IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Nasir Arslan
5-25-2024

Outline

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

Executive Summary

- Summary of methodologies
- Summary of all results

Introduction

- Project background and context

This project analyzes SpaceX Falcon 9 launch data to understand factors influencing launch success and to develop predictive models for future launches

- Problems you want to find answers

The project aims to determine which launch sites and booster versions are most reliable and to predict the success of future Falcon 9 missions.

Section 1

Methodology

Methodology

Executive Summary

- Data collection methodology:
    - The data for the SpaceX Falcon 9 launch analysis was collected using the SpaceX API for detailed launch information and web scraping techniques to gather additional payload data, ensuring comprehensive and structured datasets for further analysis
    - Perform data wrangling
    - The data wrangling consider the missing values with the mean, a data is conversion is performed for the Feature Engineering for the features which need to be selected also data for only rocket type Falcon 9 is used.
- Perform exploratory data analysis (EDA) using visualization and SQL
- Perform interactive visual analytics using Folium and Plotly Dash
- Perform predictive analysis using classification models
    - Classification models were built and tuned using GridSearchCV to find the best hyperparameters, and their performance was evaluated based on accuracy and confusion matrices, with the Decision Tree model achieving the highest accuracy

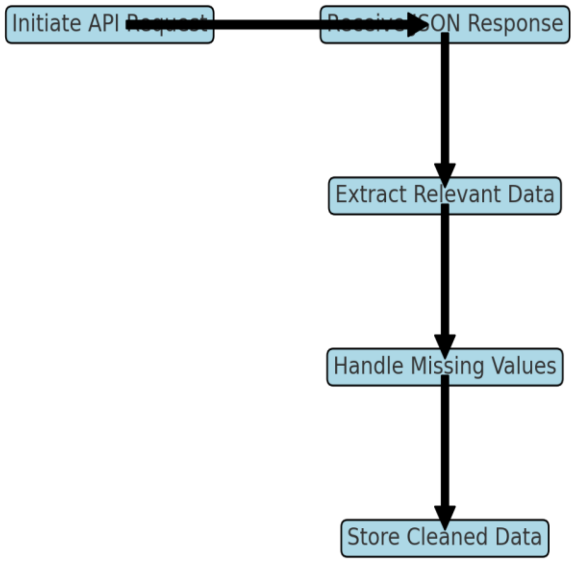
Data Collection

Data Collection through API and WebScraping

- API Request: We used the SpaceX API to fetch detailed data on Falcon 9 launches, including launch dates, payload masses, and landing outcomes.
- GET Request: The data was obtained by sending a GET request to the SpaceX API endpoint.
- Data Parsing: The response data was parsed and converted into a structured format for further processing.
- Handling Missing Values: We handled missing values by replacing them with appropriate statistical measures (e.g., mean).You need to present your data collection process use key phrases and flowcharts

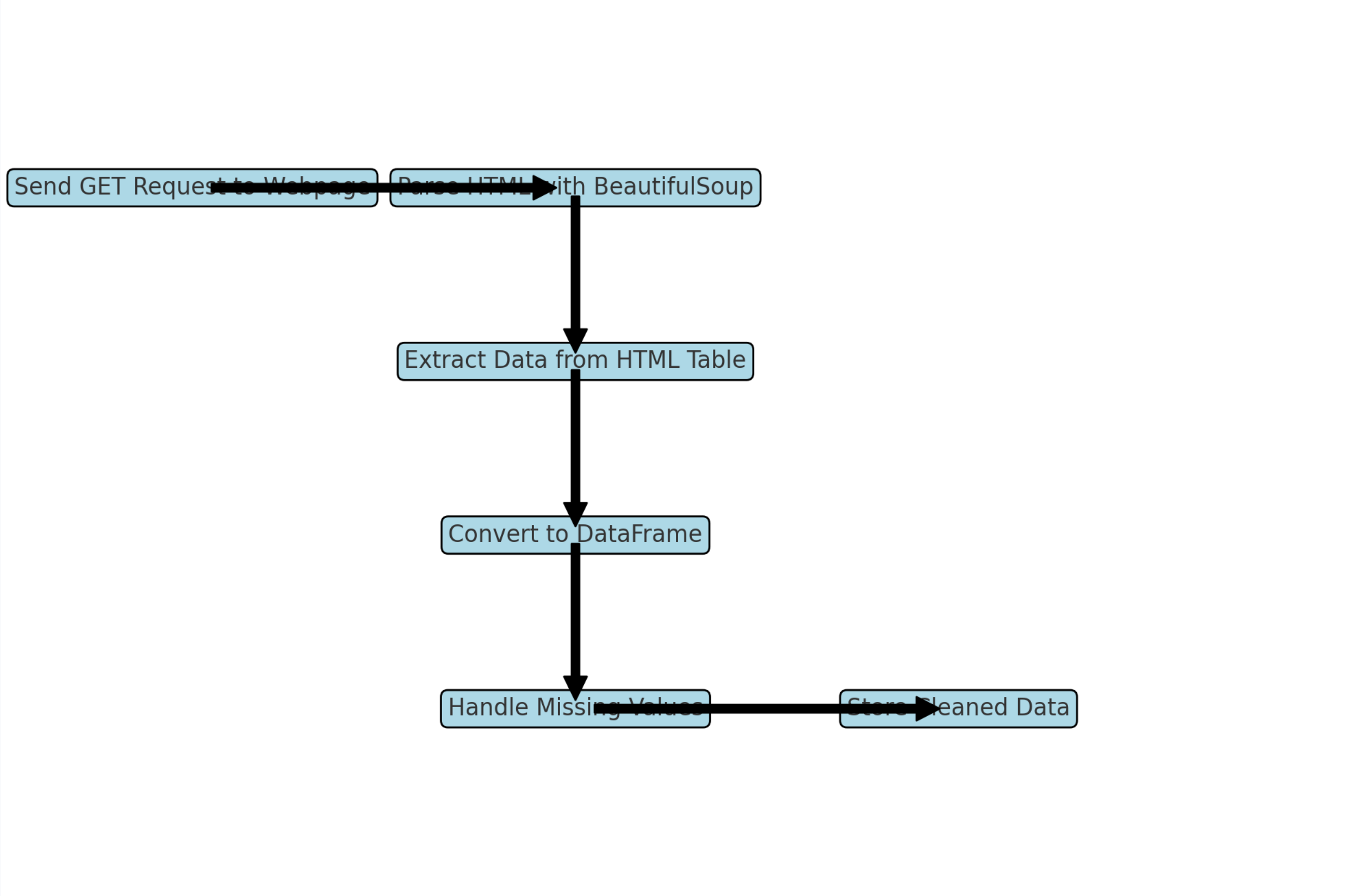
Data Collection – SpaceX API

- response = requests.get(static_json_url)
- data = pd.json_normalize(response.json())
- [IBM_capestone/M1-jupyter-labs-spacex-data-collection-api.ipynb at develop · nasirarslan35/IBM_capestone (github.com)]



Data Collection - Scraping

- Use Requests and BeautifulSoup Library
- response = requests.get(static_url)
- soup = BeautifulSoup(response.text, 'html.parser')
- [IBM_capestone/M1-jupyter-labs-webscraping.ipynb at develop · nasirarslan35/IBM_capestone (github.com)]

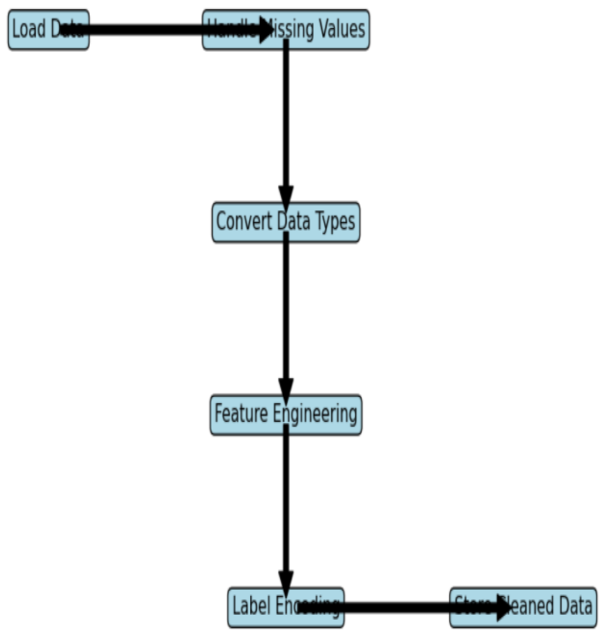


Data Wrangling

Overview:

The data wrangling process involves cleaning and transforming the collected data to prepare it for analysis and model building. This process includes handling missing values, converting data types, creating new features, and ensuring data consistency.

- Key Phrases:
- Data Cleaning: The process of identifying and handling missing or inconsistent data.
- Handling Missing Values: Replacing missing values with appropriate substitutes, such as mean values.
- Feature Engineering: Creating new features from the existing data to enhance predictive modeling.
- Data Transformation: Converting data into suitable formats for analysis.
- Data Storage: Saving the cleaned and processed data into a CSV file for further use.
- [IBM_capestone/M1-labs-jupyter-spacex-Data wrangling.ipynb at develop · nasirarslan35/IBM_capestone (github.com)]



EDA with Data Visualization

1. **Scatter Plot of Flight Number vs. Launch Site**

- To visualize the relationship between flight numbers and launch sites. To identify any patterns or trends related to specific launch sites.

2. Scatter Plot of Payload Mass vs. Launch Site

- To analyze the distribution of payload mass across launch sites. To identify any significant differences in payload mass for specific launch sites.

3. Bar Chart of Success Rate for Each Orbit Type

- To compare the success rates of launches for different orbit types. To identify which orbit types have higher success rates.

4. Scatter Plot of Flight Number vs. Orbit Type

- To explore the relationship between flight numbers and orbit types. To identify any trends or patterns in the data related to specific orbit types.

5. Scatter Plot of Payload Mass vs. Orbit Type

- To analyze the distribution of payload mass across orbit types. To identify any significant differences in payload mass for specific orbit types.

6. Line Chart of Yearly Average Success Rate

- To track the yearly trends in the success rate of Falcon 9 launches. To identify any improvements or declines in success rates over the years.
- IBM_capestone/M2-jupyter-labs-eda-dataviz.ipynb at develop · nasirarslan35/IBM_capestone (github.com)

EDA with SQL

1. Retrieve All Launch Sites

- SELECT DISTINCT Launch_Site FROM SPACEX;

2. Retrieve Launch Site Records for Specific Site

- SELECT * FROM SPACEX WHERE Launch_Site = 'CCAFS SLC 40';

3. Calculate Total Payload Carried by SpaceX

- SELECT SUM(PAYLOAD_MASS__KG_) FROM SPACEX;

4. Calculate Average Payload Mass for Specific Booster Version

- SELECT AVG(PAYLOAD_MASS__KG_) FROM SPACEX WHERE Booster_Version = 'F9 v1.1';

5. Find Dates of First Successful Landing Outcomes on Drone Ship

- SELECT Date FROM SPACEX WHERE Landing_Outcome = 'Success (drone ship)' ORDER BY Date ASC LIMIT 1;

EDA with SQL

6. List Boosters with Successful Drone Ship Landings and Specific Payload Range

- SELECT Booster_Version FROM SPACEX WHERE Landing_Outcome = 'Success (drone ship)' AND PAYLOAD_MASS__KG_ BETWEEN 4000 AND 6000;

7. Calculate Total Number of Successful and Failed Mission Outcomes

- SELECT Landing_Outcome, COUNT(*) FROM SPACEX GROUP BY Landing_Outcome;

8. List Boosters with Maximum Payload Mass

- SELECT Booster_Version FROM SPACEX ORDER BY PAYLOAD_MASS__KG_ DESC LIMIT 1;

9. Retrieve Launch Records for 2015

- SELECT Date, Booster_Version, Launch_Site FROM SPACEX WHERE DATE LIKE '2015%' AND Landing_Outcome LIKE 'Success (ground pad)';

10. Rank Landing Outcomes Between Specific Dates

- SELECT Landing_Outcome, COUNT(*) AS Count FROM SPACEX WHERE Date BETWEEN '2010-06-04' AND '2017-03-20' GROUP BY Landing_Outcome ORDER BY Count DESC;

- [IBM_capestone/M2-jupyter-labs-eda-sql-edx_sqllite.ipynb at develop · nasirarslan35/IBM_capestone (github.com)]

Build an Interactive Map with Folium

In this lab, several map objects such as markers, circles, and lines were added to a Folium map to visualize the launch sites and their proximities. Here are the objects created and the reasons for adding them

1. Markers for Launch Sites

- Purpose: To pinpoint the exact locations of SpaceX launch sites on the map. Each launch site was marked with a folium.Marker object to provide a clear visual reference of its locationAdd the GitHub URL of your completed interactive map with Folium map, as an external reference and peer-review purpose

2. Marker Clusters for Launch Outcomes

- Purpose: To visualize the success and failure outcomes of launches at each site. Details: Using MarkerCluster, markers were added to show the locations of successful and failed launches.

3. Circles to Represent Launch Site Proximity

- Purpose: To highlight the area around each launch site and indicate its proximity to important landmarks. Circles were added with specific radii to show the proximity to nearby cities, coastlines, highways, and railways.

Build an Interactive Map with Folium

4. Lines to Show Distances to Proximities

- Purpose: To visualize the distances from launch sites to nearby cities, coastlines, highways, and railways. Lines (folium.PolyLine) were drawn to represent the connections between launch sites and their proximities.2. Marker Clusters for Launch Outcomes

5. Mouse Position for Coordinate Display

- Purpose: To provide interactive feedback on the map, showing the latitude and longitude of the cursor position. The MousePosition plugin was added to display coordinates as the mouse moves over the map.

- IBM_capestone/M3-lab_jupyter_launch_site_location.jupyterlite.ipynb at develop · nasirarslan35/IBM_capestone (github.com)

Build a Dashboard with Plotly Dash

Several plots and interactive elements such as Drop Down, Pie Chart, Payload Mass Selection Range, and Scatter Plot were added to a Plotly Dash dashboard to visualize the SpaceX launch data

1. Dropdown Menu for Launch Sites

- To allow users to filter the data by selecting a specific launch site or viewing data from all launch sites. A dcc.Dropdown component was used to create the dropdown menu with options for each launch site.

2. Success Pie Chart

- To visualize the success rate of launches for each selected launch site.

3. Payload Range Slider

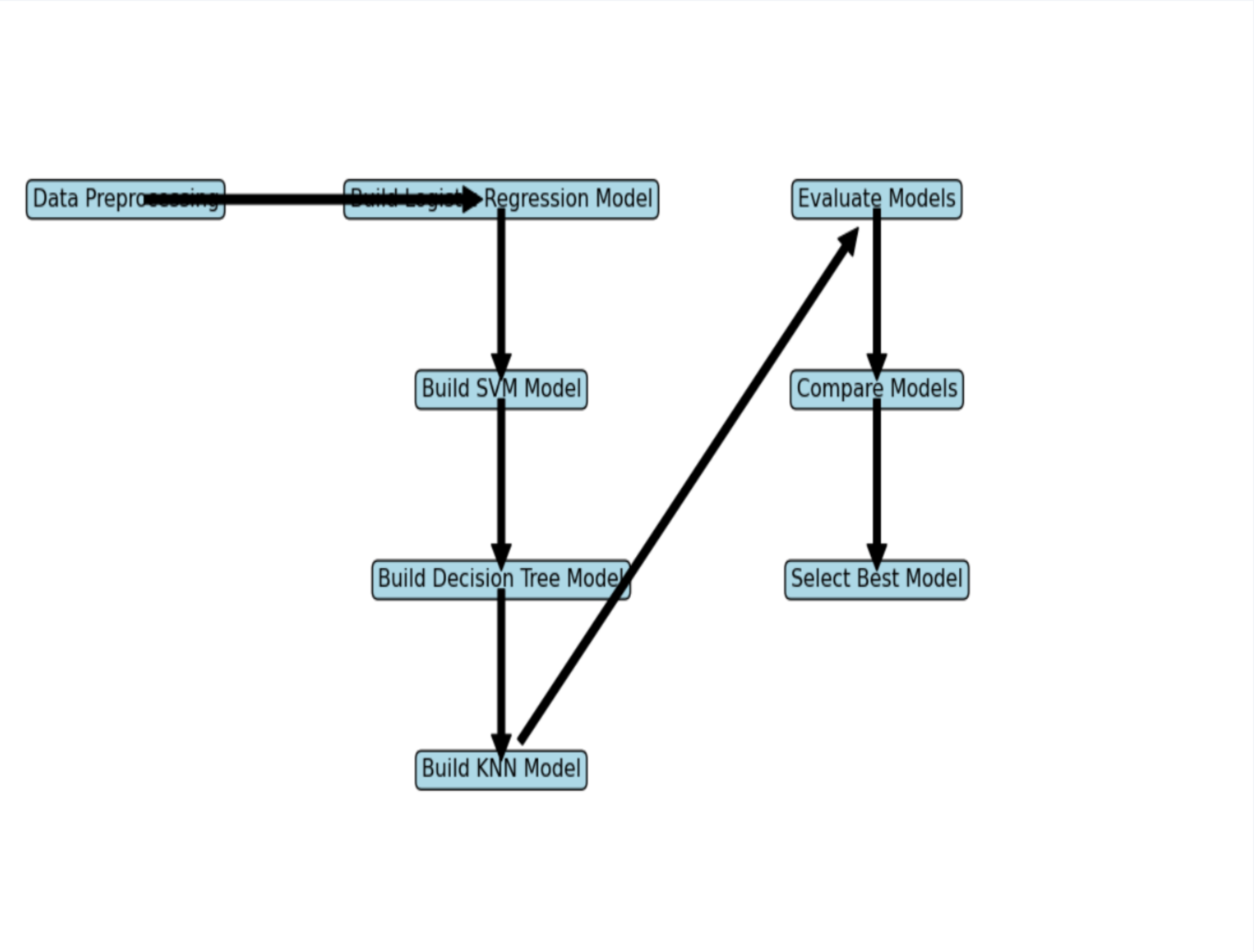
- To allow users to filter the data based on a range of payload masses. A dcc.RangeSlider component was used to create a slider for selecting the payload mass range.

4. Success-Payload Scatter Plot

- To visualize the correlation between payload mass and launch success for the selected launch site and payload range.
- IBM_capestone/M3-.ipynb at develop · nasirarslan35/IBM_capestone (github.com)

Predictive Analysis (Classification)

- The process of building, evaluating, improving, and finding the best-performing classification model for predicting the Falcon 9 first stage landing involves several steps. Here is a summary of the process steps and a flowchart. Please take a look at the flow chart. Detail steps are included in the next slide.



Predictive Analysis (Classification)

1. Data Preprocessing

- Create Class Labels: Created a column for the class label indicating successful (1) and unsuccessful (0) landings.
- Standardize Data: Standardized the data to have a mean of 0 and a standard deviation of 1.
- Train-Test Split: Split the data into training and test sets.

2. Model Building and Hyperparameter Tuning

- Logistic Regression: Used GridSearchCV to find the best hyperparameters for logistic regression.
- Support Vector Machine (SVM): Tuned SVM hyperparameters using GridSearchCV.
- Decision Tree: Applied GridSearchCV to find the best parameters for a decision tree classifier.
- K-Nearest Neighbors (KNN): Utilized GridSearchCV to optimize KNN hyperparameters.

3. Model Evaluation

- Accuracy Calculation: Calculated the accuracy of each model on the test data.
- Confusion Matrix: Plotted confusion matrices for each model to evaluate performance.

Predictive Analysis (Classification)

4. Model Comparison and Selection

- Compare Models: Compared the accuracy of all models to identify the best-performing one.
- Select Best Model: Selected the model with the highest accuracy for final use.

- IBM_capestone/M4-SpaceX_Machine_Learning_Prediction_Part_5.jupyterlite.ipynb at develop · nasirarslan35/IBM_capestone (github.com)

Results

- Exploratory data analysis results

Exploratory data analysis revealed key insights such as high success rates at launch sites KSC LC-39A and CCAFS LC-40, and demonstrated that payload mass does not strongly correlate with launch success. Additionally, certain orbits like ES-L1 and GEO showed perfect success rates, while GTO and ISS had more challenges

- Interactive analytics demo in screenshots

All the findings will be provided in detail into the next coming slides

- Predictive analysis results

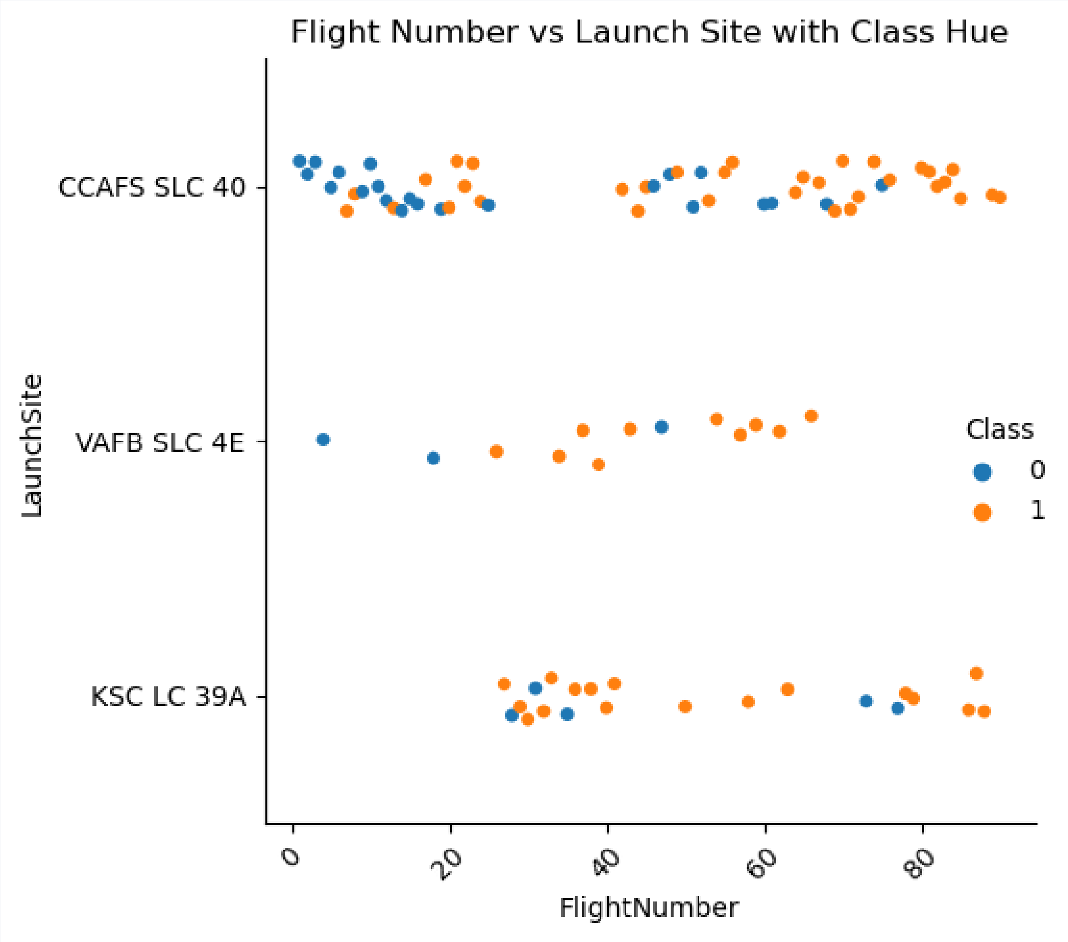
The Decision Tree model emerged as the best-performing classification model with an accuracy of 88.75%. It effectively predicts successful landings, with high precision and perfect recall, although it has some false positives that indicate room for improvement.

Section 2

Insights drawn from EDA

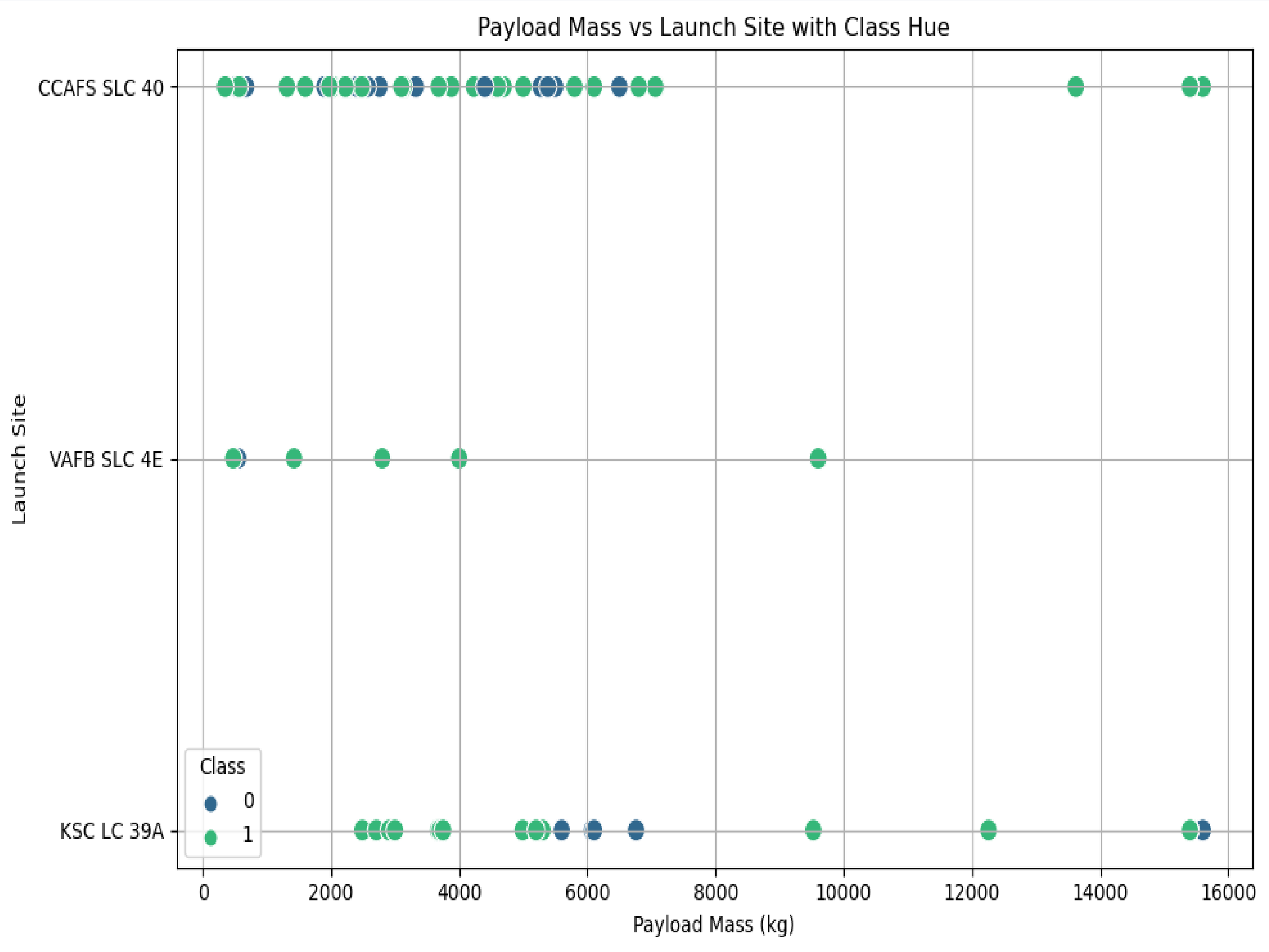
Flight Number vs. Launch Site

- CCAFS SLC 40 shows consistent usage and a balanced success rate.
- VAFB SLC 4E has fewer launches but a relatively higher success rate.
- KSC LC 39A appears to be more successful in terms of launch outcomes.



Payload vs. Launch Site

- CCAFS SLC 40: Shows a balanced success rate across a wide range of payload masses. This site handles a variety of payload sizes with a consistent performance.
- VAFB SLC 4E: Has fewer launches but a higher proportion of successful launches. The site handles moderate to high payload masses.
- KSC LC 39A: Indicates a higher success rate for lower payload masses, with mixed outcomes for higher payloads. This suggests that while the site can handle larger payloads, the success rate varies.



Success Rate vs. Orbit Type

High Success Orbits:

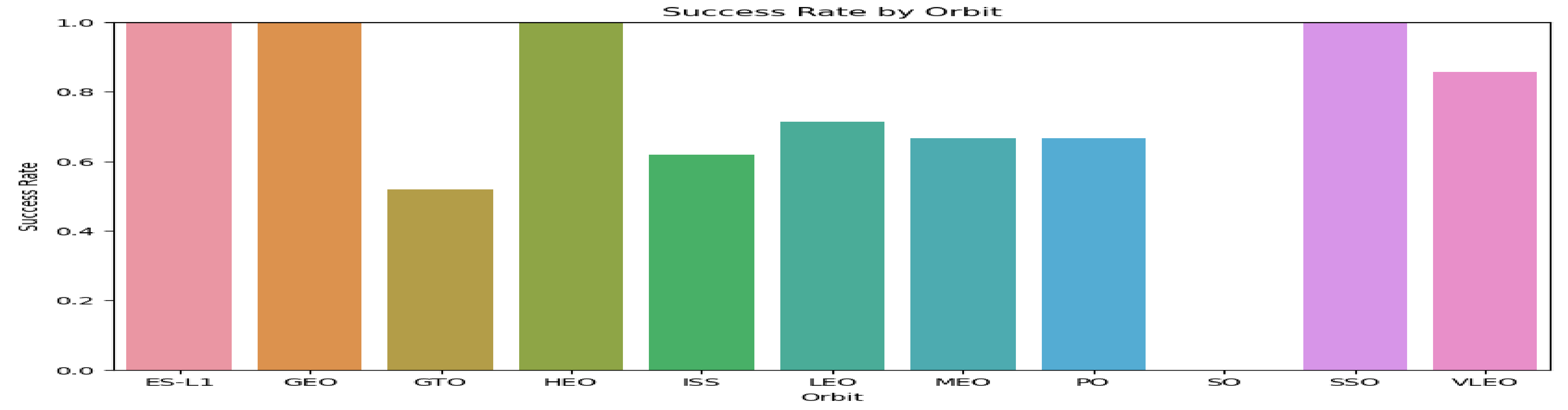
- ES-L1, GEO, HEO, SSO, VLEO: The perfect success rates indicate these orbits are consistently reliable for SpaceX launches. The technology and processes for these missions are likely well-optimized.

Challenging Orbits:

- GTO: The lower success rate highlights challenges in reaching this orbit. Potential reasons could include the complex maneuvers required for geostationary transfer.
- ISS: Docking with the ISS involves precision and timing, leading to a lower success rate compared to other orbits.

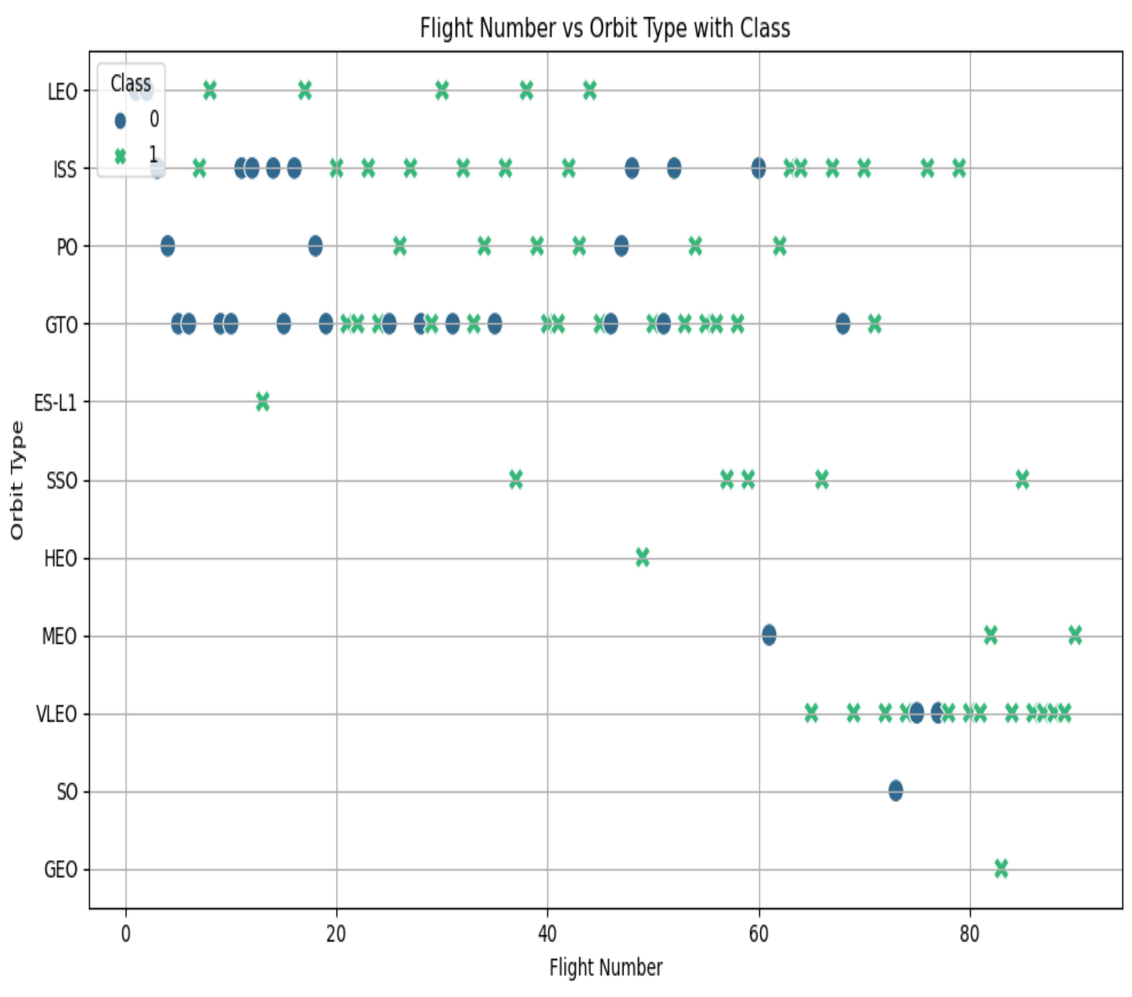
Frequent and Variable Success Orbits:

- LEO and MEO: These orbits have high success rates, though not perfect. The frequent use of LEO for various missions, including satellite deployments, may contribute to its high but varied success.
- PO: The variability in polar orbit success might be due to the specific trajectory and environmental conditions required for these missions.



Flight Number vs. Orbit Type

- The plot indicates a trend where early missions have a higher failure rate, but as the flight number increases, the success rate improves. This reflects operational learning and process optimization over time.



Payload vs. Orbit Type

High Success Orbits (LEO, SSO, ES-L1):

- These orbits show a high success rate across a range of payload masses, indicating robust mission planning and execution.

Challenging Orbits (GTO, ISS):

- GTO missions have a balanced mix of successes and failures, suggesting inherent challenges with these missions.

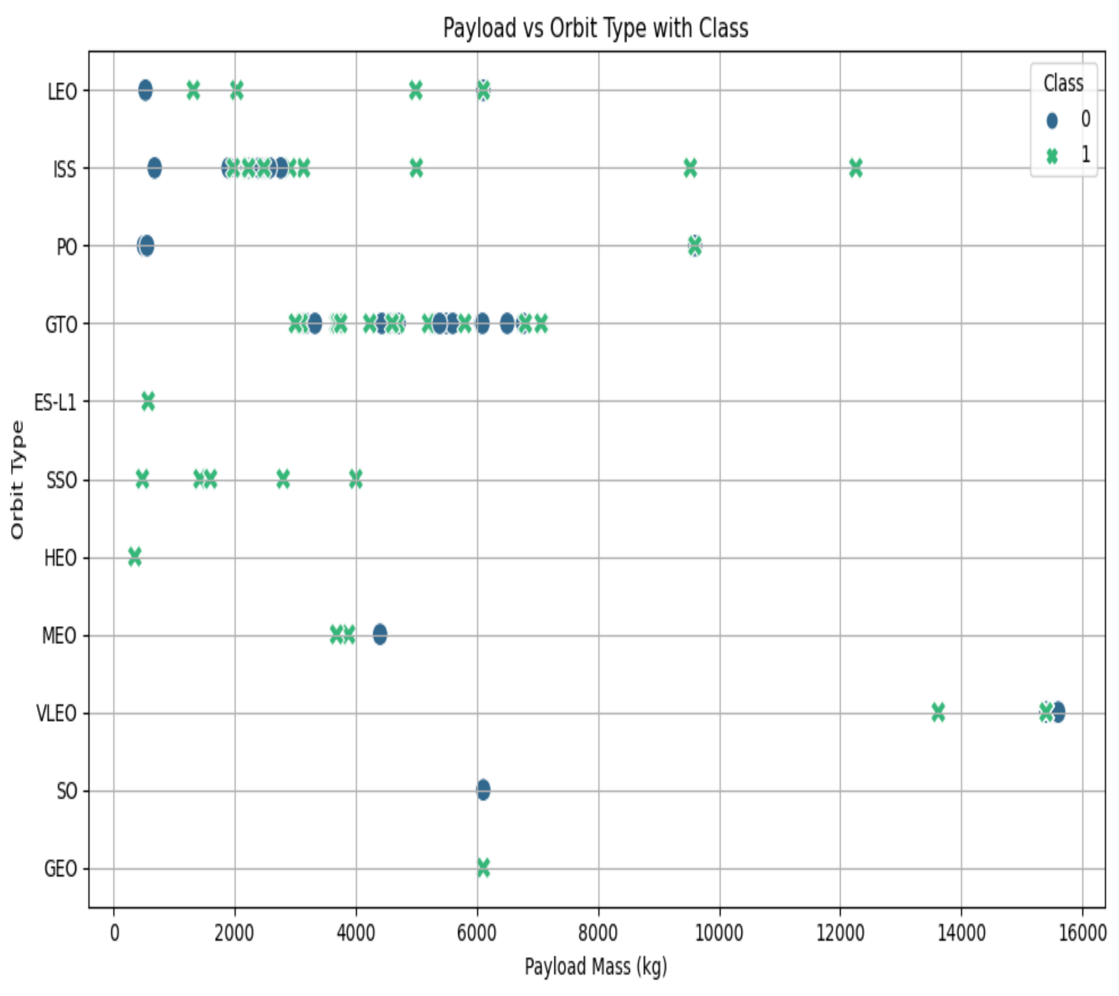
- ISS missions have a mix of outcomes, likely due to the precision required for docking with the ISS.

Moderate Payload Orbits (PO, VLEO, MEO):

- These orbits generally handle moderate payloads and show a good success rate. They demonstrate consistent performance across the payload range.

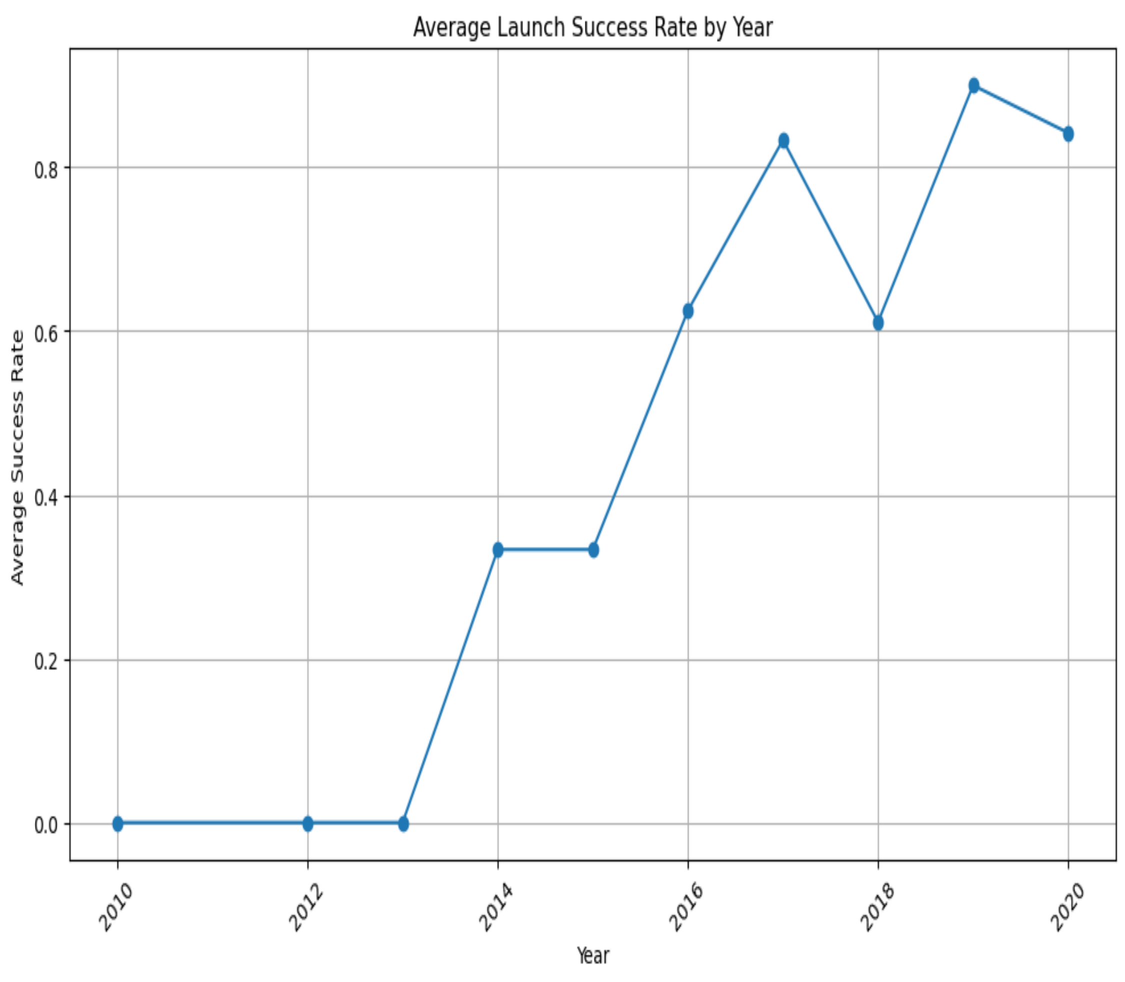
Higher Payload Orbits (GEO):

- GEO missions handle some of the highest payloads with mixed success rates, indicating that heavier payloads might introduce additional challenges.



Launch Success Yearly Trend

- The overall trend indicates that SpaceX has developed a resilient and reliable launch system, capable of maintaining high success rates even with occasional setbacks.



All Launch Site Names

- SQL command is used to extract the launch site name from the SPACEXTBL and IN “Lauch_Site” of the table.
  “SELECT DISTINCT "Launch_Site" FROM SPACEXTBL”

Below is the output of the query.

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

Launch Site Names Begin with 'KSC'

- Find 5 records where launch sites' names start with `KSC`

  SELECT * FROM SPACEXTBL WHERE "Launch_Site" LIKE 'KSC%' LIMIT 5

- The query retrieves up to 5 rows from the SPACEXTBL table where the Launch_Site column's value begins with 'KSC

- ('2017-02-19', '14:39:00', 'F9 FT B1031.1', 'KSC LC-39A', 'SpaceX CRS-10', 2490, 'LEO (ISS)', 'NASA (CRS)', 'Success', 'Success (ground pad)')
- ('2017-03-16', '6:00:00', 'F9 FT B1030', 'KSC LC-39A', 'EchoStar 23', 5600, 'GTO', 'EchoStar', 'Success', 'No attempt')
- ('2017-03-30', '22:27:00', 'F9 FT B1021.2', 'KSC LC-39A', 'SES-10', 5300, 'GTO', 'SES', 'Success', 'Success (drone ship)')
- ('2017-05-01', '11:15:00', 'F9 FT B1032.1', 'KSC LC-39A', 'NROL-76', 5300, 'LEO', 'NRO', 'Success', 'Success (ground pad)')
- ('2017-05-15', '23:21:00', 'F9 FT B1034', 'KSC LC-39A', 'Inmarsat-5 F4', 6070, 'GTO', 'Inmarsat', 'Success', 'No attempt')

Total Payload Mass

- Calculate the total payload carried by boosters from NASA

SELECT SUM("PAYLOAD_MASS__KG_") AS Total_Payload_Mass FROM SPACEXTBL WHERE "Customer" = 'NASA (CRS)'

- The query calculates the total payload mass for all launches where the customer is 'NASA (CRS)’. The SUM function then adds up the payload masses of these filtered rows.

Result: Total payload mass carried by boosters launched by NASA (CRS): 45596 kg

Average Payload Mass by F9 v1.1

- Calculate the average payload mass carried by booster version F9 v1.1

```
SELECT AVG("PAYLOAD_MASS__KG_") AS Average_Payload_Mass
FROM SPACEXTBL WHERE "Booster_Version" = 'F9 v1.1'
```

- This query calculates the average payload mass for all launches using the booster version 'F9 v1.1' from the SPACEXTBL table. The result is labeled as Average_Payload_Mass
- **Average payload mass carried by booster version F9 v1.1: 2928.4 kg**

First Successful Ground Landing Date

- Find the dates of the first successful landing outcome on drone ship.

```
SELECT MIN("Date") AS Earliest_Successful_Landing FROM SPACEXTBL
WHERE "Landing_Outcome" = 'Success (drone ship)’
```

- This query retrieves the earliest date of a successful landing using minimum function, on a drone ship from the SPACEXTBL table. The result is labeled as Earliest_Successful_Landing
- **The earliest successful drone ship landing was on: 2016-04-08**

Successful Ground Ship Landing with Payload between 4000 and 6000

- List the names of boosters which have successfully landed on Ground Ship and had payload mass greater than 4000 but less than 6000

SELECT "Booster_Version“ FROM SPACEXTBL WHERE "Landing_Outcome" = 'Success (ground pad)’ AND "PAYLOAD_MASS__KG_" > 4000 AND "PAYLOAD_MASS__KG_" < 6000

- This query retrieves the booster versions from the SPACEXTBL table where the landing outcome was a success on a ground pad and the payload mass was between 4000 and 6000 kg.

F9 FT B1032.1

F9 B4 B1040.1

F9 B4 B1043.1

Total Number of Successful and Failure Mission Outcomes

- Calculate the total number of successful and failure mission outcomes

```
SELECT "Mission_Outcome", COUNT(*) AS Outcome_Count FROM SPACEXTBL
WHERE "Mission_Outcome" LIKE '%Success%' OR "Mission_Outcome" LIKE '%Failure%'
GROUP BY "Mission_Outcome"
```

- This query counts the number of occurrences of each mission outcome that contains the words 'Success' or 'Failure' in the SPACEXTBL table. The results are grouped by Mission_Outcome and labeled as Outcome_Count
- **Outcome: Failure (in flight), Count: 1**
- **Outcome: Success, Count: 98**
- **Outcome: Success , Count: 1**
- **Outcome: Success (payload status unclear), Count: 1**

Boosters Carried Maximum Payload

- List the names of the booster which have carried the maximum payload mass

  SELECT "Booster_Version“ FROM SPACEXTBL WHERE "PAYLOAD_MASS__KG_" = (SELECT MAX("PAYLOAD_MASS__KG_") FROM SPACEXTBL)

- This query retrieves the booster version associated with the maximum payload mass from the SPACEXTBL table
- **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**

2017 Launch Records

- List the records which will display the month names, succesful landing_outcomes in ground pad ,booster versions, launch_site for the months in year 2017

SELECT substr(Date, 6, 2) AS Month, "Booster_Version", "Launch_Site", "Landing_Outcome“ FROM SPACEXTBL WHERE substr(Date, 1, 4) = '2017' AND "Landing_Outcome" = 'Success (ground pad)'

- This query retrieves the month, booster version, launch site, and landing outcome for launches in the year 2017 where the landing outcome was a success on a ground pad. The month is extracted from the Date column
- **Month, Booster Version, Launch Site, Landing Outcome**
- **February, F9 FT B1031.1, KSC LC-39A, Success (ground pad)**
- **May, F9 FT B1032.1, KSC LC-39A, Success (ground pad)**
- **June, F9 FT B1035.1, KSC LC-39A, Success (ground pad)**
- **August, F9 B4 B1039.1, KSC LC-39A, Success (ground pad)**
- **September, F9 B4 B1040.1, KSC LC-39A, Success (ground pad)**
- **December, F9 FT B1035.2, CCAFS SLC-40, Success (ground pad)**

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

SELECT "Landing_Outcome", COUNT(*) AS Outcome_Count FROM SPACEXTBL WHERE Date BETWEEN '2010-06-04' AND '2017-03-20' GROUP BY "Landing_Outcome" ORDER BY Outcome_Count DESC

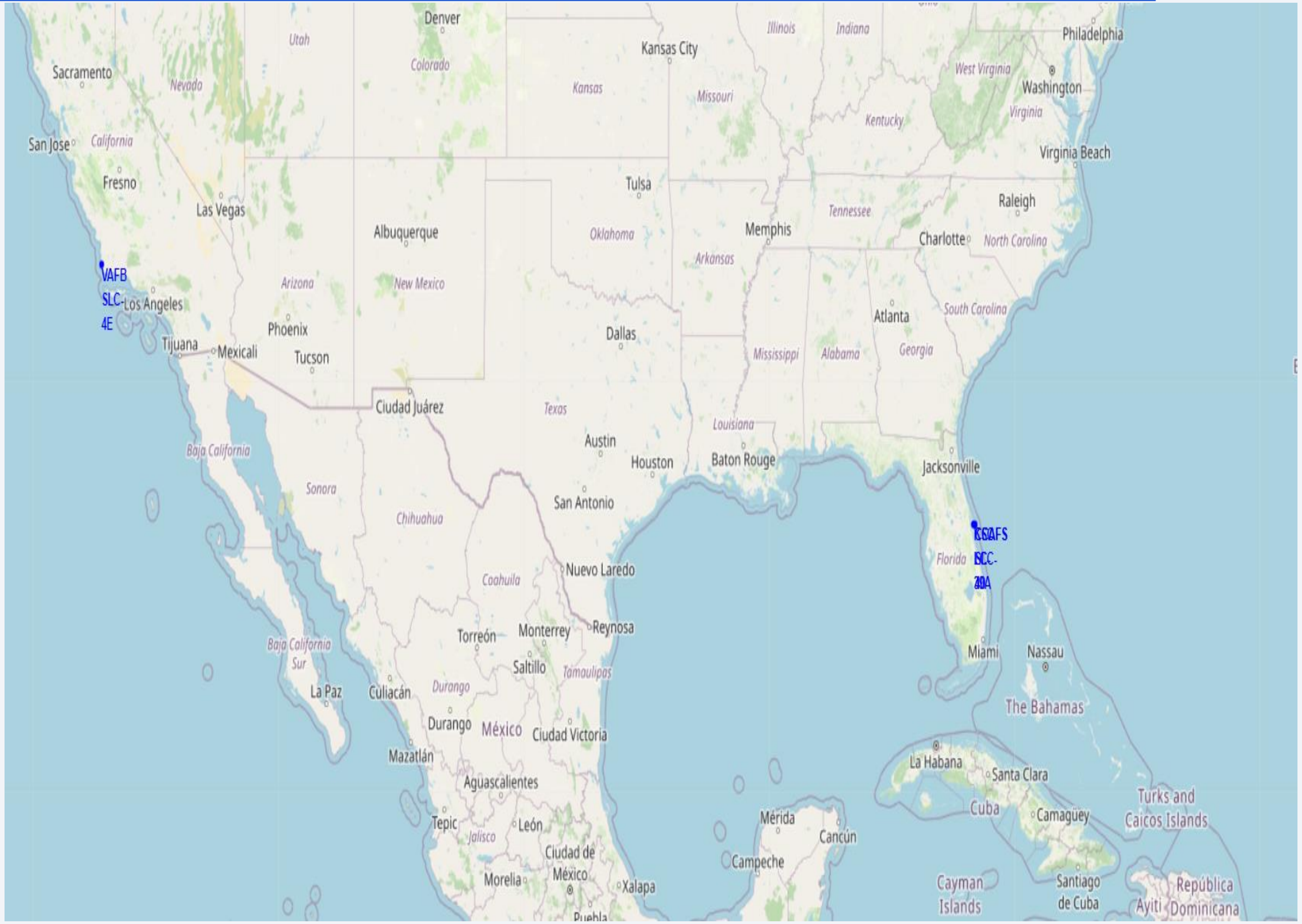
- This query counts the number of each landing outcome from the SPACEXTBL table for launches between June 4, 2010, and March 20, 2017. The results are grouped by Landing_Outcome, labeled as Outcome_Count, and sorted in descending order by Outcome_CountLanding Outcome, Count
- **No attempt, 10**
- **Success (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

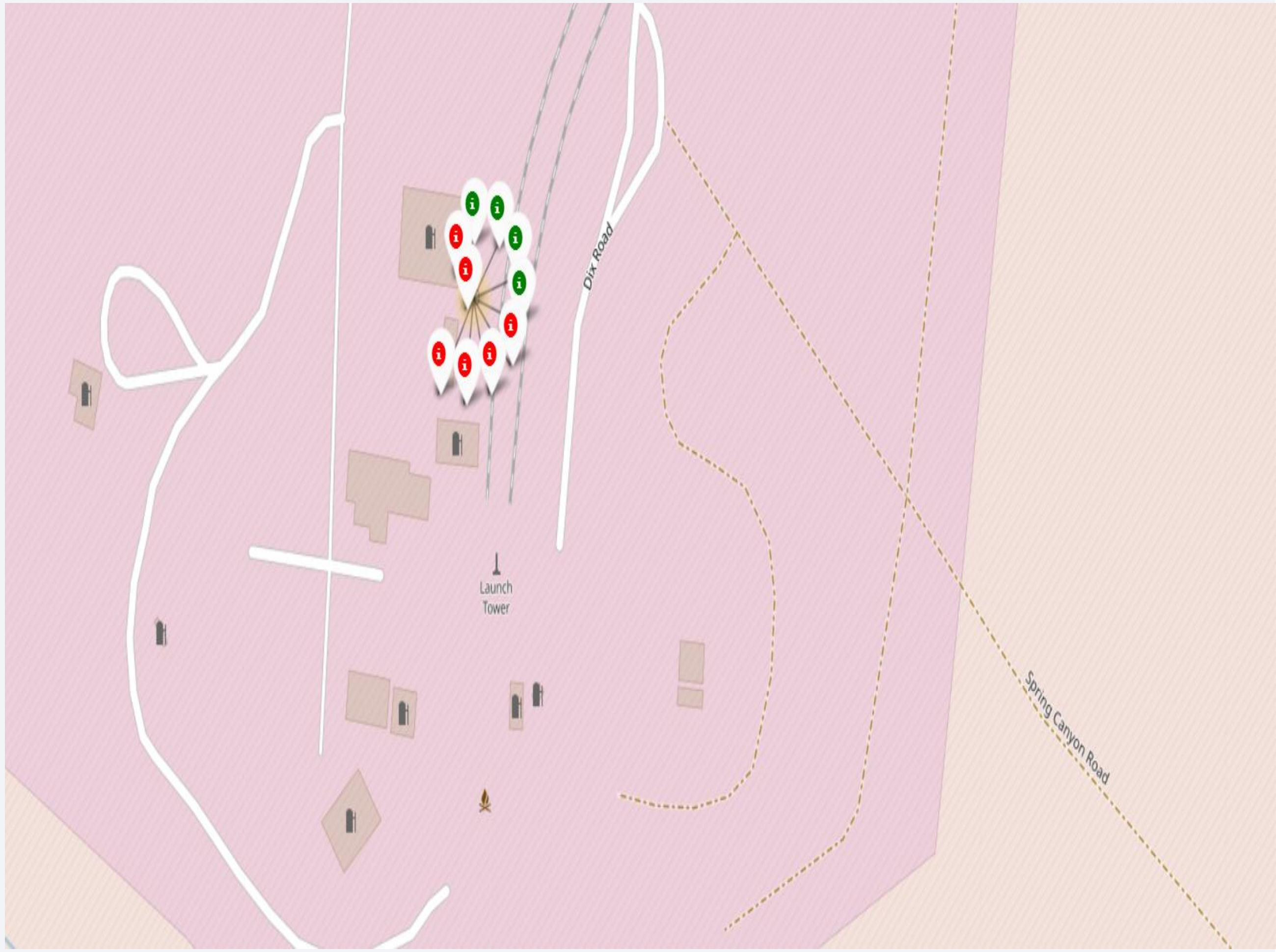
Interactive World map with launch sites

- Graph shows a simple markings on launch sites on the World map by working with the Longitude and Latitude of the Launch sites
  Lauch Sites are close to the Coast line.



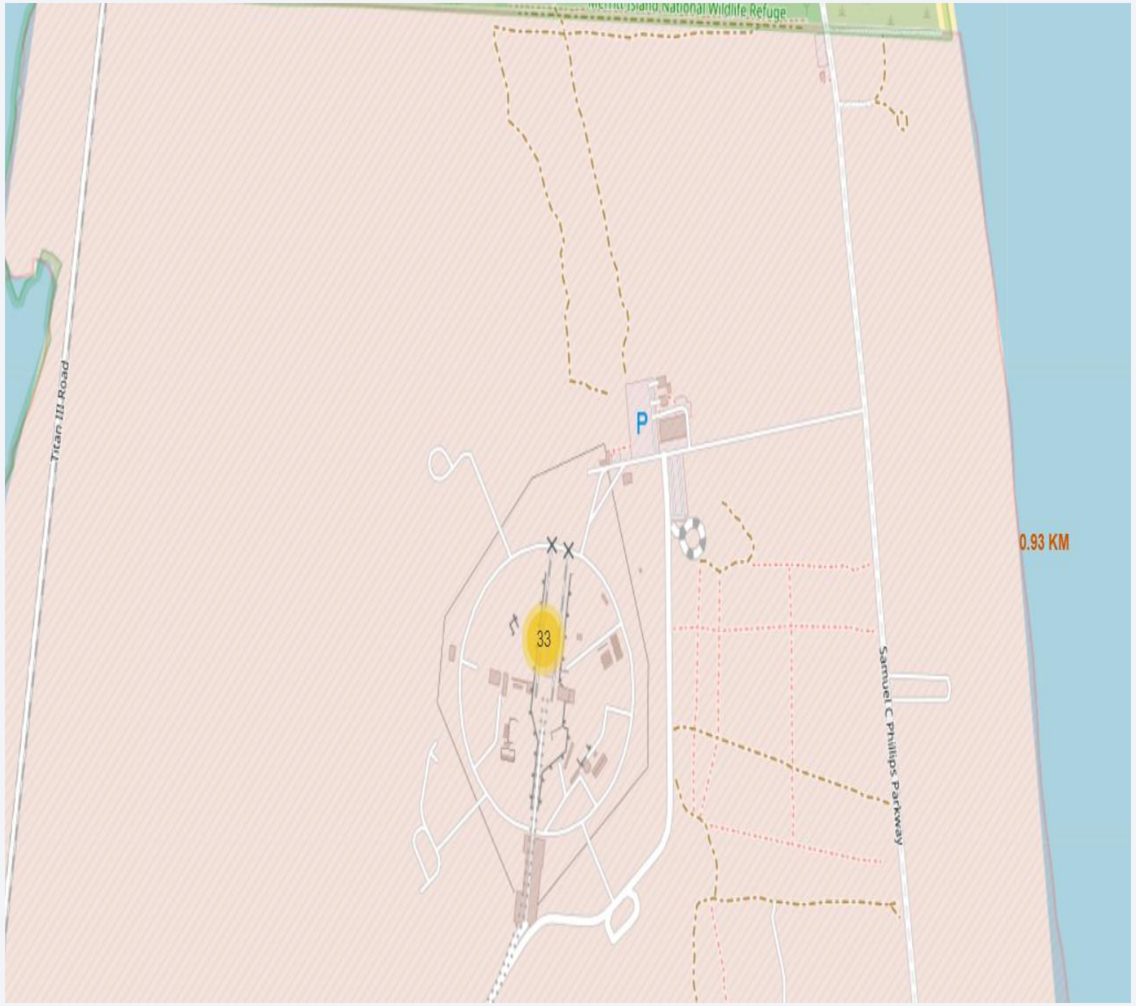
LauchSite with 10 Launches near LA

- The red and green markers provide a quick visual representation of the success and failure rates at this launch site.
- Total of 10 launches are performed from here with the success rate of 40%.



Distance to Coast Line

- In this graph the distance of launch site is measured from the Coast Line.
- Direct distance based on the longitude and latitude of launch sites and coastline is used.
- In this case distance found to be 0.93 kilo meters

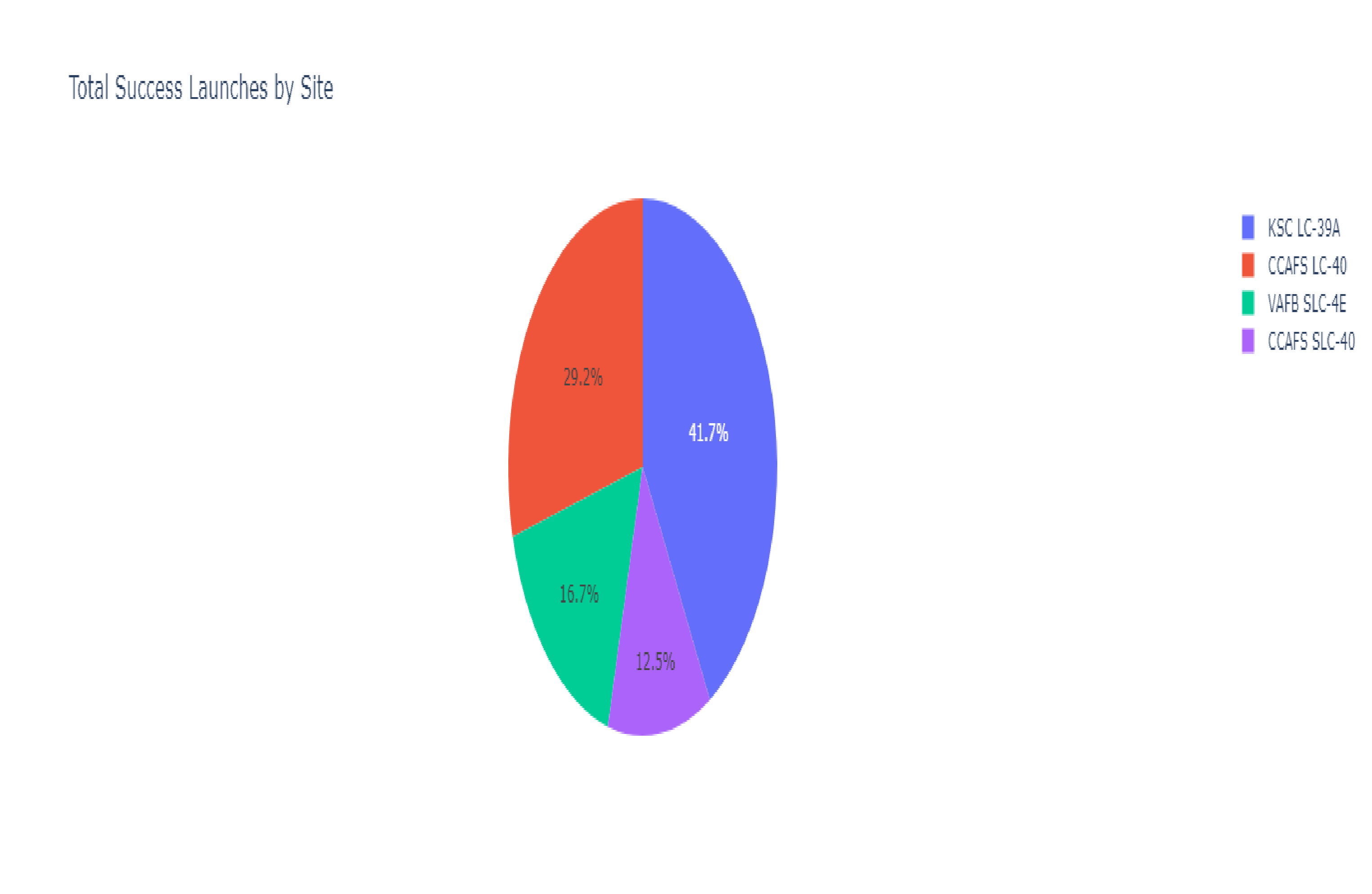


Section 4

Build a Dashboard with Plotly Dash

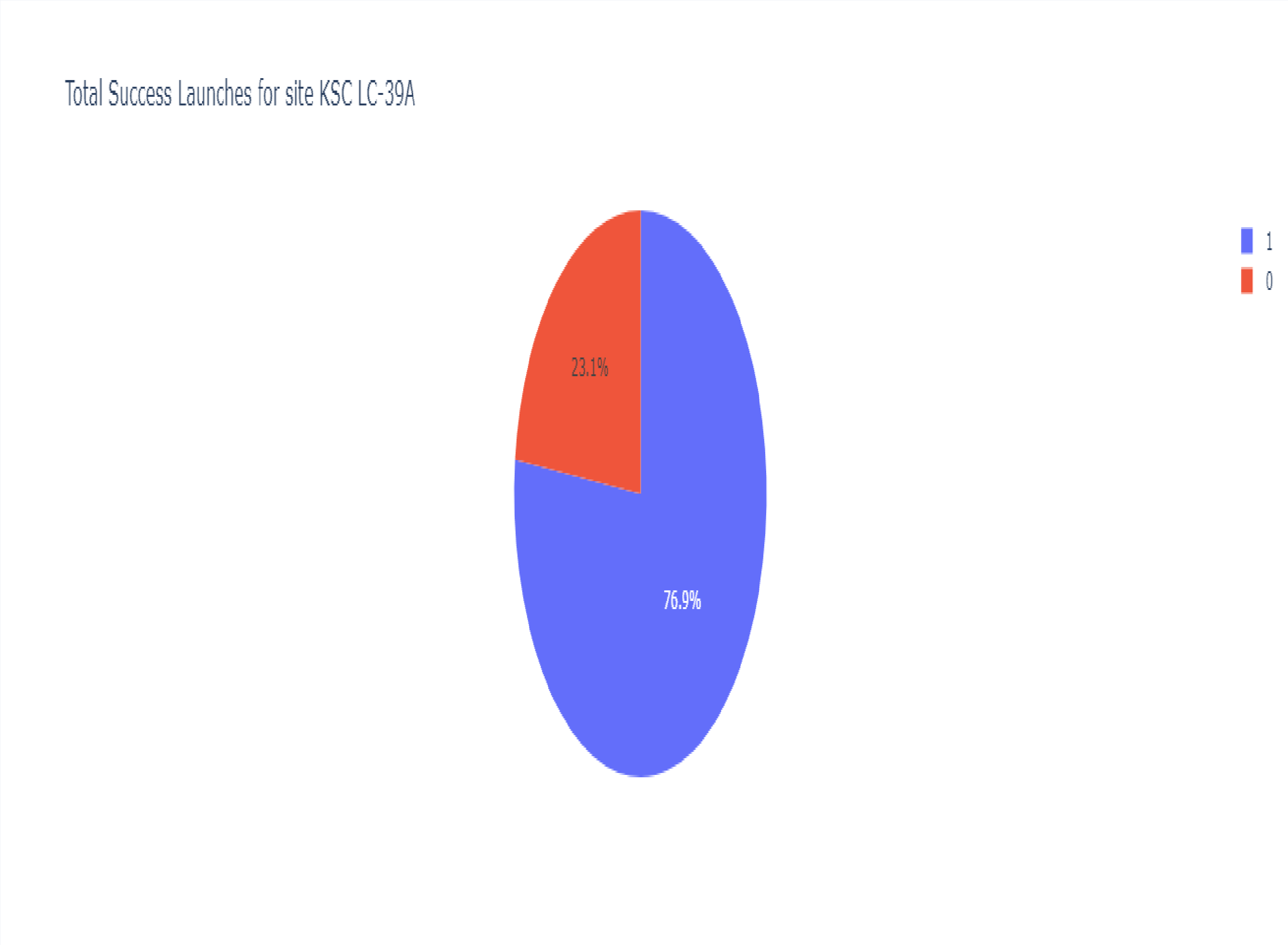
Plotly Pie Cart of Success for ALL Lauch Sites

- The pie chart provides a clear visual representation of the distribution of successful launches across different sites.
- KSC LC-39A and CCAFS LC-40 together account for over 70% of successful launches, indicating that these two sites are critical for SpaceX's operations.



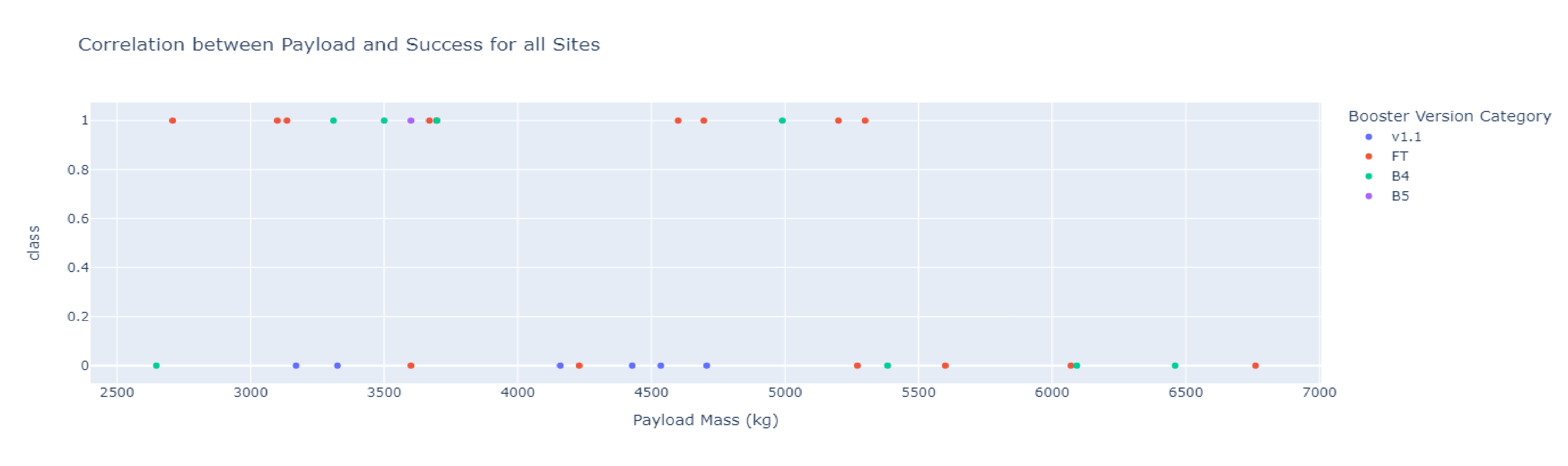
Highest Success Rate KSC LC-39A

- Out of all Four Launch Sites KC LC-39A has achieved a highest success rate of 76.9% as shown in Graph.



Scatter Plot Payload vs Launch Outcome with payload mass of 2500kg to 7500kg

- FT and B5 booster versions appear to have higher success rates and handle a broader range of payload masses.
- v1.1 and B4 booster versions show fewer data points, indicating either fewer launches or less data available for these versions.
- The high success rates for FT and B5 boosters across various payloads suggest that these versions are reliable and versatile.
- The scatter plot does not show a strong correlation between payload mass and success, indicating other factors (e.g., launch conditions, mission specifics) might play a more critical role in determining launch outcomes.

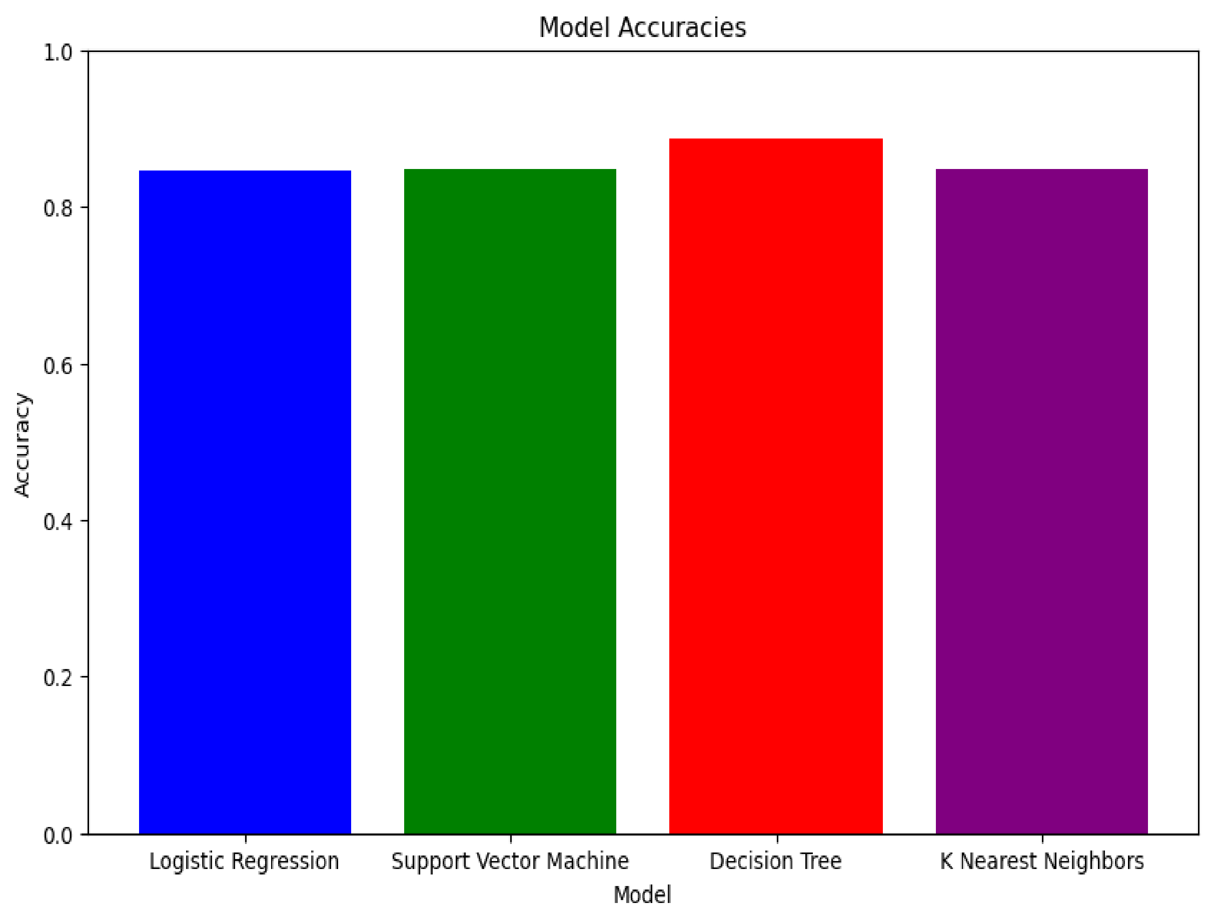


Section 5

Predictive Analysis (Classification)

Classification Accuracy

- The best performing model is Decision Tree with an accuracy of 88.75%.



Confusion Matrix

Performance Metrics:

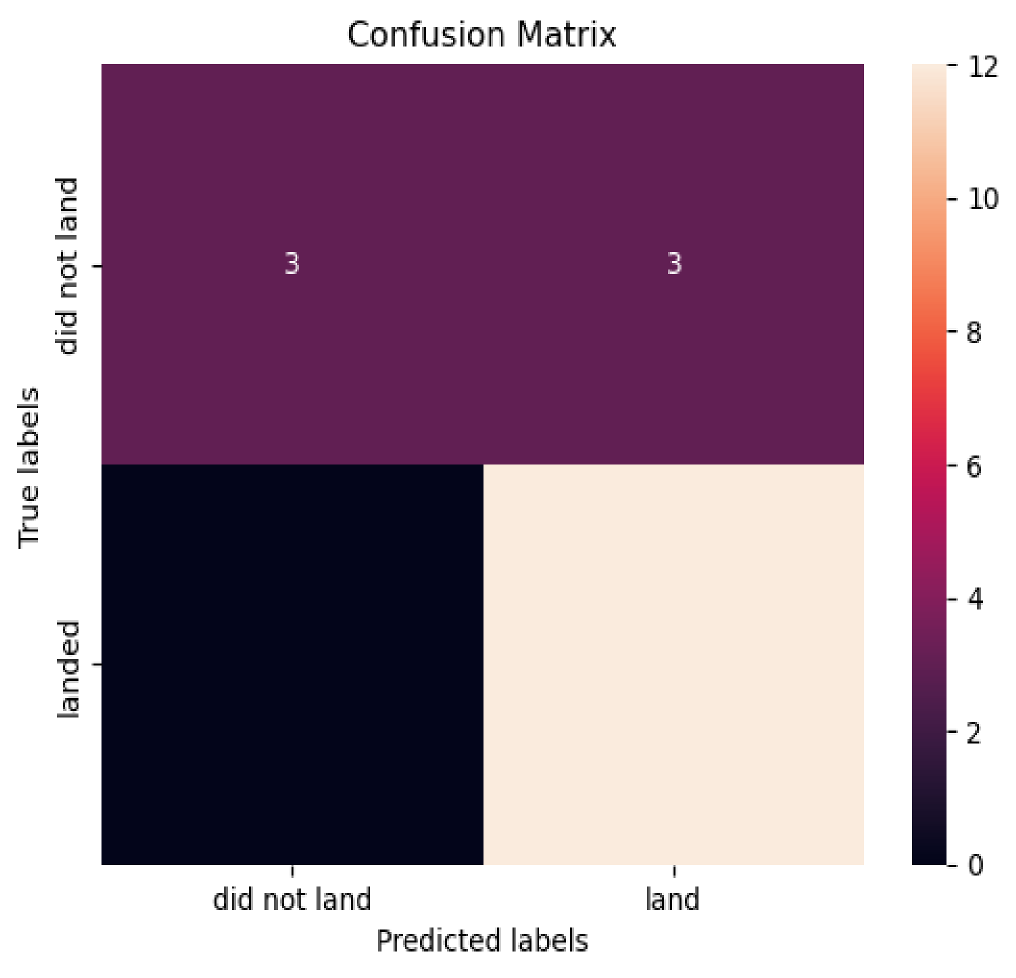
- Accuracy: The proportion of correct predictions (both TP and TN) out of all predictions.
- Accuracy = (TP + TN) / (TP + TN + FP + FN)
- Accuracy = (12 + 3) / (12 + 3 + 3 + 0) = 15 / 18 ≈ 0.833 or 83.3%

Precision for Class 1 (landed): The proportion of true positive predictions out of all positive predictions.

- Precision = TP / (TP + FP)
- Precision = 12 / (12 + 3) = 12 / 15 = 0.8 or 80%

Recall for Class 1 (landed): The proportion of true positive predictions out of all actual positives.

- Recall = TP / (TP + FN)
- Recall = 12 / (12 + 0) = 12 / 12 = 1 or 100%
- The Decision Tree model shows a strong performance in predicting successful landings with high precision and perfect recall. However, it has some false positives, where it predicts a successful landing that did not occur. The overall accuracy is 83.3%, indicating good performance with room for improvement, especially in reducing false positives.



Conclusions

- KSC LC-39A and CCAFS LC-40 together account for over 70% of successful launches, highlighting their critical importance to SpaceX operations. KSC LC-39A has the highest success rate of 76.9%, making it the most reliable launch site.
- The scatter plots indicated that the success of launches is not strongly correlated with payload mass alone. Both high and low payload masses saw successful outcomes across various booster versions, suggesting that other factors, such as mission specifics and launch conditions, play a more significant role in determining success.
- Orbits such as ES-L1, GEO, HEO, SSO, and VLEO exhibit perfect success rates, indicating reliable performance for these missions. In contrast, GTO and ISS have lower success rates, likely due to the complexity and precision required for these orbits.
- The analysis of launch success rates over the years shows a clear trend of improvement. Early missions had higher failure rates, but as flight numbers increased, the success rates improved significantly, reflecting operational learning and process optimization.
- Among the classification models built, the Decision Tree model performed the best with an accuracy of 88.75%. It showed high precision and perfect recall for successful landings, indicating strong predictive capability with some room for improvement in reducing false positives.
- The analysis revealed that FT and B5 booster versions have higher success rates and handle a broader range of payload masses compared to v1.1 and B4 versions. This suggests that newer booster versions are more reliable and versatile.
- The interactive map visualizations showed that launch sites are strategically located close to coastlines, which is beneficial for both safety and logistical reasons. Proximity to key infrastructure like highways and railways also supports efficient transportation and logistics.

Appendix

- All the files and code is presented at the GitHub repository and the link is given below.
  [nasirarslan35/IBM_capestone: final capestone (github.com)]

Thank you!