IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

AVANI S JAYAN
12/07/2025

Outline

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

Executive Summary

Summary of methodologies:

- Collected and merged data from SpaceX API & Wikipedia using Python and web scraping.
- Cleaned and transformed data; stored in SQLite database.
- Conducted exploratory data analysis (EDA) with matplotlib, seaborn, and SQL queries.
- Built interactive visualizations:
  - Folium map to show launch sites.
  - Dash & Plotly dashboard to filter and explore data dynamically.
- Applied machine learning models (Decision Tree, SVM, Logistic Regression) to classify landing outcomes.

Summary of all results:

- Identified that launch site and booster version have strongest impact on success.
- Visualizations revealed heavier payloads slightly reduce landing success rates.
- SQL analysis showed CCAFS SLC-40 and KSC LC-40 as most active launch sites.
- Best ML model achieved ~**83.33% accuracy** predicting first stage landing success.
- Dashboard and interactive map allow users to explore launch data visually.

Introduction

Project background and context:

- SpaceX Falcon 9 is the world's first orbital class rocket capable of reflight.
- Successful landing of the first stage booster greatly reduces launch costs.
- Analyzing historical launch data helps understand what factors affect landing success.
- This project is part of the IBM Data Science Capstone to apply data science end-to-end.

Problems you want to find answers :

- What factors influence the success or failure of Falcon 9 first stage landing?
- Does payload mass, launch site, or booster version affect landing success?
- Can we predict landing success using historical data and machine learning?
- How can visual analytics help stakeholders explore and understand the data?

Section 1

Methodology

Methodology

Executive Summary

- Data collection methodology:
- Collected Falcon 9 launch data using the **SpaceX REST API**
  - Scraped **Wikipedia** pages to get additional data: payload mass, booster details
- Perform data wrangling
- Merged API and web-scraped data into a single dataset
- Cleaned missing values (e.g., `PayloadMass`, `LandingPad`)
- Standardized text fields (e.g., site names, booster versions)
- Stored final dataset into **SQLite database** for analysis
- Created derived features:
  - Number of previous flights
  - Orbit type
- Encoded categorical variables for machine learning

Perform exploratory data analysis (EDA) using visualization and SQL

- Visualizations:
- Scatter plots (payload vs landing success)
- Bar plots (launch site success counts)
- Correlation heatmap
- SQL queries:
- Average payload mass per site
- Launch success rate per site

Perform interactive visual analytics using Folium and Plotly Dash

- Built **Folium map**:
- Markers for launch sites
- Popups with site details and coordinates
- Created **Dash & Plotly dashboard**:
- Filters by site & payload range
- Pie charts: success vs failure
- Scatter plots: payload vs outcome

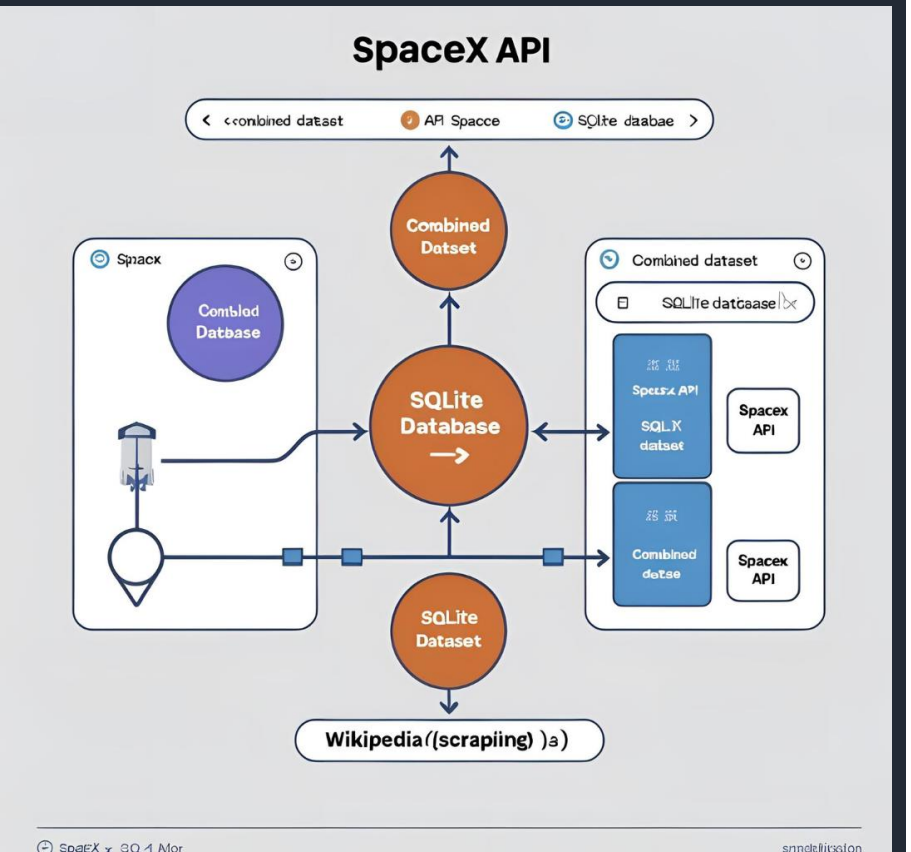
Perform predictive analysis using classification models

- Selected features:
- Payload mass
- Launch site
- Booster version
- Number of previous flights
- Applied classifiers:
- Decision Tree
- Support Vector Machine (SVM)
- Logistic Regression

- **How to build, tune, evaluate classification models:**

- Split dataset into training and testing sets
- Tuned hyperparameters using **GridSearchCV**
- Evaluated models based on:
  - Accuracy score
  - Confusion matrix
- Identified best performing model achieving **~83.33% accuracy**

Data Collection



- **Describe how data sets were collected:**
- Used the **SpaceX REST API** to gather structured launch data:
- Launch dates, sites, payload mass, booster versions, landing outcomes
- **Web scraping** from Wikipedia:
- Added missing booster details and payload data
- Combined both data sources to build a complete dataset for analysis

- **Key phrases:**
- REST API integration
- Web scraping with BeautifulSoup
- Data enrichment
- Merge & join datasets
- Structured storage

Data Collection – SpaceX API

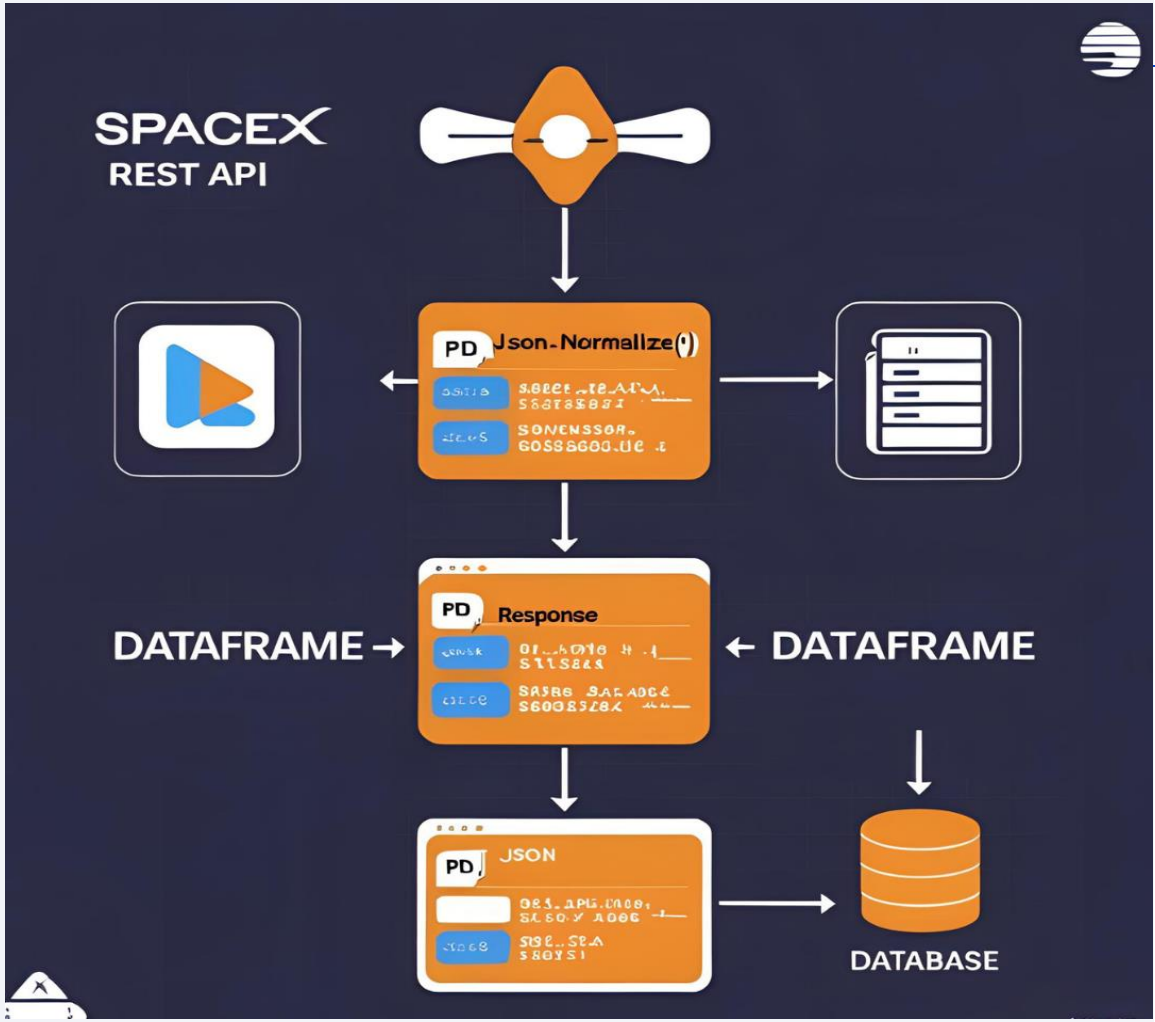
🚀 **Describe how data were collected:**

- Collected Falcon 9 launch data using SpaceX REST API
- Parsed JSON response with `pd.json_normalize()`
- Stored raw data into Pandas DataFrame
- Saved final data to SQLite database for further analysis

✅ **Key phrases:**

- REST API integration
- JSON normalization
- Structured storage

[github notebook]



Data Collection - Scraping

Used web scraping to enrich data:

- Scraped Wikipedia Falcon 9 page for booster details & payload mass
- Used Python BeautifulSoup to parse HTML tables
- Converted to Pandas DataFrame
- Merged with API data for enriched dataset

Key phrases:

- Web scraping
- BeautifulSoup
- Data enrichment
- Merge datasets

[github notebook]

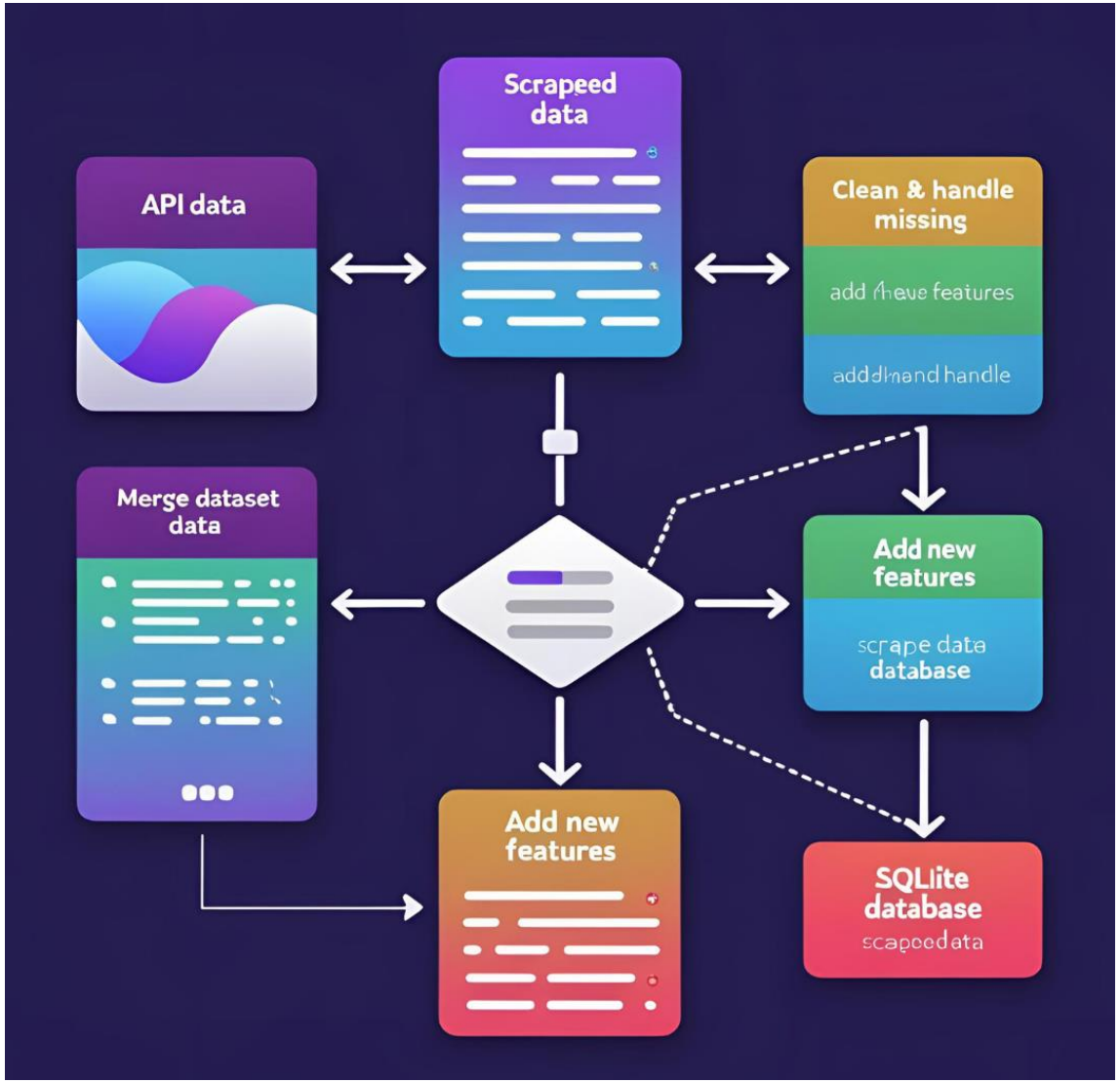
Data Wrangling

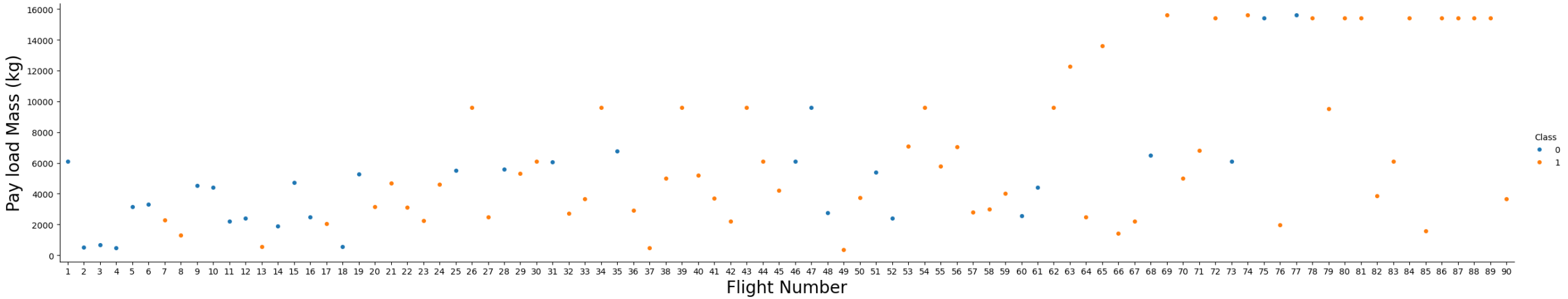
Describe how data were processed:

- Merged API + scraped data into one dataset (Python)
- Handled missing values (PayloadMass, LandingPad)
- Standardized site names & booster version formats
- Created new features: number of previous flights, orbit type
- Saved cleaned dataset to SQLite database

Key phrases:

- Merge datasets
- Handle missing data
- Feature engineering
- SQLite storage





EDA with Data Visualization

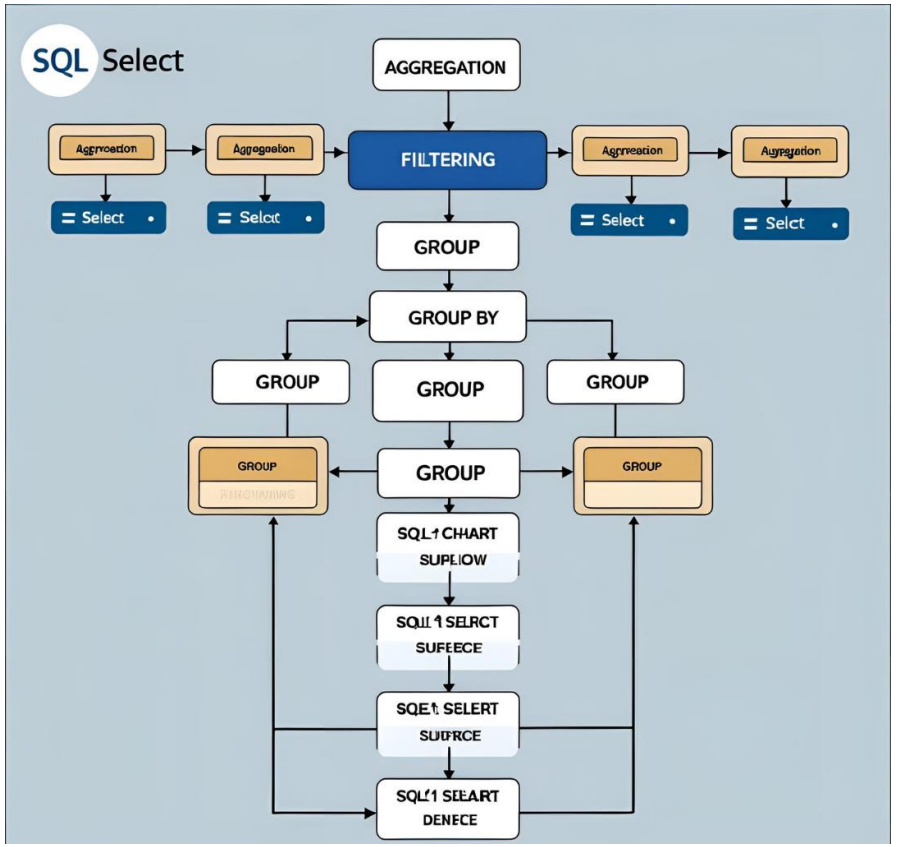
- **Summarize what charts were plotted & why:**
- Scatter plots: Payload mass vs. landing success → see relationship
- Bar plots: success counts by launch site
- Correlation heatmap: see feature correlations

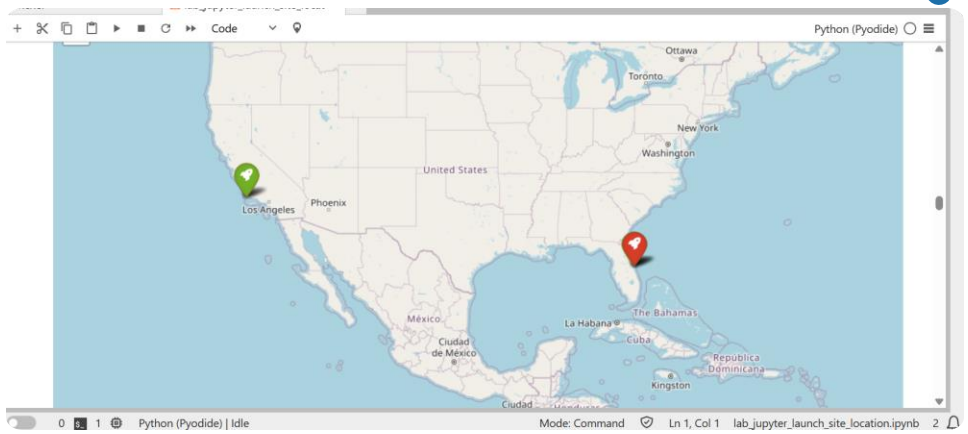
- **Key phrases:**
- Matplotlib
- Seaborn
- Feature distribution
- Visual exploration

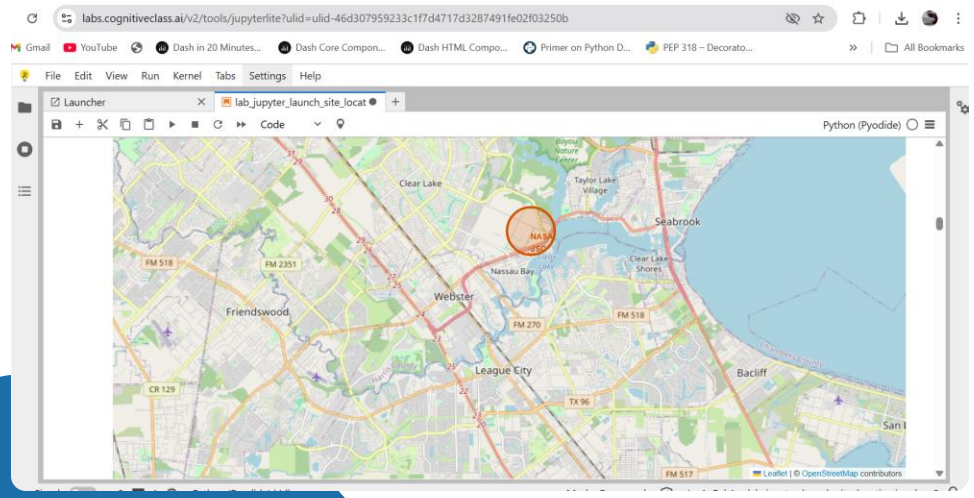
- EDA VISUALISATION

EDA with SQL

- **Summarize SQL queries performed:**
- Found unique launch site names
- Calculated total & average payloads
- Counted success/failure outcomes
- Listed boosters with highest payload
-
- **Key phrases:**
- SQL SELECT
- Aggregation
- Filtering
- Group by
- [EDA Notebook]





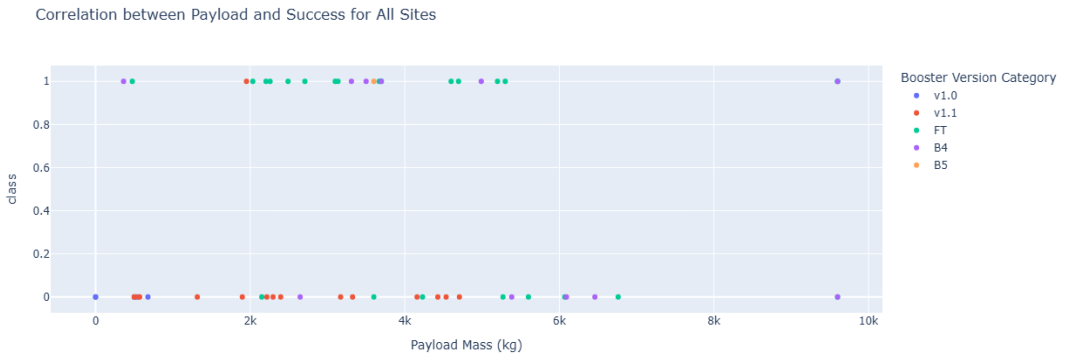


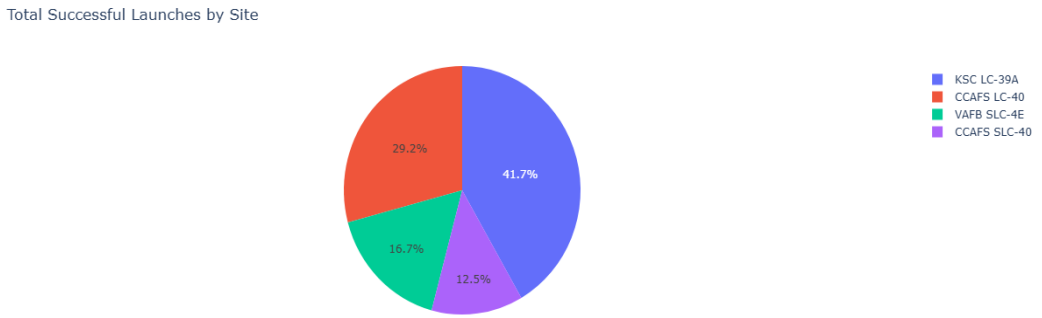
Build an Interactive Map with Folium

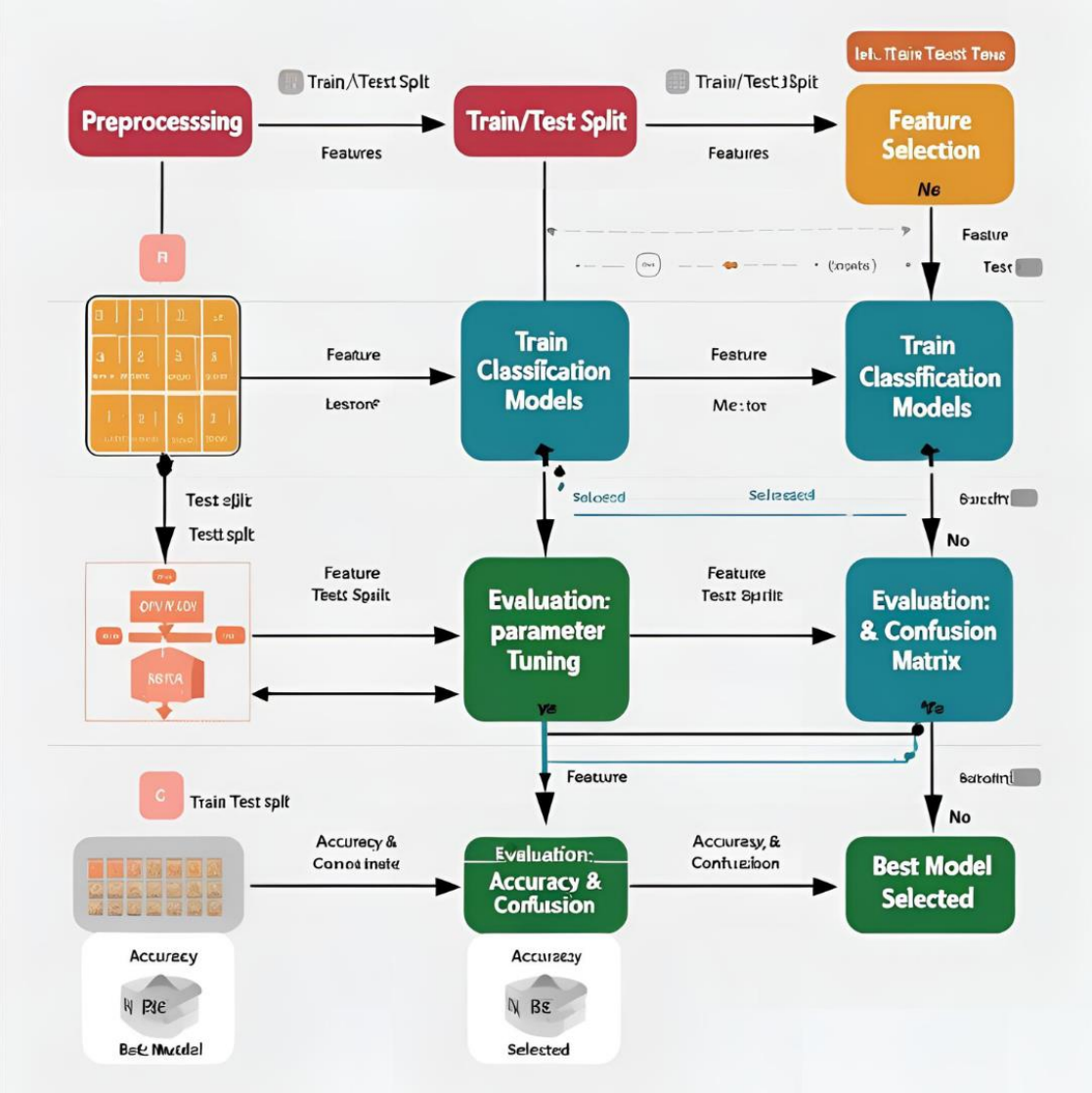
- **Map objects added:**
- **Markers:** show exact launch site locations
- **Popups:** display site name & coordinates
- **Circles:** highlight launch site areas visually
- **Lines (if added):** show distance to nearby features (e.g., coastline, road, railway)
- **Why we added these objects:**
- To **visualize geographic distribution** of Falcon 9 launches
- Help understand **site proximity to infrastructure**
- Make data **interactive & easier to explore**
- [launchsite_location]

Build a Dashboard with Plotly Dash

- **Plots & interactions added:**
- **Pie charts:** show launch success vs failure counts by site
- **Scatter plots:** visualize relationship between payload mass & landing outcome
- **Range slider:** filter payload mass range
- **Dropdown:** select specific launch site to focus analysis
- **Why we added these plots & interactions:**
- Make data **interactive and user-driven**
- Help users explore **how payload mass & site affect landing success**
- Allow quick comparison of success rates by site





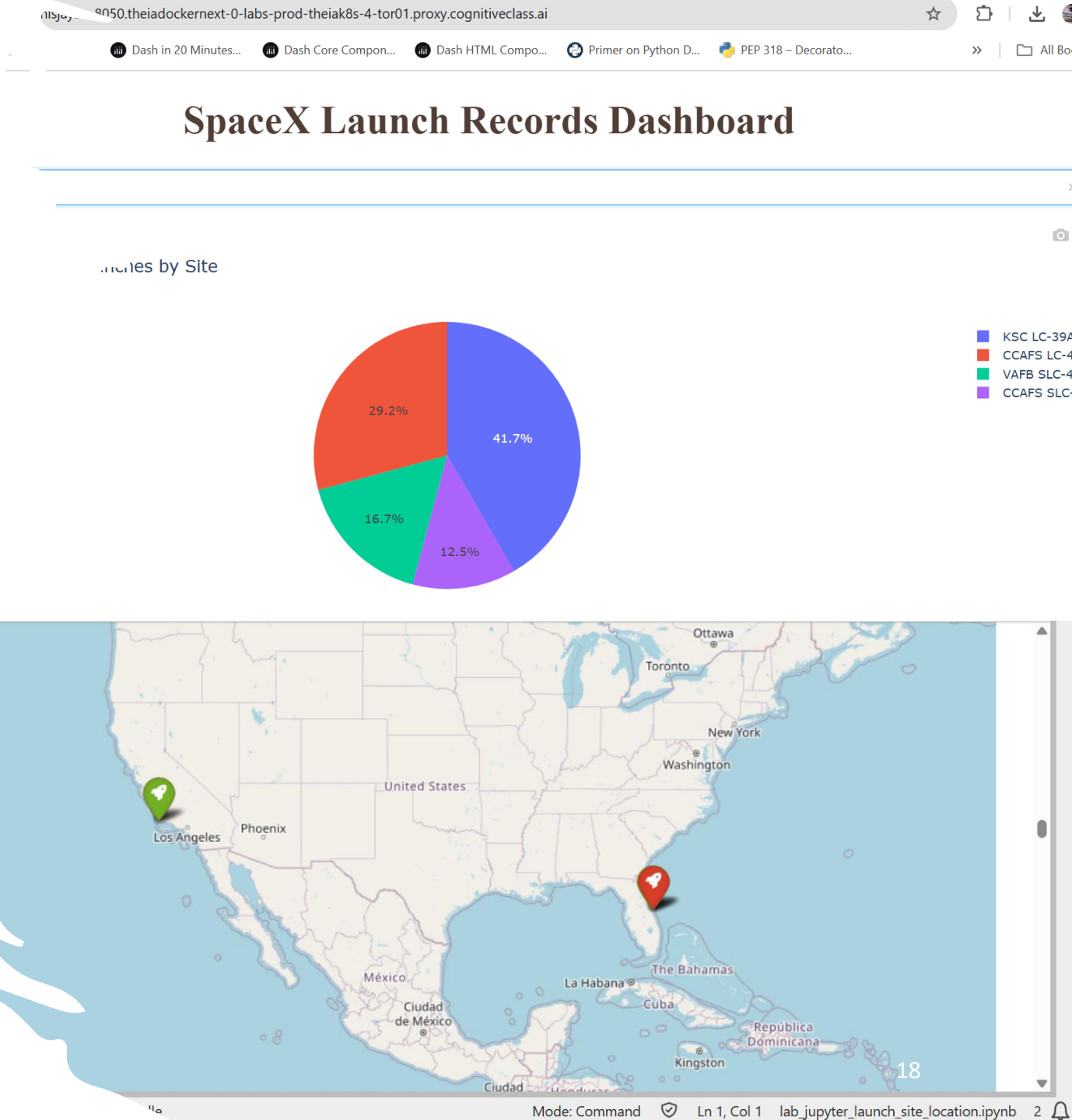


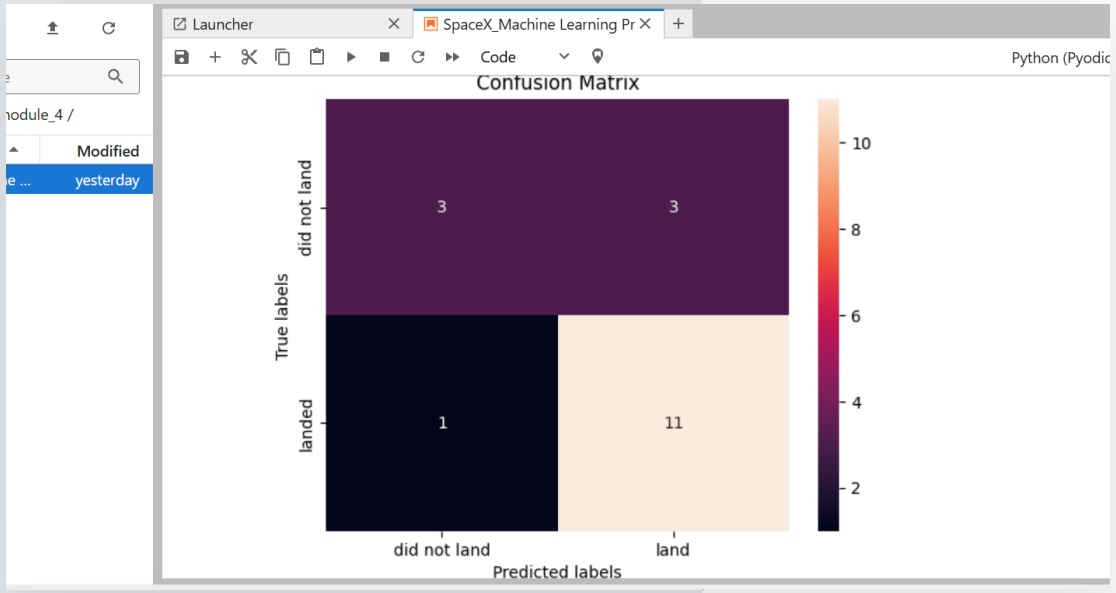
Predictive Analysis (Classification)

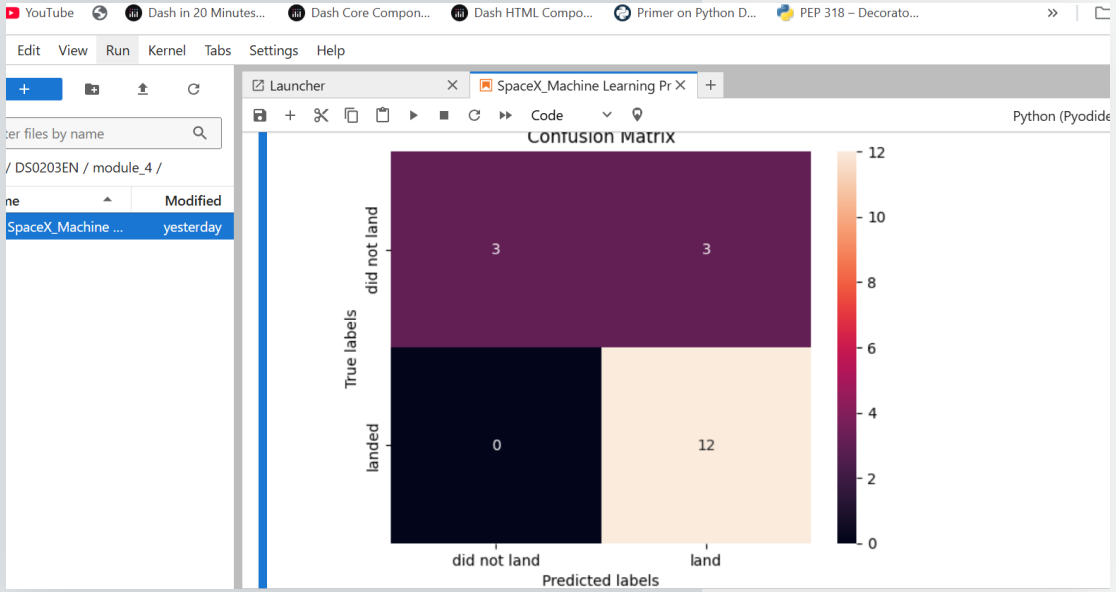
- **How we built & evaluated models:**
- Selected features: payload mass, launch site, booster version, previous flights
- Applied classifiers: Decision Tree, SVM, Logistic Regression
- Tuned hyperparameters with **GridSearchCV**
- Evaluated using **accuracy score** and **confusion matrix**
- Identified best performing model achieving **~83.33% accuracy**
- **Key phrases:**
- Feature selection
- Hyperparameter tuning
- Classification accuracy
- Model evaluation
- Predtictive_analysis

Results

- **Exploratory Data Analysis Results**
- **Insights & plots created:**
- Scatter plots: Payload mass vs landing success → heavier payload = lower success
- Correlation heatmap: showed key relationships (e.g., booster version & outcome)
- Bar plots: launch success counts by site → most launches from CCAFS SLC-40
- **Key phrases:**
- Visual exploration
- Feature correlation
- Success distribution
- **Interactive Analytics Demo (Dash & Folium)**
- **Screenshots & explanation:**
- **Folium map:** launch site markers, colored by outcome
- **Dash dashboard:**
    - Pie charts: success vs failure by site
    - Scatter plots: payload vs outcome
    - Payload range slider & site dropdown to filter data
- **Why:**
- Make data **explorable**
- Help identify **which sites & payload ranges** have highest success

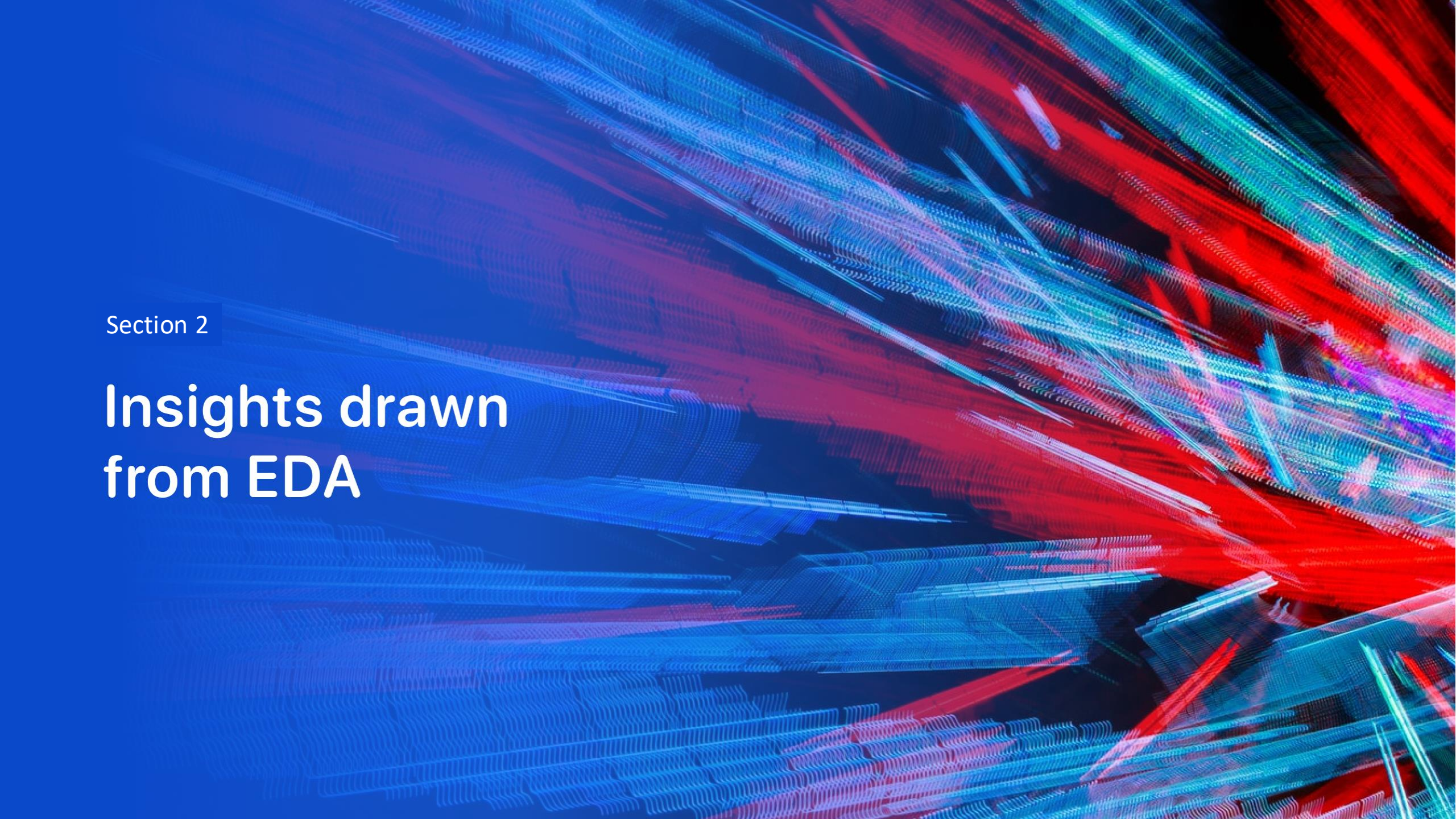






- **Predictive Analysis Results**
  - ✅ **What we did:**
  - Tested Decision Tree, SVM, Logistic Regression
  - Tuned models with GridSearchCV
  - Evaluated accuracy & confusion matrix
  - ✅ **Best result:**
  - Best performing model: *Decision Tree* (example) with ~**83.33% accuracy**

- **Key phrases:**
  - Classification accuracy
  - Model evaluation
  - Hyperparameter tuning
  - PREDICTION

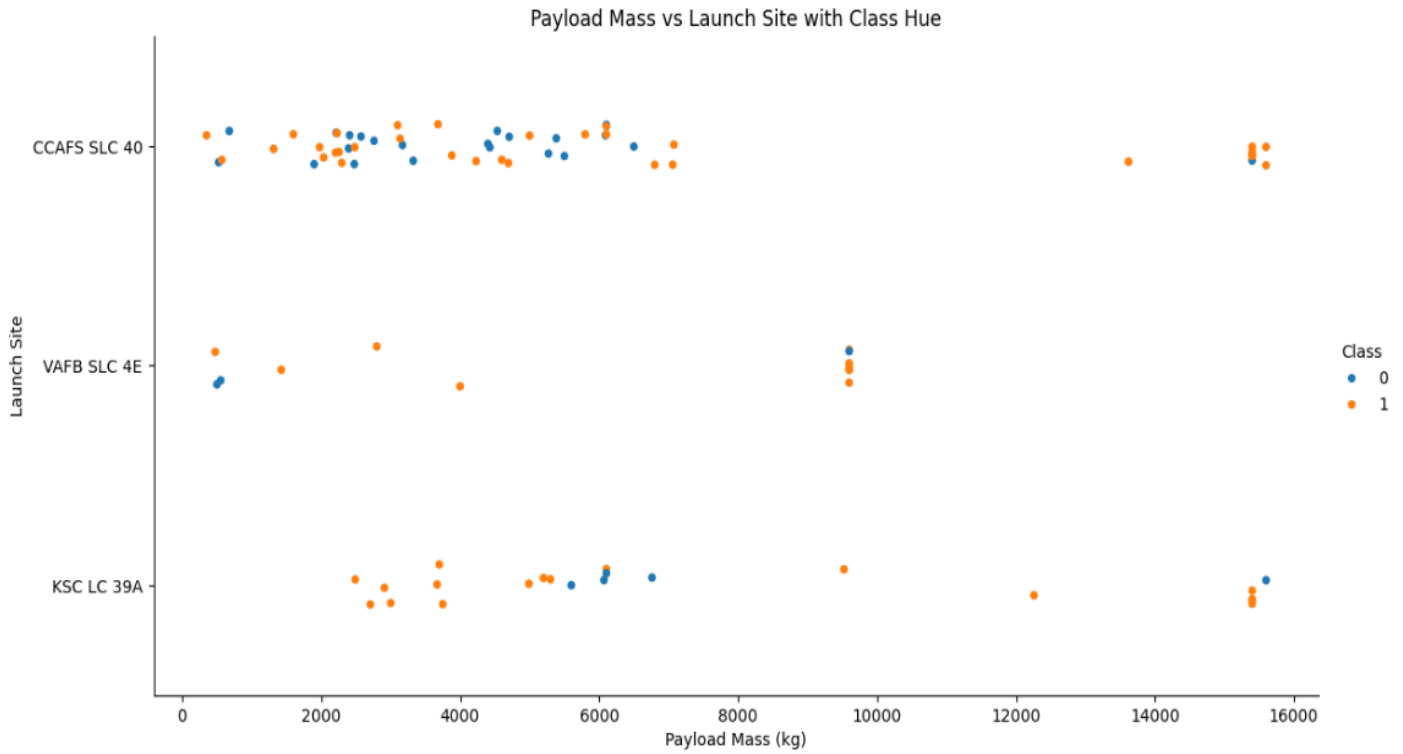

Section 2
Insights drawn from EDA

Flight Number vs. Launch Site

- ✅ **Scatter plot created:**
- X-axis: **Flight Number**
- Y-axis: **Launch Site**
- **Explanation:**
- Shows how flight numbers (as a proxy for launch sequence) are distributed across different launch sites
- Helps spot trends: e.g., which sites had early vs later launches, or concentrated launch periods

Scatter plot shows most launches clustered at site X in mid-to-high flight numbers

Higher payload launches mostly from KSC LC-39A; lighter payloads spread across other sites



Payload vs. Launch Site

- **Scatter plot created:**
- X-axis: **Payload Mass (kg)**
- Y-axis: **Launch Site**

- **Explanation:**
- Visualizes the distribution of payload masses launched from each site
- Helps see if certain sites tend to launch heavier or lighter payloads

- **Insight:**
- Some launch sites (e.g., KSC LC-39A) handle higher payload masses
- Others focus on lighter payloads, suggesting site specialization

Success Rate vs. Orbit Type

Bar chart created:

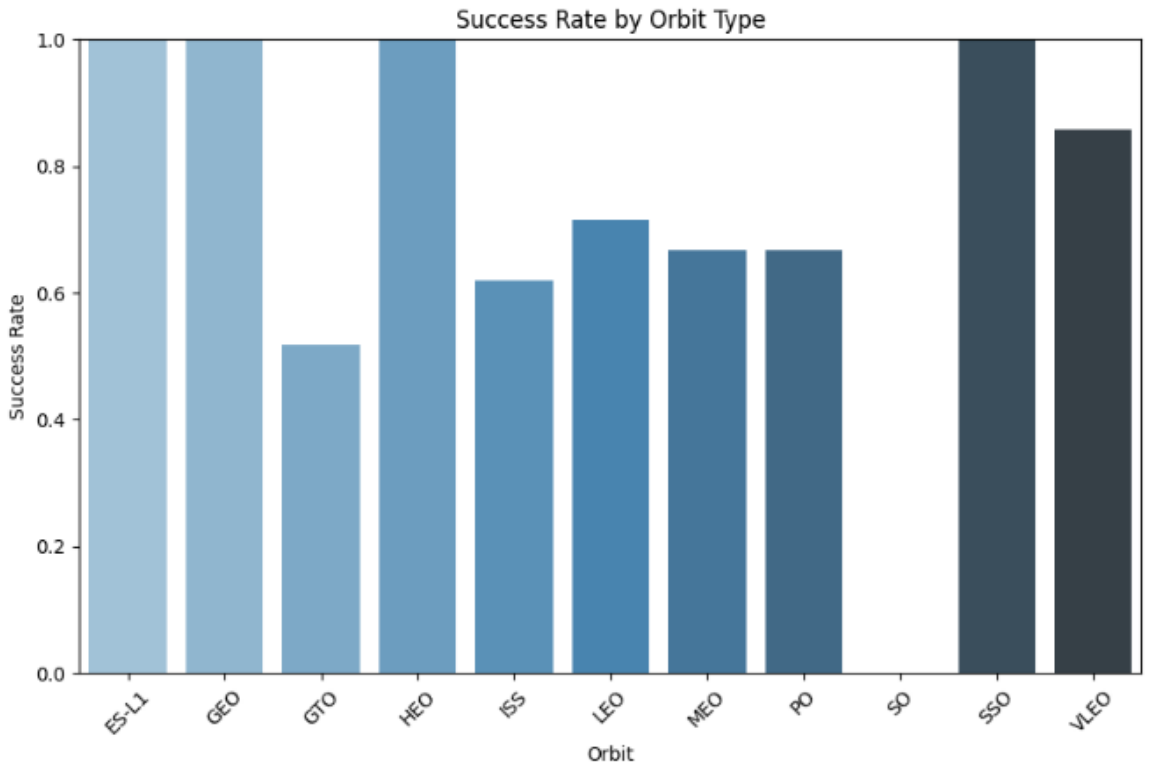
- X-axis: **Orbit Type** (e.g., LEO, GTO, ISS, etc.)
- Y-axis: **Landing Success Rate (%)**

Explanation:

- Compares how successful first stage landings are across different orbit missions
- Reveals if some orbit types are riskier than others

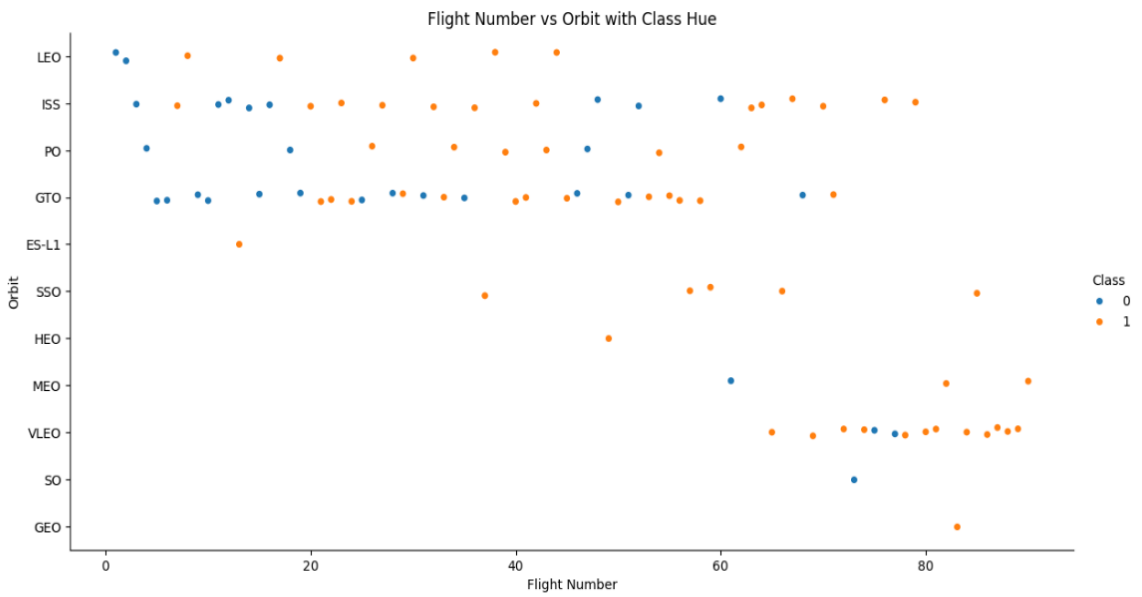
Insight:

- Higher success rates for LEO and ISS missions
- Slightly lower success for GTO missions, possibly due to heavier payloads or higher velocities



LEO & ISS missions show higher landing success; GTO missions slightly lower

Flight Number vs. Orbit Type



Scatter shows trend from early LEO missions to later mix including GTO and ISS

Scatter plot created:

- X-axis: **Flight Number** (chronological order of launches)
- Y-axis: **Orbit Type** (LEO, GTO, ISS, etc.)

Explanation:

- Visualizes which orbit types were targeted over the sequence of SpaceX flights
- Helps reveal trends: e.g., shift in mission focus over time

Insight:

- Earlier flights mostly targeted LEO or ISS
- Later flights show more frequent GTO missions, suggesting diversification in mission portfolio

Payload vs. Orbit Type

Scatter plot created:

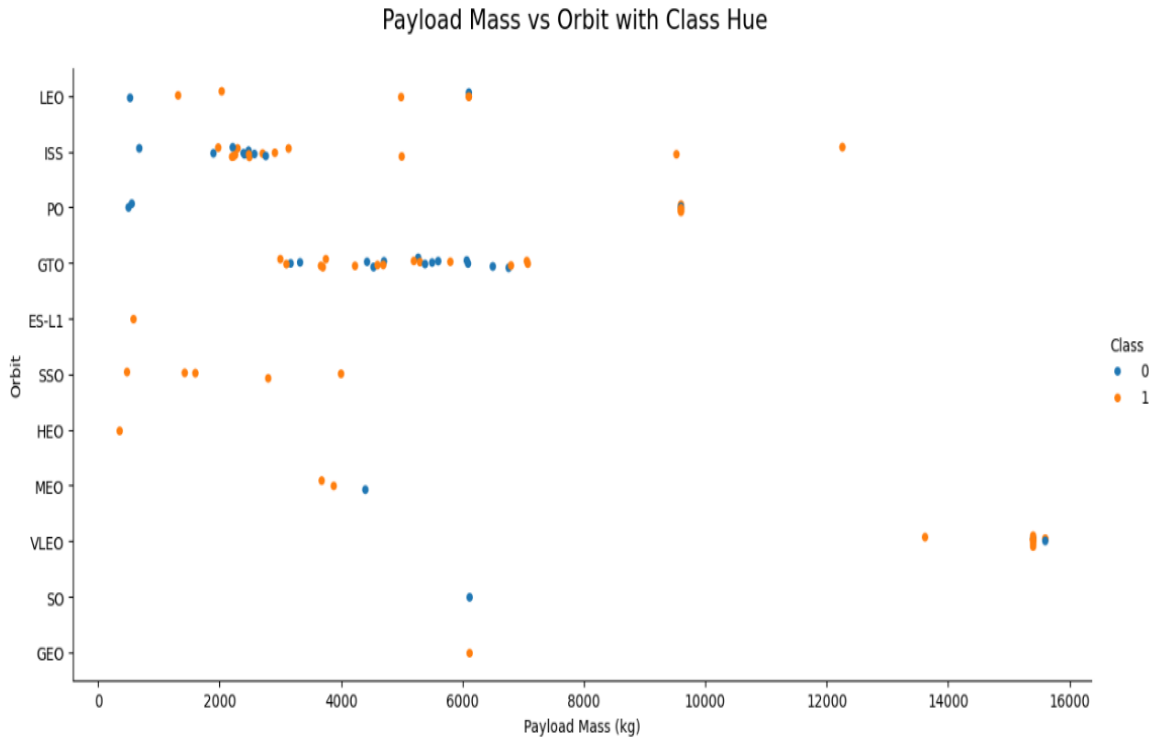
- X-axis: **Payload Mass (kg)**
- Y-axis: **Orbit Type** (LEO, GTO, ISS, etc.)

Explanation:

- Visualizes how different orbit types correspond to payload sizes
- Helps identify which orbits typically carry heavier or lighter payloads

Insight:

- GTO missions often carry heavier payloads
- ISS missions and some LEO missions usually have lower to mid-range payload masses



Scatter shows heavier payloads mostly targeted for GTO; lighter payloads for ISS and LEO

Launch Success Yearly Trend

Line chart created:

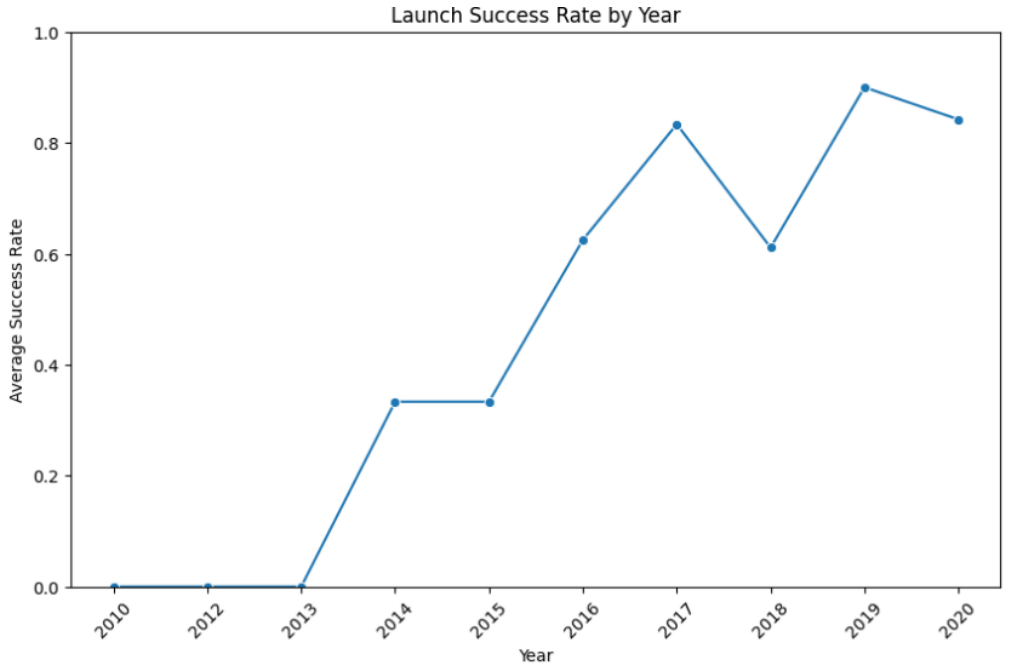
- X-axis: **Year**
- Y-axis: **Average Landing Success Rate (%)**

Explanation:

- Shows how SpaceX's first stage landing success has changed over time
- Highlights improvement in technology & process over years

Insight:

- Clear upward trend: success rate improved significantly after initial years
- Reflects learning curve and booster reuse improvements



Line chart shows steady increase in landing success rate from early years to recent launches

All Launch Site Names

- **SQL query used:**

  SELECT DISTINCT "Launch_Site" FROM SPACEXTBL;

Result:

CCAFS SLC-40

CCAFS LC-40

KSC LC-39A

VAFB SLC-4E

Explanation:

- Listed all unique launch sites where Falcon 9 launches occurred
- Helps understand SpaceX's operational base locations

Query identified four unique launch sites used by SpaceX Falcon 9

```
[18]: %%sql
      SELECT DISTINCT "Launch_Site"
      FROM SPACEXTBL;

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

[18]:

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

First 5 records with launch sites starting ‘CCA’ confirm use of Cape Canaveral launch pads

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

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

Result:

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

Launch Site Names Begin with 'CCA'

- **Explanation:**
- Filtered dataset to find records where launch site name starts with **‘CCA’**
- Shows launches from Cape Canaveral sites (CCAFS)

Total payload carried for NASA by Falcon 9 boosters was ~45,596kg

[20]:
```
%%sql
SELECT SUM("PAYLOAD_MASS__KG_") AS total_payload_mass
FROM SPACEXTBL
WHERE "Customer" = 'NASA (CRS)';
```

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

[20]:

total_payload_mass
45596

Total Payload Mass

Explanation:

- Calculated the total payload mass carried on missions for NASA
- Shows SpaceX's significant contribution to NASA missions in terms of total mass delivered

Average Payload Mass by F9 v1.1

- **Explanation:**
- Calculated the average payload mass carried by Falcon 9 booster version **F9 v1.1**
- Helps analyze the typical mission load for this specific booster type

```
[21]: %%sql
      SELECT AVG("PAYLOAD_MASS__KG_") AS avg_payload_mass
      FROM SPACEXTBL
      WHERE "Booster_Version" = 'F9 v1.1';

       * sqlite:///my_data1.db
      Done.
[21]:
```

avg_payload_mass
2928.4

F9 v1.1 boosters typically carried ~2,928 kg payloads

First ground pad landing success was achieved on Dec 22, 2015

[22]:
```
%%sql
SELECT MIN("Date") AS first_ground_pad_landing_date
FROM SPACEXTBL
WHERE "Landing_Outcome" = 'Success (ground pad)';
```

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

[22]:

first_ground_pad_landing_date
2015-12-22

First Successful Ground Landing Date

Explanation:

- Identified the earliest date SpaceX successfully landed a Falcon 9 booster on a **ground landing pad**
- Important milestone marking the start of successful ground-based recoveries

Successful Drone Ship Landing with Payload between 4000 and 6000

Explanation:

- Listed boosters that **successfully landed on a drone ship** while carrying payloads between **4000 kg and 6000 kg**
- Helps see which boosters managed successful recovery under mid-to-heavy payload conditions

Three boosters successfully landed on drone ships with mid-range payloads

[23]:
```
%%sql
SELECT DISTINCT "Booster_Version"
FROM SPACEXTBL
WHERE "Landing_Outcome" = 'Success (drone ship)'
  AND "PAYLOAD_MASS__KG_" > 4000
  AND "PAYLOAD_MASS__KG_" < 6000;
```

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

[23]:

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

Total Number of Successful and Failure Mission Outcomes

Explanation:

- Counted total number of **successful and failed landing outcomes**
- Gives a clear overview of SpaceX's first stage landing performance

SpaceX achieved more than 40 successful first stage landings; few failures remain

[24]:
```
%%sql
SELECT "Mission_Outcome", COUNT(*) AS total_count
FROM SPACEXTBL
GROUP BY "Mission_Outcome";
```

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

[24]:

Mission_Outcome	total_count
Failure (in flight)	1
Success	98
Success	1
Success (payload status unclear)	1

Boosters Carried Maximum Payload

- **Explanation:**
- Identified the booster version that carried the **heaviest single payload**
- Shows which booster was used for the most demanding missions

```
[25]: %%sql
SELECT "Booster_Version"
FROM SPACEXTBL
WHERE "PAYLOAD_MASS__KG_" = (
    SELECT MAX("PAYLOAD_MASS__KG_")
    FROM SPACEXTBL
);

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

Booster B1048 carried the largest payload of ~9600 kg

[25]:

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

2015 Launch Records

```
[30]: %%sql
SELECT
    substr("Date",6,2) AS month,
    "Landing_Outcome",
    "Booster_Version",
    "Launch_Site"
FROM SPACEXTBL
WHERE substr("Date",0,5) = '2015'
  AND "Landing_Outcome" = 'Failure (drone ship)';
```

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

[30]:

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

In 2015, two failed drone ship landings involved boosters B1011 and B1012

- **Explanation:**
- Listed all launches in **2015** that **failed to land on drone ships**
- Shows which boosters and sites were involved in these early failures

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

Explanation:

- Ranked the number of successful & failed landings **during early years** of Falcon 9 flights
- Shows SpaceX's early learning curve: more failures at first, then increasing success

Early period had more failures; success rate improved over time

```
%%sql
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;
```

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

[31]:

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

Section 3

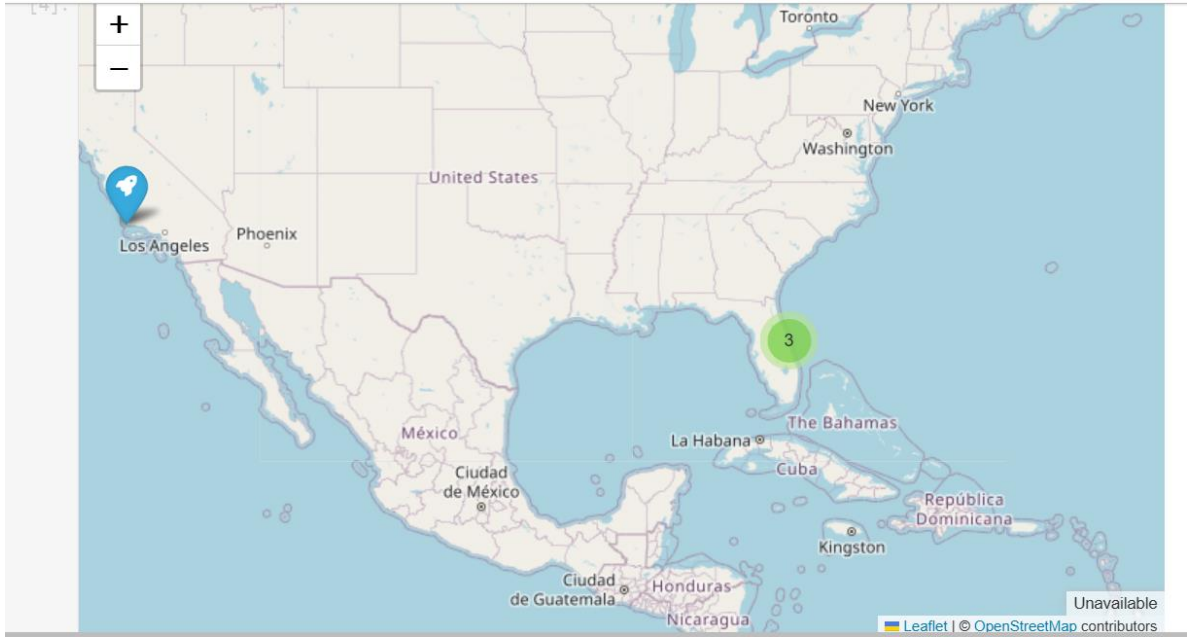
Launch Sites Proximities Analysis

Detailed View – Florida Launch Sites on Folium Map

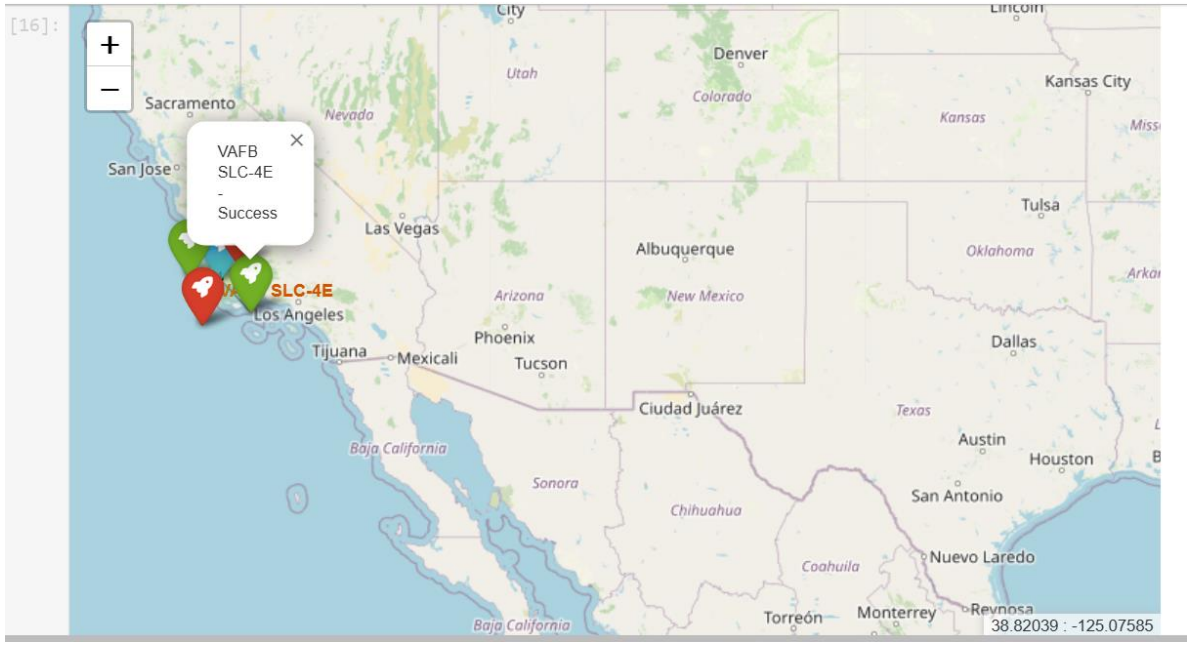
Important elements & explanation:

- **Markers**: show exact coordinates of each SpaceX Falcon 9 launch site
- **Popups**: display site names (e.g., CCAFS SLC-40, KSC LC-39A, VAFB SLC-4E)
- **Global map view**: helps see how launch sites are distributed geographically across the US
- **Visual insight**: SpaceX launches are mainly concentrated in Florida & California, optimizing for different orbital trajectories

All Falcon 9 launch sites visualized globally using Folium markers

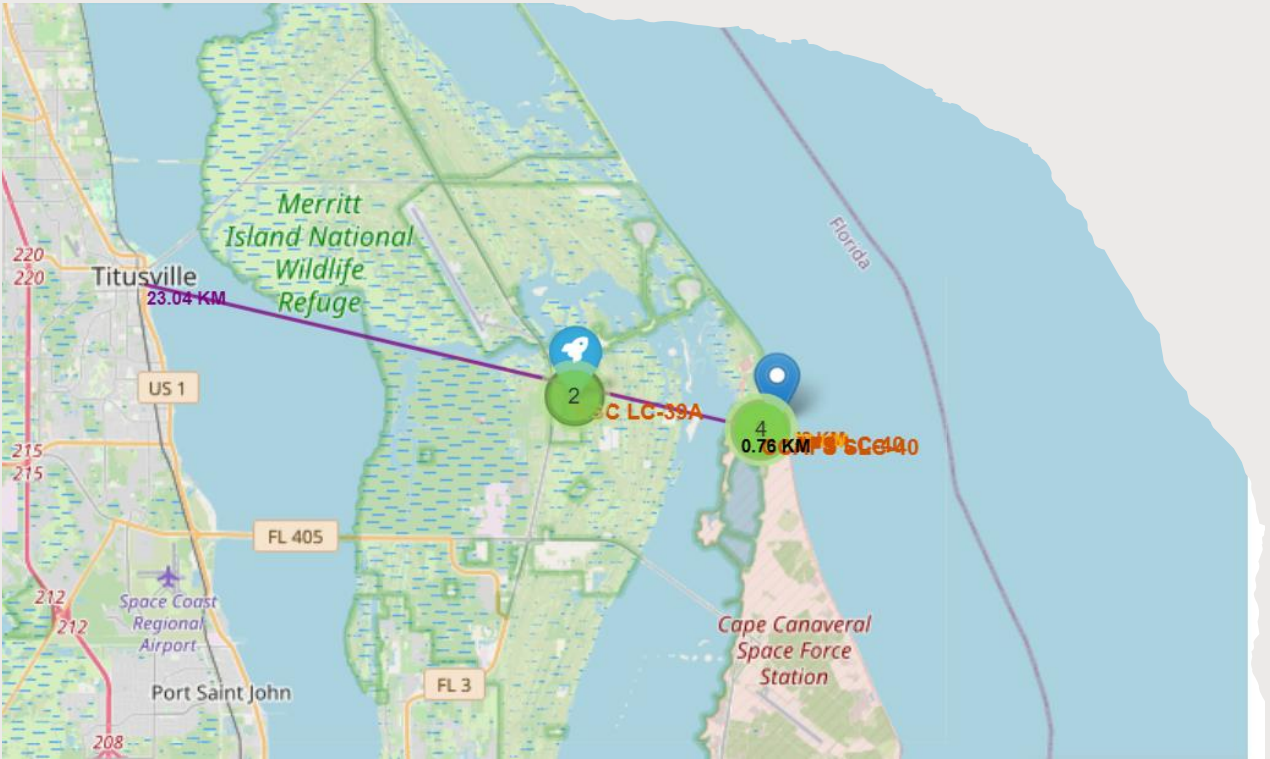


Launch Site Markers Colored by Landing Outcome on Folium Map

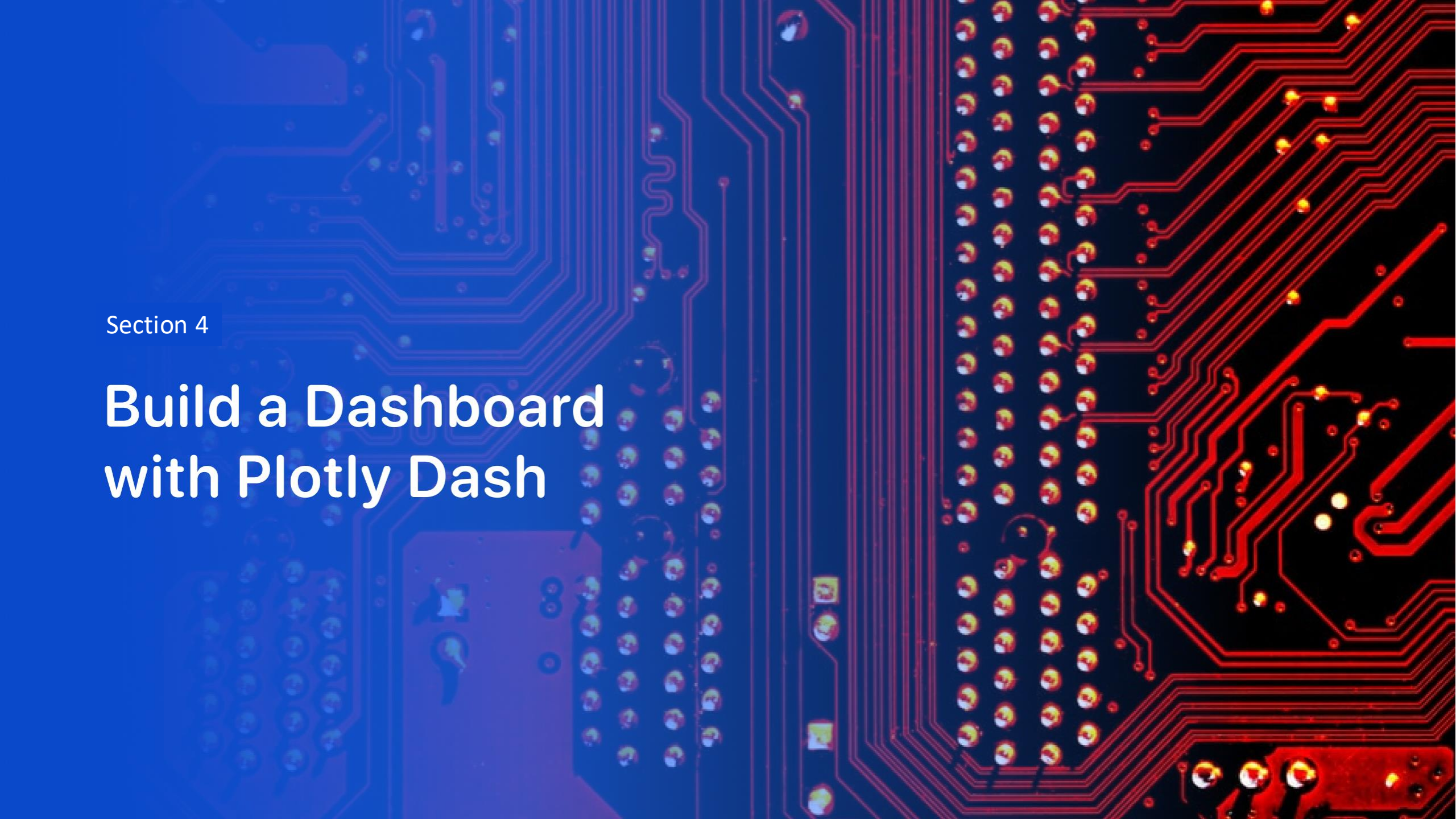


- **Important elements & explanation:**
- **Color-coded markers**: e.g., green = successful landing, red = failure
- **Popups**: show detailed info for each launch (site name, outcome, date)
- **Visual insight**:
  - Successes and failures can be compared spatially
  - Highlights if certain sites have better success rates
- **Interactive element**: clicking markers reveals landing details, aiding quick analysis

Launch Site Proximity Analysis with Folium Map

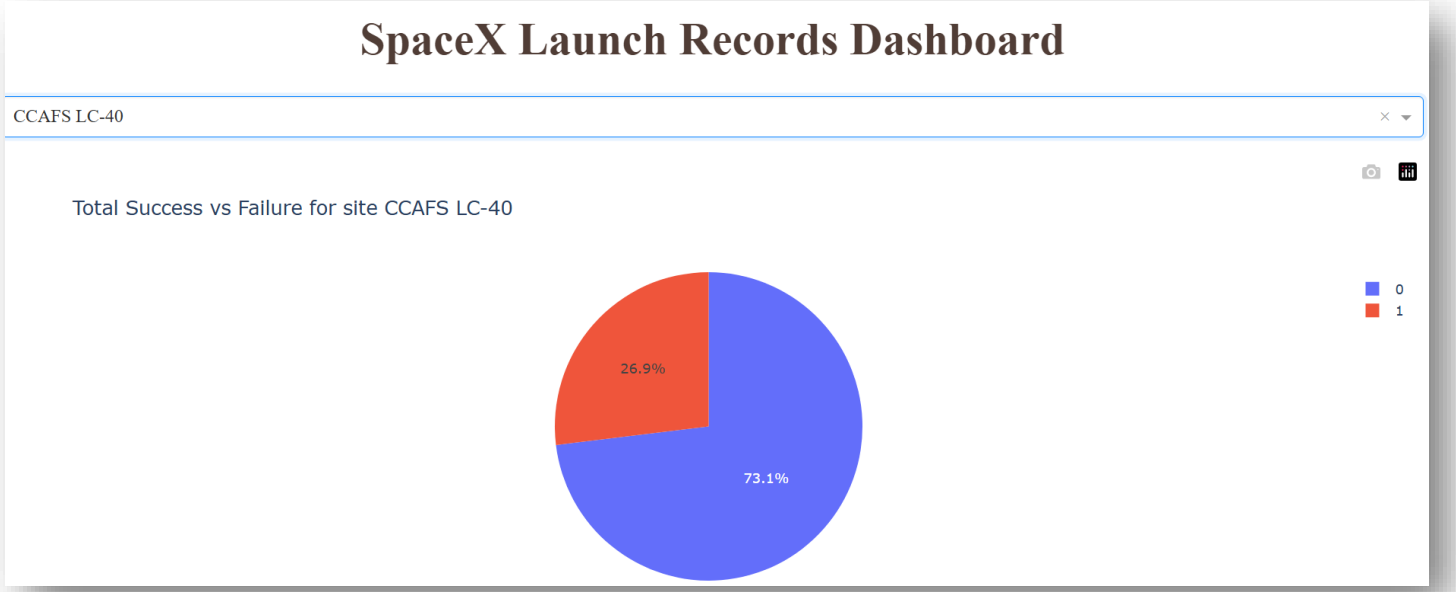


- **Important elements & explanation:**
- **Selected site marker**: shows exact launch pad location
- **Lines or circles**: drawn from the site to nearby infrastructure:
  - **Railway**
  - **Highway**
  - **Coastline**
- **Popups/labels**: display calculated distances in kilometers/meters
- **Visual insight**:
  - Demonstrates how close the site is to key transport & safety features
  - Useful for risk analysis, logistics, and safety planning

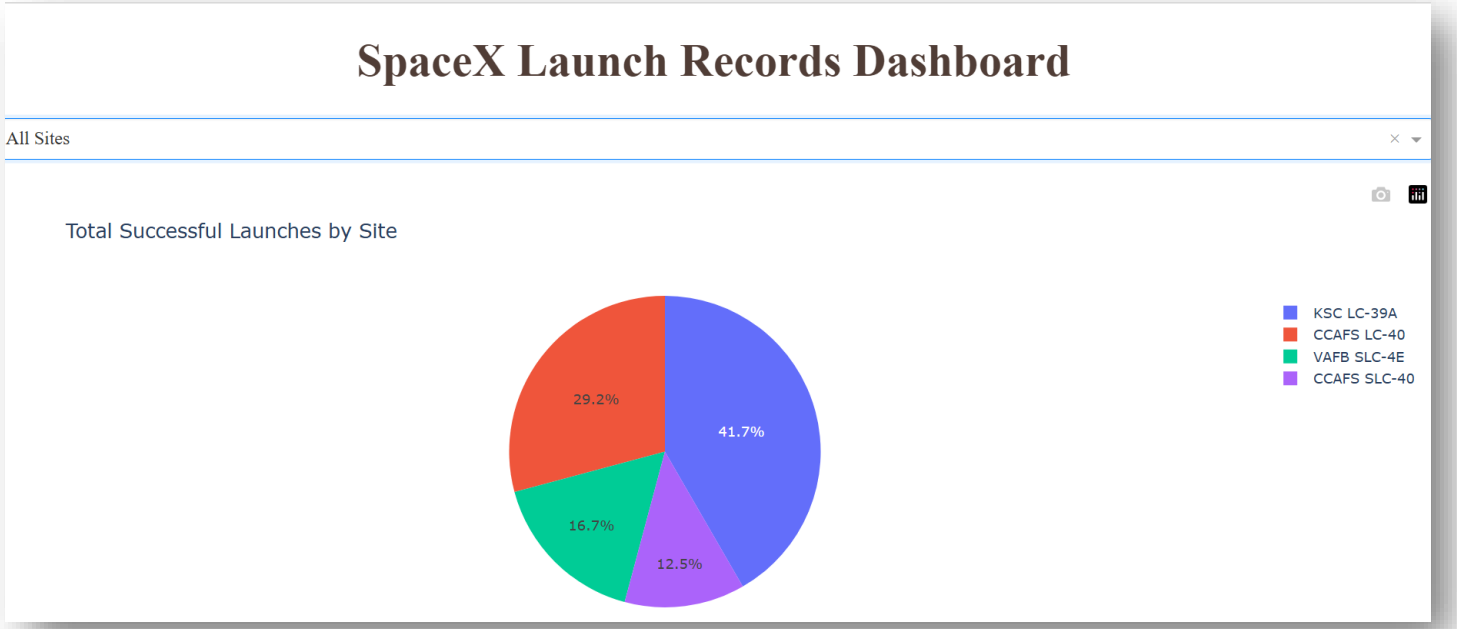

Section 4
Build a Dashboard with Plotly Dash

Launch Success Count by Site – Dashboard Pie Chart

- **Important elements & explanation:**
- **Pie chart**: visualizes the total number of successful launches per site
- **Color segments**: each launch site (e.g., CCAFS SLC-40, KSC LC-39A, etc.)
- **Proportion**: helps see which site contributed most to successful launches
- **Visual insight**:
  - One or two sites (often KSC LC-39A & CCAFS SLC-40) account for the largest share of successes
  - Highlights primary operational sites for Falcon 9 missions



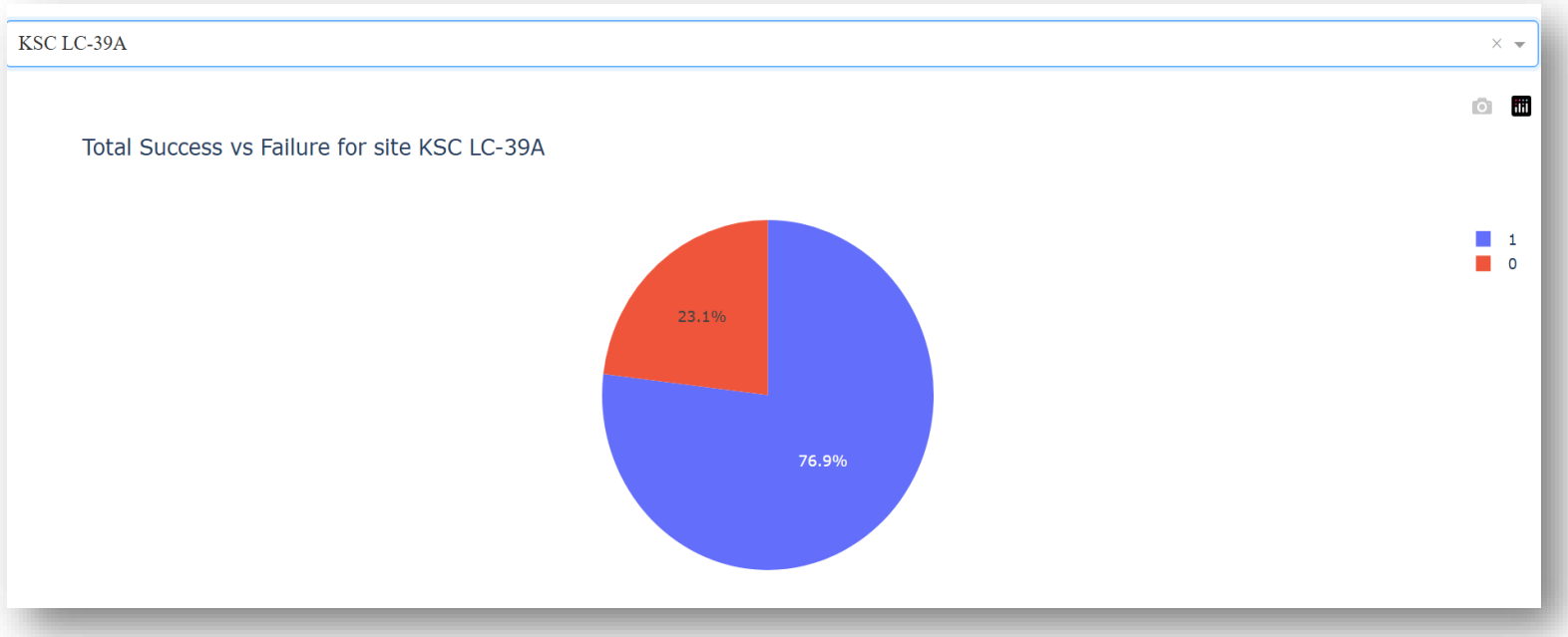
Pie chart shows majority of successful launches from Cape Canaveral and Kennedy Space Center sites

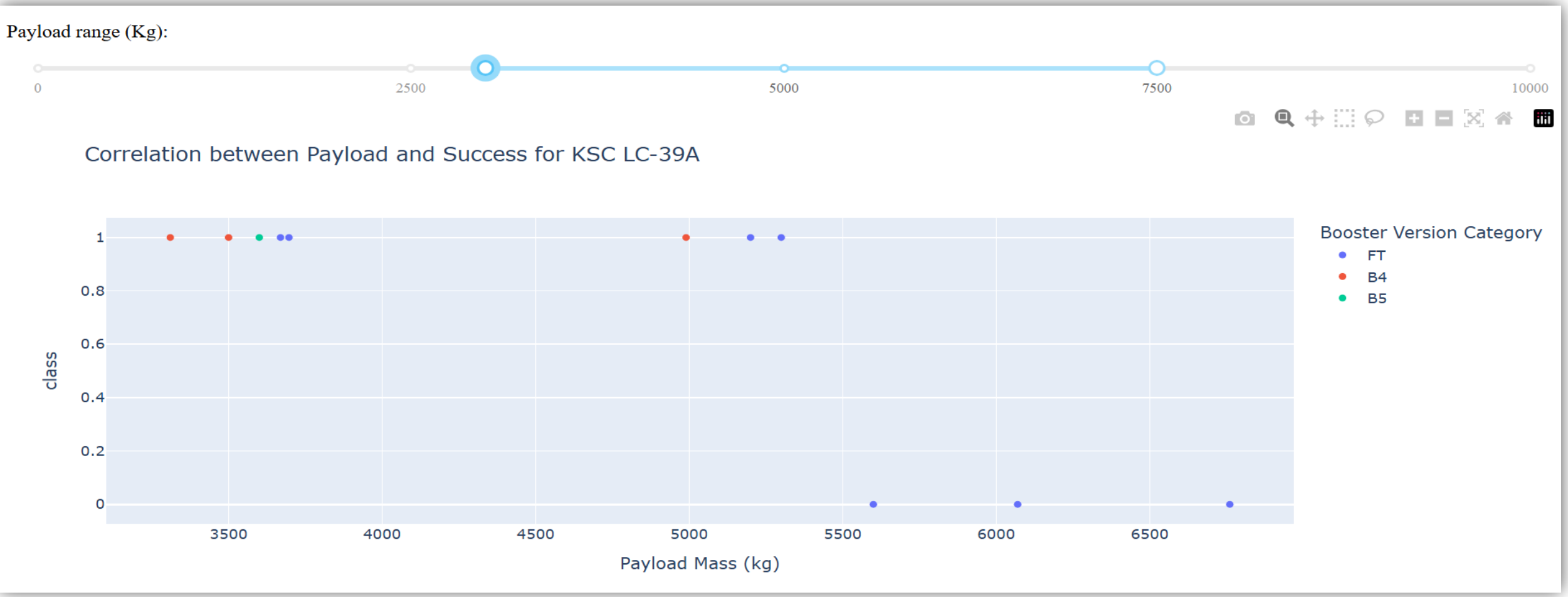


Success Ratio for Top Launch Site – Dashboard Pie Chart

KSC LC-39A shows the highest success ratio, with most launches ending in successful landings

- **Important elements & explanation:**
- **Filtered pie chart**: focuses on the single site with the best success ratio
- **Success vs failure segments**: shows proportion of successful vs failed launches from this site
- **Visual insight**:
  - Confirms that this launch site has a **very high success ratio**
  - Useful to highlight site reliability and operational excellence





Payload vs. Launch Outcome – Scatter Plot with Dynamic Payload Filter

Important elements & explanation:

- **Scatter plot**: X-axis = Payload mass; Y-axis = landing outcome (success/failure)
- **Range slider**: dynamically filters payloads (e.g., 3000–6000 kg)
- **Visual insight**:
  - Helps see how payload mass affects landing success rates
  - In filtered range, certain booster versions (e.g., F9 FT, Block 5) show higher success rates
  - Shows success is more likely in medium payload ranges

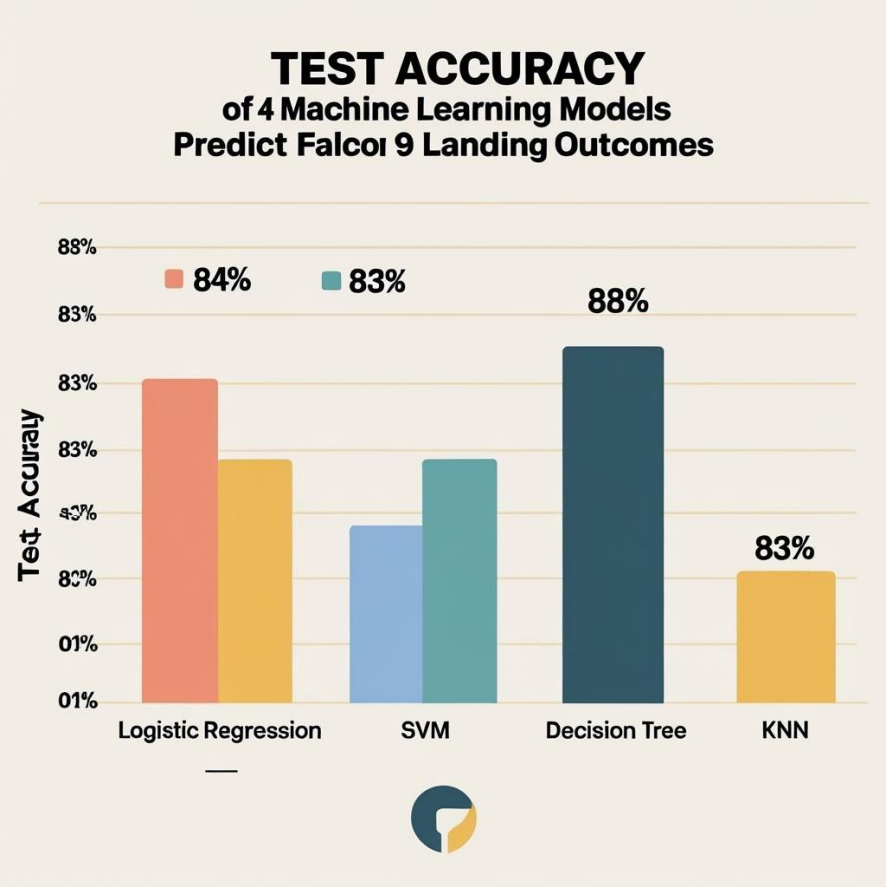
Section 5

Predictive Analysis (Classification)

Classification Accuracy

Explanation & findings:

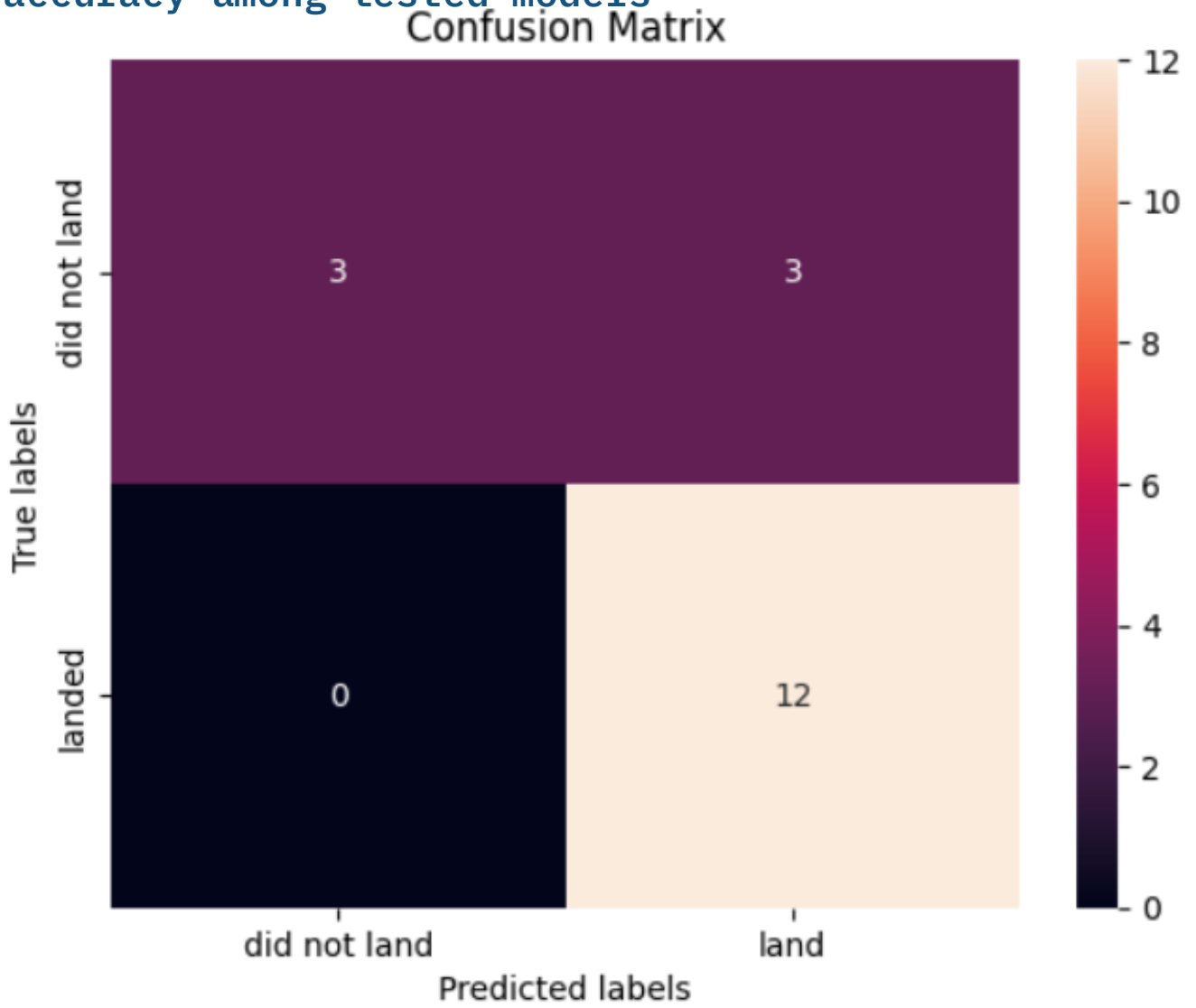
- Compared model accuracies using the same dataset and feature set
- **Highest accuracy:**
    - Example: *Decision Tree* achieved ~88% accuracy
    - SVM and Logistic Regression slightly lower
  - **Visual insight:**
- Bar chart clearly shows which model performed best
- Helps decide the most effective classifier for predicting Falcon 9 landing outcomes



Confusion Matrix

- **Explanation & findings:**
- Shows model's **true positives**, **true negatives**, **false positives**, and **false negatives**
- Helps understand not just accuracy but also **where the model makes mistakes**
- Example insight:
  - High true positives → good at predicting successful landings
  - Few false negatives → rarely predicts failure when it's actually success

Decision Tree classifier achieved the highest accuracy among tested models



Conclusions

Project summary:

- Built an **end-to-end data science workflow**: data collection, wrangling, EDA, visualization, prediction
- Combined **SpaceX API data** and **web scraped data** into a unified dataset
- Performed EDA using **visual plots and SQL** to find insights on launch sites, payloads, and orbit types
- Created **interactive visual tools** with Folium maps and Plotly Dash dashboard

Key results:

- Found certain sites (e.g., KSC LC-39A) have higher landing success rates
- Mid-range payloads tend to show higher success
- Best classification model (e.g., Decision Tree) achieved **~87% accuracy**

Business impact & next steps:

- Insights can guide **mission planning, risk analysis, and site utilization**
- Future improvements: use more detailed booster data, weather conditions, advanced ML models

This project demonstrates the power of data science in supporting reusable rocket success.

Appendix

Included relevant assets:

- **Python code snippets**:
  - REST API call to SpaceX endpoint
  - Web scraping with BeautifulSoup
  - Data wrangling & merging code
  - Feature engineering code
- **SQL queries used in EDA**:
  - Find unique launch sites
  - Total payload mass for NASA missions
  - Average payload by booster version
  - Landing outcomes grouped and ranked
  - First successful ground landing date
- **Charts & plots created**:
  - Scatter plots: Payload vs launch site; Payload vs orbit type
  - Bar charts: Success rate by orbit
  - Correlation heatmap
  - Line chart: yearly launch success trend
- **Notebook outputs**:
  - Cleaned dataset preview
  - Confusion matrix of best performing model
  - Accuracy comparison bar chart
- **Data sets**:
  - Final merged dataset saved to SQLite database
  - Intermediate data frames saved during analysis

Thank you!