IBM Developer
SKILLS NETWORK

# Winning Space Race with Data Science

<Name>
<Date>

# Outline

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

# Executive Summary

- Summary of methodologies
  - Data Collection
  - Data Wrangling
  - Exploratory Data Analysis (EDA)
  - Interactive Visual Analytics
  - Predictive Analysis
- Summary of all results
  - KSC LC-39A has highest success rate
  - Highest success rate is with a payload less than 3K
  - Decision Tree model offered most effective predictive accuracy

# Introduction

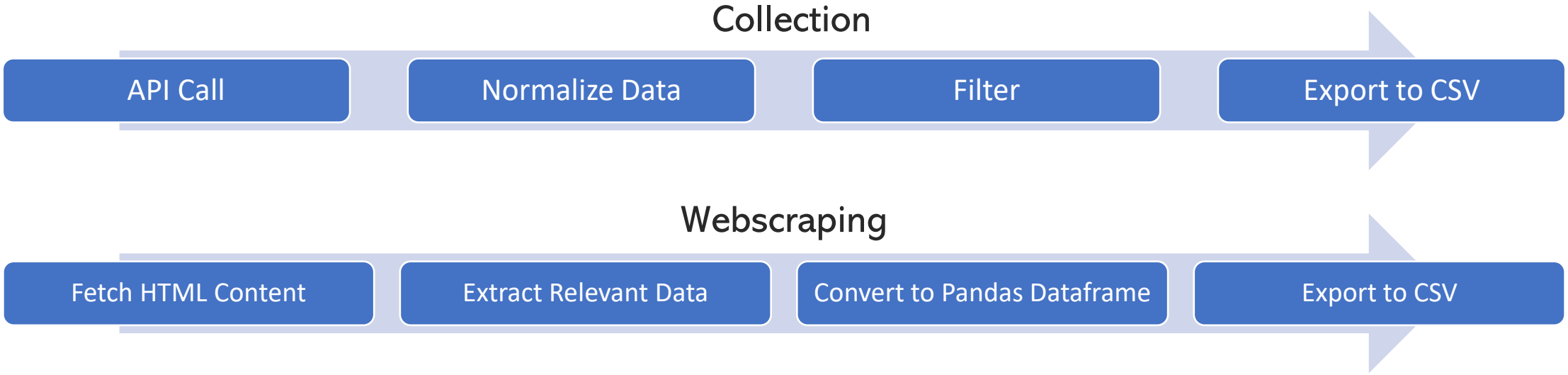
- Project background and context
  - This project provides insight into the cost efficiency of SpaceX's Falcon 9 rockets ability to reuse the first stage in multiple launches.
- Problems you want to find answers
  - The intent is to predict which rocket types and locations provide the most cost-effective means to reuse rockets and parts used for space exploration to potentially create competition in the space launch market.

Section 1

# Methodology

# Data Collection – SpaceX API

Data was collected using SpaceX API and through webscraping to collect Falcon 9 launch records from Wikipedia.




Github Links:

https://github.com/schinook84/Applied-Data-Science-Capstone/blob/main/Data%20Collection%20API.ipynb

https://github.com/schinook84/Applied-Data-Science-Capstone/blob/main/Webscraping.ipynb

# Data Wrangling

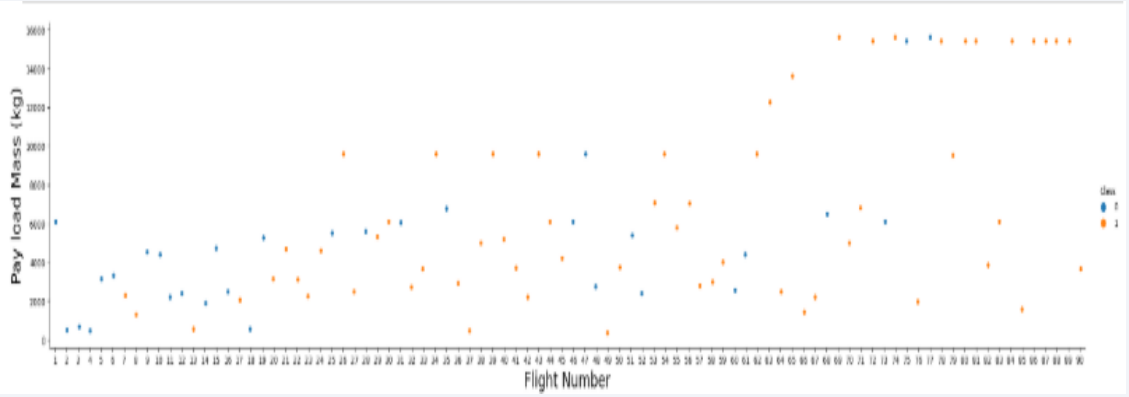
Patterns were sought in the data to determine what would be useful in determining success rates of launches. Data was separated into successful and failed launches.

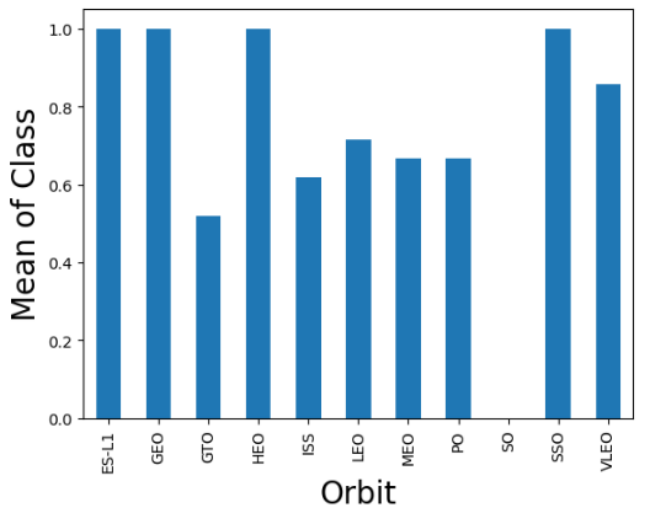

Github Link:

https://github.com/schinook84/Applied-Data-Science-Capstone/blob/main/Data%20Wrangle.ipynb

# EDA with Data Visualization

- Charts Used and Reasoning:
  - Scatter Chart
    - Shows feature correlations
  - Bar Chart
    - Illustrates success rate comparisons by orbit




Github Link:

https://github.com/schinook84/Applied-Data-Science-Capstone/blob/main/Exploring%20and%20Preparing%20Data.ipynb

# EDA with SQL

- Performed query to count number of launches at each site
- Ran query to count number of each orbit type
- Performed query on amount of successful and failed launches

Github Link:

https://github.com/schinook84/Applied-Data-Science-Capstone/blob/main/SQL%20Notebook.ipynb

# Build an Interactive Map with Folium

- Created markers and circles to show launch sites
- Drew lines to show distance to landmarks

Github Link:

https://github.com/schinook84/Applied-Data-Science-Capstone/blob/main/Launch%20Sites%20Locations%20Analysis%20with%20Folium.ipynb

# Build a Dashboard with Plotly Dash

- Added a pie and scatter chart to show relationship between payload and success rate and different launch sites.

Github Link:

https://github.com/schinook84/Applied-Data-Science-Capstone/blob/main/spacex_dash_app.py

# Predictive Analysis (Classification)

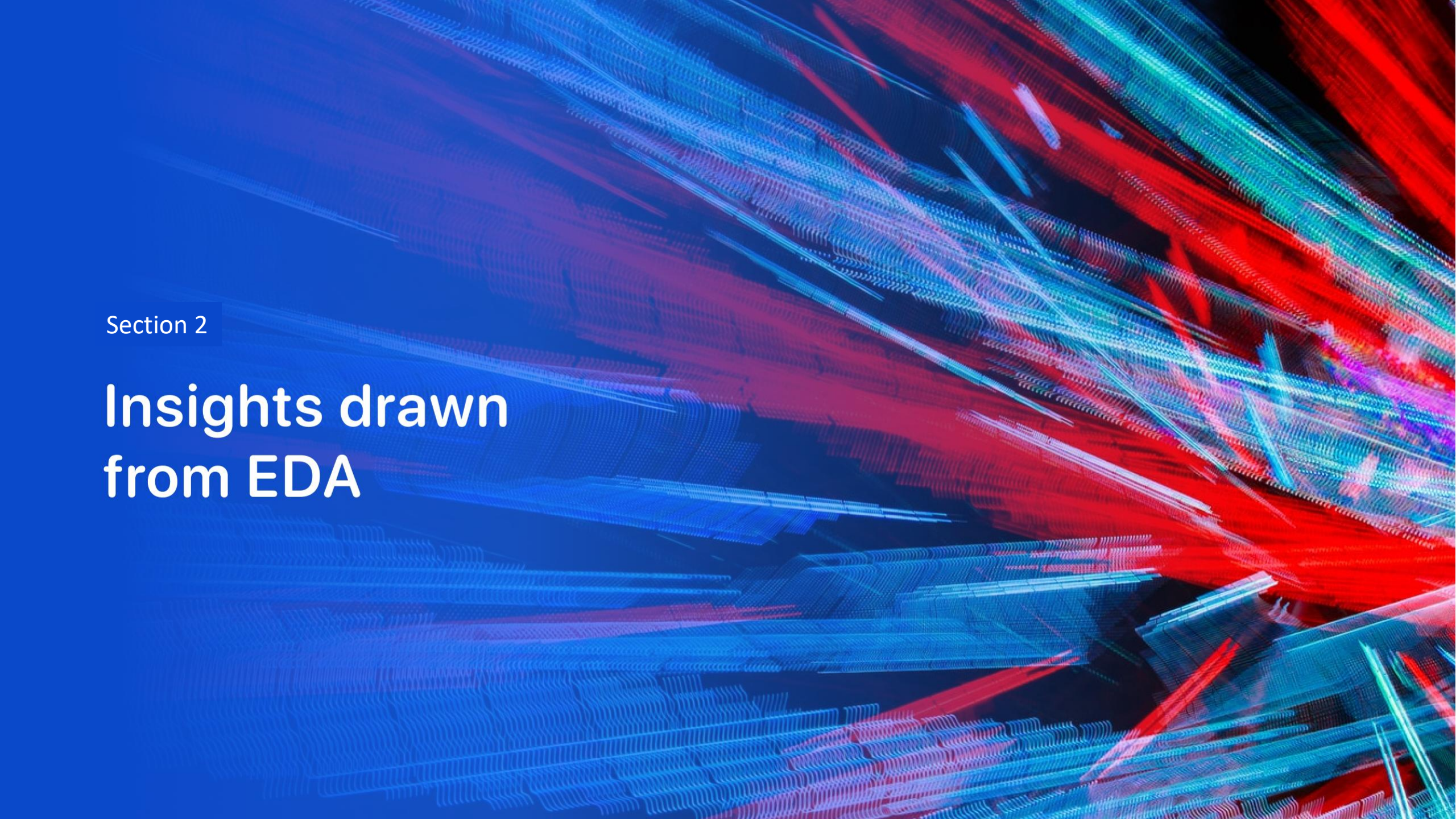
- Multiple models used to include logistic regressing, support vector machine, and decision tree to predict mission success.
- Models were evaluated with a confusion matrix to evaluate model performance to show true positives and negatives as well false positives and negatives.

Github Link:

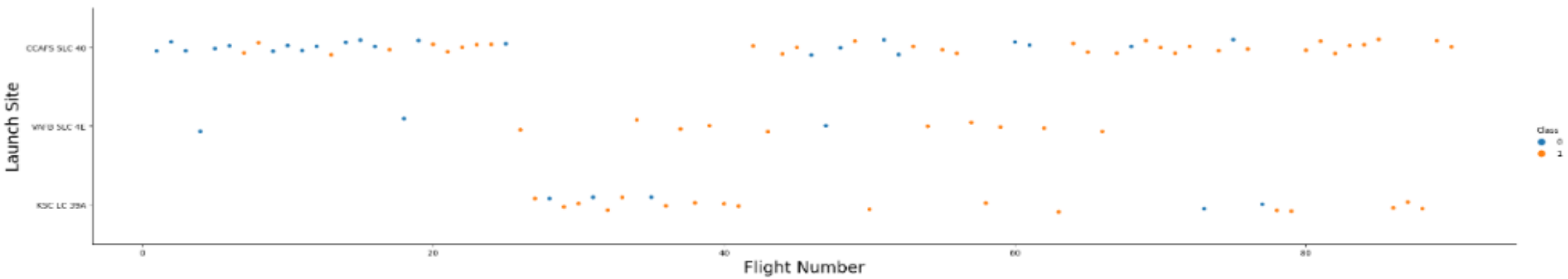
https://github.com/schinook84/Applied-Data-Science-Capstone/blob/main/SpaceX_Machine%20Learning%20Prediction.ipynb

# Results

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

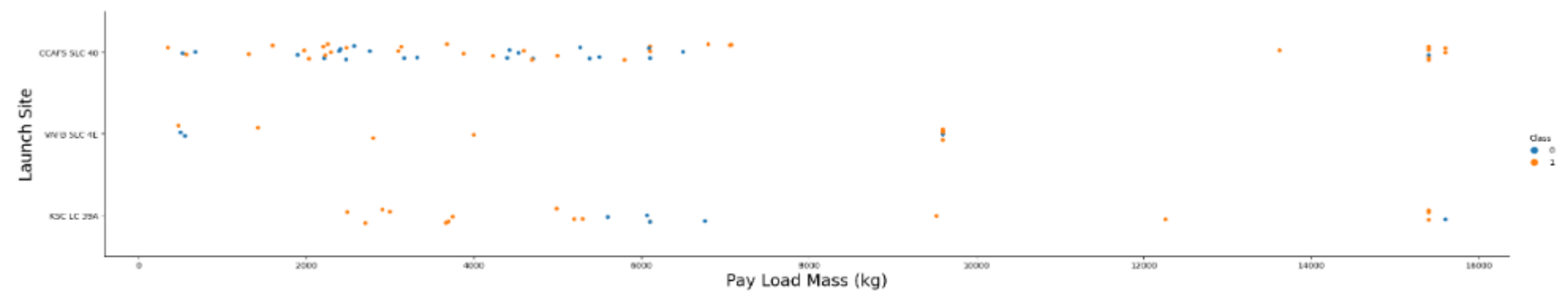

Section 2
Insights drawn from EDA

# Flight Number vs. Launch Site



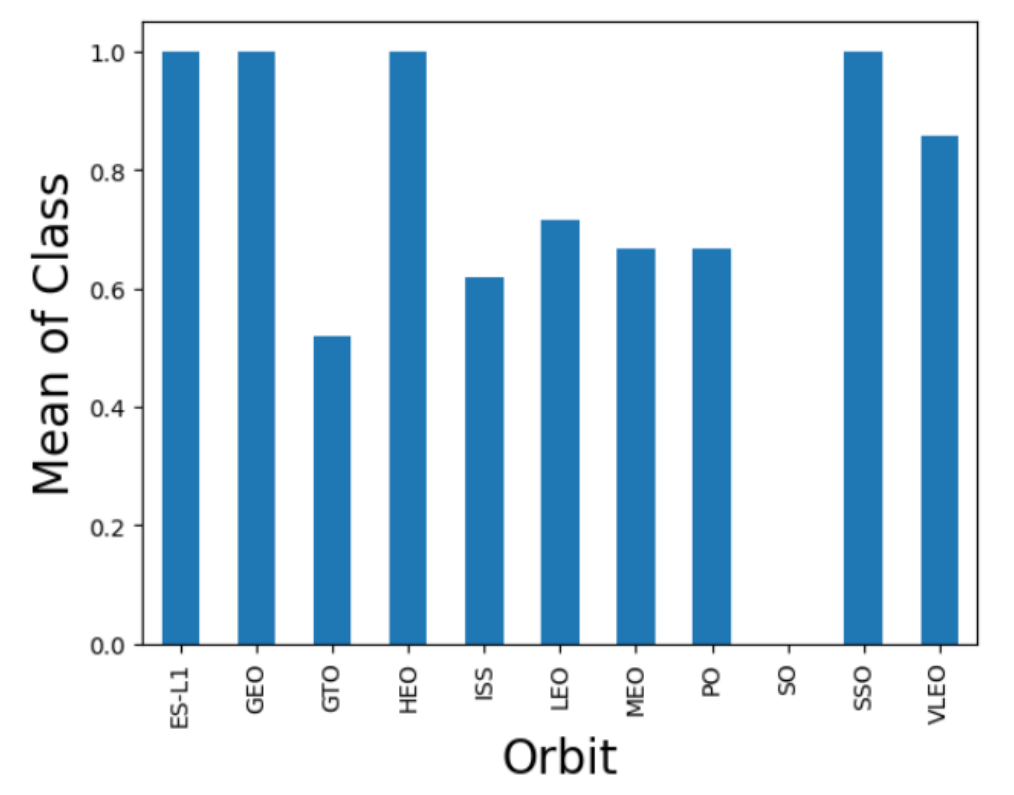

Launch success increases at each site as they have more launches.

# Payload vs. Launch Site



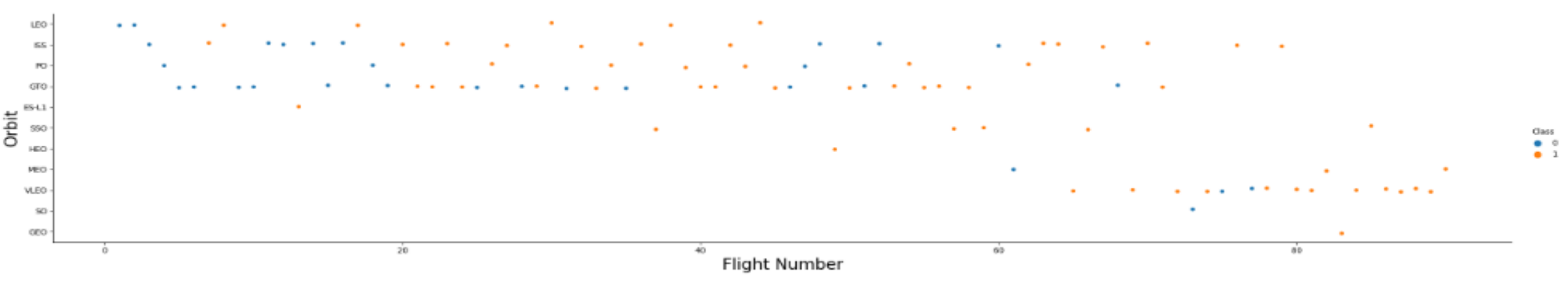

KSC LC 39A conducts a substantial amount of launched with heave payloads.

# Success Rate vs. Orbit Type



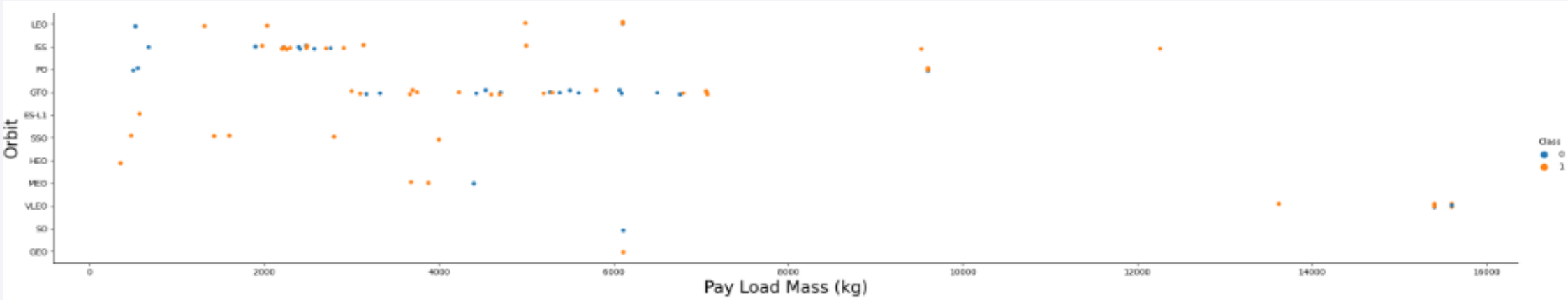

GTO has the lowest success rate while ES-L1, GEO, HEO, and SSO have the highest.

# Flight Number vs. Orbit Type



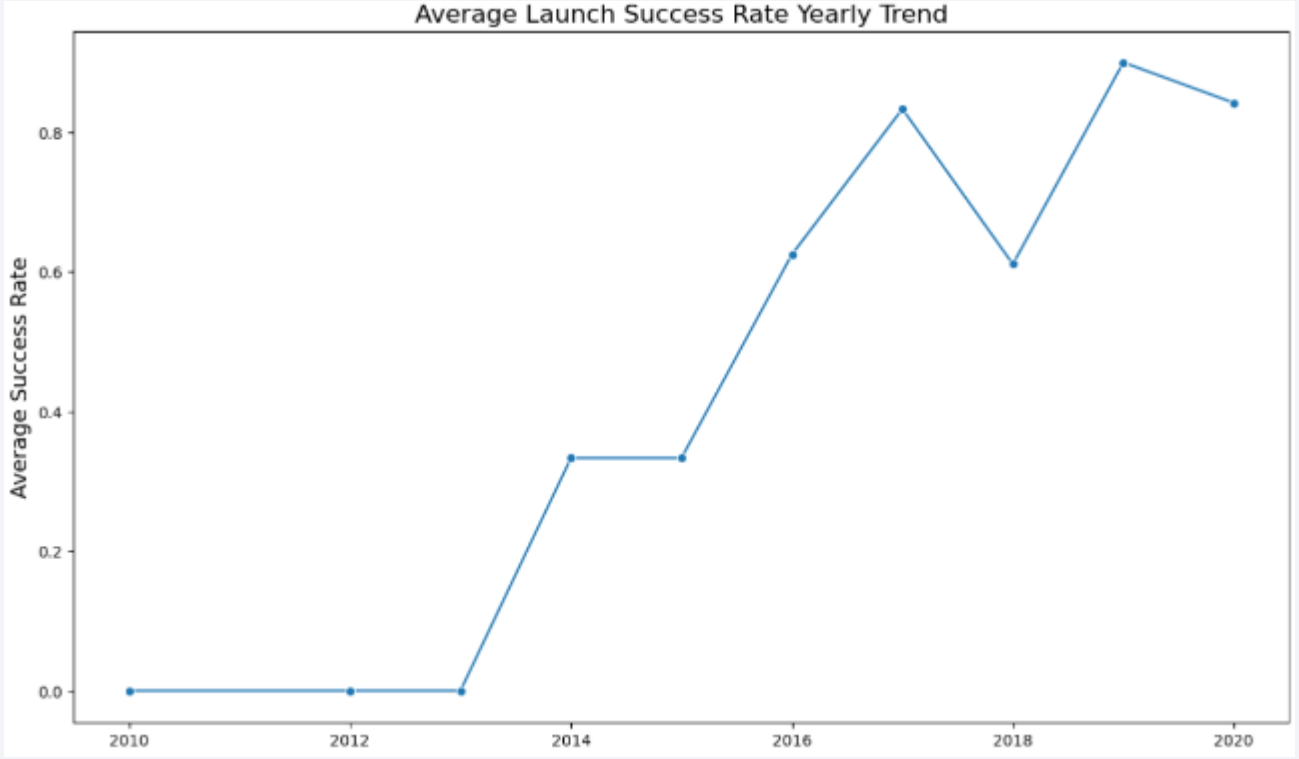

HEO, MEO, VLEO, SO, and GEO were not sought to be launched into until there had been over 50 launches.

# Payload vs. Orbit Type




VLEO seems to be the location of the highest payload launches.

# Launch Success Yearly Trend




Launch success was on a good path until 2017.

# All Launch Site Names

- The launch sites were CCAFS LC-40, CCAFS SLC-40, KSC LC 39A and VAFB SLC-4E.

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

# Launch Site Names Begin with 'CCA'

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

| | |
|---|---|
| **49** | CCAFS SLC-40 |
| **50** | CCAFS SLC-40 |
| **51** | CCAFS SLC-40 |
| **52** | CCAFS SLC-40 |
| **53** | CCAFS SLC-40 |

# Total Payload Mass

- Calculate the total payload carried by boosters from NASA

| SUM(PAYLOAD_MASS__KG_) |
| --- |
| 48213 |

# Average Payload Mass by F9 v1.1

- Calculate the average payload mass carried by booster version F9 v1.1

| AVG(PAYLOAD_MASS__KG_) |
| --- |
| 2534.6666666666665 |

# First Successful Ground Landing Date

- Find the dates of the first successful landing outcome on ground pad

| MIN("Date") |
| --- |
| 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

| 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 outcomes

| COUNT(Mission_Outcome) |
|---|
| 1 |
| 98 |
| 1 |
| 1 |

# Boosters Carried Maximum Payload

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

| Year | Month | Booster_Version | Launch_Site | Landing_Outcome |
|---|---|---|---|---|
| 2015 | 01 | F9 v1.1 B1012 | CCAFS LC-40 | Failure (drone ship) |
| 2015 | 04 | F9 v1.1 B1015 | CCAFS LC-40 | Failure (drone ship) |

# 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

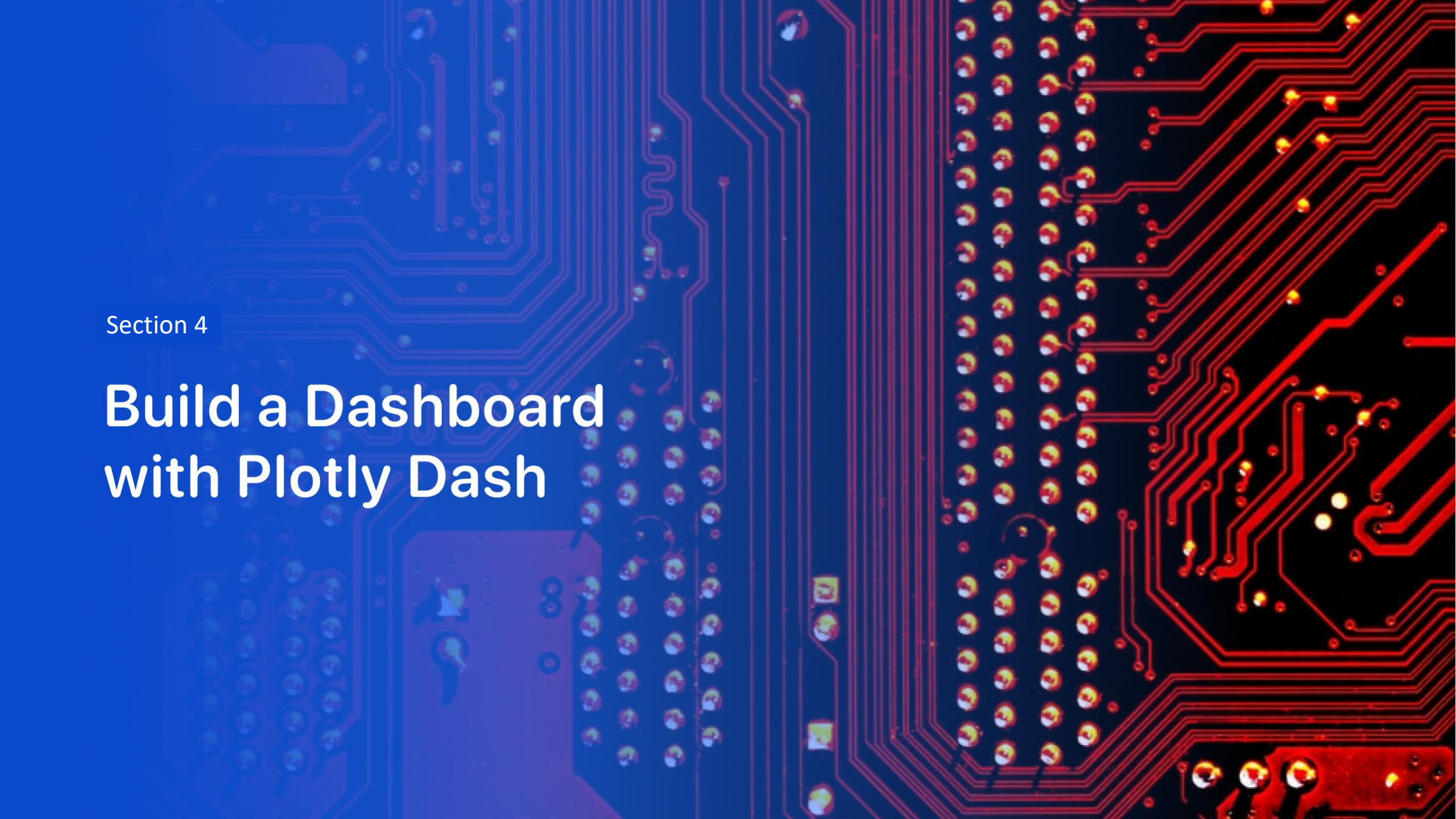
| Landing_Outcome | Outcome_Count |
|---|---|
| No attempt | 10 |
| Success (drone ship) | 5 |
| Failure (drone ship) | 5 |
| Success (ground pad) | 3 |
| Controlled (ocean) | 3 |
| Uncontrolled (ocean) | 2 |
| Failure (parachute) | 2 |
| Precluded (drone ship) | 1 |

Section 3

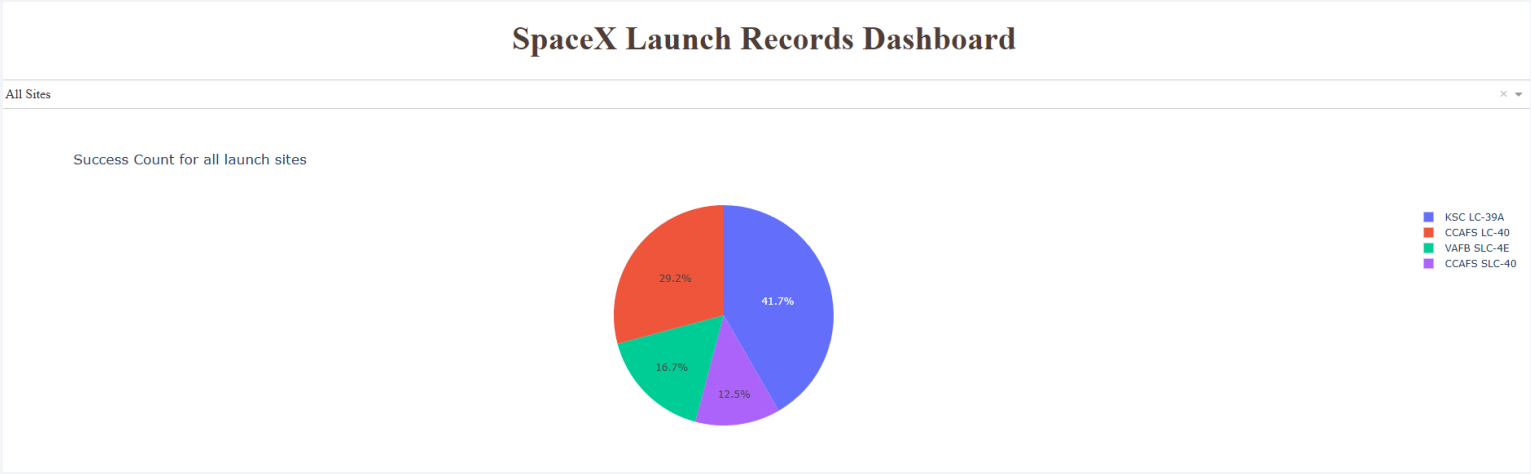
# Launch Sites Proximities Analysis

# SpaceX Launch Sites Map

- Launch sites are near the coast.

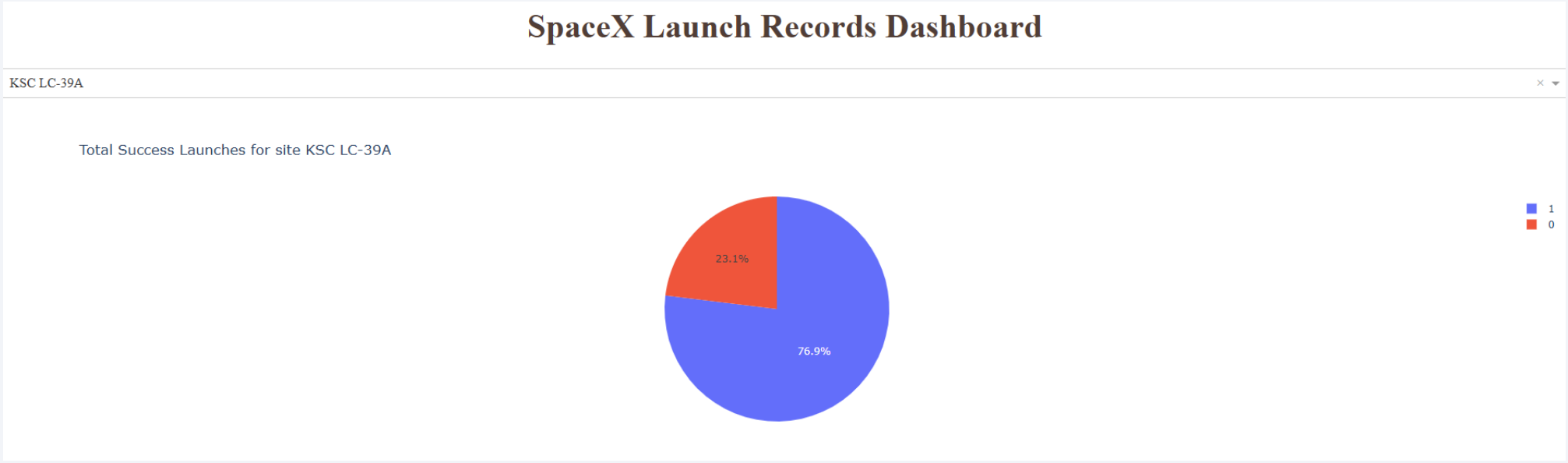

Section 4
Build a Dashboard with Plotly Dash

# All Sites Launch Success



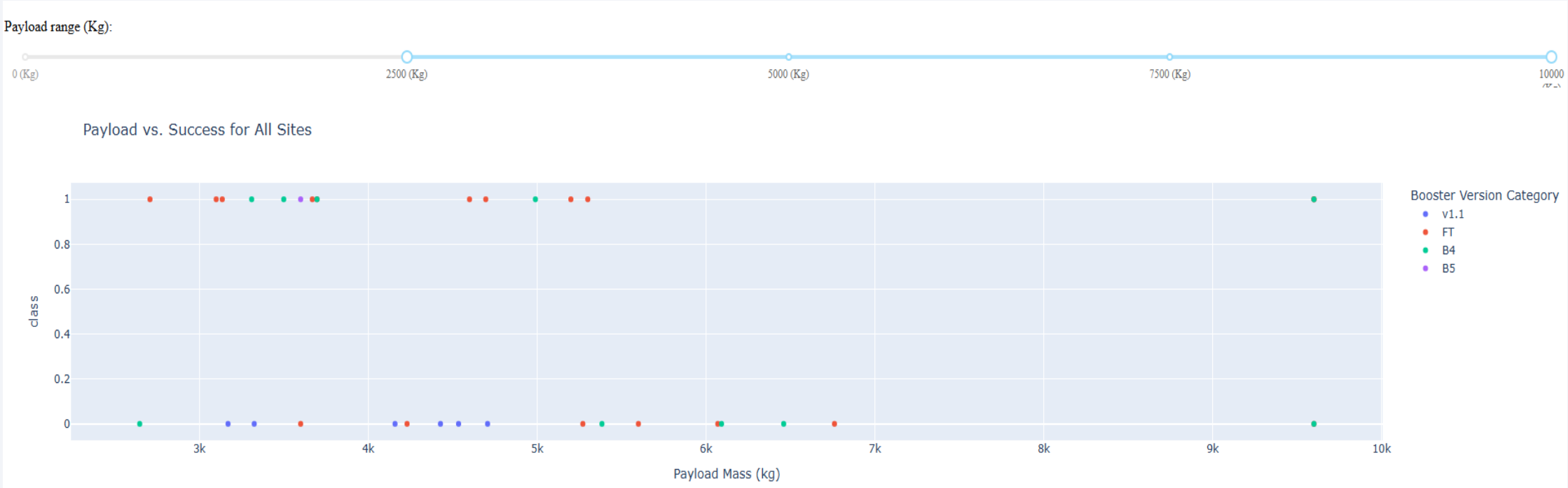

- KSC LC 39-A has the most successful launches.

# KSC LC-39A Launch Site




- This site had a 76.9% launch success rate.

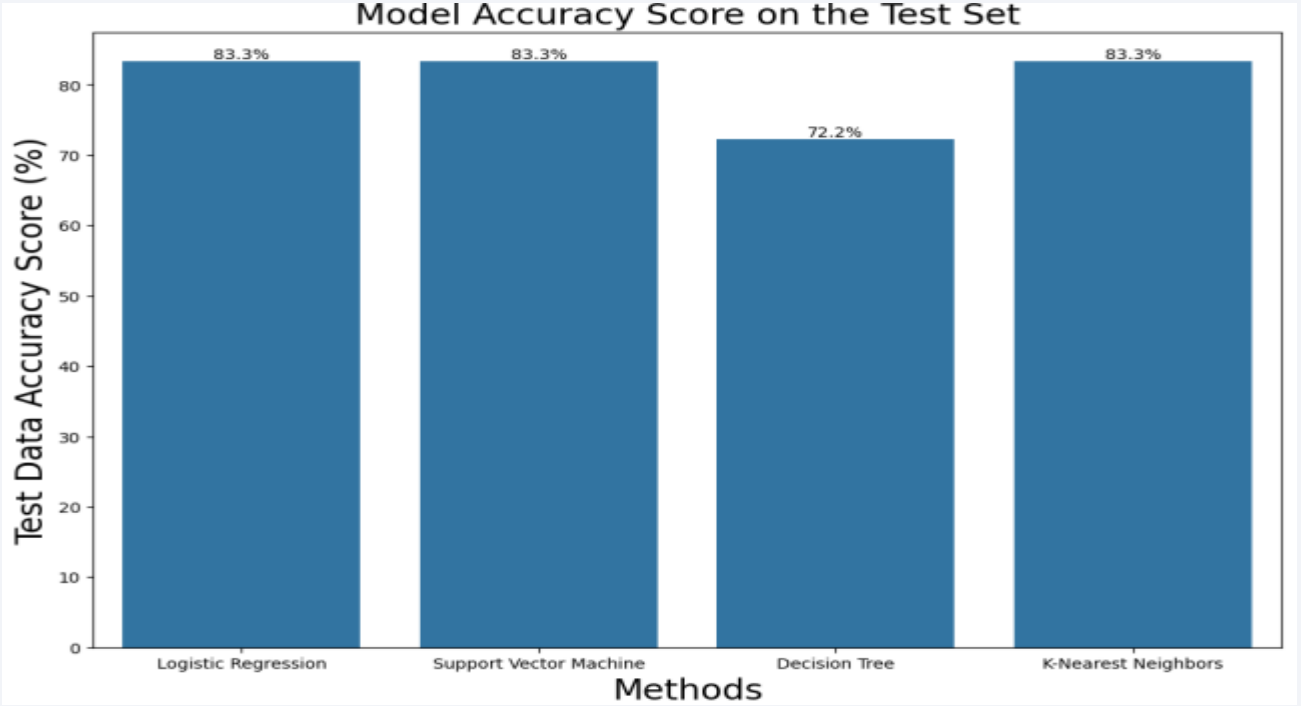
# Payload vs. Launch Scatter Chart



Section 5

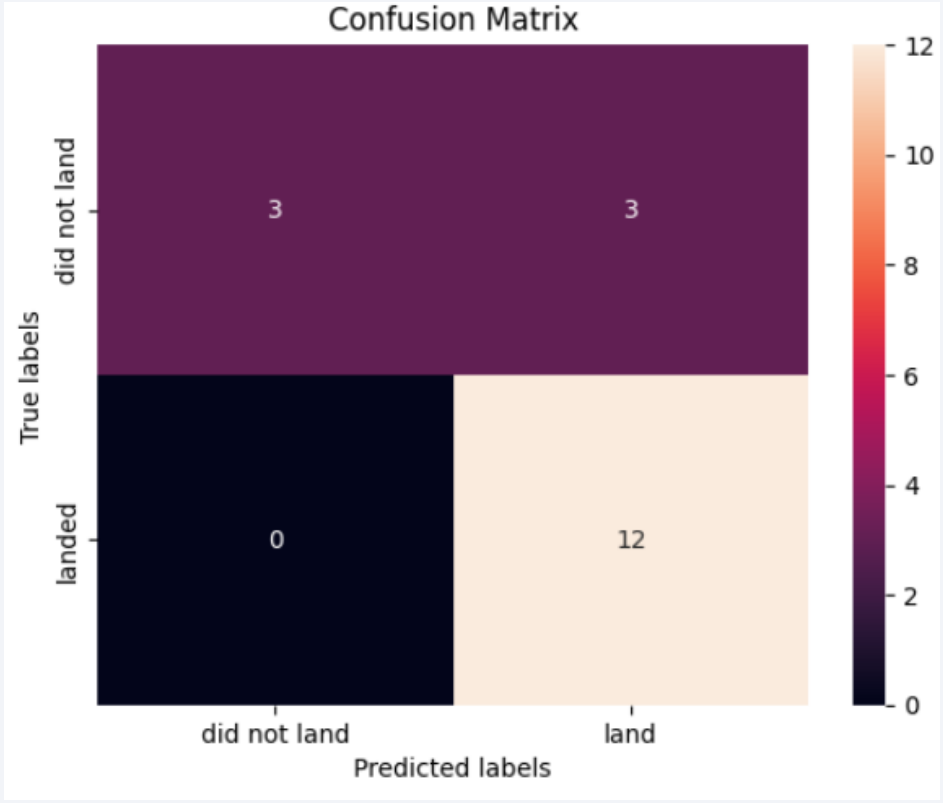
# Predictive Analysis (Classification)

# Classification Accuracy




- 3 out of 4 methods had the same accuracy score.

# Confusion Matrix




- The predictions indicate a high likelihood of successful launches.

# Conclusions

- Point 1
  - KSC LC-39A had the highest success rate among all launch sites.
- Point 2
  - Payloads below 3K are the most successful.
- Point 3
  - Booster FT has the highest likelihood of successful recovery.

# Thank you!