



SpaceX Falcon 9 First Stage Landing Prediction

Igor Yurov

11-03-2022

OUTLINE



- Executive Summary
- Introduction
- Data Collection and Data Wrangling Methodology
- EDA and Interactive visual Analytics Methodology
- Predictive Analysis Methodology
- EDA with Visualization Results
- EDA with SQL Results
- Interactive Map with Folium Results
- Plotly Dash Dashboard Results
- Predictive Analysis (Classification) Results
- Conclusion

EXECUTIVE SUMMARY



- The purpose of this project is to predict if the Falcon 9 first stage will land successfully
- Predictive analysis can be applied to predict if the Falcon 9 first stage will land successfully
- Best machine learning prediction model is **Decision tree classifier** with accuracy of 0.8892
- Success rate is significantly improved since 2017
- Success rate is higher for heavy payloads

INTRODUCTION



- SpaceX advertises Falcon 9 rocket launches with a cost of 62 million dollars. Other providers cost upward of 165 million dollars each. Much of the savings is because SpaceX can reuse the first stage.
- The purpose of this project is to predict if the Falcon 9 first stage will land successfully and therefore determine the cost of a launch.
- This information can be used if an alternate company wants to bid against SpaceX for a rocket launch.

Data Collection and Data Wrangling Methodology



- Request and parse the SpaceX launch data using the GET request
- Clean the requested data
 - Filter the dataframe to only include Falcon 9 launches
 - Deal with Missing Values
 - Create a landing outcome label from Outcome column

EDA and Interactive Visual Analytics Methodology



- Exploratory data analysis (EDA) is used to analyze and investigate data sets and summarize their main characteristics, often employing data visualization methods.
- It helps determine how best to manipulate data sources to get the answers needed, making it easier to discover patterns, spot anomalies, test a hypothesis, or check assumptions.

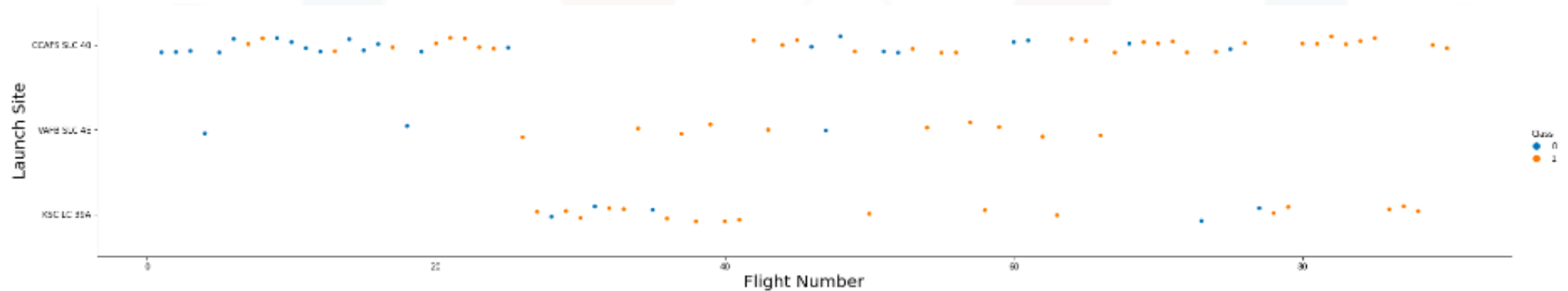
Predictive Analysis Methodology



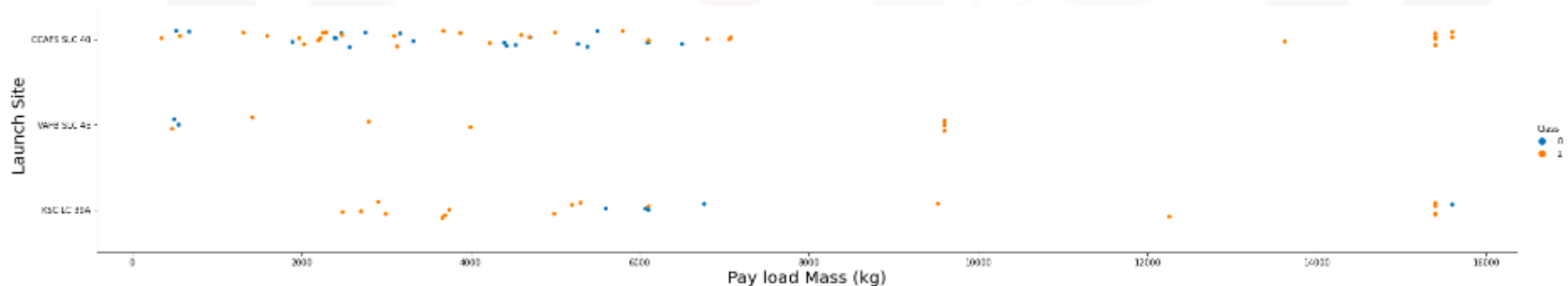
- Predictive analysis is a branch of advanced analytics that makes predictions about future outcomes using historical data combined with statistical modeling, data mining techniques and machine learning.
- Companies employ predictive analytics to find patterns in this data to identify risks and opportunities.

EDA with Visualization Results - 1

- Visualize the relationship between Flight Number and Launch Site

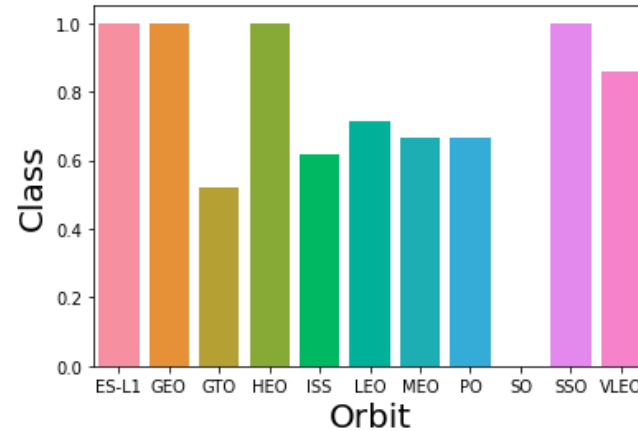


- Visualize the relationship between Payload and Launch Site

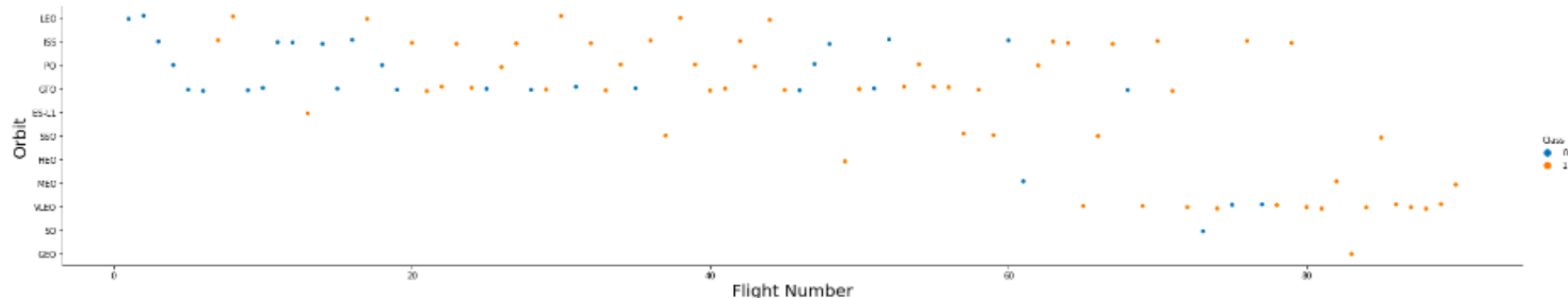


EDA with Visualization Results – 2

- Visualize the relationship between success rate of each orbit type

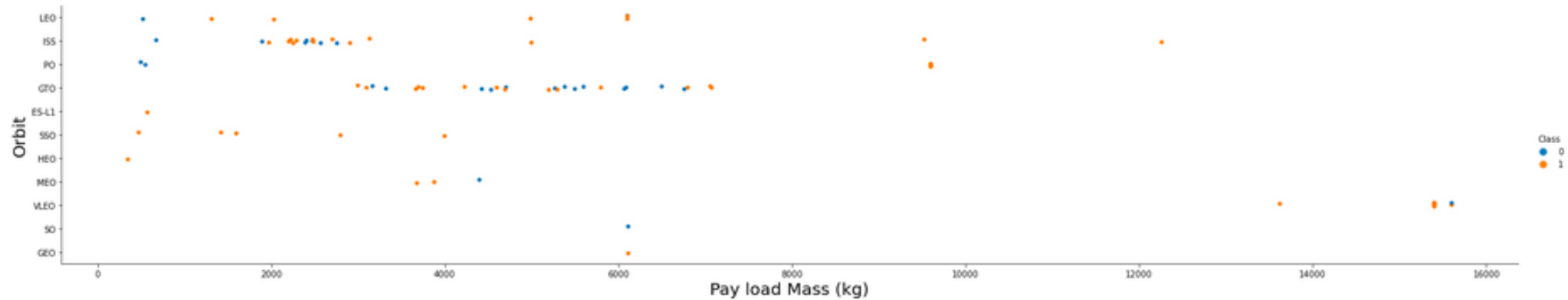


- Visualize the relationship between FlightNumber and Orbit type

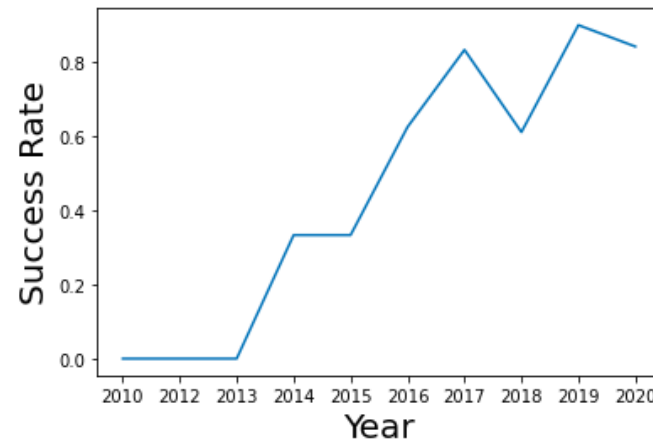


EDA with Visualization Results - 3

- Visualize the relationship between Payload and Orbit type



- Visualize the relationship between FlightNumber and Orbit type



EDA with SQL Results – 1

- Display the names of the unique launch sites in the space mission

launch_site

CCAFS LC-40

CCAFS SLC-40

KSC LC-39A

VAFB SLC-4E

- Display 5 records where launch sites begin with the string '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

EDA with SQL Results – 2

- Display the total payload mass carried by boosters launched by NASA (CRS)

Total Payload Mass

45596

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

Average Payload

2928

EDA with SQL Results – 3

- List the date when the first successful landing outcome in ground pad was achieved.

1

2015-12-22

- List the names of the boosters which have success in drone ship and have 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

EDA with SQL Results – 4

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

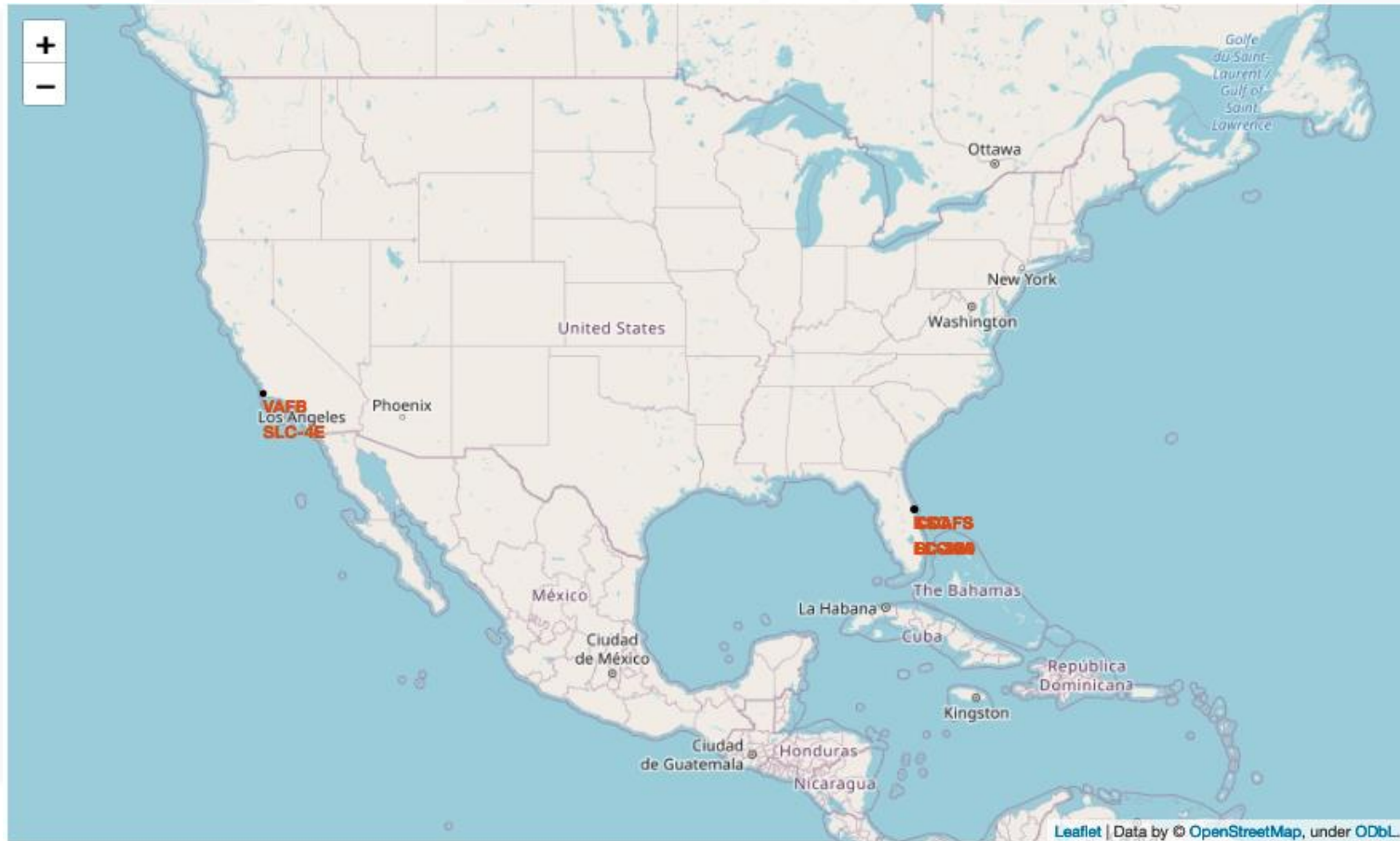
landing__outcome	booster_version	launch_site
Failure (drone ship)	F9 v1.1 B1012	CCAFS LC-40
Failure (drone ship)	F9 v1.1 B1015	CCAFS LC-40

- 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	count
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

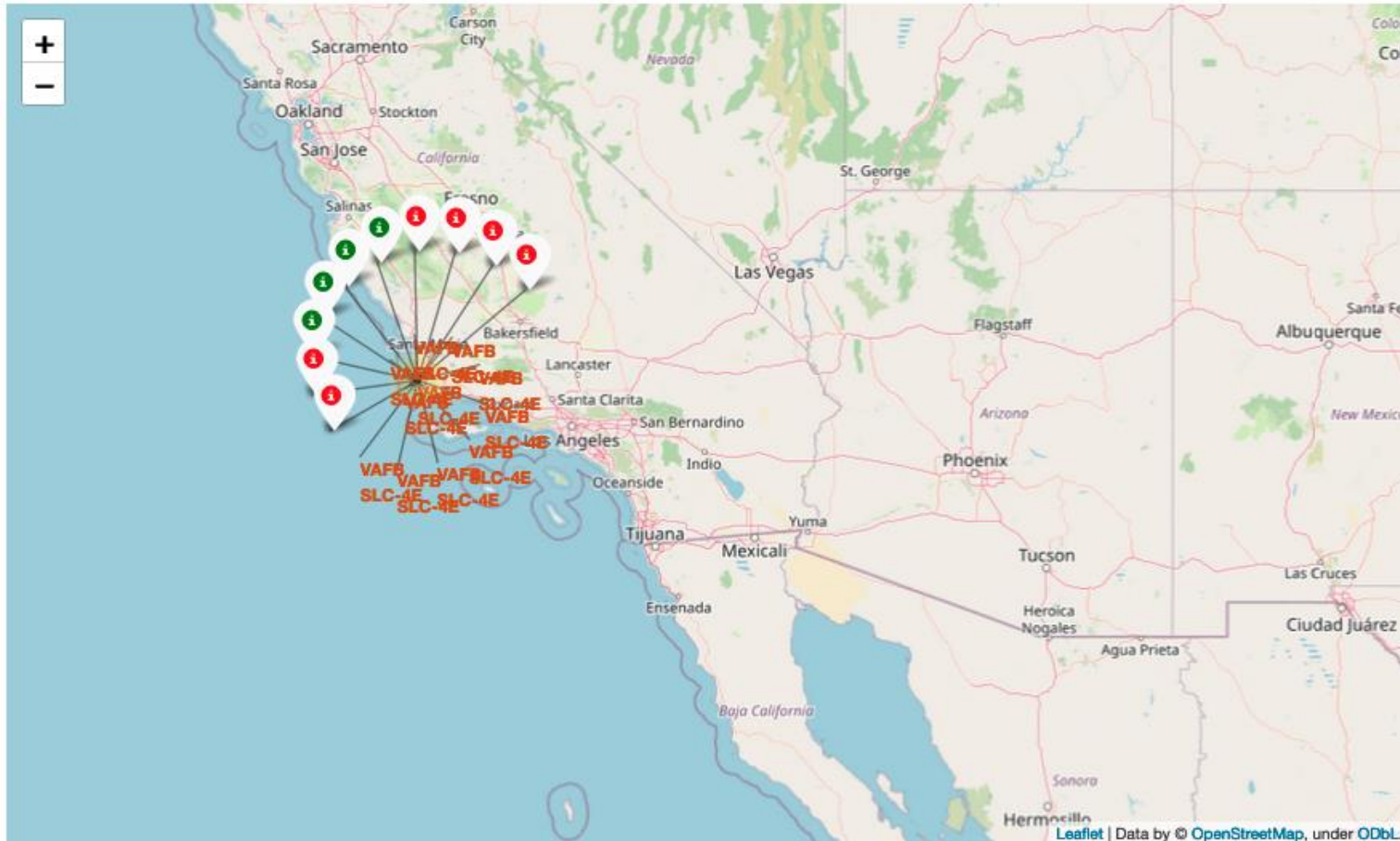
Interactive Map with Folium Results - 1

- Mark all launch sites on a map



Interactive Map with Folium Results - 2

- Mark the success/failed launches for each site on the map



Interactive Map with Folium Results - 3

- Calculate the distances between a launch site to its proximities



Plotly Dash Dashboard Results - 1

- Dropdown list to enable Launch Site selection

SpaceX Launch Records Dashboard

All Sites

All Sites

CCAFS LC-40

VAFB SLC-4E

KSC LC-39A

CCAFS SLC-40

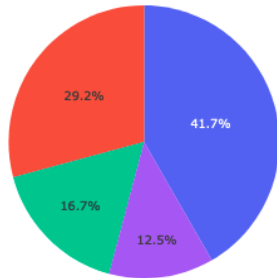
Plotly Dash Dashboard Results - 2

- Pie chart to show the total successful launches count for all sites
- If a specific launch site was selected, show the Success vs. Failed counts for the site

SpaceX Launch Records Dashboard

All Sites

Total Success Launches By Site

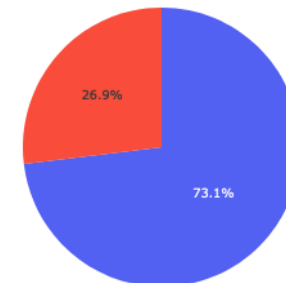


■ KSC LC-39A
■ CCAFS LC-40
■ VAFB SLC-4E
■ CCAFS SLC-40

SpaceX Launch Records Dashboard

CCAFS LC-40

Total Success Launches for site CCAFS LC-40



■ 0
■ 1

Plotly Dash Dashboard Results - 3

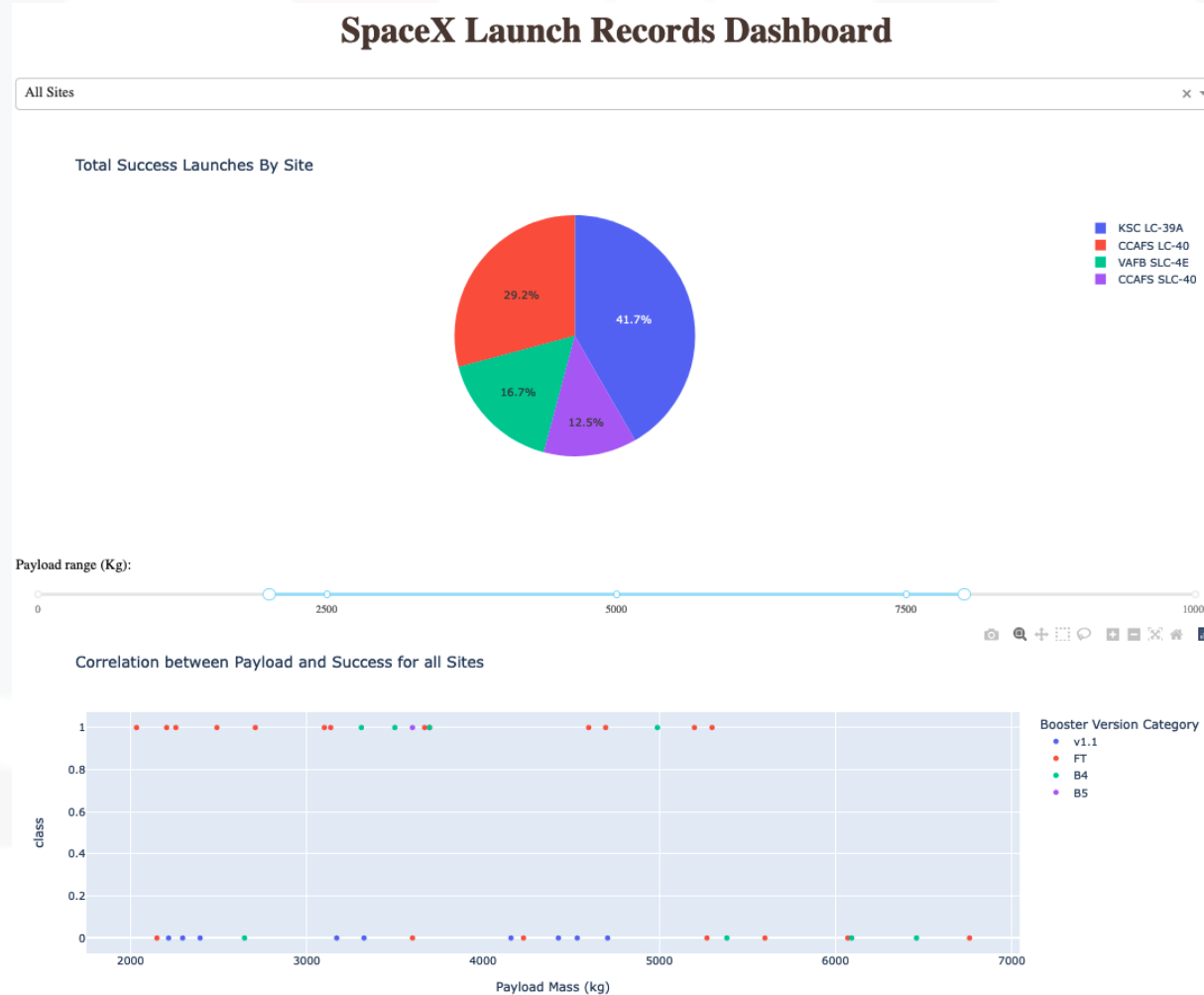
- Slider to select payload range

Payload range (Kg):



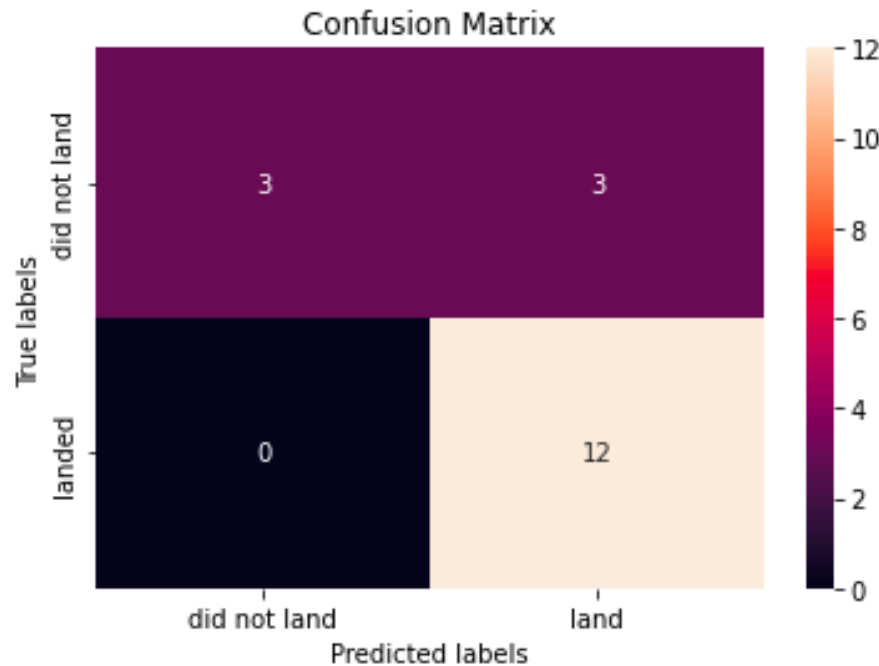
Plotly Dash Dashboard Results - 4

- Scatter chart to show the correlation between payload and launch success



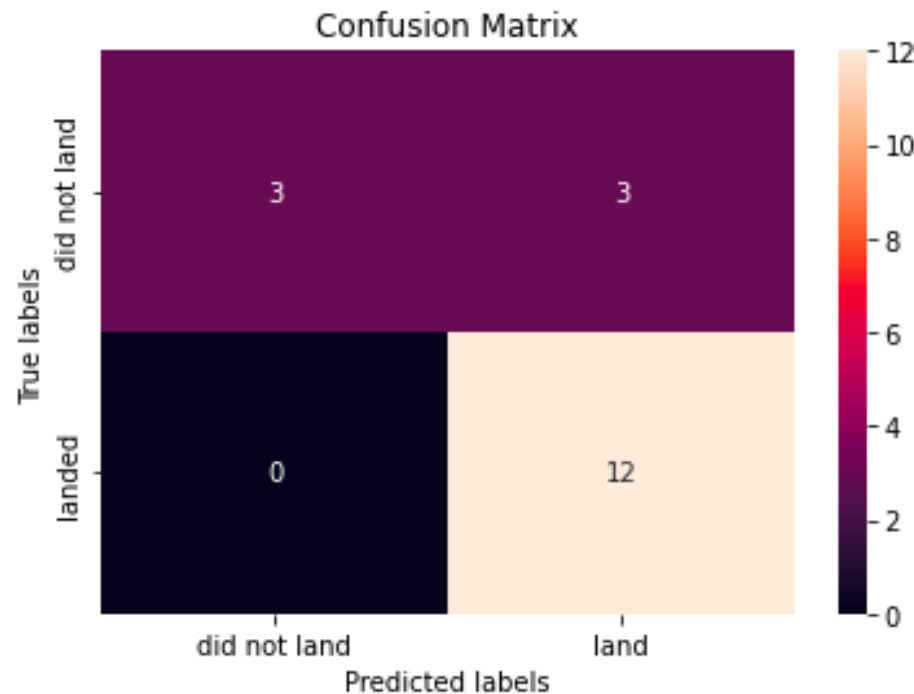
Predictive Analysis (Classification) Results - 1

- Logistic regression model:
 - tuned hyperparameters :(best parameters) {'C': 0.01, 'penalty': 'l2', 'solver': 'lbfgs'}
 - accuracy : 0.8464285714285713



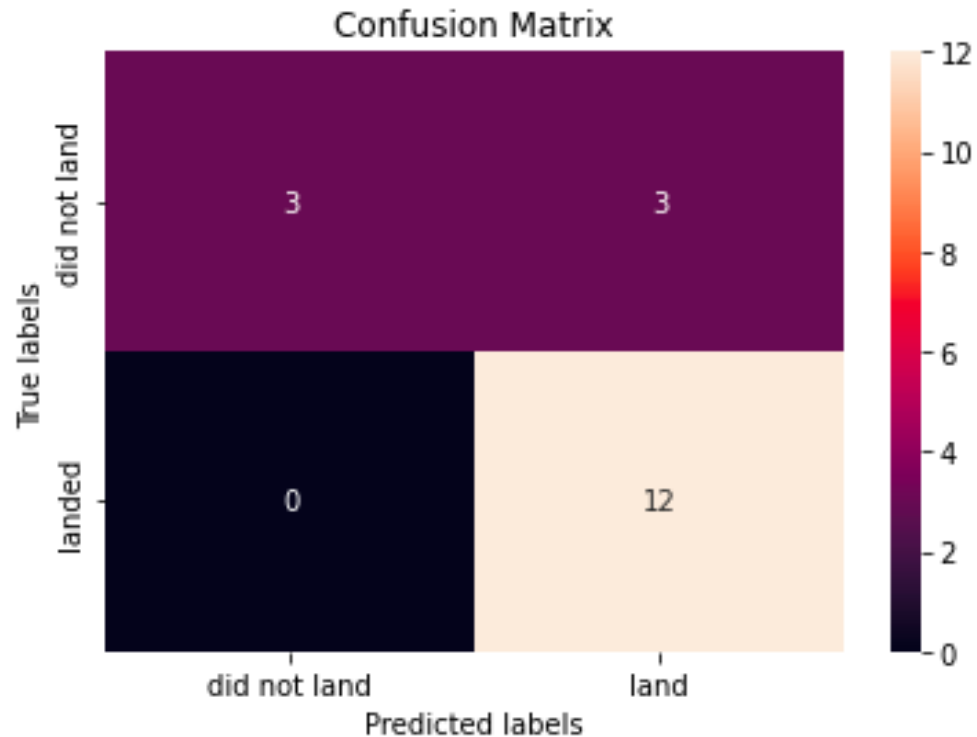
Predictive Analysis (Classification) Results - 2

- Support vector machine model:
 - tuned hyperparameters :(best parameters) {'C': 1.0, 'gamma': 0.03162277660168379, 'kernel': 'sigmoid'}
 - accuracy : 0.8482142857142856



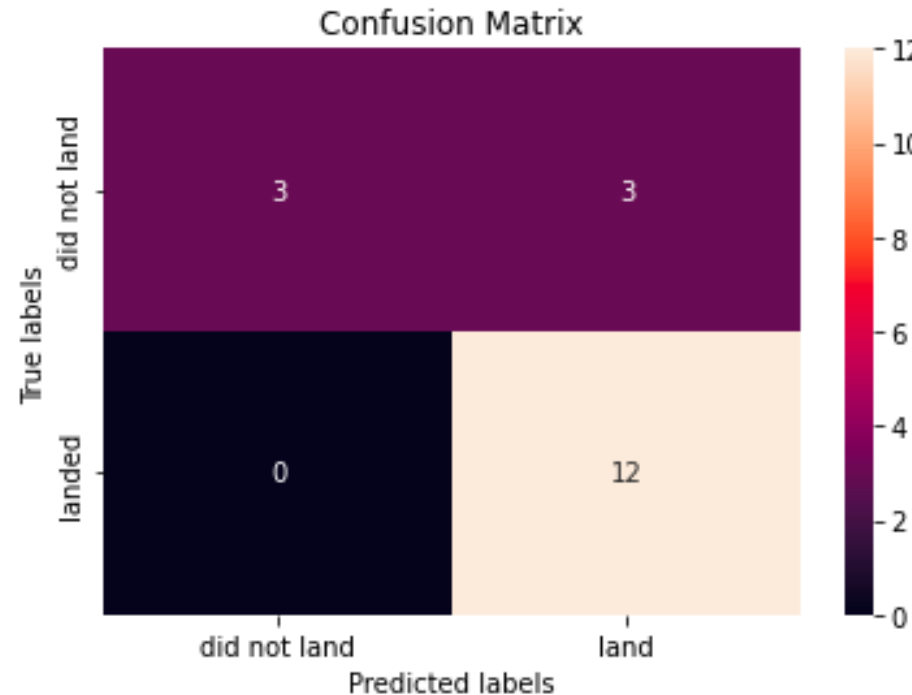
Predictive Analysis (Classification) Results - 3

- Decision tree classifier model:
 - tuned hyperparameters :(best parameters) {'criterion': 'gini', 'max_depth': 14, 'max_features': 'auto', 'min_samples_leaf': 1, 'min_samples_split': 10, 'splitter': 'random'}
 - accuracy : 0.8892857142857142



Predictive Analysis (Classification) Results - 4

- K nearest neighbors model:
 - tuned hyperparameters :(best parameters) {'algorithm': 'auto', 'n_neighbors': 10, 'p': 1}
 - accuracy : 0.8482142857142858



CONCLUSION



- Predictive analysis can be applied to predict if the Falcon 9 first stage will land successfully
- Best machine learning prediction model is **Decision tree classifier** with accuracy of 0.8892
- Success rate is significantly improved since 2017
- Success rate is higher for heavy payloads