IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Iustinian Chirila
23.06.2023

Outline

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

Executive Summary

- Falcon 9 Stage One is good for light payloads (between 2000kg and 4000kg)
- Best launching site is KSC LC-39A
- Most successful launches are after 2017
- The best method to recover is via drone ship
- Model can predict the recovery with accuracy of 83,33%

Introduction

- Falcon 9 is a partially reusable medium-lift launch vehicle that can carry cargo and crew into Earth orbit, designed, manufactured and launched by American aerospace company SpaceX.
- With Falcon 9, SpaceX can recover and reuse a large and expensive component of the rocket called Stage One.
- We will predict if SpaceX Stage One recovery will be successful or not.

Section 1

Methodology

Methodology

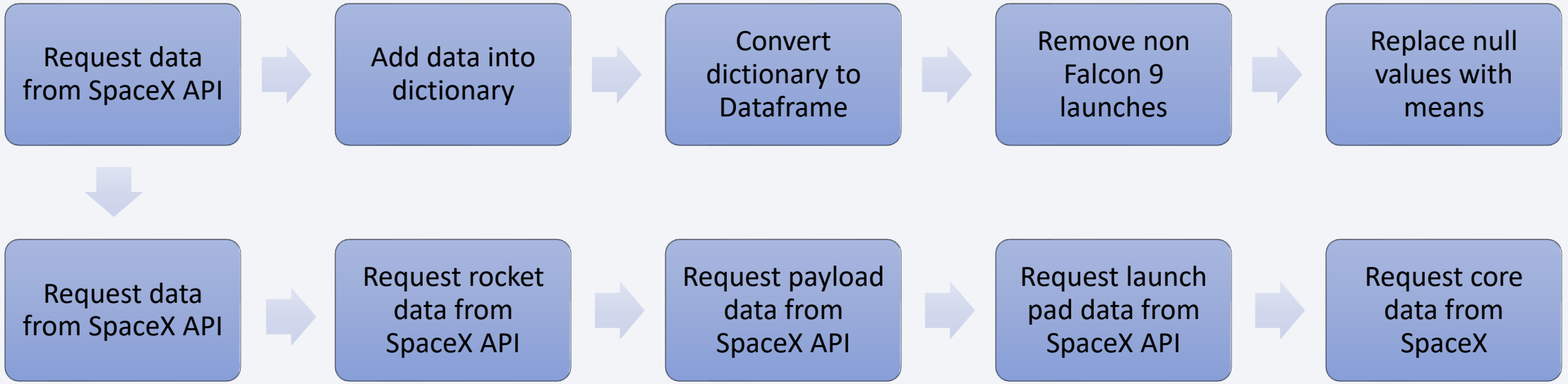
Executive Summary

- Data collection methodology:
  - The data was collected from SpaceX API
- Perform data wrangling
  - Modifying and transforming data we can use for training models
- Perform exploratory data analysis (EDA) using visualization and SQL
- Perform interactive visual analytics using Folium and Plotly Dash
- Perform predictive analysis using classification models
  - Training and test data

Data Collection

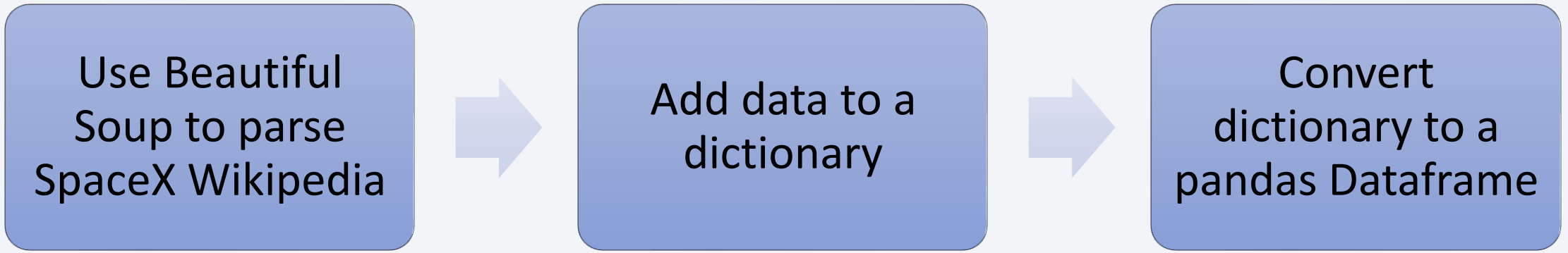
- For collecting data we need a .json file
- The file contains data with rocket launches from SpaceX
- Download the file
- Transform it into Dataframes
- Null values are converted to means

Data Collection – SpaceX API



- GitHub

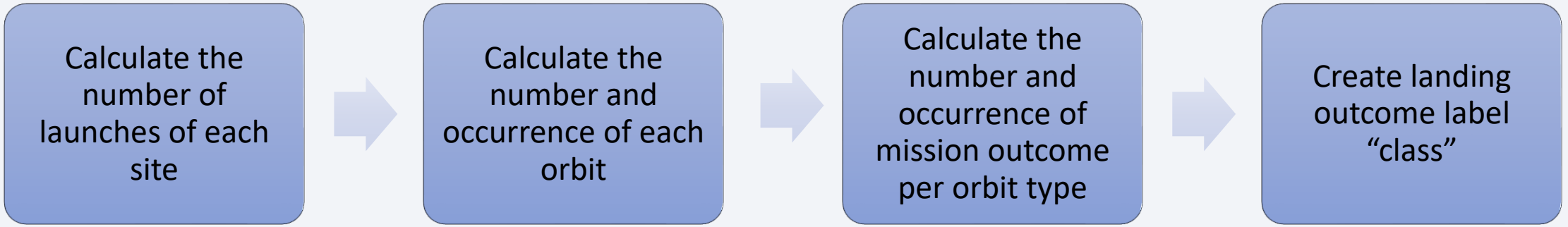
Data Collection - Scraping



- GitHub

Data Wrangling

- We have to determine if the recovery is successful or not
- In the dataframe are 8 types of outcomes and 2 values for each, success or failure.
- True ASDS, True RTLS and True Ocean as successful landings
- None ASDS, False ASDS, False RTLS, False Ocean and None None as failed landings
- We could create a column “class” to differentiate between successful and unsuccessful
- 1 for successful and 0 for unsuccessful recover



- GitHub

EDA with Data Visualization

- Plots showing relationships between different variables:
  - Flight Number vs. Payload (Cat plot)
  - Flight Number vs. Launch Site (Cat plot)
  - Launch Site vs. Payload (Scatter plot)
  - Success Rate vs. Orbit type (Bar plot)
  - Orbit type vs. Flight Number (Scatter plot)
  - Orbit type vs. Payload (Scatter plot)
  - Success rate vs. Time in years (Line plot)

- GitHub

EDA with SQL

- We made some queries on our data for the relationship between variables
  - Launch Site
  - Payload Mass
  - Mission Outcome
  - Booster Version
  - Date

- GitHub

Build an Interactive Map with Folium

- We used Folium package to create a map where we can:
  - See the distance to the closest coastline, city, railway and highway
  - See where each Falcon 9 is located
  - Learn how many launches have occurred at each location
  - Green markers is for successful recovery and red markers for failed ones

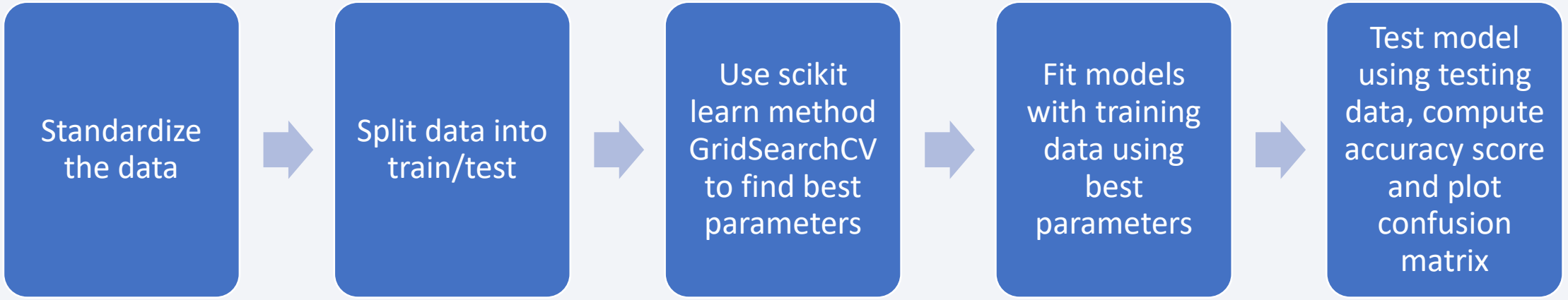
- GitHub

Build a Dashboard with Plotly Dash

- We made a dashboard through Ploty Dash
  - To see the proportion of successful recovery to unsuccessful ones we used a pie chart
  - A scatter plot Recovery Outcomes vs Payload Mass with a range 0-10000kg with bounds that can be changed
- The dashboard provides insight into the launch sites and payload masses relationships with the recovery outcomes

- GitHub

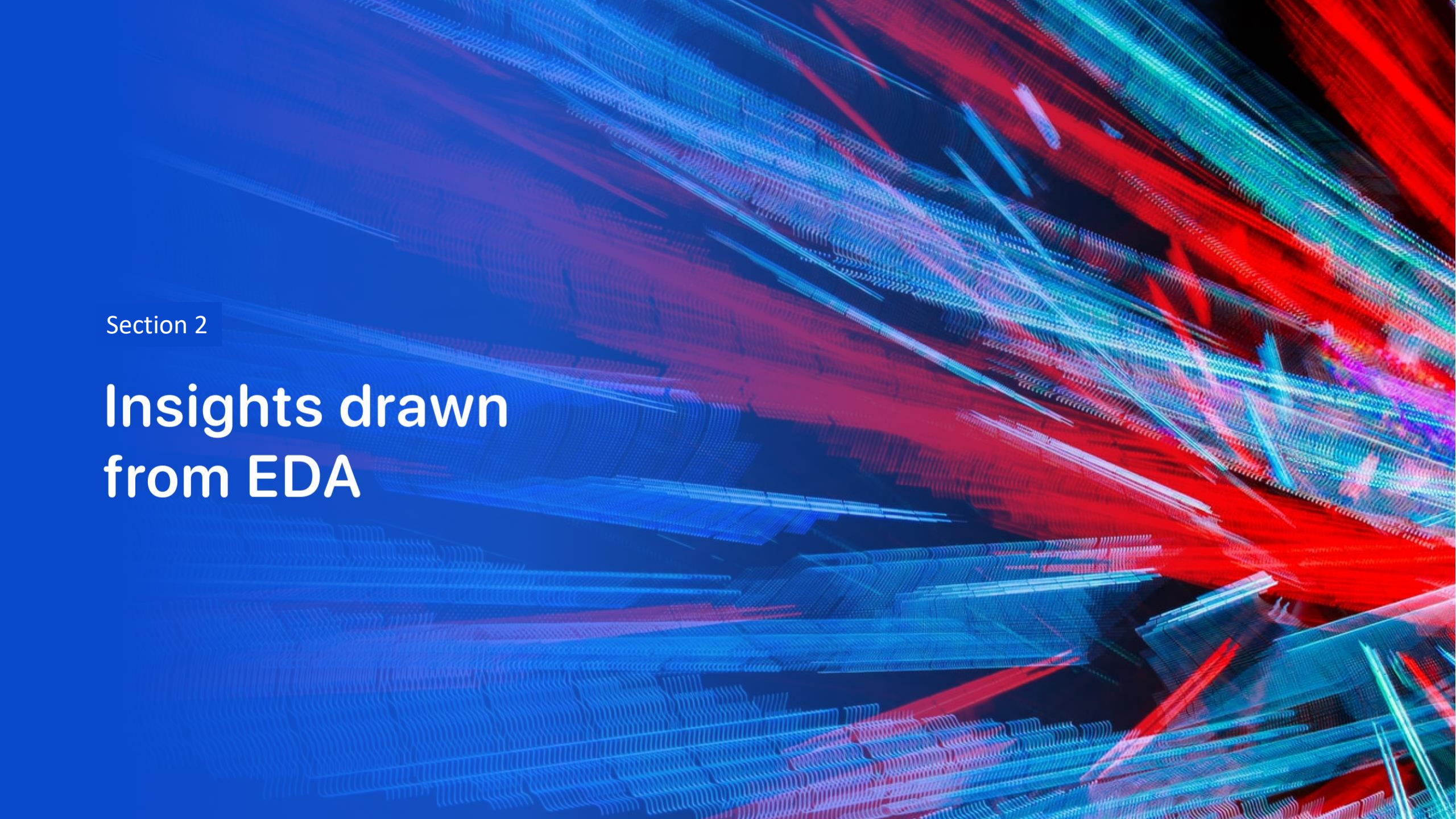
Predictive Analysis (Classification)



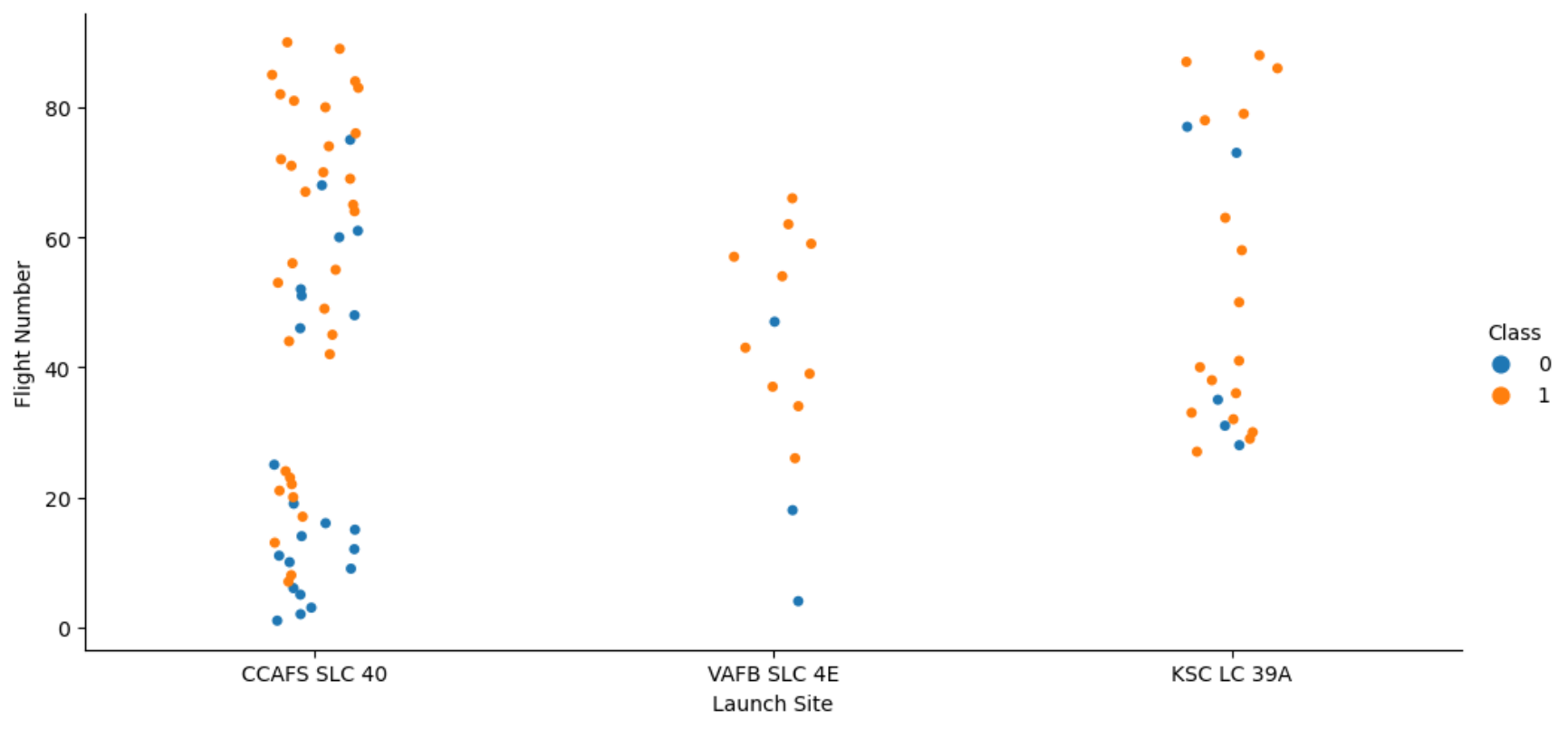
- GitHub

Results

- Exploratory data analysis results
  - Launches are most successful after year 2017
  - Light payloads are easier to recover (mass between 2000kg and 4000kg)
  - KSC LC-39A seems to be the ideal launch site because of success rate of over 75%
  - Drone ships are the best recovery methods
- Predictive analysis results
  - Each model performed almost equally predicting a recovery outcome rate of 83.33%

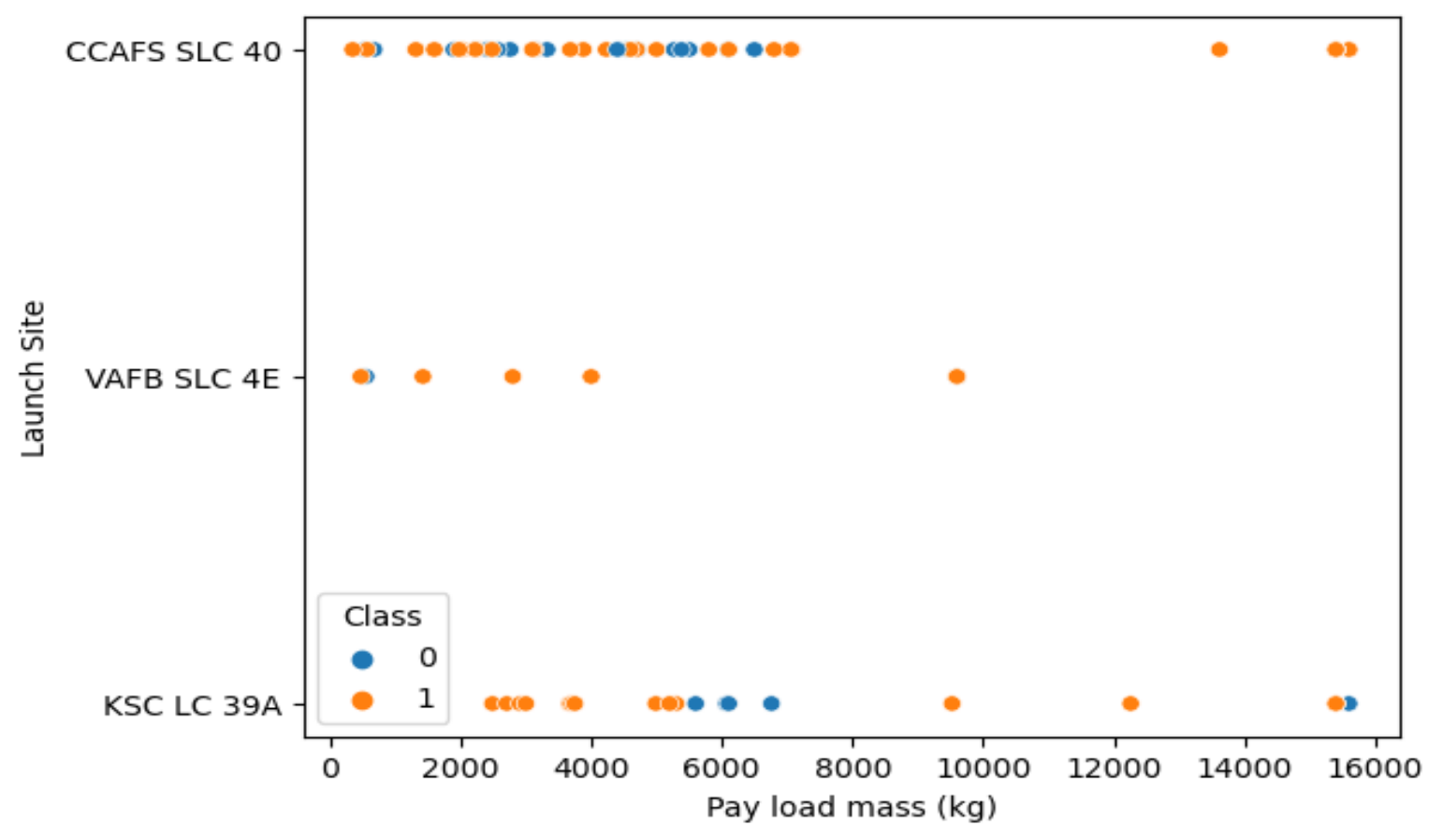

Section 2
Insights drawn from EDA

Flight Number vs. Launch Site



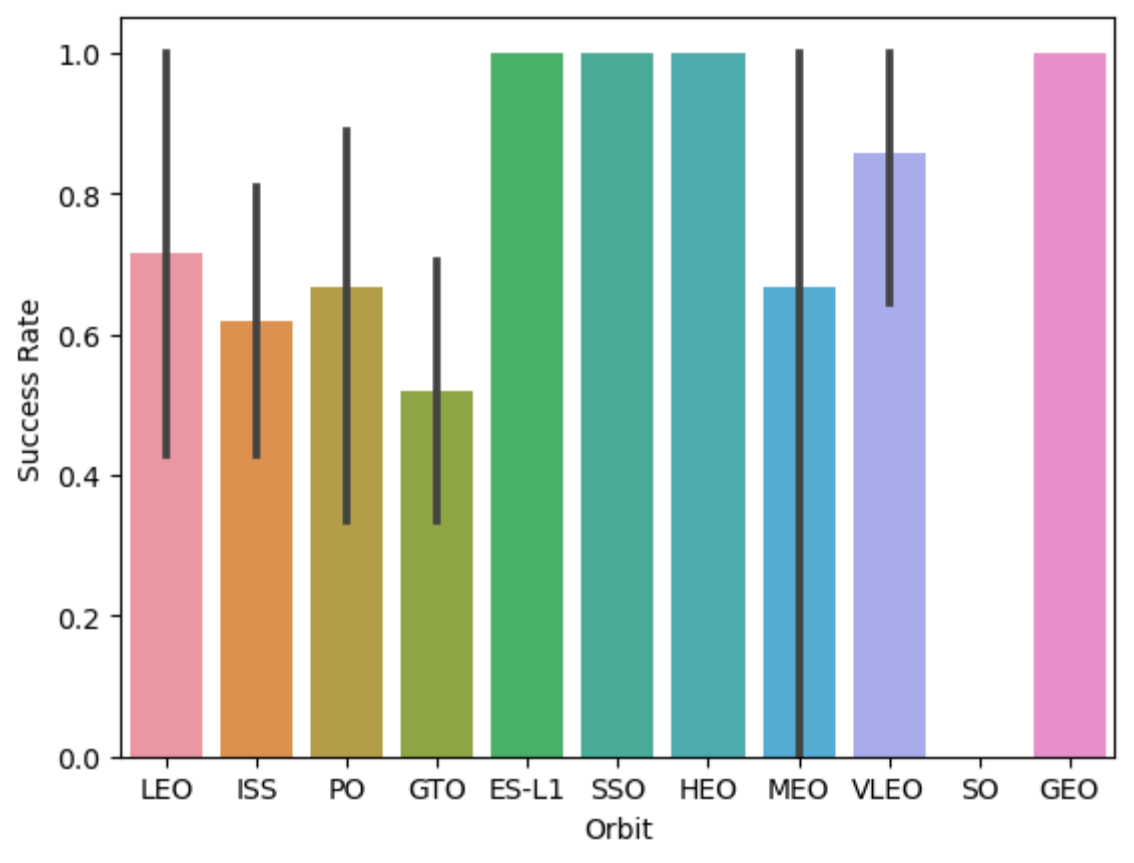
- Success rate has grown over time for each site, KSC LC-39A is the most consistent

Payload vs. Launch Site



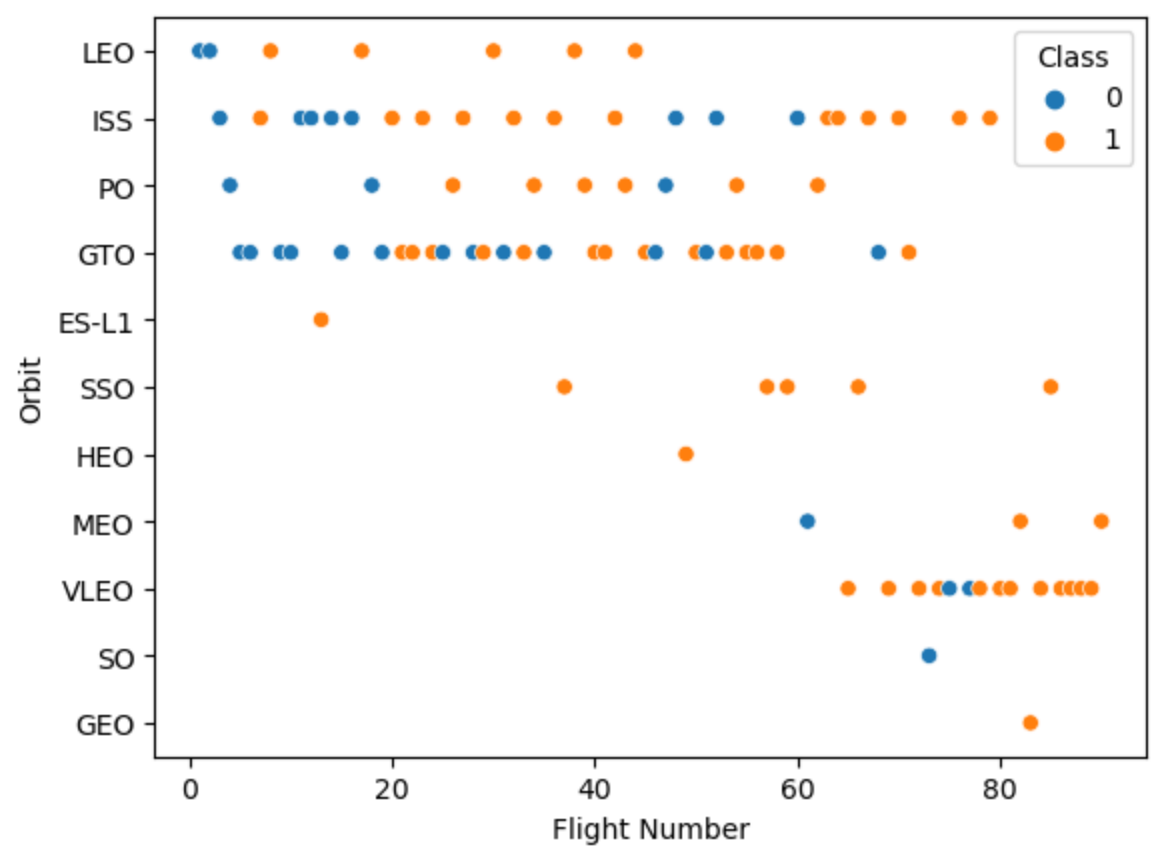
- Smaller payloads correlate with higher success rate (<6500kg-7000kg)

Success Rate vs. Orbit Type



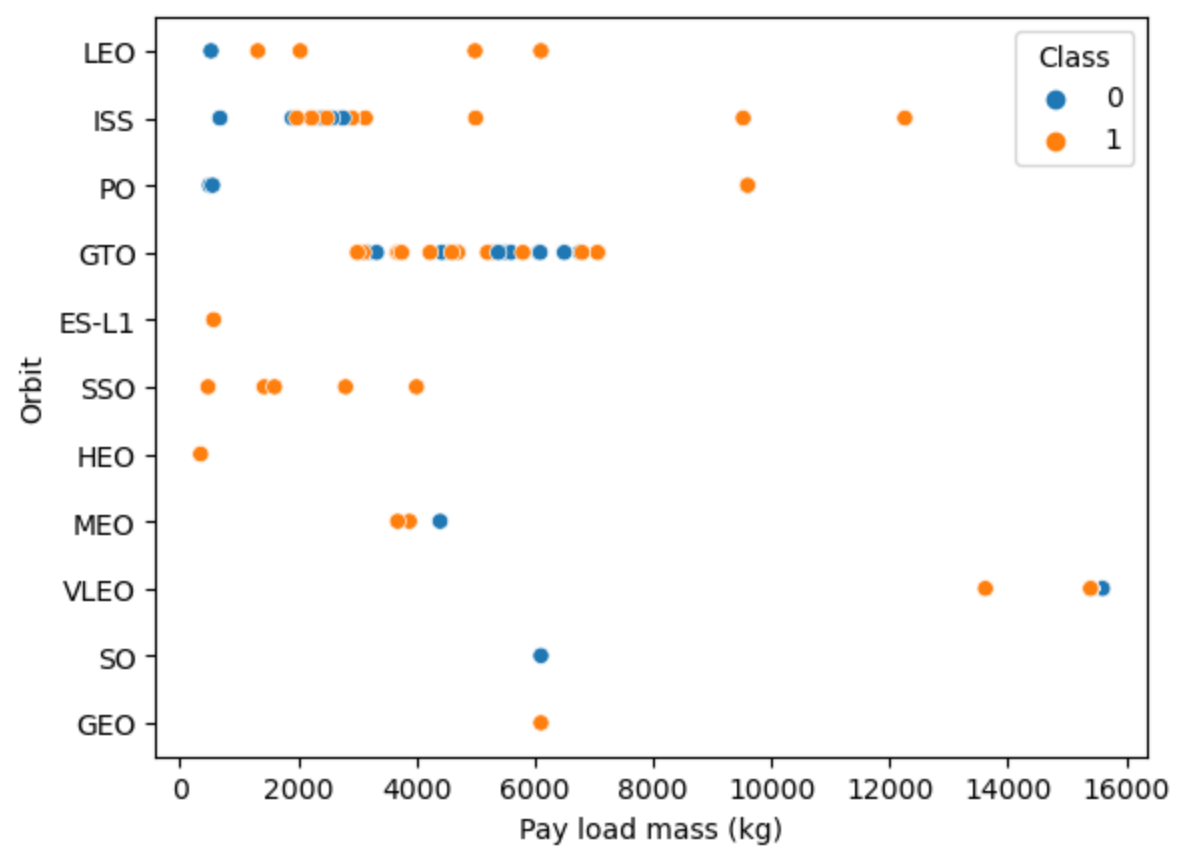
- ES-L1, SSO, HEO and GEO have the highest success rate

Flight Number vs. Orbit Type



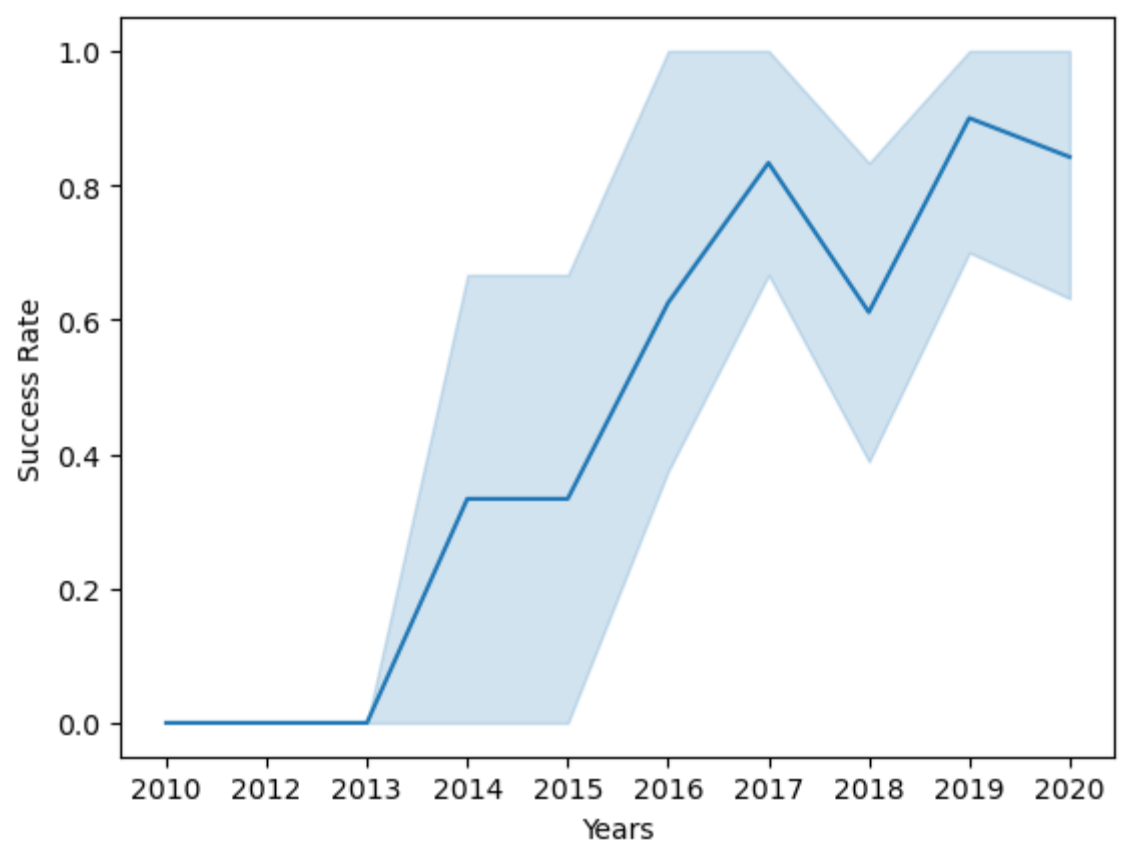
- LEO, SSO and VLEO have high success rate while having sample size
- Each ES-L1, HEO and GEO have only one site

Payload vs. Orbit Type



- LEO and SSO orbits success due light payload

Launch Success Yearly Trend



- Success rate has increased from 2013 to 2017

All Launch Site Names

- We have 4 launch sites
  - CCAFS LC-40
  - CCAFS SLC-40
  - KSC LC-39A
  - VAFB SLC-4E

Launch Site Names Begin with 'CCA'

- These are the firs 5 launch sites that begin with CCA

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	07:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525	LEO (ISS)	NASA (COTS)	Success	No attempt
2012-10-08	00:35:00	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

Total Payload Mass

- Total payload mass carried by boosters from NASA is 111.268

Average Payload Mass by F9 v1.1

- Average payload mass carried by booster version F9 v1.1 is 2534

First Successful Ground Landing Date

- First successful landing outcome on ground pad occurred on 04.06.2010

Successful Drone Ship Landing with Payload between 4000 and 6000

- These are the names of boosters which have successfully landed on drone ship and had payload mass greater than 4000 but less than 6000
  - F9 FT B1022
  - F9 FT B1026
  - F9 FT B1021.2
  - F9 FT B1031.2

Total Number of Successful and Failure Mission Outcomes

- The total number of successful and failure mission outcomes is 101 from where 100 with success and 1 failure

Boosters Carried Maximum Payload

- Booster which have carried the maximum payload mass are:
  - 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

- The failed landing outcomes in drone ship, their booster versions, and launch site names in year 2015 are:

DATE	landing__outcome	booster_version	launch_site
2015-01-10	Failure (drone ship)	F9 v1.1 B1012	CCAFS LC-40
2015-04-14	Failure (drone ship)	F9 v1.1 B1015	CCAFS LC-40

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

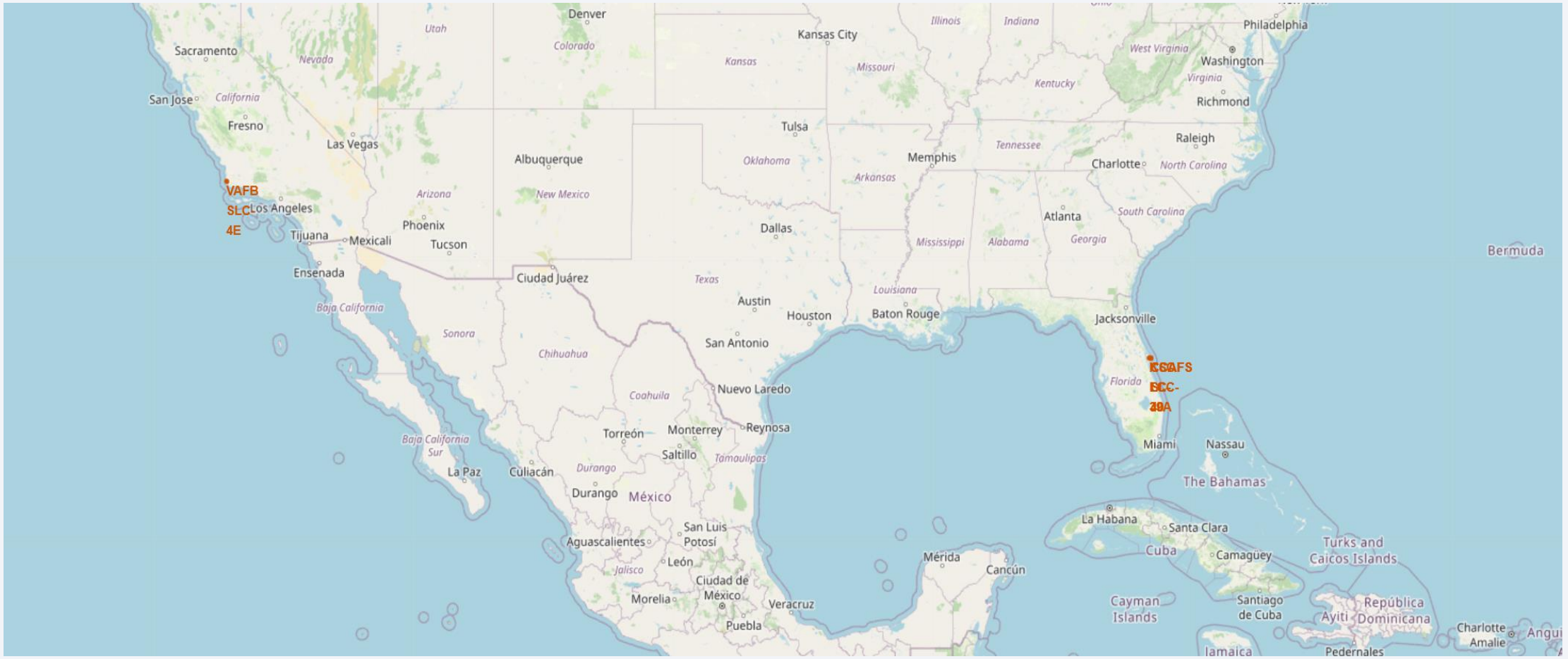
- Rank of landing outcomes from 2010-06-04 and 2017-03-20, in descending order

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

Section 3

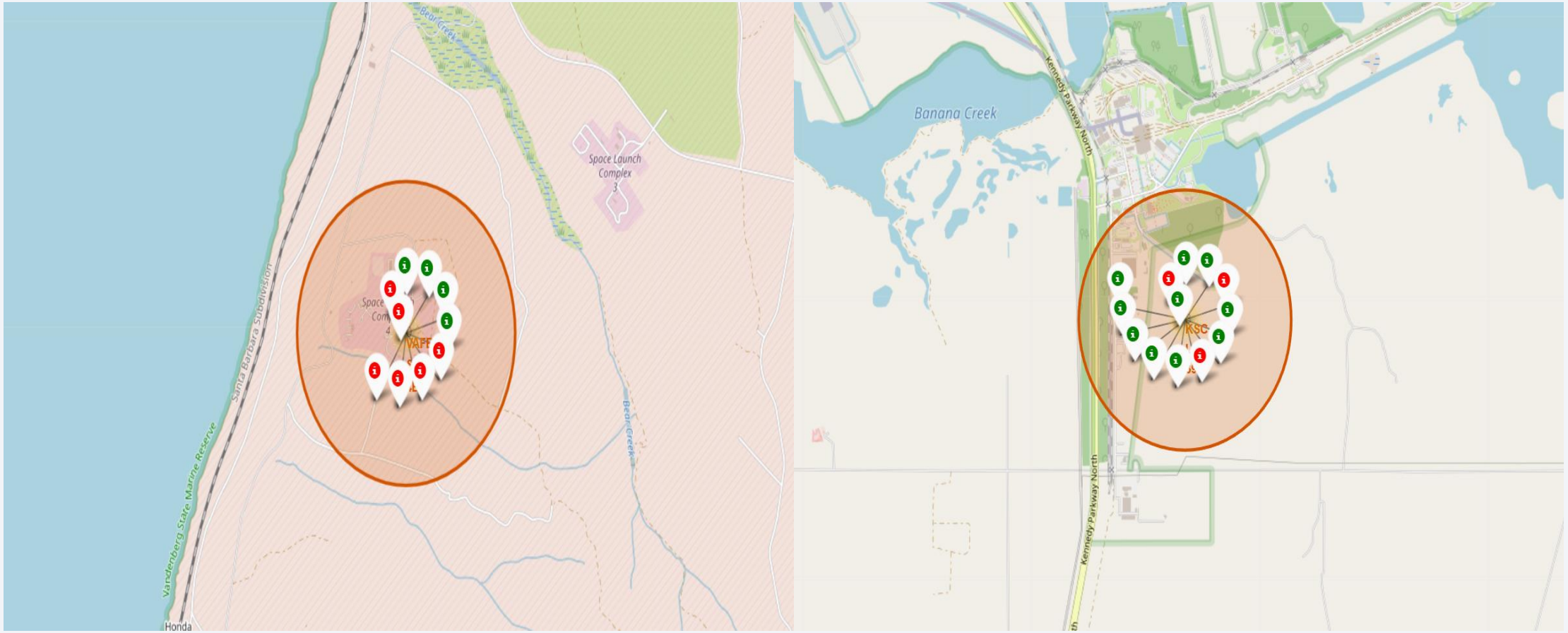
Launch Sites Proximities Analysis

Launch Site Locations Map



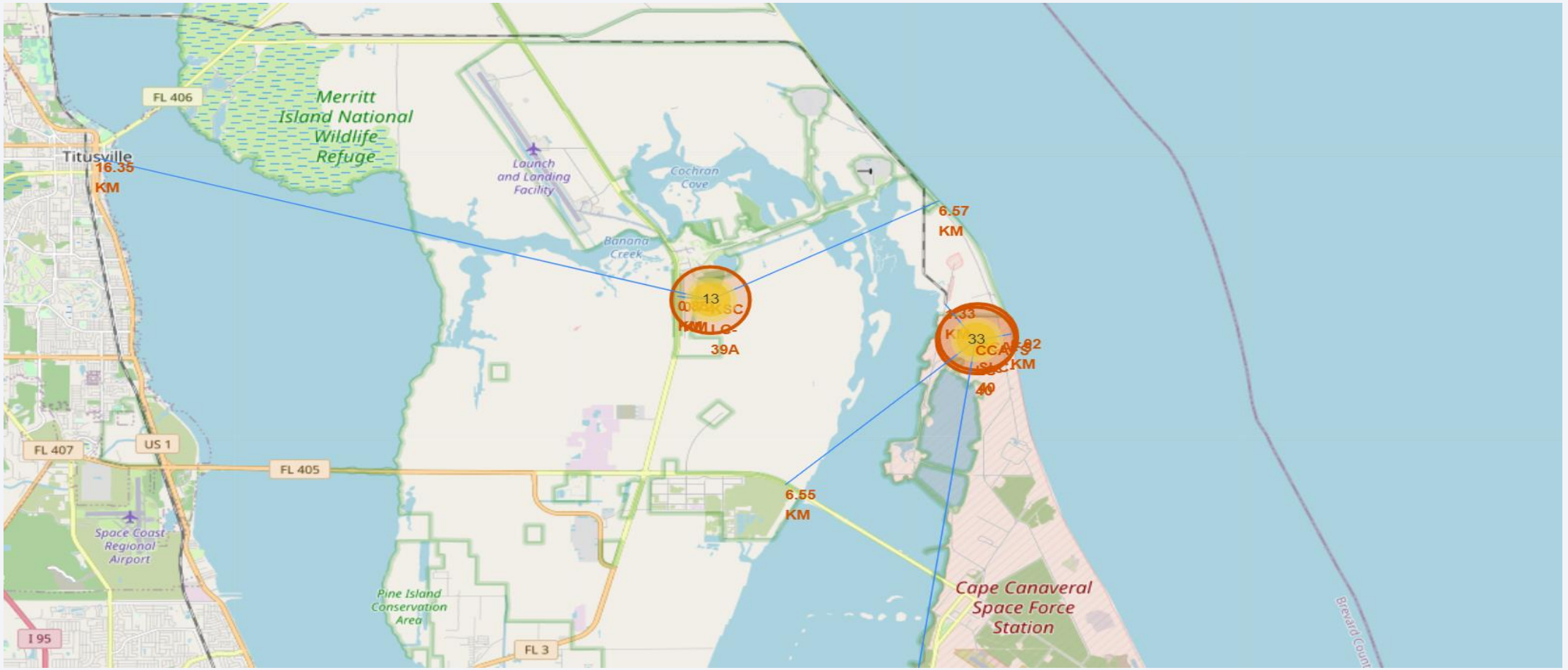
- SpaceX has one launch site on Pacific coast (California) and the rest on Atlantic coast (Florida)

VAFB SLC-4E and KSC LC-39 Recovery Outcomes

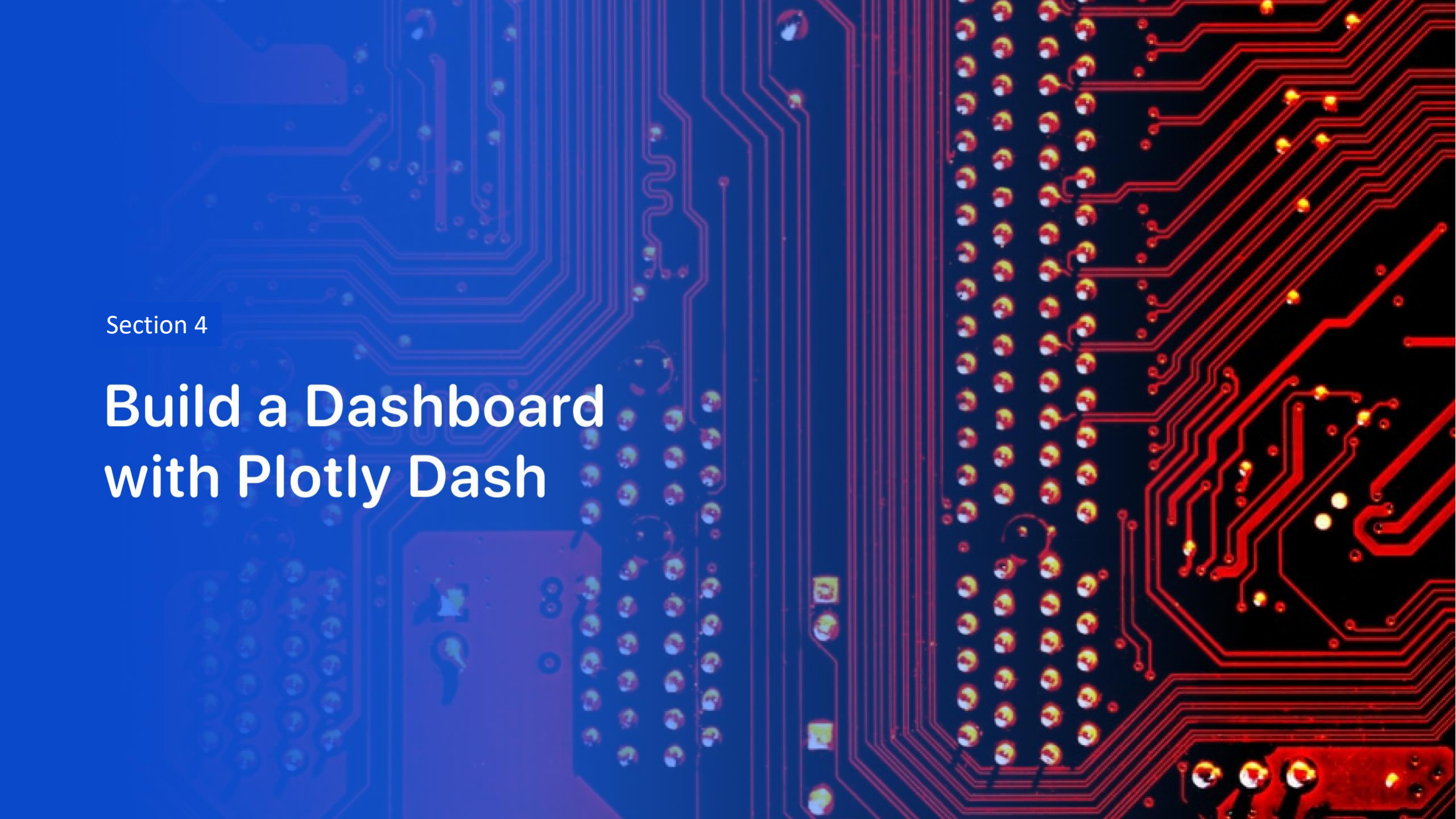


- Green markers are successful recovery of Falcon 9 Stage One and red ones are unsuccessful recovery

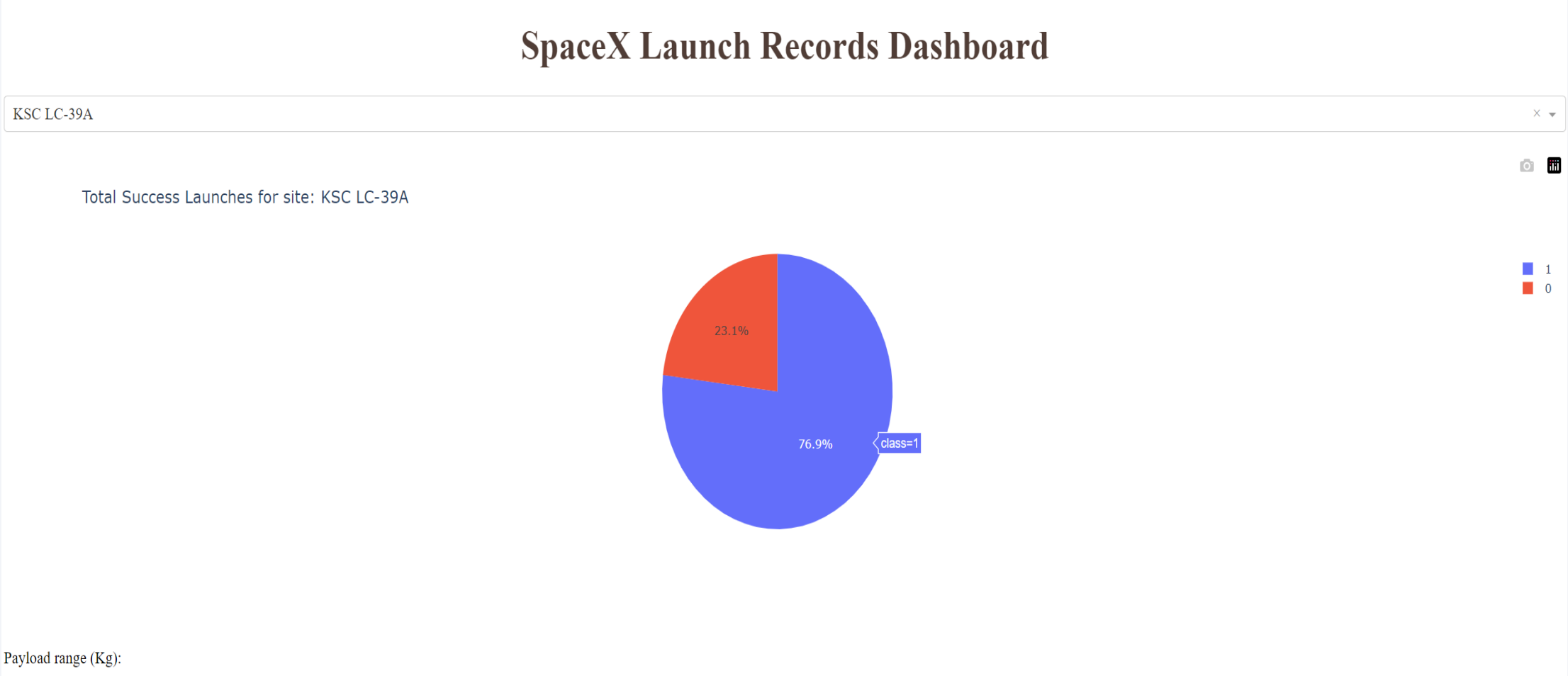
KSC LC-39A, CCAFS LC-40 and CCAFS SLC-40 Nearby Locations



- Blue lines represents the distance to the nearest coastline, city/town, railway and highway

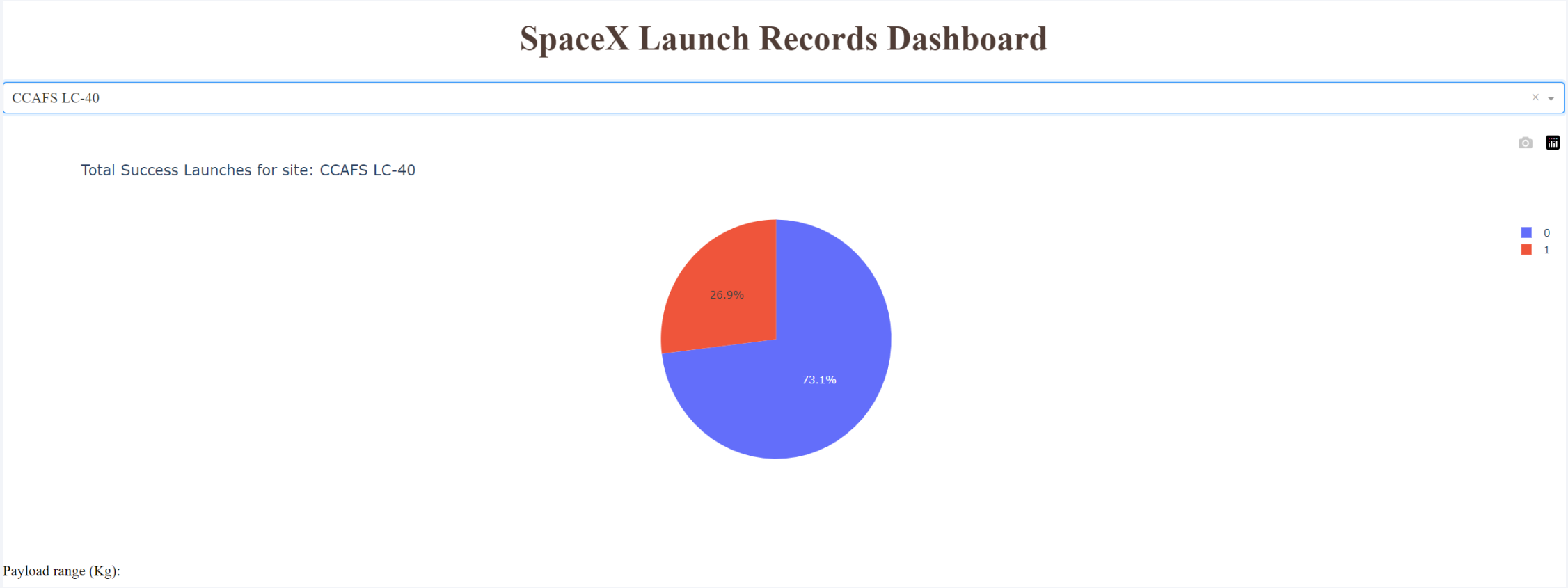

Section 4
Build a Dashboard with Plotly Dash

KSC LC-39A Successful Launches



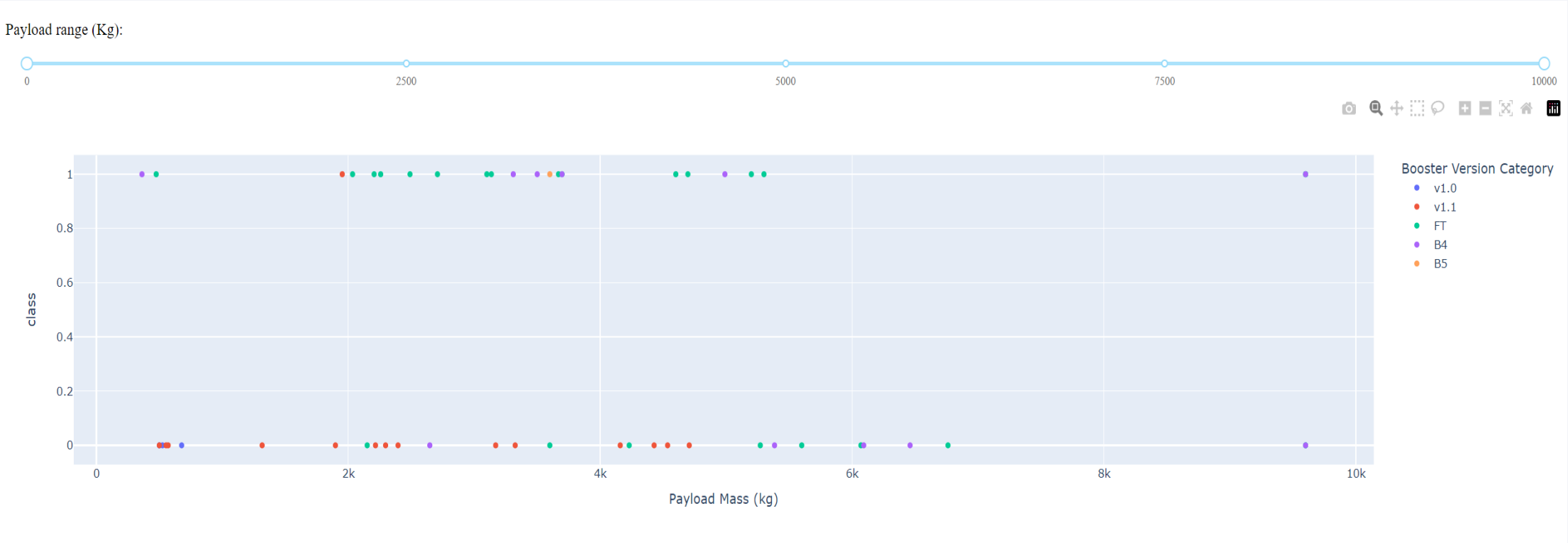
- KSC LC-39A have fewer launches but the success rate is high

CCAFS LC-40 Successful Launches



- CCAFS LC-40 have the most launches and it has the most successful launches

Recovery Outcome vs Payload Mass



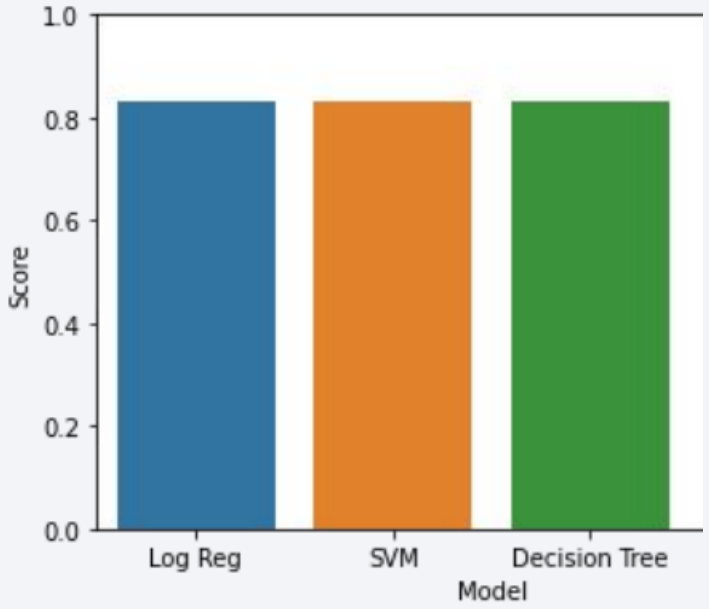
- Between 2000 and 4000 are the most successful payload range
- FT booster is very effective and v1.1 have the most failure

Section 5

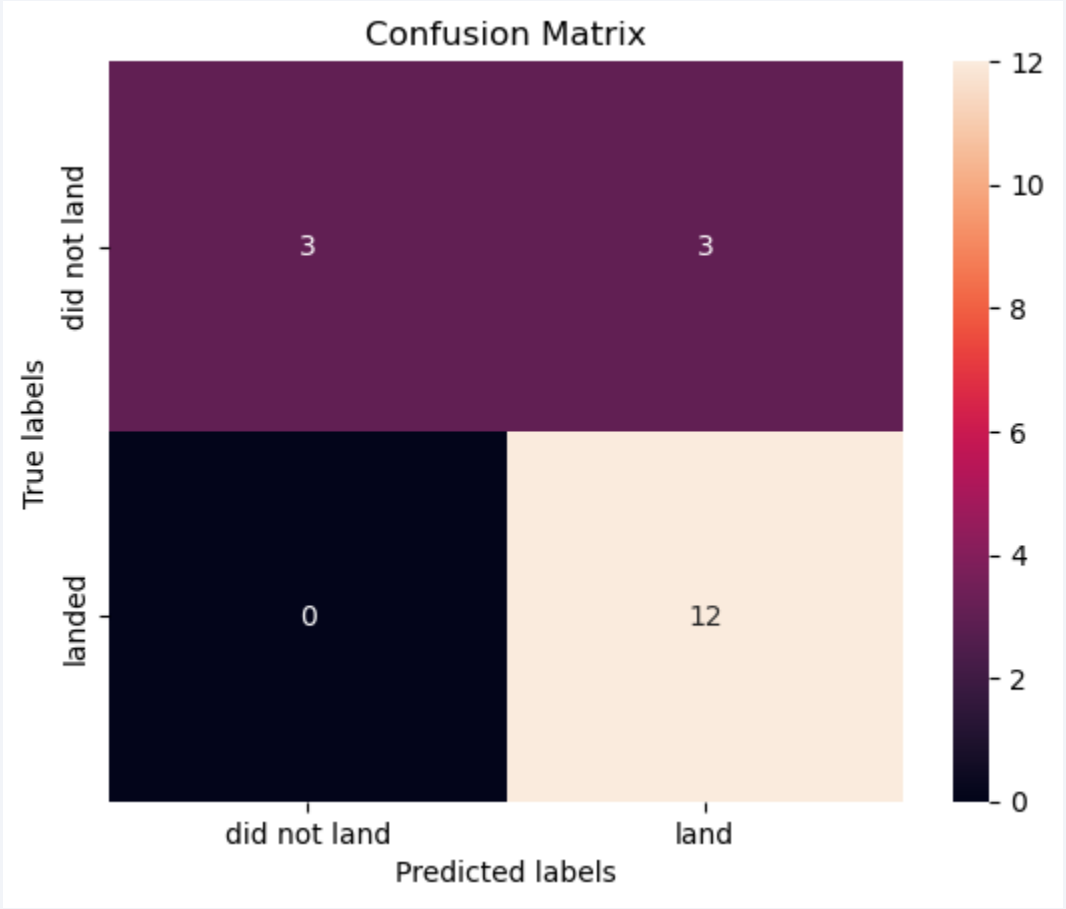
Predictive Analysis (Classification)

Classification Accuracy

- Log Reg has 83,33% accuracy
- SVM has 83,33% accuracy
- Decision Tree has 77,77% accuracy
- KNN has 83,33% accuracy
- So Log Reg, SVM and KNN have the same and the best accuracy



Confusion Matrix



- Overall the model is mostly accurate, however it wrongly predicted 3 landings as successful but they are not

Conclusions

- Falcon 9 Stage One is good for light payloads (between 2000kg and 4000kg)
- Best launching site is KSC LC-39A
- Most successful launches are after 2017
- The best method to recover is via drone ship
- Model can predict the outcome of the given recovery with a accuracy of 83,33%

Thank you!