SpaceX First Stage Reuse

Andre Eustaquio

July 31, 2024

IBM Developer

SKILLS NETWORK

Presentation Contents

- Executive Summary
- Introduction
- Methodology
- Results
  - EDA with Visualization
  - EDA with SQL
  - Interactive Maps with Folium
  - Dashboard with Plotly
  - Predictive Analysis
- Conclusion

EXECUTIVE SUMMARY

- The research attempts to identify the factors of a successful rocket landing.
- Determining first stage landing success:
  - Creating a success/fail outcome variable.
  - Considering factors such as payload, launch site, flight number, yearly trend.
  - Calculating success rates for each factor.
- Launch success has improved over time, KSC-LC-39A with the highest success rate.
- Most launch sites are significantly close to the equator and coastline.
- Decision tree model performed best for predictive analysis.

INTRODUCTION

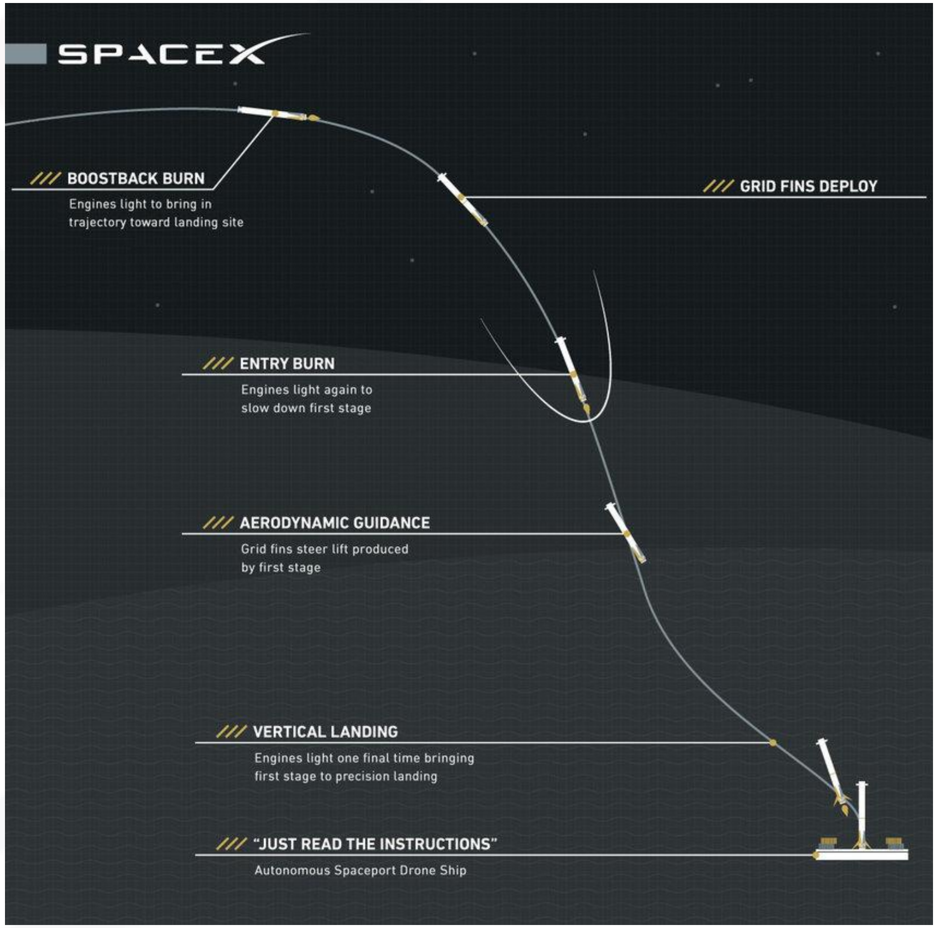
- SpaceX advertises Falcon 9 rocket launches with the cost of 62 million dollars where other providers cost upward to 165 million dollars each.
- The savings is mostly due to reusing the first stage.
- By determining if the first stage will land, we can determine the cost of launch.
- Factors to determine:
  - Payload mass, launch site, number of flights, orbits affects first stage landing success
  - Rate of landing over time.
  - Launch site location.
  - Best performing predictive model for successful landing.

METHODOLOGY

- Collect data using SpaceX REST API and webscraping.
- Wrangle data by filtering data, handling missing values, applying one hot encoding to prepare data for analysis and modeling.
- Exploratory data analysis using SQL database and data visualization techniques.
- Interactive visual analytics
  - Folium for mapping launch sites.
  - Dashboard for launch data
- Build predictive models for landing outcomes.

Data Collection

- Request date from SpaceX REST API
- Decode using .json_normalize() to convert to a dataframe (JSON file)
- Request information for launch data from SpaceX
- Create a dictionary from the data
- Create a dataframe from the dictionary
- Filter for only Falcon 9 launch data from dataframe
- Replace missing values of payload mass by calculating mean
- Export data to CSV file



Data Collection-Web Scraping

2020 [edit]

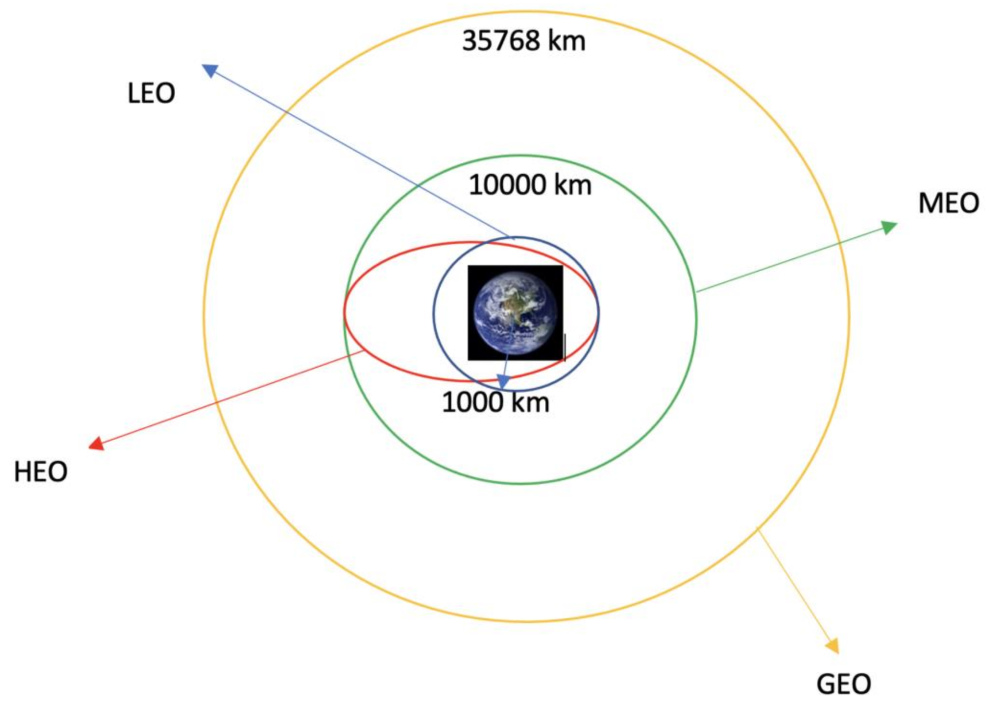
In late 2019, Gwynne Shotwell stated that SpaceX hoped for as many as 24 launches for Starlink satellites in 2020,[490] in addition to 14 or 15 non-Starlink launches. At 26 launches, 13 of which for Starlink satellites, Falcon 9 had its most prolific year, and Falcon rockets were second most prolific rocket family of 2020, only behind China's Long March rocket family.[491]

[hide] Flight No.	Date and time (UTC)	Version, Booster[b]	Launch site	Payload[c]	Payload mass	Orbit	Customer	Launch outcome	Booster landing
78	7 January 2020, 02:19:21[492]	F9 B5 ♺ B1049.4	CCAFS, SLC-40	Starlink 2 v1.0 (60 satellites)	15,600 kg (34,400 lb)[5]	LEO	SpaceX	Success	Success (drone ship)
	Third large batch and second operational flight of Starlink constellation. One of the 60 satellites included a test coating to make the satellite less reflective, and thus less likely to interfere with ground-based astronomical observations.[493]								
79	19 January 2020, 15:30[494]	F9 B5 ♺ B1046.4	KSC, LC-39A	Crew Dragon in-flight abort test[495] (Dragon C205.1)	12,050 kg (26,570 lb)	Sub-orbital[496]	NASA (CTS)[497]	Success	No attempt
	An atmospheric test of the Dragon 2 abort system after Max Q. The capsule fired its SuperDraco engines, reached an apogee of 40 km (25 mi), deployed parachutes after reentry, and splashed down in the ocean 31 km (19 mi) downrange from the launch site. The test was previously slated to be accomplished with the Crew Dragon Demo-1 capsule;[498] but that test article exploded during a ground test of SuperDraco engines on 20 April 2019.[419] The abort test used the capsule originally intended for the first crewed flight.[499] As expected, the booster was destroyed by aerodynamic forces after the capsule aborted.[500] First flight of a Falcon 9 with only one functional stage — the second stage had a mass simulator in place of its engine.								
80	29 January 2020, 14:07[501]	F9 B5 ♺ B1051.3	CCAFS, SLC-40	Starlink 3 v1.0 (60 satellites)	15,600 kg (34,400 lb)[5]	LEO	SpaceX	Success	Success (drone ship)
	Third operational and fourth large batch of Starlink satellites, deployed in a circular 290 km (180 mi) orbit. One of the fairing halves was caught, while the other was fished out of the ocean.[502]								
81	17 February 2020, 15:05[503]	F9 B5 ♺ B1056.4	CCAFS, SLC-40	Starlink 4 v1.0 (60 satellites)	15,600 kg (34,400 lb)[5]	LEO	SpaceX	Success	Failure (drone ship)
	Fourth operational and fifth large batch of Starlink satellites. Used a new flight profile which deployed into a 212 km × 386 km (132 mi × 240 mi) elliptical orbit instead of launching into a circular orbit and firing the second stage engine twice. The first stage booster failed to land on the drone ship[504] due to incorrect wind data.[505] This was the first time a flight proven booster failed to land.								
82	7 March 2020, 04:50[506]	F9 B5 ♺ B1059.2	CCAFS, SLC-40	SpaceX CRS-20 (Dragon C112.3 ♺)	1,977 kg (4,359 lb)[507]	LEO (ISS)	NASA (CRS)	Success	Success (ground pad)
	Last launch of phase 1 of the CRS contract. Carries *Bartolomeo*, an ESA platform for hosting external payloads onto ISS.[508] Originally scheduled to launch on 2 March 2020, the launch date was pushed back due to a second stage engine failure. SpaceX decided to swap out the second stage instead of replacing the faulty part.[509] It was SpaceX's 50th successful landing of a first stage booster, the third flight of the Dragon C112 and the last launch of the cargo Dragon spacecraft.								
83	18 March 2020, 12:16[510]	F9 B5 ♺ B1048.5	KSC, LC-39A	Starlink 5 v1.0 (60 satellites)	15,600 kg (34,400 lb)[5]	LEO	SpaceX	Success	Failure (drone ship)
	Fifth operational launch of Starlink satellites. It was the first time a first stage booster flew for a fifth time and the second time the fairings were reused (Starlink flight in May 2019).[511] Towards the end of the first stage burn, the booster suffered premature shut down of an engine, the first of a Merlin 1D variant and first since the CRS-1 mission in October 2012. However, the payload still reached the targeted orbit.[512] This was the second Starlink launch booster landing failure in a row, later revealed to be caused by residual cleaning fluid trapped inside a sensor.[513]								
84	22 April 2020, 19:30[514]	F9 B5 ♺ B1051.4	KSC, LC-39A	Starlink 6 v1.0 (60 satellites)	15,600 kg (34,400 lb)[5]	LEO	SpaceX	Success	Success (drone ship)

- Request Falcon 9 launch data from Wikipedia
- Create BeautifulSoup object from HTML response.
- Extract column names from HTML table header.
- Collect data by parsing HTML tables
- Create dictionary from data
- Create dataframe from dictionary
- Export data to CSV file

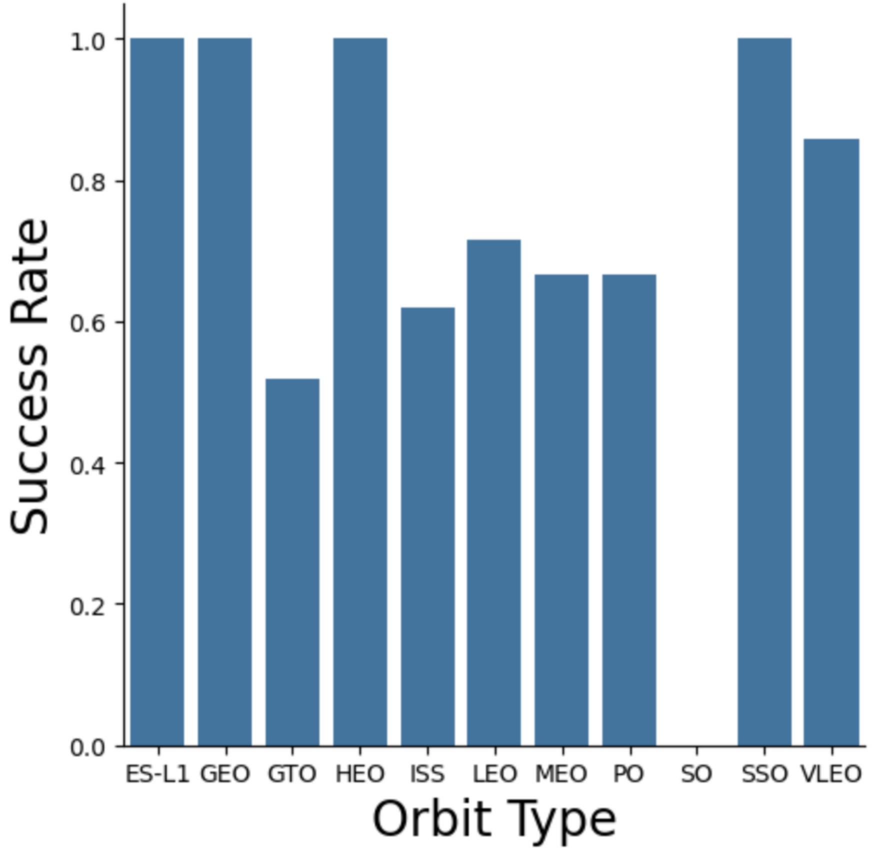
Data Wrangling

- Perform exploratory data analysis to determine training labels
- Calculate:
  - Number of launches for each site
  - Number and occurrence of orbit type
  - Number and occurrence of mission outcome of the orbit types
- Create landing outcome label from outcome column
- Export data to CSV file



EDA with Visualization

- Create charts:
  - Flight Number vs. Payload
  - Flight Number vs. Launch Site
  - Payload Mass (kg) vs. Launch Site
  - Payload Mass (kg) vs. Orbit type
- View relationship using scatter plot.
- Comparisons among discrete categories and measured value using bar charts.



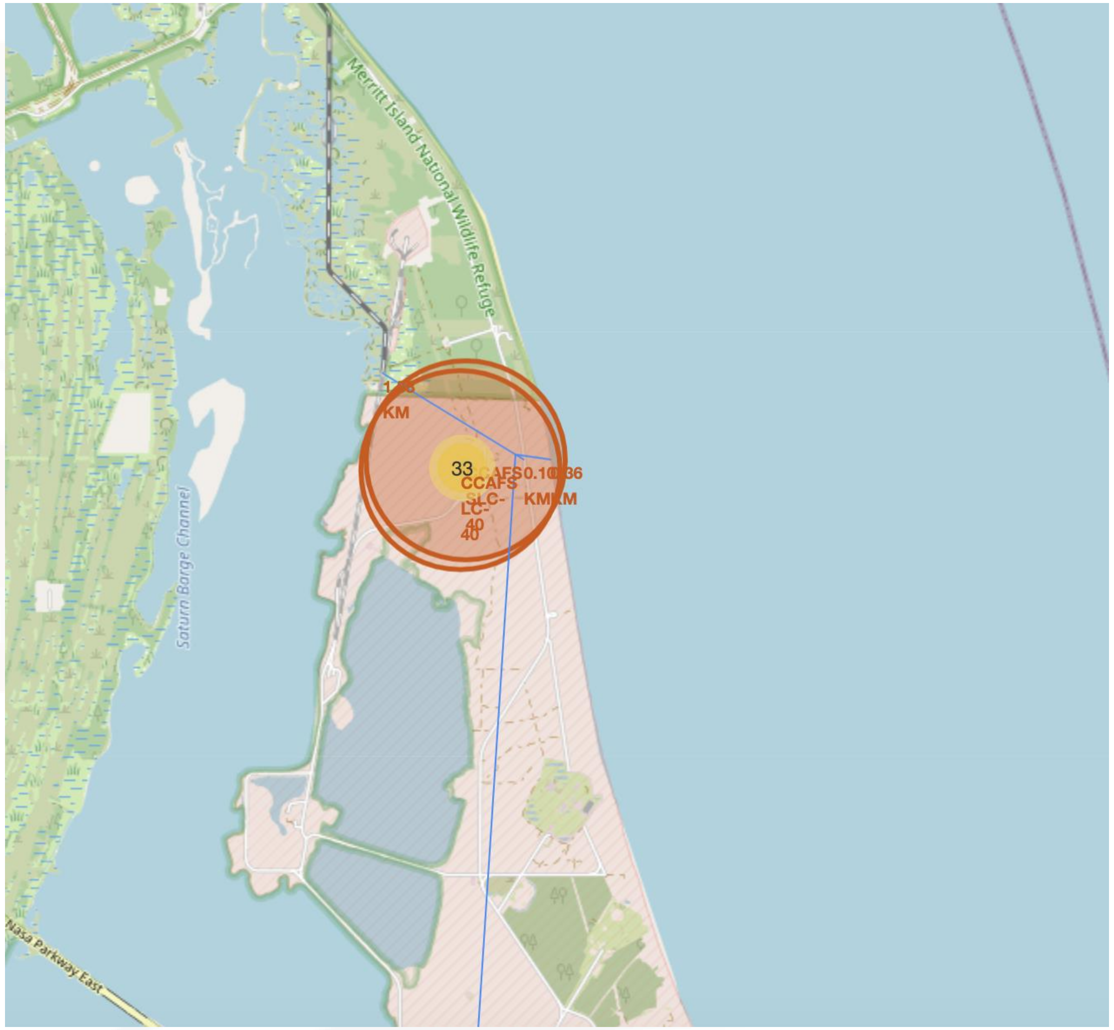
EDA with SQL

- Display
  - Names of each launch site beginning with 'CCA'
  - Total payload mass carried by boosters launched by NASA (CRS)
  - Average payload mass carried by booster version F9 v1.1

- List
  - Date when first successful landing outcome in ground pad was achieved
  - Names of boosters with success in drone ship and have a payload mass greater than 4000 but less than 6000
  - Total number of successful and failure mission outcomes
  - Names of booster versions that carried the maximum payload mass
  - Failed landing outcomes on drone ship including booster version and launch site for each month in 2015
  - Count landing outcomes between 2010-06-04 and 2017-03-20 in desc.

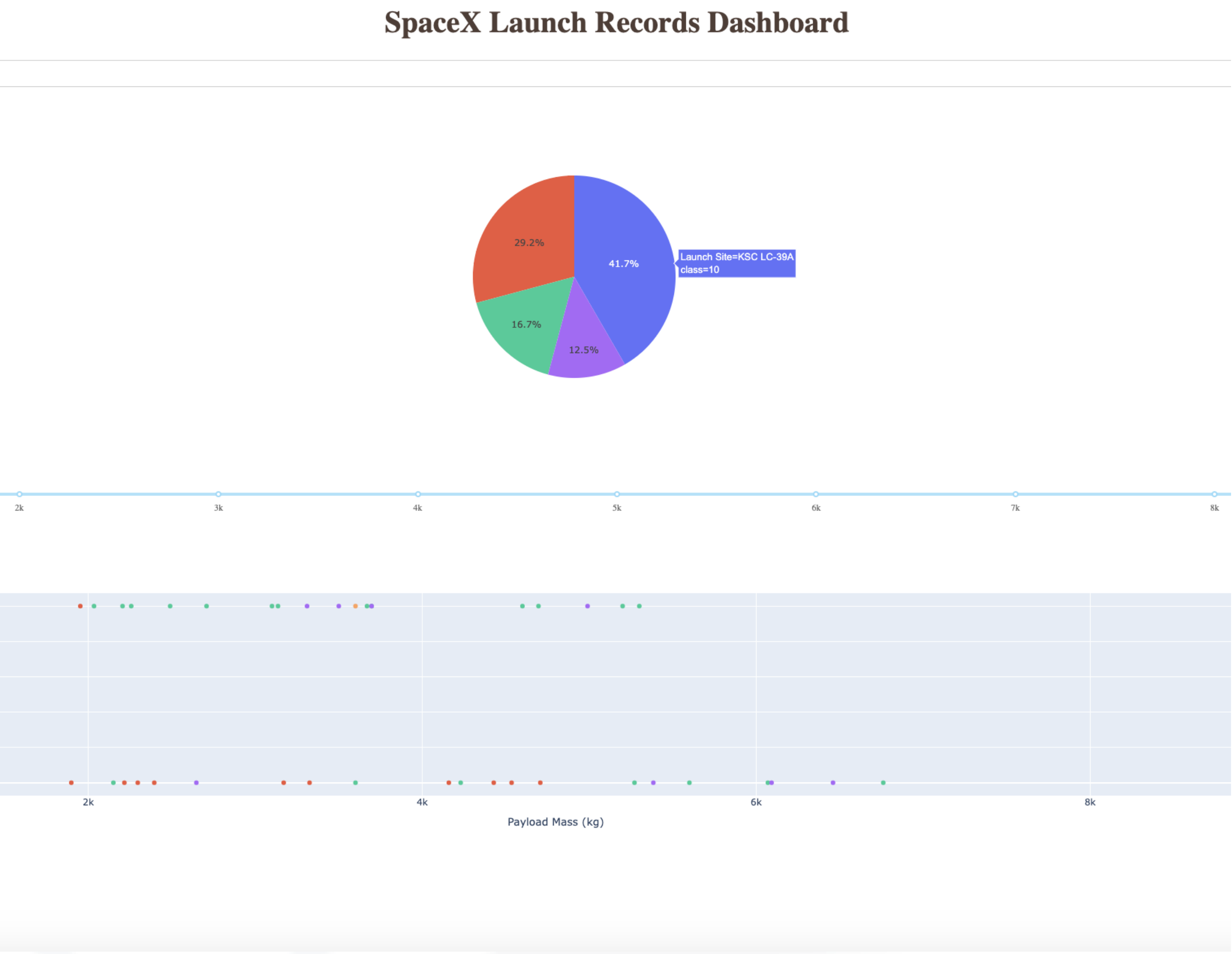
Map with Folium

- Mark all launch sites on a map
  - Added blue circle at NASA Johnson Space Center with pop up label coordinates.
  - Added red circles for all launch sites with pop up labels showing location name and coordinates.
- Mark the success/failed launches for each site on the map
  - Added additional colored markers successful (green) and unsuccessful (red) launches based on success rate.
- Calculate the distances between a launch site to its proximities
  - Added blue lines displaying distance (km) from launch site CCAFS SLC-40 and its proximity to the nearest coastline, railway, and city.



Dashboard with Plotly Dash

- Dropdown list with launch sites
  - Allows users to select all launch sites or specific launch sites
- Pie Chart displaying successful launches
  - Allows users to see successful/unsuccessful launches as percent of the total
- Slider Range for Payload Mass
  - Allows users to select payload mass range
- Scatter Chart displaying Payload Mass vs. Success Rate by Booster Version
  - Allows users to see the correlation between payload and launch success



Predictive Analysis

- Create NumPy array from the Class column
- Standardize the data using StandardScaler. Fit and transform the data.
- Split the data using train_test_split
- Create a GridSearchCV object with cv=10 for parameter optimization
- Apply GridSearchCV to:
  - Logistic Regression
  - Support Vector Machine (SVM)
  - Decision Tree
  - K-Nearest Neighbors
- Calculate accuracy on the test data using score
- Assess confusion matrix for all models
- Identify the best model

Results Summary

- Exploratory Data Analysis
  - KSC LC-39A had the highest success rate
  - Launch success rate increased since 2013 until 2020
  - Orbits ES-L1, GEO, HEO and SSO have highest success rates
- Interactive Visual Analytics
  - Most launch sites were near the Equator and coastline
  - Launch sites proximity:
    - Far from nearest city away from the population due to potential launch hazards
    - Close to highways and railroads for accessibility for launch support
- Predictive Analytics
  - Decision tree is the best predictive model

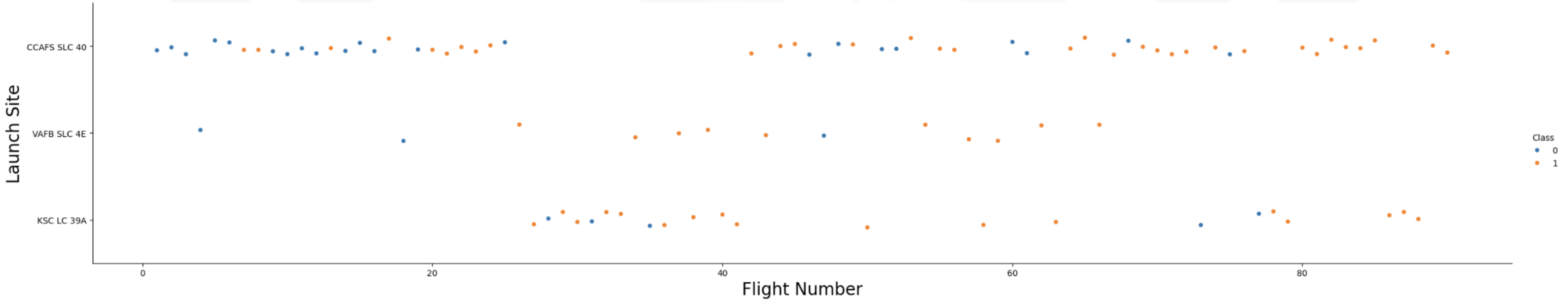
EDA and Visualization Results

IBM Developer

SKILLS NETWORK

Flight Number vs. Launch Site

- About half of launches were from CCAFS SLC 40 launch site
- VAFB SLC 4E and KSC LC 39A have higher success rates
- We can infer that newer launches have a higher success rate

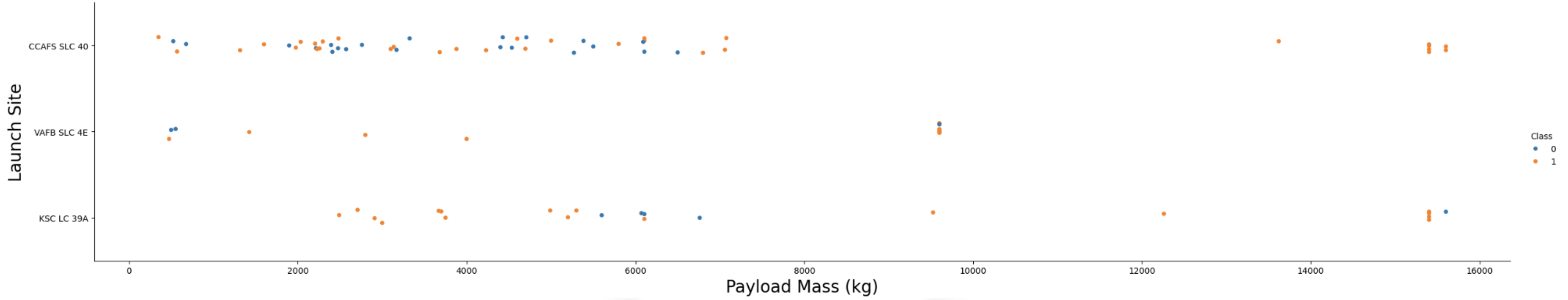


IBM Developer

SKILLS NETWORK

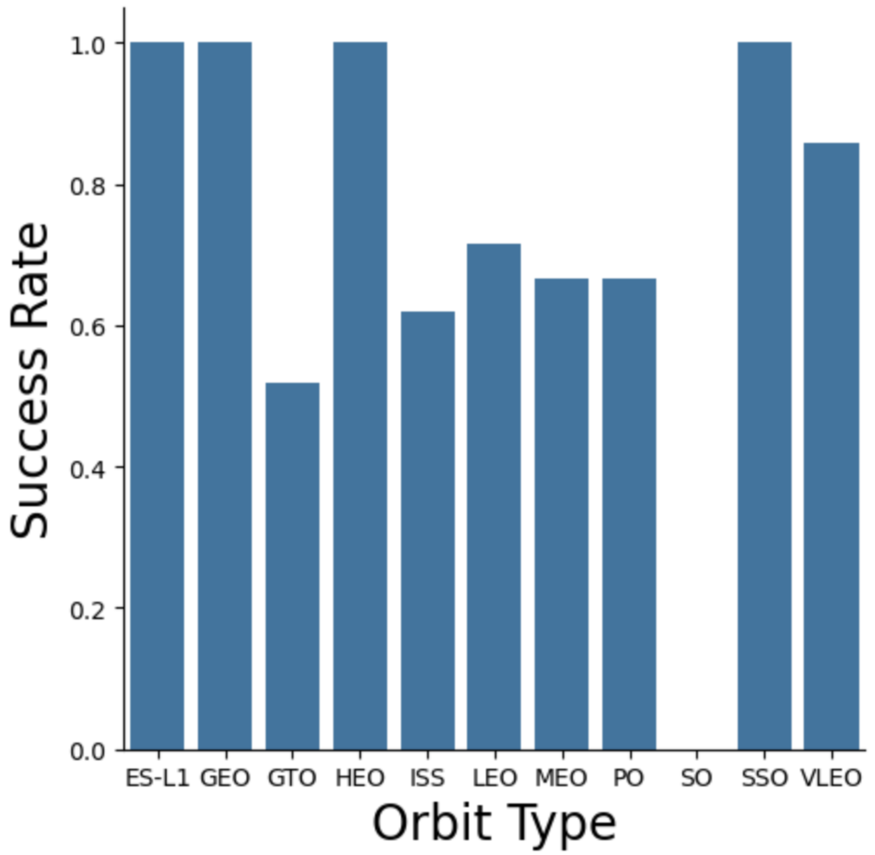
Payload vs. Launch Site

- Most launches with a payload greater than 7,000 kg were successful
- KSC LC 39A has a high success rate for payloads less than 5,500 kg and over 9,000 kg
- CCAFS SLC 40 has a high success rate for payloads greater than 6,000kg
- VAFB-SLC has no rockets that launched a payload greater than 10,000kg



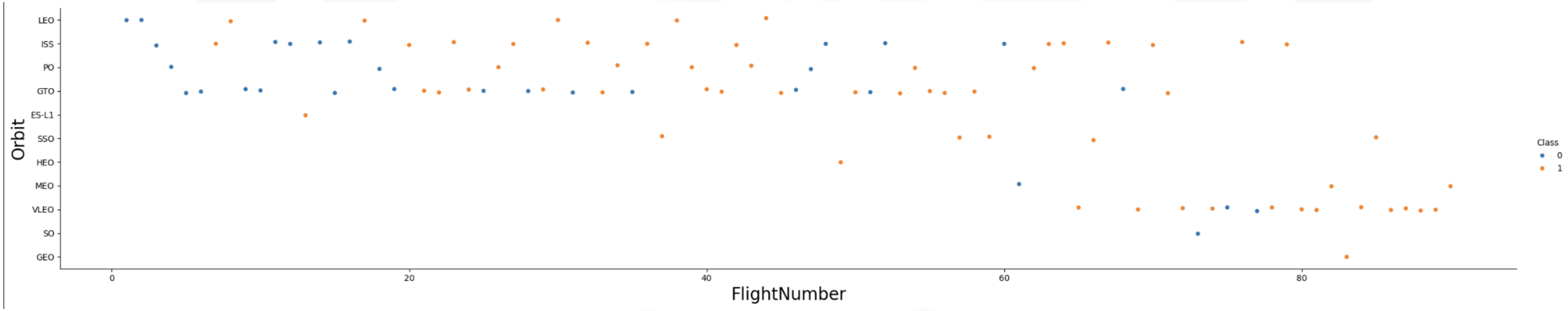
Success Rate By Orbit

- 100% Success Rate
  - ES-L1
  - GEO
  - HEO
  - SSO
- 50%-80% Success Rate
  - GTO
  - ISS
  - LEO
  - MEO
  - PO
- 0% Success Rate
  - SO



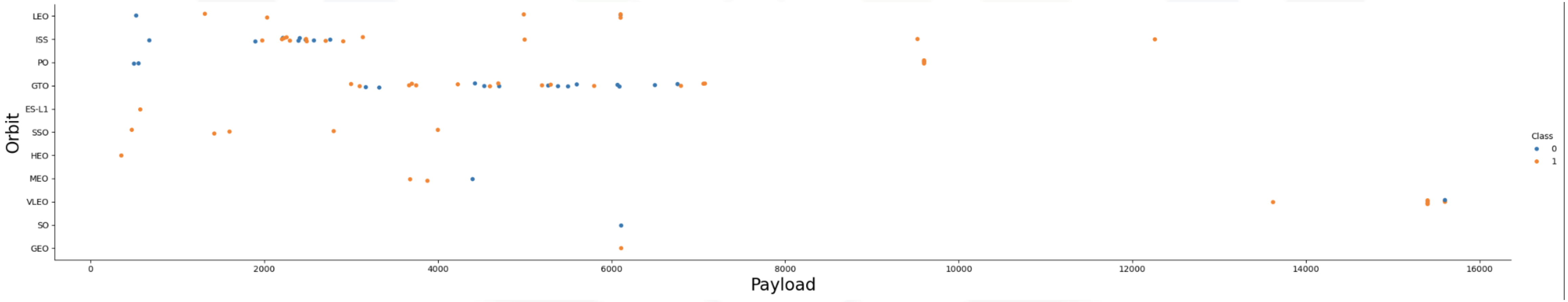
Flight Number vs. Orbit

- ES-L1, SSO, HEO has the least amount of flights with 100% success rate, all in the newer second half of flight numbers
- LEO orbit displays an apparent relationship
- The success rate is higher with the newer (higher) flight numbers



Payload vs. Orbit

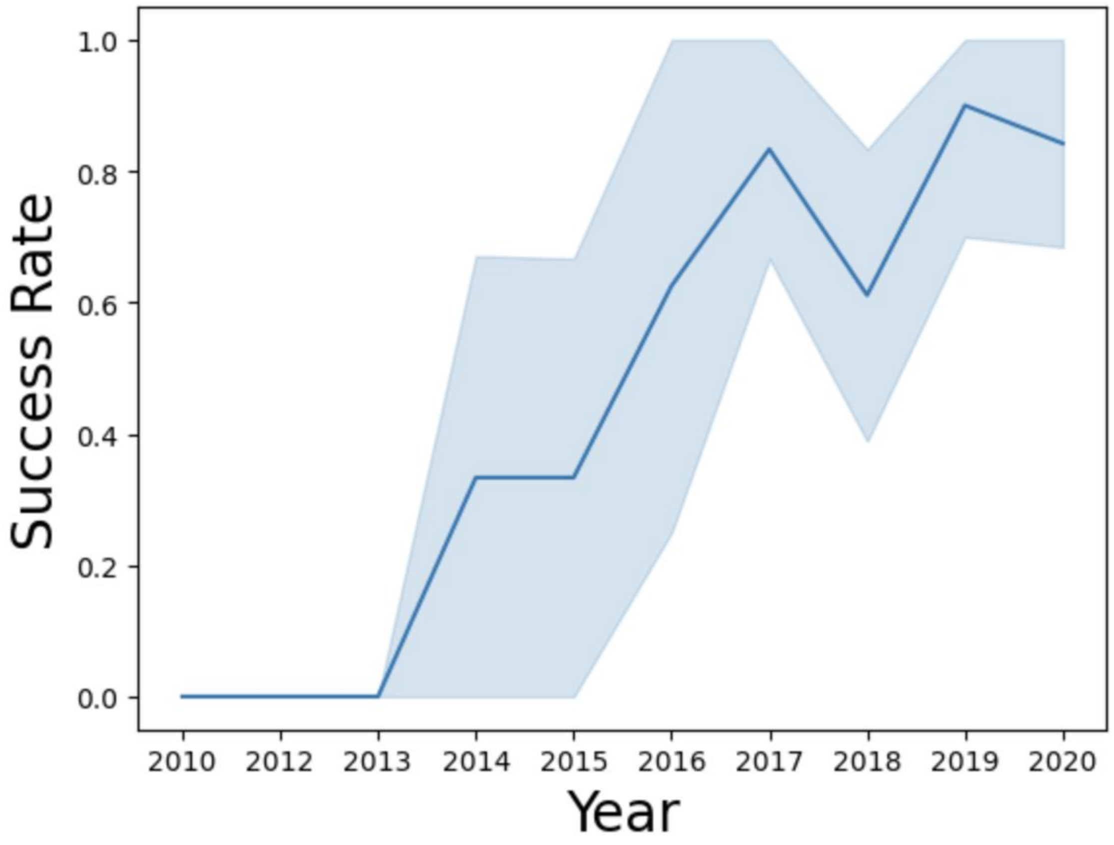
- LEO, ISS, PO orbits were successful with heavy payloads
- GTO and ISS shows mixed success with payloads less than 7,000 kg



IBM Developer

SKILLS NETWORK

Launch Success Over Time



- The success rate steadily increased from year 2013-2017
- The success rate decreased between the year of 2017-2018
- Overall, the success rate increased from 2013-2020

EDA with SQL

IBM Developer

SKILLS NETWORK

Launch Site Info

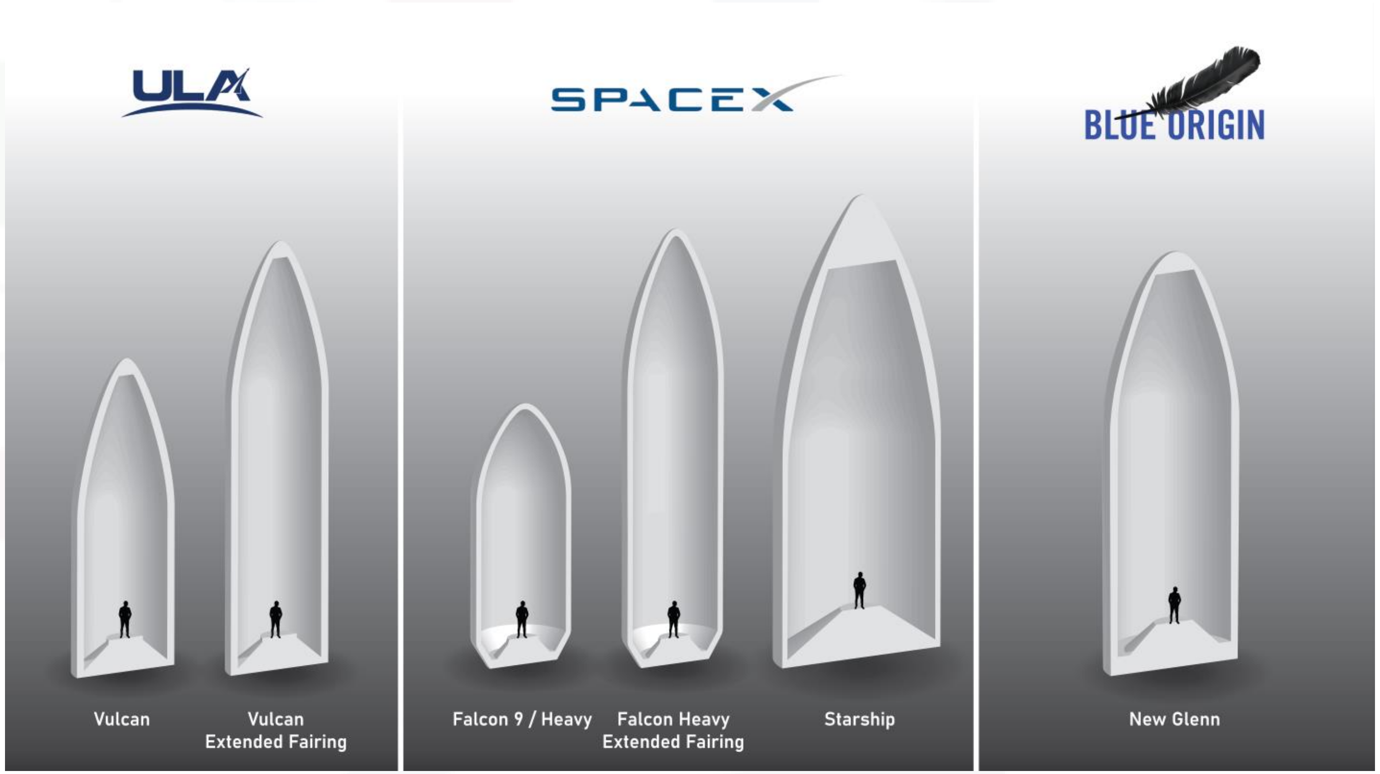
Launch Site Names

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

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

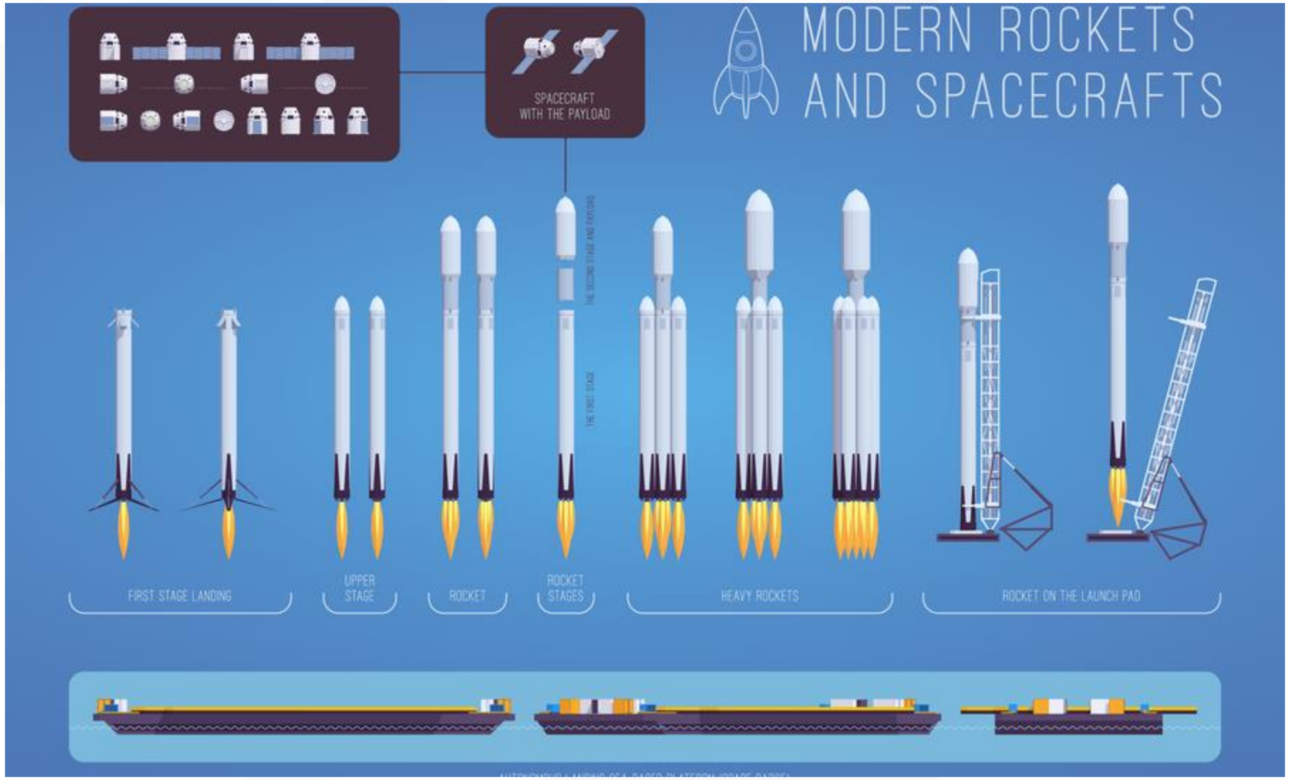
Payload Mass

- Total payload mass carried by boosters launched by NASA (CRS)
    - 45596 kg
- Average payload mass carried by booster version F9 v1.1
    - 2928.4 kg



Landing Information

- First successful landing in ground pad
  - 12/22/2015
- Booster versions that have success in drone ship and have payload mass greater than 4,000 kg but less than 6,000 kg
  - F9 FT B1022
  - F9 FT B1026
  - F9 FT B1021.2
  - F9 FT B1031.2
- Total mission success and failure outcomes
  - Success: 99
  - Failure: 1
  - Success (payload status unclear): 1



Booster Versions Carrying Max Payload Mass

- 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

Failed Landings on Drone Ship

- In year 2015
  - 1/10/2015
  - 4/14/2015

MONTH	Date	Booster_Version	Launch_Site	Landing_Outcome
01	2015-01-10	F9 v1.1 B1012	CCAFS LC-40	Failure (drone ship)
04	2015-04-14	F9 v1.1 B1015	CCAFS LC-40	Failure (drone ship)

Count of Successful Landing

- Ranking the count of successful landing outcomes between 2010-06-04 and 2017-03-20 in descending order

Landing_Outcome	COUNT_OUTCOMES
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

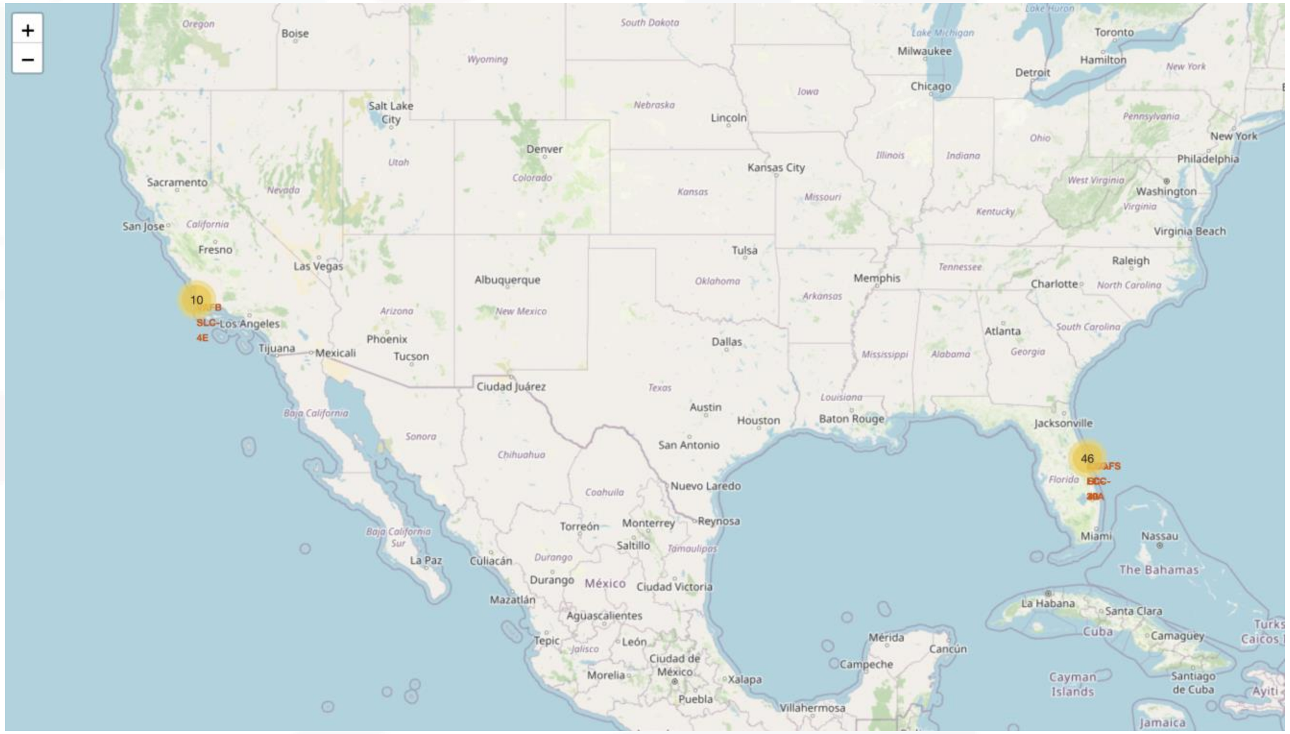
Interactive Maps with Folium

IBM Developer

SKILLS NETWORK

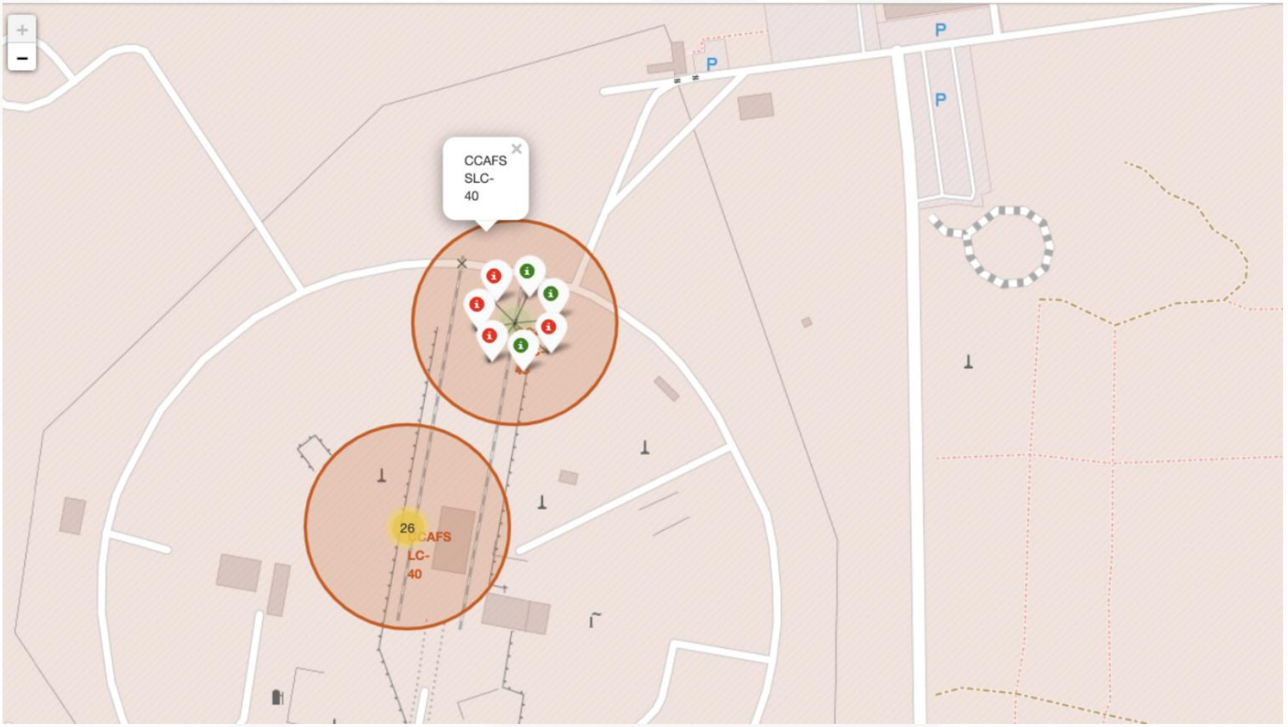
Launch Sites

- All launch sites are in proximity to the Equator where it takes less fuel to launch into space
- All launch sites are close proximity to the coastline due to safety concerns, especially attempting to land first stage



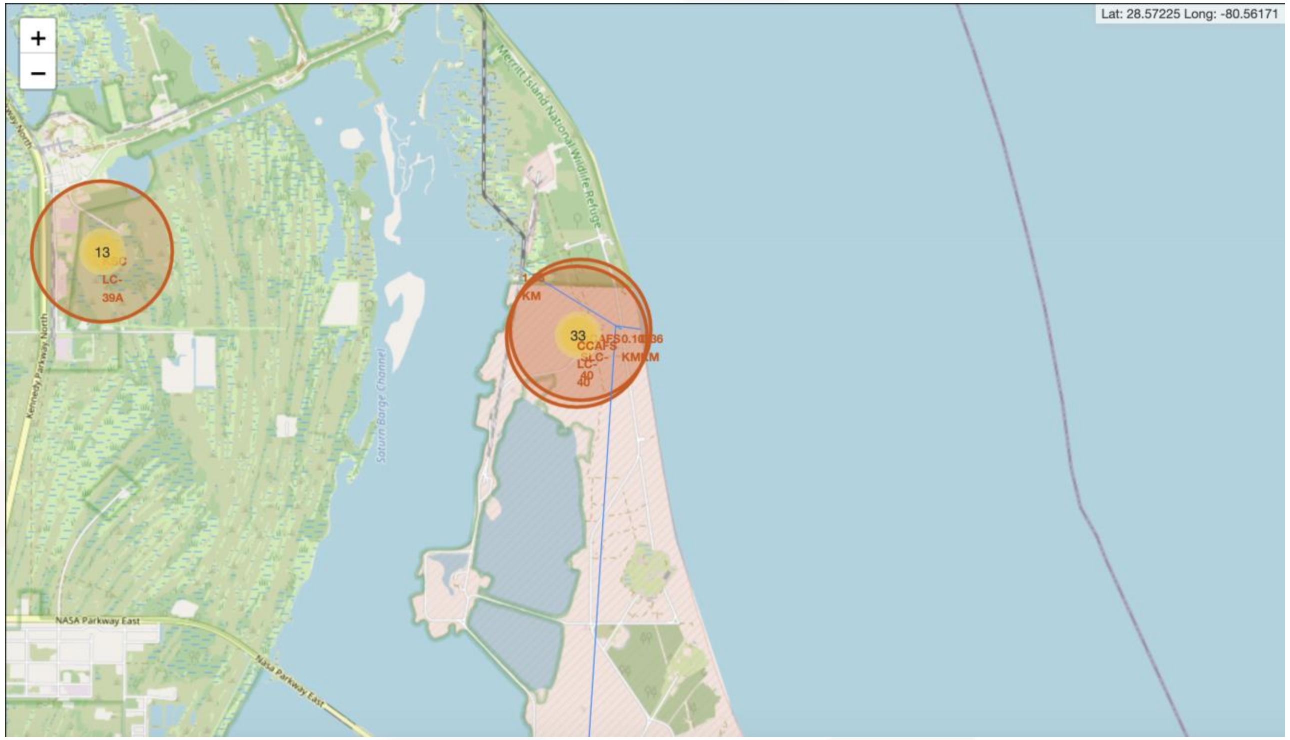
Launch Outcomes

- Green markers: successful launch
- Red markers: unsuccessful launch
- Launch site CCAFS SLC-40 displaying low success rate



Calculating Distance Between Launch Site and its Proximities

- Launch sites are close to proximity to railways, ~ 1.5849991272972073 km, for transportation of necessary resources.
- Launch sites are very close to proximity to highways, ~0.10004284785972558 km, for accessibility for the launch mission.
- Launch sites are fairly close to the coastline, ~0.35845932283653603 km, closer to a body water for safety of potential danger.
- Launch sites are kept far away from nearby cities compared to the other variables, ~54.12533041924254 km, to keep any danger and noise away from highly populated civilian areas.



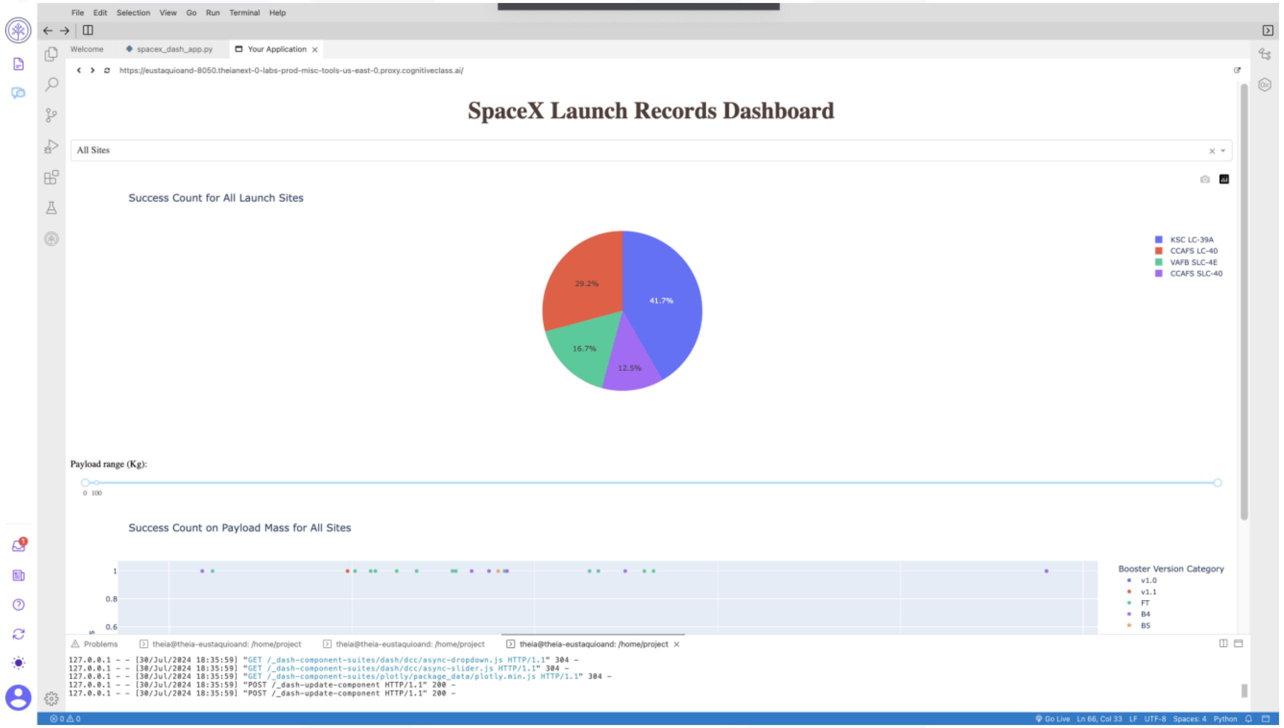
Dashboard with Plotly

IBM Developer

SKILLS NETWORK

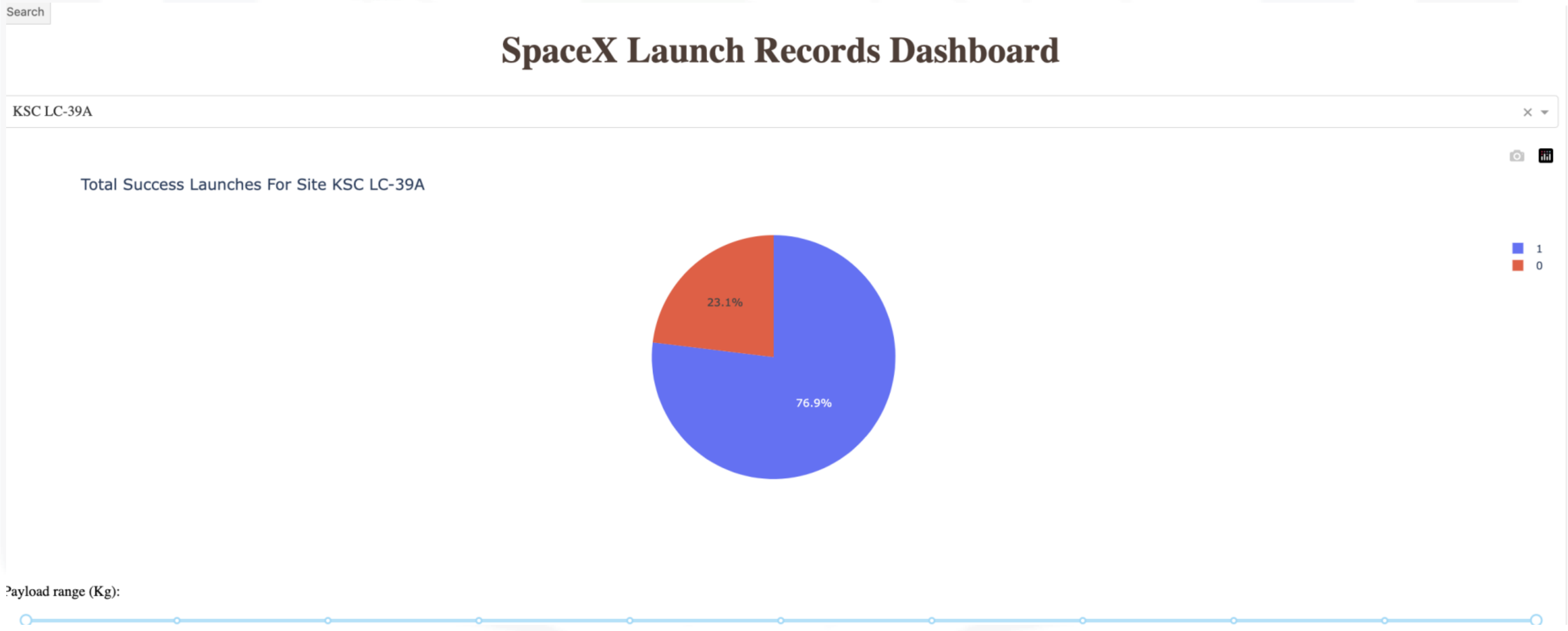
Launch Success by Site

- KSC LC-39A has the highest success rate percentage



Launch Success for KSC LC-29A

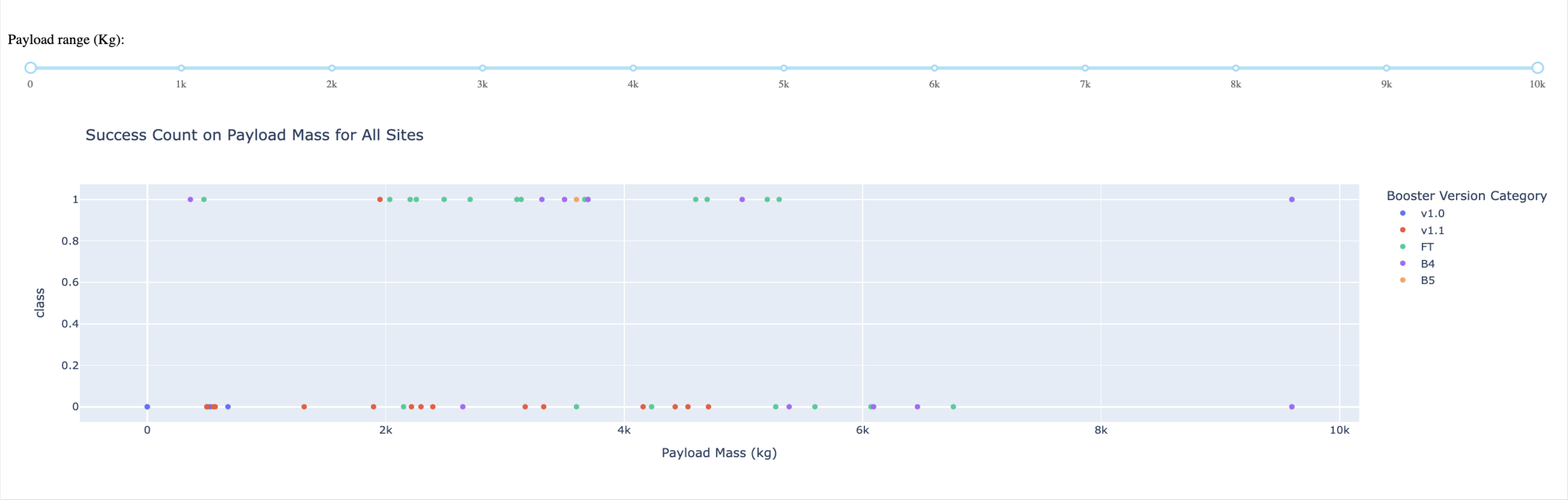
- Highest success rate among launch sites at 76.9%



Payload Mass and Success

By Booster Version Category:

- Payloads between 2,000 kg and 6,000 kg have the highest success rate



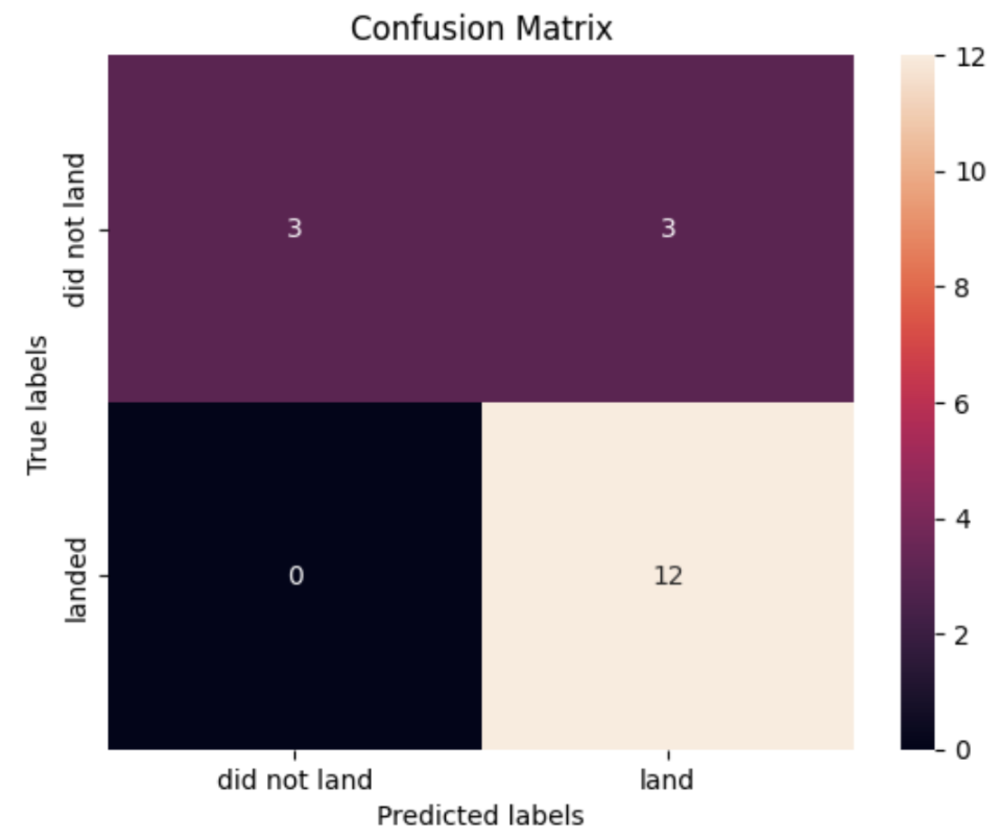
Predictive Analysis

IBM Developer

SKILLS NETWORK

Classification

- All models performed similar with a test set accuracy score of 0.8333
- All methods confusion matrices were identical, Type 1 Error
- The Decision Tree method slightly outperforms with a score of 0.875 when looking at best score
- Best score is the average of all cv folds for a combination of best parameters



Conclusion

- Launch success increases over time in newer flight numbers
- All launch sites are close to the coastline and equator
- Launch site KSC LC-39A has the highest success rate among other launch sites
- Orbits ES-L1, GEO, HEO, SSO have 100% success rate
- The higher the payload mass (kg) the higher the success rate
- Decision tree model for predictive analysis slightly outperforms other models
- A larger dataset may help when building predictive analytics results