Winning Space Race with Data Science

S Shin

Jul 03 2024

Outline

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

Executive Summary

- Summary of methodologies

Performed data wrangling, EDA, and built predictive models using machine learning techniques.

- Summary of all results

Identified factors influencing launch success rates.

Developed interactive dashboards for data visualization.

Introduction

- Project background and context
- This project aims to analyze SpaceX launch data to understand the factors affecting launch success and predict future launches.

- Problems you want to find answers

Falcon 9 and its achievements and future

Section 1

Methodology

Methodology

Executive Summary

- Data collection methodology:

Data Sources:SpaceX API, Public datasets

- Perform data wrangling
    - Loaded datasets using pandas.Cleaned and transformed data for analysis.Merged and integrated multiple data sources.
- Perform exploratory data analysis (EDA) using visualization and SQL
- Perform interactive visual analytics using Folium and Plotly Dash
- Perform predictive analysis using classification models
    - How to build, tune, evaluate classification models

Data Collection

- Models Used:Logistic Regression
- Support Vector Machine (SVM)
- Decision Tree
- K-Nearest Neighbors (KNN)
- Grid Search and Cross-Validation:Used GridSearchCV for hyperparameter tuning.
- Evaluated models using cross-validation.

Data Collection – SpaceX API

- Present your data collection with SpaceX REST calls using key phrases and flowcharts

- Add the GitHub URL of the completed SpaceX API calls notebook (must include completed code cell and outcome cell), as an external reference and peer-review purpose

https://github.com/hoonbravo/DS-Certificate

Data Collection - Scraping

- Present your web scraping process using key phrases and flowcharts

- Add the GitHub URL of the completed web scraping notebook, as an external reference and peer-review purpose

https://github.com/hoonbravo/DS-Certificate

Data Wrangling

- Describe how data were processed
- You need to present your data wrangling process using key phrases and flowcharts
- Add the GitHub URL of your completed data wrangling related notebooks, as an external reference and peer-review purpose
- https://github.com/hoonbravo/DS-Certificate

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
- https://github.com/hoonbravo/DS-Certificate

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
- https://github.com/hoonbravo/DS-Certificate

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
- Add the GitHub URL of your completed interactive map with Folium map, as an external reference and peer-review purpose
- https://github.com/hoonbravo/DS-Certificate

Build a Dashboard with Plotly Dash

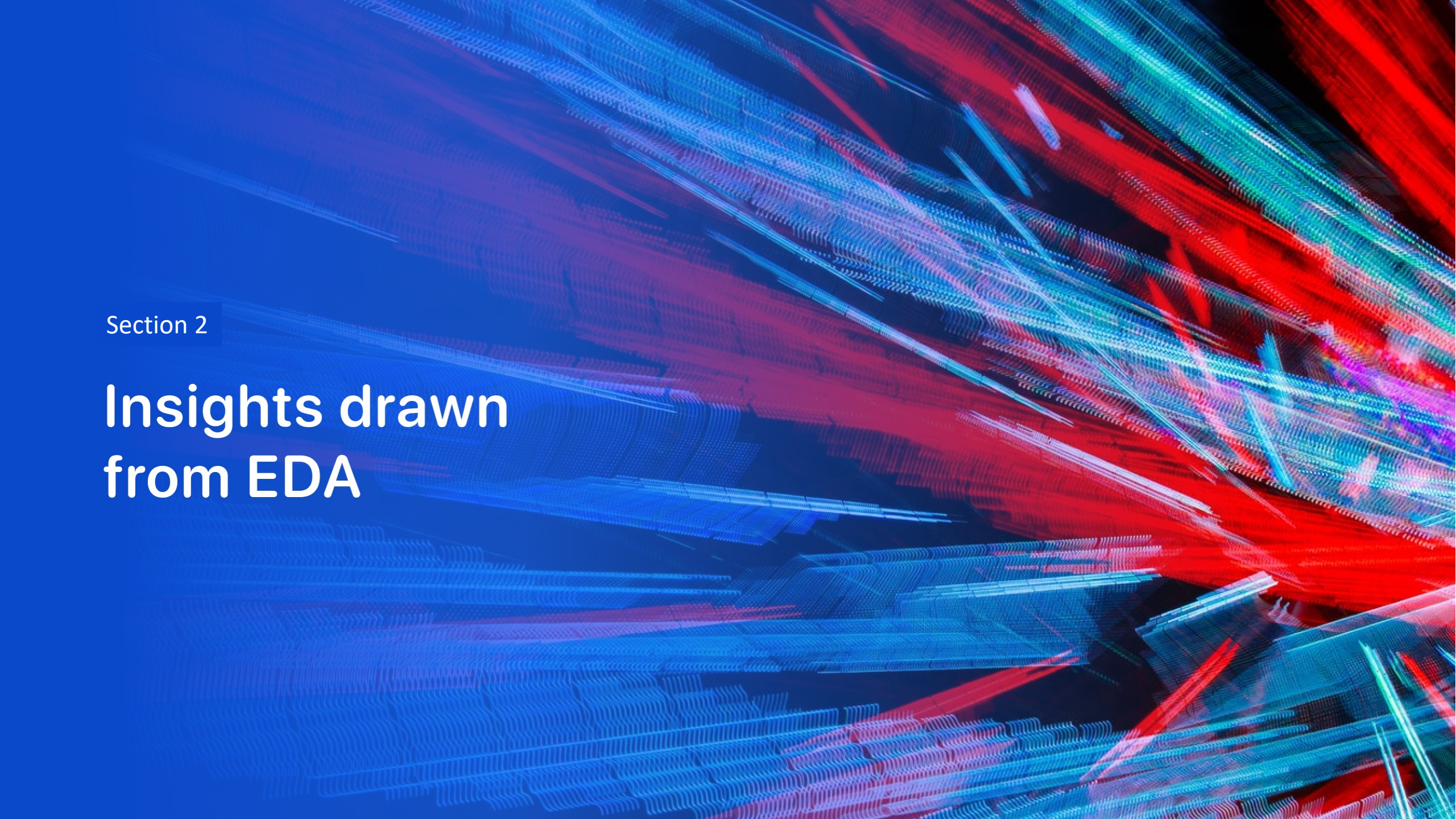
- Summarize what plots/graphs and interactions you have added to a dashboard
- Explain why you added those plots and interactions
- Add the GitHub URL of your completed Plotly Dash lab, as an external reference and peer-review purpose
- https://github.com/hoonbravo/DS-Certificate

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
- https://github.com/hoonbravo/DS-Certificate

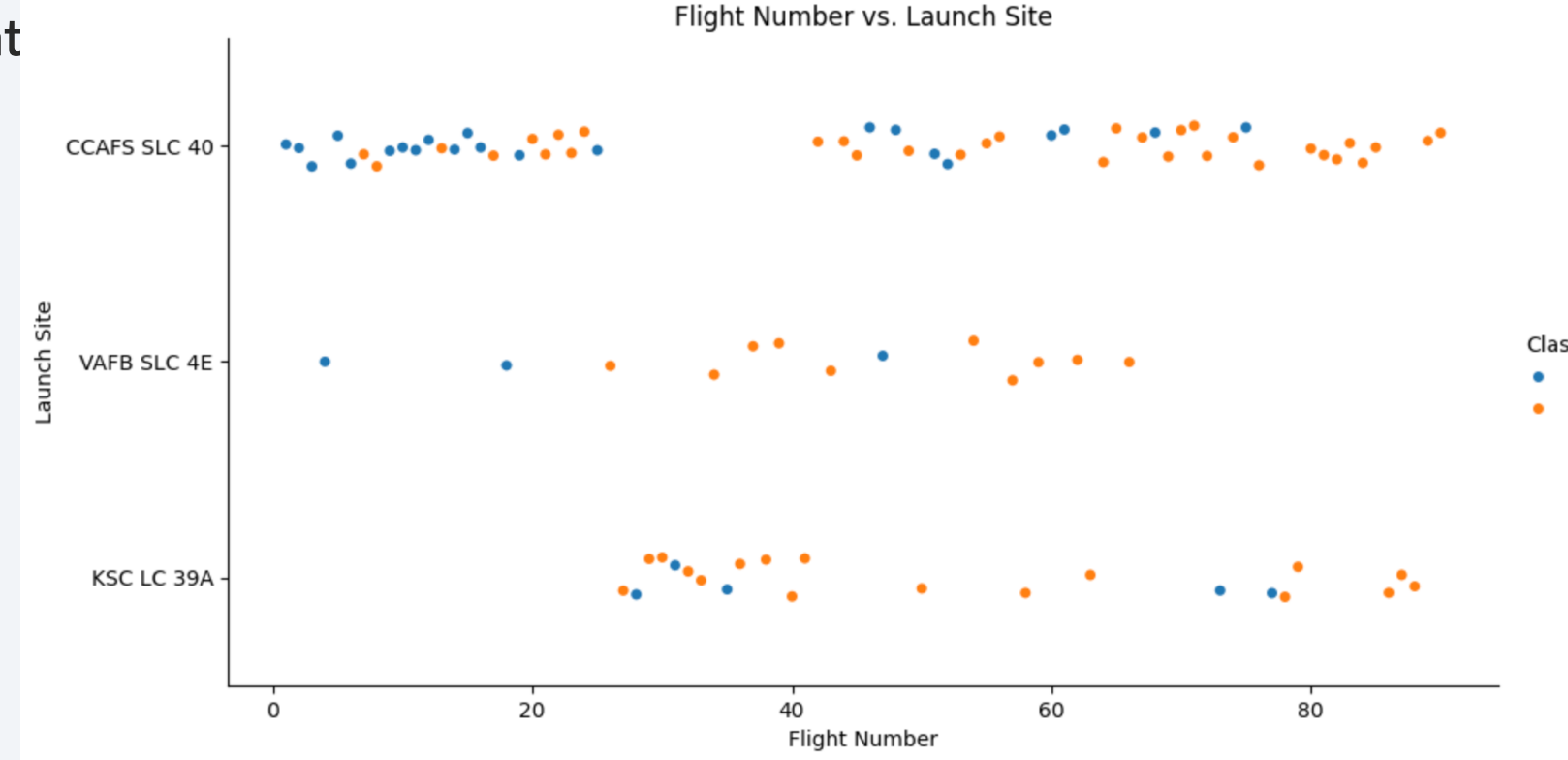
Results

- Exploratory data analysis results
- Interactive analytics demo in screenshots
- Predictive analysis results
- https://github.com/hoonbravo/DS-Certificate


Section 2
Insights drawn from EDA

Flight Number vs. Launch Site

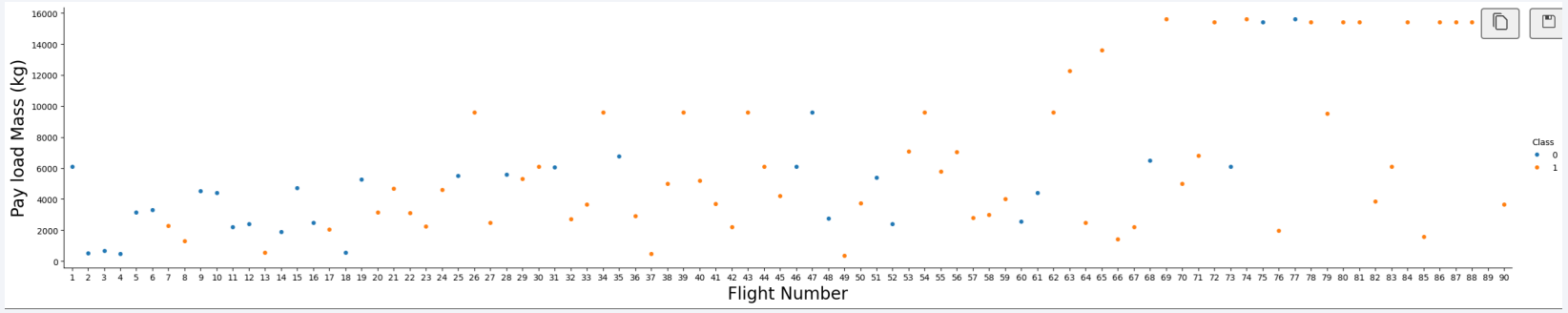
- Show a scatter plot of Flight Number vs. Launch Site
- Show the screenshot of the scatter plot with explanations



Payload vs. Launch Site

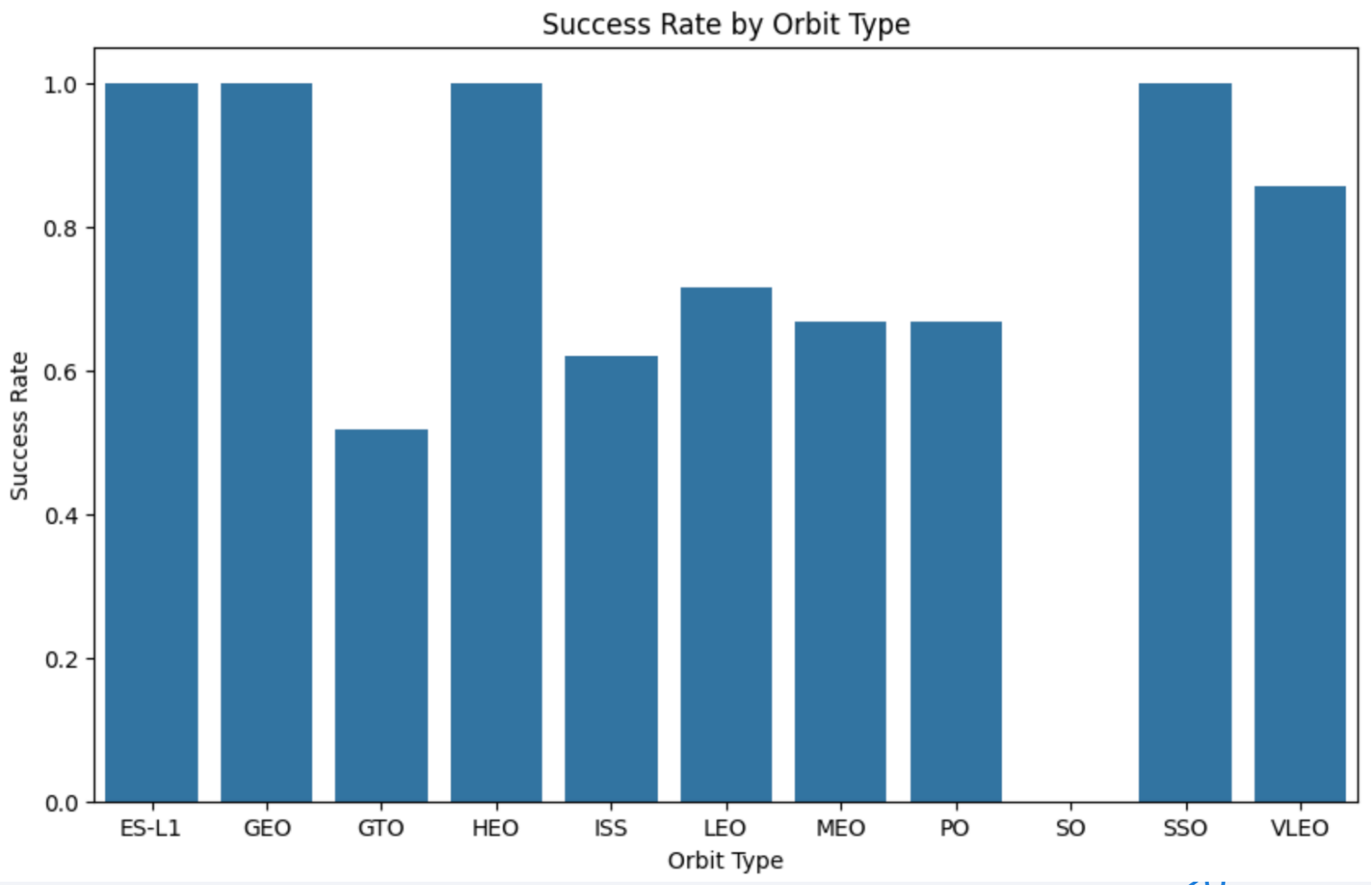
- Show a scatter plot of Payload vs. Launch Site

- Show the screenshot of the scatter plot with explanations



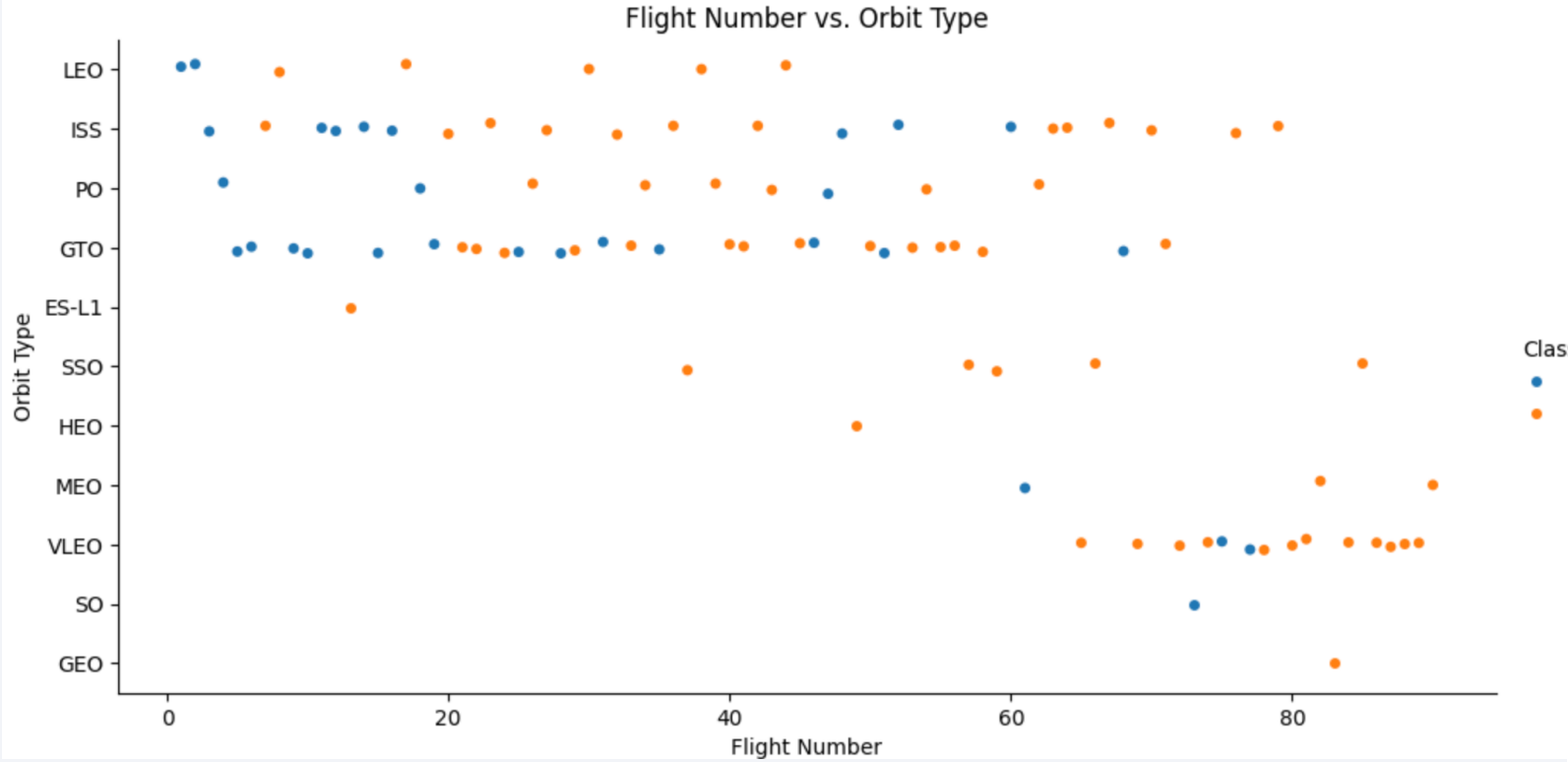
Success Rate vs. Orbit Type

- Show a bar chart for the success rate of each orbit type
- Show the screenshot of the scatter plot with explanations



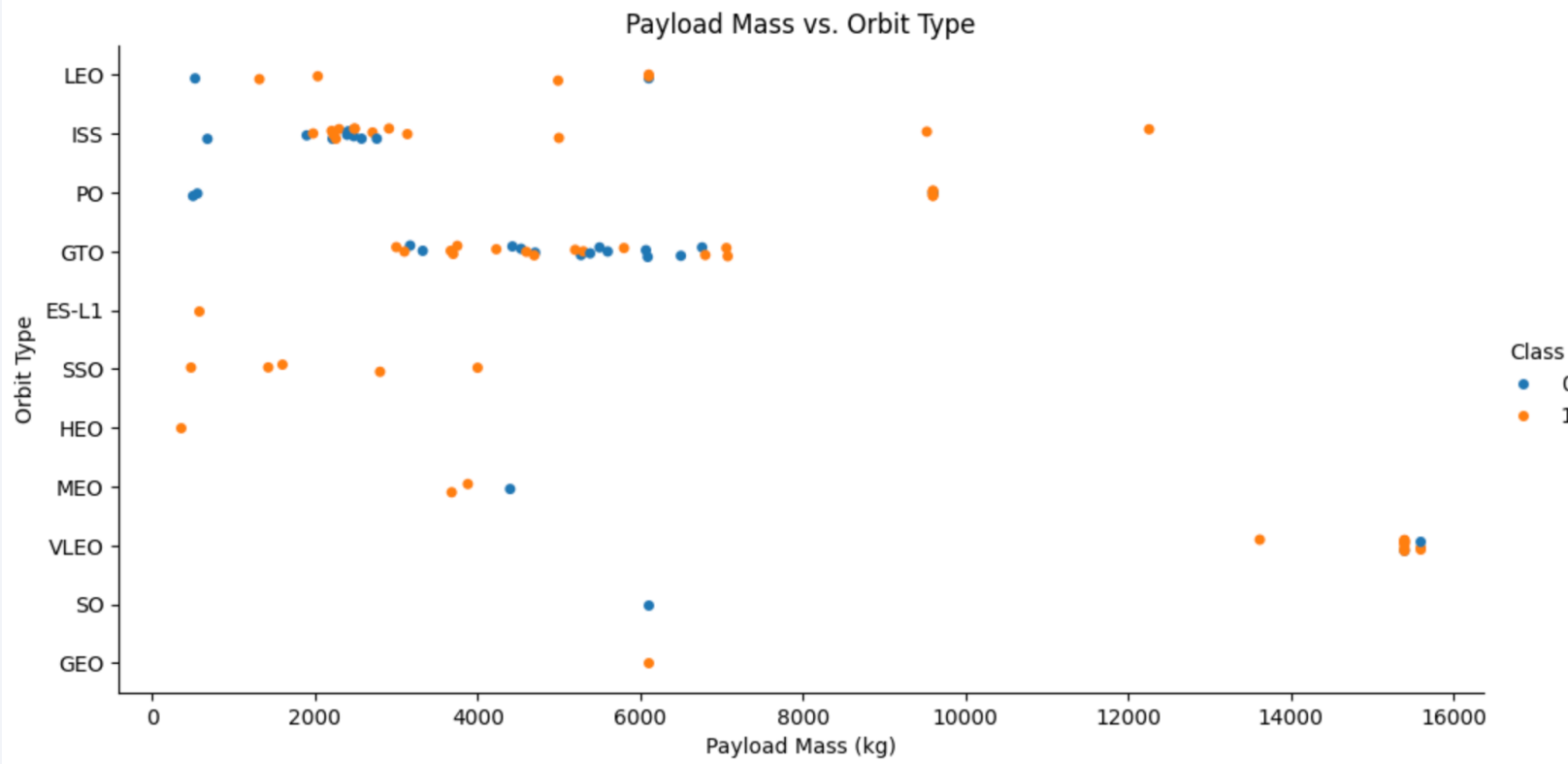
Flight Number vs. Orbit Type

- Show a scatter point of Flight number vs. Orbit type
- Show the screenshot of the scatter plot with explanations



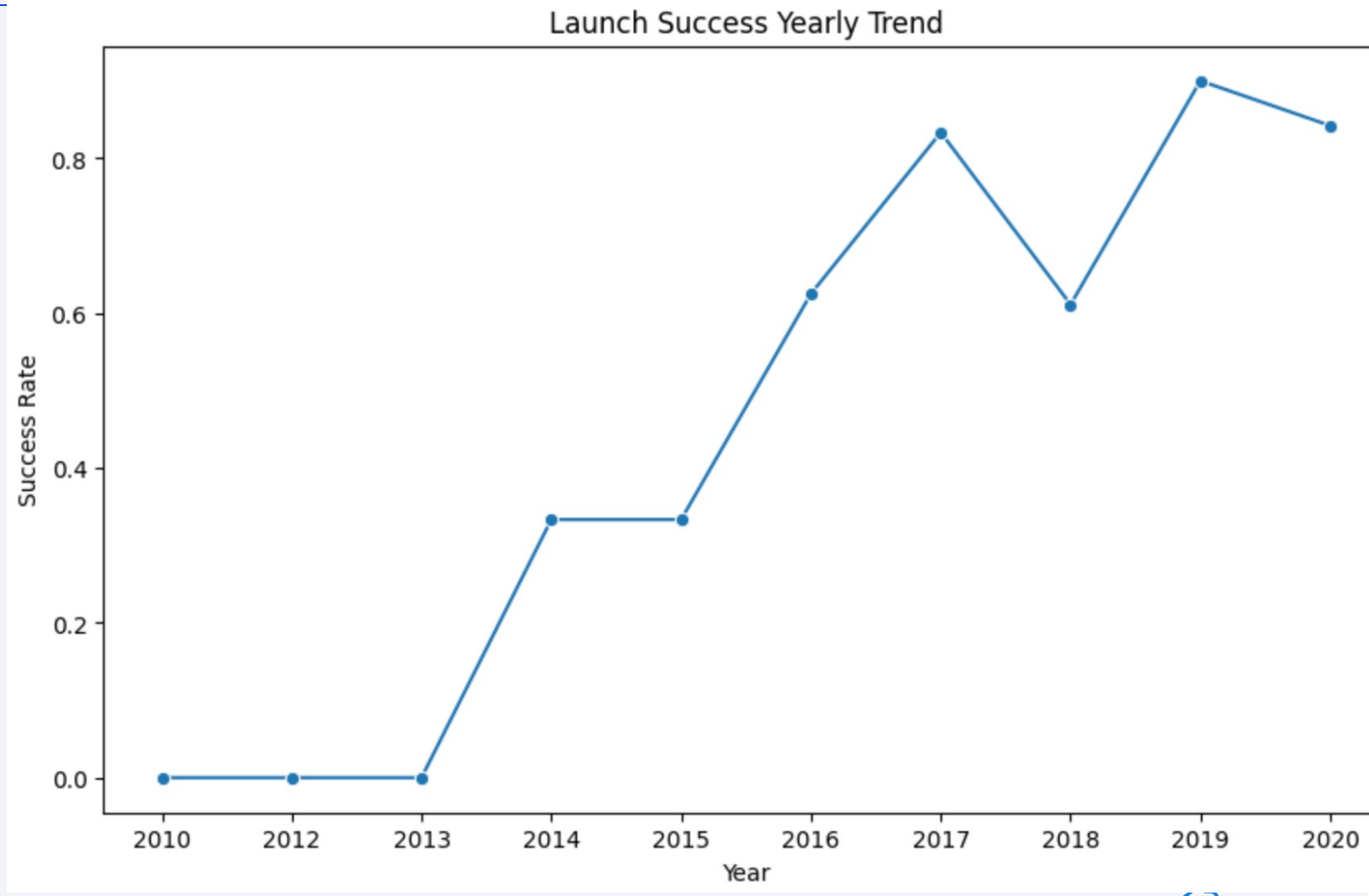
Payload vs. Orbit Type

- 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
0	CCAFS LC-40
1	VAFB SLC-4E
2	KSC LC-39A
3	CCAFS SLC-40

Launch Site Names Begin with 'CCA'

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

	Launch_Site	launch_count
0	CCAFS LC-40	26
1	CCAFS SLC-40	34
2	KSC LC-39A	25
3	VAFB SLC-4E	16

Total Payload Mass

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

	total_payload_mass
0	45596

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

	avg_payload_mass
0	2928.4

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

	first_successful_landing
0	2015-12-22

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

	Booster_Version
0	F9 FT B1022
1	F9 FT B1026
2	F9 FT B1021.2
3	F9 FT B1031.2

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

	Landing_Outcome	count
0	Controlled (ocean)	5
1	Failure	3
2	Failure (drone ship)	5
3	Failure (parachute)	2
4	No attempt	21
5	No attempt	1
6	Precluded (drone ship)	1
7	Success	38
8	Success (drone ship)	14
9	Success (ground pad)	9
10	Uncontrolled (ocean)	2

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

	Booster_Version
0	F9 B5 B1048.4
1	F9 B5 B1049.4
2	F9 B5 B1051.3
3	F9 B5 B1056.4
4	F9 B5 B1048.5
5	F9 B5 B1051.4
6	F9 B5 B1049.5
7	F9 B5 B1060.2
8	F9 B5 B1058.3
9	F9 B5 B1051.6
10	F9 B5 B1060.3
11	F9 B5 B1049.7

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

	Month	Landing_Outcome	Booster_Version	Launch_Site
0	01	Failure (drone ship)	F9 v1.1 B1012	CCAFS LC-40
1	04	Failure (drone ship)	F9 v1.1 B1015	CCAFS LC-40

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

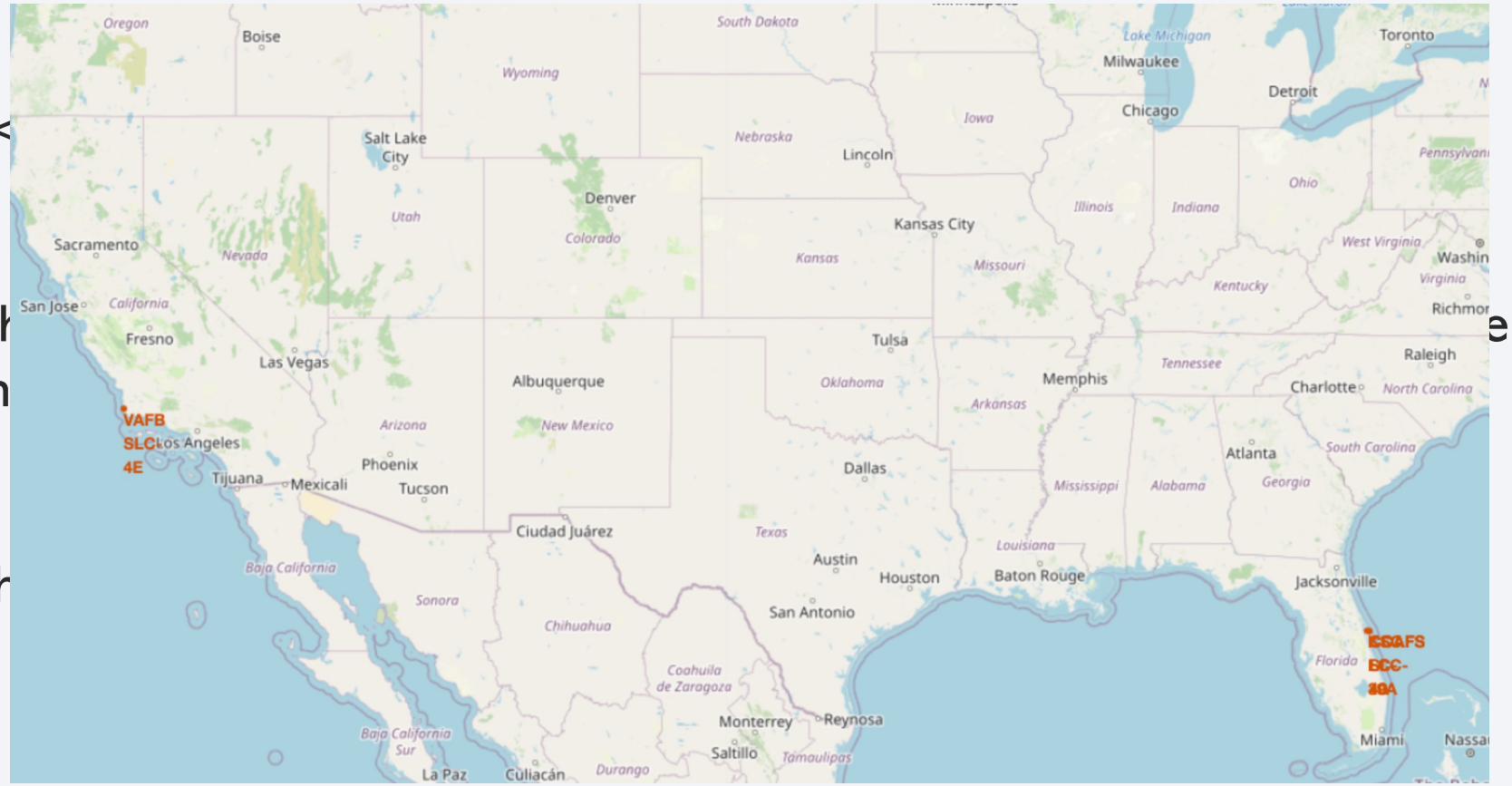
- Present your query result with a short explanation here

	Landing_Outcome	Outcome_Count
0	Failure (drone ship)	5
1	Success (ground pad)	3


Section 3
Launch Sites
Proximities Analysis

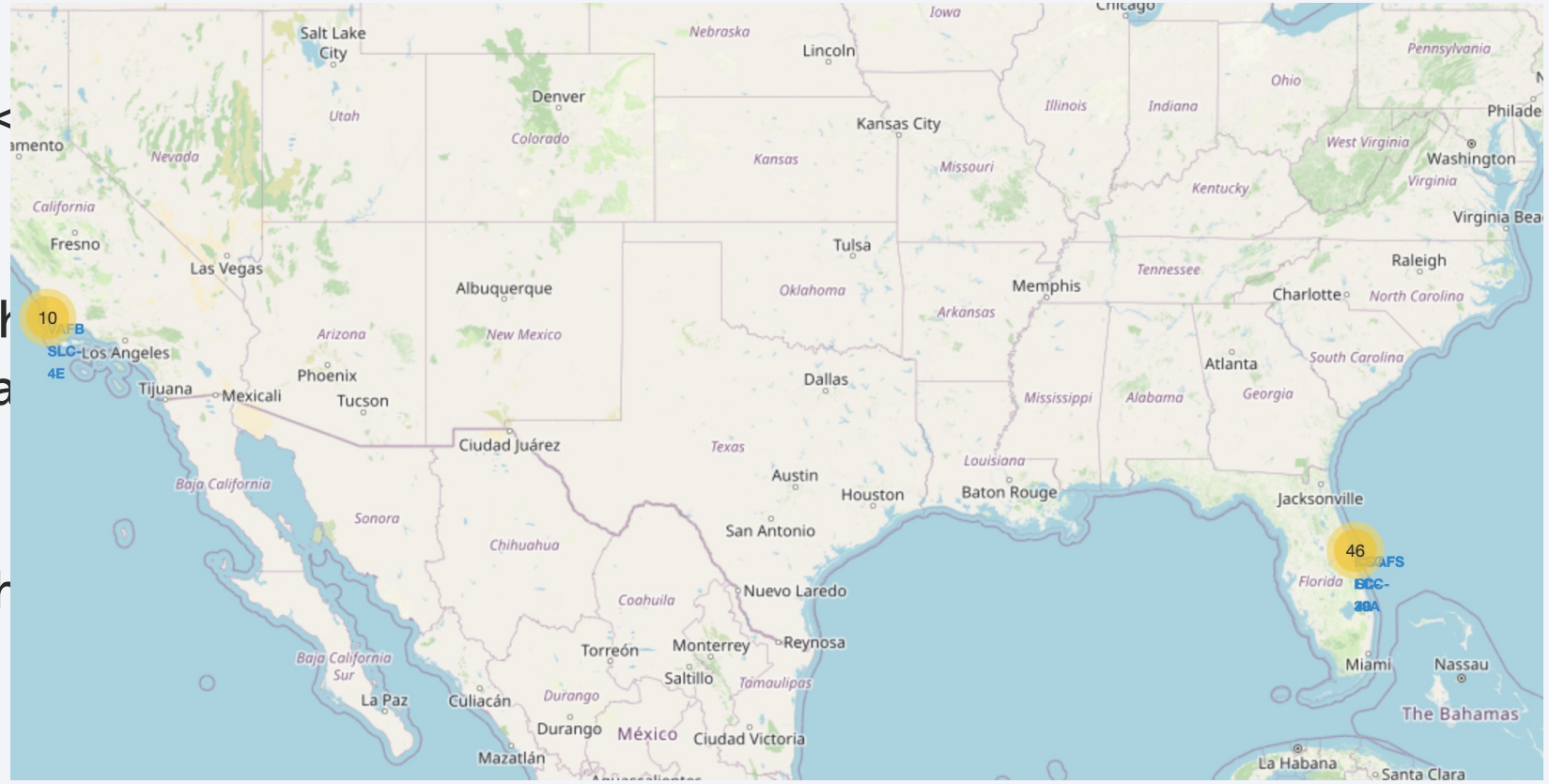
<Folium Map Screenshot 1>

- Replace <
- Explore th
  all launch
- Explain th



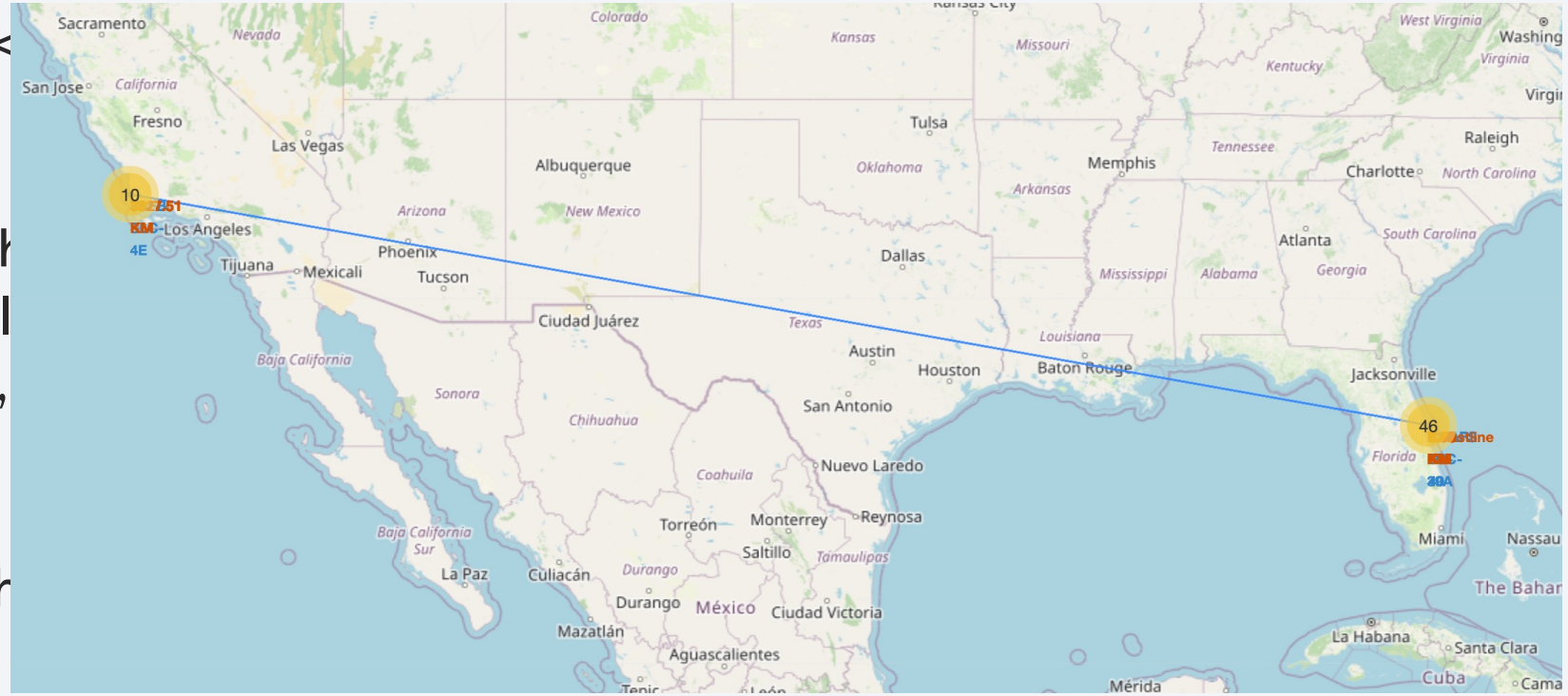
<Folium Map Screenshot 2>

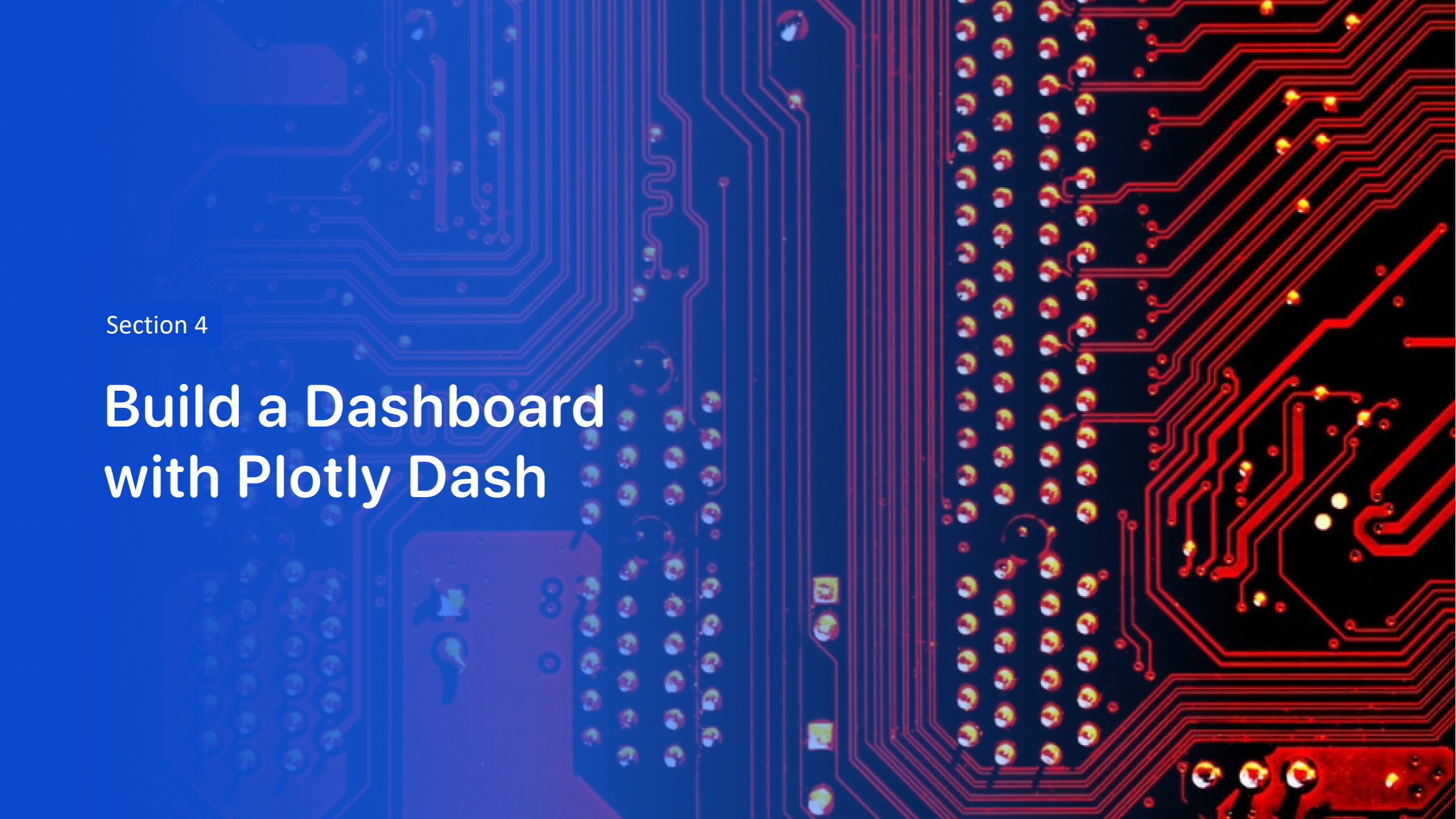
- Replace <
- Explore th
  labeled la
- Explain th



<Folium Map Screenshot 3>

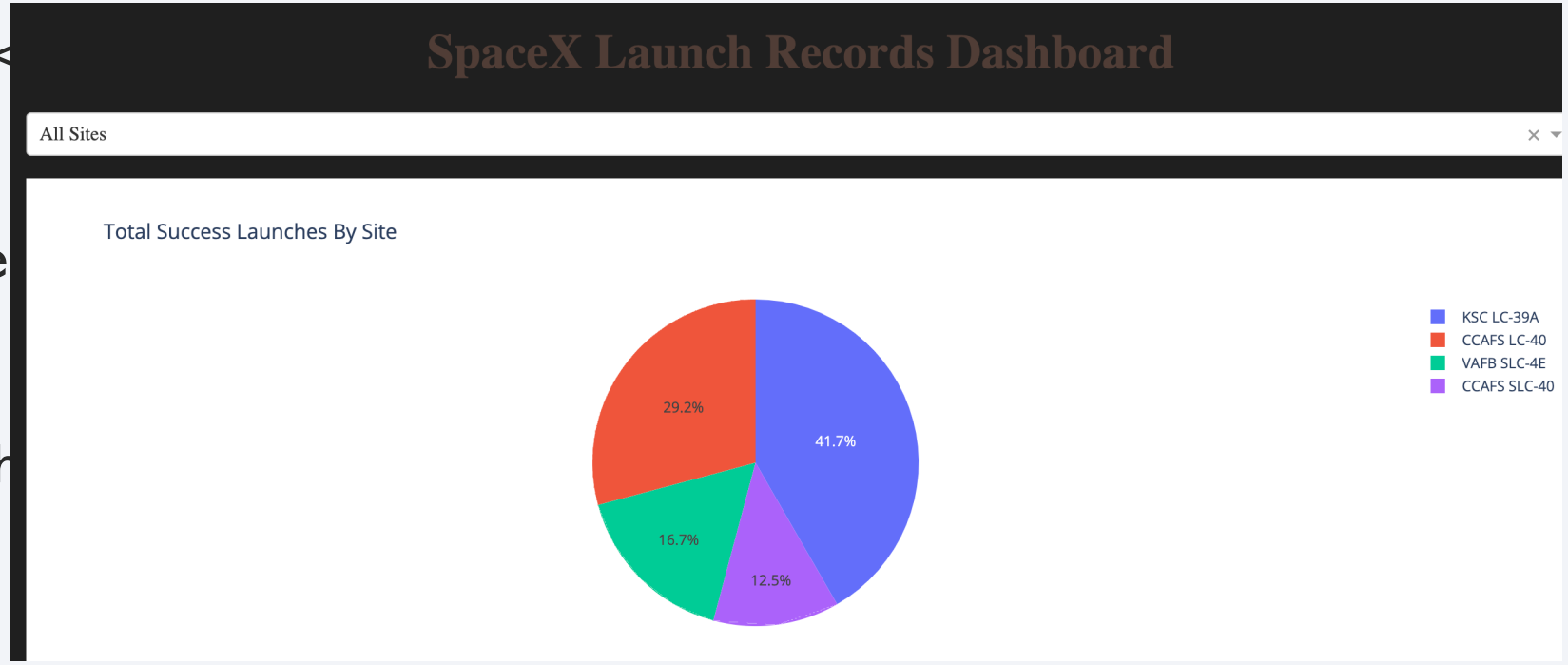
- Replace <
- Explore th
  selected l
  coastline,
- Explain th




Section 4
Build a Dashboard with Plotly Dash

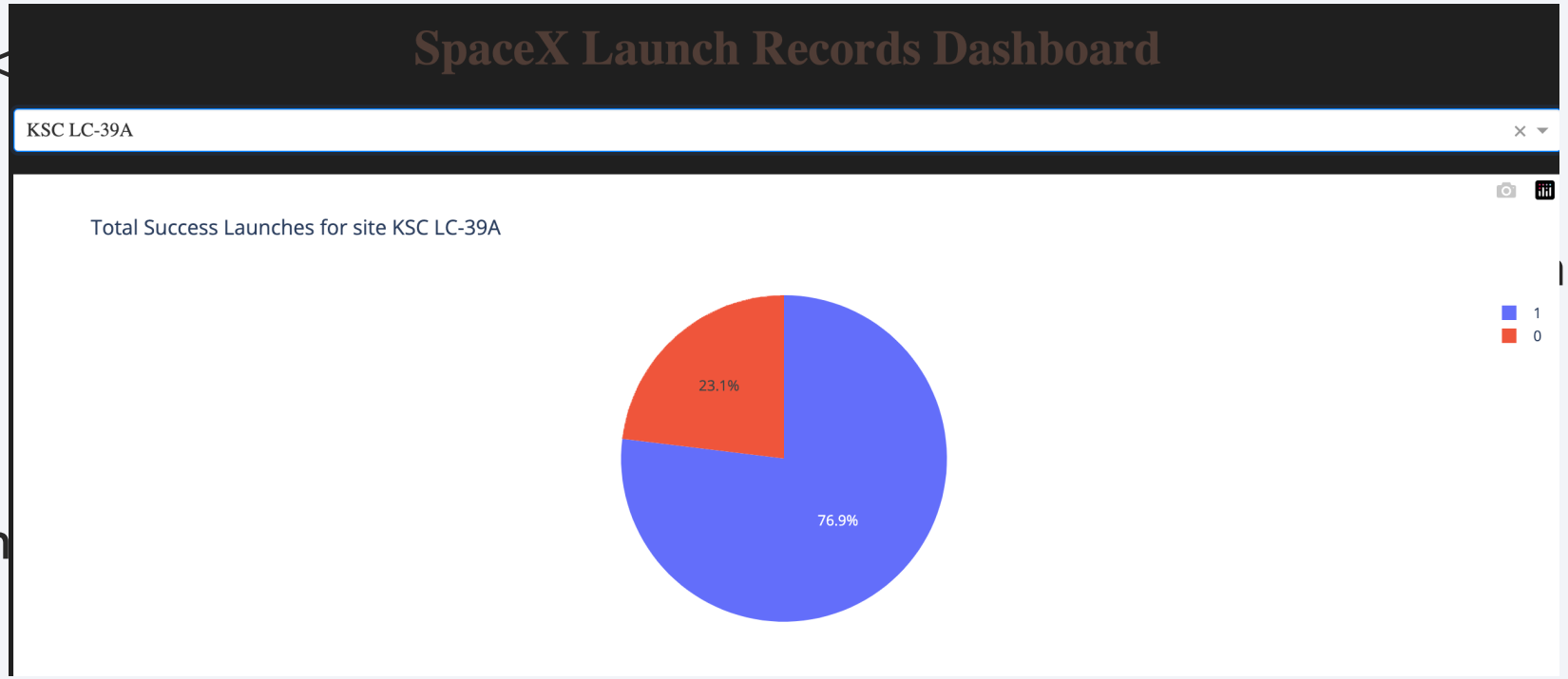
<Dashboard Screenshot 1>

- Replace <
- Show the
- Explain th



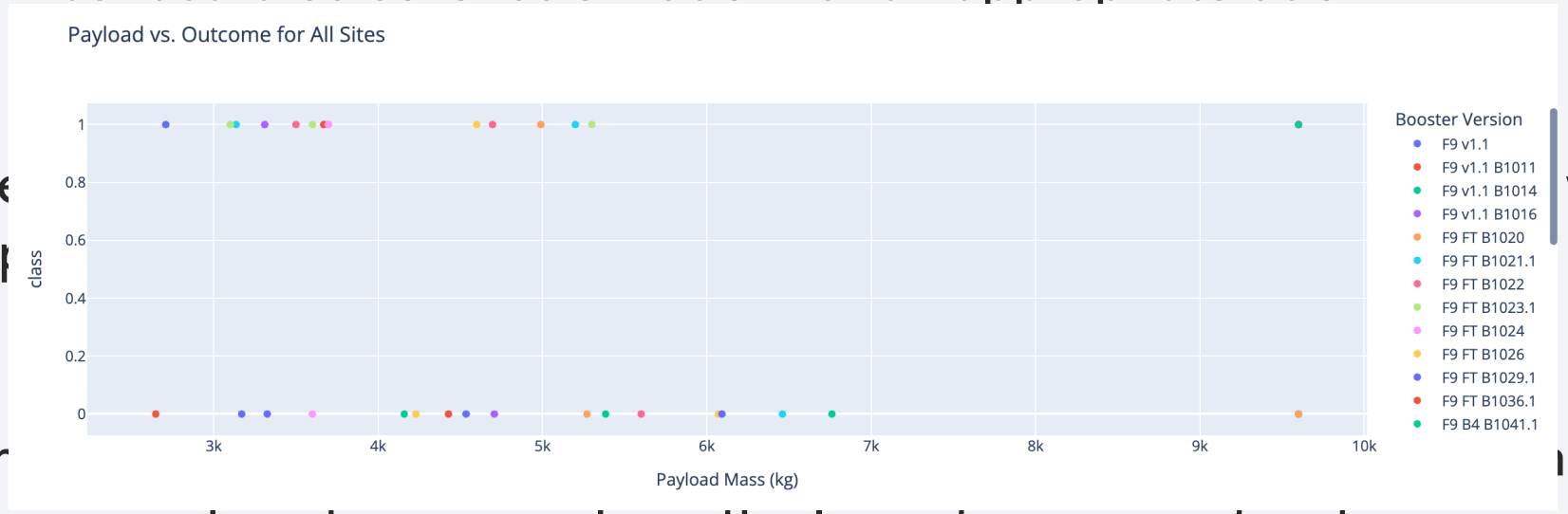
<Dashboard Screenshot 2>

- Replace <
- Show the success ratio
- Explain th



<Dashboard Screenshot 3>

- Replace <Dashboard screenshot 3> title with an appropriate title
- Show scre with different p
- Explain th ich payload range or booster version have the largest success rate, etc.



Section 5

Predictive Analysis (Classification)

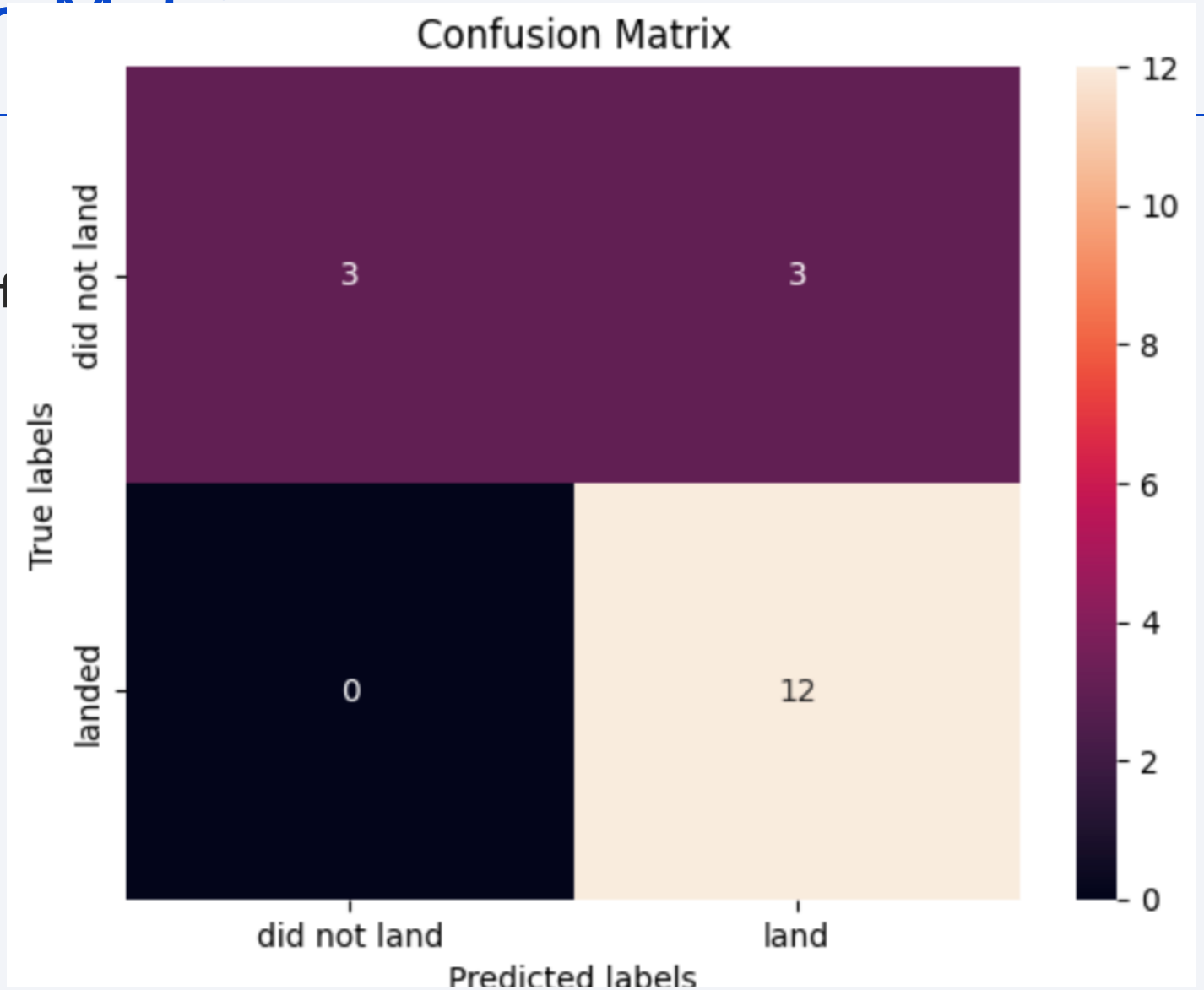
Classification Accuracy

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

- Find which model has the highest classification accuracy

Confusion Matrix

- Show the conf
  explanation



Conclusions

- **Summary of Findings:**Successfully analyzed SpaceX launch data.
- Developed models with high accuracy for predicting launch success.
- **Implications:**The insights gained can help improve future launch outcomes and operational efficiency.

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!