IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Serhat Eren Onal
03.03.2023

Outline

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

Executive Summary

- Data collection with API
- Data collection with Web Scraping
- Data Wrangling
- Exploratory Data analysis with SQL
- Exploratory Data analysis with Visualization
- Interactive Visual Analytics with Folium
- Interactive Visual analytics with Dashboard
- Machine Learning Prediction
- SUM OF ALL RESULST
- Exploratory Data analysist result
- Prediction analytics result

Introduction

- Project background and context
- SpaceX try to introduce the Falcon 9 on its website because Falcon 9 cheaper than other launches and they can reusable. So we can determine the percetange the successfully landed so we can determine real cost the launch . Project aim is c use the machine learning and try to forecast the successfully landed.

- Problems you want to find answers
- what does success depend on(location,launcher,launch model etc.)
- is it worth or just some asparagus

Section 1

Methodology

Methodology

Executive Summary

- Data collection methodology:
  - Data was collected using api,webscarping from WP
- Perform data wrangling
  - One-hot-encoding
- 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

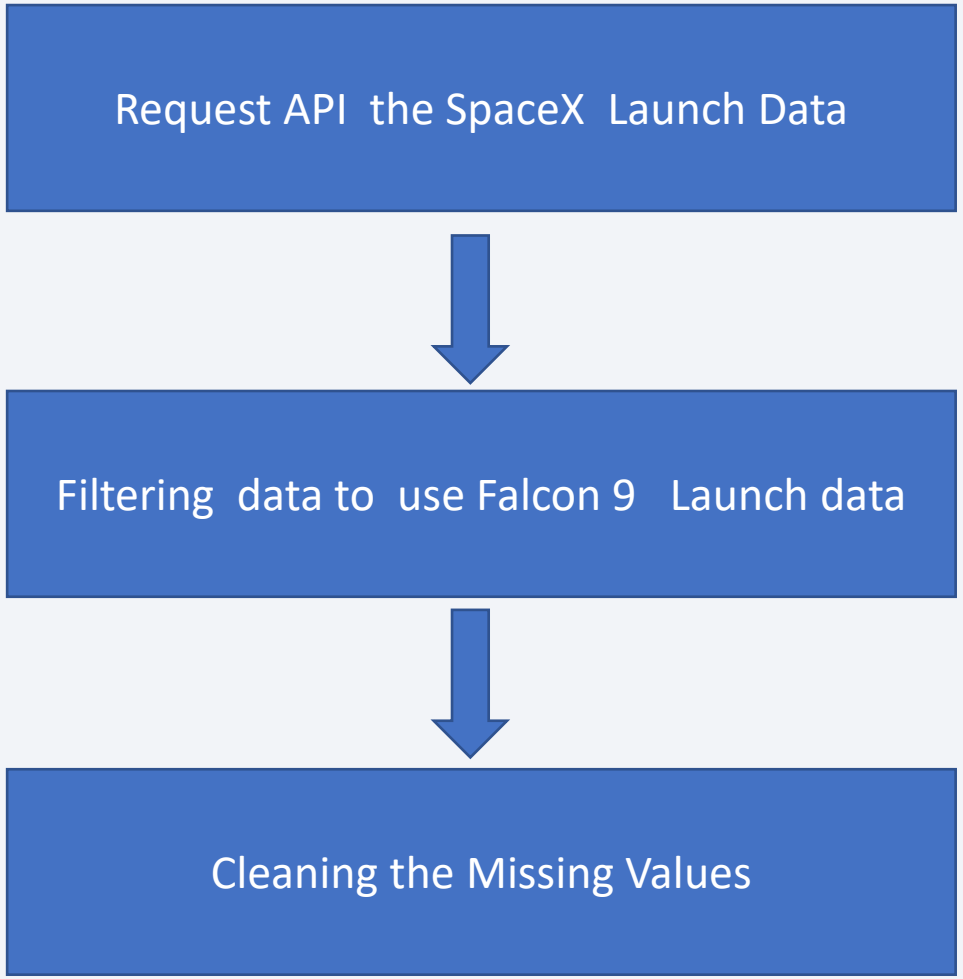
Data Collection

- Data sets were collected from SpaceX API and Wikipedia using WebScraping

Data Collection – SpaceX API

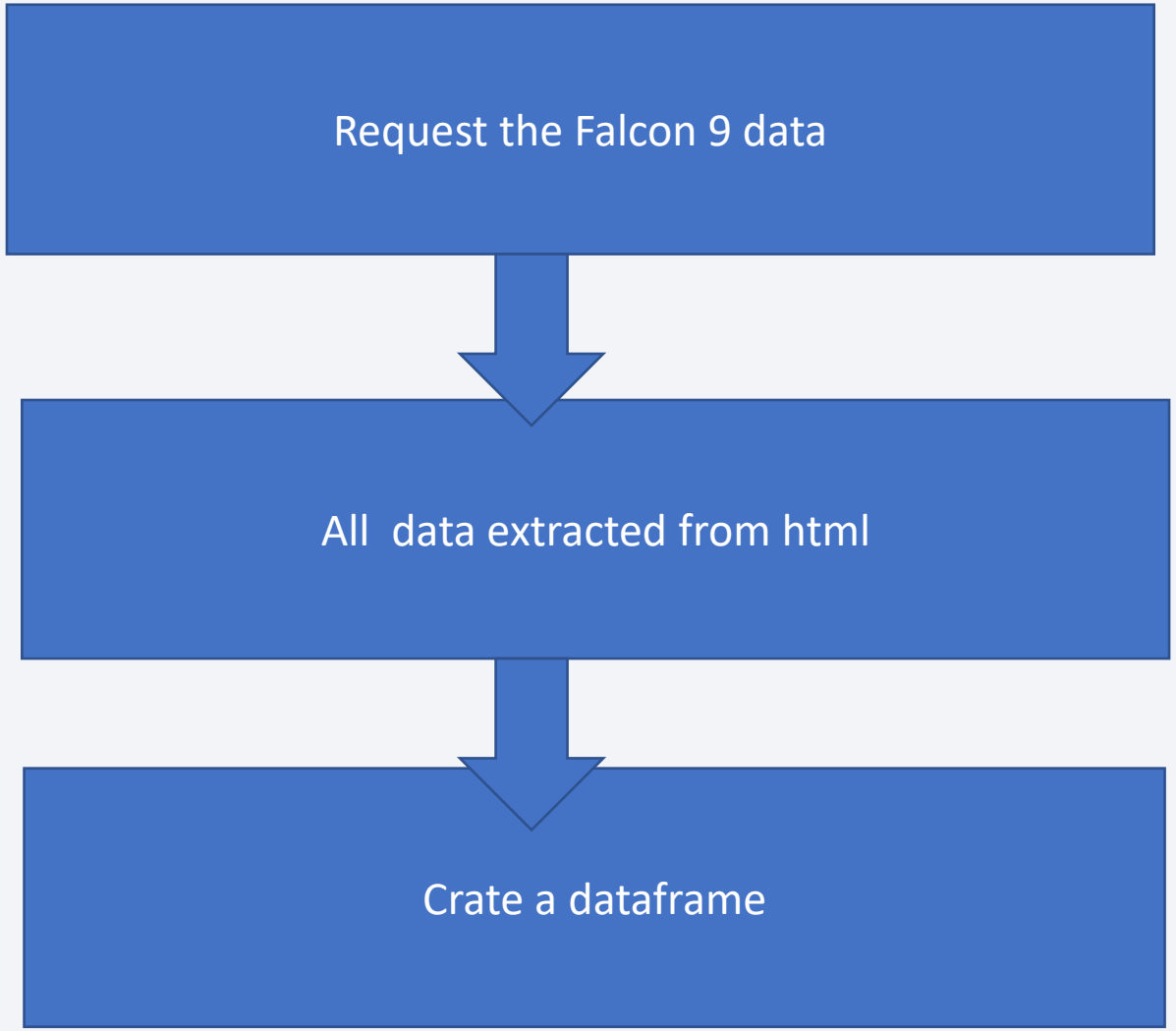
- We used the get request to the get data

- 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



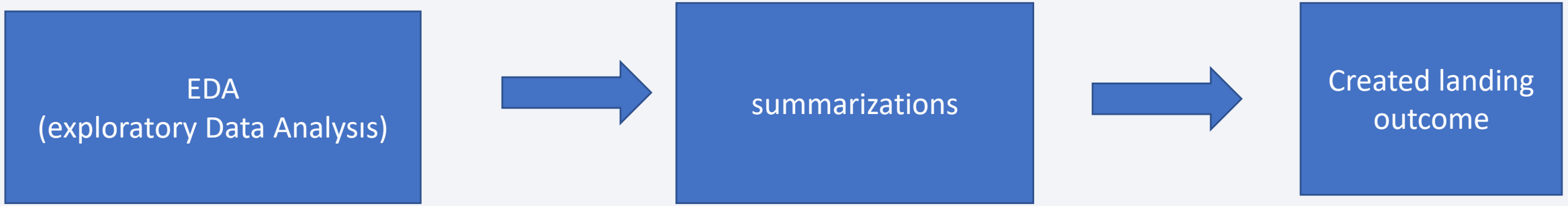
Data Collection - Scraping

- We use the BeautifulSoup to WebScaping to reach Falcon 9 Data
- Add the GitHub URL of the completed web scraping notebook, as an external reference and peer-review purpose



Data Wrangling

- We calculated the number of launches at various features to orbit range
- Created landing outcome later save as the csv
- Add the GitHub URL of your completed data wrangling related notebooks, as an external reference and peer-review purpose



EDA with Data Visualization

- Scatterplots and barplots were used to visualization

Link is the EDA with DATA vis.

EDA with SQL

- Names of the unique launch sites in the space mission
- Top 5 launch start with CCA
- Total payload which booster launched by CRS
- Average payload mass F9 v1.1
- Total number of successful and failure mission outcomes

Link for code

Build an Interactive Map with Folium

- Marker circles lines and marker cluster were used with folium maps package
- Marker indicate points launch sites
- Circles indicate highlated Booster location
- Marker cluster used to groups of each coordinate

HERE CODE

Build a Dashboard with Plotly Dash

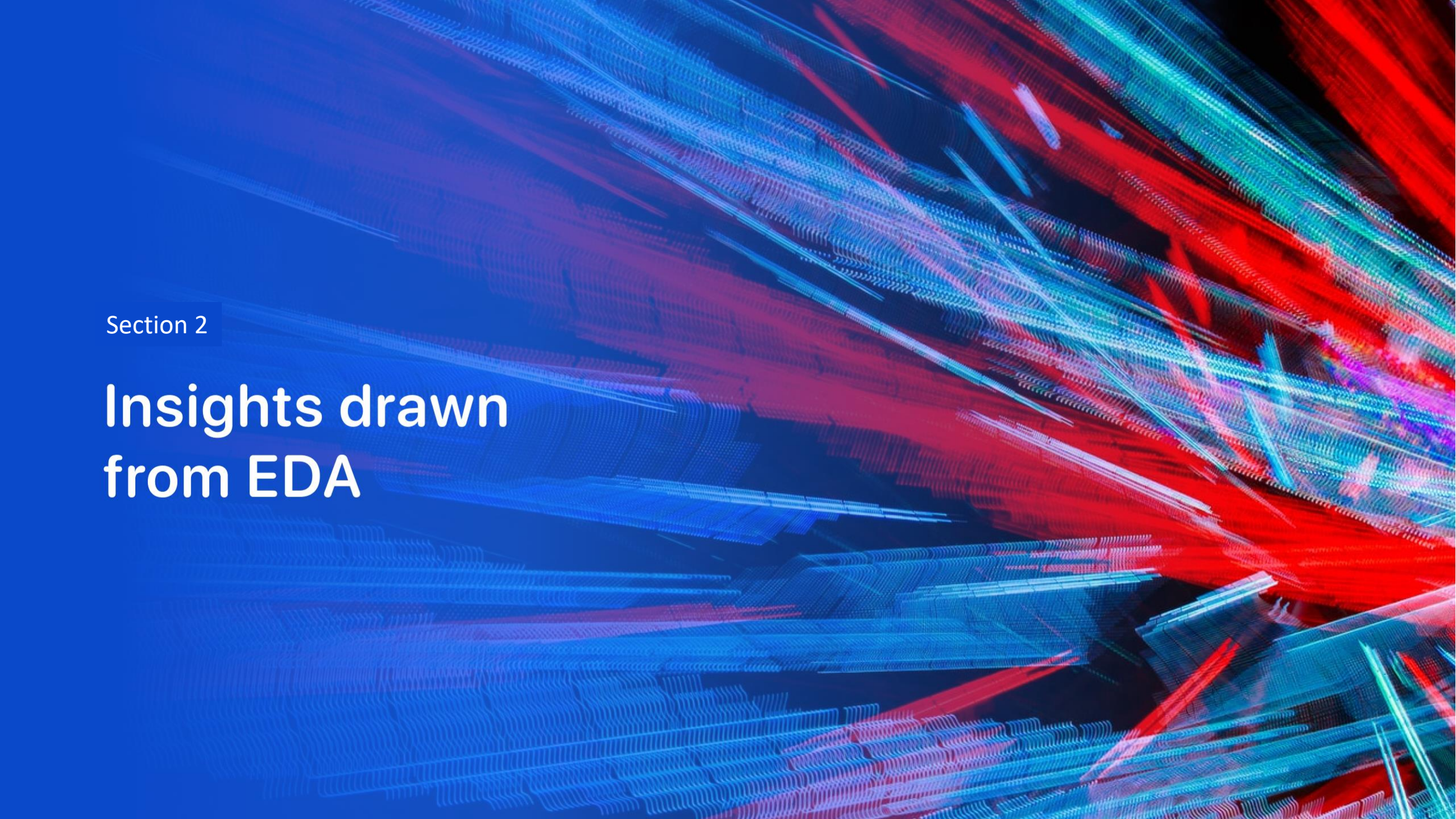
- Summarizing graph and plots to use the create a dashboard to visualization
- Build interactive dashboard with plotly
- Pie char to total launch by sites
- Scatter plot to showing relation outcome and payload mass
- link

Predictive Analysis (Classification)

- Loaded the data using nummpy,pandas and transformed the data split our data train and test
- Built different machine learning models and try to find best parameters with GridSearchCV
- Used accuracy to scoring and try to improve with this scoring
- Finally found the best parameters and model to classification model to out project
- link

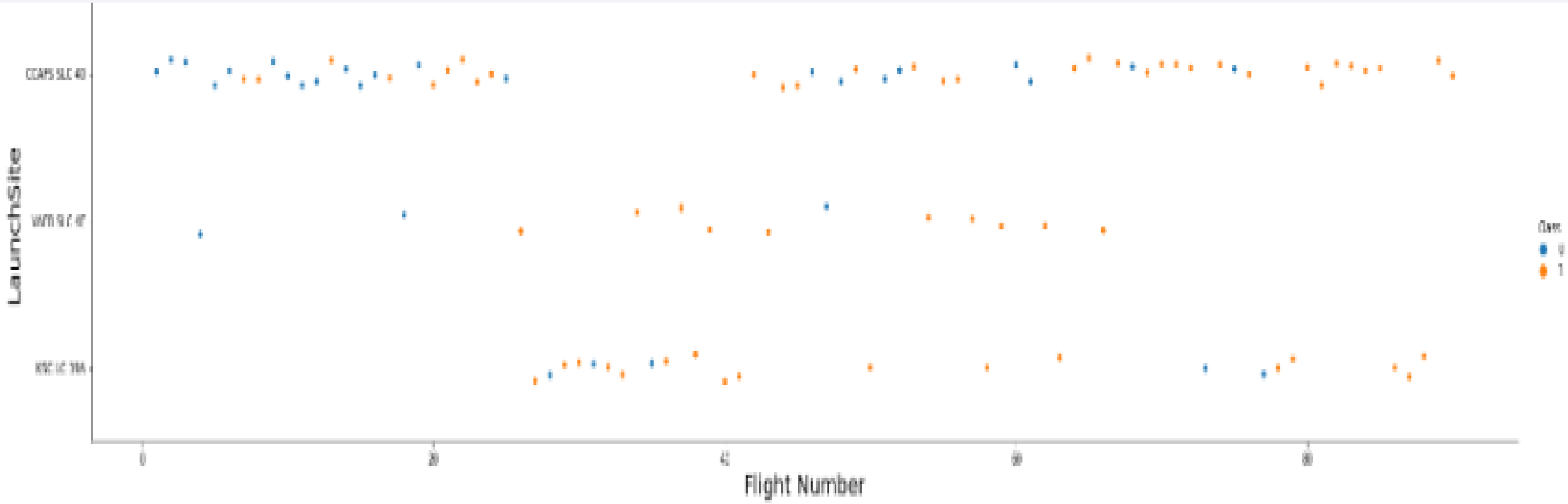
Results

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


Section 2
Insights drawn from EDA

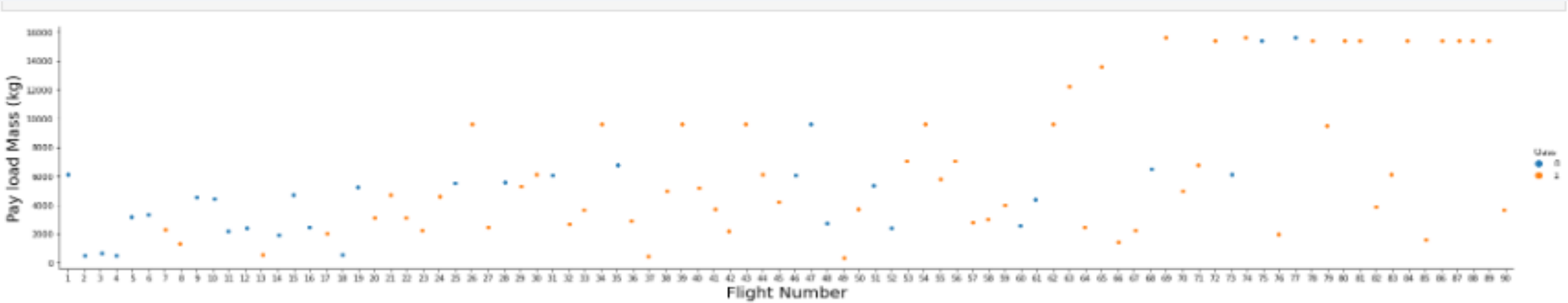
Flight Number vs. Launch Site

The plot shows that best launch is CCAF5 SLC 40 because this launch have more succces rate than others



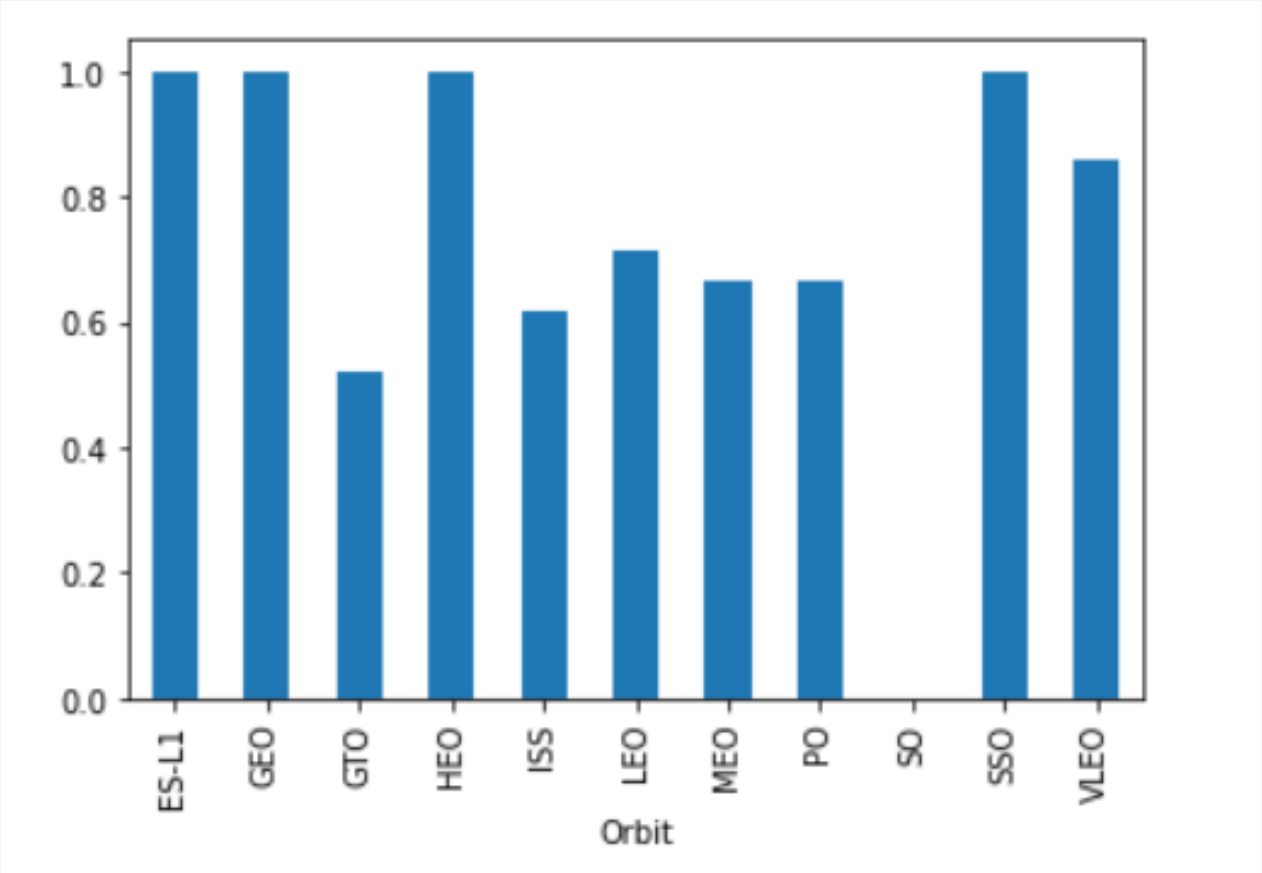
Payload vs. Launch Site

- Over the 9000kg have perfect success rate
- More payload is get higher success rate



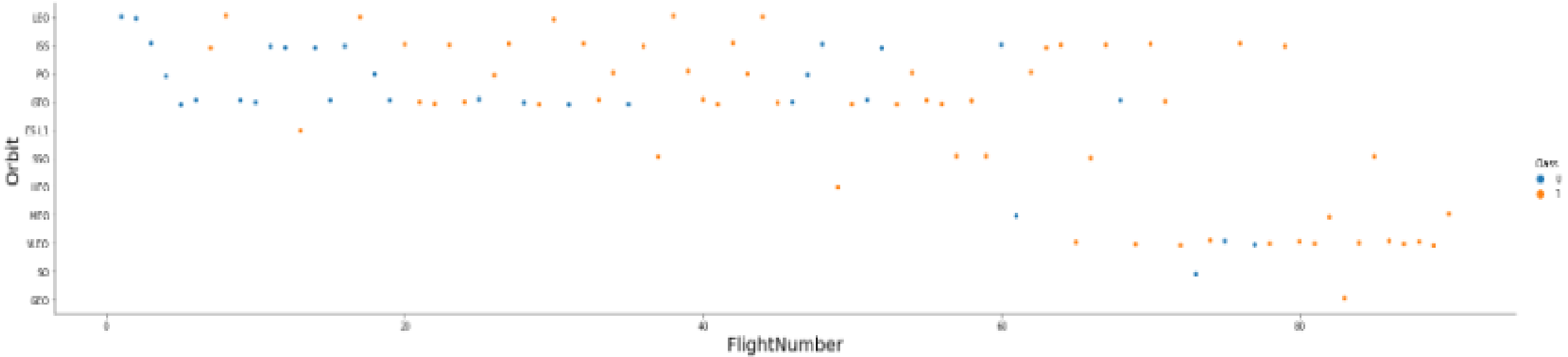
Success Rate vs. Orbit Type

- Succes Rate
- 1-ES-L1
- GEO
- HEO
- SSO



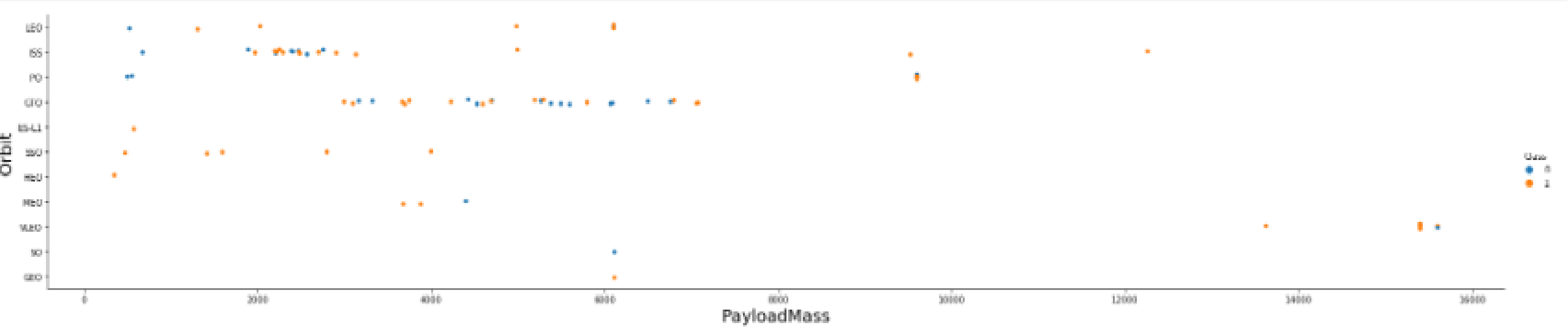
Flight Number vs. Orbit Type

- Over the time success rate get higher and in the recent time VLEO is using more oftenly than others



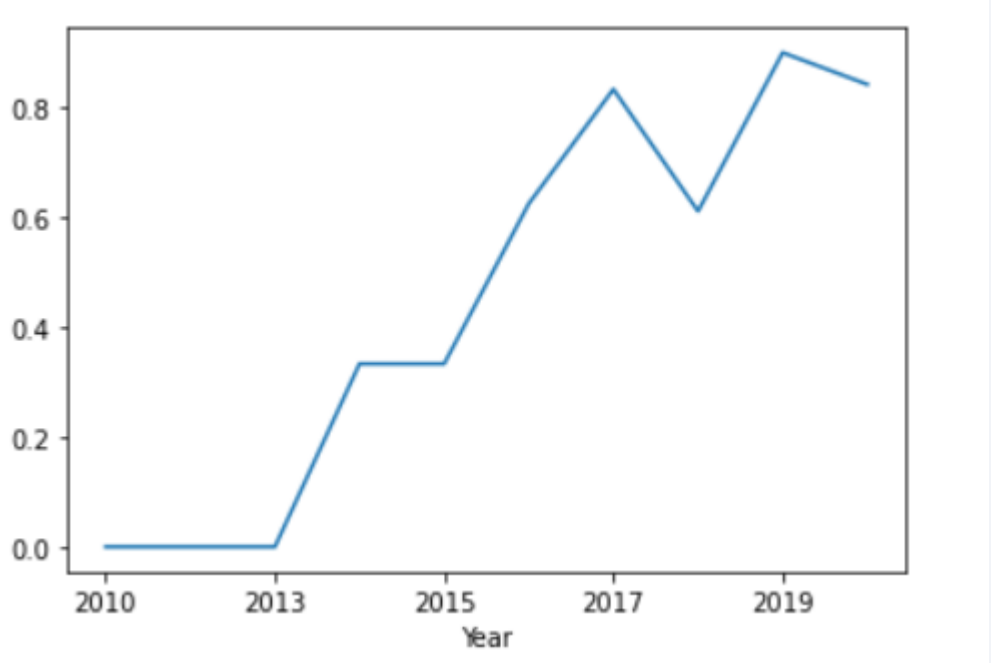
Payload vs. Orbit Type

- Payload and Orbit type have no Relationship



Launch Success Yearly Trend

- From the plot we can oberve that success rate generally always get higher



All Launch Site Names

- There are 4 launchsite
- Launch_site values

Launch Site
CCAFS LC-40
CCAFS SLC-40
KSC LC-39A
VAFB SLC-4E

Launch Site Names Begin with 'CCA'

- 5 records where launch sites 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	**CCA**FS LC-40	Dragon Spacecraft Qualification Unit	0	LEO	SpaceX	Success	Failure (parachute)
2010-12-08	15:43:00	F9 v1.0 B0004	**CCA**FS 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	**CCA**FS LC-40	Dragon demo flight C2	525	LEO (ISS)	NASA (COTS)	Success	No attempt
2012-10-08	00:35:00	F9 v1.0 B0006	**CCA**FS LC-40	SpaceX CRS-1	500	LEO (ISS)	NASA (CRS)	Success	No attempt
2013-03-01	15:10:00	F9 v1.0 B0007	**CCA**FS LC-40	SpaceX CRS-2	677	LEO (ISS)	NASA (CRS)	Success	No attemp

Total Payload Mass

- Total patload carried by boosters from nasa

Total Payload (kg)
111.268

Average Payload Mass by F9 v1.1

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

- Average mass is the 2928 kg for F9 v1.1

Avg Payload (kg)
2.928

First Successful Ground Landing Date

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

Booster Version
F9 FT B1021.2
F9 FT B1031.2
F9 FT B1022
F9 FT B1026

Total Number of Successful and Failure Mission Outcomes

- total number of successful and failure mission outcomes

Mission Outcome	Occurrences
Success	99
Success (payload status unclear)	1
Failure (in flight)	1

Boosters Carried Maximum Payload

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

Booster Version (...)
F9 B5 B1048.4
F9 B5 B1048.5
F9 B5 B1049.4
F9 B5 B1049.5
F9 B5 B1049.7
F9 B5 B1051.3

Booster Version
F9 B5 B1051.4
F9 B5 B1051.6
F9 B5 B1056.4
F9 B5 B1058.3
F9 B5 B1060.2
F9 B5 B1060.3

2015 Launch Records

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

Booster Version	Launch Site
F9 v1.1 B1012	CCAFS LC-40
F9 v1.1 B1015	CCAFS LC-40

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

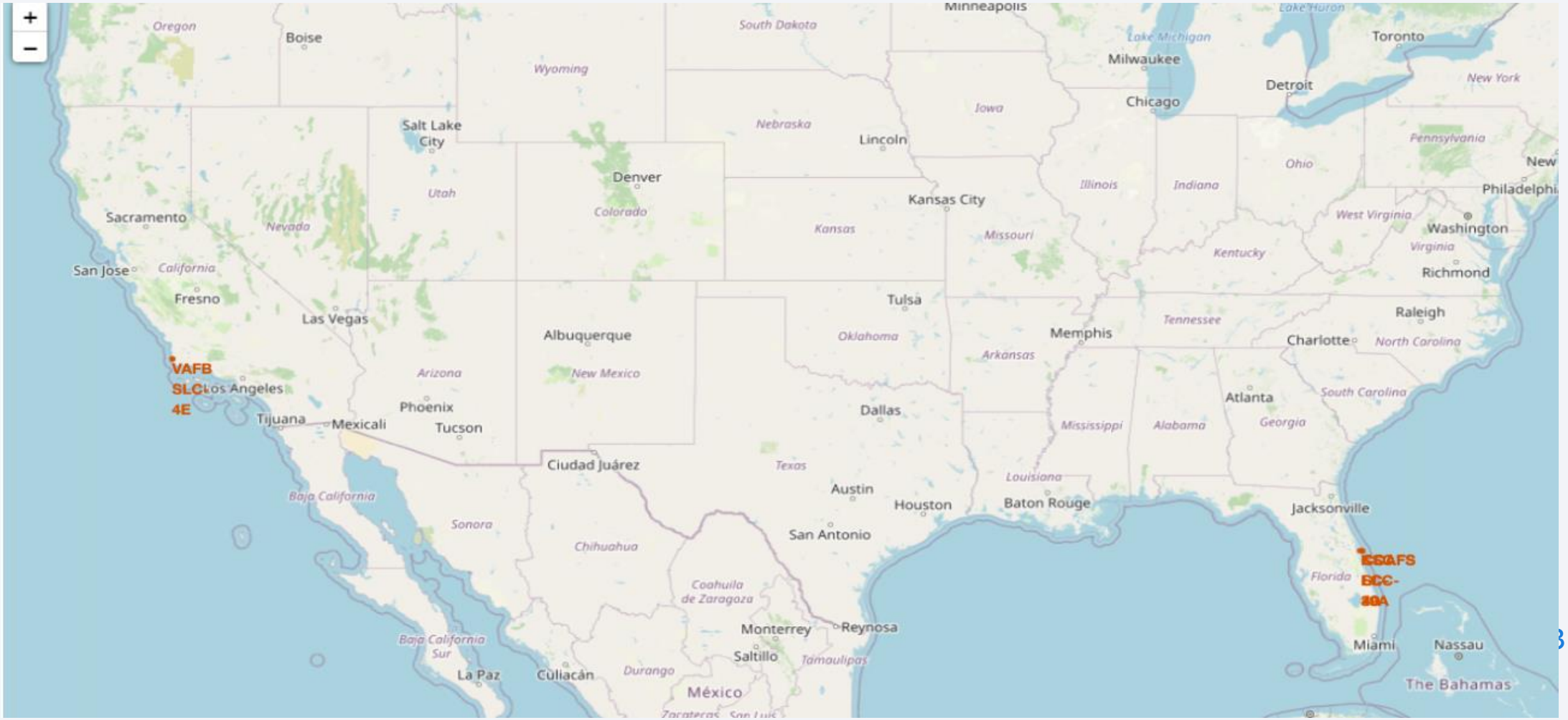
- Ranking of landing outcomes (such as Failure (drone ship) or Success (ground pad)) between the date 2010-06-04 and 2017-03-20, in descending order

Landing Outcome	Occurrences
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

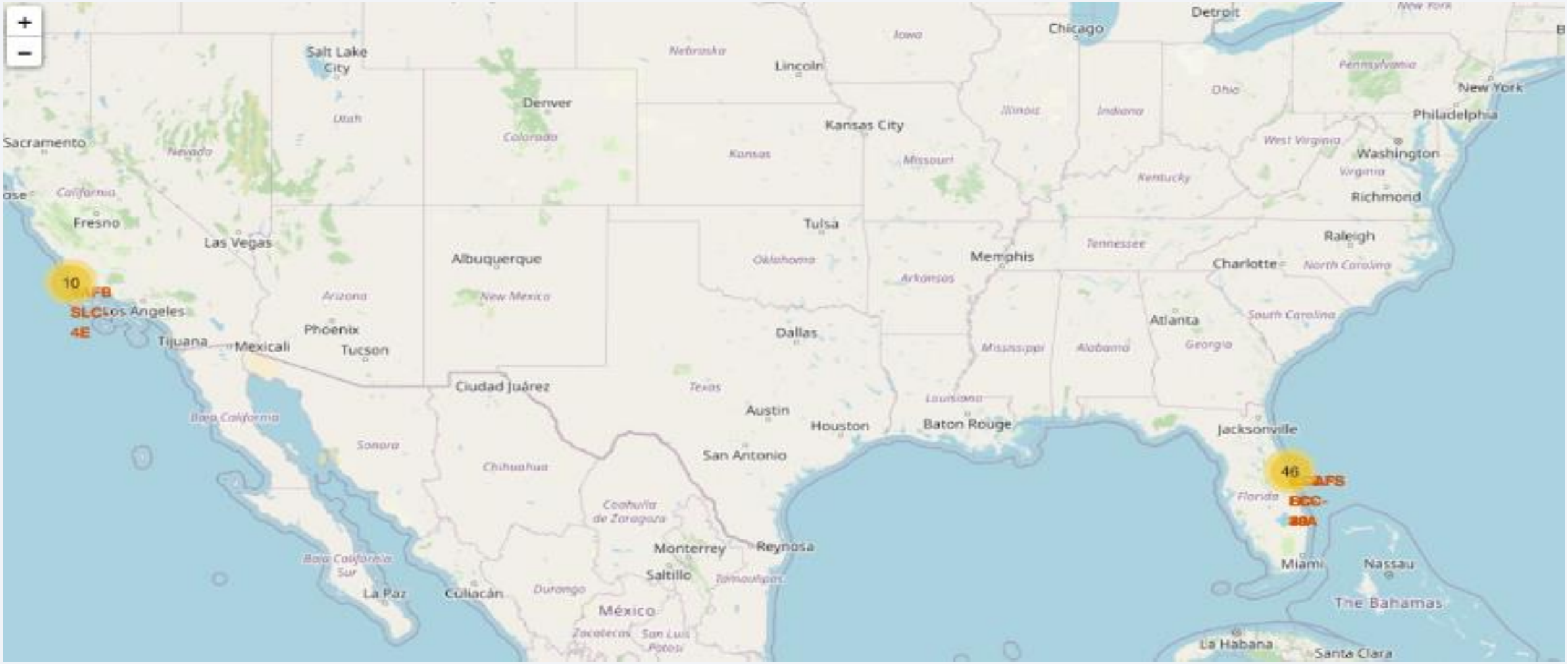

Section 3
Launch Sites
Proximities Analysis

Launch Sites

- Launchsites are near the ocean because of the safety

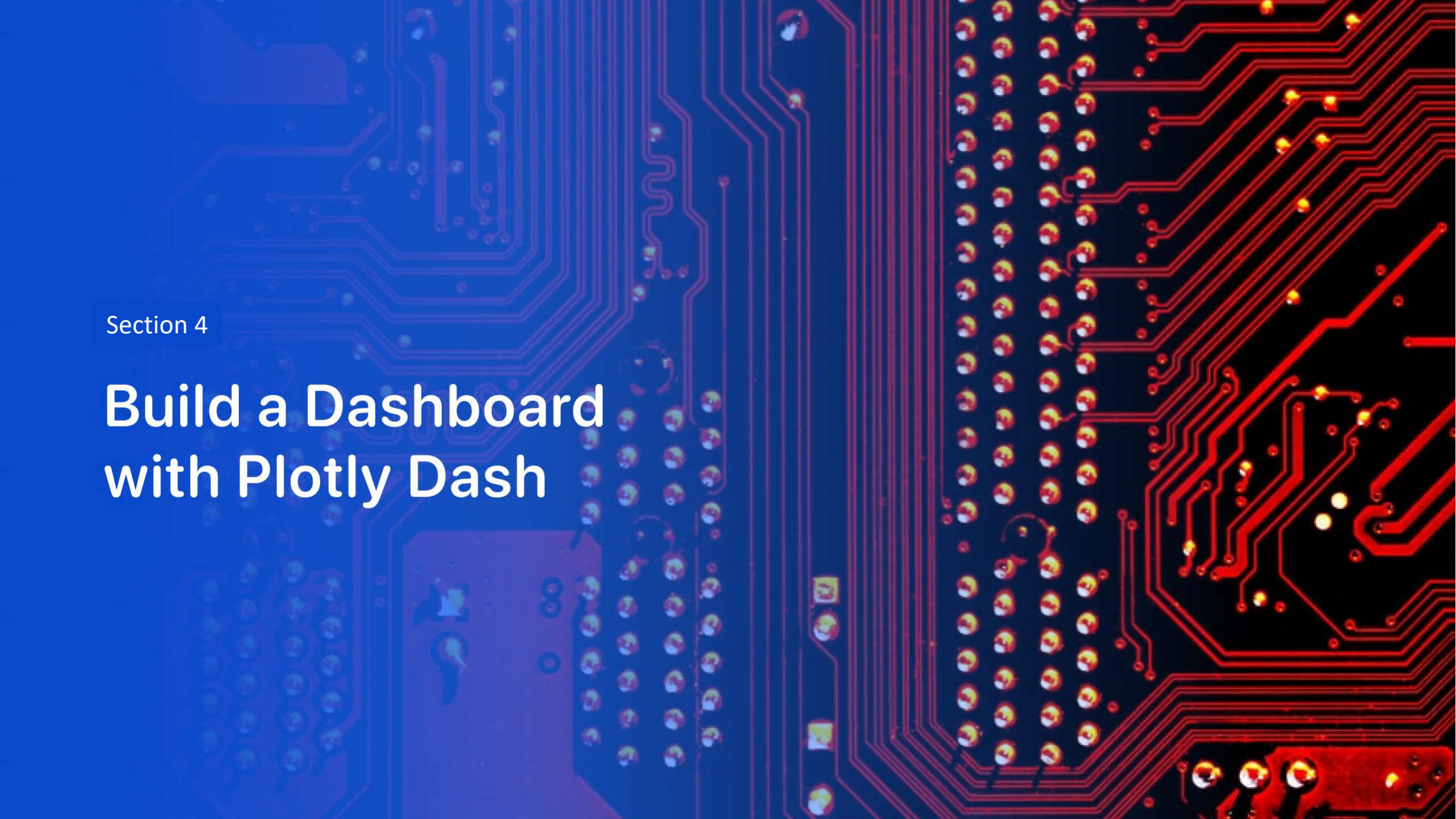


Outcome

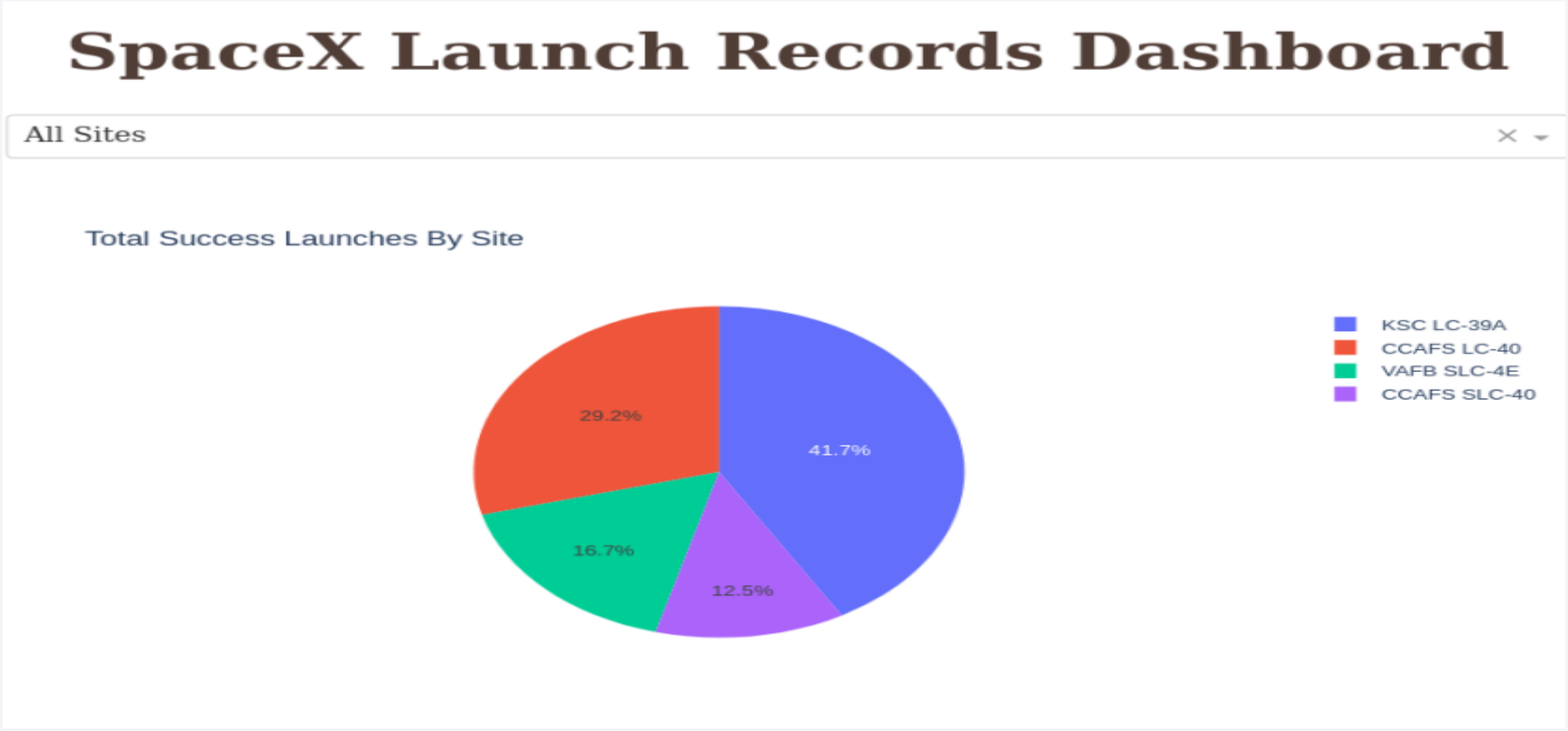

Salt Lake City
Denver
Sacramento
Fresno
Las Vegas
Albuquerque
Los Angeles
Phoenix
Tijuana
Mexicali
Tucson
Ciudad Juárez
Dallas
Austin
Houston
San Antonio
Baton Rouge
Jacksonville
Monterrey
Reynosa
Saltillo
La Paz
Culiacán
México
Miami
Nassau
The Bahamas
La Habana
Santa Clara
Chicago
Detroit
Lincoln
Kansas City
Tulsa
Memphis
Atlanta
Charlotte
Raleigh
Richmond
Washington
Philadelphia
New York
10
46
SLC
4E
BCC
39A
AFS

Distance

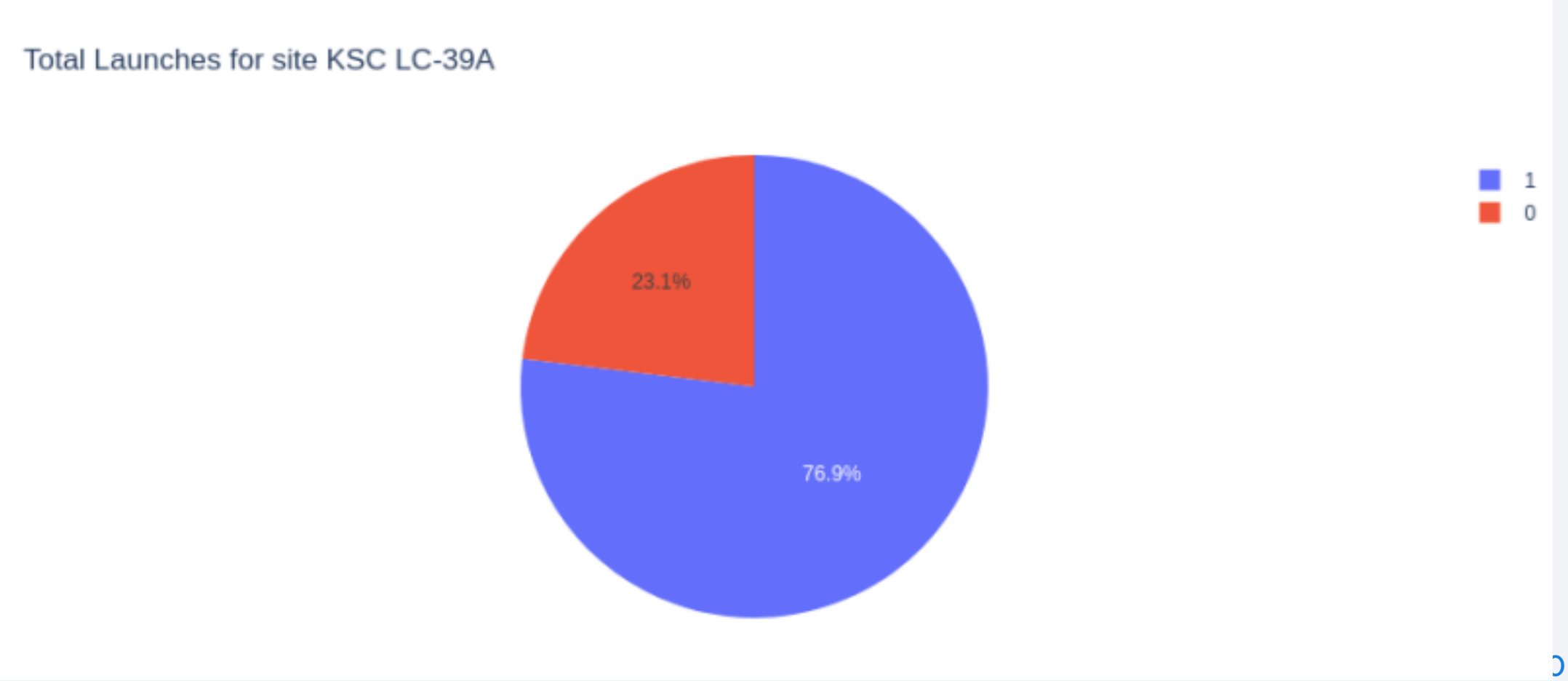

P
P
7
AFS
SLC-
40
26
AFS
LC-
40
Centaur Road
Samuel C Phillips Pkwy
0.90
KM


Section 4
Build a Dashboard with Plotly Dash

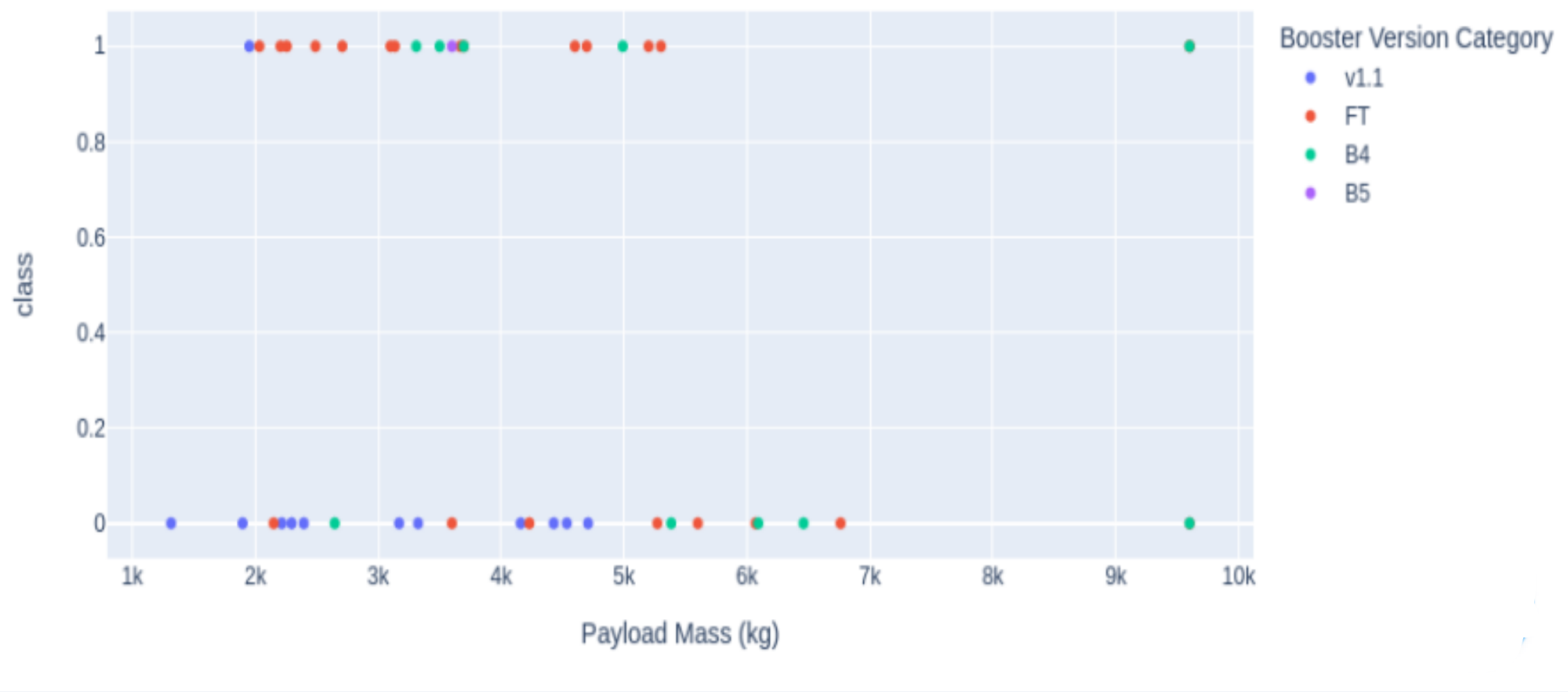
Launch Records



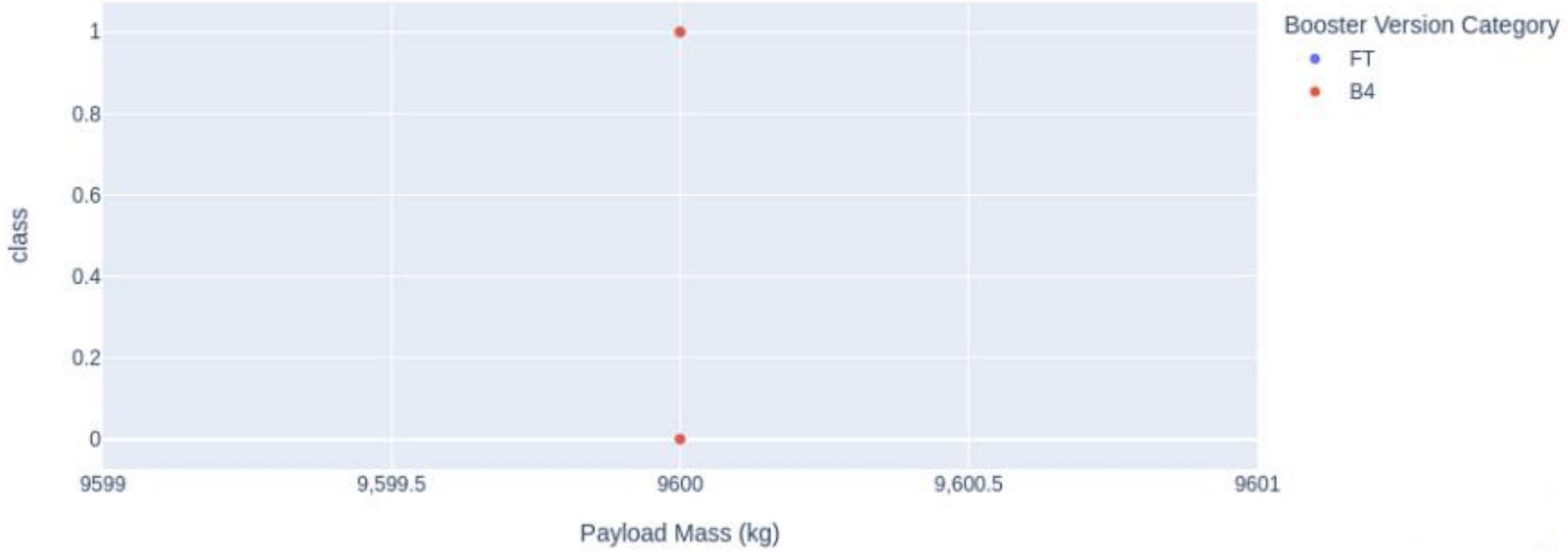
Succes Ratio of launch



Payload / Launch Outcome



Payload/Launch outcome

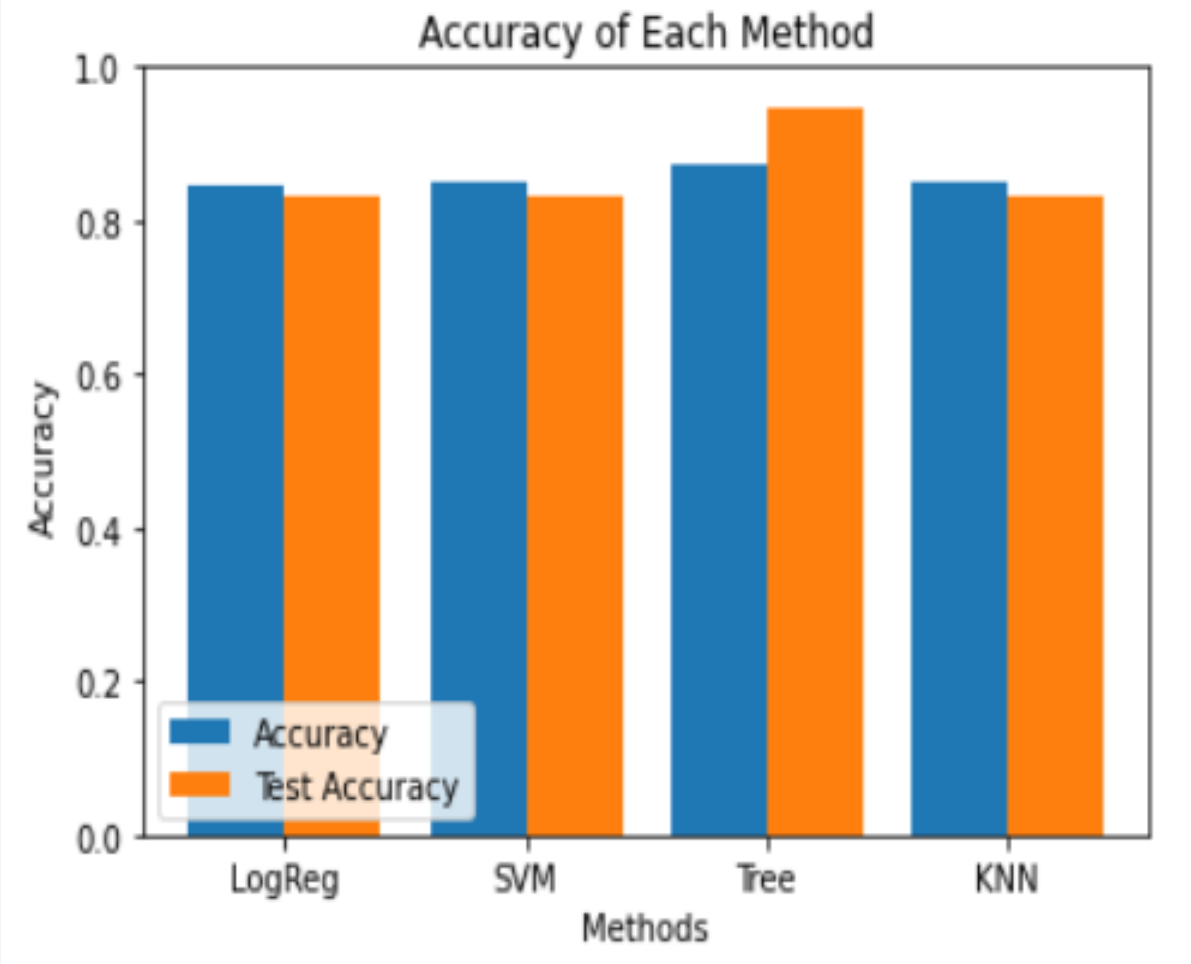


Section 5

Predictive Analysis (Classification)

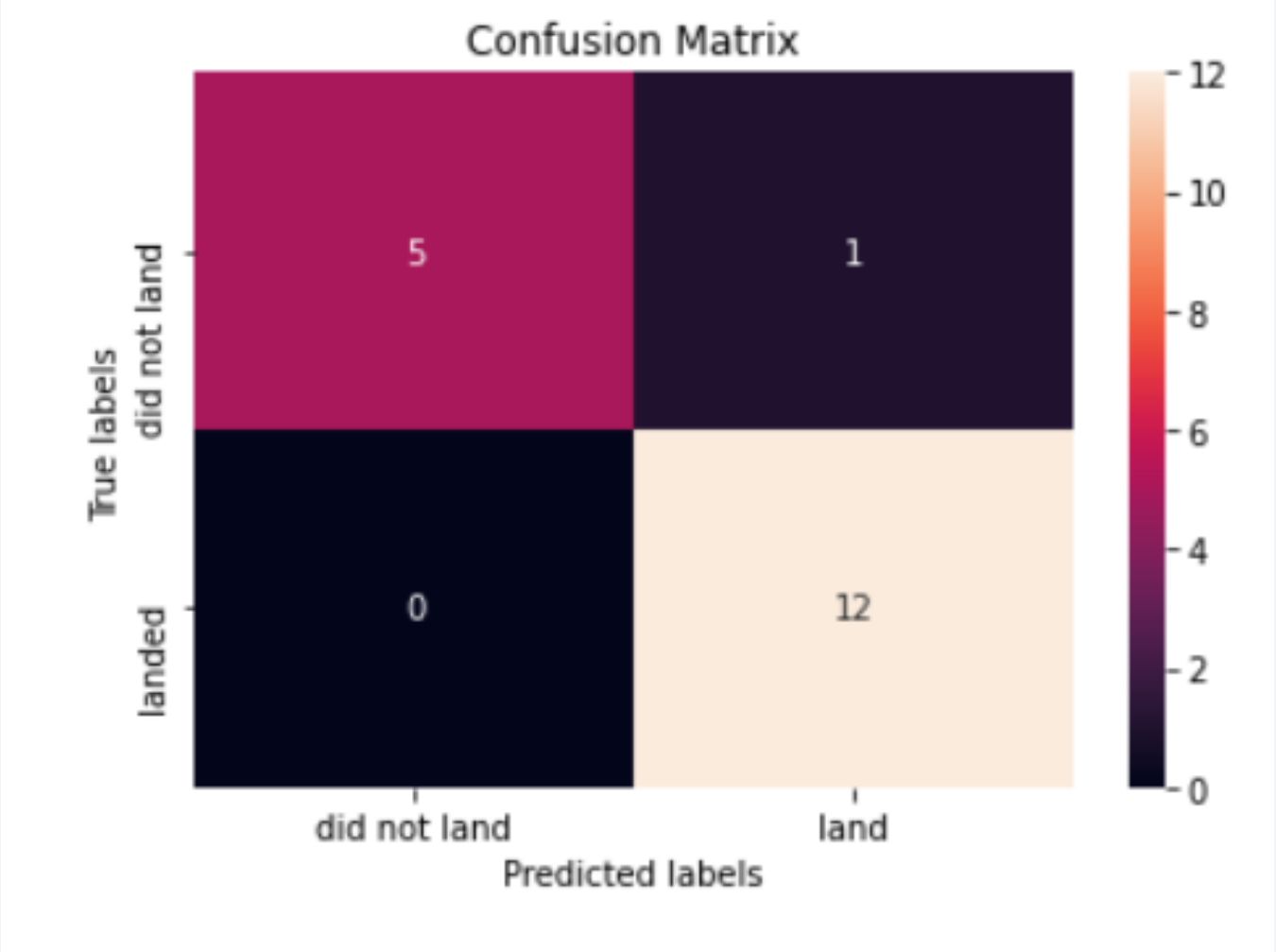
Classification Accuracy

- Tree is the most successful algorithm to this project with over the %88 accuracy



Confusion Matrix

- In the 18 data Tree model is the predict the 17 true 1 false prediction



Conclusions

- Launch rate increased 2013 to 2020

- ES-L1 orbit is have the most successful rate
- KSC LC 39A is the most successful launch site

Thank you!