IBM Developer
SKILLS NETWORK

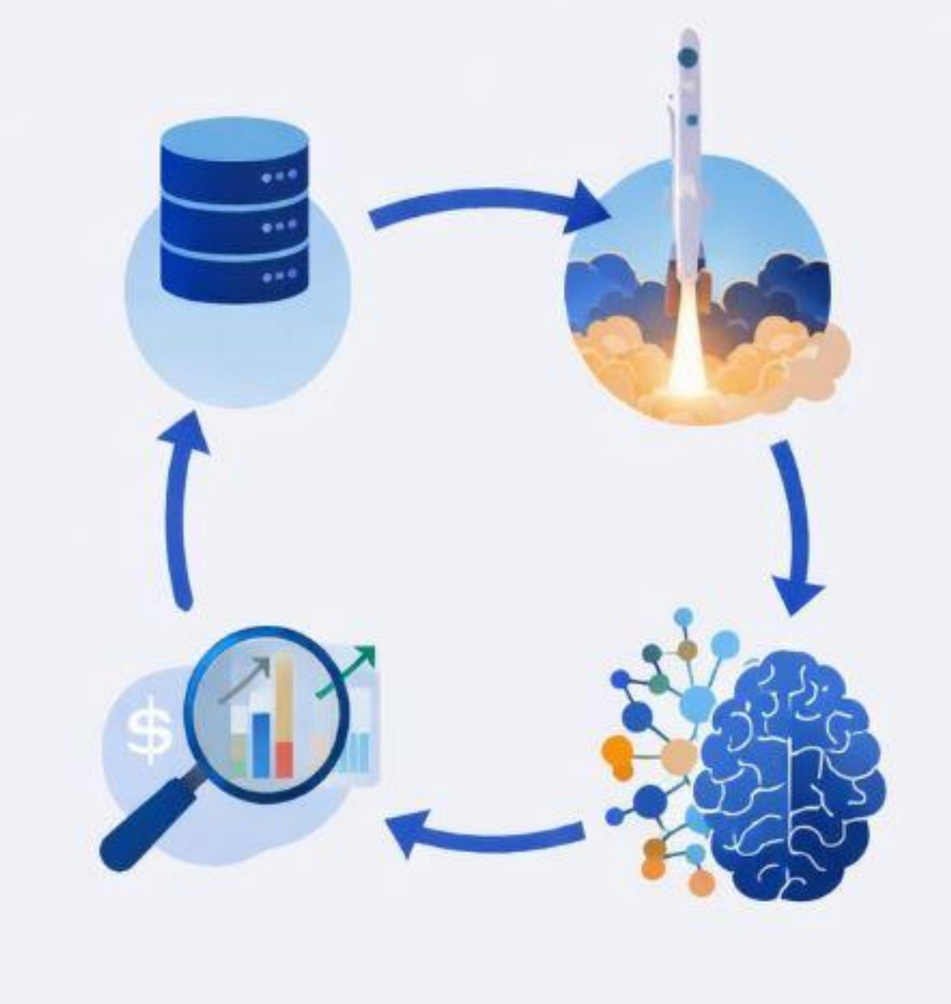
Winning Space Race with Data Science

Guido Antonio Acosta Laso
January 2026

Outline

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

Executive Summary



- This project analyzes SpaceX launch data using data science techniques.
- Data was collected, cleaned, and explored to understand launch outcomes.
- Exploratory data analysis was performed using visualization and SQL queries.
- Interactive maps and dashboards were created to identify patterns.
- Several machine learning models were built to predict landing success.
- The results show that data science can help understand and predict launch performance.

Introduction

- This project analyzes historical SpaceX launch data using data science techniques.
- The main goal is to understand which factors influence the success of rocket launches.
- We explore launch outcomes, payloads, launch sites, and rocket characteristics.
- Using data analysis and machine learning, we try to predict whether a launch will be successful or not.

Section 1

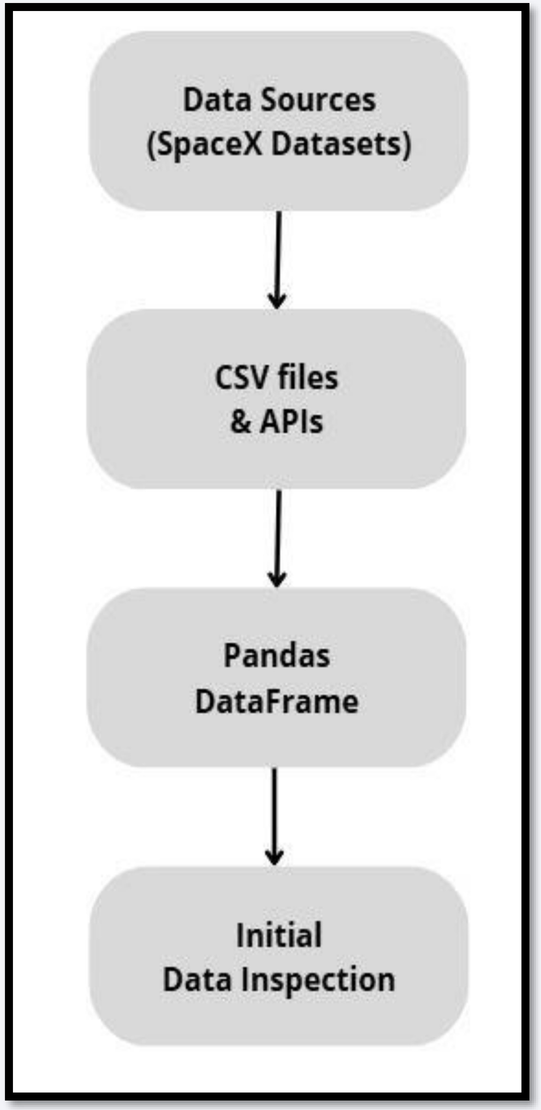
Methodology

Methodology

- Data collection methodology:
    - Data was collected from SpaceX launch records and public datasets.
- Perform data wrangling
    - The data was cleaned, transformed, and prepared for analysis.
- Perform exploratory data analysis (EDA) using visualization and SQL
    - Visual analysis and SQL queries were used to explore trends and patterns.
- Perform interactive visual analytics using Folium and Plotly Dash
    - Interactive maps and dashboards were created using Folium and Plotly Dash.
- Perform predictive analysis using classification models
    - Classification models were built, tuned, and evaluated to predict launch success

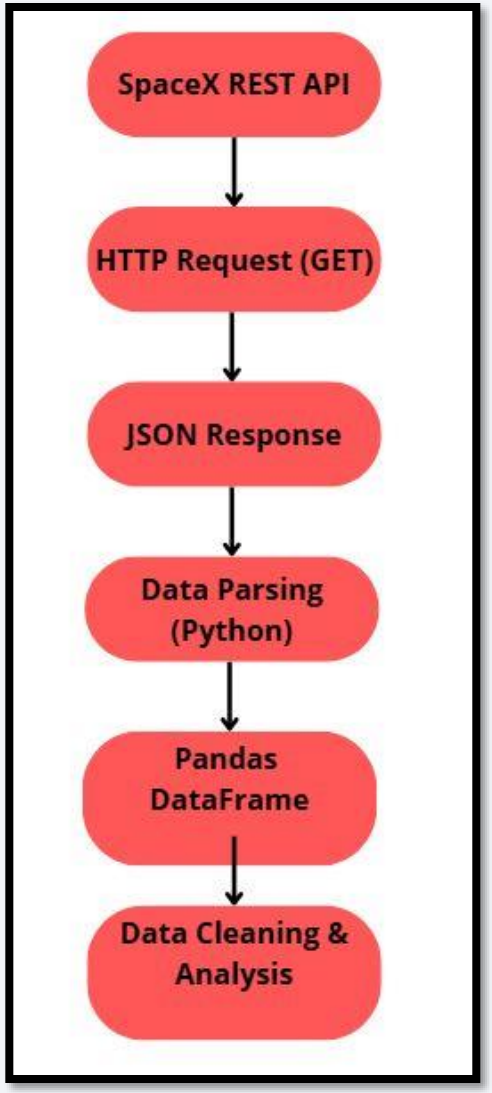
Data Collection

- Launch data was collected from SpaceX public datasets provided by IBM Skills Network.
- Multiple datasets were loaded using CSV files and APIs.
- The data includes:
    1. Launch dates and locations
    2. Rocket and payload information
    3. Launch outcomes (success or failure)
- The data collection process followed these steps:
    1. Data source identification
    2. Data loading into Pandas DataFrames
    3. Initial data inspection and validation.



Data Collection – SpaceX API

- SpaceX launch data was collected using the SpaceX REST API.
- API requests were sent using HTTP GET calls.
- The response data was returned in JSON format.
- JSON data was parsed and converted into Pandas DataFrames.
- The cleaned data was used for further analysis and modeling.

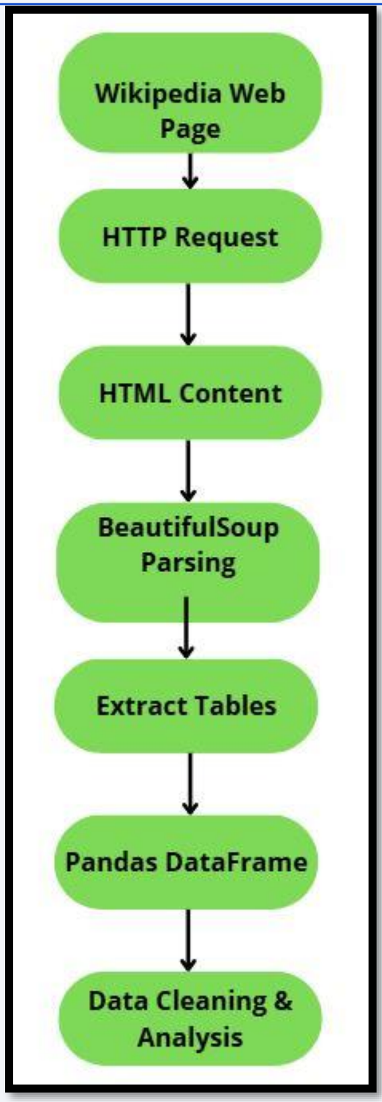


GitHub Link:

https://github.com/Guidoku84/IBM_Data_Science_Capstone_Project/blob/main/jupyter-labs-spacex-data-collection-api.ipynb

Data Collection - Scraping

- Falcon 9 launch records were collected using web scraping techniques.
- HTML pages were requested from Wikipedia.
- Launch tables were extracted from the webpage.
- Data was parsed using BeautifulSoup.
- The extracted data was converted into Pandas DataFrames.
- The cleaned dataset was used for further analysis.ng

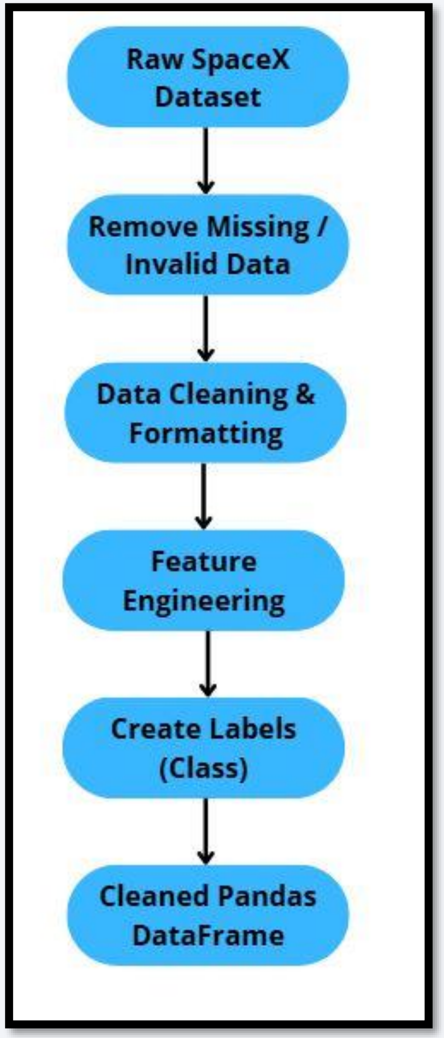


GitHub Link:

https://github.com/Guidoku84/IBM_Data_Science_Capstone_Project/blob/main/jupyter-labs-webscraping.ipynb

Data Wrangling

- Raw SpaceX launch data was cleaned and preprocessed.
- Missing values were handled and inconsistent records were removed.
- Categorical variables were encoded for analysis.
- New features and labels were created for machine learning models.
- The cleaned dataset was prepared for exploratory data analysis and modeling.



GitHub Link:

https://github.com/Guidoku84/IBM_Data_Science_Capstone_Project/blob/main/labs-jupyter-spacex-Data%20wrangling.ipynb

EDA with Data Visualization

- Multiple charts were created to explore relationships between launch outcomes and key variables.
- Bar charts were used to compare launch success rates across launch sites and orbit types.
- Scatter plots were used to analyze payload mass versus landing success.
- Line and bar charts were used to analyze trends over time.
- These visualizations helped identify patterns and factors influencing successful Falcon 9 landings.

GitHub Link:

https://github.com/Guidoku84/IBM_Data_Science_Capstone_Project/blob/main/edadataviz.ipynb

EDA with SQL

- SQL queries were used to explore SpaceX launch data stored in a SQLite database.
- Queries were executed to analyze launch success rates by launch site.
- SQL was used to examine relationships between payload mass and landing success.
- Aggregate functions (COUNT, AVG) were applied to summarize launch outcomes.
- Filtering and grouping techniques were used to identify patterns in mission success.

GitHub Link:

https://github.com/Guidoku84/IBM_Data_Science_Capstone_Project/blob/main/jupyter-labs-eda-sql-coursera_sqllite.ipynb

Build an Interactive Map with Folium

- An interactive map was created using Folium to visualize SpaceX launch sites.
- Markers were added to represent each launch site location.
- Circle markers were used to indicate launch success and failure outcomes.
- Distance lines and circles were added to show proximity to coastlines, railways, and cities.
- These map objects helped analyze geographical factors affecting launch success.

GitHub Link:

https://github.com/Guidoku84/IBM_Data_Science_Capstone_Project/blob/main/lab_jupyter_launch_site_location.ipynb

Build a Dashboard with Plotly Dash

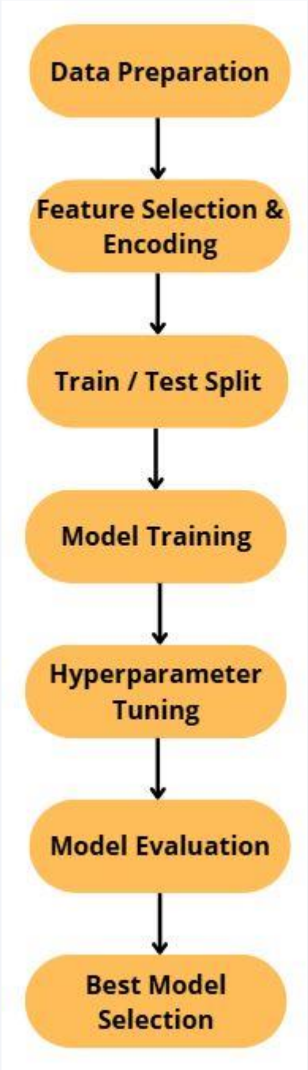
- An interactive dashboard was built using Plotly Dash.
- Pie charts were used to show launch success and failure rates.
- Scatter plots were used to analyze payload mass versus landing success.
- Dropdown filters allow users to select launch sites.
- Range sliders allow interactive analysis by payload mass.
- These interactions help explore factors affecting Falcon 9 landing success.

GitHub Link:

https://github.com/Guidoku84/IBM_Data_Science_Capstone_Project/blob/main/spacex-dash-app.py

Predictive Analysis (Classification)

- A supervised machine learning approach was used to predict the success of Falcon 9 landings.
- The dataset was prepared by selecting relevant features, encoding categorical variables, and normalizing numerical values.
- The data was split into training and testing sets to evaluate model performance.
- Multiple classification algorithms were trained and compared, including Logistic Regression, Support Vector Machine (SVM), Decision Tree, and K-Nearest Neighbors (KNN).
- GridSearchCV was applied to tune hyperparameters and improve model performance.
- Models were evaluated using accuracy, confusion matrix, and classification report metrics.
- The best-performing model was selected based on overall accuracy and generalization performance on the test dataset.



GitHub Link:

https://github.com/Guidoku84/IBM_Data_Science_Capstone_Project/blob/main/SpaceX_Machine%20Learning%20Prediction_Part_5.ipynb

Results

Exploratory Data Analysis Results

- Launch success rates vary across different launch sites.
- Higher payload mass generally reduces landing success.
- Certain orbit types show higher success rates.
- Falcon 9 landing success improved over time.

Results

Interactive Analytics Results

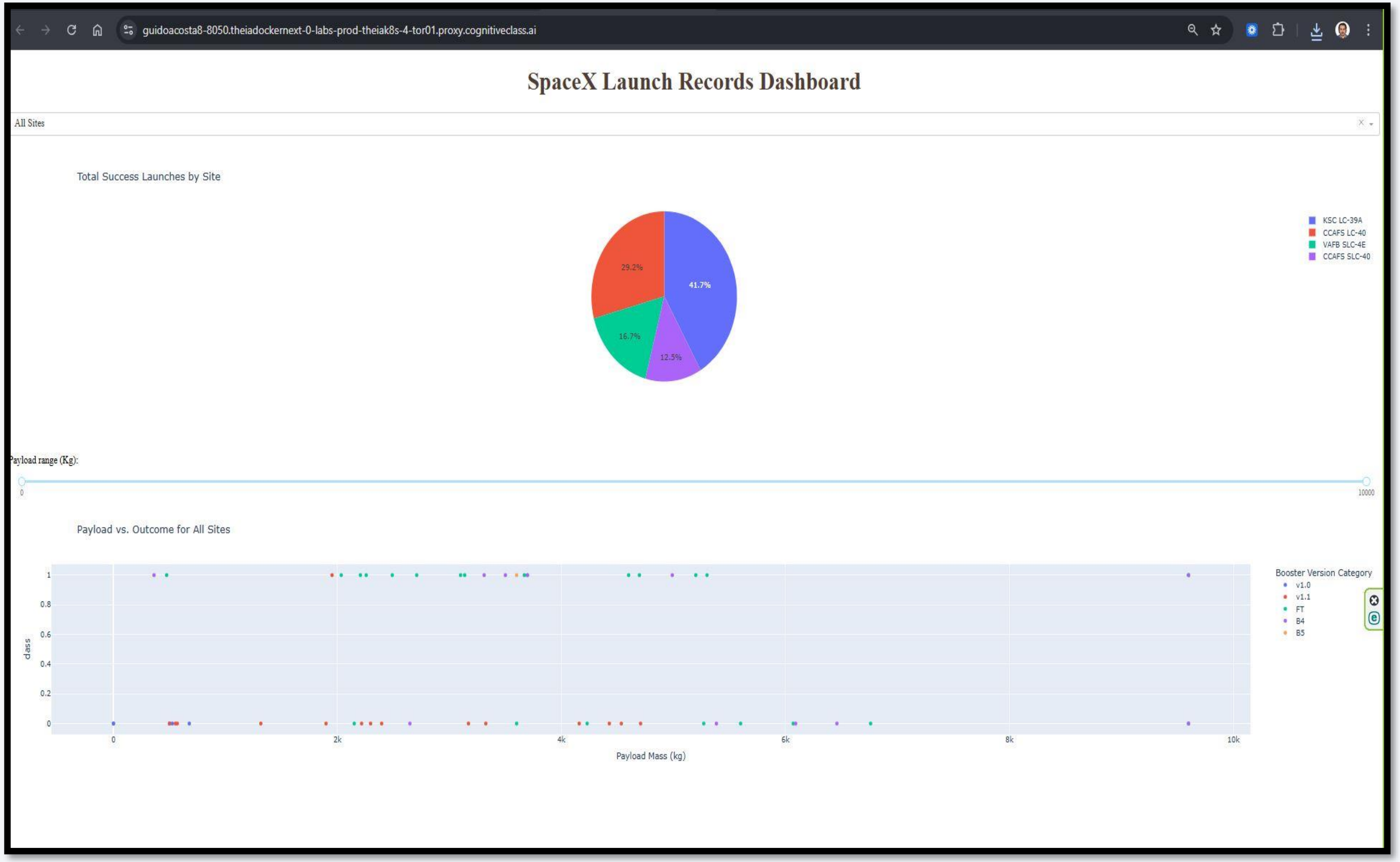
- An interactive dashboard was built to explore launch data.
- Users can filter results by launch site and payload range.
- Pie charts and scatter plots update dynamically based on user selection.
- The dashboard helps visualize success patterns clearly.

Results

Predictive Analysis Results

- Several classification models were trained and evaluated.
- Hyperparameter tuning improved model performance.
- The best-performing model achieved good accuracy.
- The model can predict Falcon 9 landing success reasonably well.

Results

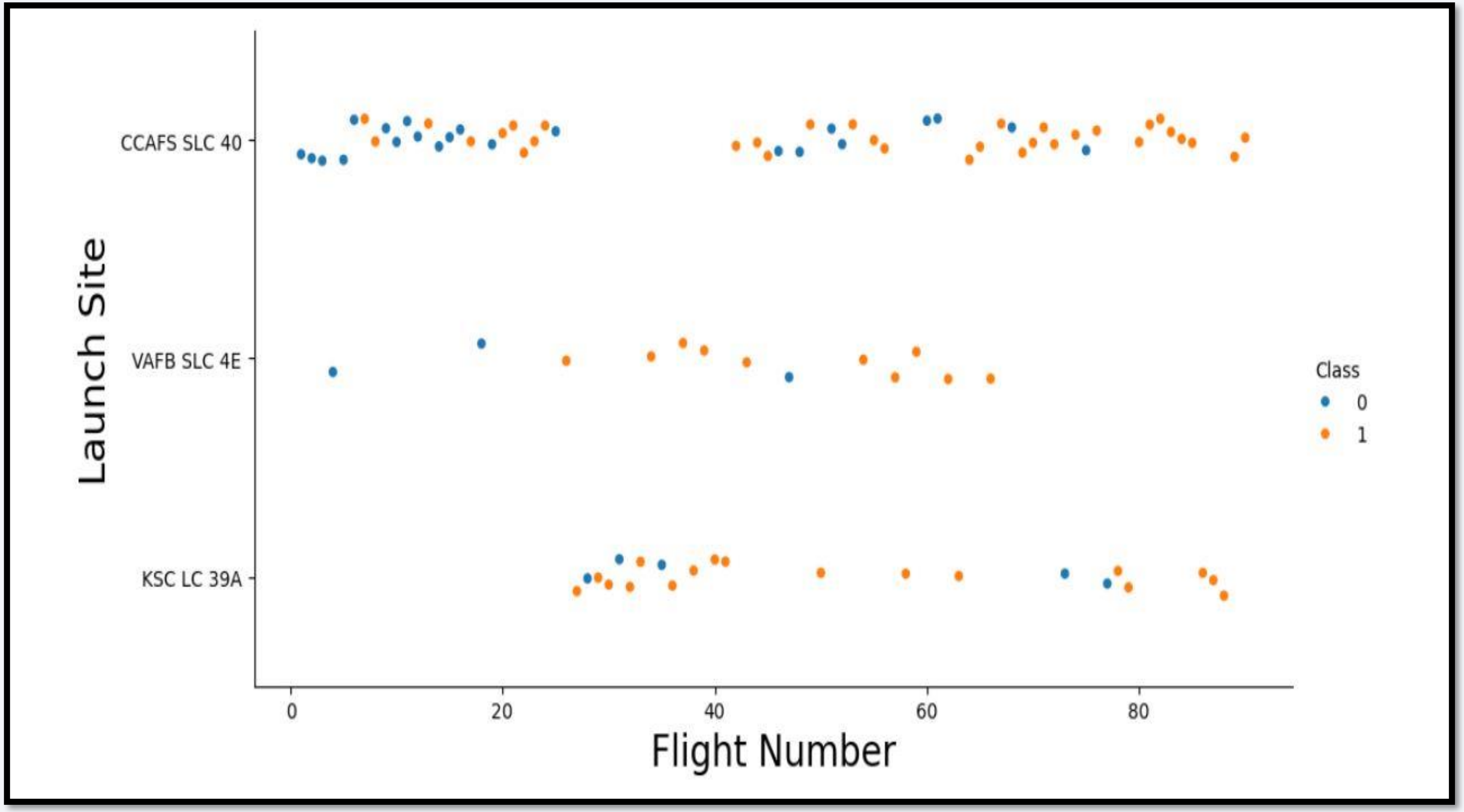


guidoacosta8-8050.theiadockernext-0-labs-prod-theiak8s-4-tor01.proxy.cognitiveclass.ai
SpaceX Launch Records Dashboard
All Sites
Total Success Launches by Site
41.7%
29.2%
16.7%
12.5%
KSC LC-39A
CCAFS LC-40
VAFB SLC-4E
CCAFS SLC-40
Payload range (Kg):
0
10000
Payload vs. Outcome for All Sites
class
Payload Mass (kg)
Booster Version Category
v1.0
v1.1
FT
B4
B5

Section 2

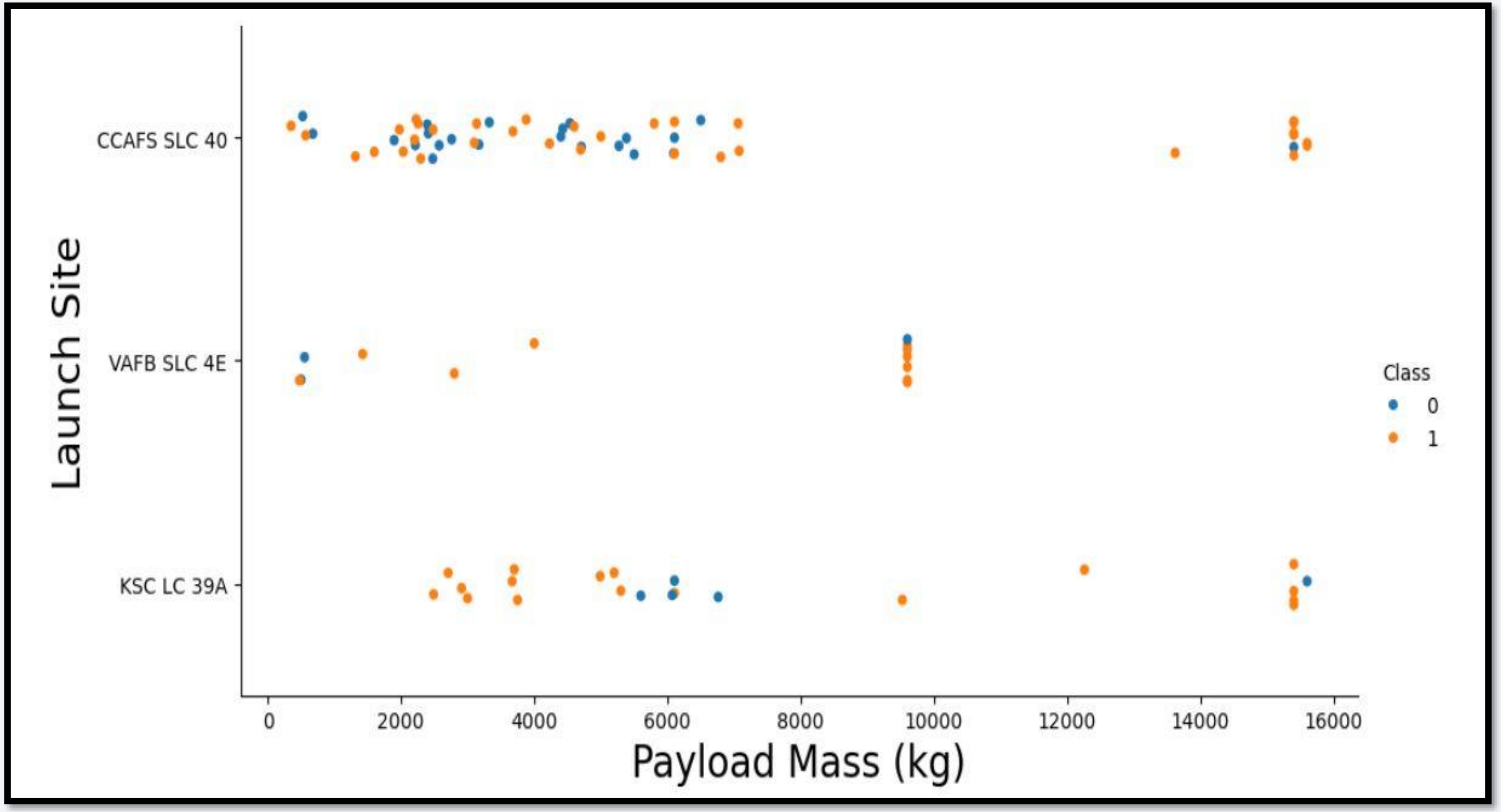
Insights drawn from EDA

Flight Number vs. Launch Site



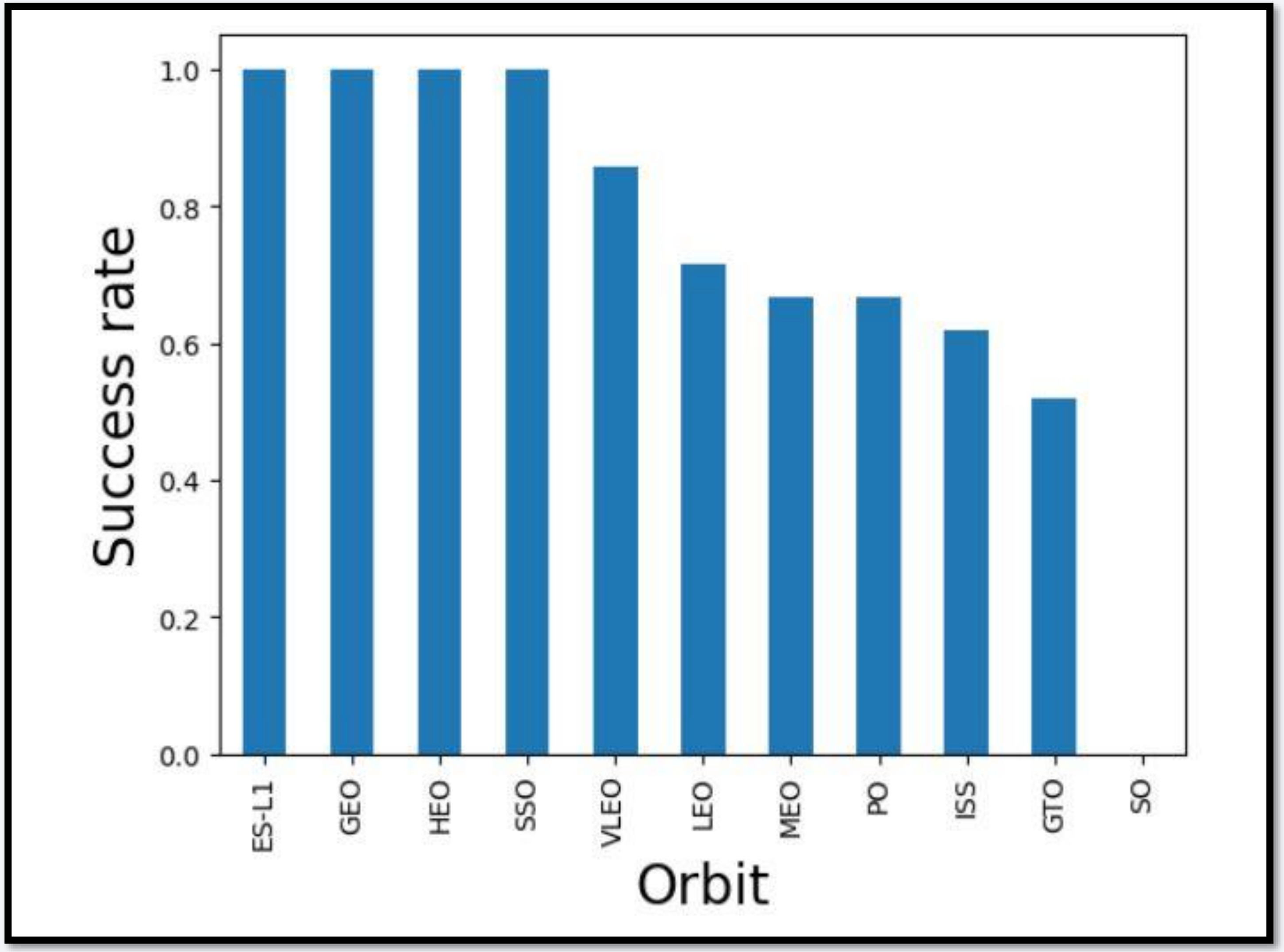
- Launch outcomes (success and failure) are observed across all launch sites, indicating that the launch site alone does not determine mission success.
- Flights are distributed over a wide range of flight numbers for all sites, showing consistent launch activity over time.
- Both successful and unsuccessful landings occur throughout the program, suggesting other factors influence landing success.

Payload vs. Launch Site



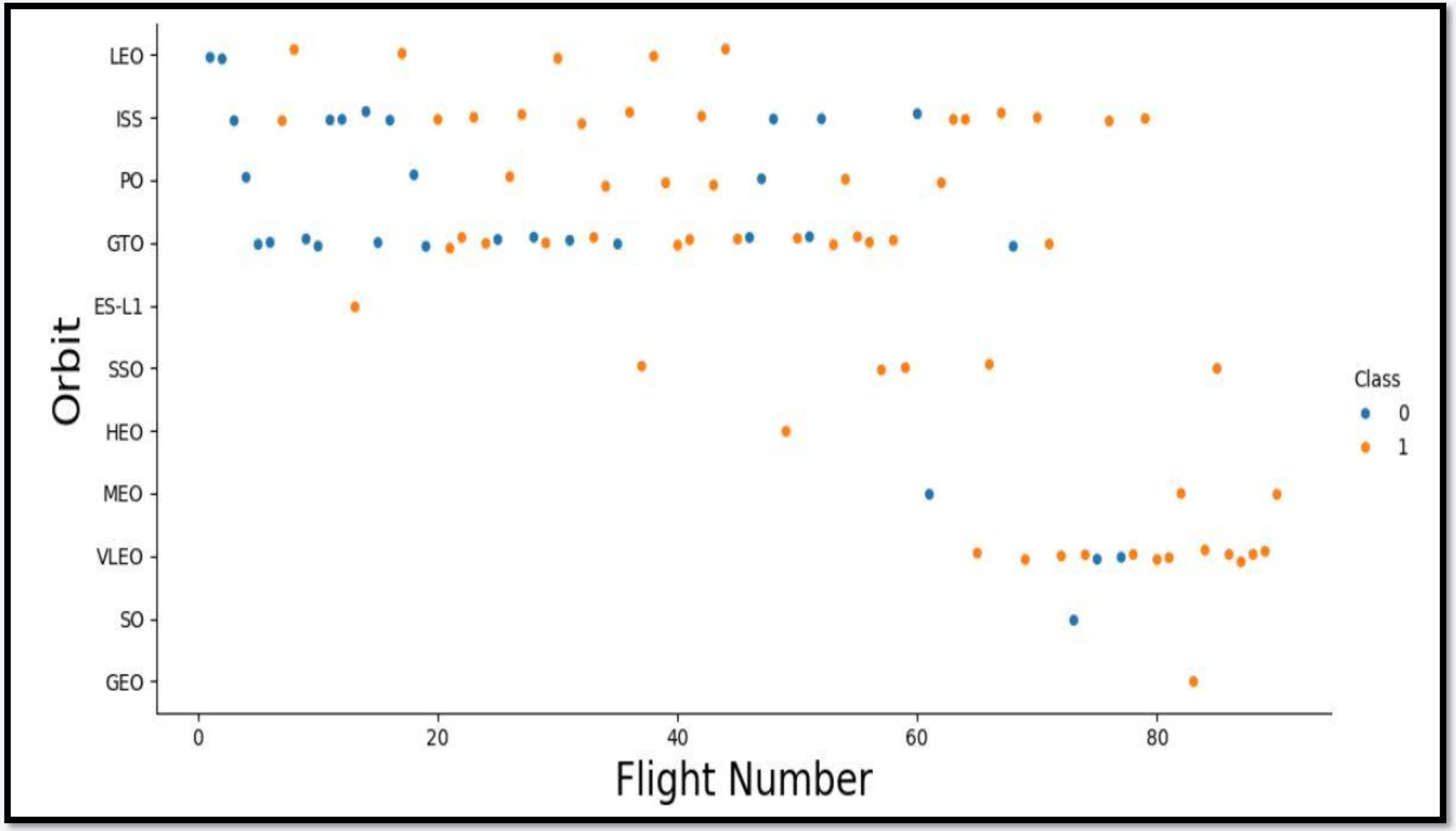
- Payload mass varies significantly across different launch sites, with KSC LC-39A and CCAFS SLC-40 handling a wider range of payload weights, including heavier missions.
- VAFB SLC-4E shows a more limited payload range, suggesting it is used for more specific mission profiles.
- Successful and unsuccessful landings occur across similar payload ranges at all sites, indicating that payload mass alone does not fully determine landing success.
- The distribution suggests that mission complexity and other operational factors play an important role alongside payload weight.

Success Rate vs. Orbit Type



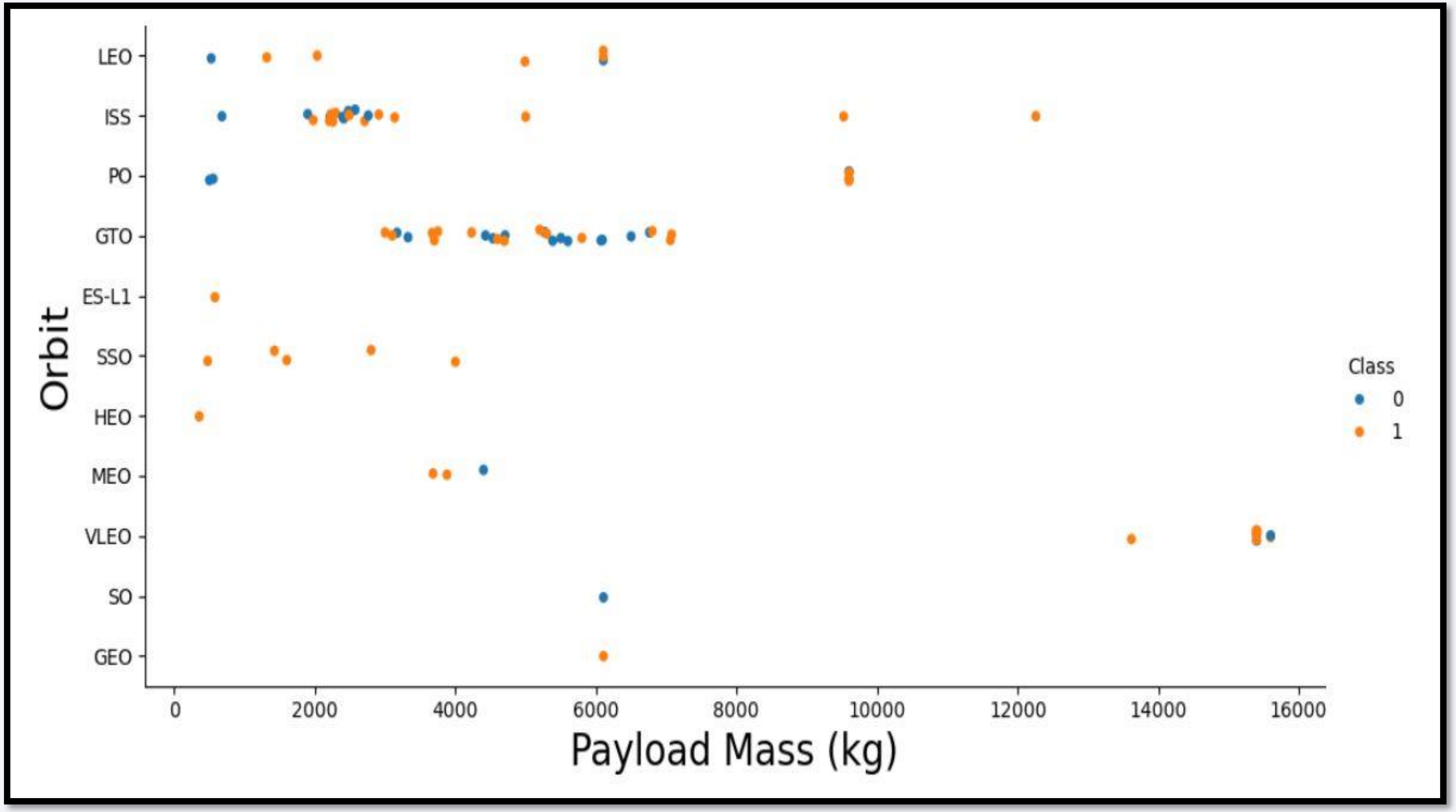
- High success rates for most orbit types: Missions to LEO, MEO, ES-L1, GEO, HEO, SSO, and VLEO show very high success rates, in several cases close to or equal to 100%, indicating strong reliability of Falcon 9 first-stage landings for these orbits.
- Lower success rate for GTO missions: The GTO (Geostationary Transfer Orbit) shows a noticeably lower success rate compared to other orbit types, suggesting that missions to GTO involve higher complexity and operational challenges.
- Orbit complexity influences landing success: The variation in success rates across orbit types indicates that mission profile and orbital requirements play a significant role in landing outcomes, not just launch frequency.

Flight Number vs. Orbit Type



- Increased Success Over Time: The success rate of Falcon 9 launches improves as the flight number increases, indicating that operational experience and iterative improvements contribute to better landing outcomes.
- Orbit-Specific Performance: Early missions to orbits such as GTO and ISS show mixed outcomes, while later flights demonstrate a higher proportion of successful landings.
- Learning Curve Effect: The visualization suggests that SpaceX benefits from a learning curve, achieving more consistent success across multiple orbit types in later missions.

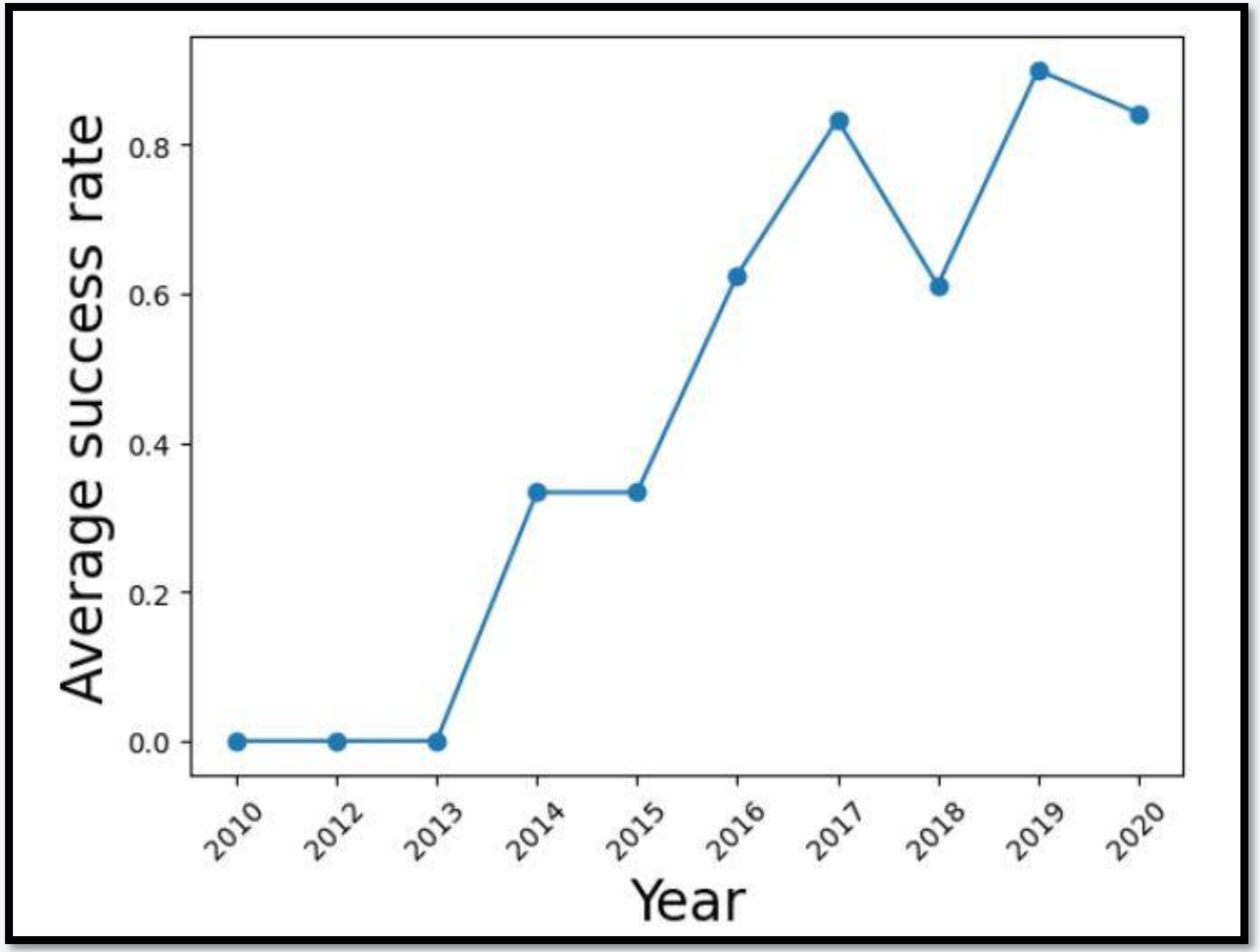
Payload vs. Orbit Type



- Higher success frequency at lower payloads: Successful landings are more frequent across most orbit types when the payload mass is below approximately 6000 kg, suggesting that lighter payloads are easier to recover successfully.
- Mixed outcomes for heavier payloads: Missions with higher payload masses (above 10,000 kg) show a mix of successful and failed landings across different orbit types, indicating increased technical difficulty for heavier missions.
- Orbit-dependent behavior: Certain orbit types such as LEO and ISS show a higher concentration of successful landings at moderate payload ranges, while more demanding orbits (e.g., GTO) present greater variability in outcomes.

Launch Success Yearly Trend

- Long-Term Improvement: The annual launch success rate shows a clear upward trend from 2013 onward, exceeding 80% by 2020.
- Temporary Variations: Although there is a slight dip around 2018, the overall trend remains positive.
- Operational Maturity: The increasing success rate reflects SpaceX's growing operational experience and improvements in launch and landing technologies.



All Launch Site Names

Task 1

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

[12]:
```
%%sql
PRAGMA table_info(SPACEXTABLE);
```

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

[12]:

cid	name	type	notnull	dflt_value	pk
0	Date	TEXT	0	None	0
1	Time (UTC)	TEXT	0	None	0
2	Booster_Version	TEXT	0	None	0
3	Launch_Site	TEXT	0	None	0
4	Payload	TEXT	0	None	0
5	PAYLOAD_MASS__KG_	INT	0	None	0
6	Orbit	TEXT	0	None	0
7	Customer	TEXT	0	None	0
8	Mission_Outcome	TEXT	0	None	0
9	Landing_Outcome	TEXT	0	None	0

[13]:
```
%%sql
SELECT DISTINCT "Launch_Site" FROM SPACEXTABLE;
```

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

[13]:

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

- Using an SQL query (SELECT DISTINCT Launch_Site FROM SPACEXTABLE), we identified all unique launch sites used in the SpaceX missions dataset.
- The results show that SpaceX has conducted launches from four main sites: CCAFS LC-40, KSC LC-39A, VAFB SLC-4E, and CCAFS SLC-40.
- This confirms that launch operations are concentrated in a limited number of strategic locations, which are later analyzed to understand their impact on mission success.

Launch Site Names Begin with 'CCA'

Task 2

Display 5 records where launch sites begin with the string 'CCA'

[14]:
```
%%sql
SELECT *
FROM SPACEXTABLE
WHERE "Launch_Site" LIKE 'CCA%'
LIMIT 5;
```

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

[14]:

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

- Using an SQL query with the LIKE 'CCA%' condition, five launch records were retrieved where the launch site name starts with “CCA”.
- The results confirm that CCAFS LC-40 is one of the primary launch sites used by SpaceX, especially during the early Falcon 9 missions.

Total Payload Mass

- Using an SQL aggregation query, the total payload mass delivered by Falcon 9 boosters for NASA Commercial Resupply Services (CRS) missions was calculated.
- The results show that Falcon 9 rockets have transported a total of 45,596 kg of payload for NASA CRS missions.

Task 3

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

[23]:
```
%load_ext sql
```

```
The sql extension is already loaded. To reload it, use:
  %reload_ext sql
```

[24]:
```
%sql sqlite:///my_data1.db
```

[26]:
```
%%sql
SELECT SUM("PAYLOAD_MASS__KG_") AS Total_Payload_Mass
FROM SPACEXTABLE
WHERE "Customer" = 'NASA (CRS)';
```

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

[26]:

Total_Payload_Mass
45596

Average Payload Mass by F9 v1.1

Task 4

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

[18]:
```
%%sql
SELECT AVG("PAYLOAD_MASS__KG_") AS Avg_Payload_Mass
FROM SPACEXTABLE
WHERE "Booster_Version" = 'F9 v1.1';
```

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

[18]:

Avg_Payload_Mass
2928.4

- The average payload mass carried by the Falcon 9 v1.1 booster is approximately 2928.4 kg. This value represents the typical payload capacity for missions using this booster version and reflects its operational performance during the analyzed launches.

First Successful Ground Landing Date

- Using an SQL query with the MIN(Date) function, we identified the earliest date when a successful ground pad landing was achieved.
- The first successful ground landing occurred on December 22, 2015, marking a major milestone in Falcon 9 reusability and SpaceX's landing technology.

```
[24]: %%sql
      SELECT MIN("Date") AS First_Successful_Ground_Landing
      FROM SPACEXTABLE
      WHERE "Landing_Outcome" = 'Success (ground pad)';

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

[24]:

First_Successful_Ground_Landing
2015-12-22

Successful Drone Ship Landing with Payload between 4000 and 6000

Task 6

List the names of the boosters which have success in drone ship and have payload mass greater than 4000 but less than 6000

[28]:
```
%%sql
SELECT DISTINCT "Booster_Version"
FROM SPACEXTABLE
WHERE "Landing_Outcome" = 'Success (drone ship)'
  AND "PAYLOAD_MASS__KG_" > 4000
  AND "PAYLOAD_MASS__KG_" < 6000;
```

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

[28]:

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

- This query identifies the Falcon 9 booster versions that successfully landed on a drone ship while carrying payloads between 4000 kg and 6000 kg.
- The results show that multiple booster versions were able to achieve successful drone ship landings within this payload range, demonstrating Falcon 9's capability to handle medium-heavy payloads with high landing reliability at sea.

Total Number of Successful and Failure Mission Outcomes

- This query calculates the total number of missions grouped by their outcome (success or failure).The results show that the majority of SpaceX missions were successful, indicating a high overall mission reliability and consistent performance over time.

Task 7

List the total number of successful and failure mission outcomes

[33]:

```
%%sql
SELECT "Mission_Outcome", COUNT(*) AS "Total"
FROM SPACEXTABLE WHERE "Mission_Outcome" IN ('Success', 'Failure')
GROUP BY "Mission_Outcome";
```

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

[33]:

Mission_Outcome	Total
Success	98

Boosters Carried Maximum Payload

Task 8

List all the booster_versions that have carried the maximum payload mass, using a subquery with a suitable aggregate function.

```
[31]: %%sql
SELECT DISTINCT "Booster_Version"
FROM SPACEXTABLE
WHERE "PAYLOAD_MASS_KG" = (
    SELECT MAX("PAYLOAD_MASS_KG")
    FROM SPACEXTABLE
);
```

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

[31]:

Booster_Version
F9 v1.0 B0003
F9 v1.0 B0004
F9 v1.0 B0005
F9 v1.0 B0006
F9 v1.0 B0007
F9 v1.1 B1003
F9 v1.1
F9 v1.1 B1011
F9 v1.1 B1010
F9 v1.1 B1012
F9 v1.1 B1013
F9 v1.1 B1014
F9 v1.1 B1015
F9 v1.1 B1016
F9 v1.1 B1018
F9 FT B1019
F9 v1.1 B1017
F9 FT B1020

- This query identifies the booster versions that carried the maximum payload mass recorded in the dataset.
- By using a subquery to find the highest payload value and filtering boosters that match it, we can determine which Falcon 9 boosters were capable of handling the heaviest missions.

2015 Launch Records

- The query identifies missions in 2015 that resulted in failed drone ship landings, along with their corresponding booster versions and launch sites.
- Results show that landing failures during this period occurred mainly during the early development phase of Falcon 9 reusability, highlighting the technological challenges SpaceX was still overcoming at that time.

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.

[32]:
```
%%sql
SELECT
    substr("Date",6,2) AS Month,
    "Booster_Version",
    "Launch_Site",
    "Landing_Outcome"
FROM SPACEXTABLE
WHERE substr("Date",1,4) = '2015'
  AND "Landing_Outcome" = 'Failure (drone ship)';
```

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

[32]:

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

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

Task 10

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.

[33]:
```sql
%%sql
SELECT
    "Landing_Outcome",
    COUNT(*) AS Total
FROM SPACEXTABLE
WHERE "Date" BETWEEN '2010-06-04' AND '2017-03-20'
GROUP BY "Landing_Outcome"
ORDER BY Total DESC;
```

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

[33]:

Landing_Outcome	Total
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

- This query ranks the different landing outcomes between June 4, 2010 and March 20, 2017 based on their frequency. The results show that No attempt was the most common outcome during this period, followed by Success (drone ship) and Failure (drone ship). This ranking helps to understand how landing performance evolved in the early years of SpaceX launches.

Section 3

Launch Sites Proximities Analysis

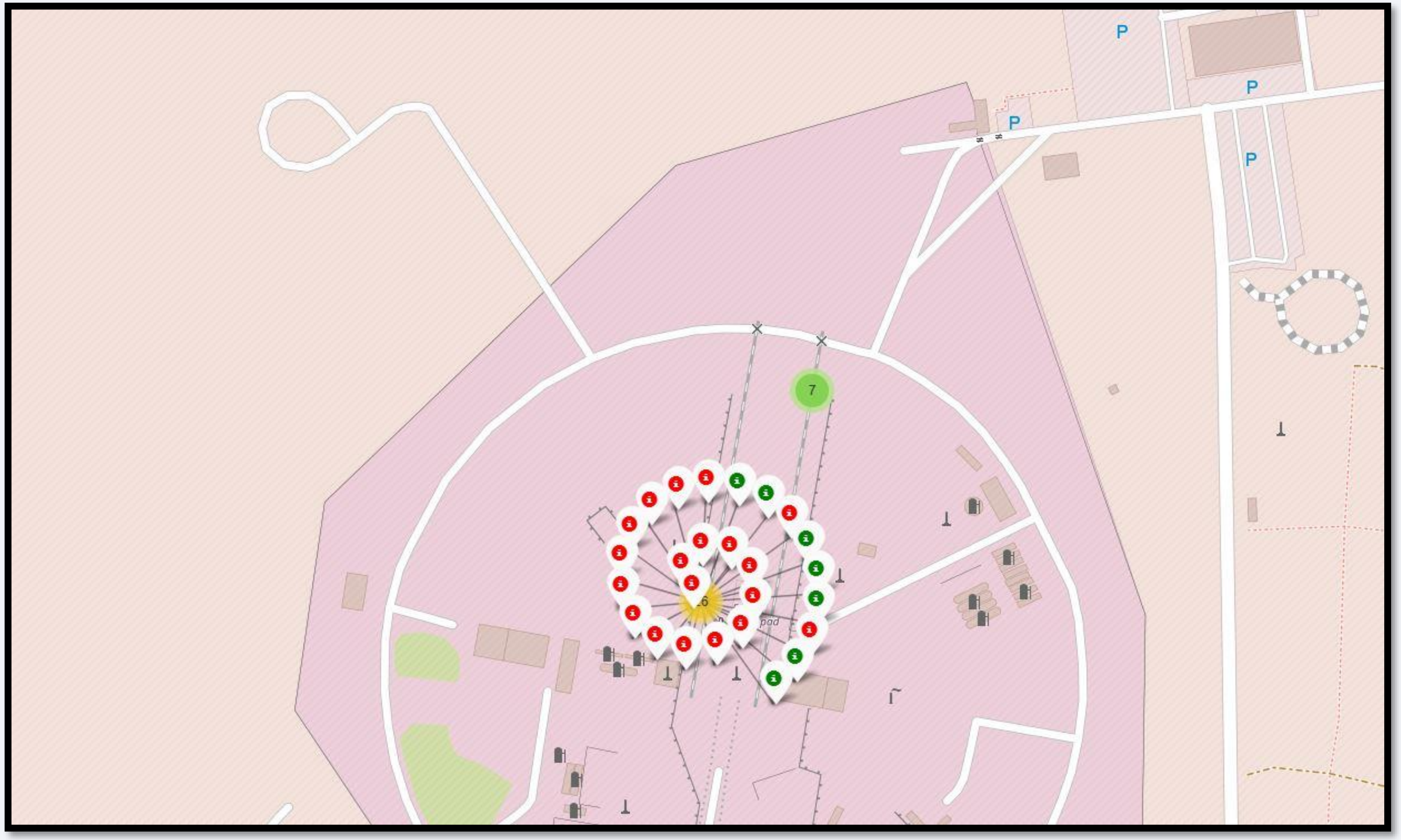
Global Map of SpaceX Launch Sites

Global Map of SpaceX Launch Sites

- The map displays the geographic locations of all SpaceX launch sites using an interactive Folium map.
- The launch sites are mainly concentrated in the United States, specifically in Florida (Cape Canaveral and Kennedy Space Center) and California (Vandenberg Space Force Base).
- All launch sites are located near the coastline, which allows safer launch trajectories over the ocean and reduces risk to populated areas.
- The map provides a clear spatial overview of launch site distribution, which is useful for further proximity and geographic analysis.

Launch Outcomes by Site

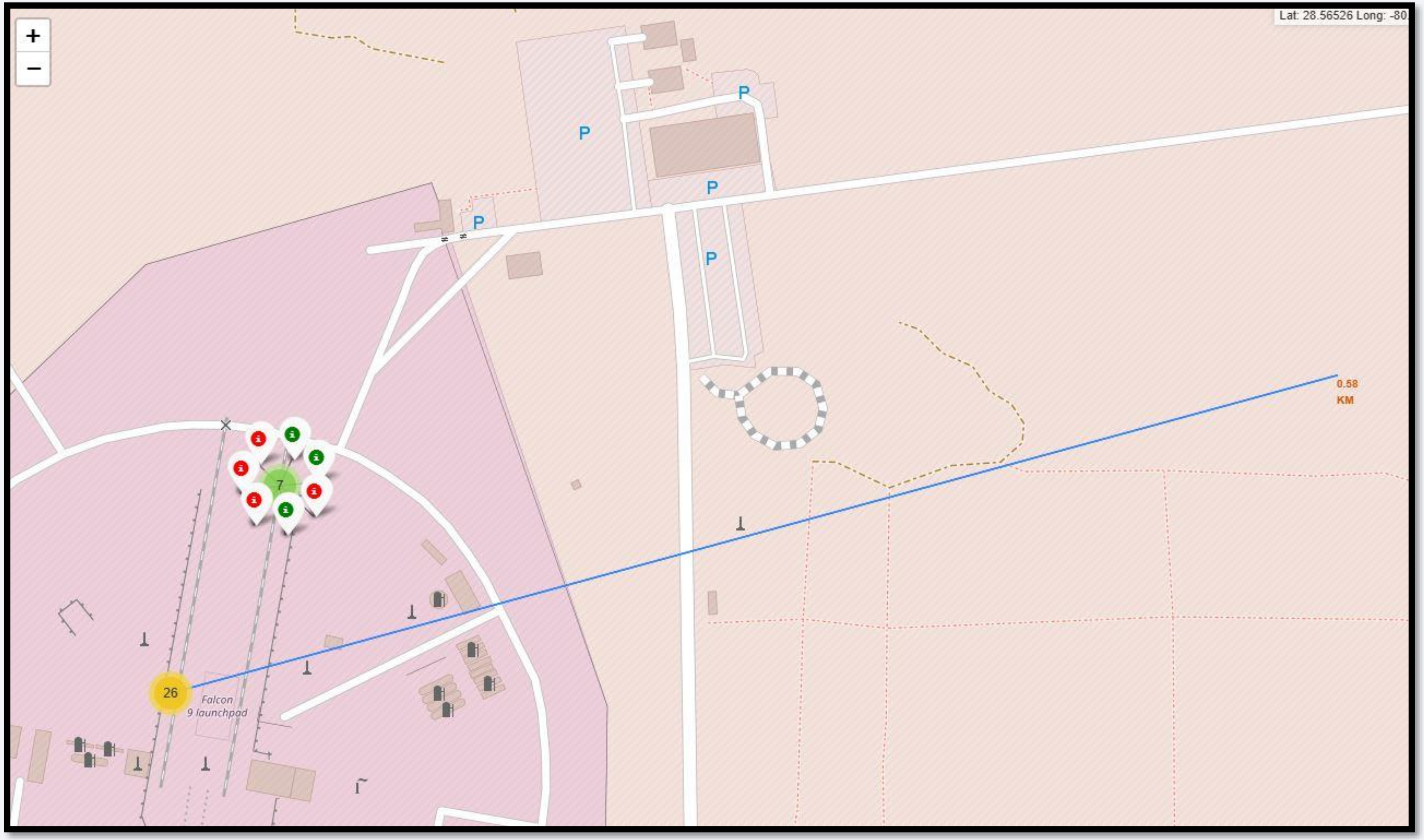
Launch Outcomes by Site



Launch Outcomes by Site

- This folium map visualizes SpaceX launch sites with color-coded markers representing launch outcomes. Green markers indicate successful landings, while red markers represent failed landings.
- The clustered markers help to clearly identify patterns in launch performance at each site. By observing the distribution of colors, it is evident that some launch sites, such as CCAFS LC-40 and KSC LC-39A, show a higher concentration of successful launches, indicating improved reliability over time.
- This visualization provides an intuitive way to compare success and failure rates across different launch sites and highlights how performance varies by location.

Launch Site Proximity Analysis (Coastline, Railway, Highway)



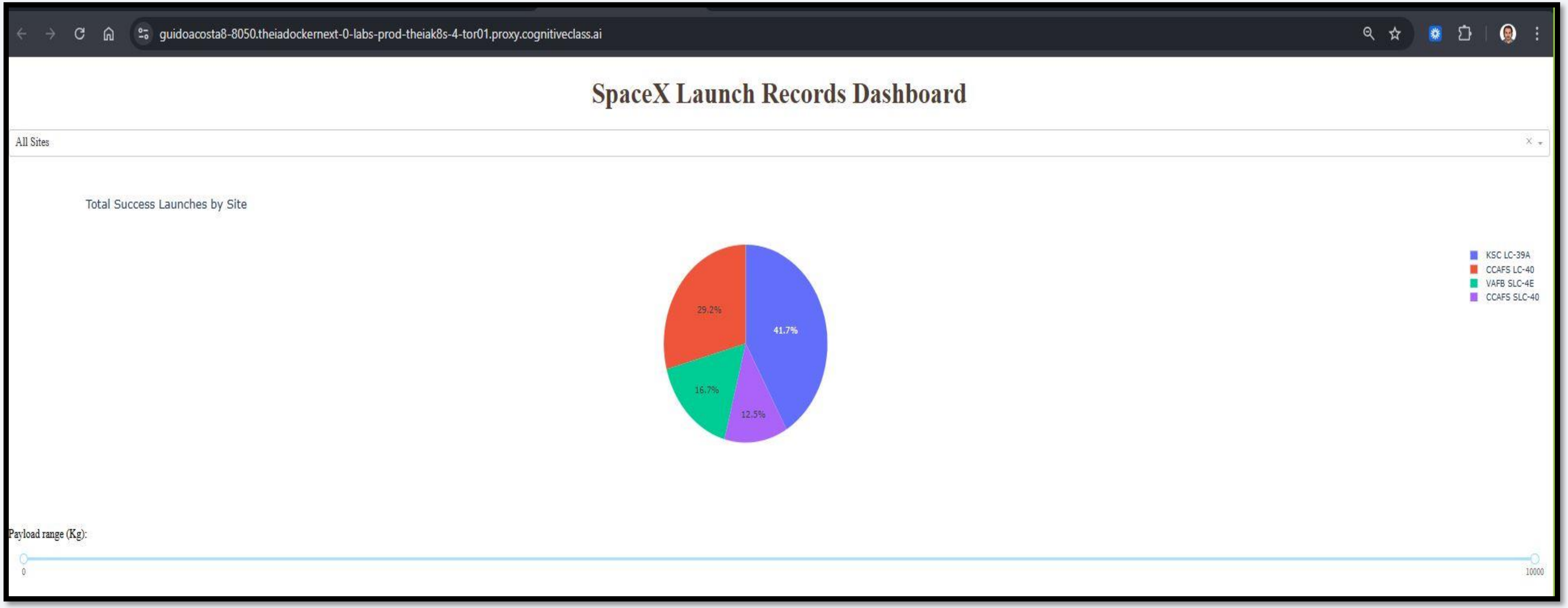
Launch Site Proximity Analysis (Coastline, Railway, Highway)

- This map focuses on a selected launch site and displays its proximity to key infrastructure such as the coastline, highways, and railways. Distance lines are drawn from the launch site to the nearest points of interest, with calculated distances shown directly on the map. The visualization highlights that launch sites are strategically located close to the coastline, which supports safer over-water flight paths and recovery operations while maintaining distance from densely populated areas.

Section 4

Build a Dashboard with Plotly Dash

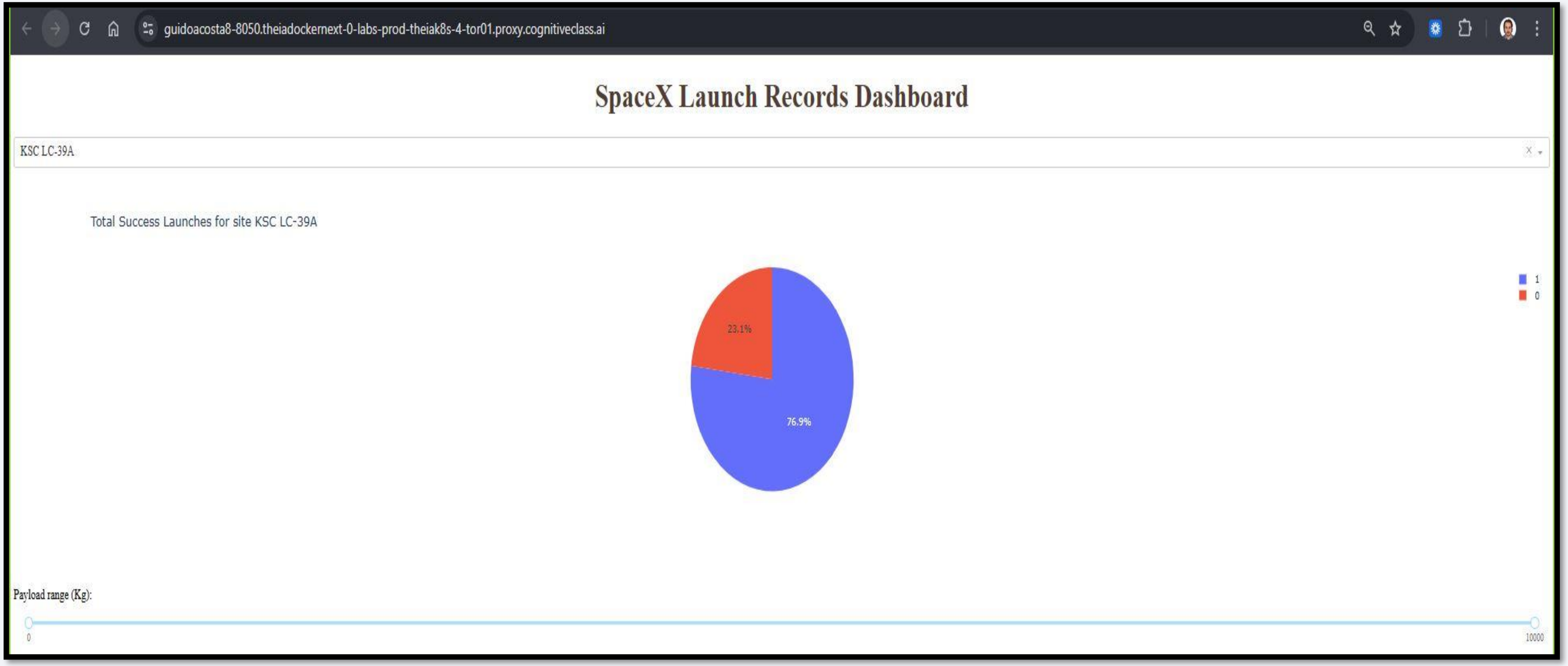
Launch Success Count for All Sites



Launch Success Count for All Sites

- This dashboard presents a pie chart showing the distribution of successful SpaceX launches across all launch sites. Each slice represents the proportion of successful launches contributed by a specific launch site.
- The visualization highlights that KSC LC-39A accounts for the highest percentage of successful launches (approximately 41.7%), indicating it is the most frequently and reliably used launch site. CCAFS LC-40 follows with around 29.2%, while VAFB SLC-4E and CCAFS SLC-40 contribute smaller shares.
- Overall, this chart provides a clear comparison of launch success distribution and demonstrates how launch activity and success are concentrated at specific sites.

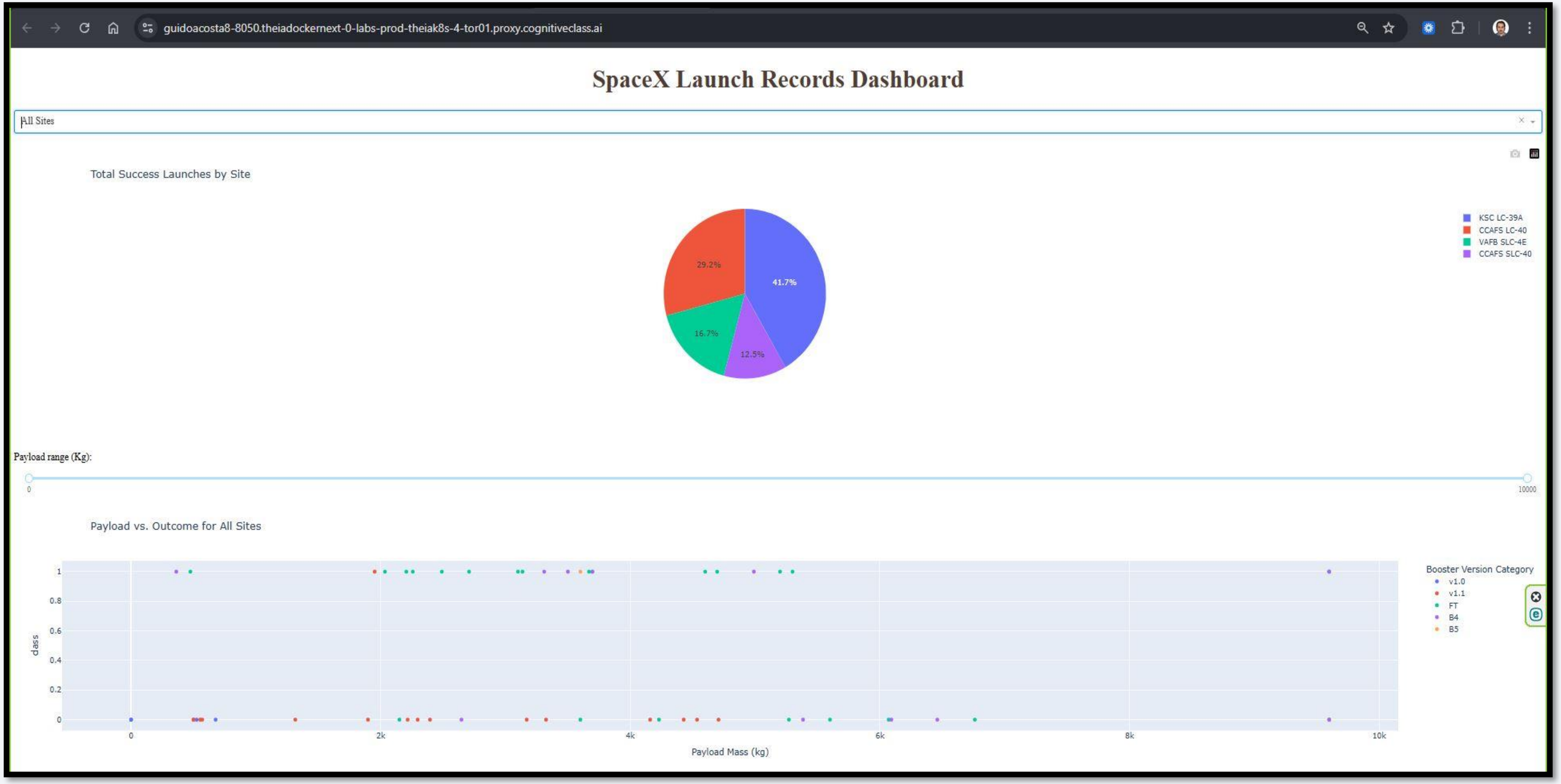
Launch Success Ratio for the Site with Highest Performance (KSC LC-39A)



Launch Success Ratio for the Site with Highest Performance (KSC LC-39A)

- This pie chart focuses on KSC LC-39A, the launch site with the highest launch success ratio. The chart shows that 76.9% of launches were successful, while 23.1% were unsuccessful.
- The high proportion of successful launches highlights the strong performance, reliability, and effectiveness of KSC LC-39A compared to other launch sites, making it a key contributor to SpaceX's overall mission success.

Payload Mass vs. Launch Outcome Across All Sites



Payload Mass vs. Launch Outcome Across All Sites

- This scatter plot shows the relationship between payload mass and launch outcome for all launch sites. Each point represents a launch, colored by booster version, with payload mass on the x-axis and launch outcome on the y-axis.
- By adjusting the payload range slider, patterns can be observed across different payload intervals. The visualization suggests that medium payload ranges tend to have higher success rates, while extremely high payloads show more variability.
- Additionally, certain booster versions, such as FT, appear more frequently and demonstrate higher success rates across a wide range of payload masses.

Section 5

Predictive Analysis (Classification)

Classification Accuracy

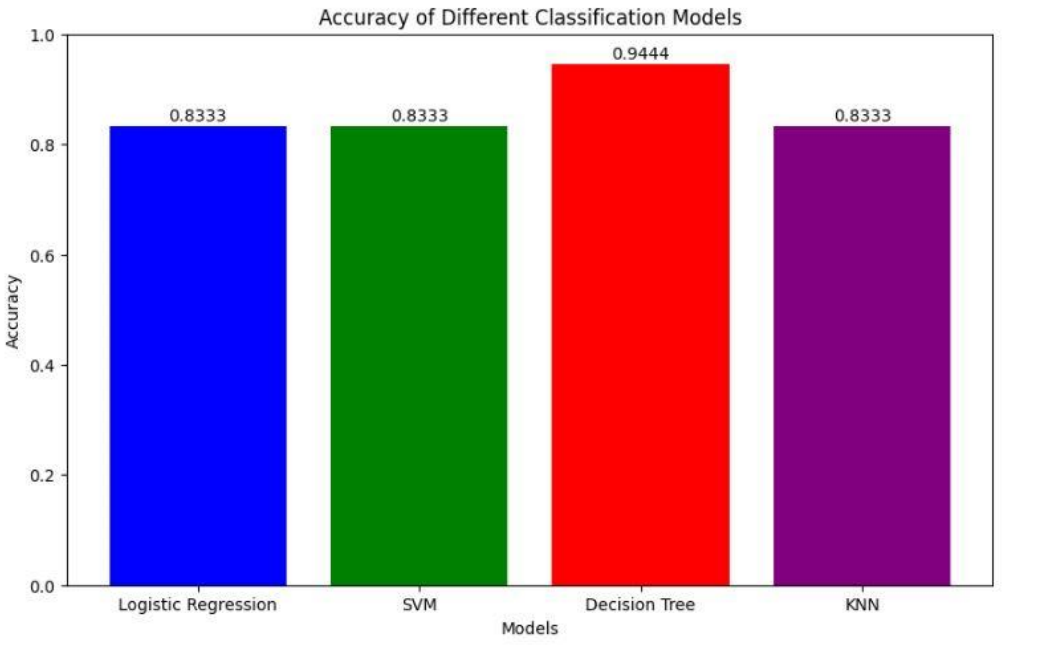
```python
# Model names and their corresponding accuracies
models = ['Logistic Regression', 'SVM', 'Decision Tree', 'KNN']
accuracies = [0.8333, 0.8333, 0.9444, 0.8333]

# Create a bar chart
plt.figure(figsize=(10, 6))
plt.bar(models, accuracies, color=['blue', 'green', 'red', 'purple'])

# Add title and labels
plt.title('Accuracy of Different Classification Models')
plt.xlabel('Models')
plt.ylabel('Accuracy')
plt.ylim(0, 1)  # Setting the y-axis range from 0 to 1

# Add accuracy values on top of the bars
for i, accuracy in enumerate(accuracies):
    plt.text(i, accuracy + 0.01, f'{accuracy:.4f}', ha='center')

# Show the plot
plt.show()
```



This bar chart compares the classification accuracy of four machine learning models used to predict the success of SpaceX rocket landings: Logistic Regression, Support Vector Machine (SVM), Decision Tree, and K-Nearest Neighbors (KNN).

The results show that the Decision Tree model achieved the highest accuracy, with a score of 0.9444, outperforming the other models. Logistic Regression, SVM, and KNN each obtained an accuracy of 0.8333.

This indicates that the Decision Tree model is better suited for capturing the patterns and relationships present in this dataset. Its higher accuracy suggests it can more effectively handle non-linear decision boundaries and interactions between features related to launch success.

Based on these results, the Decision Tree model was selected as the best-performing classification model for predicting SpaceX launch outcomes.

Confusion Matrix

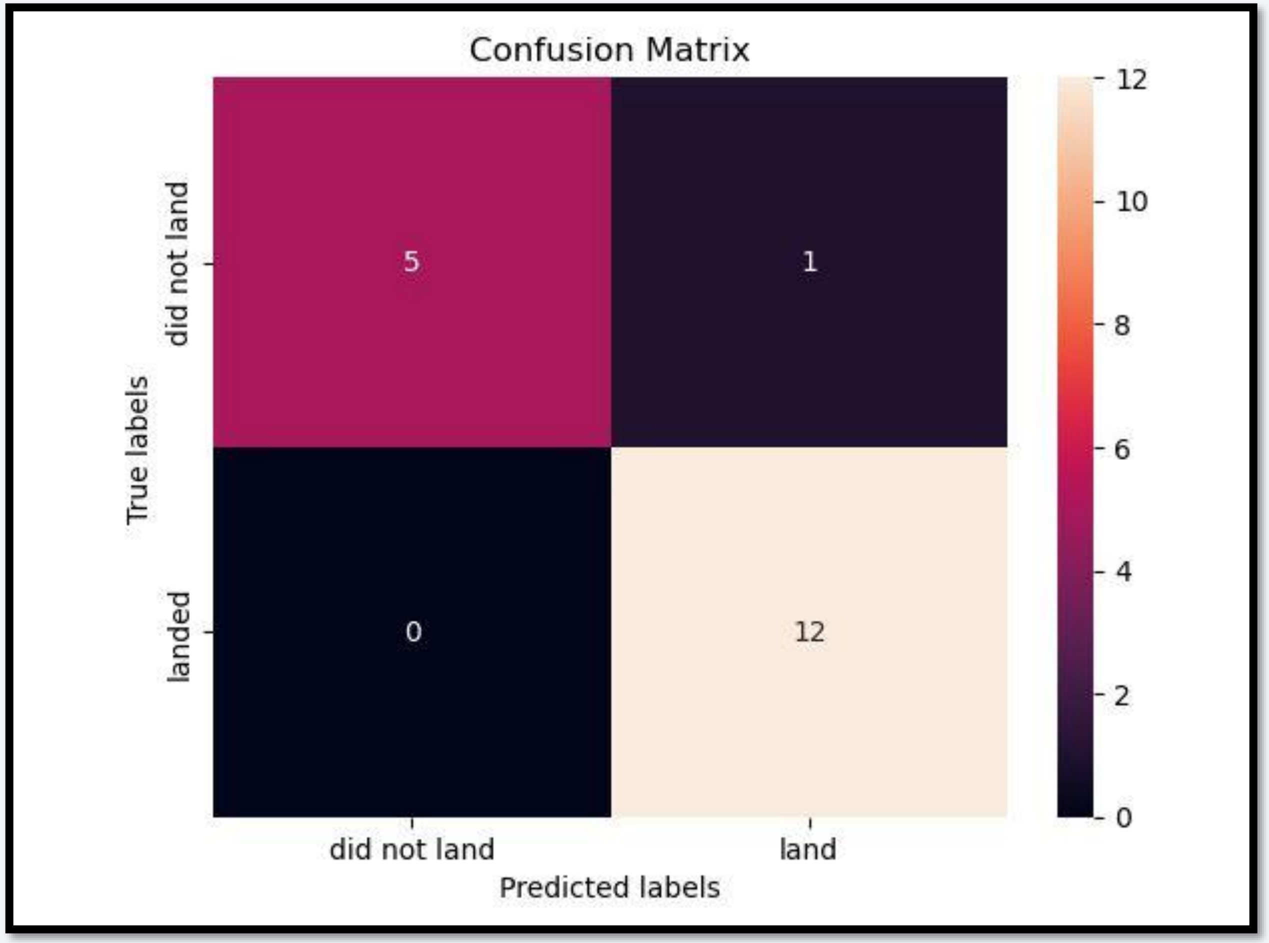
The confusion matrix corresponds to the Decision Tree model, which achieved the highest classification accuracy (94.44%) among all evaluated models.

The model correctly classified a high number of launches:

- 12 True Positives (successful landings correctly predicted)
- 5 True Negatives (failed landings correctly predicted)

There are no false negatives, meaning the model did not miss any actual successful landings. This is critical in aerospace applications, where failing to predict a successful landing could have high operational risks.

The model produced only 1 false positive, which is generally more manageable than false negatives, as it leads to conservative decision-making.



Conclusions

- **Point 1: Launch Site Performance**

The analysis showed that CCAFS LC-40 has the highest launch success rate, accounting for approximately 43.7% of all successful launches. This suggests that this site benefits from favorable operational conditions and well-established launch procedures.

- **Point 2: Booster Version Reliability**

The Falcon 9 FT booster version demonstrated a consistently high success rate across different payload masses, indicating strong reliability and robustness compared to other booster versions.

- **Point 3: Payload Mass Impact**

No clear relationship was observed between higher payload mass and lower launch success rates. This indicates that payload mass alone is not a decisive factor, and that operational factors such as launch site and booster version play a more significant role.

- **Point 4: Value of Data Visualization and Modeling**

The use of interactive visualizations (Folium and Plotly Dash) combined with predictive classification models provided valuable insights into SpaceX launch performance. These tools enhance decision-making by enabling a deeper understanding of geographical, technical, and operational factors affecting launch success.

Thank you!