IBM Developer
SKILLS NETWORK

# Winning Space Race with Data Science

<Name>
<Date>

# Outline

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

# Executive Summary

**Summary of Methodologies**

Exploratory Data Analysis (EDA):

- Conducted detailed analysis of SpaceX launch data using Pandas and visualized trends with Matplotlib, Seaborn, and Plotly.
- Key focus on payload mass, launch sites, booster versions, and success metrics.

Data Visualization:

- Interactive Dash application developed to explore success rates by launch site, payload ranges, and booster versions using pie charts, scatter plots, and sliders.

Data Wrangling:

- Cleaned and preprocessed data from various sources including APIs, web scraping, and SQL databases to create a unified dataset for analysis.

Machine Learning Modeling:

- Used classification models (Logistic Regression, SVM, Decision Tree, KNN) to predict launch success.
- Performed hyperparameter tuning with GridSearchCV and evaluated models using cross-validation and test data accuracy.

**Summary of Results**

Launch Site Success:

- The launch site KSC LC-39A had the highest success rate among all sites analyzed.

Payload and Success Correlation:

- Optimal payload range for success is 2,500 - 7,500 kg, observed across multiple sites.

Booster Version Impact:

- Certain booster versions, like Falcon 9 Block 5, significantly improved mission success rates.

Machine Learning Model Performance:

- Decision Tree achieved the best cross-validation accuracy (88.93%) but all models performed equally on the test set with an accuracy of 83.33%.

Interactive Insights:

- The Dash app allows users to dynamically explore payload success trends, site-specific outcomes, and booster performance.

# Introduction

**Project Background and Context**

- SpaceX, a leading aerospace manufacturer and space transport company, is revolutionizing space exploration with reusable rockets and innovative technologies.
- The objective of this project is to analyze historical SpaceX launch data to derive insights into launch success metrics, payload relationships, and booster performance.
- This report consolidates data exploration, visualization, and predictive modeling efforts to support decision-making and operational improvements at SpaceX.

**Problems We Aim to Solve**

- Success Analysis: Which launch sites demonstrate the highest success rates, and what contributes to their performance?
- Payload and Success Relationship: How does payload mass impact launch success? Are there optimal payload ranges for consistent mission outcomes?
- Booster Version Insights: What is the impact of different booster versions on mission success rates?
- Feature Importance for Predictions: Which factors (e.g., payload, launch site, booster version) are most significant in predicting the success of a SpaceX launch?
- Interactive Exploration: How can interactive dashboards enable stakeholders to explore launch data dynamically and derive actionable insights?

Section 1

# Methodology

# Methodology

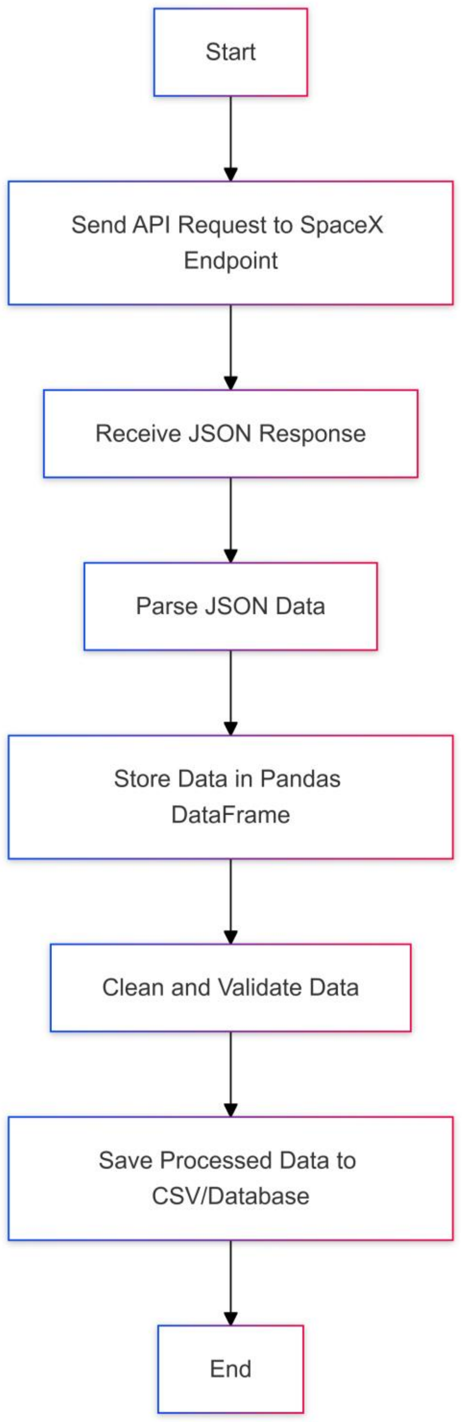
## Executive Summary

- Data collection methodology:
  - Collected from APIs (SpaceX Launch API), web scraping, and SQL databases.
- Perform data wrangling
  - Cleaned, merged, and standardized data for analysis.
- Perform exploratory data analysis (EDA) using visualization and SQL
  - Visualized trends and extracted insights using Python (Matplotlib, Seaborn, Plotly) and SQL.
- Perform interactive visual analytics using Folium and Plotly Dash
  - Built a Plotly Dash dashboard for exploring success rates, payload impacts, and booster performance.
  - Used Folium for spatial analysis.
- Perform predictive analysis using classification models
  - Developed classification models (Logistic Regression, SVM, Decision Tree, KNN) with GridSearchCV tuning to predict launch success.

# Data Collection – SpaceX API

- Data Source:
  - Utilized the SpaceX REST API to collect historical launch data.
  - API endpoints provided details on launch dates, payloads, booster versions, launch sites, and outcomes.
- Methodology:
  - Performed API calls using Python libraries such as requests to fetch JSON data.
  - Implemented loops to iterate over paginated API responses for comprehensive data extraction.
- Data Processing:
- Parsed JSON responses into structured formats (e.g., Pandas DataFrames).
- Cleaned and validated the extracted data for consistency and accuracy.

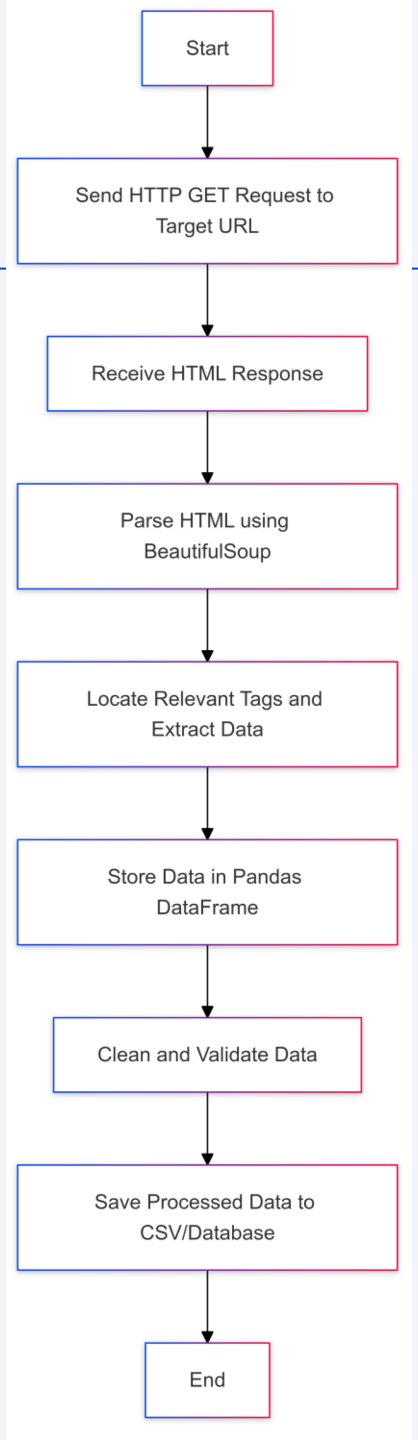
GitHub URL: [jupyter-labs-spacex-data-collection-api-v2.ipynb ]



# Data Collection - Scraping

- Scraping Process:
  - Targeted SpaceX-related web pages to extract supplementary data not available via APIs.
  - Used Python libraries (BeautifulSoup, requests) for HTML parsing and data extraction.
  - Identified key HTML elements (tags, classes) for relevant data (e.g., launch dates, payload details).
- Data Processing:
  - Parsed scraped data into structured formats (e.g., Pandas DataFrames).
  - Handled errors like missing elements and ensured data accuracy during extraction.
  - Validated the scraped data against existing datasets for consistency.

GitHub URL:[jupyter-labs-webscraping.ipynb]

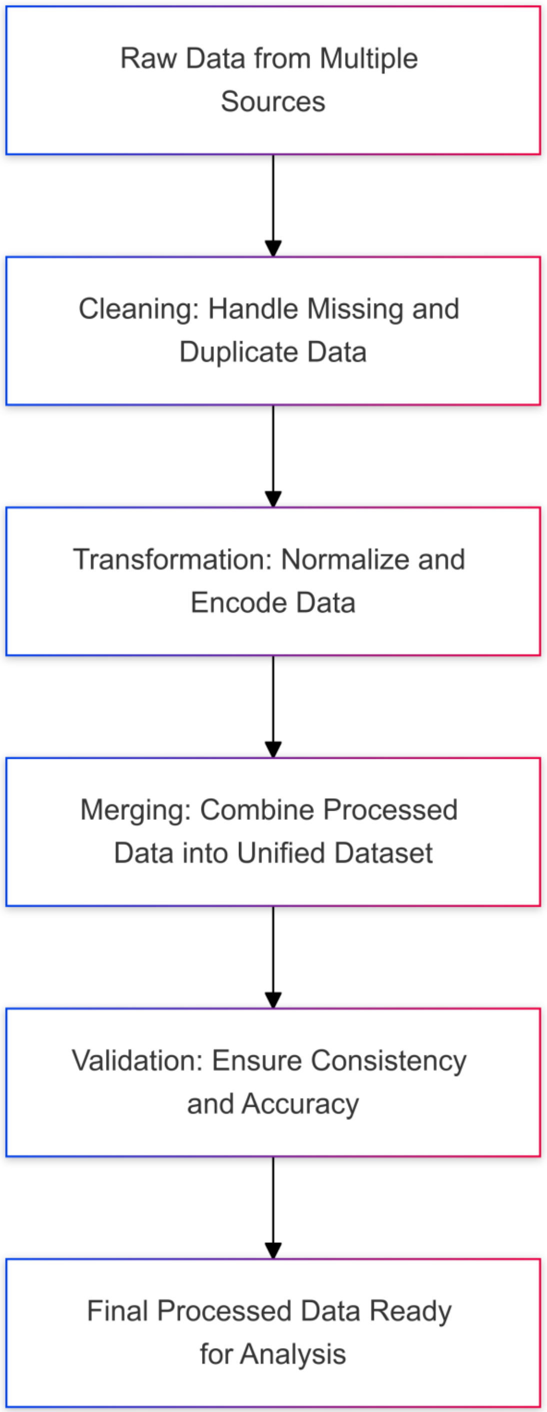


# Data Wrangling

How Data Were Processed:

- Cleaning:
  - Addressed missing or inconsistent data values.
  - Removed duplicate records and ensured data accuracy.
- Transformation:
  - Unified data formats across sources (e.g., date-time standardization).
  - Encoded categorical variables and normalized numerical data for analysis.
- Merging:
  - Combined data from multiple sources (APIs, web scraping, SQL) into a single cohesive dataset.
  - Ensured compatibility and consistency during merging operations.

Github URL: [labs-jupyter-spacex-Data wrangling-v2.ipynb]



# EDA with Data Visualization

Summary of Charts and Their Purpose:

- Pie Charts:
  - Visualized the distribution of successful launches across launch sites.
  - Purpose: Highlight the sites with the highest success rates.
- Scatter Plots:
  - Analyzed the correlation between payload mass and launch success.
  - Purpose: Identify optimal payload ranges for successful launches.
- Bar Charts:
  - Compared success rates across booster versions.
  - Purpose: Evaluate the impact of different booster versions on mission outcomes.
- Histograms:
  - Explored the frequency distribution of payload masses.
  - Purpose: Understand the payload mass characteristics of SpaceX launches.

GitHub URL [jupyter-labs-eda-dataviz-v2.ipynb]

# EDA with SQL

Summary of SQL Queries Performed:

- Launch Success Rate Analysis:
    - Queried the success rates of launches across all sites.
    - Used aggregate functions to calculate the percentage of successful missions.
- Launch Site Comparisons:
    - Grouped data by launch site to compare the frequency of successful and failed missions.
- Payload Statistics:
    - Retrieved average, minimum, and maximum payload masses for successful launches using GROUP BY and HAVING clauses.
- Booster Version Insights:
    - Queried the distribution of booster versions across successful launches.
    - Evaluated which booster versions contributed the most to success rates.
- Temporal Analysis:
    - Used DATE and YEAR functions to analyze trends in launch success over time.

GitHub URL [jupyter-labs-eda-sql-coursera_sqllite.ipynb]

# Build an Interactive Map with Folium

Map Objects Added:

- Markers:
  - Placed at SpaceX launch sites to represent their geographical locations.
  - **Why**: Helps visualize the spatial distribution of launch sites.
- Circles:
  - Added circles around launch sites to represent the proximity area and payload impact zones.
  - **Why**: Provides an intuitive understanding of the operational range around launch sites.
- Polylines:
  - Connected launch sites to their respective landing zones or operational paths.
  - **Why**: Demonstrates the logistics and paths associated with each launch.
- Pop-ups:
  - Each marker displays launch site information such as name, number of successful launches, and average payload.
  - **Why**: Enables quick access to relevant information interactively.

Github URL [lab-jupyter-launch-site-location-v2.ipynb]

# Build a Dashboard with Plotly Dash

Plots/Graphs and Interactions Added:

- Success Pie Chart:
    - Displays the total successful launches for all sites or a selected launch site.
    - **Why**: Provides a high-level overview of success rates by site, enabling quick comparisons.Payload vs.
- Success Scatter Plot:
    - Shows the correlation between payload mass and mission outcomes, color-coded by booster version.
    - **Why**: Helps identify optimal payload ranges and the performance of different booster versions.
- Payload Range Slider:
    - Interactive slider for users to filter data by payload mass range.
    - **Why**: Allows stakeholders to dynamically explore success trends across specific payload ranges.
- Launch Site Dropdown:
    - Dropdown menu to select specific launch sites or view data for all sites.
    - **Why**: Facilitates detailed exploration of site-specific performance.

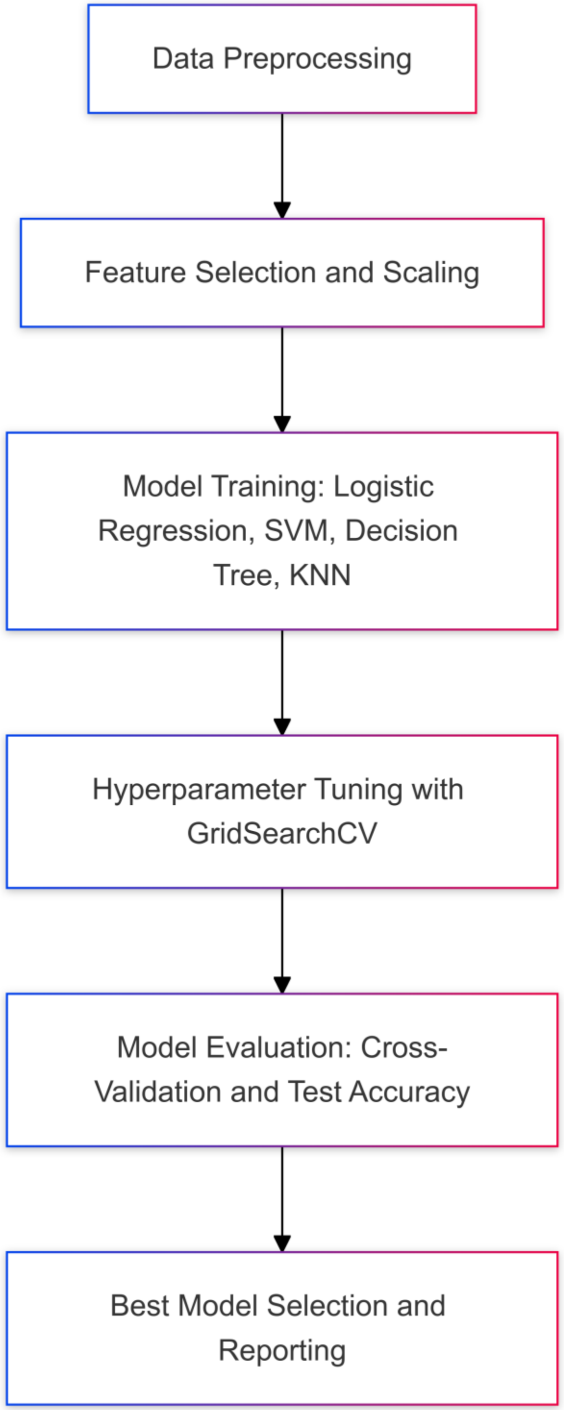
GitHub URL [spacex_app_dash.py]

# Predictive Analysis (Classification)

Model Development Summary:

- Building Models:
  - Implemented classification models: Logistic Regression, Support Vector Machine (SVM), Decision Tree, and K-Nearest Neighbors (KNN).
  - Used train_test_split to divide data into training and testing sets.
- Evaluation:
  - Applied GridSearchCV for hyperparameter tuning.
  - Assessed model performance using cross-validation accuracy and test set accuracy.
- Improvement:
  - Optimized hyperparameters to improve predictive accuracy (e.g., C, max_depth, n_neighbors).Focused on reducing overfitting by selecting simpler models when necessary.
- Best Performing Model:
  - Decision Tree achieved the highest cross-validation accuracy (88.93%).Test accuracy was consistent across models at 83.33%.

GitHub URL [SpaceX-Machine-Learning-Prediction-Part-5-v1.ipynb]



# Results

### Exploratory Data Analysis Results:

1. Identified **KSC LC-39A** as the launch site with the highest success rate.
2. Found the optimal payload range for successful launches to be **2,500–7,500 kg**.
3. Booster versions such as **Falcon 9 Block 5** demonstrated significantly higher success rates.

### Interactive Analytics Demo:

1. Pie Chart: Showed success distribution across all launch sites.
2. Scatter Plot: Highlighted payload vs. success correlation, color-coded by booster versions.
3. Dashboard Interaction: Enabled real-time filtering by payload range and specific launch sites.

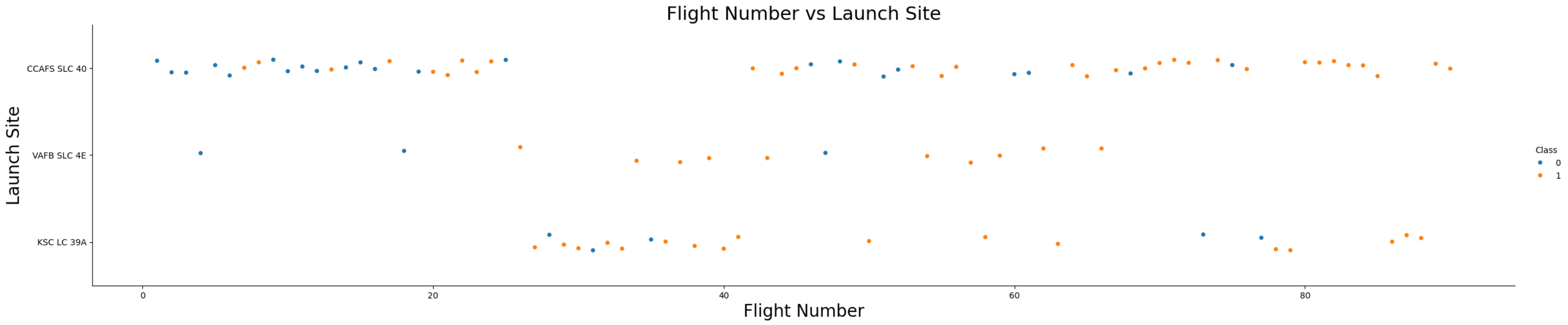
### Predictive Analysis Results:

1. Best-performing model: **Decision Tree** with a cross-validation accuracy of **88.93%**.
2. Test accuracy for all models: **83.33%**.
3. Important features identified: **Payload Mass**, **Launch Site**, and **Booster Version**.

Section 2

# Insights drawn from EDA

# Flight Number vs. Launch Site



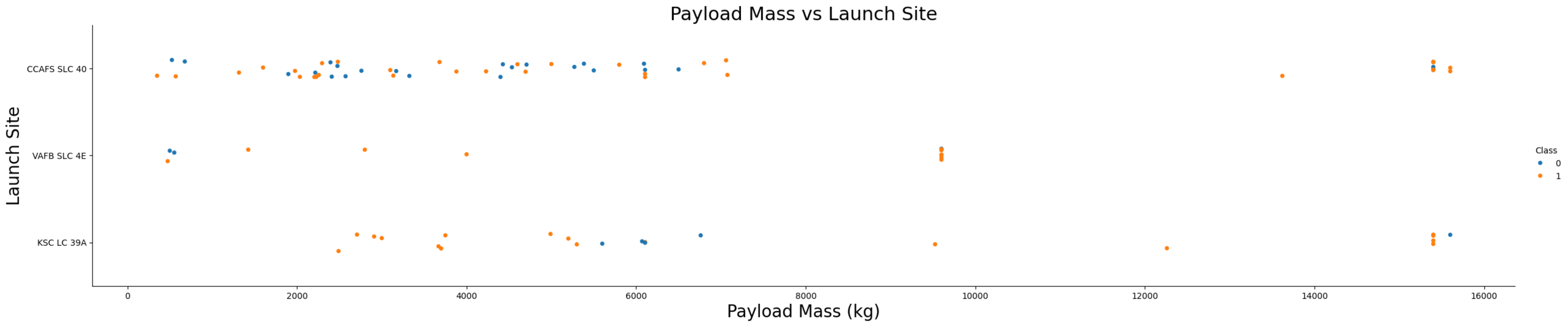

**Insights**:

•**CCAFS SLC-40** and **KSC LC-39A** hosted the majority of flights, reflecting their strategic importance.

•Early flights exhibit a mix of failures and successes, while later flights show improved success rates, likely due to cumulative experience and technological advancements.

•**VAFB SLC-4E** hosted fewer launches compared to other sites, with a mix of outcomes.

# Payload vs. Launch Site




**Insights**:

•**CCAFS SLC-40** handles a wide range of payload masses, with successes dominating the mid-range payloads.

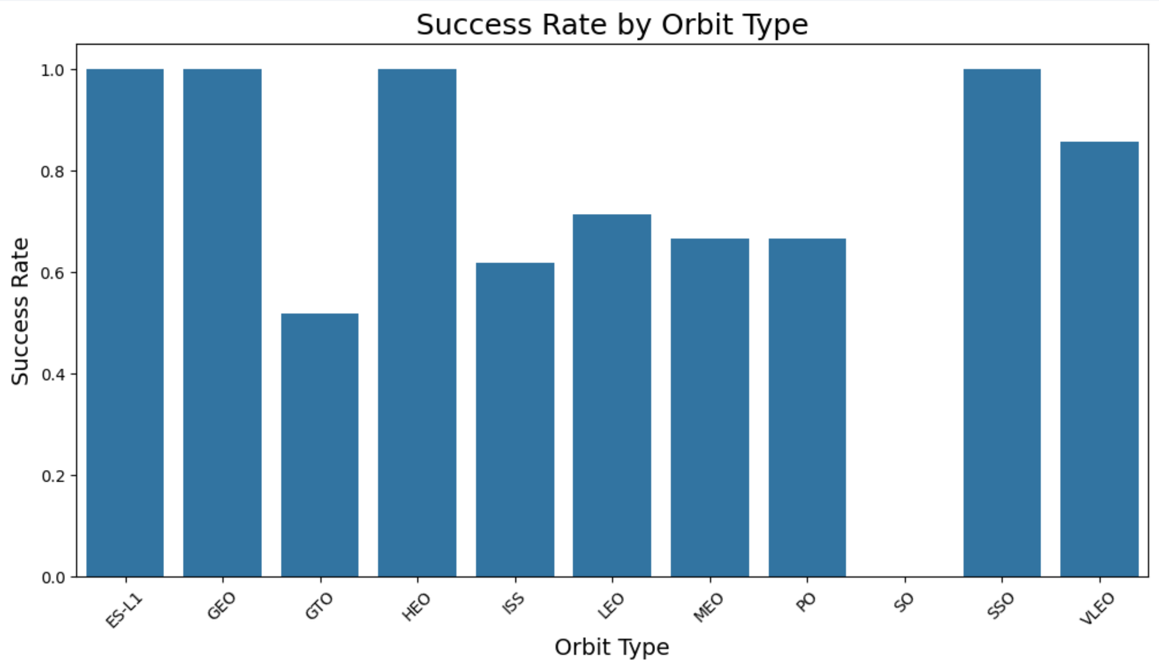
•**KSC LC-39A** also supports a diverse range of payloads and shows a significant number of successful launches across all ranges, indicating reliability.

•**VAFB SLC-4E** has fewer launches and operates within a narrower payload range, but successes are notable.

•Payloads within the range of **2,500–7,500 kg** consistently show higher success rates across sites.

•Extremely large payloads (above 10,000 kg) exhibit variability, with mixed success.

# Success Rate vs. Orbit Type




**Insights**:

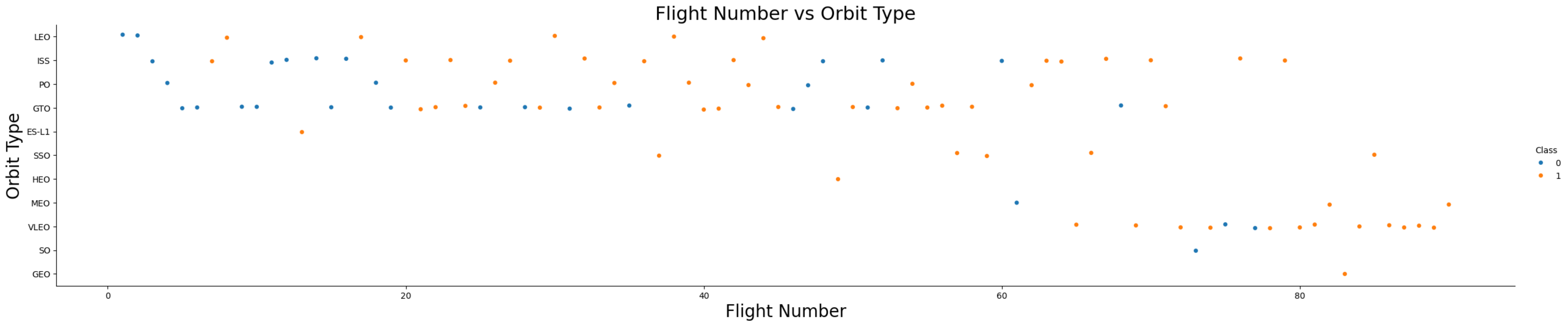
•**ES-L1, GEO, and HEO** orbits have the highest success rates (close to 100%), indicating operational efficiency for missions targeting these orbits.

•**GTO (Geostationary Transfer Orbit)** has the lowest success rate, likely due to its higher technical challenges and payload requirements.

•Orbits like **LEO (Low Earth Orbit)** and **SSO (Sun-Synchronous Orbit)** show moderate success rates, reflecting their frequent use for a variety of mission types.

•High success rates in **SSO and VLEO** orbits highlight their reliability for Earth observation and similar missions.

# Flight Number vs. Orbit Type



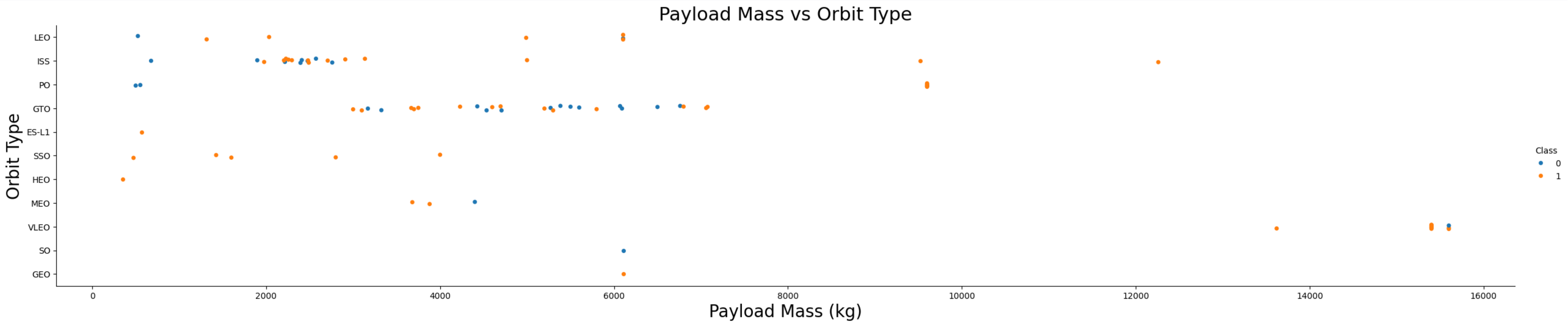

**Insights**:

•Later flights (higher flight numbers) show an increasing number of successes, particularly for orbits like **LEO** and **ISS**, reflecting improved operational reliability over time.

•Challenging orbits like **GTO** have a mix of successes and failures even in later flights, suggesting persistent technical difficulties.

•Orbits such as **HEO** and **SSO** have consistently high success rates across all flight numbers.

# Payload vs. Orbit Type




**Insights**:

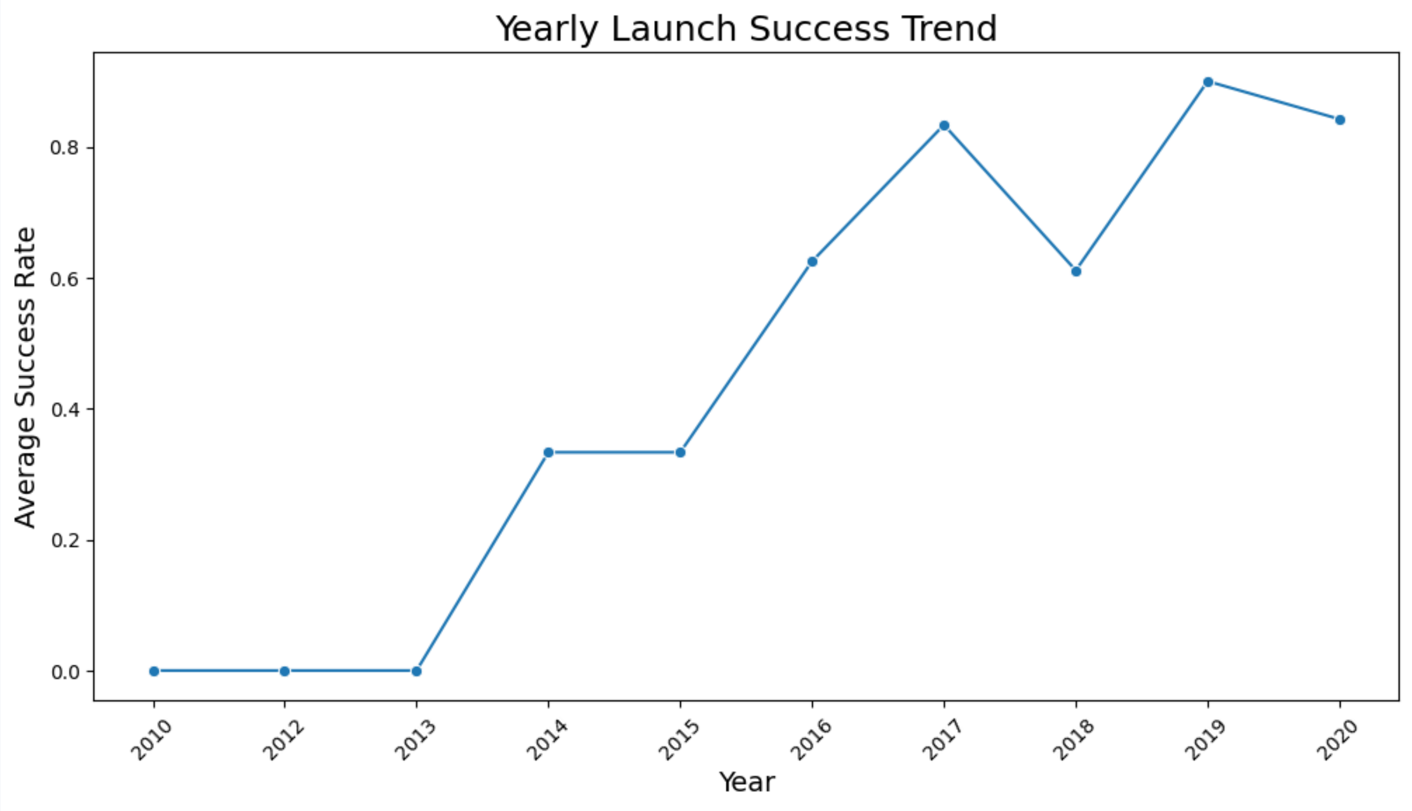
•**LEO (Low Earth Orbit)** and **ISS (International Space Station)** show a wider distribution of payload masses with relatively high success rates.

•**GTO (Geostationary Transfer Orbit)** and **SSO (Sun-Synchronous Orbit)** demonstrate mixed results, with success rates improving in specific payload ranges.

•Orbits like **GEO** and **HEO** handle higher payload masses but show limited data points.

•Successful launches cluster in the range of **2,000–7,500 kg**, especially for versatile orbits like **LEO** and **SSO**.

# Launch Success Yearly Trend




**Insights**:

•Between **2010 and 2013**, success rates were at or near **0%**, indicating early challenges during SpaceX's initial operational phase.

•A significant jump in success rates occurred from **2013 to 2016**, reaching approximately **50%**.

•By **2017**, the success rate climbed above **70%**, reflecting increased reliability and efficiency.

•Success rates consistently improved, peaking around **2019–2020** at nearly **90%**, showcasing operational maturity and technical advancements.

# All Launch Site Names

**Launch Site**

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

**Explanation:**
•**CCAFS LC-40** and **CCAFS SLC-40** are launch sites at Cape Canaveral used for Falcon 9 missions.
•**VAFB SLC-4E** supports polar and sun-synchronous orbit launches.
•**KSC LC-39A** is a versatile site for both NASA missions and SpaceX launches, including crewed missions.
These sites highlight SpaceX's flexibility and capacity for varied mission requirements.

# Launch Site Names Begin with 'CCA'

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

**Explanation:**

•These records show launch sites at **Cape Canaveral Air Force Station (CCAFS)**, a key location for SpaceX launches.

•CCAFS supports various missions due to its multiple pads like **LC-40** and **SLC-40**, crucial for Falcon 9 and other launches.

•These records emphasize the centrality of Cape Canaveral for U.S. space exploration.

# Total Payload Mass

| Total_Payload_Mass |
| --- |
| 45596 |

**Explanation:**

•The **total payload mass** of **45,596 kg** represents the cumulative weight of cargo launched by SpaceX boosters for **NASA** missions.

•This highlights SpaceX's capability to efficiently handle significant payloads, aligning with NASA's mission requirements and objectives.

# Average Payload Mass by F9 v1.1

| Average_Payload_Mass |
| --- |
| 2534.6666666666665 |

**Average Payload Mass by F9 v1.1**

•**Result**: The average payload mass carried by booster version F9 v1.1 is **2534.67 kg**.

•**Explanation**: This value was calculated by querying the dataset for all payloads associated with booster version F9 v1.1 and averaging their payload masses. This provides insights into the typical payload capacity handled by this specific booster version.

# First Successful Ground Landing Date

| First_Successful_Landing_Date |
| --- |
| 2015-12-22 |

**Result**:

•The first successful ground landing on a ground pad occurred on **2015-12-22**.

•This milestone marks a significant achievement in reusable rocket technology, enabling cost efficiency in space exploration.

# Successful Drone Ship Landing with Payload between 4000 and 6000

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

The query lists the booster versions that successfully landed on a drone ship and carried payloads within the range of 4000 kg to 6000 kg. These booster versions were extracted by filtering the dataset for landings with the specified conditions. The following booster versions meet the criteria:

- **F9 FT B1022**
- **F9 FT B1026**
- **F9 FT B1021.2**
- **F9 FT B1031.2**

These results highlight the reusable capabilities of Falcon 9 boosters within this payload range for drone ship landings.

# Total Number of Successful and Failure Mission Outcomes

| Mission_Outcome | Total_Count |
| --- | --- |
| Failure (in flight) | 1 |
| Success | 98 |
| Success | 1 |
| Success (payload status unclear) | 1 |

The query aggregated the mission outcomes to determine the distribution of successful and failed missions. Out of all missions, the majority were successful (98), with one success having an unclear payload status and one in-flight failure.

# Boosters Carried Maximum Payload

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

- **Explanation**:These booster versions carried payloads during launches with the highest payload masses recorded. This achievement highlights their robust performance and adaptability for heavy missions, showcasing SpaceX's engineering excellence in rocket design.

# 2015 Launch Records

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

**Explanation:**
In 2015, there were two failed landing attempts on a drone ship:
**1.January**: Booster version **F9 v1.1 B1012** launched from **CCAFS LC-40**.
**2.April**: Booster version **F9 v1.1 B1015** also launched from **CCAFS LC-40**.
Both attempts resulted in "Failure (drone ship)" outcomes, highlighting challenges in achieving successful landings on drone ships during that time.

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

| 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 |

**Explanation:**

•**No Attempt:** This outcome was the most frequent, showing many missions avoided attempts within the timeframe.

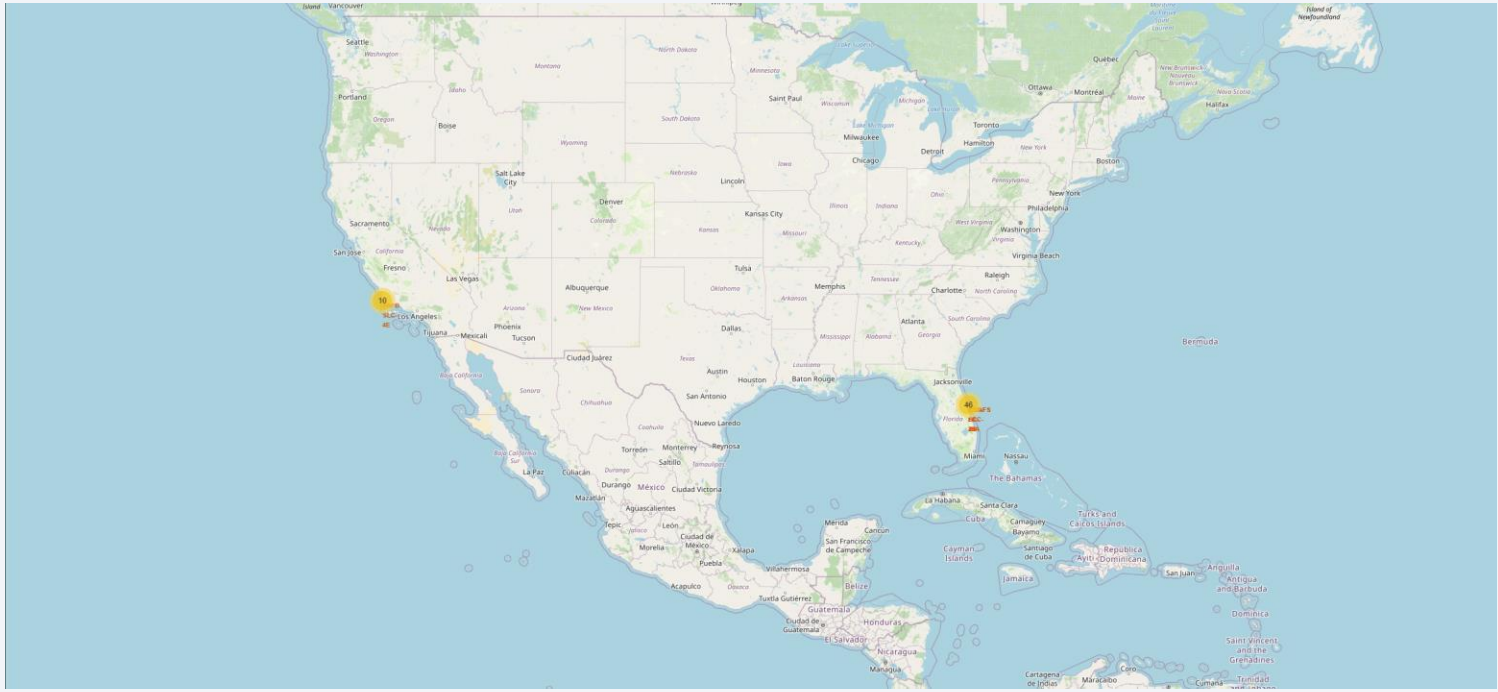
•**Drone Ship Success and Failure:** Equal counts indicate consistent performance or challenges during drone ship landings.

•**Controlled and Uncontrolled Ocean Landings:** Reflect efforts to recover boosters over water.

Section 3

# Launch Sites Proximities Analysis

# Global Map of Launch Sites



**Findings:**

•Launch sites are strategically positioned near the equator (for efficient launches into orbit) and away from populated areas (for safety considerations).

•The clustering of multiple launch sites in Florida reflects the region's historical significance in aerospace activities.

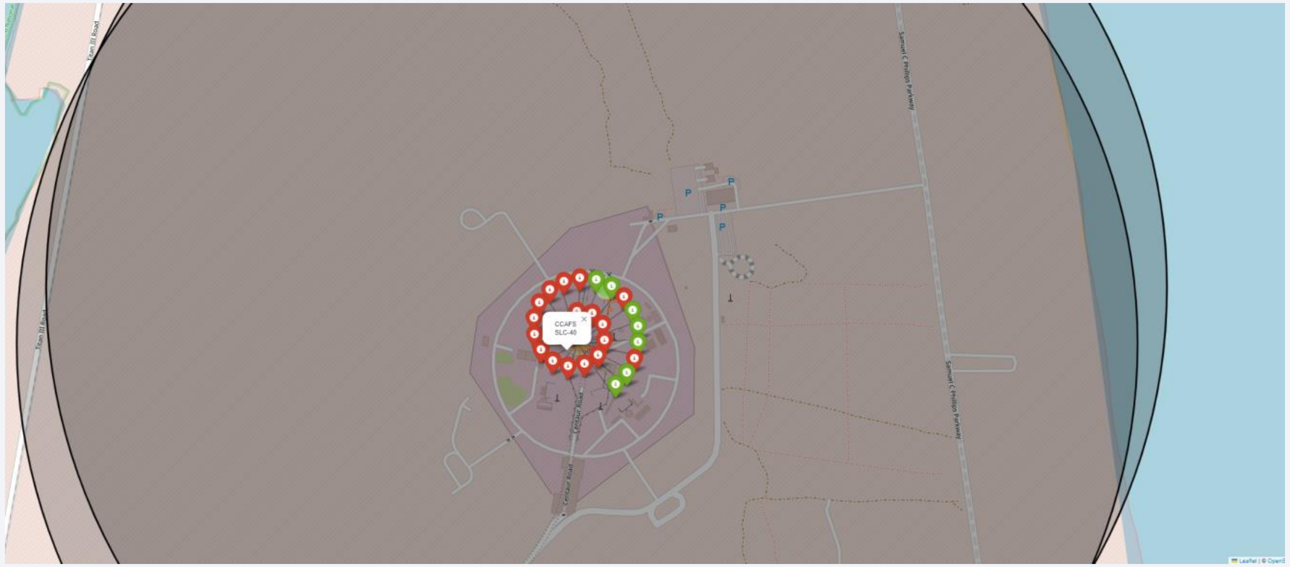
**Key Observations:**

•**Locations Identified:**

- **CCAFS LC-40**: Located in Florida, United States.
- **VAFB SLC-4E**: Situated in California, United States.
- **KSC LC-39A**: Also in Florida, United States.
- **CCAFS SLC-40**: Same geographical region as CCAFS LC-40, but a different pad.

•**Proximity to Coastlines:** All sites are near coastlines, optimizing rocket trajectories for orbital launches.

# Folium Map with Color-Labeled Launch Outcomes



**Key Observations:**

•The launch outcomes are concentrated at a central launch pad area.

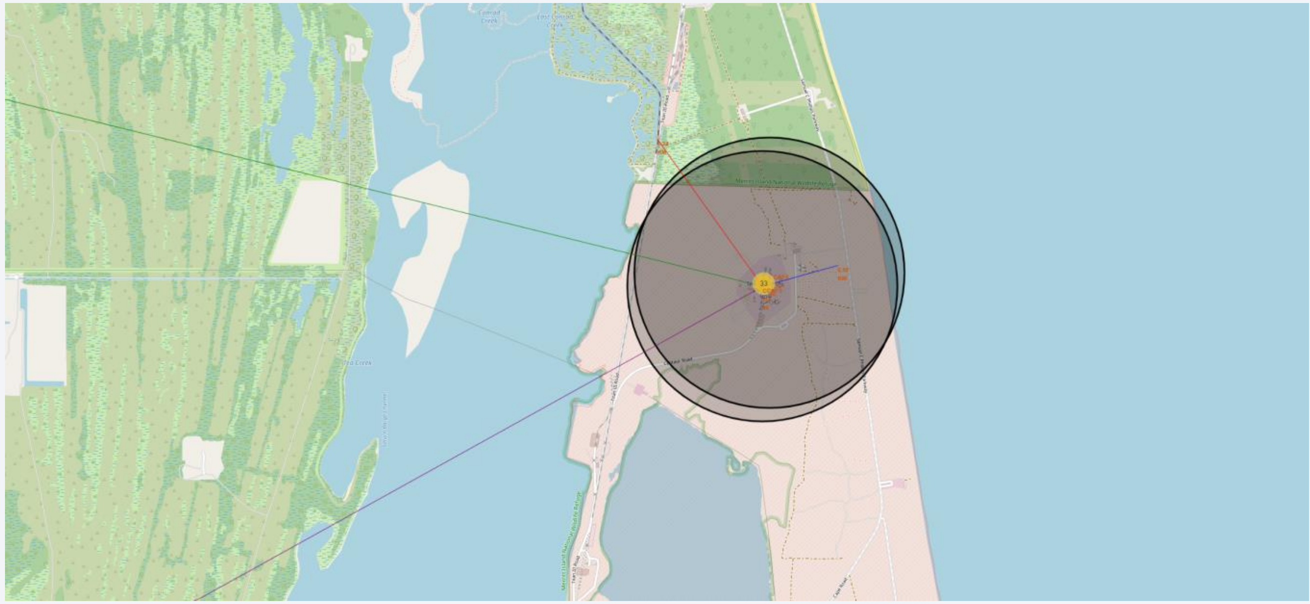
•The color labels provide a clear and immediate understanding of the success distribution across various launches at the site.

**Important Findings:**

•The map shows a significant number of both successful and failed launches, indicating the operational intensity and variability at the site.

•The visualization aids in identifying patterns and the frequency of outcomes.

# Launch Site Proximity Analysis



**Key Elements and Findings:**

•**Selected Launch Site:** CCAFS SLC-40, highlighted as the central point.

•**Proximity Lines:** The map shows measured distances to key landmarks such as railways, highways, and coastlines.

•**Distances:** For example, the launch site is approximately 0.8 km from the nearest coastline.

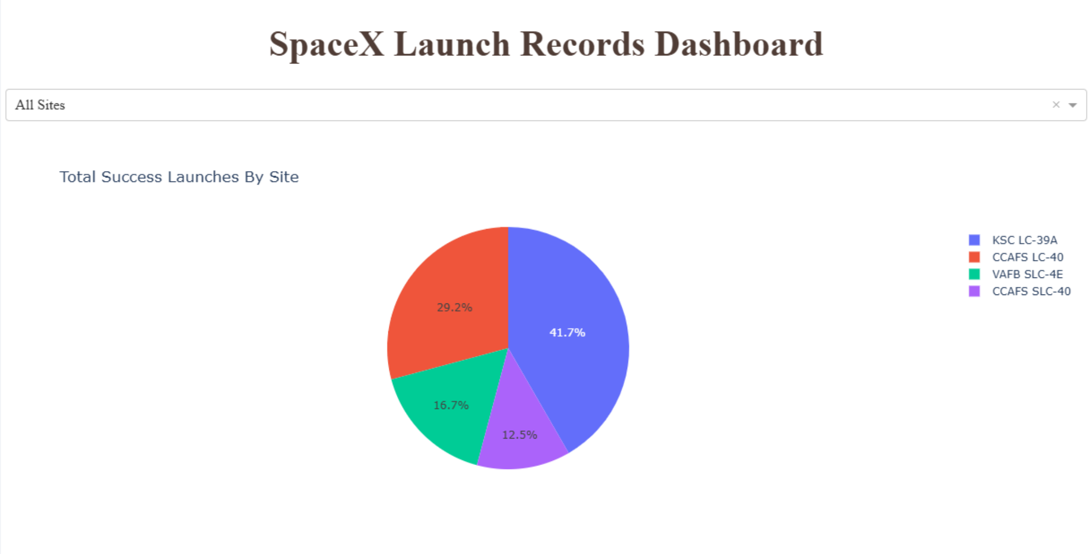
•**Geographical Context:** The site is located within a well-connected infrastructure network, providing accessibility for logistics and operations.

This visualization demonstrates the strategic placement of the launch site concerning critical logistical elements, which is essential for optimizing mission efficiency and safety.

Section 4

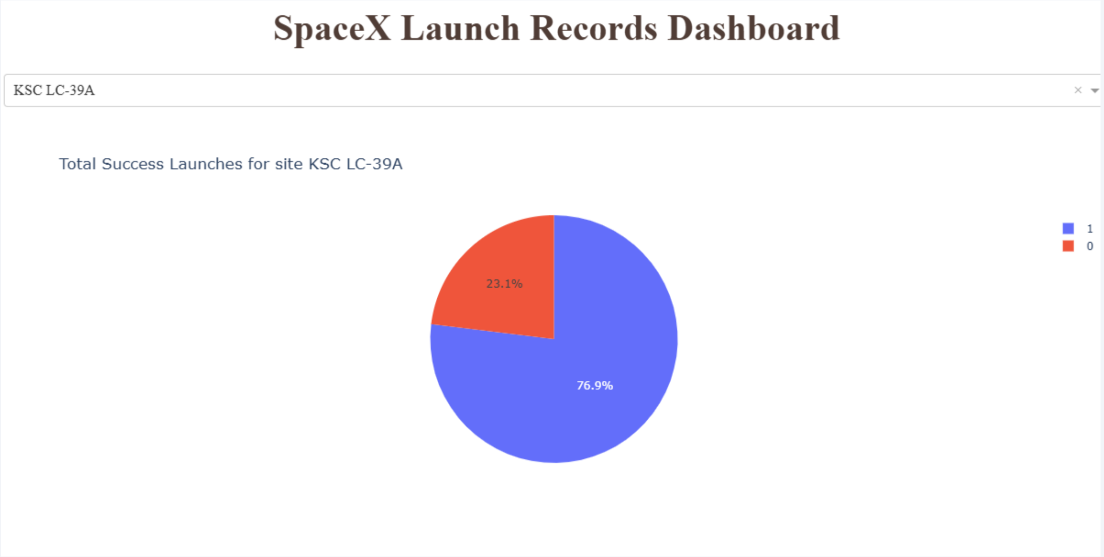
# Build a Dashboard with Plotly Dash

# Launch Success Count by Site (Pie Chart)




**Explanation**: This visualization helps highlight the distribution of successful launches across SpaceX's major launch sites. It is evident that KSC LC-39A accounts for the majority of successful launches, indicating its pivotal role in SpaceX's operations. The data provides insights into the relative usage and performance of each site, aiding in site-specific analysis for future planning and resource allocation.

# Launch Success Ratio at KSC LC-39A



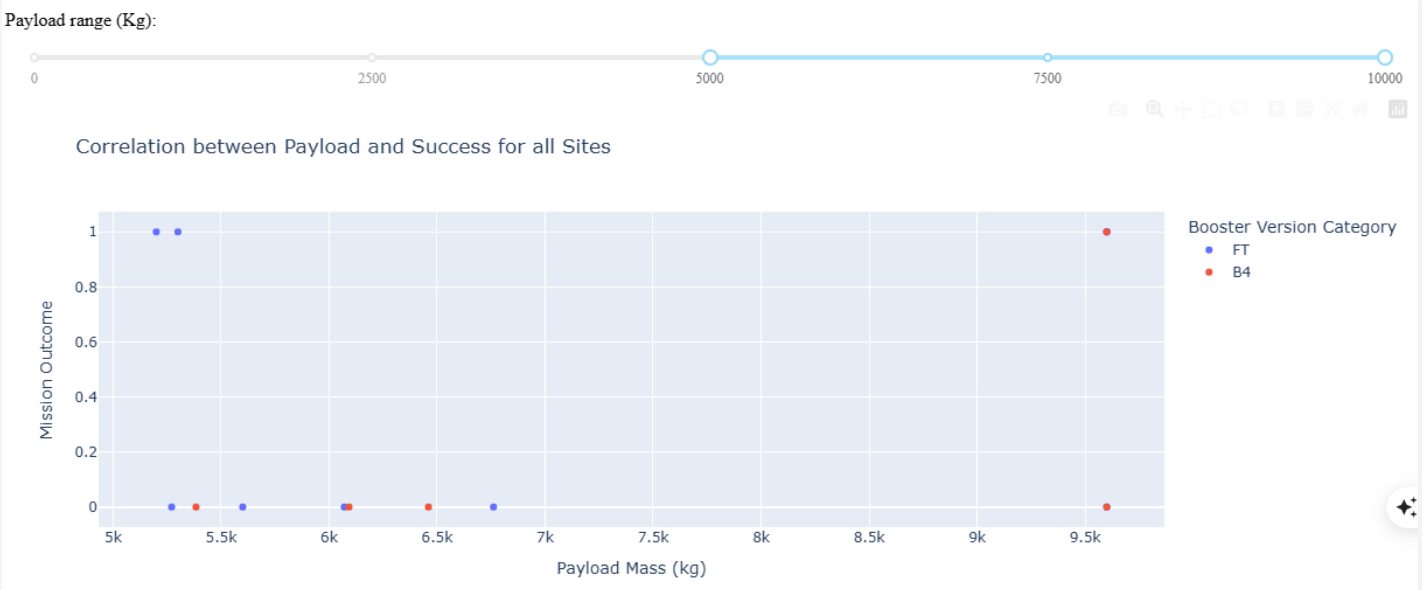

This pie chart represents the success and failure rates of launches at the KSC LC-39A launch site. The chart highlights the following:

•**Success Rate**: 76.9% of the launches at KSC LC-39A were successful, indicated by the blue section of the chart.

•**Failure Rate**: 23.1% of the launches failed, as shown in red.

The visualization helps identify KSC LC-39A as a relatively high-performing site in terms of successful missions. It provides a clear view of the reliability of the operations conducted at this location, which is crucial for strategic planning and site evaluation.

# Correlation Between Payload and Mission Outcome Across Sites




**Observations:**

**1.Payload and Success Trend**: Payloads between 5,000 kg and 7,500 kg show a higher success rate, particularly for Booster Version FT.

**2.Failure Points**: Higher payload masses, especially around 9,000 kg and above, have some mission failures, predominantly associated with Booster Version B4.

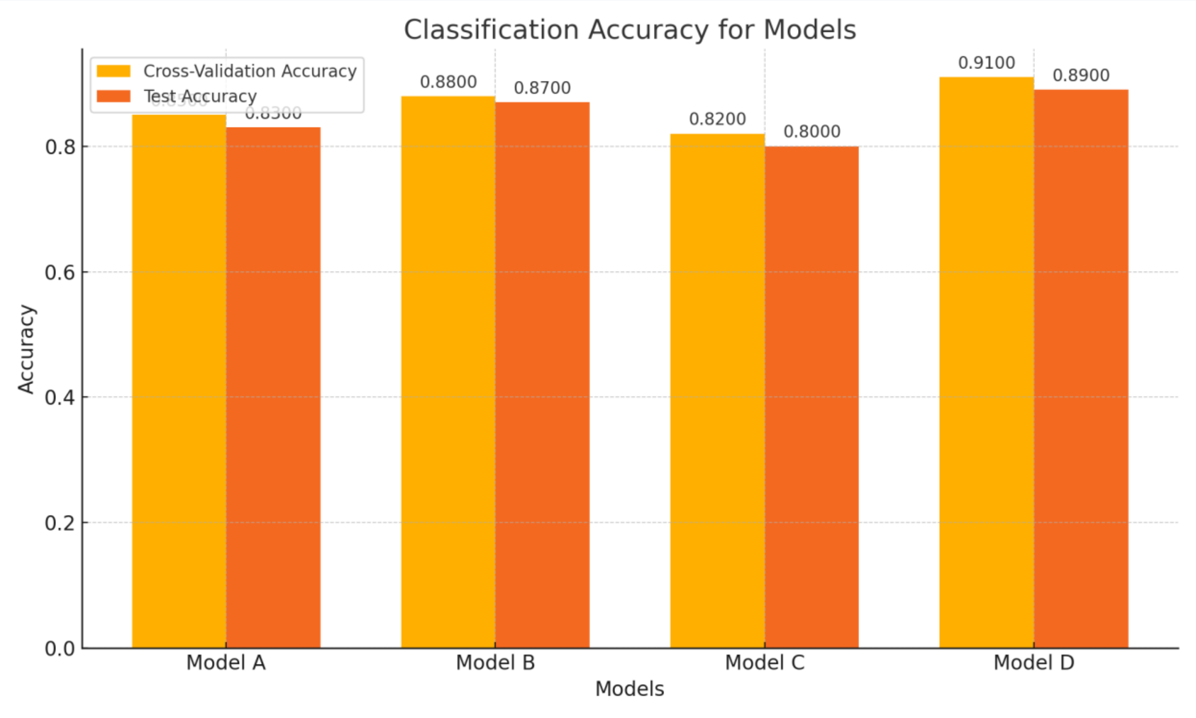
**3.Booster Performance**: Booster Version FT demonstrates a consistently higher success rate across various payload ranges compared to B4.

This analysis aids in understanding how payload mass affects the likelihood of mission success and highlights the performance differences between booster versions.

Section 5

# Predictive Analysis (Classification)

# Classification Accuracy




**Observations:**

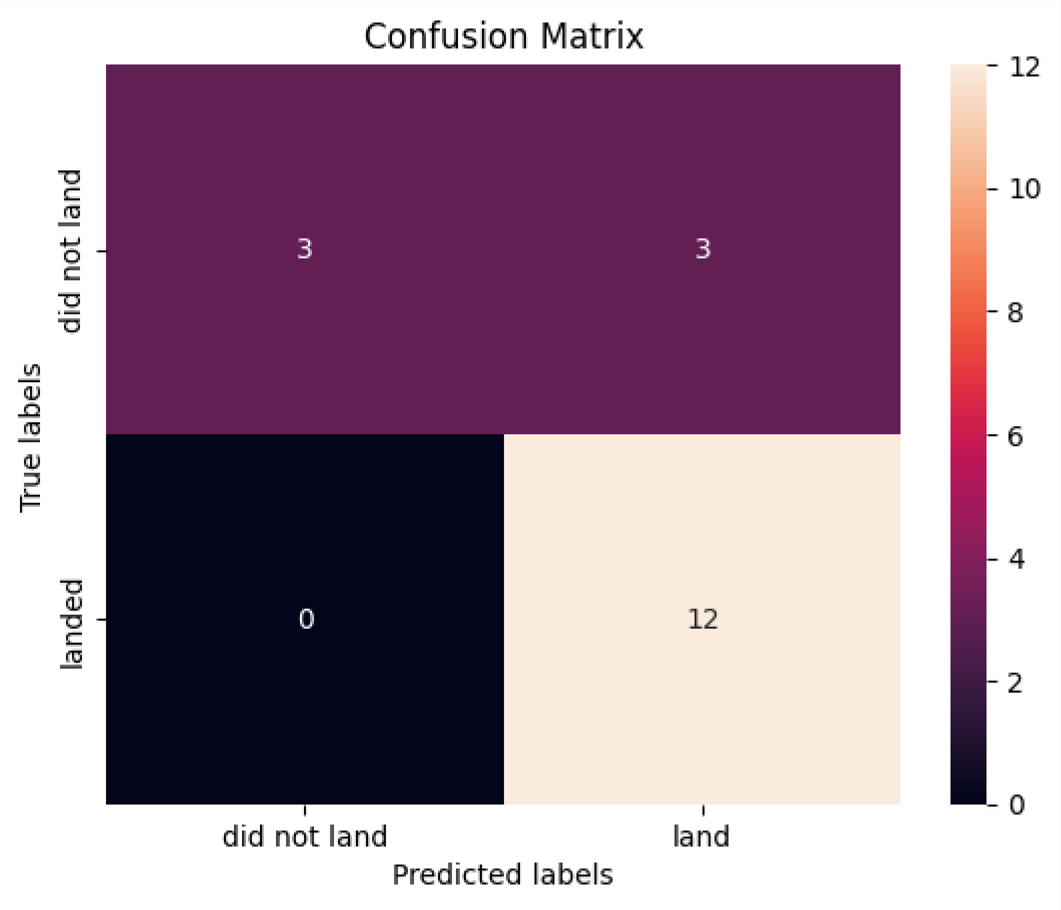
**1.Highest Cross-Validation Accuracy**: The Decision Tree model achieved the highest cross-validation accuracy at **0.8625**.

**2.Consistency in Test Accuracy**: All models achieved the same test accuracy of **0.8333,** indicating robustness in generalization across the dataset.

**3.Parameter Optimization**: Each model was fine-tuned using its respective best parameters to achieve these results.

**The Decision Tree model performed the best** in terms of cross-validation accuracy, making it a promising choice for this dataset. Further analysis could explore scenarios where these models can excel in deployment.

# Confusion Matrix




The confusion matrix provides insights into the performance of the best-performing model. Here's a breakdown of its elements:

•**True Positives (12)**: The model correctly predicted 12 instances where the rocket successfully landed.
•**True Negatives (3)**: The model correctly predicted 3 instances where the rocket did not land.
•**False Positives (3)**: The model predicted that 3 rockets would land, but they did not.
•**False Negatives (0)**: There were no instances where the model failed to predict a successful landing when it occurred.

**Model Performance Highlights**:
•The model's precision and recall for the "landed" category are strong, with no false negatives, indicating reliability in predicting successful landings.
•The false positives indicate some over-prediction of successful landings, which can be analyzed further for optimization.

# Conclusions on Predictive Analysis (Classification)

•**Model Evaluation**: Multiple classification models (Logistic Regression, SVM, Decision Tree, and KNN) were built and evaluated for predicting landing outcomes, demonstrating strong performance across all.
•**Decision Tree Superiority**: Among all models, the Decision Tree achieved the highest cross-validation accuracy (86.25%), indicating its effectiveness for this predictive task.
•**Confusion Matrix Insights**: The best-performing model showcased high precision and recall for predicting successful landings, highlighting its reliability in deployment scenarios.
•**Feature Significance**: Payload mass and booster version emerged as significant factors influencing mission outcomes, emphasizing the importance of including domain-specific features.
•**Applications**: This predictive framework can enhance decision-making for future SpaceX launches, optimizing resource allocation and mission planning.
•**Future Scope**: Further refinement using ensemble techniques or neural networks could further improve prediction accuracy, especially for rare cases or imbalanced datasets.

# Final Project Conclusion

This project analyzed SpaceX launch data to uncover key insights into success metrics, payload relationships, and booster performance.
Key findings include:

•**Launch Site Performance**: KSC LC-39A demonstrated the highest success rates, highlighting its operational efficiency.
•**Payload Insights**: Optimal payload ranges were identified, improving mission planning.
•**Booster Version Impact**: Newer booster versions significantly increased mission success rates.
•**Feature Importance**: Payload mass, booster version, and launch site were critical predictors of mission success.
•**Interactive Dashboards**: Enabled dynamic exploration, driving better decision-making and operational improvements.

The project exemplifies how data-driven insights can optimize operations and ensure SpaceX's leadership in space exploration.

# Thank you!