IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Lau Wen Jun

2nd Feb 2024

Outline

- Executive Summary
- Introduction
- Methodology
- Results
- Conclusion

Executive Summary

- In this project, we aimed to predict the success of the first stage landing of SpaceX's Falcon 9 rocket, a critical factor influencing the cost of launches. Utilizing SpaceX's public API and web scraping techniques, we gathered comprehensive data on rocket launches, including key details such as payload masses, launch sites, and outcomes of each mission. Through meticulous data wrangling, we addressed data inconsistencies and missing values, preparing a clean dataset for analysis. The exploratory data analysis (EDA) phase revealed insightful trends, such as the evolution of launch success rates over time and the average payload capacities of different Falcon 9 boosters.
- The predictive analysis phase involved the development and comparison of multiple machine learning models, including Logistic Regression, SVM, Decision Trees, and KNN. By employing GridSearchCV for hyperparameter tuning, we sought to optimize each model's performance, using accuracy as the primary metric for evaluation. This iterative process of model refinement highlighted the Decision Tree classifier as the most effective in predicting landing success, suggesting a strong potential for accurately forecasting mission outcomes based on historical launch data.
- The culmination of this project demonstrated the feasibility of using machine learning techniques to predict the outcome of Falcon 9 first-stage landings with high accuracy. The findings not only offer SpaceX a tool to estimate launch costs more effectively but also provide valuable insights for competitors looking to bid against SpaceX for rocket launches. By leveraging the predictive power of machine learning, stakeholders in the commercial space launch industry can make informed decisions, enhancing the economic viability and safety of space missions.

Introduction

Project Background and Context:

- SpaceX has revolutionized the economics of space travel with its Falcon 9 rocket, which is advertised on its website with a launch cost of 62 million dollars. This price point is significantly lower than other providers, who typically charge upwards of 165 million dollars per launch. One of the key factors contributing to SpaceX's competitive pricing is the reusability of the Falcon 9's first stage, which can lead to considerable cost savings. The ability to predict whether this first stage will successfully land is not only a technical challenge but also has substantial financial implications. Accurate predictions can help determine the actual cost of a launch and are vital for any potential competitor or collaborator looking to bid against or work with SpaceX for rocket launch services.

Problem Statement:

- The challenge lies in developing an accurate predictive model that can assess the likelihood of the first stage of Falcon 9 landing successfully. Such a model would enable us to estimate the cost-effectiveness of the launches and provide a significant competitive edge in the commercial spaceflight market. By leveraging machine learning algorithms and a robust dataset of past launches, we aim to create a machine learning pipeline that can reliably predict the outcome of the first stage landing, thereby estimating the financial aspects of SpaceX launches for potential market entrants or partners.

Section 1

Methodology

Methodology

Executive Summary

- Data collection methodology:
    - The data was collected through APIs and web scraping, which provided a structured and automated approach to data extraction from various online sources. These methods ensured the efficient gathering of real-time and historical data necessary for the analysis.
    - Source:
        - SpaceX API (https://api.spacexdata.com/v4/rockets/)
        - Wikipedia (https://en.wikipedia.org/wiki/List_of_Falcon/_9/_and_Falcon_Heavy_launches), using web scraping.
- Perform data wrangling
    - Data wrangling involved cleaning and structuring the raw data to make it suitable for analysis. This process included handling missing values, correcting data formats, normalizing data, and removing duplicates to ensure data quality and consistency.

Methodology

Executive Summary

- Perform exploratory data analysis (EDA) using visualization and SQL
- Perform interactive visual analytics using Folium and Plotly Dash
- Perform predictive analysis using classification models
    - The predictive analysis utilized classification models such as Logistic Regression, Decision Trees, Support Vector Machines (SVM), and K-Nearest Neighbors (KNN). These models were trained and tested on the dataset to predict categorical outcomes effectively.
    - The models were built by initializing with their respective algorithms and then trained using the processed data. Hyperparameter tuning was conducted to optimize each model's performance, often using methods like cross-validation and grid search. The evaluation of the models was based on their predictive accuracy and other performance metrics, which helped in selecting the best deployment model.

Data Collection – SpaceX API

- The data collection process for SpaceX information entailed the use of the SpaceX REST API, a publicly available service that provides access to detailed information about rocket specifications, launch details, and other related data. The use of this API involves sending requests to endpoints, such as https://api.spacexdata.com/v4/rockets/, which then return data in a structured format like JSON.
- https://github.com/wenjun78/ibmdatasciencecap/blob/main/jupyter-labs-spacex-data-collection-api.ipynb

Flow:

1. Identify Data Requirements: Determine the specific type of data needed for analysis (e.g., rocket details, launch outcomes, payload masses).
2. Access SpaceX API: Utilize the RESTful endpoints provided by SpaceX to access the desired data. For example, to get information about rockets, send a GET request to the /v4/rockets endpoint.
3. Send API Request: Make the HTTP GET request to the API endpoint using a tool or library that can handle HTTP requests (like curl, Python's requests library, etc.).
4. Receive and Parse Response: The API responds with the data in JSON format, which must be parsed to extract the relevant information.

Data Collection - Scraping

- The data collection process involving web scraping from Wikipedia
- https://github.com/wenjun78/ibmdatasciencecap/blob/main/jupyter-labs-webscraping.ipynb

Flow Chart:

1. Identify Data Source: Choose the specific Wikipedia page(s) containing the relevant data about SpaceX launches.
2. Inspect HTML Structure: Use browser tools to inspect the HTML structure of the page to understand how the data is organized (e.g., tables, lists, divs).
3. Write Scraper: Develop a script using a web scraping library (such as BeautifulSoup in Python) to programmatically navigate the structure of the web page and extract the needed data.
4. Execute Scraper: Run the script to make an HTTP request to the Wikipedia page and parse the HTML content returned.
5. Extract Data: The script processes the HTML to extract the data, typically found within tables or infoboxes on Wikipedia pages.

Data Wrangling

- The data wrangling process typically involves several key steps to convert raw data into a more useful format for analysis. Here's a summary of this process using key phrases and a description that could be visualized in a flowchart:
    1. Data Inspection: Examine the raw data for quality and structure issues. Key phrases: "Identify missing values", "Detect outliers", "Spot corrupt data".
    2. Data Cleaning: Address quality issues found during inspection. Key phrases: "Handle missing data", "Correct data errors", "Remove duplicates".
    3. Data Transformation: Modify data to a suitable format or structure for analysis. Key phrases: "Normalize data", "Aggregate data", "Reshape data".
    4. Data Enrichment: Enhance data by merging with other datasets or adding derived attributes. Key phrases: "Merge datasets", "Create calculated fields".
    5. Data Reduction: Reduce the volume but produce the same or similar analytical results. Key phrases: "Feature selection", "Dimensionality reduction".
    6. Data Validation: Ensure that the data cleaning and transformation steps have been properly applied. Key phrases: "Validate against rules", "Review transformations".
    7. Data Storage: Save the wrangled data in a storage system for analysis. Key phrases: "Export to database", "Write to CSV/JSON".
- https://github.com/wenjun78/ibmdatasciencecap/blob/main/labs-jupyter-spacex-Data%20wrangling.ipynb

EDA with Data Visualization

- Pie Charts: Used to illustrate the proportion of successful launches at different sites and the success-failure ratio at a specific launch site. These charts provide a quick visual comparison of categorical data.
- Bar Charts: Deployed to compare the accuracy of different classification models. They effectively showcase differences in performance metrics like accuracy between the models.
- Scatter Plots: Utilized for visualizing the correlation (or lack thereof) between payload mass and launch success across various booster versions. Scatter plots are ideal for spotting trends, clusters, and outliers in numerical data.
- https://github.com/wenjun78/ibmdatasciencecap/blob/main/jupyter-labs-eda-dataviz.ipynb.jupyterlite.ipynb

EDA with SQL

SQL Queries Performed:

- Retrieved distinct launch site names.
- Displayed records for launch sites beginning with 'CCA'.
- Calculated total and average payload masses for specific criteria.
- Identified dates of the first successful landing on a ground pad.
- Listed boosters successful in drone ship landings within a certain payload range.
- Counted total mission outcomes.
- Found boosters carrying the maximum payload.
- Extracted specific failure landing outcomes for 2015.
- Ranked landing outcomes by frequency within a date range.

https://github.com/wenjun78/ibmdatasciencecap/blob/main/jupyter-labs-eda-sql-coursera_sqllite.ipynb

Build an Interactive Map with Folium

- Markers: Used to pinpoint precise locations such as launch sites, enabling users to identify each site's geographical coordinates on the map.
- Circles: Highlighted specific areas of interest around given coordinates, such as safety zones or areas of significant activity like the NASA Johnson Space Center, allowing for a spatial understanding of the importance or influence of these areas.
- Marker Clusters: Grouped multiple related events or points, such as multiple launches from a single launch site, to simplify the map's visual representation and enhance user readability when dealing with a high density of markers.
- Lines: Drew attention to the distance or connection paths between two points, facilitating the visualization of routes or correlations between different geographic locations.

https://github.com/wenjun78/ibmdatasciencecap/blob/main/lab_jupyter_launch_site_location.jupyterlite.ipynb

Build a Dashboard with Plotly Dash

- Interactive Dashboard: Developed using Plotly Dash to create a user-friendly interface that allows for dynamic interaction with the data.
  - Pie Charts: Incorporated to display the distribution of total launches by various launch sites, providing an immediate visual representation of each site's launch activity proportion to facilitate quick comparison and insights.
  - Scatter Graphs: Showcased to illustrate the relationship between launch outcomes and payload masses across different booster versions. This helps in identifying patterns or trends, such as whether heavier payloads have a lower or higher success rate, and how different booster versions perform in terms of launch success relative to payload mass.

https://github.com/wenjun78/ibmdatasciencecap/blob/main/spacex_dash_app.py

Predictive Analysis (Classification)

1. Data Preparation: Utilized Numpy and Pandas for data manipulation, performed necessary transformations and preprocessing steps, and split the data into training and testing sets for model evaluation.

2. Model Construction: Developed various machine learning models, including Logistic Regression, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Decision Trees.

3. Hyperparameter Tuning: Applied GridSearchCV to systematically work through multiple combinations of parameter tunes, cross-validating as it went to determine which tune gives the best performance.

4. Performance Evaluation: Selected accuracy as the primary metric to evaluate the models, as it provides a straightforward measure of the models' ability to correctly classify the outcomes.

5. Model Improvement: Conducted feature engineering to create new features or modify existing ones and fine-tuned algorithms to enhance model performance.

6. Best Model Identification: Compared the performance of all models and hyperparameter configurations to identify the most effective classifier for the given dataset.

https://github.com/wenjun78/ibmdatasciencecap/blob/main/Machine%20Learning%20Prediction.ipynb

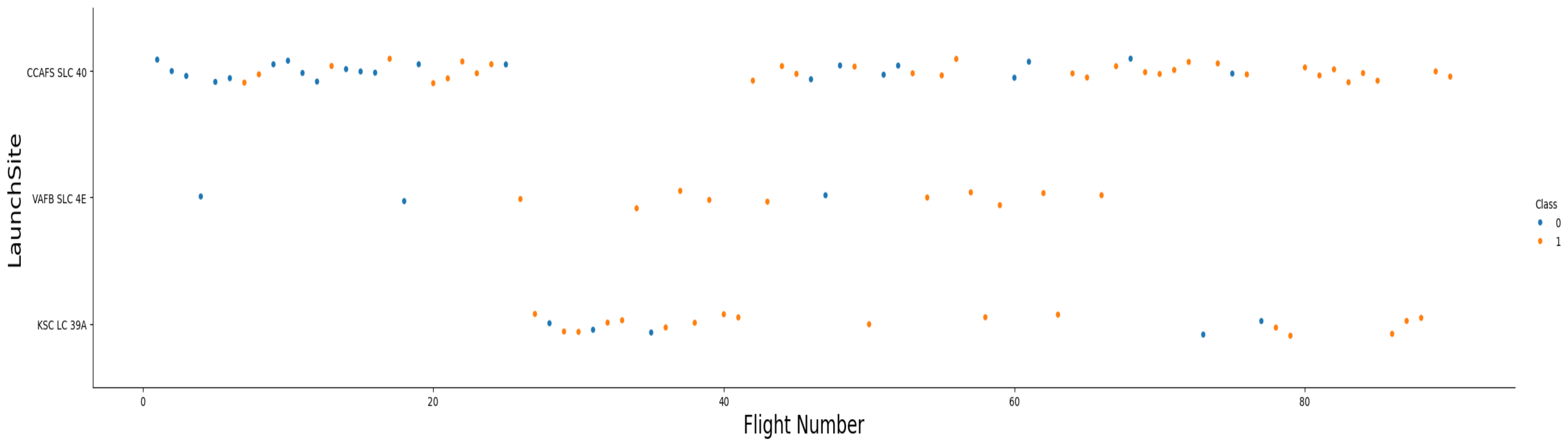
Results

- **Launch Site Usage:** Space X operates from four distinct launch sites, which underscores the organization's operational range and capabilities.
- **Initial Launch Partners:** The initial launches were conducted for Space X and NASA, highlighting early collaboration or self-sponsored missions to establish Space X's presence in the space industry.
- **Booster Payload Capacity:** The F9 v1.1 booster has an average payload capacity of 2,928 kg, indicating its substantial lift capability for various missions.
- **Landing Milestone:** The first successful booster landing on a ground pad occurred in 2015, five years after the initial launch, marking significant progress in reusable rocket technology.
- **Drone Ship Landings:** Multiple Falcon 9 booster versions have achieved successful landings on drone ships even with payloads above the average, demonstrating Space X's advancements in precision landing technology.
- **Mission Success Rate:** Nearly all mission outcomes were successful, which speaks volumes about the reliability and efficiency of Space X's launch operations.
- **Drone Ship Landing Challenges:** In 2015, two specific booster versions, F9 v1.1 B1012 and F9 v1.1 B1015, encountered failures when attempting to land on drone ships, indicating the challenges and risks associated with the early stages of developing reusable rocket technology.
- **Improvement Over Time:** There has been a clear trend of improved landing outcomes over the years, illustrating continuous learning, technological refinement, and increased experience in launch and recovery operations.

Section 2

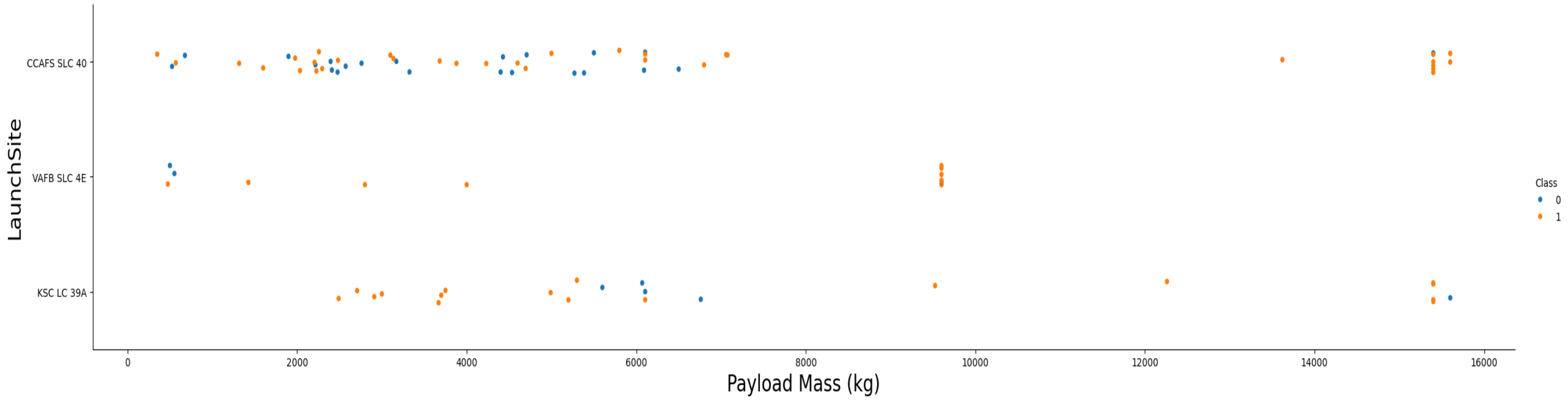
Insights drawn from EDA

Flight Number vs. Launch Site



- Data indicates a positive trend in landing success rate with the increase in flight number, suggesting advancements in technology or expertise over time.

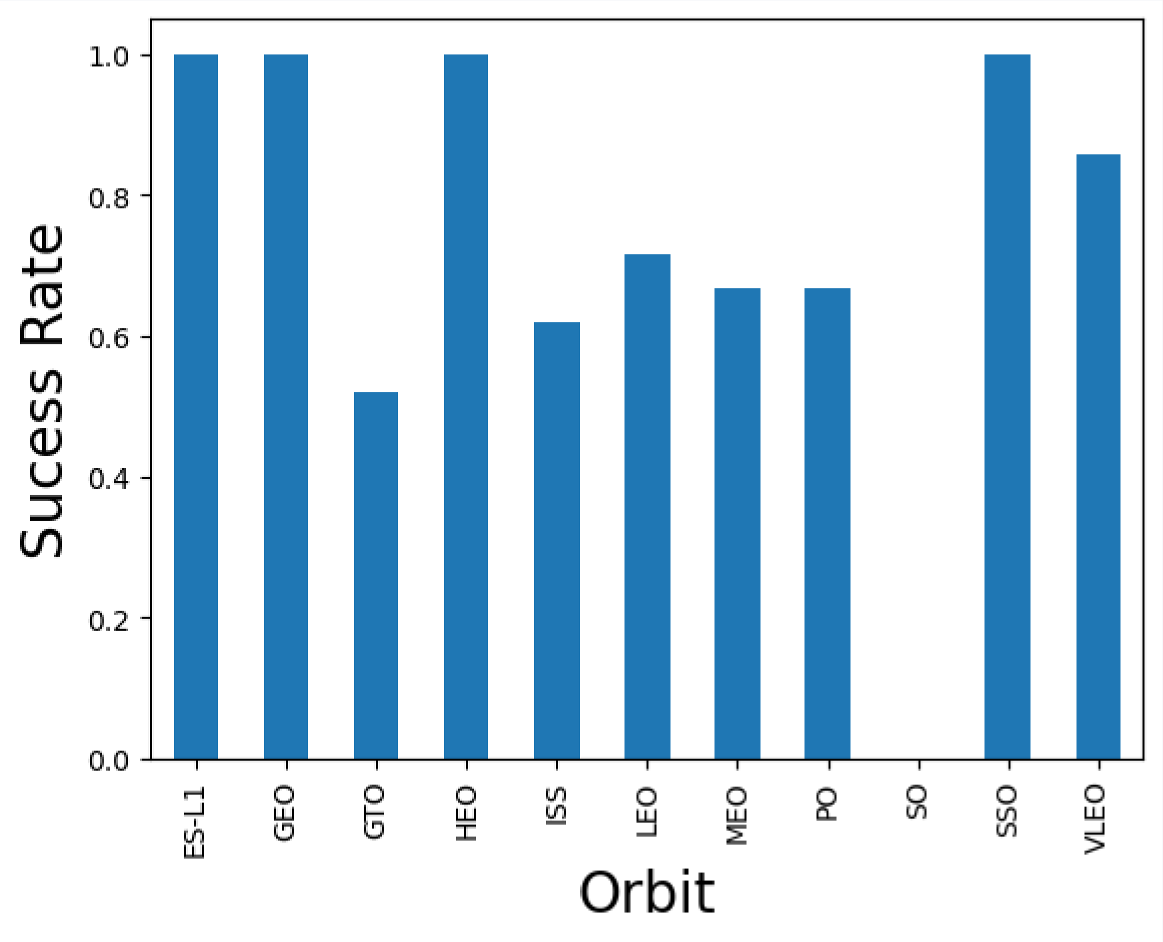
Payload vs. Launch Site



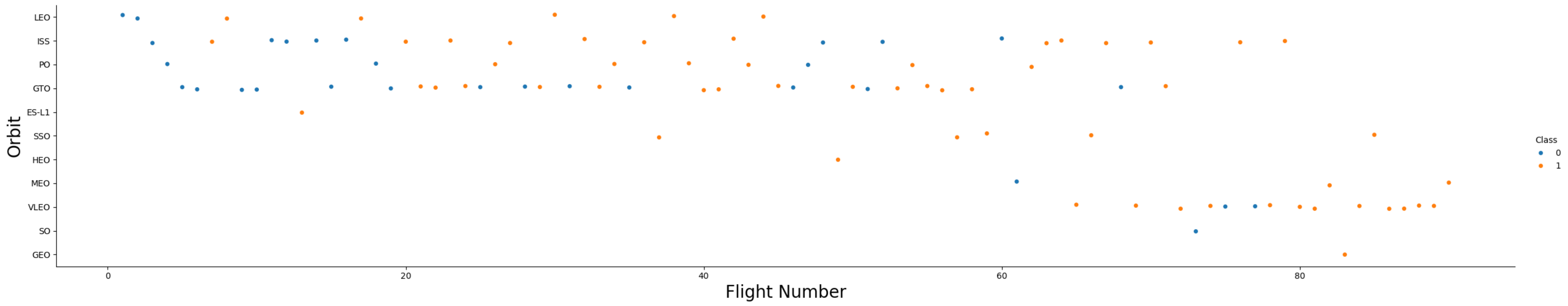
- The majority of launches took place at CCSFS SLC-40, indicating it might be a preferred site due to logistical, technical, or historical reasons, while VAFB SLC 4E saw the least, potentially due to specific mission requirements or site capabilities.

Success Rate vs. Orbit Type

- Certain orbits like ES_L1, GEO, HEO, and SSO have a record of flawless success, possibly hinting at the maturity of technology or favorable conditions for these orbits.

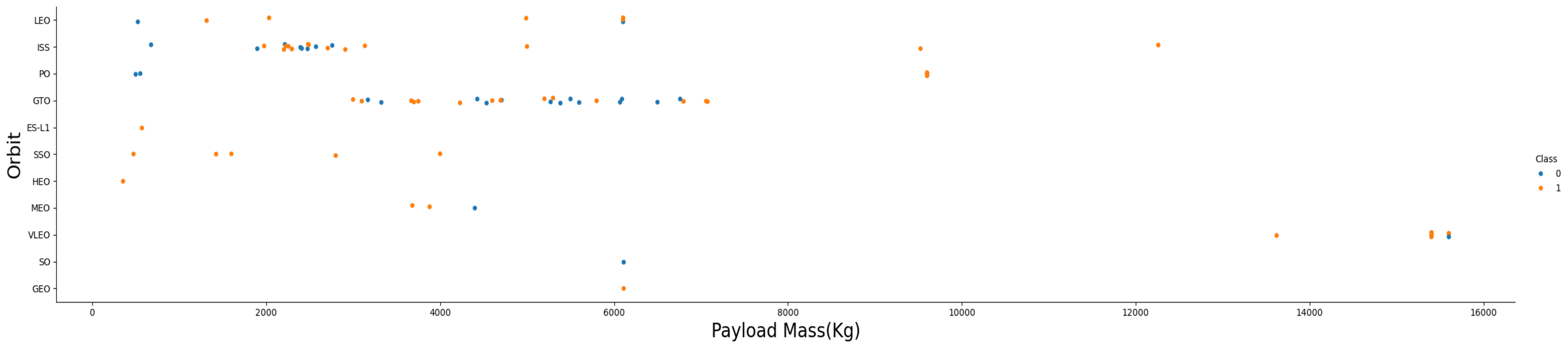


Flight Number vs. Orbit Type



- In LEO orbits, a higher number of flights correlates with success, suggesting experience improves outcomes, whereas in GTO orbits, success appears less correlated with flight frequency, indicating other factors may play a more significant role.

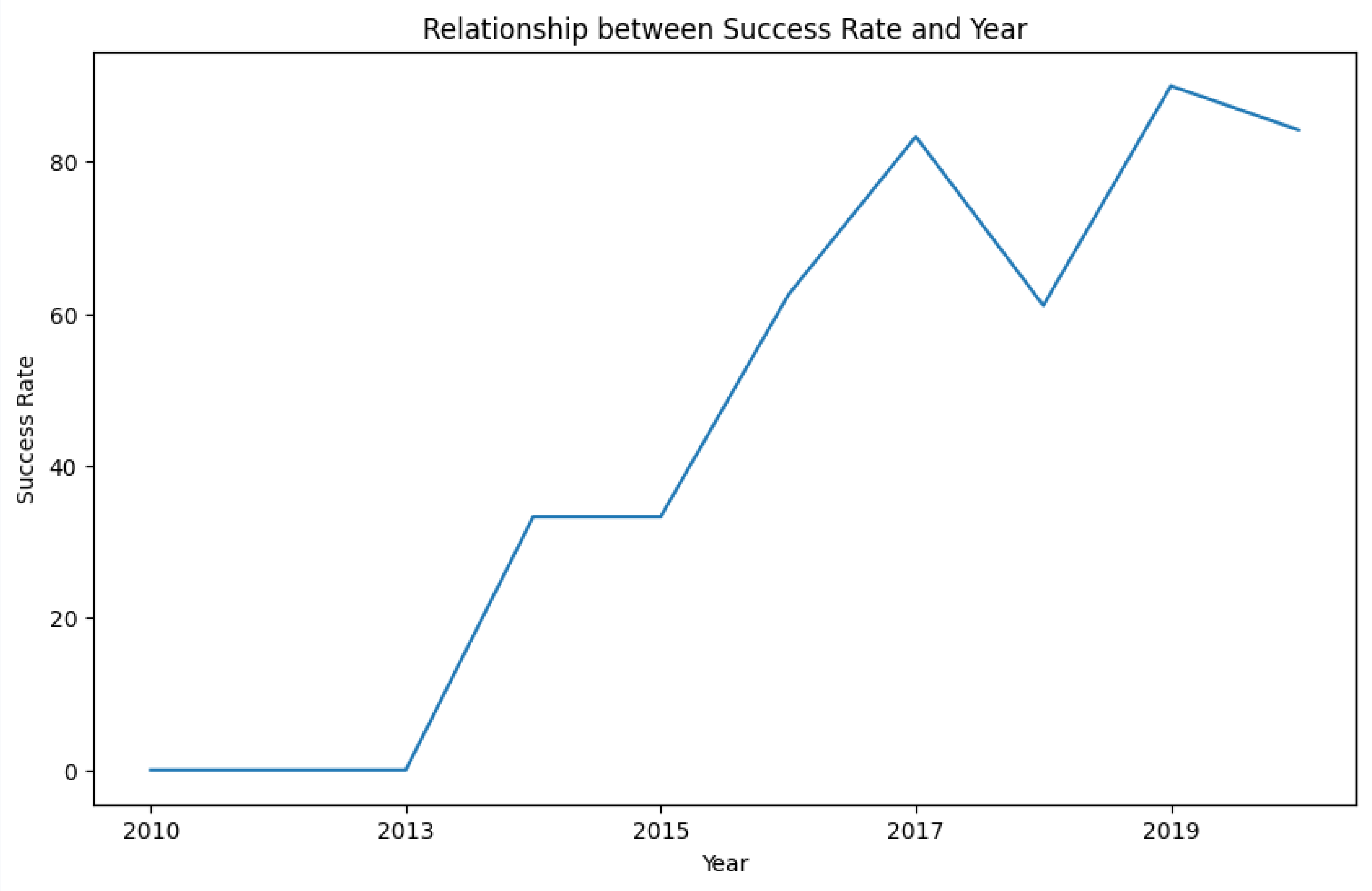
Payload vs. Orbit Type



- Heavy payloads show higher success in Polar, LEO, and ISS orbits, perhaps due to the robust design of missions to these destinations. The GTO orbit shows a mixed result, implying complexity or inherent risk in such missions.

Launch Success Yearly Trend

- There's been a noticeable upward trend in success rate from 2013 to 2020, reflecting continuous improvement in operational processes or technology.



All Launch Site Names

- The SQL query executed lists all distinct SpaceX launch sites, which are specific locations designated for launching spacecraft. The result indicates the diversity of SpaceX's operational launch platforms.

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

```
%sql select distinct(LAUNCH_SITE) from SPACEXTBL
```

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

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 SPACEXTBL 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

- This query filters and displays the first five launch records where the site name begins with 'CCA'. This could be used to analyze the frequency or outcomes of launches from a particular site or set of sites with a common prefix.

Total Payload Mass

- The total payload mass for all SpaceX missions contracted by NASA under the CRS (Commercial Resupply Services) program is calculated. This figure represents the cumulative mass that these specific missions have carried to space, reflecting the scale of cargo delivered by SpaceX for NASA.

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

```
%sql select sum(PAYLOAD_MASS__KG_) from SPACEXTBL where CUSTOMER = 'NASA (CRS)'
```

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

sum(PAYLOAD_MASS__KG_)
45596

Average Payload Mass by F9 v1.1

- The average payload mass carried by the SpaceX Falcon 9 version 1.1 booster is calculated. This metric is useful for understanding the typical carrying capacity of this version of the booster and could be compared against other versions for payload efficiency.

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

```
%sql select avg(PAYLOAD_MASS__KG_) from SPACEXTBL where BOOSTER_VERSION = 'F9 v1.1'
```

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

avg(PAYLOAD_MASS__KG_)
2928.4

First Successful Ground Landing Date

- The query retrieves the earliest date (using the min function) when SpaceX successfully landed a booster on a ground pad. The result indicates that the first successful landing on a ground pad occurred on December 22, 2015.

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

Hint:Use min function

```
%sql select min(DATE) from SPACEXTBL where Landing_Outcome = 'Success (ground pad)'
```

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

min(DATE)
2015-12-22

Successful Drone Ship Landing with Payload between 4000 and 6000

- This SQL query lists the versions of SpaceX boosters that have both successfully landed on a drone ship and carried a payload between 4000 and 6000 kg. Several booster versions meet these criteria.

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 SPACEXTBL where Landing_Outcome = '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

- The query counts the total number of mission outcomes classified as either 'Success' or 'Failure (in flight)'. It shows there have been a total of 99 such mission outcomes.

List the total number of successful and failure mission outcomes

```
%sql select count(MISSION_OUTCOME) from SPACEXTBL where MISSION_OUTCOME = 'Success' or MISSION_OUTCOME = 'Failure (in flight)'
```

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

count(MISSION_OUTCOME)
99

Boosters Carried Maximum Payload

- This query identifies the booster versions that have carried the maximum payload mass, determined by a subquery that finds the highest payload mass in the database.

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

```
%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

- The results display records from 2015, showing months and specific details related to failures in drone ship landings. The substr function is used to extract the month from the date.

Task 9

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.

Note: SQLLite does not support monthnames. So you need to use substr(Date, 6,2) as month to get the months and substr(Date,0,5)='2015' for year.

```
%sql SELECT substr(Date,6,2) as month, DATE,BOOSTER_VERSION, LAUNCH_SITE, [Landing_Outcome] \
FROM SPACEXTBL \
where [Landing_Outcome] = 'Failure (drone ship)' and substr(Date,0,5)='2015';
```

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

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

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

- This query ranks the different types of landing outcomes between specific dates, ordered by frequency. The results show the number of times each type of landing outcome occurred, such as 'No attempt', 'Success (drone ship)', 'Failure (drone ship)', and 'Success (ground pad)'.

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(*) as count_outcomes \
FROM SPACEXTBL \
WHERE DATE between '2010-06-04' and '2017-03-20' group by [Landing_Outcome] order by count_outcomes DESC;
```

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

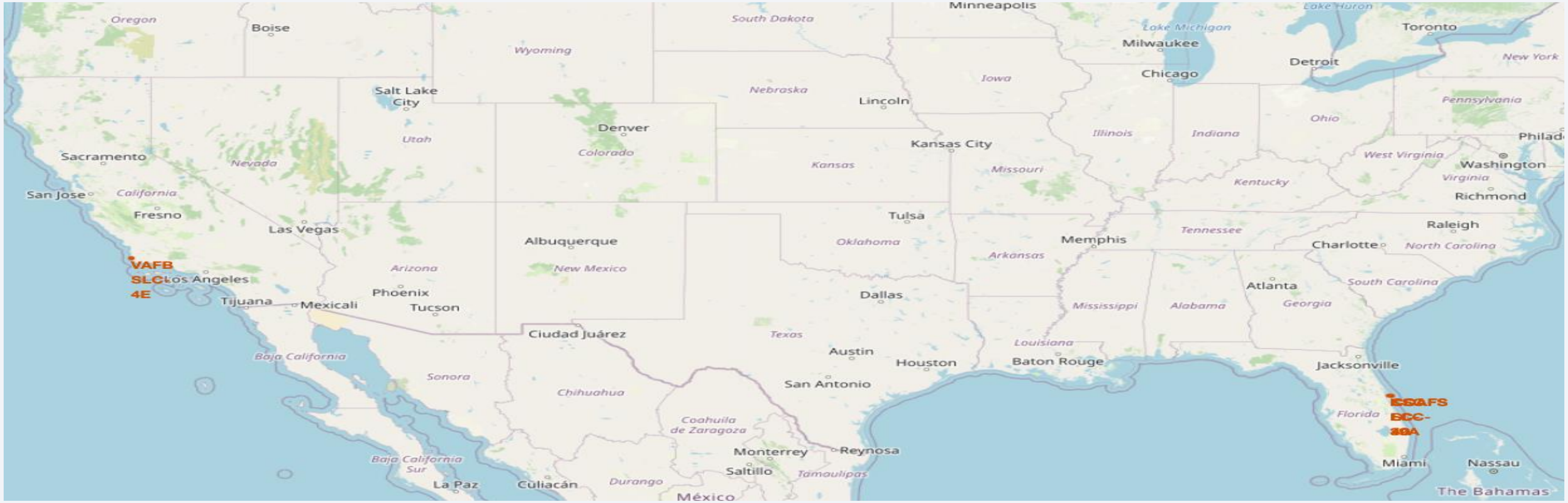
Landing_Outcome	count_outcomes
No attempt	10
Success (drone ship)	5
Failure (drone ship)	5
Success (ground pad)	3
Controlled (ocean)	3
Uncontrolled (ocean)	2
Failure (parachute)	2
Precluded (drone ship)	1

Section 3

Launch Sites Proximities Analysis

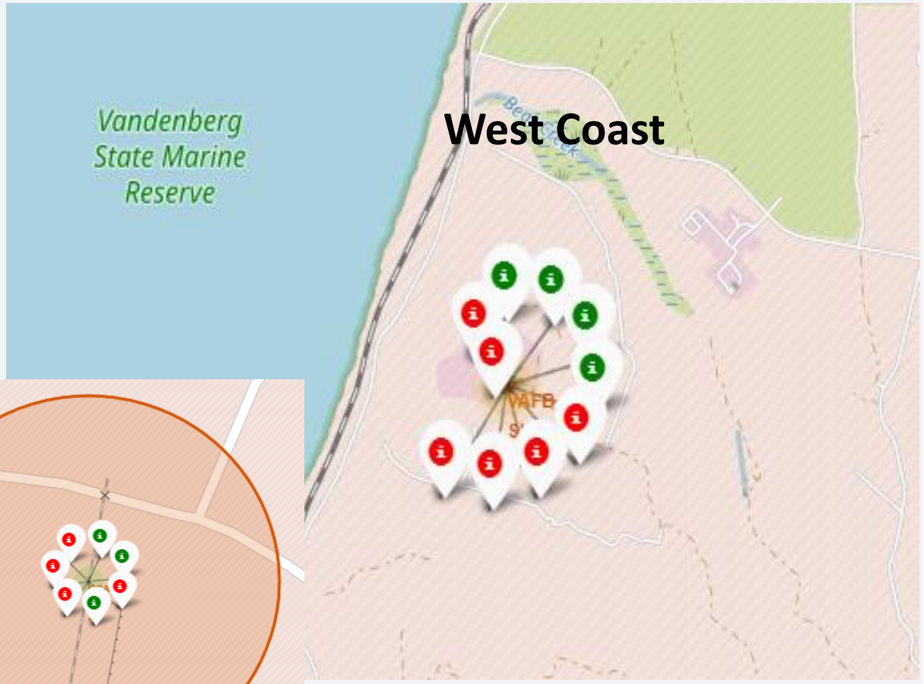
All Launch Sites on a Map

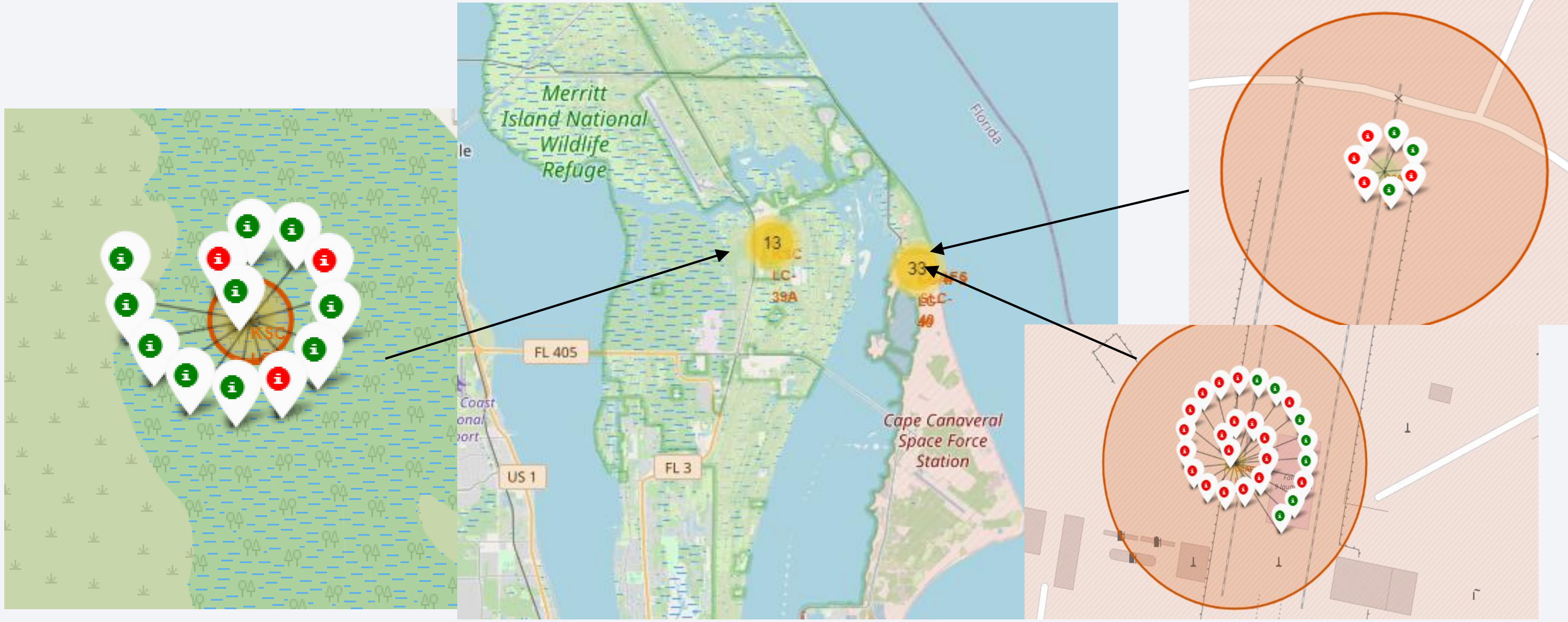
- Launch sites are often located near the equator to take advantage of the Earth's rotation. The speed of Earth's rotation is fastest at the equator, which provides an extra boost to rockets, allowing them to use less fuel to reach orbit.
- As for their placement on the east or west coast of the USA, eastern coastal sites are beneficial for launches that need to achieve an eastward trajectory, which is common for many orbits. Launching eastward over the ocean minimizes risk as it avoids overflying populated areas, ensuring that any debris from the launch will fall into the ocean rather than on land where it could cause damage or casualties. Additionally, the west coast location is favorable for launches into polar orbits, which often require a southward trajectory.



Success/ Failed launches for each site

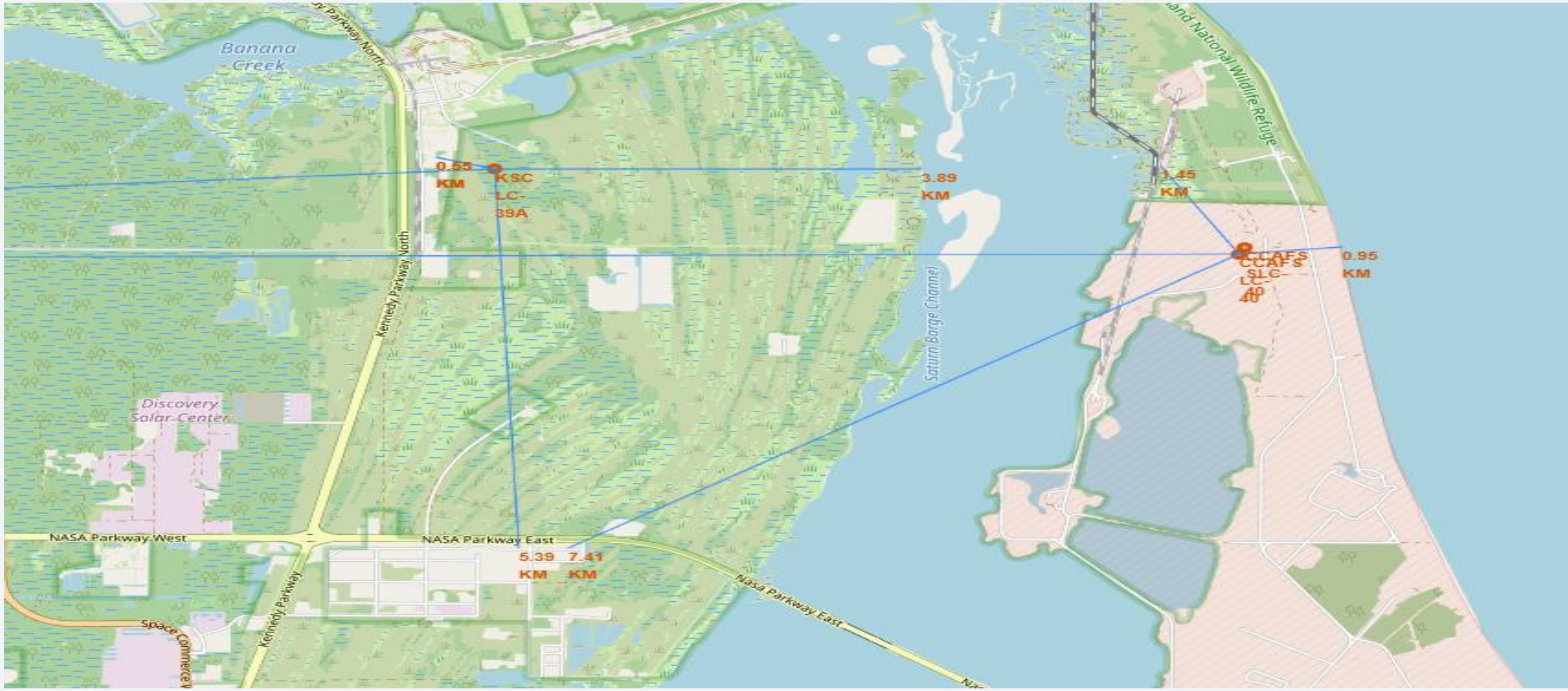
- Mount Island National Wildlife Reserve stands out for its high number of successful launches, with a success rate of approximately 77%, as 10 out of 13 launches have been successful.





Distances between a launch site to its proximities

- Launch Sites are near coast.
- Launch Sites are also close to Major Highways and Railway for logistic purposes.
- Launch sites are far from dense human habitats like cities.

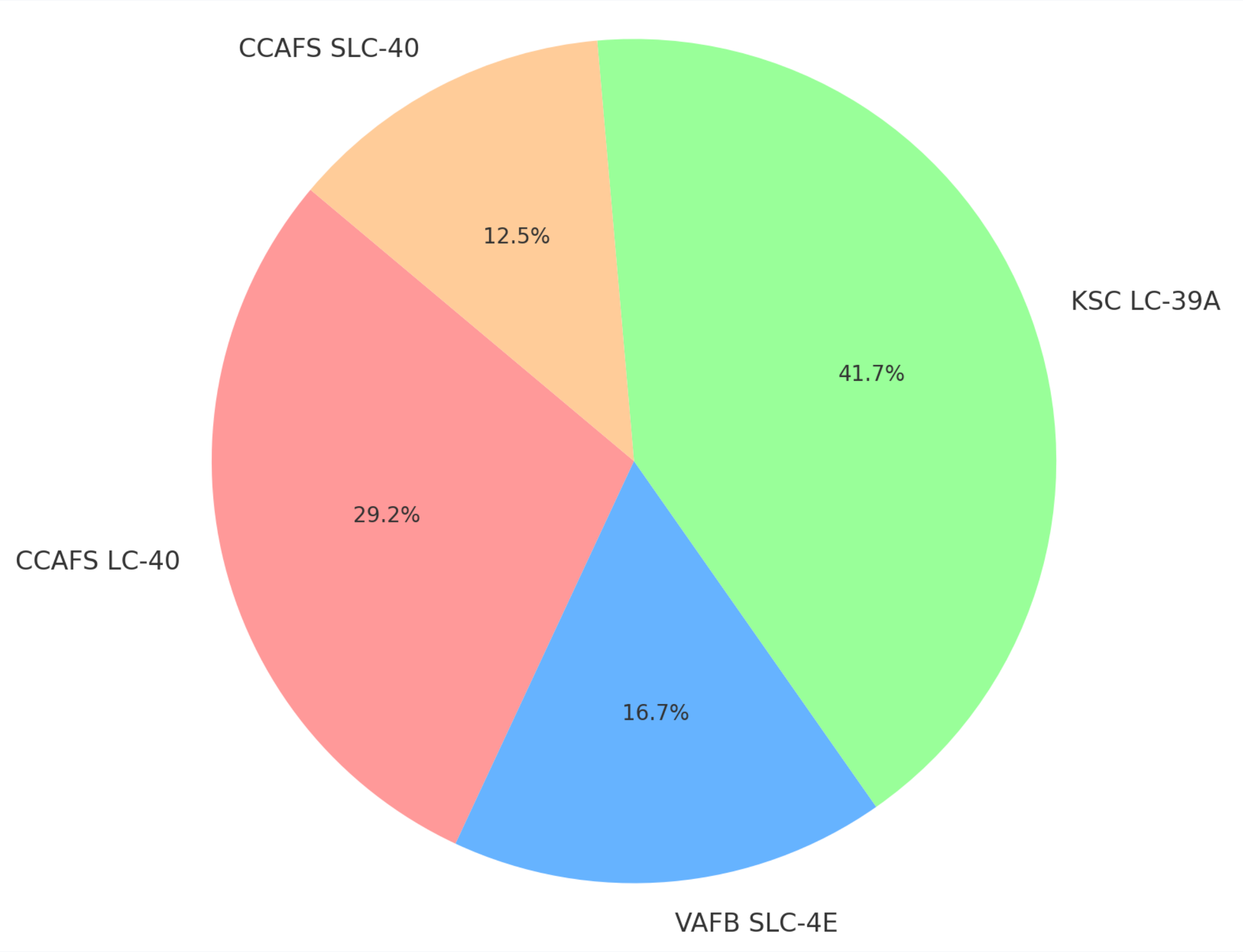


Section 4

Build a Dashboard with Plotly Dash

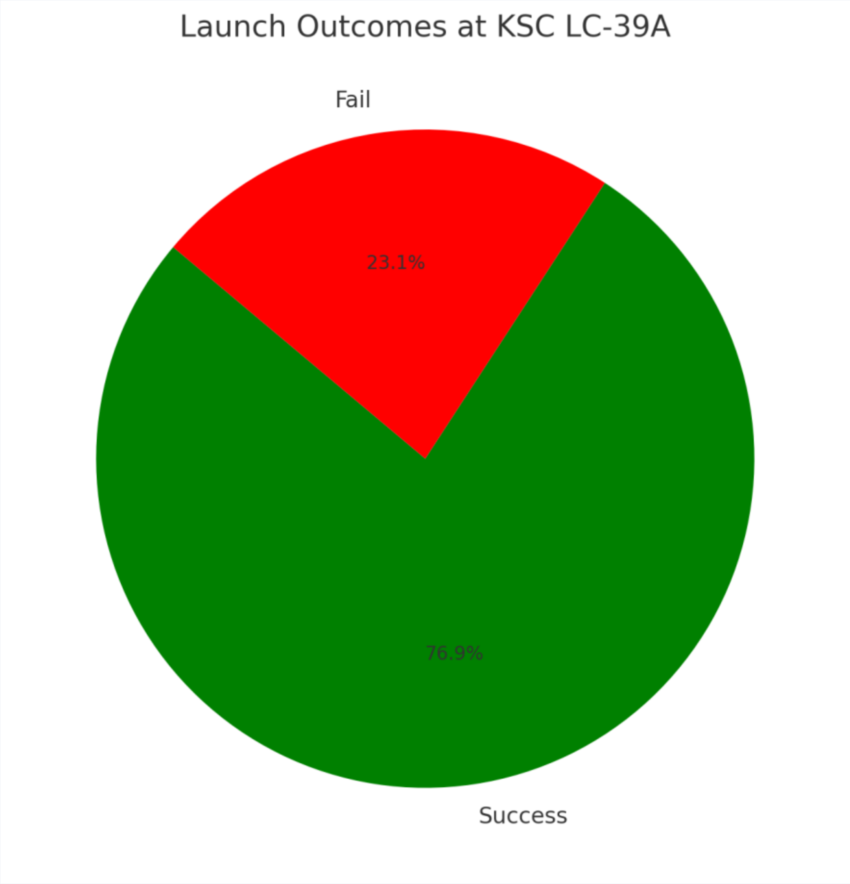
Success rate by Launch Site

- The pie chart representing the success count for each launch site. The chart clearly shows that KSC LC-39A had the highest (41.7%).



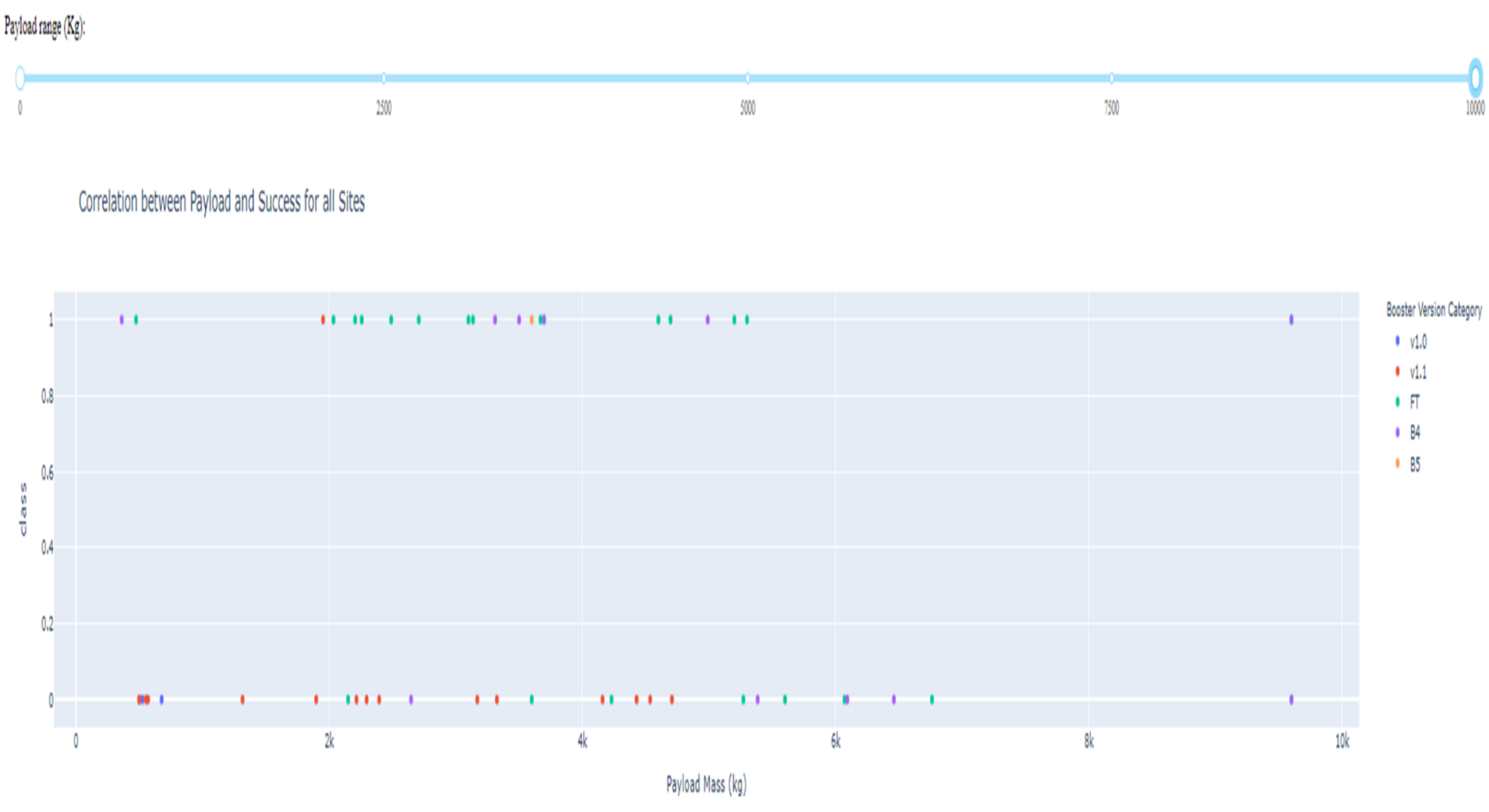
Launch site with the highest launch success ratio

- Here is the pie chart for the launch outcomes at KSC LC-39A. It shows that there have been a total of 13 launches, with 10 being successful (76.9%) and 3 failing (23.1%).



Payload vs Launch Outcome for all sites, with different payload selected

- The majority of data points are clustered at the lower end of the payload mass range, indicating that most launches carried payloads of lower mass.
- There is a mix of successes and failures at lower payload masses, but as the payload mass increases beyond approximately 4000 kg, the launches are predominantly successful.
- The highest payload masses, near the 10,000 kg mark, have been successfully launched, suggesting the capability of the launch vehicles to handle heavy payloads effectively.
- The range slider at the top suggests that the viewer can adjust the payload mass range to view a more detailed correlation within specific intervals.
- Different booster versions are represented by various colors, but there does not seem to be a distinct pattern linking booster versions to success rates based on this plot alone.

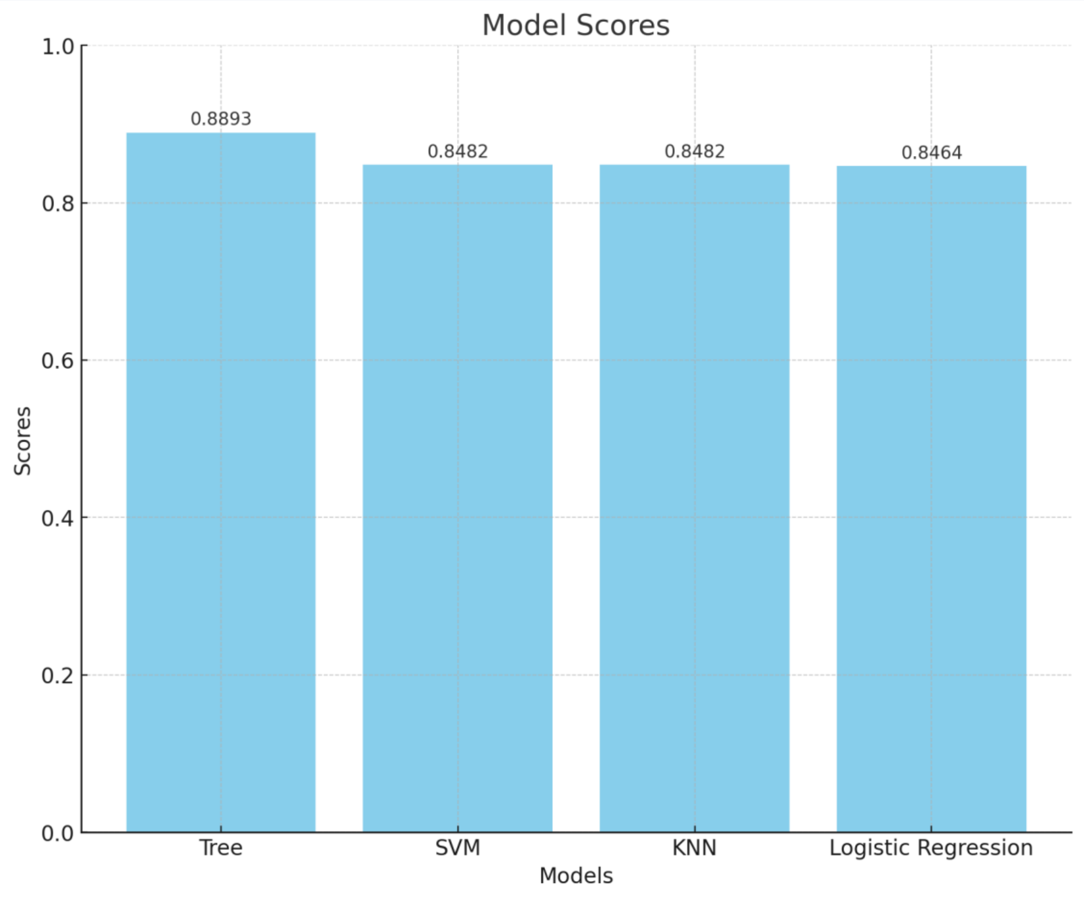


Section 5

Predictive Analysis (Classification)

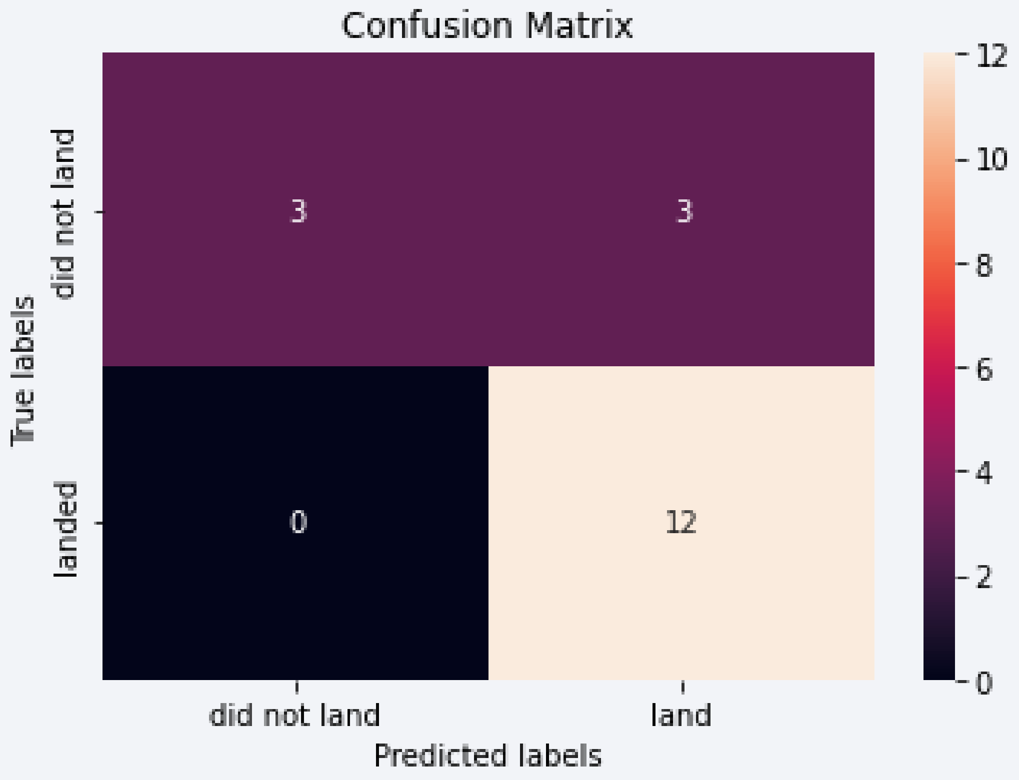
Classification Accuracy

- The bar chart displaying the scores of different models in descending order. The 'Tree' model has the highest score, followed by 'SVM' and 'KNN', which have the same score, and finally 'Logistic Regression' with the lowest score among the four. The y-axis represents the model names, and the x-axis represents the score values.



Confusion Matrix of Decision Tree

- Based on the confusion matrix, the model predicted 12 successful landings correctly (true positives) and no unsuccessful landings where there were successful ones (false negatives), which is an improvement. 3 unsuccessful landings were correctly predicted (true negatives), and 3 successful landings were incorrectly predicted as unsuccessful (false positives).
- To calculate the accuracy:
- Accuracy=15/18
- Accuracy=0.8333 or 83.33%
- This model's accuracy is approximately 83.33%, indicating that it correctly predicted the landing outcome for 83.33% of the cases in the dataset provided.



Conclusions

- Launch sites are strategically positioned near coasts to maximize rocket efficiency and safety during launches. The proximity to the equator also leverages Earth's rotation for launch efficiency.
- Access to major transportation networks such as highways and railways is crucial for logistical support, allowing for the efficient movement of materials and personnel.
- The location of launch sites away from populated areas minimizes risks to humans in the event of launch failures.
- The flight count at a launch site correlates positively with its success rate, suggesting improved proficiency with increased activity.
- There has been a steady improvement in launch success rates from 2013 to 2020.
- Certain orbits like ES-L1, GEO, HEO, SSO, and VLEO have higher success rates, which might be due to the nature of missions suited for these orbits or the technological advances addressing the challenges of these specific paths.
- KSC LC-39A stands out with the highest number of successful launches, highlighting its status as a premier launch site.
- The Decision tree classifier emerges as the most effective machine learning model for predicting launch success, indicating its suitability for this kind of classification task.

Thank you!