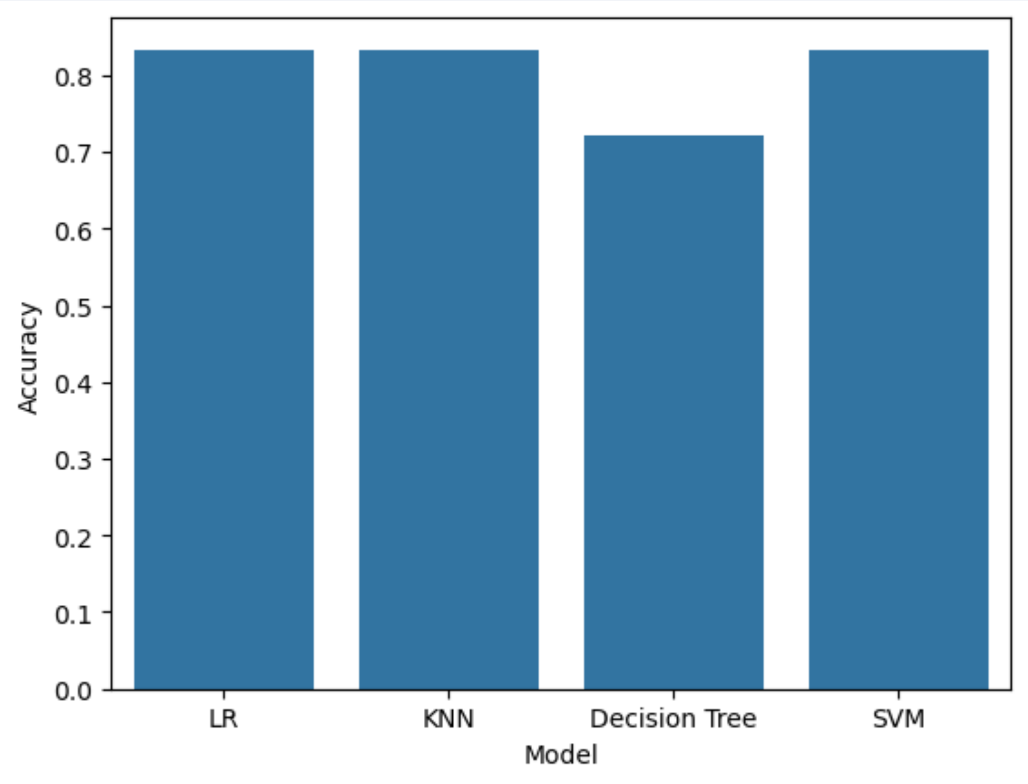


Section 5

Predictive Analysis (Classification)

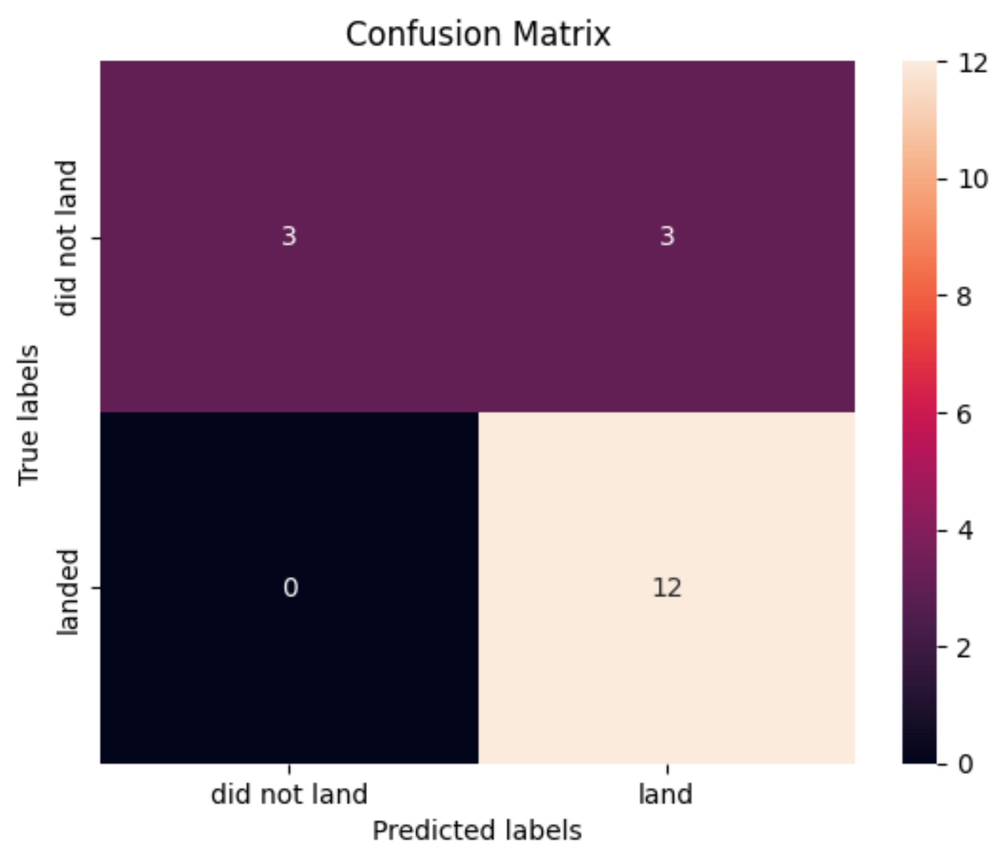
Classification Accuracy

- LR test accuracy: 0.8333333333333334
- SVM test accuracy: 0.8333333333333334
- Decision Tree test accuracy: 0.7222222222222222
- KNN test accuracy: 0.8333333333333334
- Best Model: **Logistic Regression** with accuracy: 0.8333333333333334



Confusion Matrix

- Examining the confusion matrix, we see that logistic regression can distinguish between the different classes. We see that the problem is false positives.
- Overview:
- True Postive - 12 (True label is landed, Predicted label is also landed)
- False Postive - 3 (True label is not landed, Predicted label is landed)



Conclusions

Launch Sites and Success Rates:

- Certain launch sites, like CCAFS SLC-40 and KSC LC-39A, consistently show higher success rates, suggesting site-specific factors may influence the likelihood of successful landings.

Payload Mass and Launch Success:

- There is a strong correlation between higher payload mass and successful launches, though success is possible with a wide range of payloads depending on other variables.

Logistic Regression as the Best Model:

- Logistic Regression, after hyperparameter tuning, was identified as the best performing model, achieving an accuracy of 83.33% in predicting launch outcomes.

Potential for Future Improvements:

- Further data collection and model refinement, including more complex machine learning models, could improve the accuracy of predictions and help SpaceX optimize launch performance.

Thank you!