Space X Launch Success Rate

Francis Odhiambo

21st November 2023

Outline

- Executive Summary
- Introduction
- Methodology
- Results
  - Visualization – Charts
  - Dashboard
- Discussion
  - Findings & Implications
- Conclusion
- Appendix

Executive Summary

Methodologies used in this project

- Data Collection with API and Web Scrapping
- Data Wrangling
- Exploratory Data Analysis with SQL, Pandas and Matplotlib
- Data Analysis for Data Visualization
- Interactive Visual Analytics with Folium
- Interactive dashboard with Ploty Dash
- Predictive Analysis

Project results

- Exploratory Data Analysis result
- Analysis graphs and diagrams
- Final prediction

Introduction

- Space travel has become a desire for everyone. To realize this, companies offering space travel services such as Virgin Galactic, Rocket Lab, Blue Origin and Space X have moved to make it more affordable to their users.

- With several successful rocket launches and ability to reuse some of the rocket components, Space X has seen the most successful and inexpensive rocket launches with a cost of 62million company against an upwards of 165million for the competitions. Space X has achieved this by managing to reuse the first stage, thereby cutting on overall launch cost.

- This analysis project is aimed at determining the cost of a Space X rocket launch by determining if the first stage will land successfully. This stage does most of the work and is much larger than the second stage. Here we see the first stage next to a person and several other landmarks. This stage is quite large and expensive.

Objective

The objectives of this task include:

- Gathering information about Space X and creating dashboards
- Predicting if SpaceX will reuse the first stage by will training a machine learning model and using public information
- Present the finding on a dashboard

Methodology – Data Collection

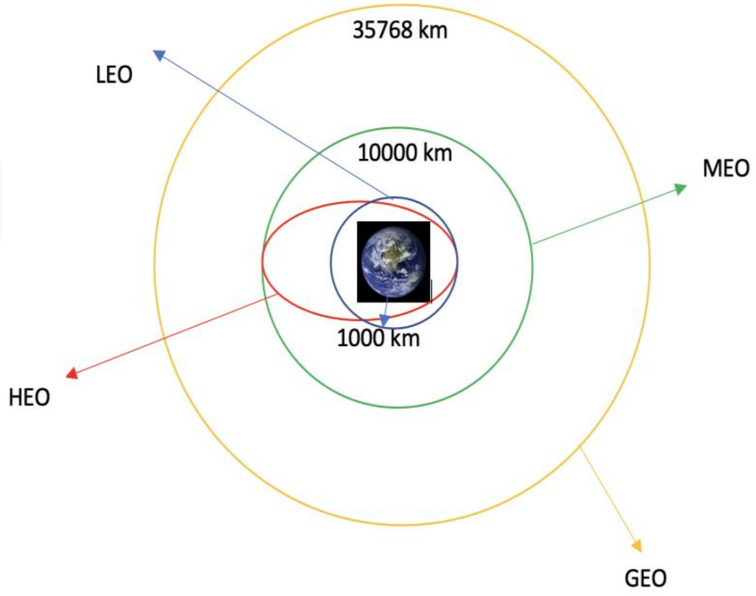
- SpaceX data was collected using API and Webscrapping methods from publicly available sources
- Static response object was used make the requested JSON results more consistent and to return a dataframe
- The data frame dataframe was filtered to only include Falcon 9 launches
- The dataframe was rectified to fill in missing values
- Web scraping was also done from Wikipedia for Falcon 9 launch records with BeautifulSoup.
  - The objective was to extract the launch records as HTML table, parse the table and convert it to a pandas dataframe for future analysis.

Methodology – Data Wrangling

Exploratory data analysis was conducted to determine:

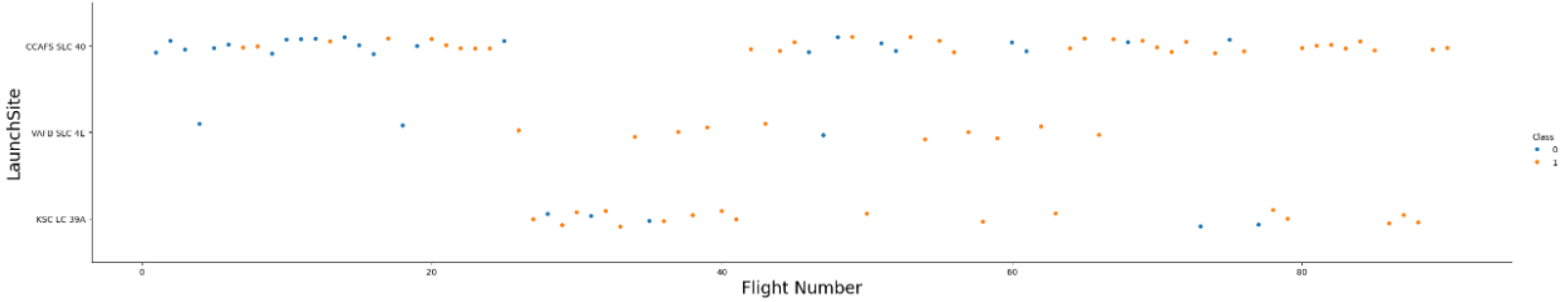
- Number of launches on each site
- The number and occurrence of each orbit
- The number and occurrence of mission outcome of the orbits

From the outcome, machine learning training labels were determined



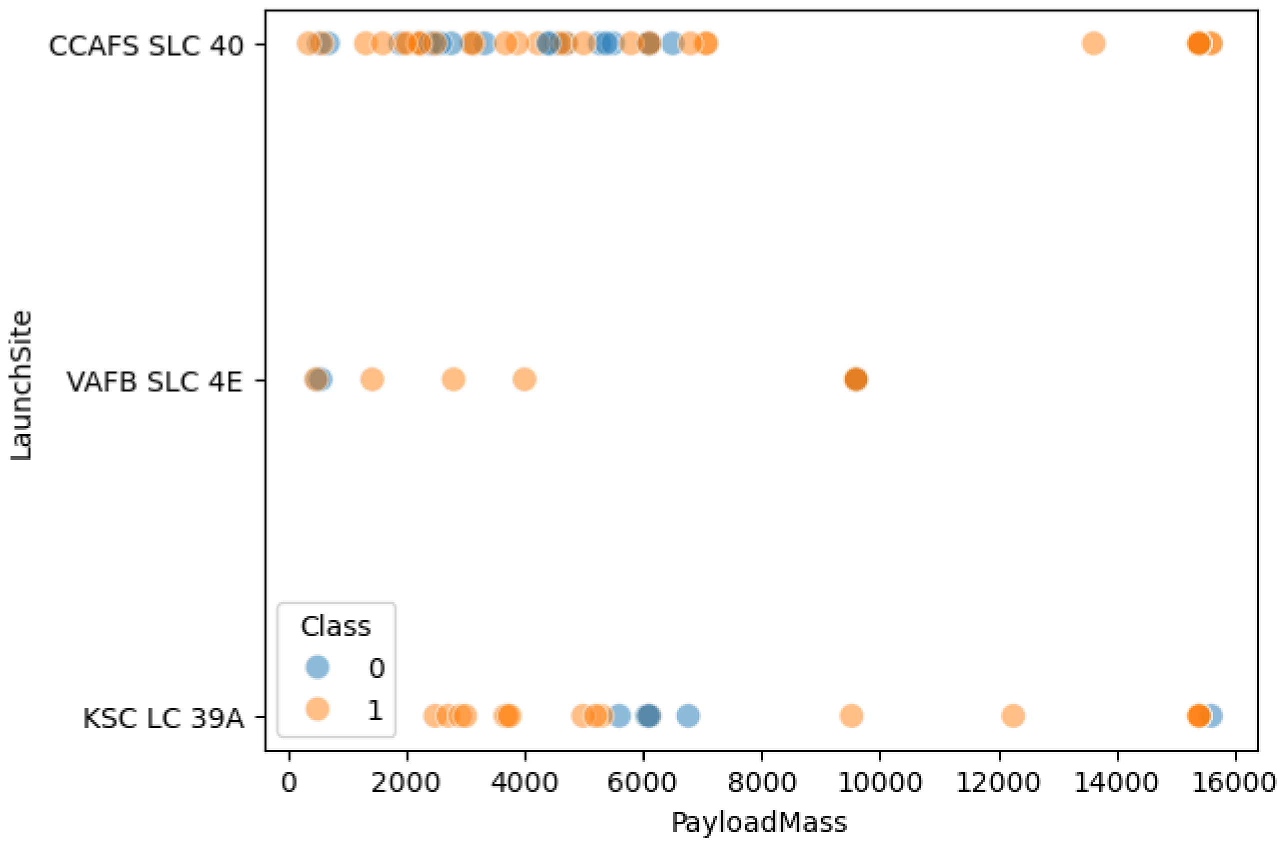
Results-EDA with Visualization

The relationship between Flight Number and Launch Site



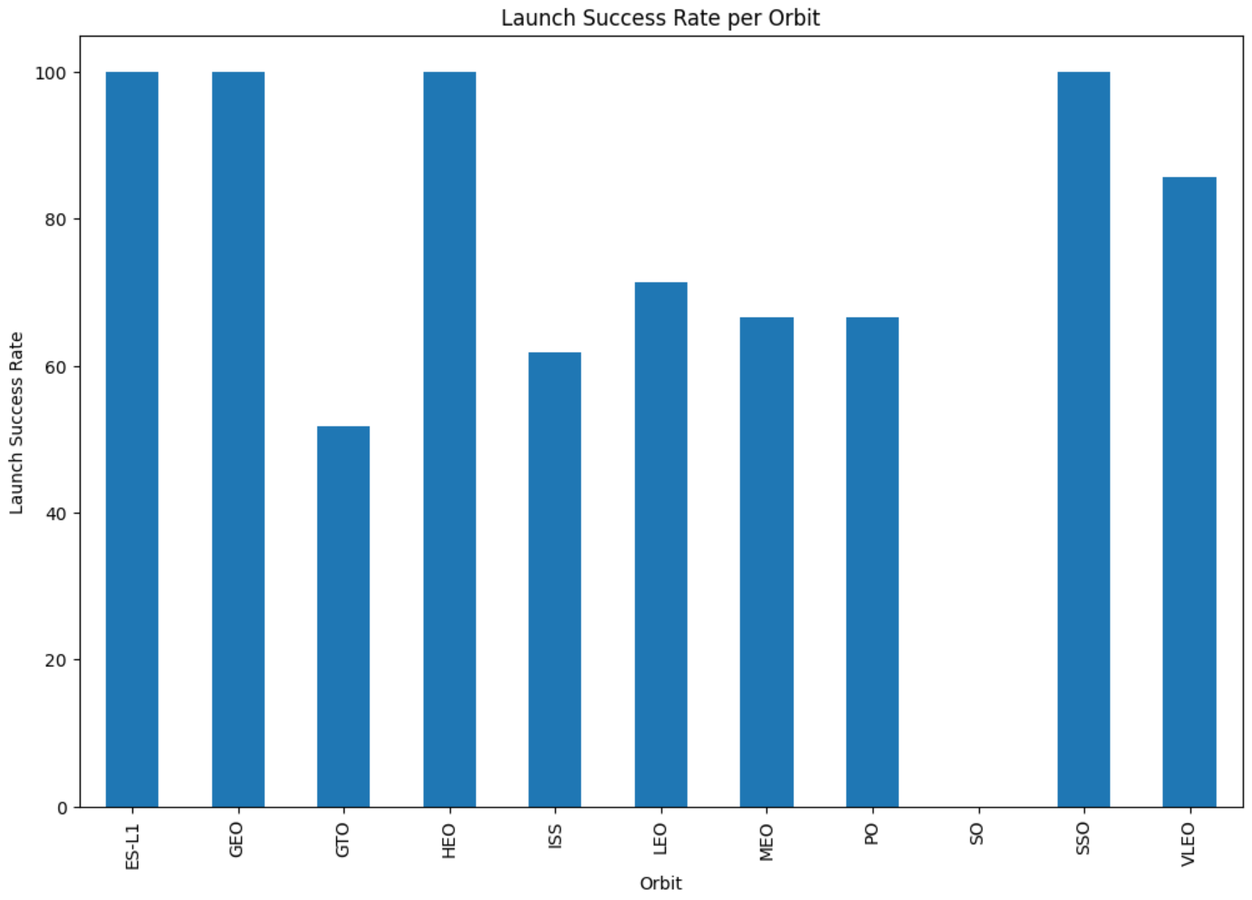
Results-EDA with Visualization

The relationship between Payload and Launch Site



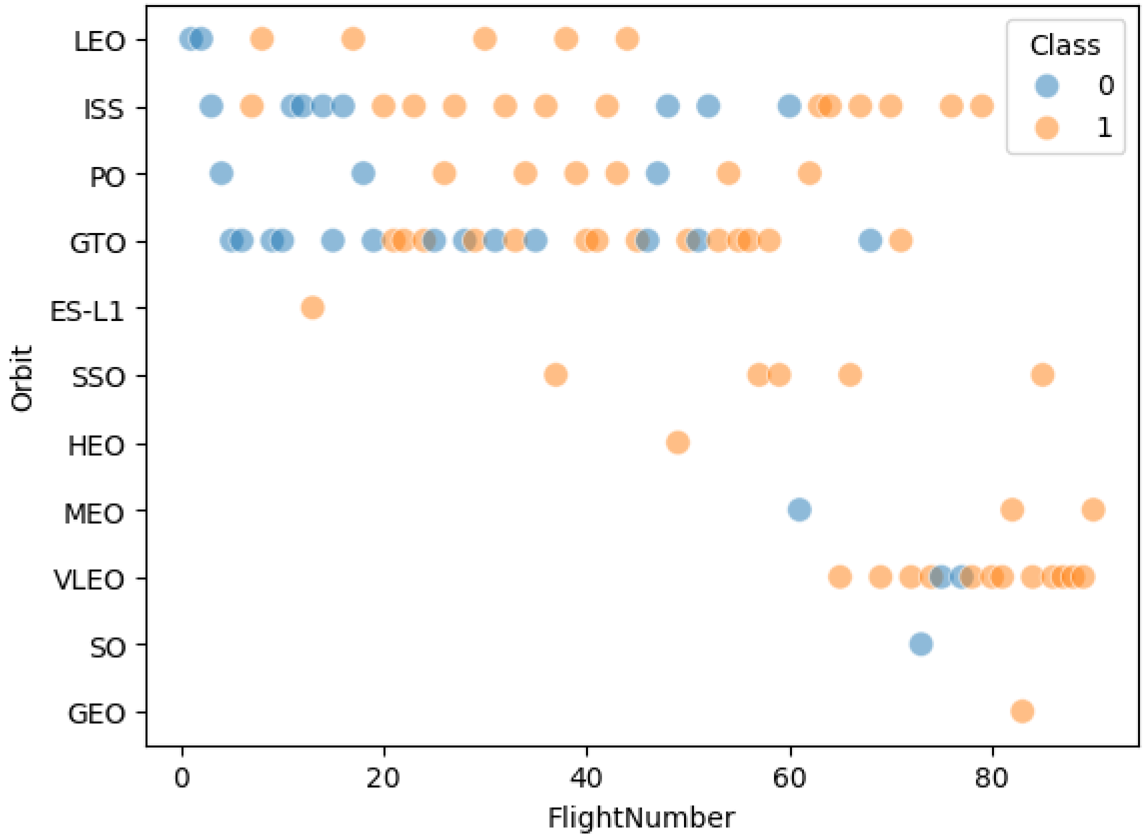
Results-EDA with Visualization

The relationship between success rate of each orbit type



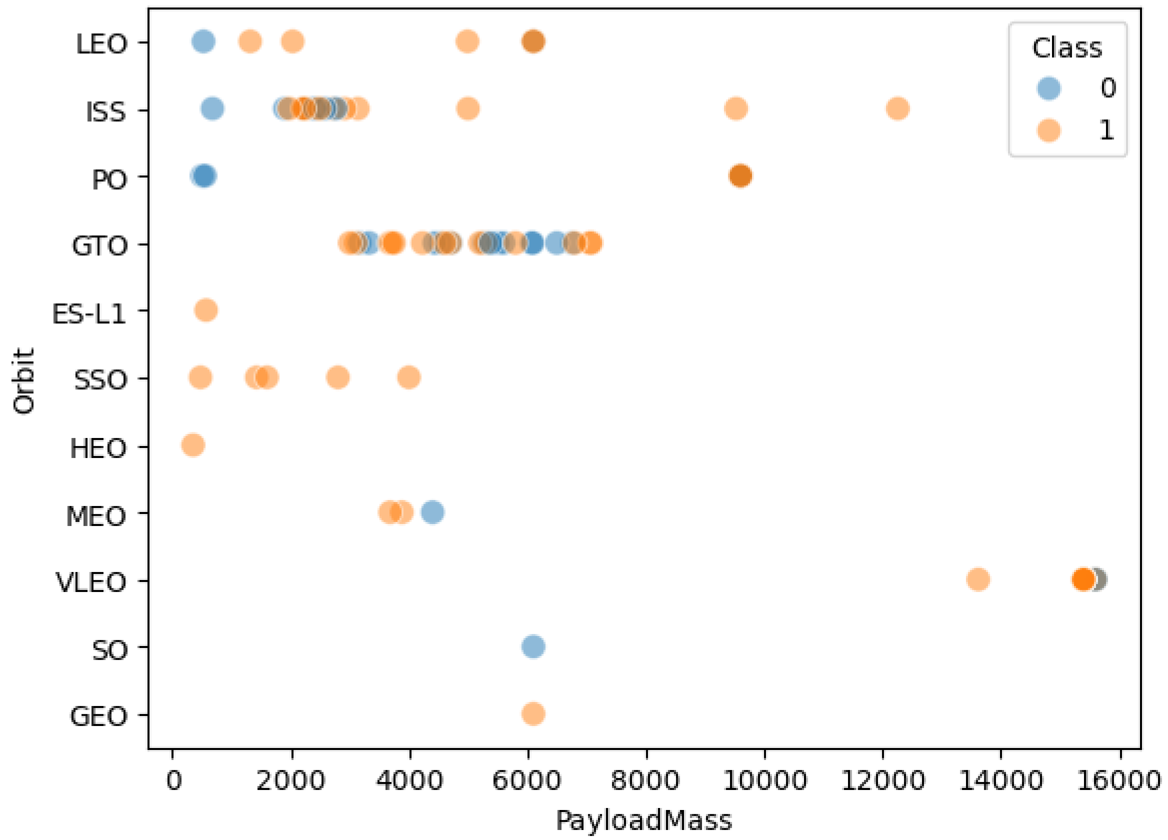
Results-EDA with Visualization

The relationship between FlightNumber and Orbit type



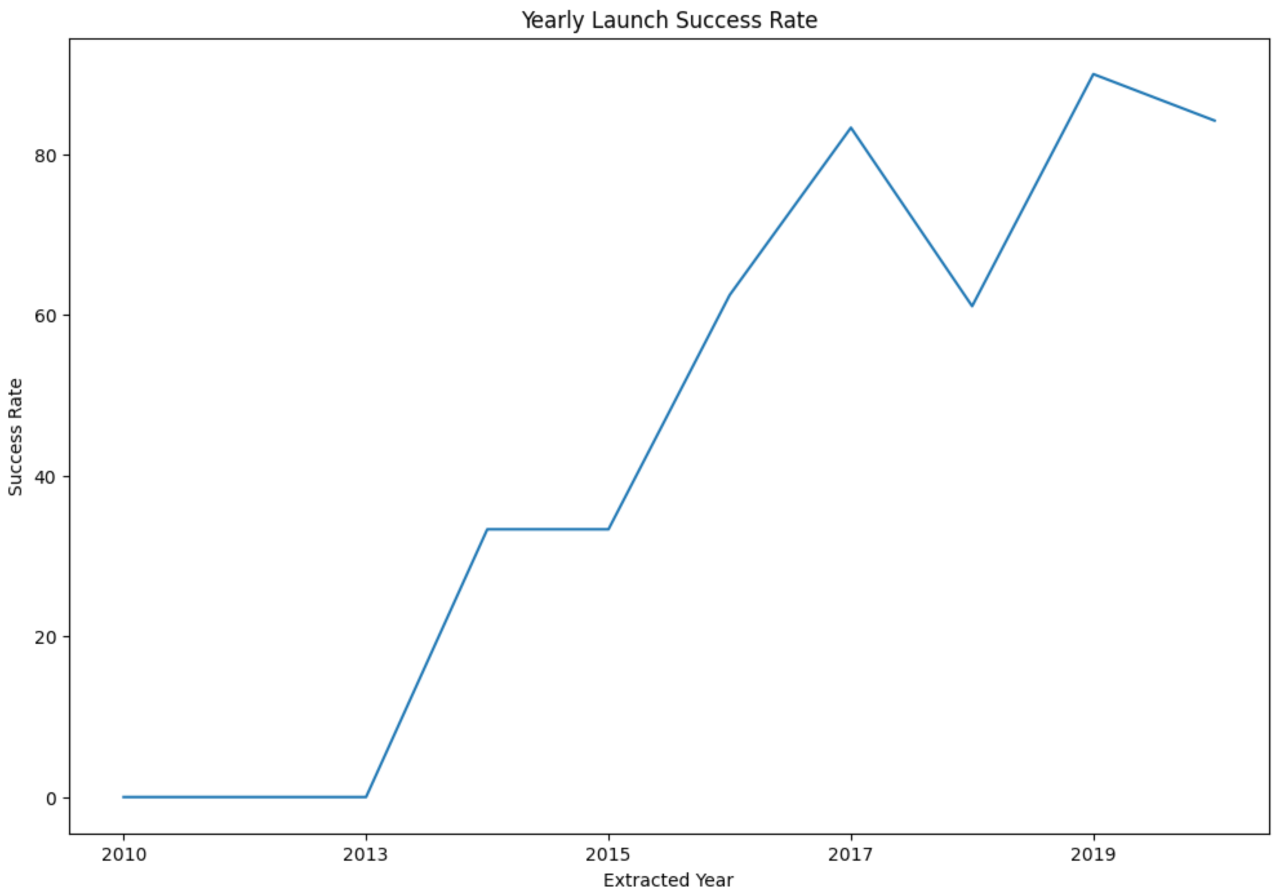
Results-EDA with Visualization

The relationship between Payload and Orbit type



Results-EDA with Visualization

The launch success yearly trend



Results-EDA with SQL

The names of the unique launch sites in the space mission

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

Results-EDA with SQL

The 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	7:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525	LEO (ISS)	NASA (COTS)	Success	No attempt
2012-10-08	0:35:00	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

Results-EDA with SQL

The total payload mass carried by boosters launched by NASA (CRS)

SUM(PAYLOAD_MASS__KG_)
45596

Results-EDA with SQL

The average payload mass carried by booster version F9 v1.1

AVG(PAYLOAD_MASS__KG_)
2928.4

Results-EDA with SQL

The relationship between Flight Number and Launch Site

1
2015-12-22

Results-EDA with SQL

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

Results-EDA with SQL

The total number of successful and failure mission outcomes

Mission_Outcome	total_number
Failure (in flight)	1
Success	98
Success	1
Success (payload status unclear)	1

Results-EDA with SQL

The names of the booster versions which have carried the maximum payload mass. Use a subquery

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

Results-EDA with SQL

The records which will display the month names, failure landing_outcomes in drone ship ,booster versions, launch_site for the months in year 2015.

DATE	time__utc_	booster_version	launch_site	payload	payload_mass__kg_	orbit	customer	mission_outcome	landing__out
2017-02-19	14:39:00	F9 FT B1031.1	KSC LC-39A	SpaceX CRS-10	2490	LEO (ISS)	NASA (CRS)	Success	Success (g
2017-01-14	17:54:00	F9 FT B1029.1	VAFB SLC-4E	Iridium NEXT 1	9600	Polar LEO	Iridium Communications	Success	Success (
2016-08-14	05:26:00	F9 FT B1026	CCAFS LC-40	JCSAT-16	4600	GTO	SKY Perfect JSAT Group	Success	Success (
2016-07-18	04:45:00	F9 FT B1025.1	CCAFS LC-40	SpaceX CRS-9	2257	LEO (ISS)	NASA (CRS)	Success	Success (g
2016-05-27	21:39:00	F9 FT B1023.1	CCAFS LC-40	Thaicom 8	3100	GTO	Thaicom	Success	Success (
2016-05-06	05:21:00	F9 FT B1022	CCAFS LC-40	JCSAT-14	4696	GTO	SKY Perfect JSAT Group	Success	Success (
2016-04-08	20:43:00	F9 FT B1021.1	CCAFS LC-40	SpaceX CRS-8	3136	LEO (ISS)	NASA (CRS)	Success	Success (
2015-12-22	01:29:00	F9 FT B1019	CCAFS LC-40	OG2 Mission 2 11 Orbcomm-OG2 satellites	2034	LEO	Orbcomm	Success	Success (g

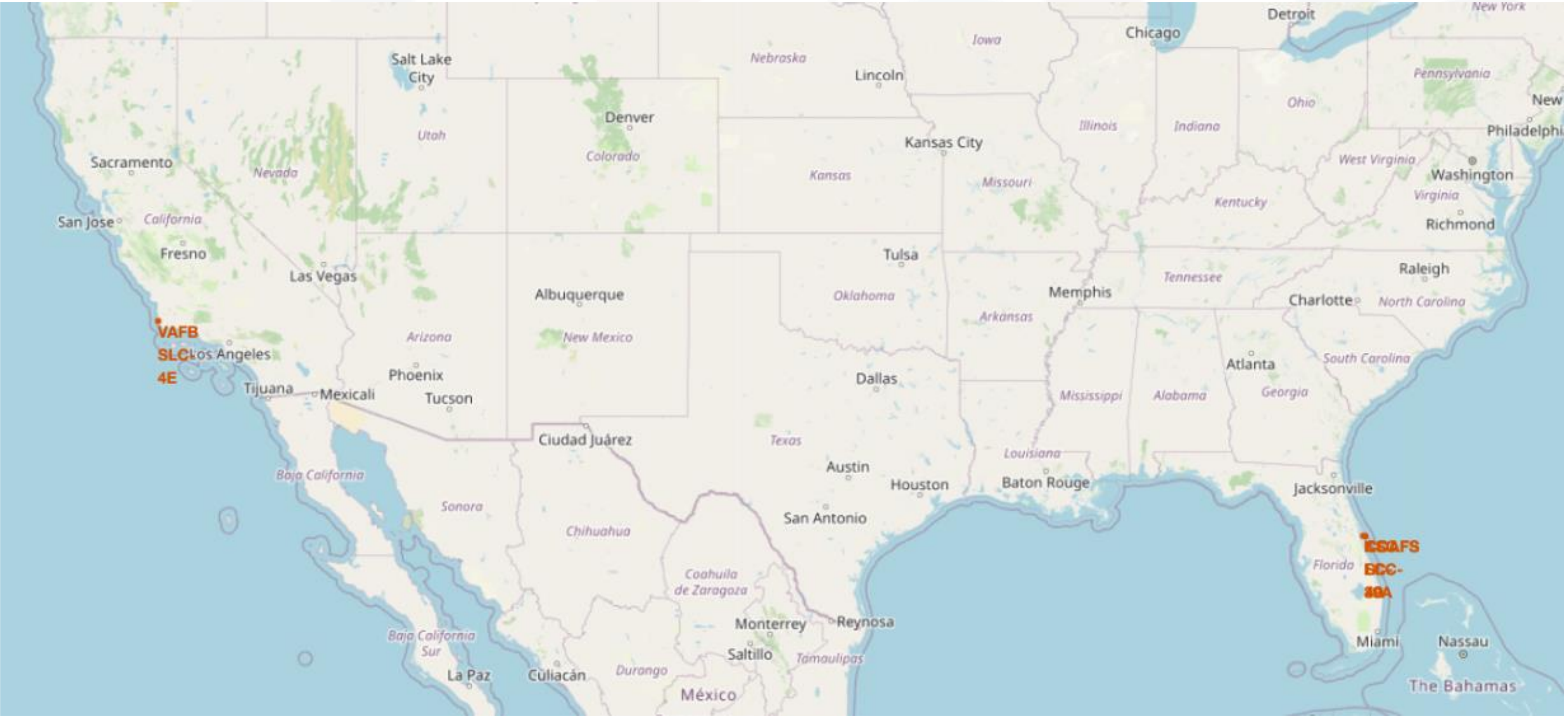
Results-EDA with SQL

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	landing__out
2017-02-19	14:39:00	F9 FT B1031.1	KSC LC-39A	SpaceX CRS-10	2490	LEO (ISS)	NASA (CRS)	Success	Success (g
2017-01-14	17:54:00	F9 FT B1029.1	VAFB SLC-4E	Iridium NEXT 1	9600	Polar LEO	Iridium Communications	Success	Success (
2016-08-14	05:26:00	F9 FT B1026	CCAFS LC-40	JCSAT-16	4600	GTO	SKY Perfect JSAT Group	Success	Success (
2016-07-18	04:45:00	F9 FT B1025.1	CCAFS LC-40	SpaceX CRS-9	2257	LEO (ISS)	NASA (CRS)	Success	Success (g
2016-05-27	21:39:00	F9 FT B1023.1	CCAFS LC-40	Thaicom 8	3100	GTO	Thaicom	Success	Success (
2016-05-06	05:21:00	F9 FT B1022	CCAFS LC-40	JCSAT-14	4696	GTO	SKY Perfect JSAT Group	Success	Success (
2016-04-08	20:43:00	F9 FT B1021.1	CCAFS LC-40	SpaceX CRS-8	3136	LEO (ISS)	NASA (CRS)	Success	Success (
2015-12-22	01:29:00	F9 FT B1019	CCAFS LC-40	OG2 Mission 2 11 Orbcomm-OG2 satellites	2034	LEO	Orbcomm	Success	Success (g

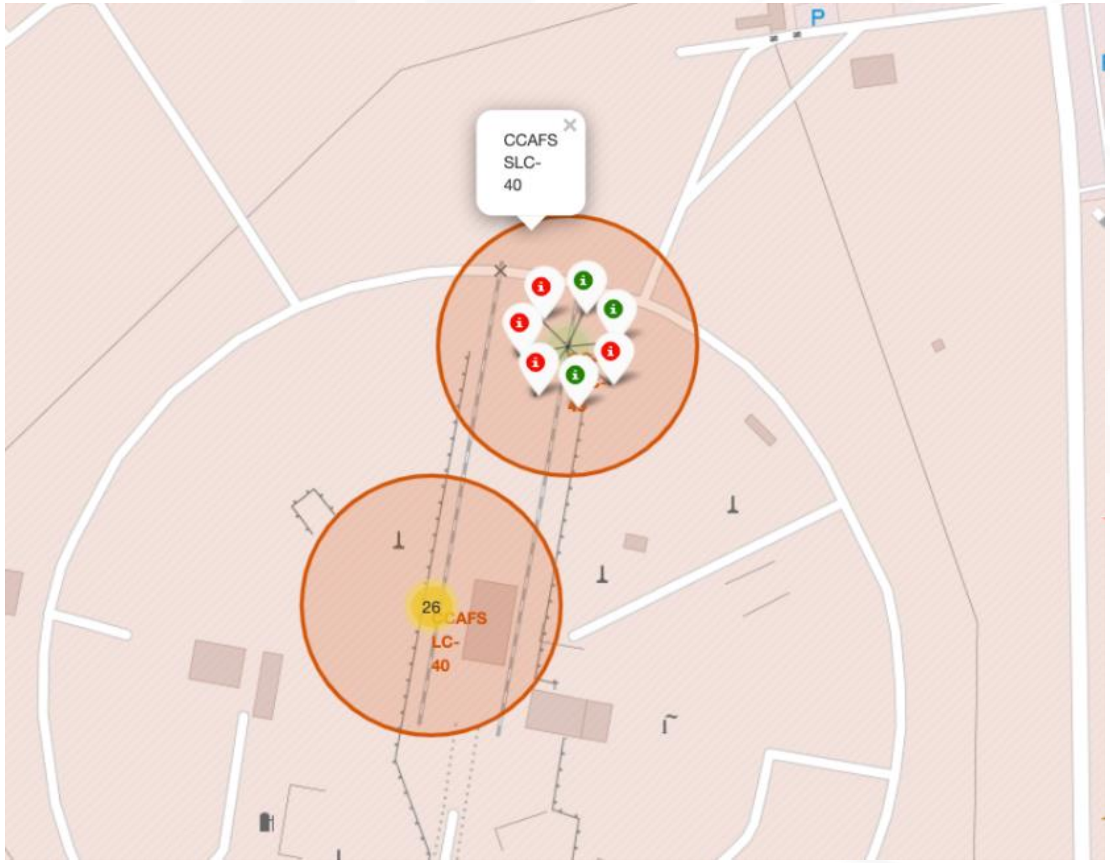
Results-Interactive Map with Folium

All Launch Sites



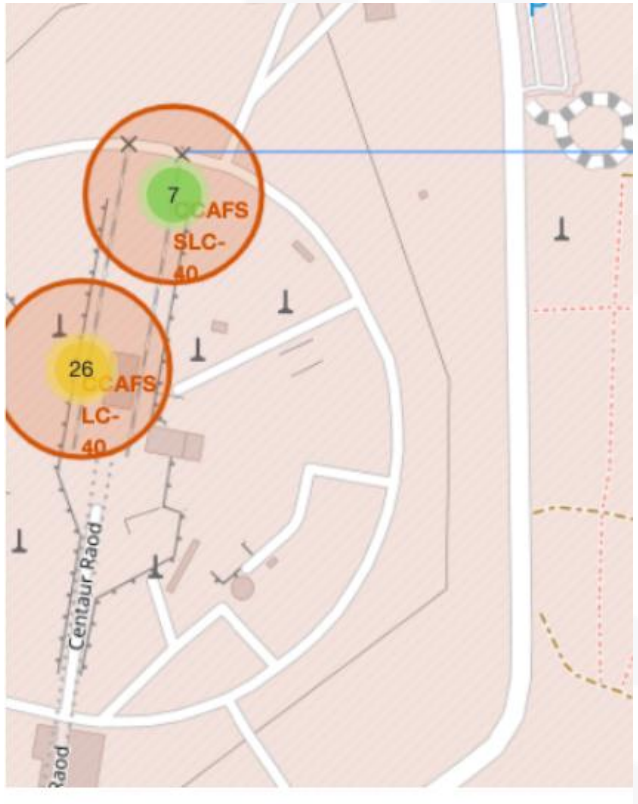
Results-Interactive Map with Folium

Success and Failed Launches



Results-Interactive Map with Folium

The distances between a launch site to its proximities



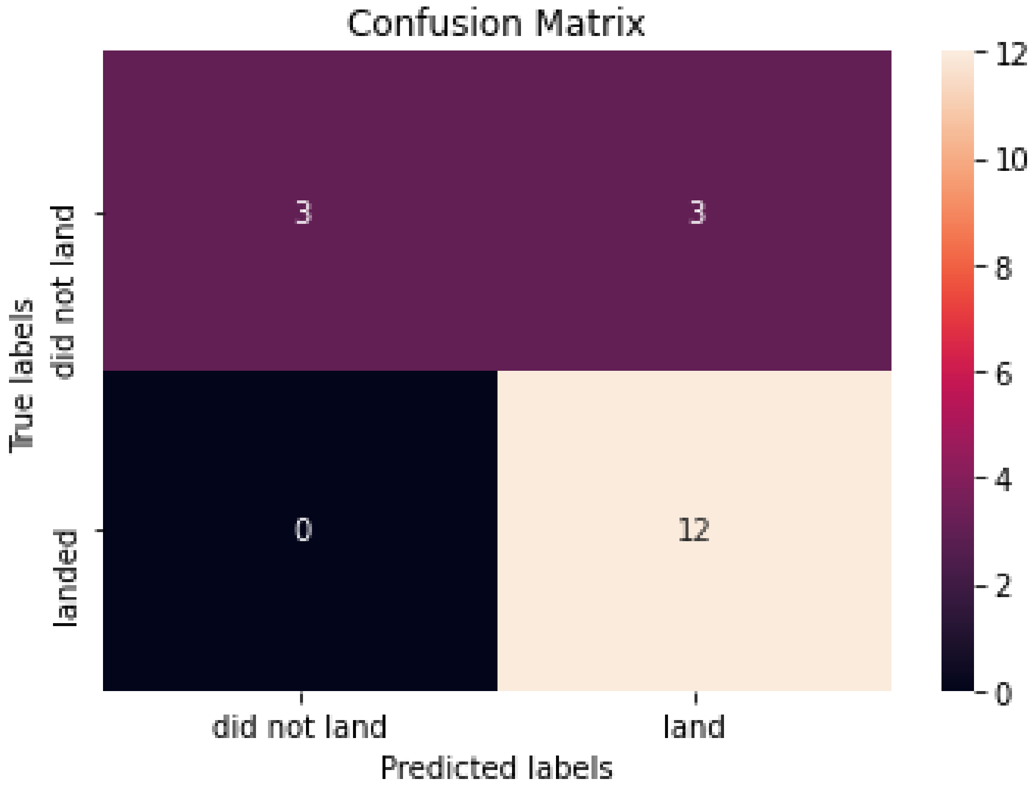
Predictive Analytics

All the models performed at about the same level and had the same scores and accuracy. This is likely due to the small dataset. The Decision Tree model slightly outperformed the rest when looking at .best_score_

	LogReg	SVM	Tree	KNN
Jaccard_Score	0.800000	0.800000	0.800000	0.800000
F1_Score	0.888889	0.888889	0.888889	0.888889
Accuracy	0.833333	0.833333	0.833333	0.833333

Confusion Matrix

The confusion matrix for the decision tree classifier



Conclusion

- Research
- Model Performance: The models performed similarly on the test set with the decision tree model slightly outperforming
- Equator: Most of the launch sites are near the equator for an additional natural boost - due to the rotational speed of earth - which helps save the cost of putting in extra fuel and boosters
- Coast: All the launch sites are close to the coast
- Launch Success: Increases over time
- KSC LC-39A: Has the highest success rate among launch sites. Has a 100% success rate for launches less than 5,500 kg
- Orbits: ES-L1, GEO, HEO, and SSO have a 100% success rate
- Payload Mass: Across all launch sites, the higher the payload mass (kg), the higher the success rate