IBM Developer
SKILLS NETWORK

# Winning Space Race with Data Science

Muhammad Mubashir Mirza
April 17, 2023.

# Outline

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

# Executive Summary

- Summary of methodologies
- Summary of all results

# Introduction

- Project background and context
- Problems you want to find answers

Section 1

# Methodology

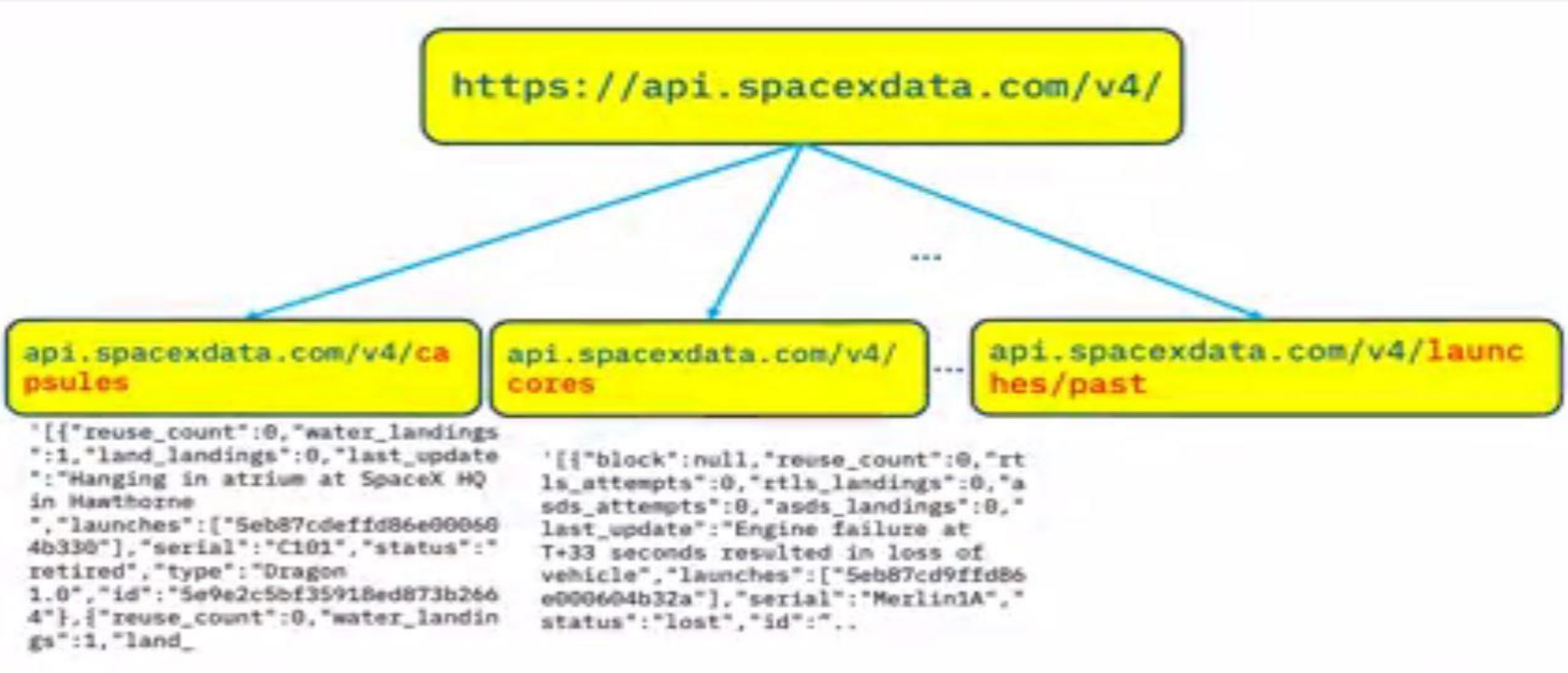
# Methodology

## Executive Summary

- Data collection methodology:
  - The data is collected by scraping the SpaceX wiki page.
- Perform data wrangling
  - The data is processedby removing unnecessary details, normalization and standardization.
- Perform exploratory data analysis (EDA) using visualization and SQL
- Perform interactive visual analytics using Folium and Plotly Dash
- Perform predictive analysis using classification models
  - Classification models are build using Logistic regression, KNN and SVM and Decision Trees. Their best parameters are calculated by Grid Search. Their results are evaluated by F1 Score and Confusion Matrix.

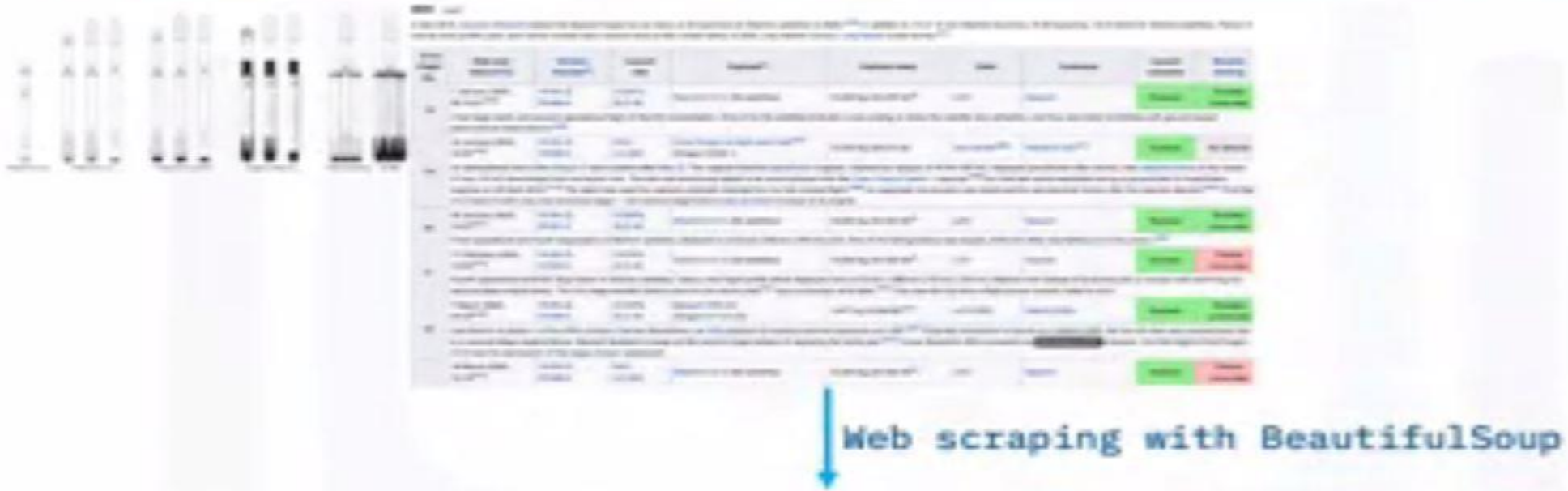
# Data Collection

- Describe how data sets were collected.
- You need to present your data collection process use key phrases and flowcharts




https://github.com/Mirza-Algo/DataScienceProjects/blob/main/Data%20Collection.ipynb

# Web scraping Falcon 9 Launch records




https://github.com/Mirza-Algo/DataScienceProjects/blob/main/WebScraping.ipynb

# Data Wrangling

https://github.com/Mirza-Algo/DataScienceProjects/blob/main/Data%20Wrangling.ipynb

# EDA with Data Visualization

https://github.com/Mirza-Algo/DataScienceProjects/blob/main/Data%20Visualization.ipynb

# EDA with SQL

https://github.com/Mirza-Algo/DataScienceProjects/blob/main/SQL%20EDA.ipynb

# Build an Interactive Map with Folium

https://github.com/Mirza-Algo/DataScienceProjects/blob/main/2.%20Folium%20Maps.ipynb

# Build a Dashboard with Plotly Dash

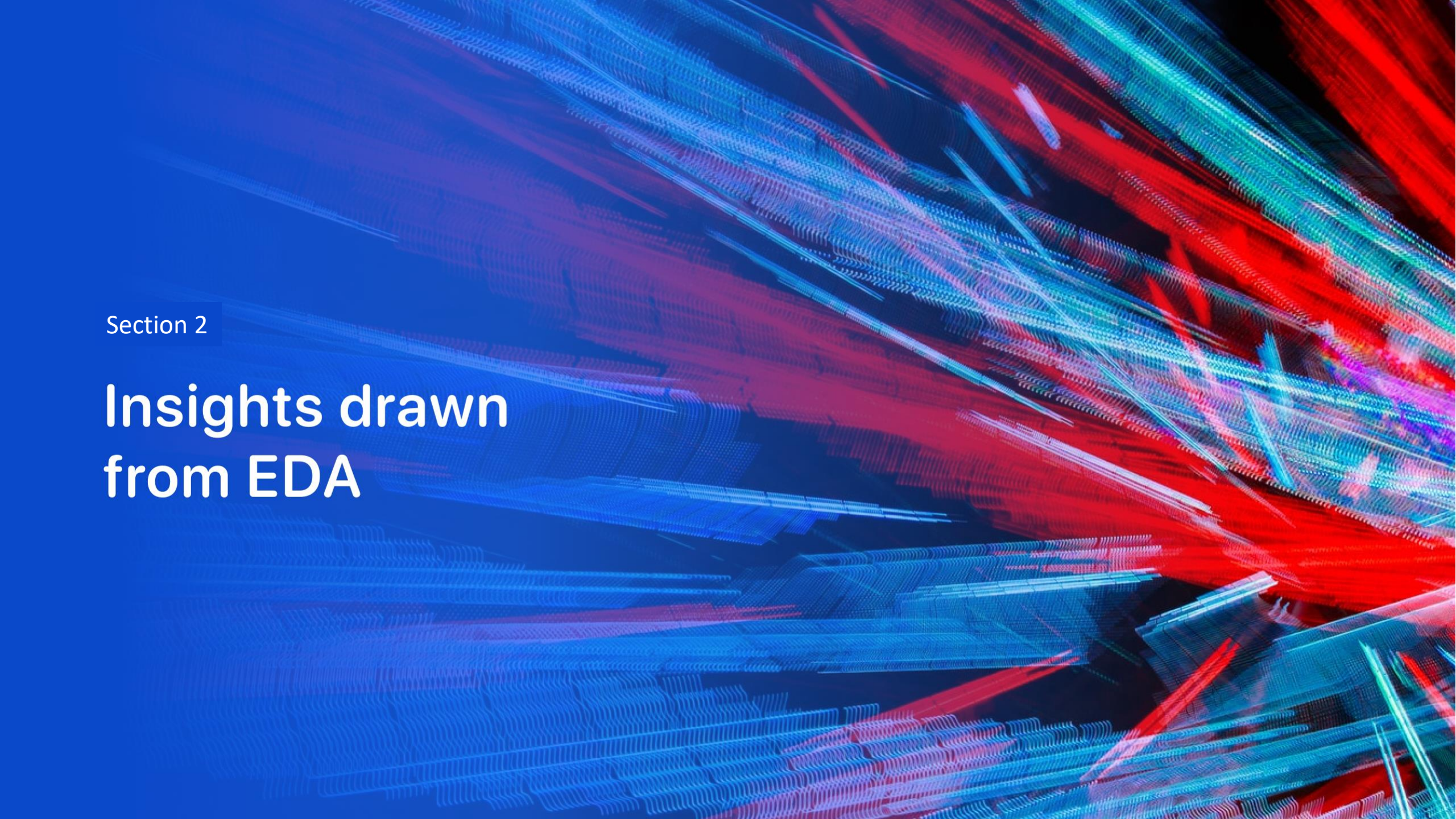
https://github.com/Mirza-Algo/DataScienceProjects/blob/main/3.%20Interactive%20Dashboard.ipynb

# Predictive Analysis (Classification)

https://github.com/Mirza-Algo/DataScienceProjects/blob/main/SpaceX_Machine_Learning_Prediction.ipynb

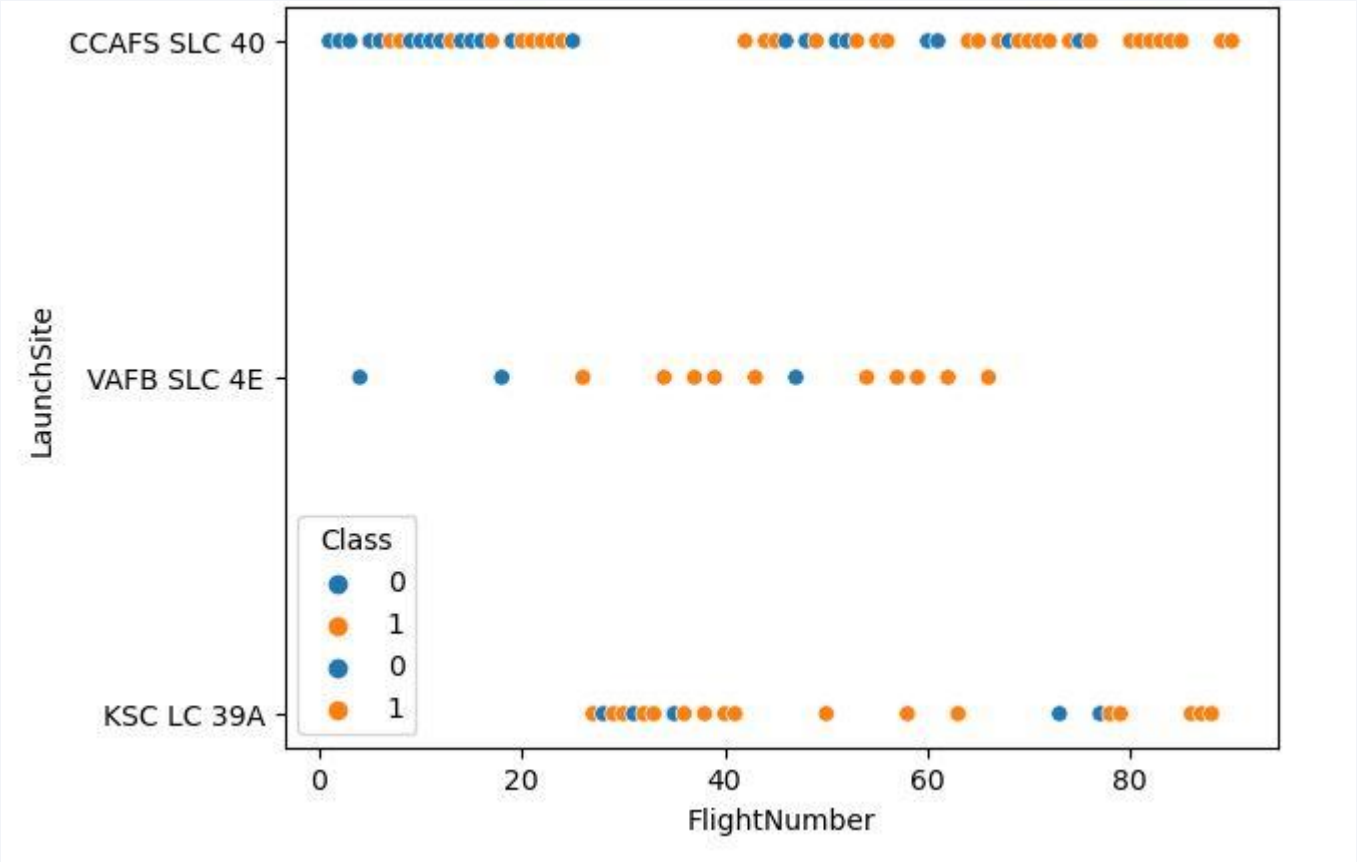
# Results

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


Section 2
Insights drawn from EDA

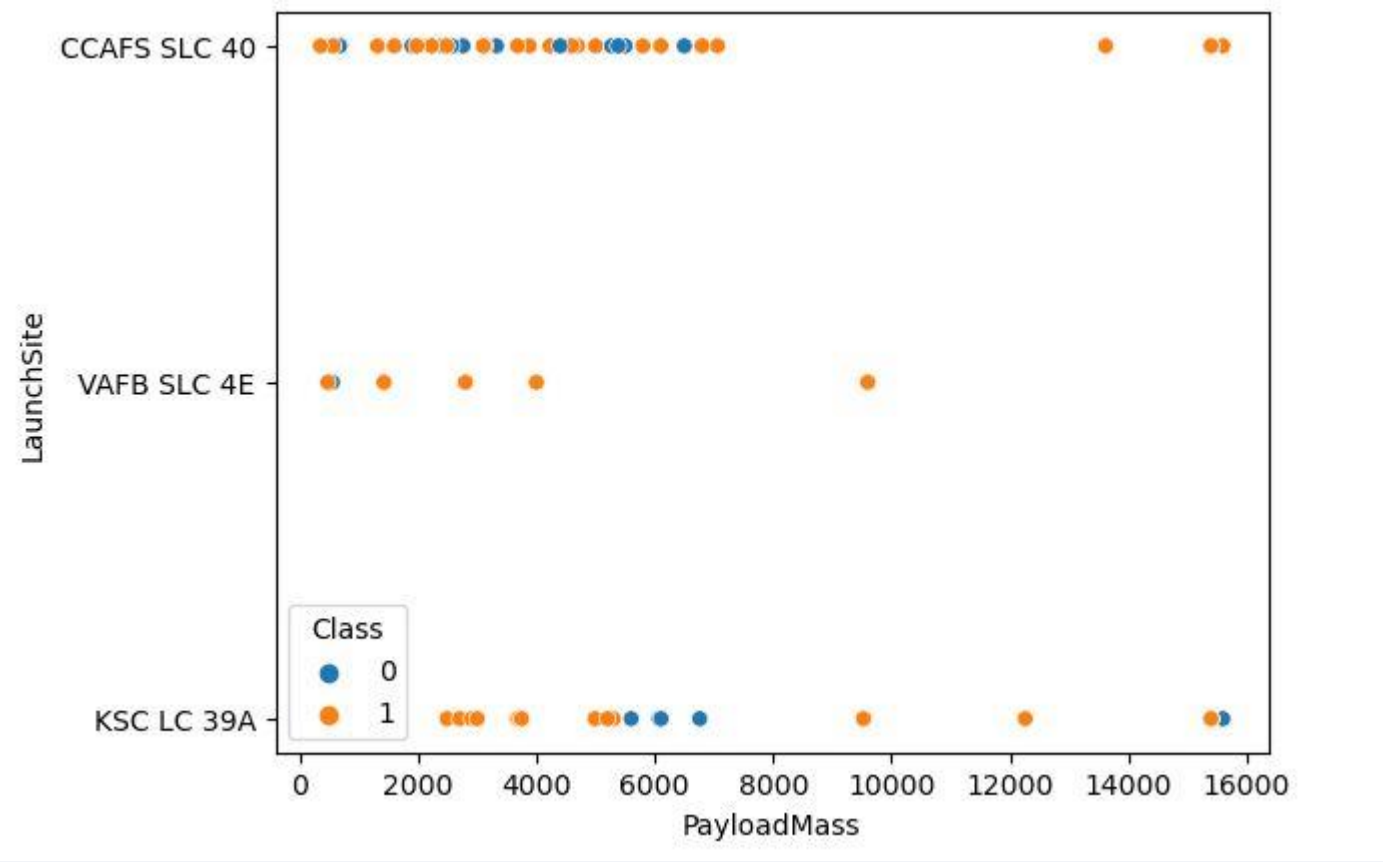
# Flight Number vs. Launch Site

- Scatter plot of Flight Number vs. Launch Site



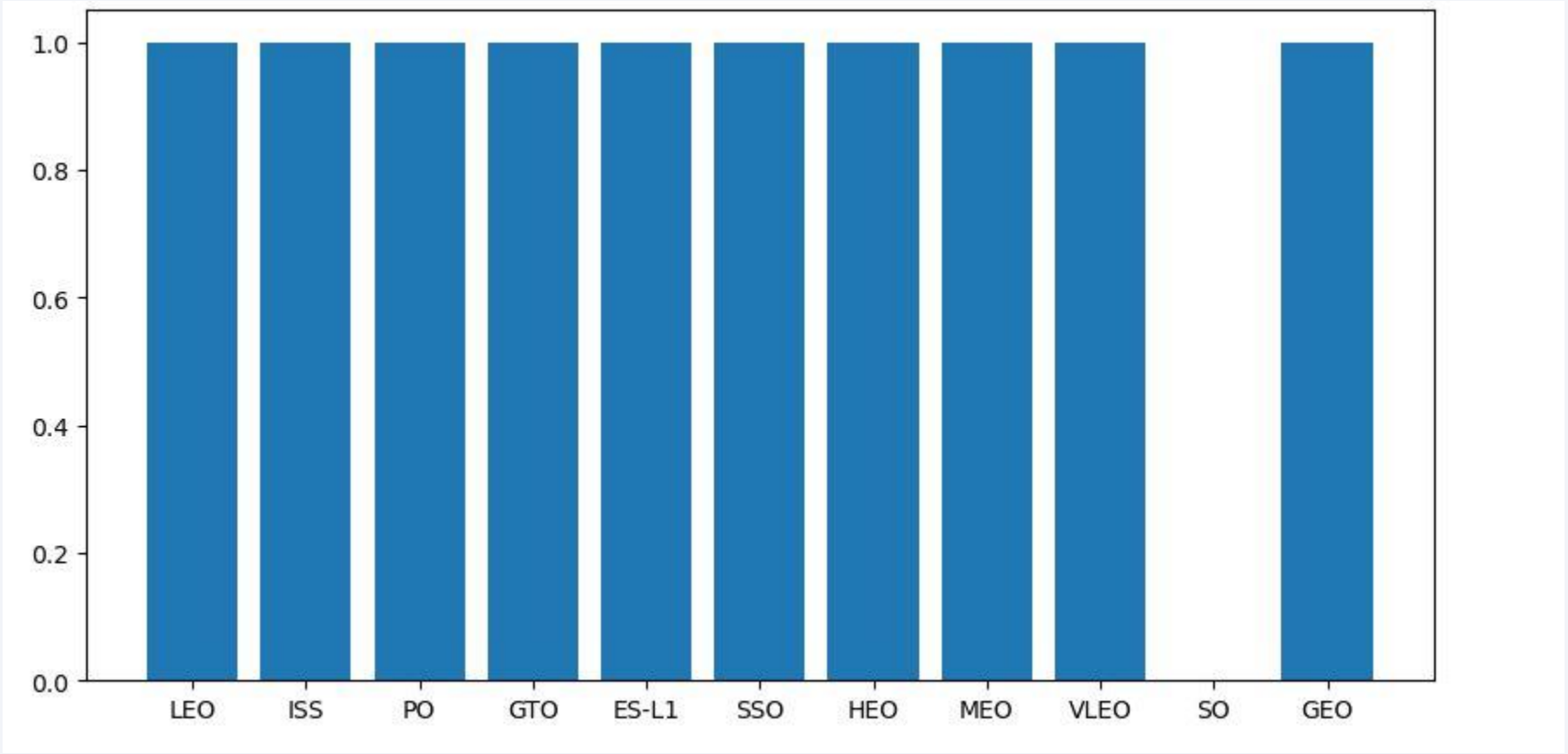
# Payload vs. Launch Site

- Scatter plot of Payload vs. Launch Site



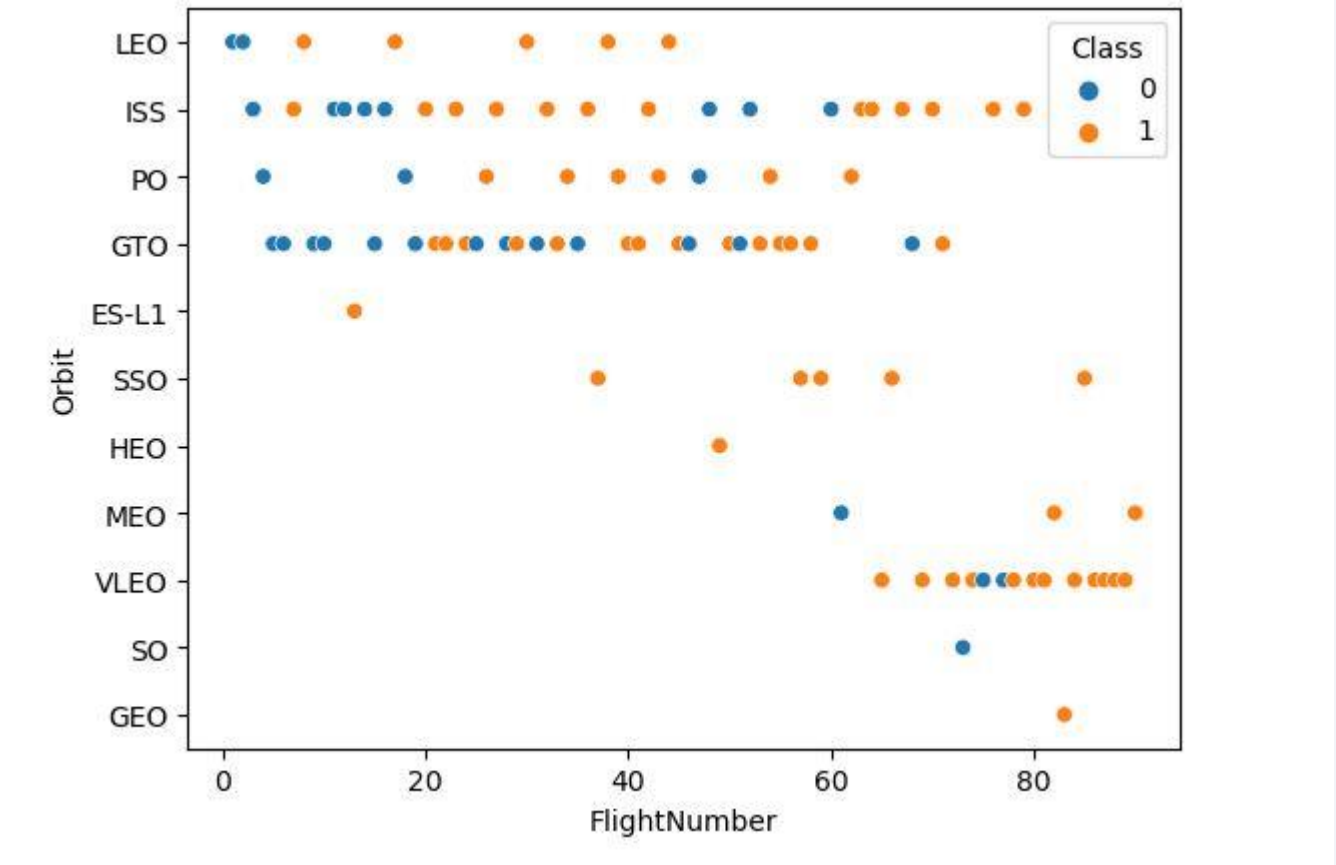
# Success Rate vs. Orbit Type

- Bar chart for the success rate of each orbit type



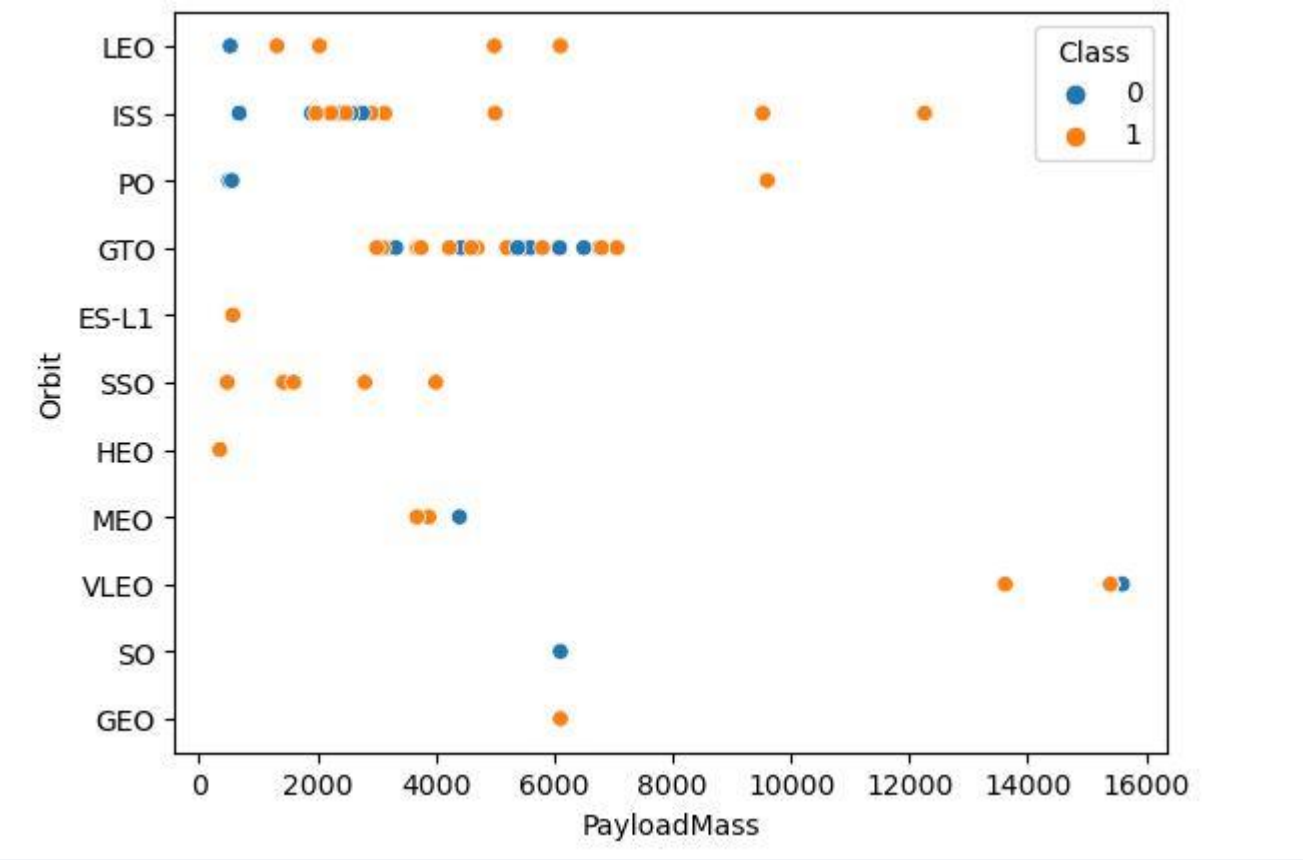
# Flight Number vs. Orbit Type

- Scatter point of Flight number vs. Orbit type



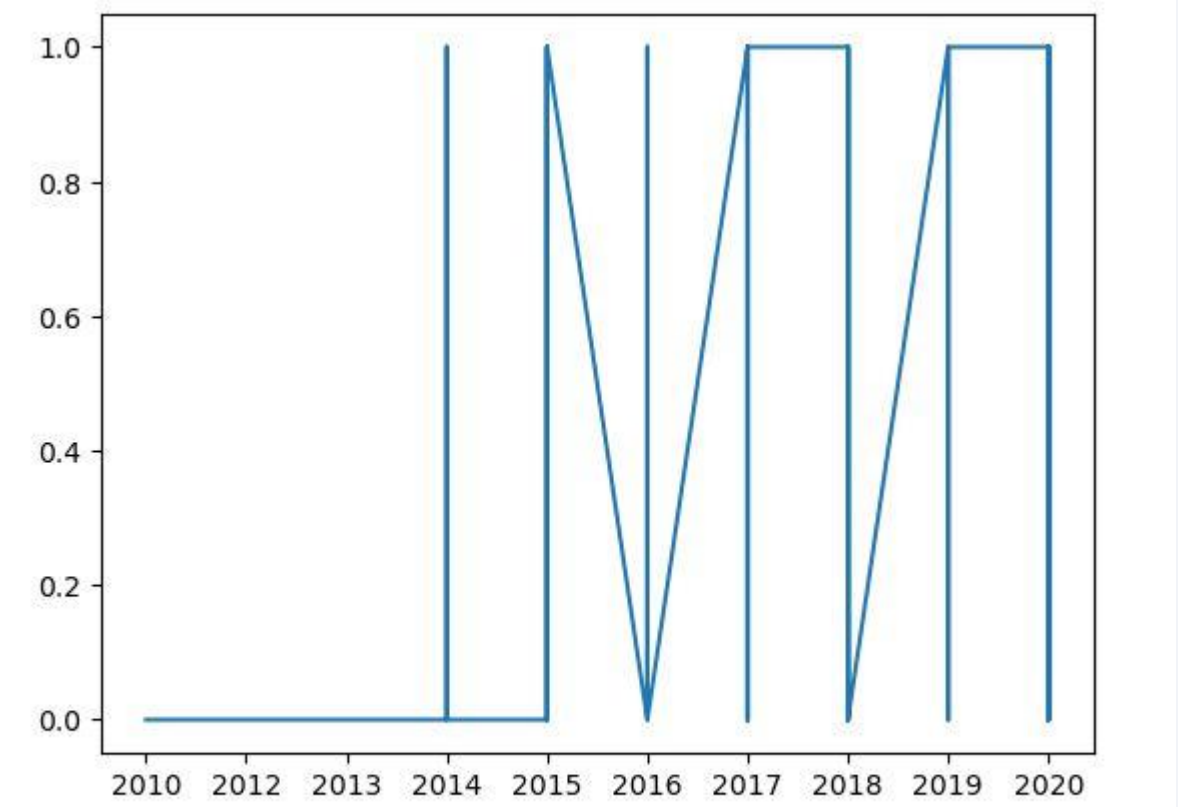
# Payload vs. Orbit Type

- Scatter point of payload vs. orbit type



# Launch Success Yearly Trend

- Line chart of yearly average success rate




you can observe that the sucess rate since 2013 kept increasing till 2020

# All Launch Site Names

- Names of the unique launch sites.

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

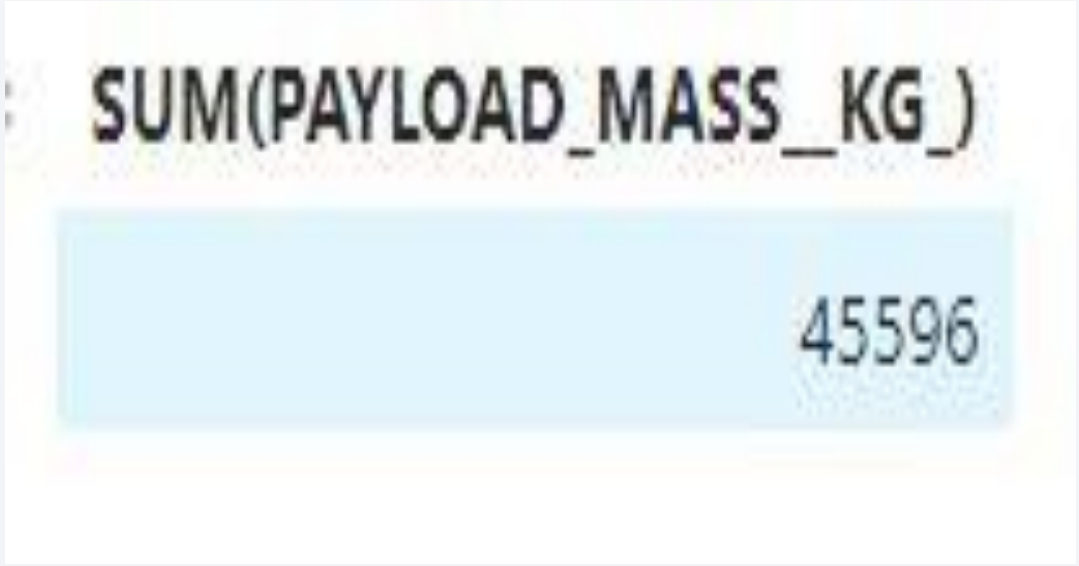
# Launch Site Names Begin with 'CCA'

- 5 records where launch sites begin with `CCA`.

| Launch_Site |
| --- |
| CCAFS LC-40 |
| CCAFS LC-40 |
| CCAFS LC-40 |
| CCAFS LC-40 |
| CCAFS LC-40 |

# Total Payload Mass

- Total payload carried by boosters from NASA.



# Average Payload Mass by F9 v1.1

- Average payload mass carried by booster version F9 v1.1.

| AVG(PAYLOAD_MASS_KG_) |
| --- |
| 2928.4 |

# First Successful Ground Landing Date

- Date of the first successful landing outcome on ground pad.

# Successful Drone Ship Landing with Payload between 4000 and 6000

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

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

# Total Number of Successful and Failure Mission Outcomes

- Calculate the total number of successful and failure mission outcome.

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

# Boosters Carried Maximum Payload

- List of the names of the booster which have carried the maximum payload mass.

| 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

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

| MONTHS | LandingOutcome | 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 |

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

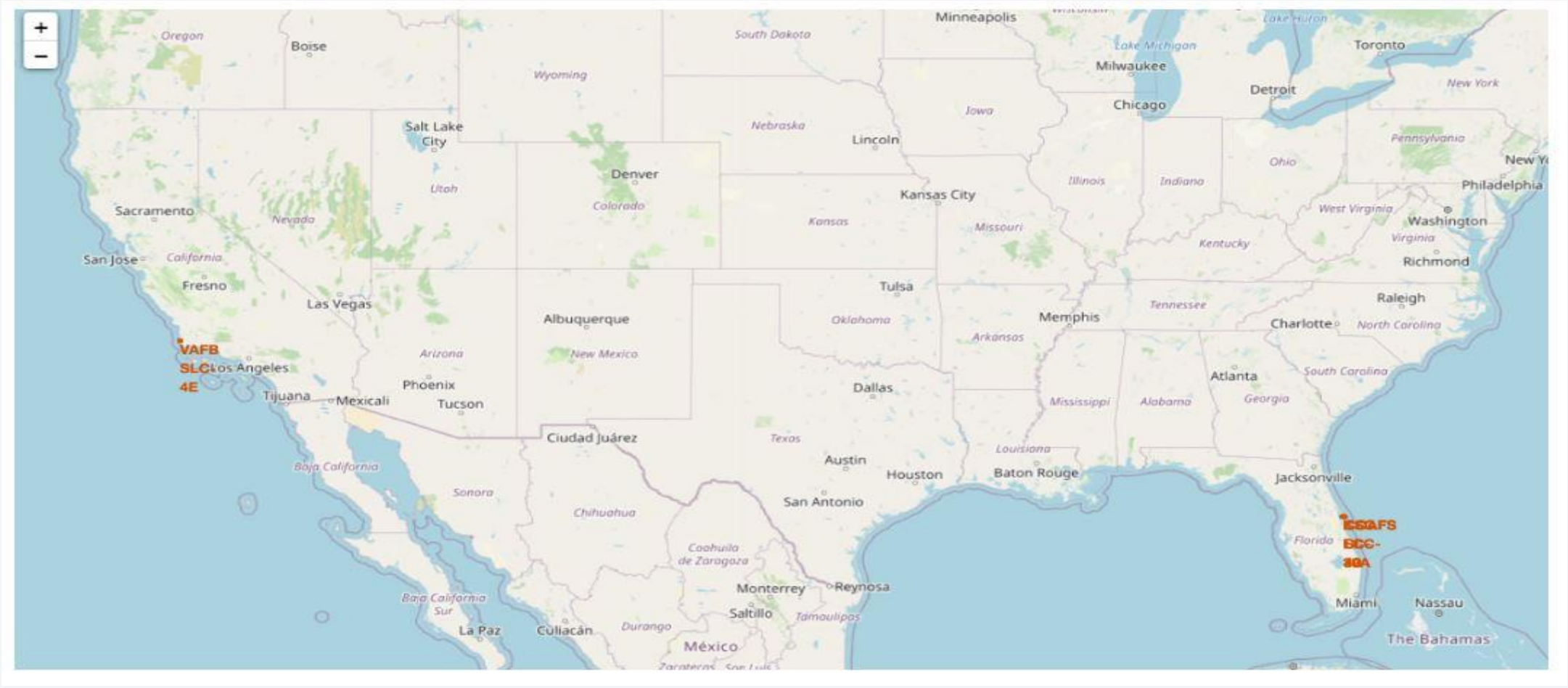
- Rank the count of landing outcomes (such as Failure (drone ship) or Success (ground pad)) between the date 2010-06-04 and 2017-03-20, in descending order

| Date | Time (UTC) | Booster_Version | Launch_Site | Payload | PAYLOAD_MASS__KG_ | Orbit | Customer | Mission_Outcome | LandingOutcome |
|---|---|---|---|---|---|---|---|---|---|
| 18-04-2014 | 19:25:00 | F9 v1.1 | CCAFS LC-40 | SpaceX CRS-3 | 2296 | LEO (ISS) | NASA (CRS) | Success | Controlled (ocean) |
| 05-12-2018 | 18:16:00 | F9 B5B1050 | CCAFS SLC-40 | SpaceX CRS-16 | 2500 | LEO (ISS) | NASA (CRS) | Success | Failure |
| 10-01-2015 | 09:47:00 | F9 v1.1 B1012 | CCAFS LC-40 | SpaceX CRS-5 | 2395 | LEO (ISS) | NASA (CRS) | Success | Failure (drone ship) |
| 04-06-2010 | 18:45:00 | F9 v1.0 B0003 | CCAFS LC-40 | Dragon Spacecraft Qualification Unit | 0 | LEO | SpaceX | Success | Failure (parachute) |
| 08-10-2012 | 00:35:00 | F9 v1.0 B0006 | CCAFS LC-40 | SpaceX CRS-1 | 500 | LEO (ISS) | NASA (CRS) | Success | No attempt |
| 06-08-2019 | 23:23:00 | F9 B5 B1047.3 | CCAFS SLC-40 | AMOS-17, Starlink 1 v1.0 | 6500 | GTO | Spacecom | Success | No attempt |
| 07-08-2018 | 05:18:00 | F9 B5 B1046.2 | CCAFS SLC-40 | Merah Putih | 5800 | GTO | Telkom Indonesia | Success | Success |
| 08-04-2016 | 20:43:00 | F9 FT B1021.1 | CCAFS LC-40 | SpaceX CRS-8 | 3136 | LEO (ISS) | NASA (CRS) | Success | Success (drone ship) |
| 18-07-2016 | 04:45:00 | F9 FT B1025.1 | CCAFS LC-40 | SpaceX CRS-9 | 2257 | LEO (ISS) | NASA (CRS) | Success | Success (ground pad) |

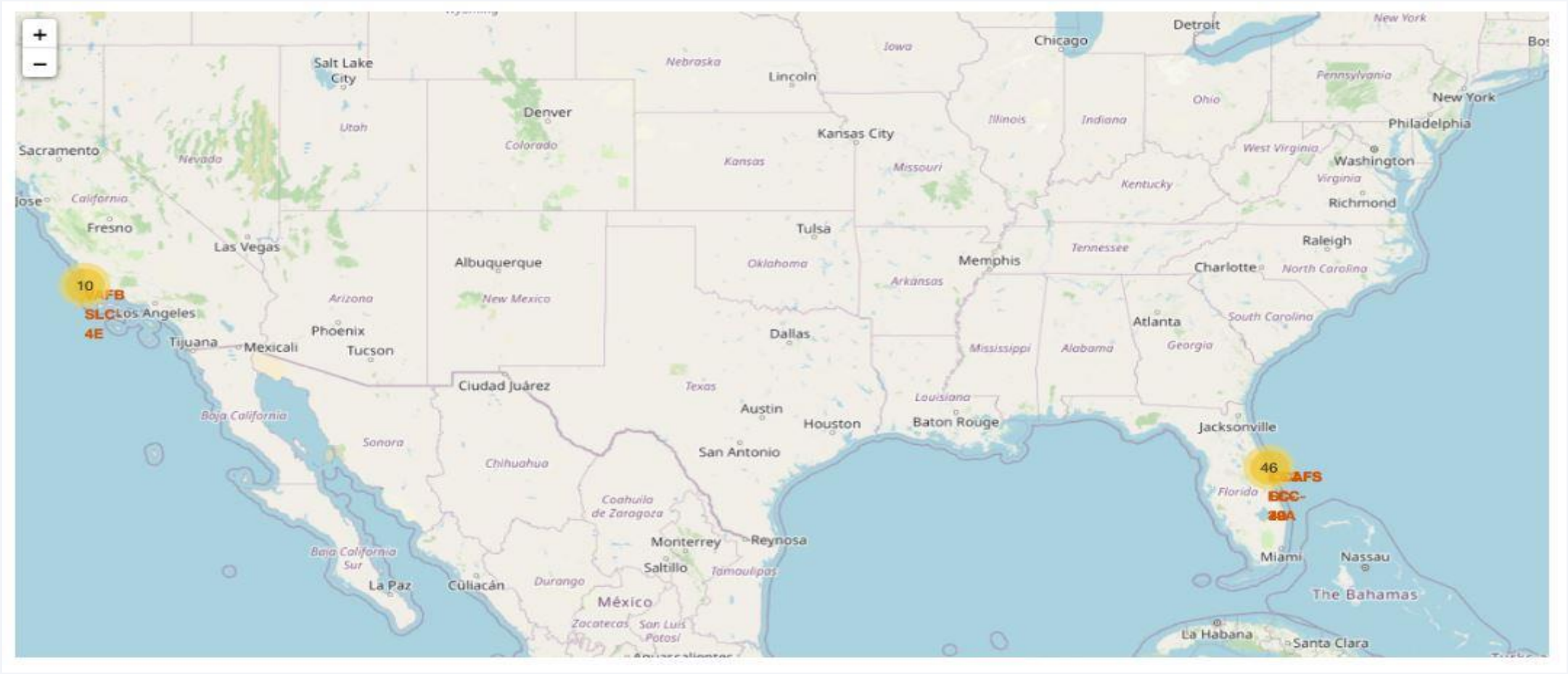
Section 3

# Launch Sites Proximities Analysis

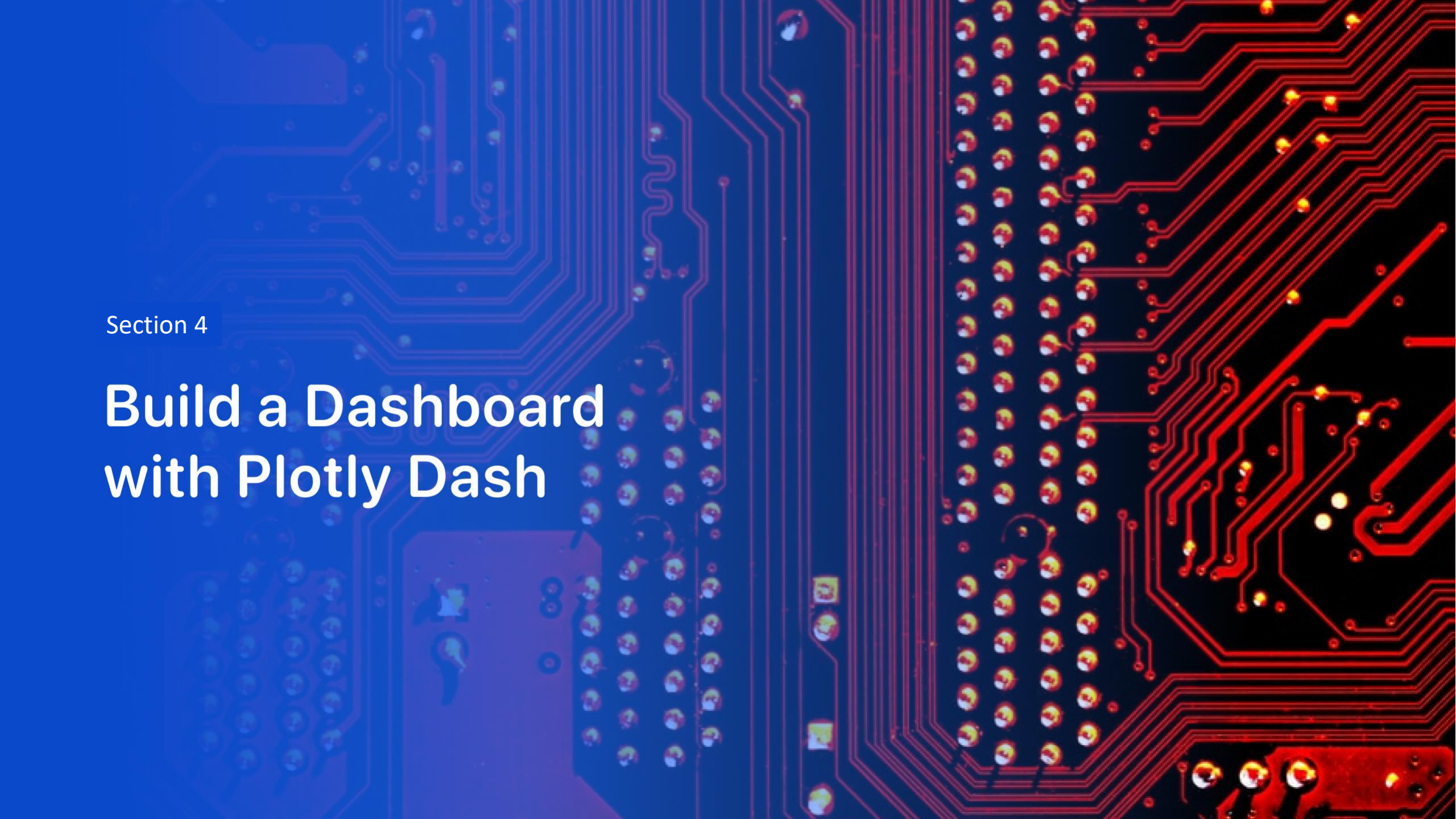
# Launch Sites on the Map



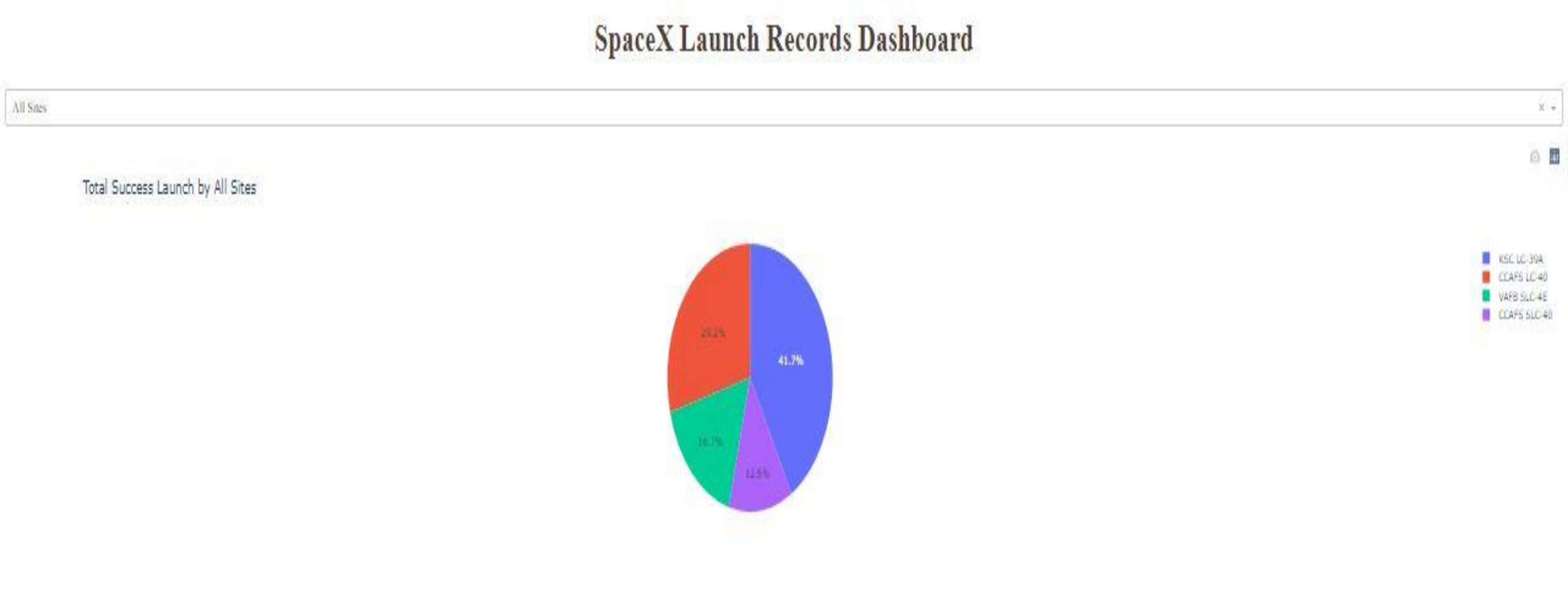
# Color Labeled Launch Outcomes



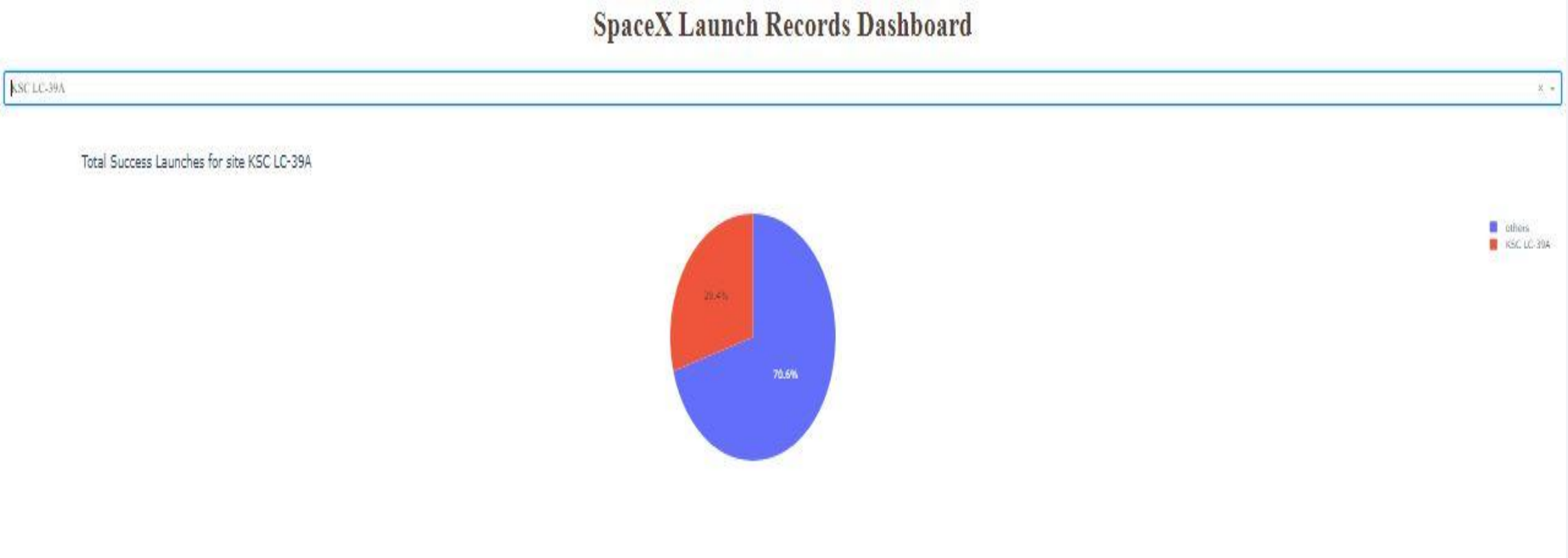
# Launch Sites Proximities


Section 4
Build a Dashboard with Plotly Dash

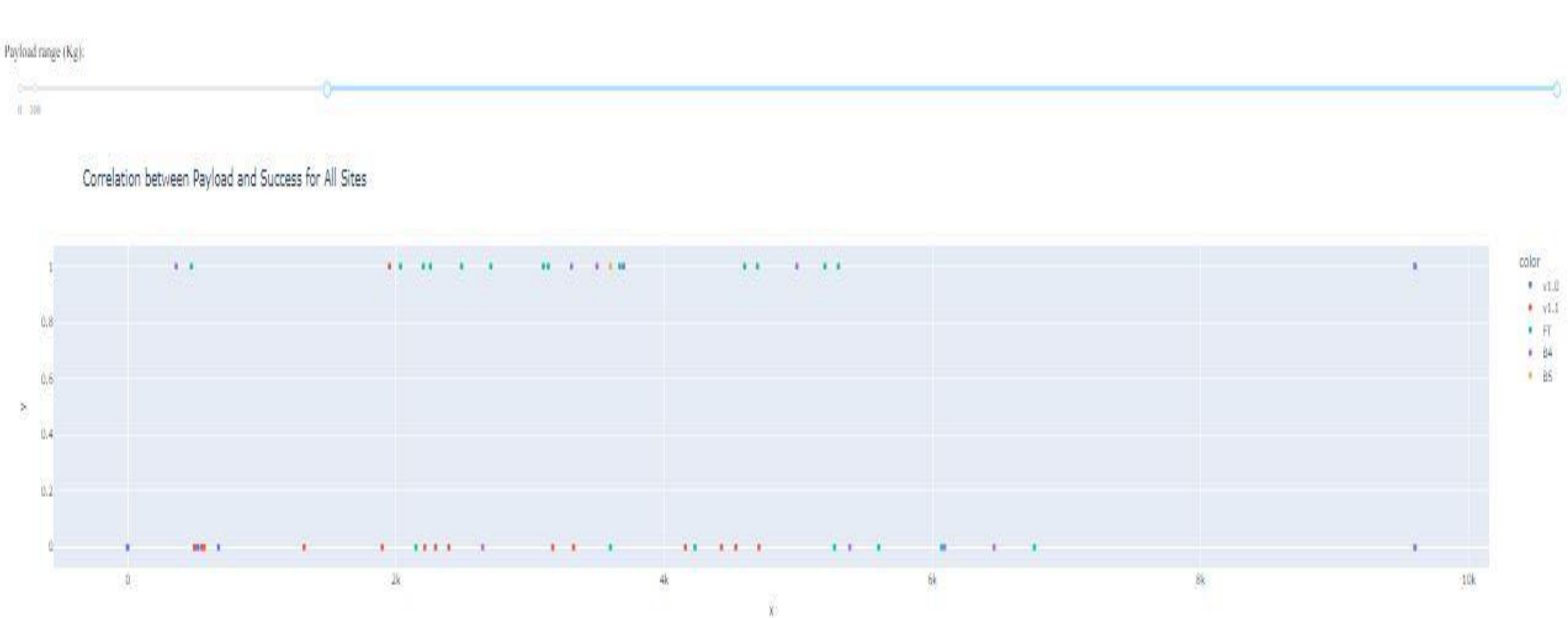
# Total Success Launch by All Sites



# Launch Site with the highest rate



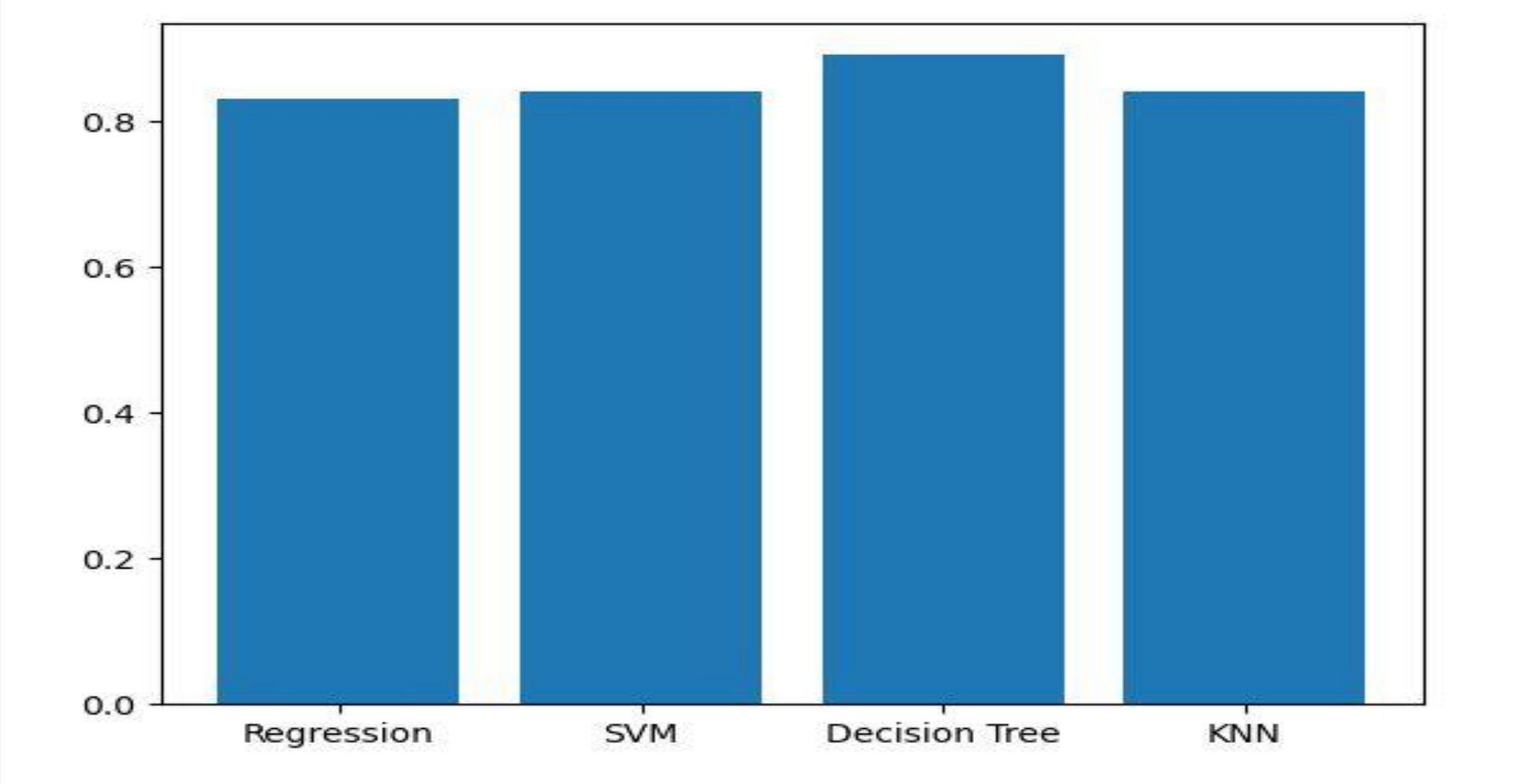
# Payload vs Launch Outcome Scatter Plot



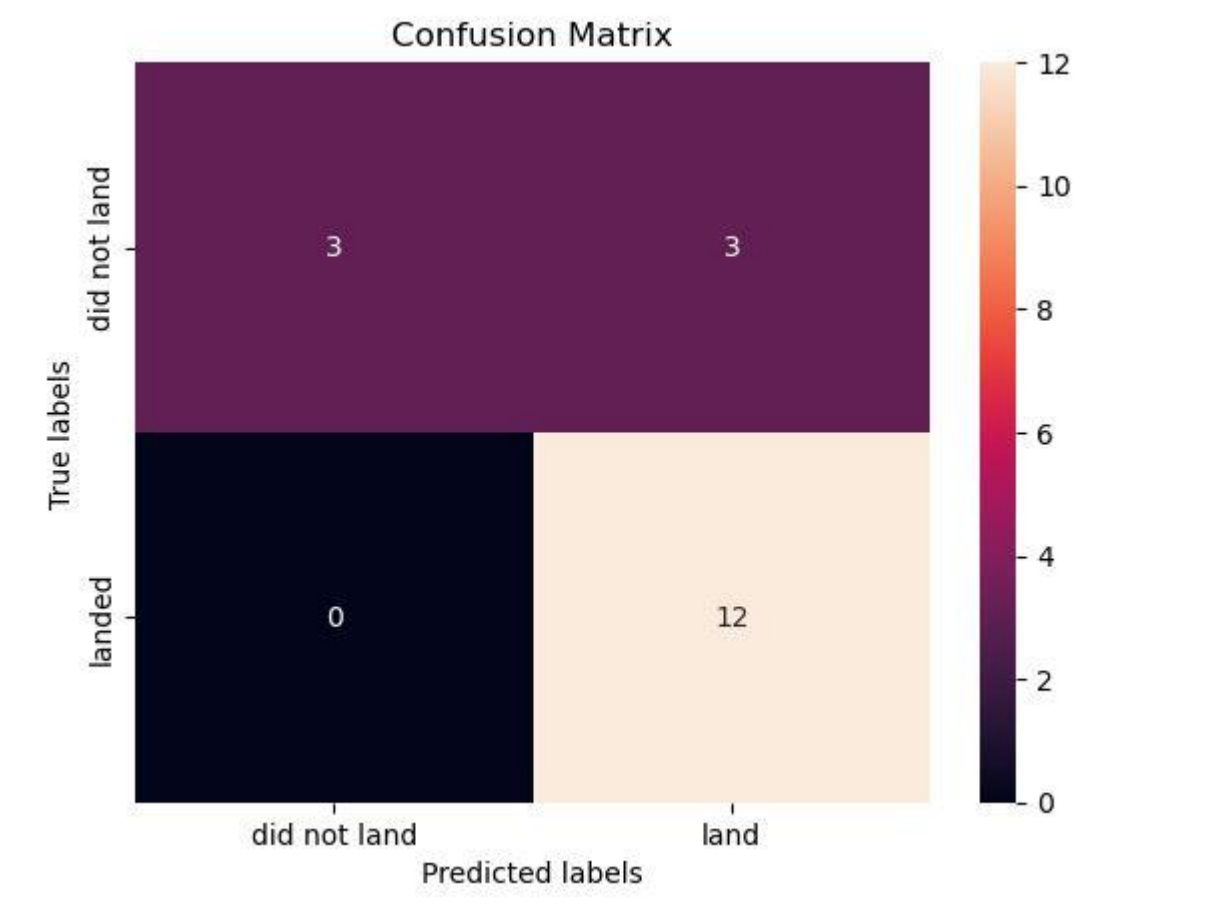
Section 5

# Predictive Analysis (Classification)

# Classification Accuracy



# Confusion Matrix



# 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!