IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Partho Sarothi Das
25th September 2025

Outline

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

Executive Summary

- **Methodologies**
  - Collected SpaceX data using REST API and web scraping
  - Performed data wrangling to clean and transform datasets
  - Conducted EDA with SQL and visualizations
  - Built interactive analytics with Folium maps and Plotly Dash
  - Developed and evaluated classification models for launch outcome prediction
- **Results**
  - Identified patterns between payload, launch site, orbit type, and mission success
  - Built interactive dashboards for visual exploration of launch performance
  - Found the best-performing classification model with high accuracy in predicting mission success

Introduction

Project Background:

SpaceX aims to reduce the cost of space travel with reusable rockets.

Problem Statement:

Can we predict launch success and understand factors (payload, orbit, site) influencing outcomes?

Objective:

Perform data-driven analysis and predictive modeling on SpaceX launches.

Section 1

Methodology

Methodology

Executive Summary

- Data collection methodology:

  Extracted launch data from SpaceX API (REST calls), Scraped additional data from SpaceX Wikipedia page, Stored results in CSV for processing

- Data wrangling

  Cleaned missing/invalid values, Created derived columns, Standardized payload mass and booster categories

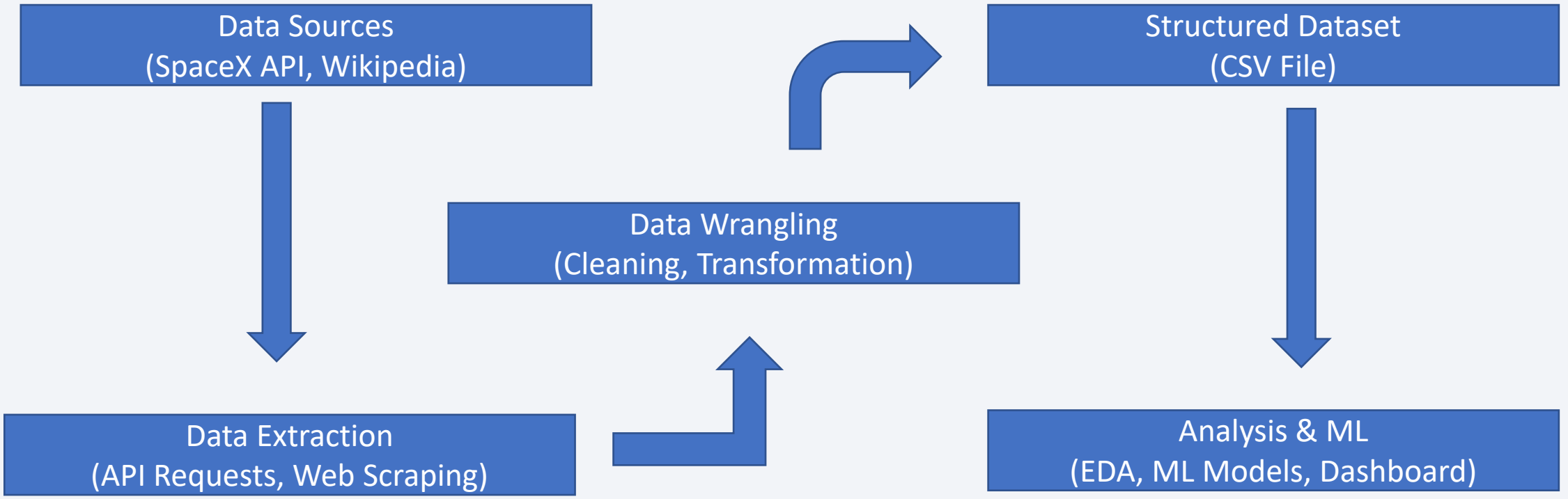
- Perform exploratory data analysis (EDA) using visualization and SQL

- Perform interactive visual analytics using Folium and Plotly Dash

- Perform predictive analysis using classification models

  Built classification models (Logistic Regression, Decision Tree, KNN, SVM), Hyperparameter tuning for best performance, Evaluated models with accuracy scores and confusion matrix.

Data Collection

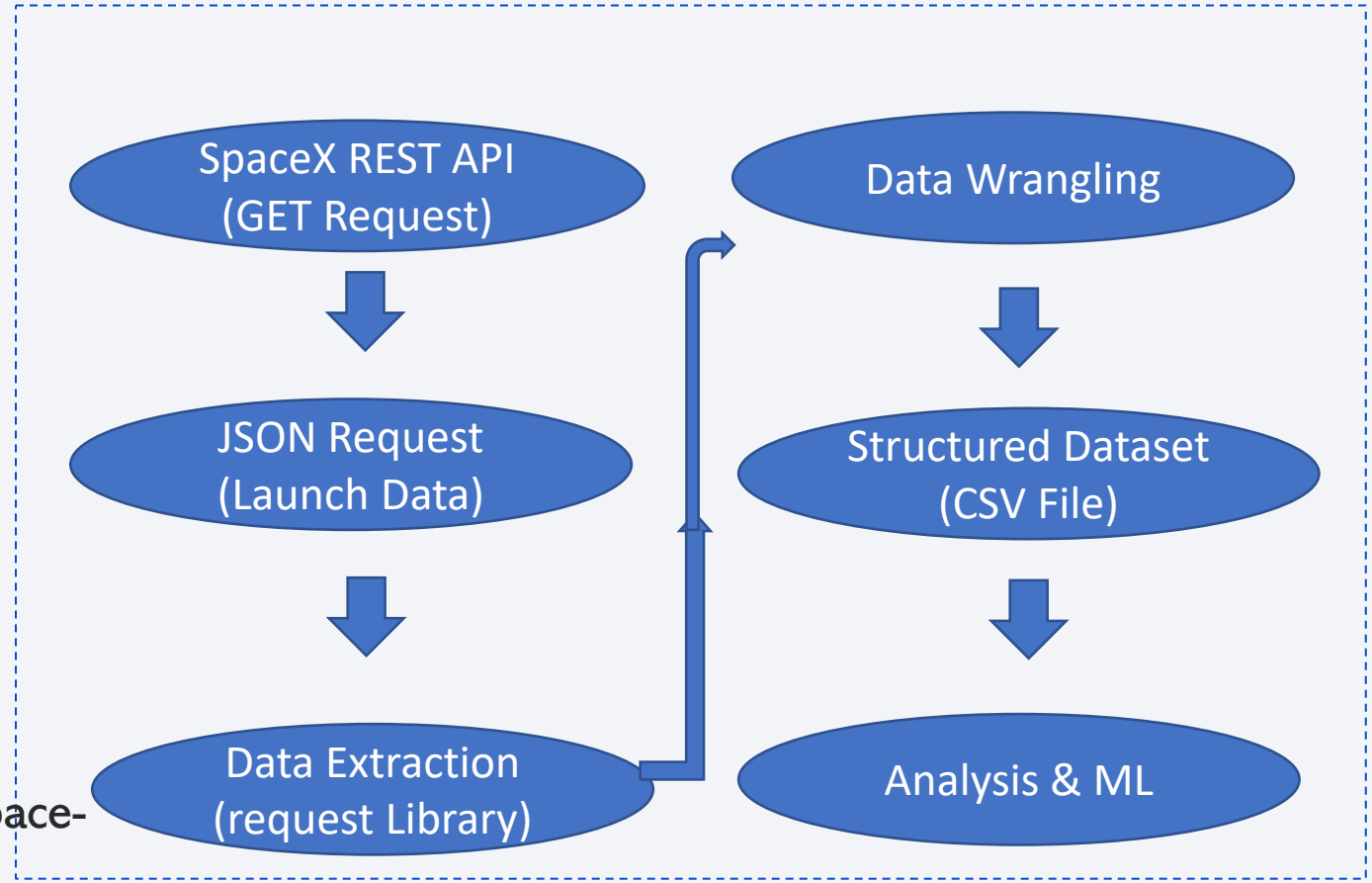


Data Collection – SpaceX API

- Key Phrases to

1. Rest API Requests (GET method)
2. JSON Response Parsing
3. Data Wrangling with Pandas
4. Data Cleaning
5. Structured Dataset Creating (CSV)

GitHub URL:
https://github.com/ParthoSarothiDas/space-race-spaceX



Data Collection - Scraping

- Present your web scraping process using key phrases and flowcharts

GitHub URL:
https://github.com/ParthoSarothiDas/space-race-spaceX

Place your flowchart of web scraping here

Data Wrangling

- Describe how data were processed
- You need to present your data wrangling process using key phrases and flowcharts

GitHub URL: https://github.com/ParthoSarothiDas/space-race-spaceX

EDA with Data Visualization

- Summarize what charts were plotted and why you used those charts
- Add the GitHub URL of your completed EDA with data visualization notebook, as an external reference and peer-review purpose

EDA with SQL

- Using bullet point format, summarize the SQL queries you performed
- Add the GitHub URL of your completed EDA with SQL notebook, as an external reference and peer-review purpose

Build an Interactive Map with Folium

- Summarize what map objects such as markers, circles, lines, etc. you created and added to a folium map
- Explain why you added those objects

GitHub URL: https://github.com/ParthoSarothiDas/space-race-spaceX

Build a Dashboard with Plotly Dash

- Summarize what plots/graphs and interactions you have added to a dashboard
- Explain why you added those plots and interactions
- Add the GitHub URL of your completed Plotly Dash lab, as an external reference and peer-review purpose

Predictive Analysis (Classification)

- Summarize how you built, evaluated, improved, and found the best performing classification model
- You need present your model development process using key phrases and flowchart
- Add the GitHub URL of your completed predictive analysis lab, as an external reference and peer-review purpose

Results

- Exploratory data analysis results
- Interactive analytics demo in screenshots
- Predictive analysis results

Section 2

Insights drawn from EDA

Flight Number vs. Launch Site

- Show a scatter plot of Flight Number vs. Launch Site

- Show the screenshot of the scatter plot with explanations

Payload vs. Launch Site

- Show a scatter plot of Payload vs. Launch Site

- Show the screenshot of the scatter plot with explanations

Success Rate vs. Orbit Type

- Show a bar chart for the success rate of each orbit type

- Show the screenshot of the scatter plot with explanations

Flight Number vs. Orbit Type

- Show a scatter point of Flight number vs. Orbit type

- Show the screenshot of the scatter plot with explanations

Payload vs. Orbit Type

- Show a scatter point of payload vs. orbit type

- Show the screenshot of the scatter plot with explanations

Launch Success Yearly Trend

- Show a line chart of yearly average success rate

- Show the screenshot of the scatter plot with explanations

All Launch Site Names

- Find the names of the unique launch sites
- Present your query result with a short explanation here

Launch Site Names Begin with 'CCA'

- Find 5 records where launch sites begin with `CCA`
- Present your query result with a short explanation here

Total Payload Mass

- Calculate the total payload carried by boosters from NASA
- Present your query result with a short explanation here

Average Payload Mass by F9 v1.1

- Calculate the average payload mass carried by booster version F9 v1.1
- Present your query result with a short explanation here

First Successful Ground Landing Date

- Find the dates of the first successful landing outcome on ground pad
- Present your query result with a short explanation here

Successful Drone Ship Landing with Payload between 4000 and 6000

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

- Present your query result with a short explanation here

Total Number of Successful and Failure Mission Outcomes

- Calculate the total number of successful and failure mission outcomes
- Present your query result with a short explanation here

Boosters Carried Maximum Payload

- List the names of the booster which have carried the maximum payload mass
- Present your query result with a short explanation here

2015 Launch Records

- List the failed landing_outcomes in drone ship, their booster versions, and launch site names for in year 2015

- Present your query result with a short explanation here

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

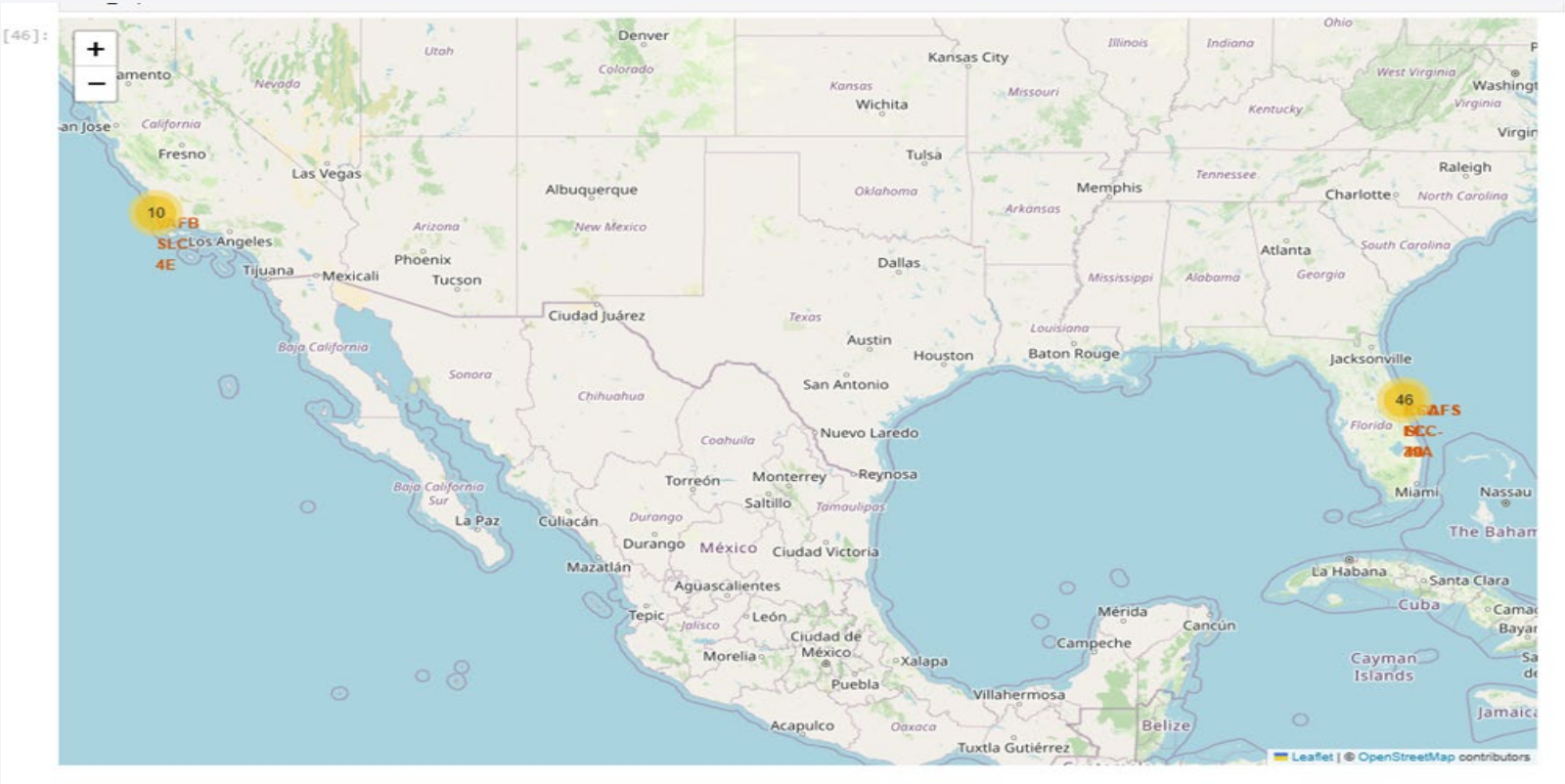
- Rank the count of landing outcomes (such as Failure (drone ship) or Success (ground pad)) between the date 2010-06-04 and 2017-03-20, in descending order

- Present your query result with a short explanation here

Section 3

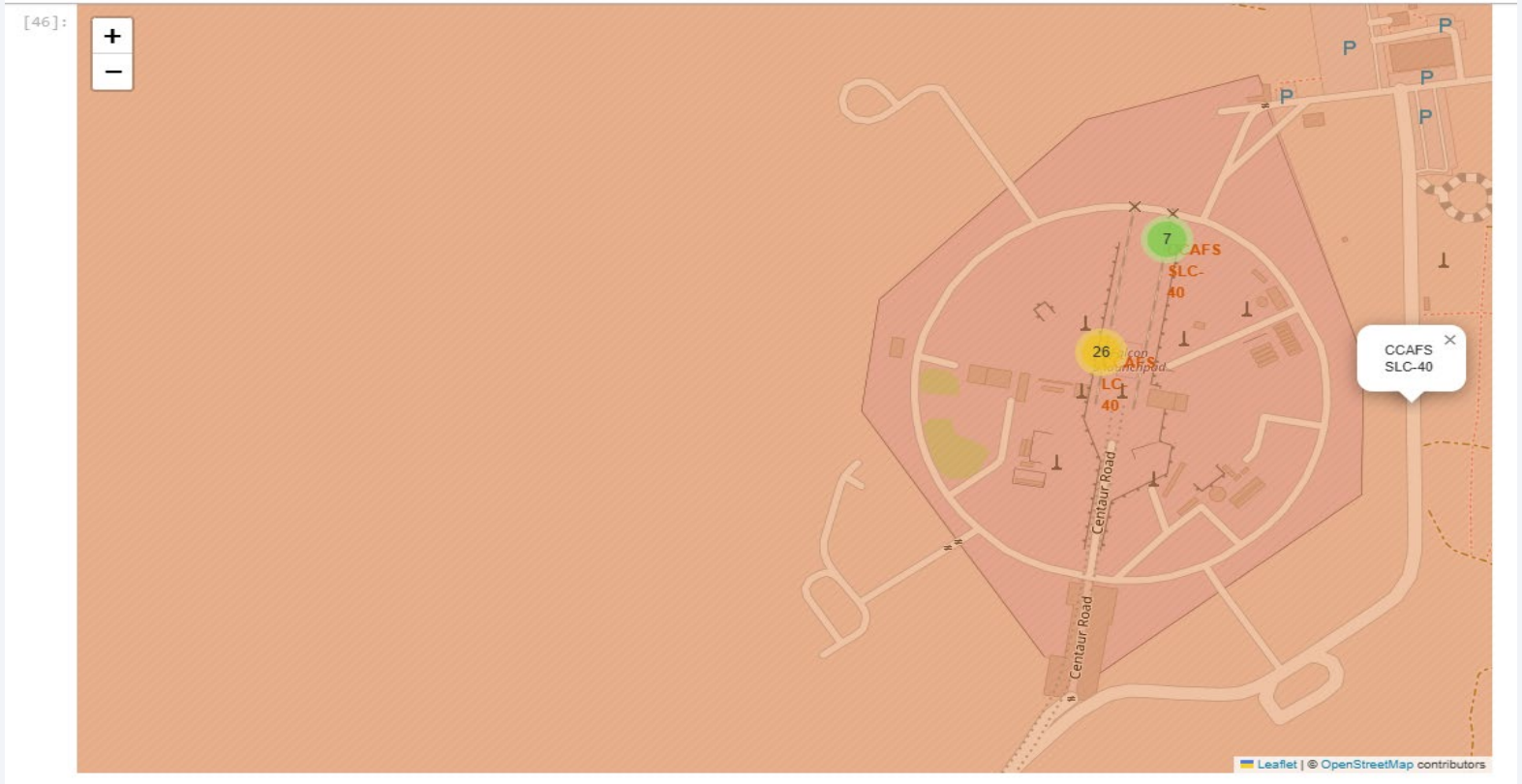
Launch Sites Proximities Analysis

All Launch Sites



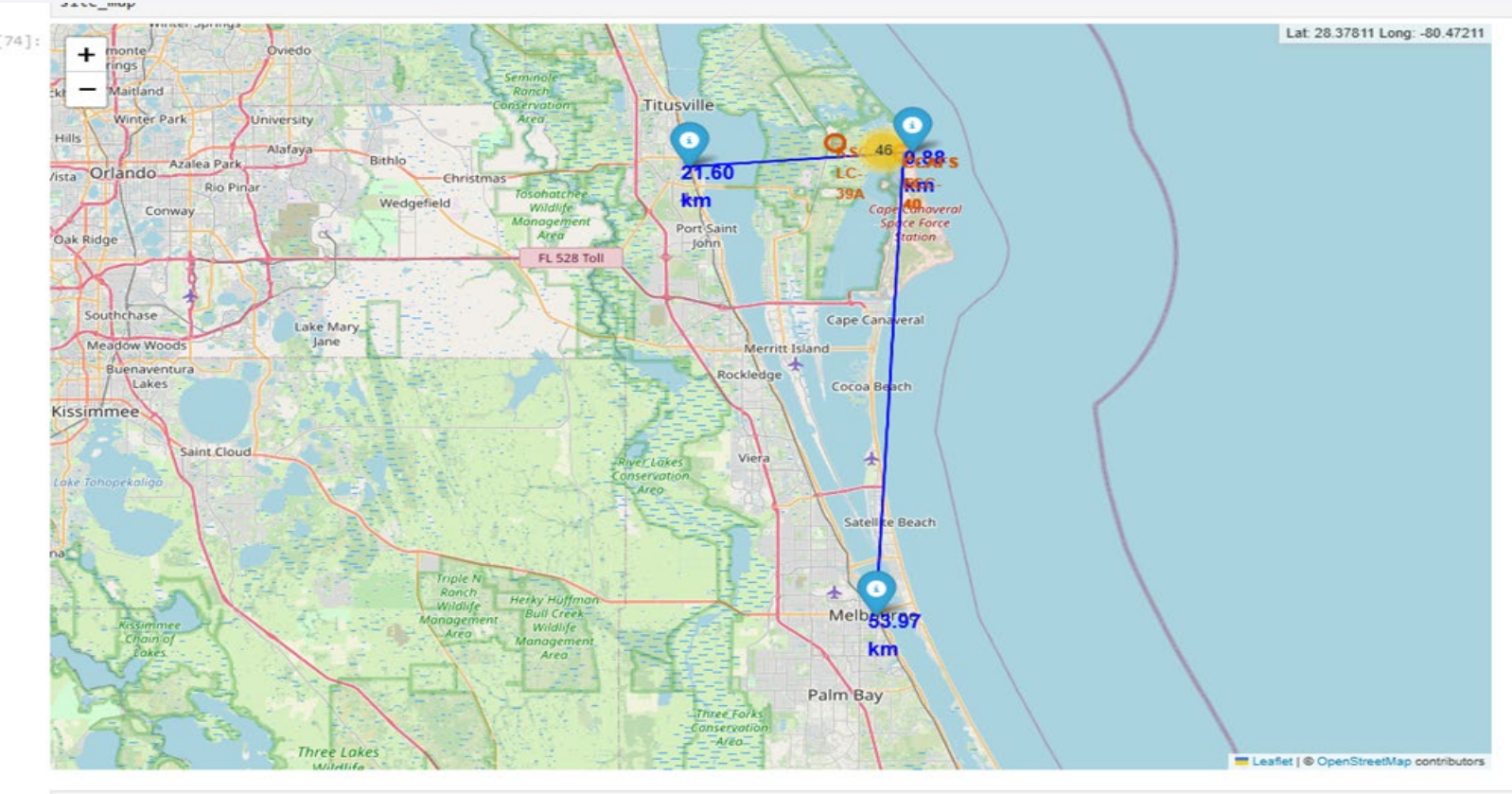
- Explain the important elements and findings on the screenshot

Color-labeled Launch Outcomes

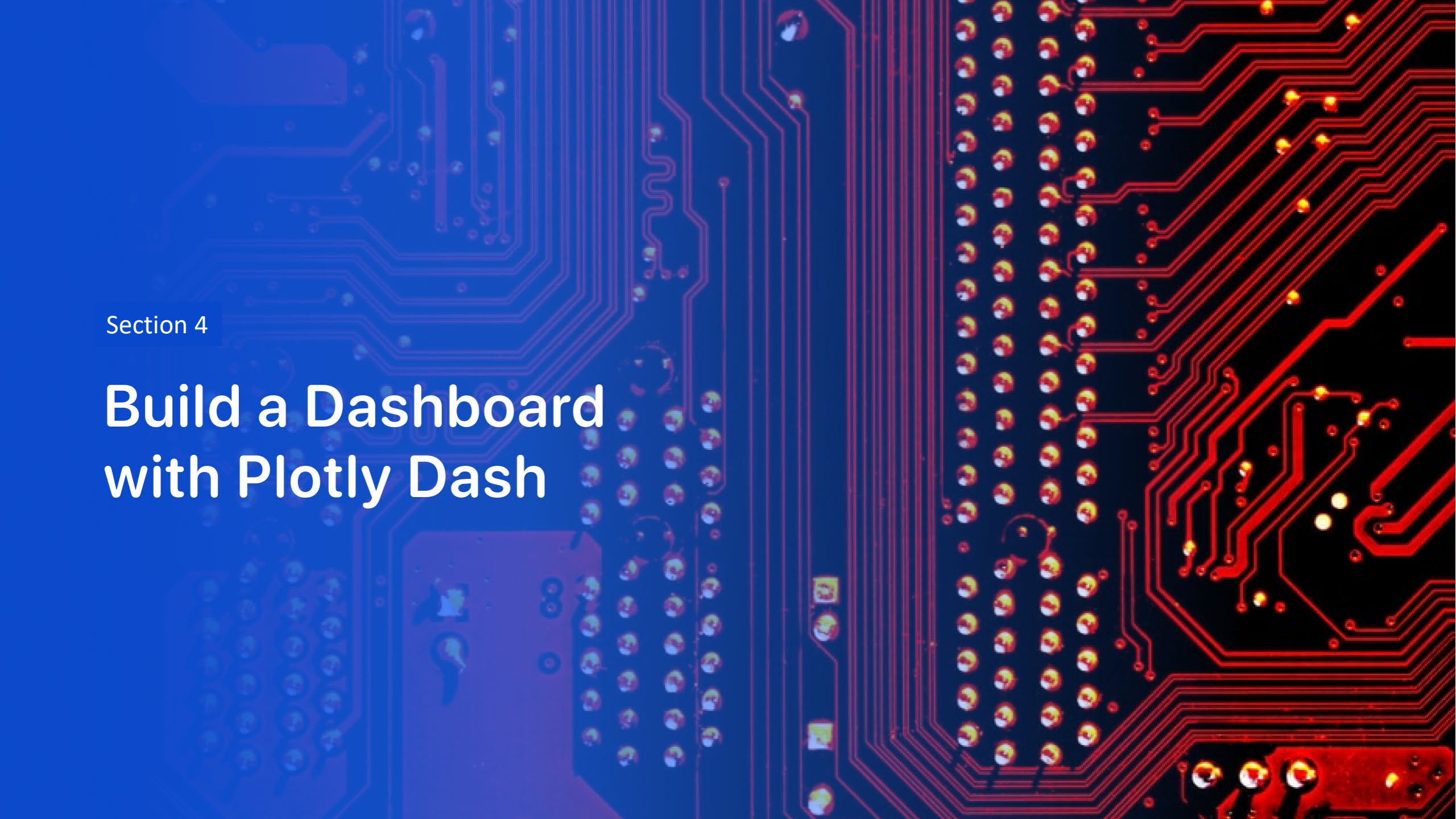


- Explain the important elements and findings on the screenshot

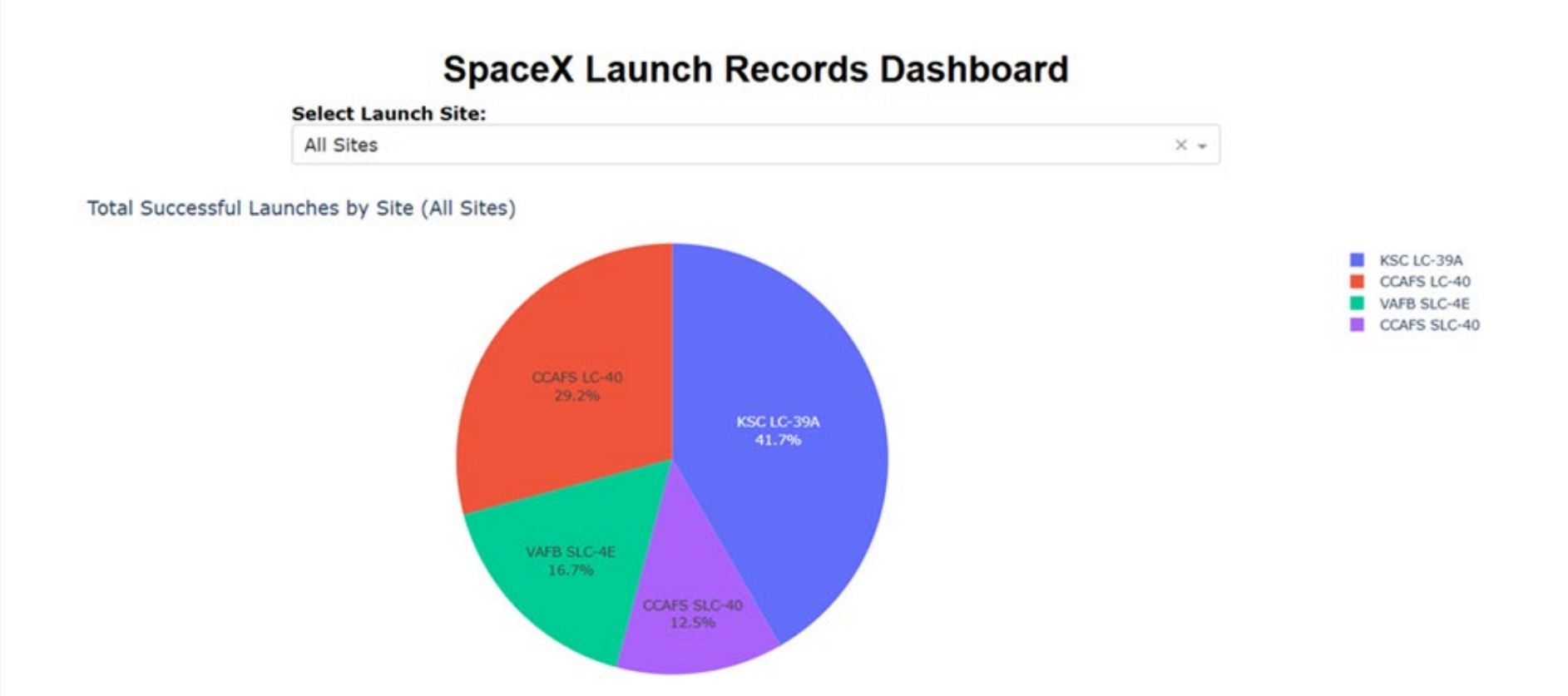
Proximities



- Explain the important elements and findings on the screenshot

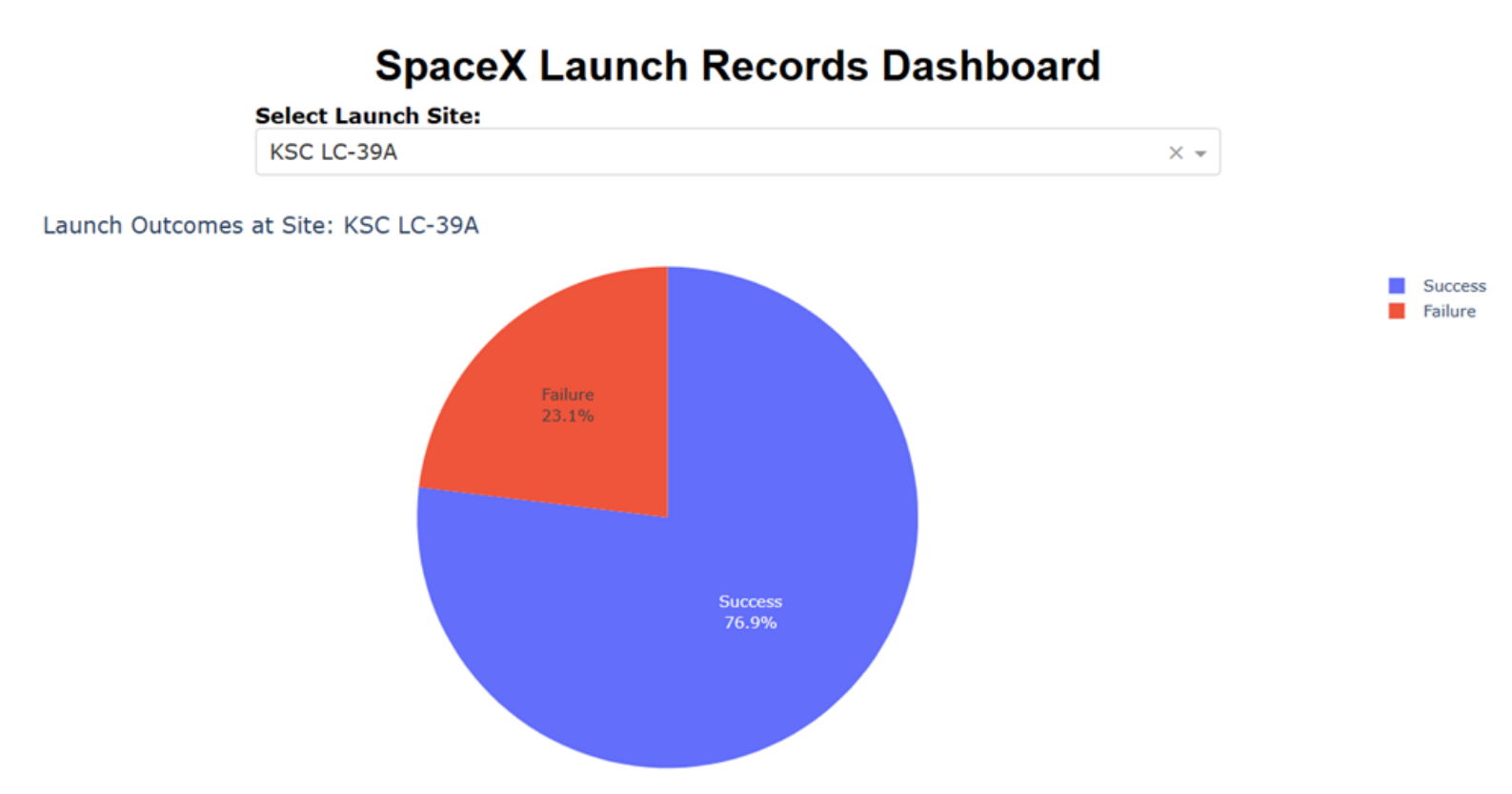

Section 4
Build a Dashboard with Plotly Dash

Total Successful Launches Pie Chart



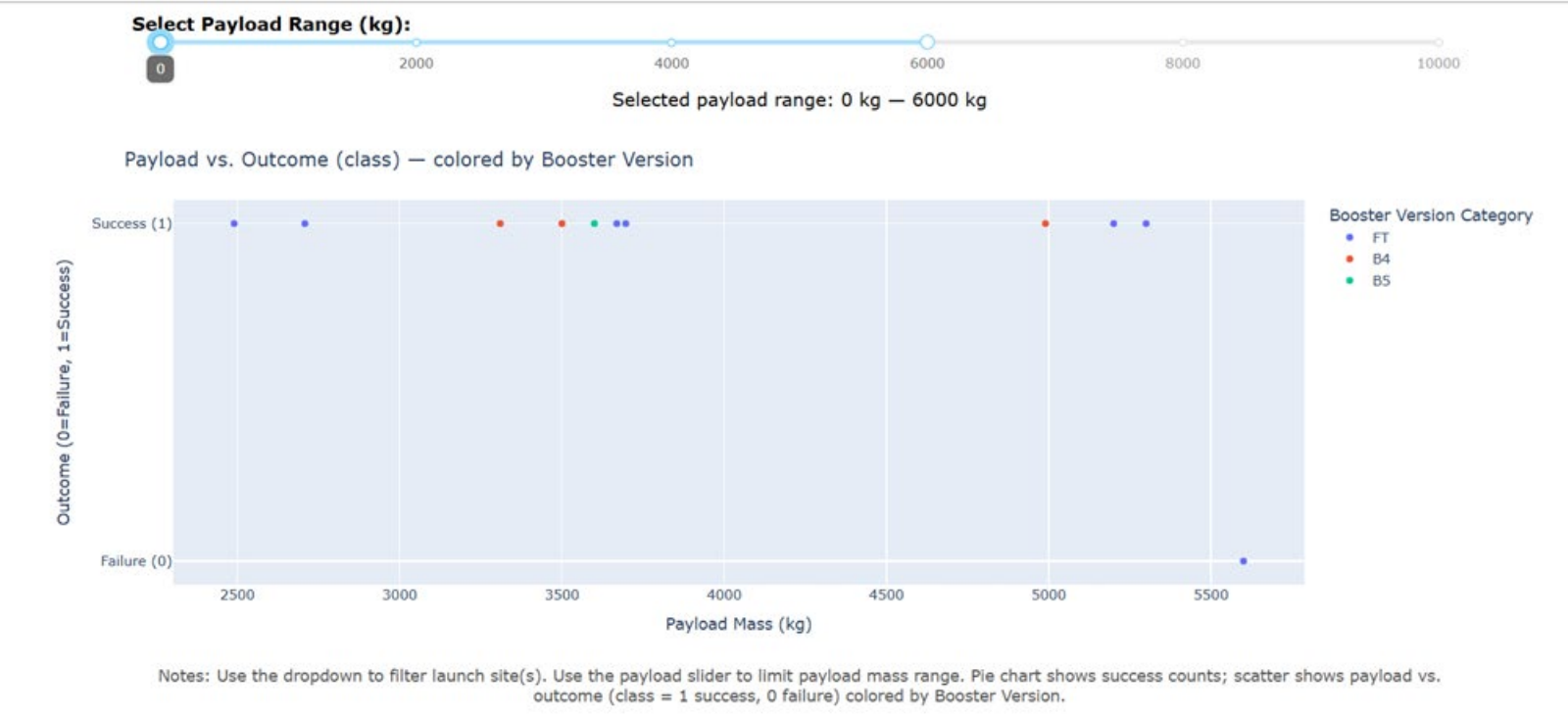
Explain the important elements and findings on the screenshot

Highest Launch Ratio



- Explain the important elements and findings on the screenshot

Payload Vs Outcome



- Explain the important elements and findings on the screenshot, such as which payload range or booster version have the largest success rate, etc.

Section 5

Predictive Analysis (Classification)

Classification Accuracy

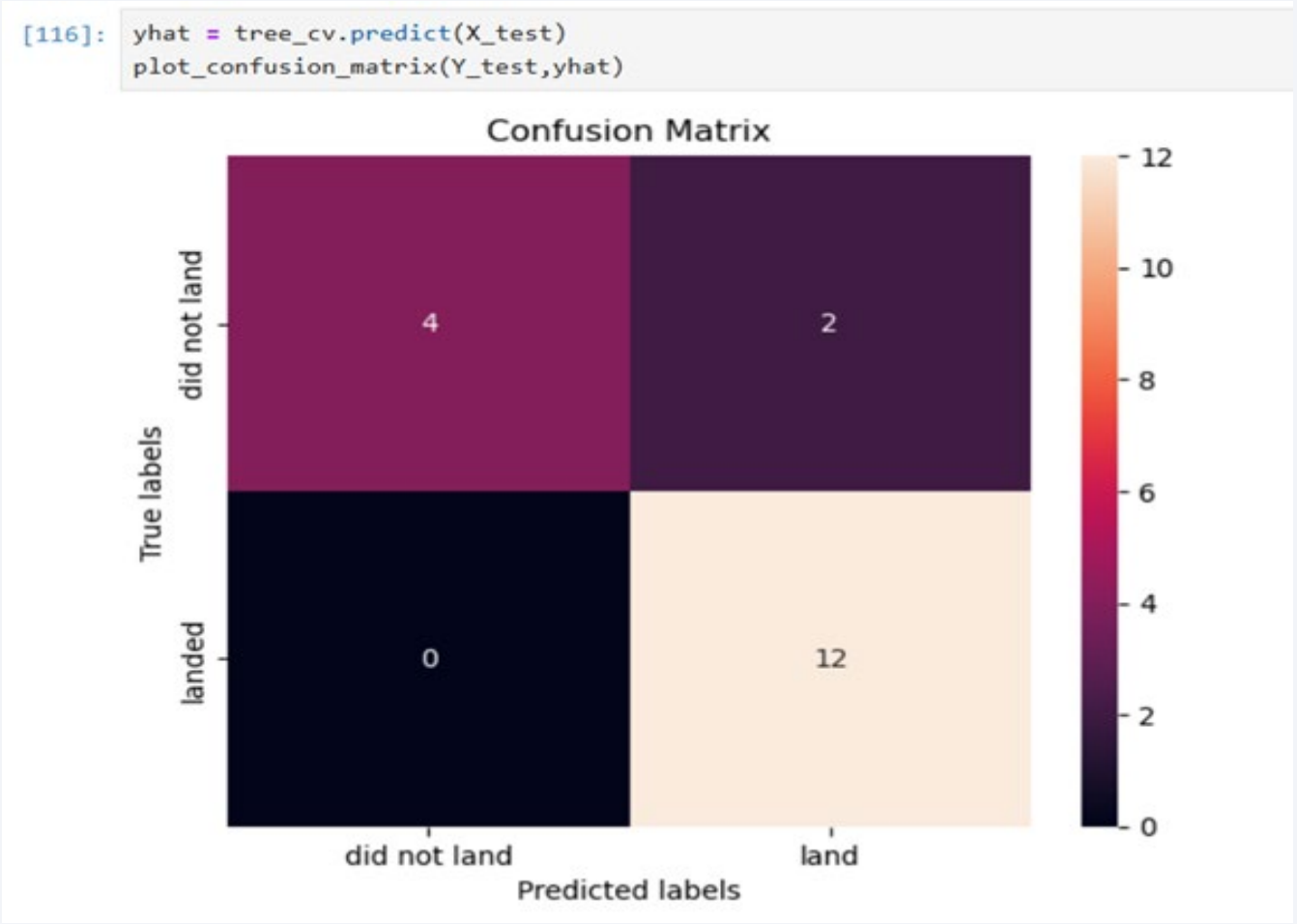
- Visualize the built model accuracy for all built classification models, in a bar chart

- Find which model has the highest classification accuracy

Confusion Matrix

- Confusion matrix of the best performing model with an explanation

```
[116]: yhat = tree_cv.predict(X_test)
       plot_confusion_matrix(Y_test,yhat)
```



Conclusions

- Point 1
- Point 2
- Point 3
- Point 4
- …

Appendix

- Include any relevant assets like Python code snippets, SQL queries, charts, Notebook outputs, or data sets that you may have created during this project

Thank you!