IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Leon Altherr
June 5th, 2023

Outline

EXECUTIVE SUMMARY

INTRODUCTION

METHODOLOGY

RESULTS

CONCLUSION

Executive Summary

This project is centred around analysing the data from SpaceX's Falcon 9 launches. The goal is to gain insights that can inform future launch strategies, contribute to improvements in launch success rates, and ultimately support SpaceX's mission to advance space exploration.

- **Summary of Methodologies**:
  - *This project employed a comprehensive methodology that included data collection, data wrangling, exploratory data analysis (EDA), interactive visual analytics, and predictive analysis.*
  - **Data Collection:** Utilised the SpaceX API and Wikipedia web scraping to gather comprehensive data on SpaceX's Falcon 9 launches.
  - **Data Wrangling:** The collected data was cleaned and processed to ensure it was ready for analysis.
  - **Exploratory Data Analysis (EDA):** Conducted an in-depth EDA using visualisation tools and SQL to understand the underlying structure of the data and identify patterns.
  - **Interactive Visual Analytics:** Created interactive visualisations and dashboards using Folium and Plotly Dash for dynamic exploration of the data.
  - **Predictive Analysis:** Built, tuned, and evaluated various classification models to predict the outcome of future launches.
- **Summary of Results**:
  - *The results of the analysis revealed key insights about SpaceX's Falcon 9 launches.*
  - **EDA Results**: Found that the success rate of launches increased over time, certain launch sites were preferred for heavier payloads, and specific orbit types had higher success rates.
  - **Interactive Analytics Results**: Provided dynamic and accessible insights into the data, enhancing the understanding of the launch data.
  - **Predictive Analysis Results:** Identified the Support Vector Machine (SVM) model as the most accurate in predicting the outcome of future launches.
- **Conclusion**:
  - The combination of these methodologies provided a comprehensive understanding of SpaceX's Falcon 9 launches. The insights gained from this analysis could inform future launch strategies, contribute to improvements in launch success rates, and ultimately support SpaceX's mission to advance space exploration.

Introduction

- **Project Background and Context:**
    - This project is centred around SpaceX's Falcon 9 launches. SpaceX, founded by Elon Musk, is a private American aerospace manufacturer and space transportation company.
    - The Falcon 9 is a reusable, two-stage rocket designed and manufactured by SpaceX for the reliable and safe transport of satellites and the Dragon spacecraft into orbit.
    - This project aims to analyse the data from these launches to gain insights and make predictions about future launches.
- **Problem Statement:**
    - Despite the advancements in space technology, there are still challenges in achieving successful launches and landings. These challenges include dealing with variable payload weights, selecting the appropriate launch sites, and determining the most successful orbit types.
    - This project aims to address these challenges by analysing the data from past launches to identify patterns and trends that could inform future launch strategies.
- **Objectives:**
    - Understand the factors that contribute to the success or failure of SpaceX Falcon 9 launches.
    - Develop interactive visualisations that allow for a dynamic exploration of the data.
    - Build predictive models that can accurately predict the outcome of future launches.

Section 1

Methodology

Methodology

Executive Summary

- **Data collection methodology:**
  - I utilised a combination of the SpaceX API and Wikipedia web scraping techniques to gather comprehensive data on SpaceX's Falcon 9 launches.
- **Data Wrangling Methodology:**
  - Data Wrangling: After collection, the data underwent a rigorous wrangling process to ensure it was clean, structured, and ready for analysis. This involved handling missing values, correcting inconsistencies, and transforming data into a suitable format. For instance, categorical data was encoded, and numerical data was normalised or standardised as needed.
  - Data Processing: The cleaned data was then processed to extract meaningful features and insights. This involved aggregating data, creating new variables, and preparing the data for analysis.
- **Perform exploratory data analysis (EDA) using visualization and SQL**
- **Perform interactive visual analytics using Folium and Plotly Dash**
- **Perform predictive analysis using classification models**
  - Model Building, Tuning, and Evaluation: I used cross-validation and hyperparameter tuning to optimise the models. The models' performance was evaluated based on accuracy and confusion matrices. The best-performing model was then selected for further analysis.
- **Conclusion**:
  - My methodology was designed to provide a comprehensive analysis of SpaceX's Falcon 9 launches. By combining rigorous data collection and processing techniques with advanced analytics and predictive modelling, I was able to extract meaningful insights and make accurate predictions about future launches.

Data Collection

SpaceX API

API Request of SpaceX launch data → Convert .json to dataframe → Extract Relevant Columns → Convert to csv

Web Scraping from Wikipedia

HTTP GET → Create Beautiful Soup Object → Extract relevant columns to dataframe → Convert to csv

Data Collection – SpaceX API

1. Send an API request and read the response to obtain the necessary data.
2. Declare and initialize global variables to store relevant information.
3. Call appropriate helper functions, passing the retrieved API data as parameters.
4. Construct structured data using a dictionary, organizing the gathered information.
5. Convert the dictionary into a dataframe and export it as a CSV file.

Data Collection - Scraping

1. Initiate an HTTP GET request to retrieve the HTML page.
2. Generate a Beautiful Soup object to parse and navigate the HTML Structure.
3. Extract the column names from the HTML table header.
4. Create a dictionary using the extracted column names as keys.
5. Call helper functions to populate the record dictionary with launch information.
6. Convert the Dictionary to a dataframe for further analysis and visualization.

Data Wrangling

Load dataframe from CSV

Exploratory Data Analysis of launch patterns

Create Training Labels: Successful landing (1) vs unsuccessful landing (0)

EDA with Data Visualization

- **1. Scatter Plots:** Scatter plots were used to depict the relationships between two numerical variables and allow us to identify any correlations, patterns or trends between them. They are also beneficial in spotting outliers and understanding the data distribution. In this analysis, I utilised scatter plots for:
  - Visualising the correlation between *Flight Number* and *Launch Site*. This provided insight into whether specific launch sites were more commonly used during certain time periods.
  - Analyzing the relationship between *Payload Mass* and *Launch Site*, which helped us determine if heavier payloads were typically launched from specific sites.
  - Studying the association between *Flight Number* and *Orbit Type*. This clarified if certain orbit types were more prevalent in specific stages of the space mission timeline.
  - Exploring the correlation between *Payload Mass* and *Orbit Type*, which enabled us to understand if certain orbit types were more suited to handle specific payload masses.

- **2. Bar Charts:** Bar charts are effective tools for comparing the sizes of different groups and are frequently used when dealing with categorical data. The lengths of the bars are proportional to the values they represent, making it easy to compare groups against each other. In this case, I plotted a bar chart to:
  - Visualise the success rate of each orbit type. This showed the relative success rates for different orbit types, providing a comprehensive comparison between them.

- **3. Line Charts:** Line charts excel in illustrating trends and changes over time, making them an ideal choice when dealing with time-series data. In this analysis, I deployed a line chart to:
  - Trace the average launch success rate over the years. This helps us understand the progress in mission success over time, indicating whether improvements in technology or processes had a positive impact on mission success rates.

EDA with SQL

- **Task 1:** Display the names of the unique launch sites in the space mission
  - SELECT DISTINCT Launch_Site FROM SPACEXDATASET
- **Task 2:** Display 5 records where launch sites begin with the string 'CCA'
  - SELECT * FROM SPACEXDATASET WHERE Launch_Site LIKE 'CCA%' LIMIT 5
- **Task 3:** Display the total payload mass carried by boosters launched by NASA (CRS)
  - SELECT SUM(PAYLOAD_MASS__KG_) FROM SPACEXDATASET WHERE Customer = 'NASA (CRS)'
- **Task 4:** Display average payload mass carried by booster version F9 v1.1
  - SELECT AVG(PAYLOAD_MASS__KG_) FROM SPACEXDATASET WHERE Booster_Version LIKE 'F9 v1.1'
- **Task 5:** List the date when the first successful landing outcome in ground pad was achieved
  - SELECT MIN(DATE) FROM SPACEXDATASET WHERE Landing_Outcome = 'Success (ground pad)'
- **Task 6:** List the names of the boosters which have succeeded in landing on the first ground pad for the first time
  - SELECT Booster_Version FROM SPACEXDATASET WHERE Landing__Outcome = 'Success (ground pad)' GROUP BY Booster_Version HAVING MIN(DATE)
- **Task 7:** List the total number of successful and failure mission outcomes
  - SELECT Mission_Outcome, COUNT(*) FROM SPACEXDATASET GROUP BY Mission_Outcome
- **Task 8:** List the names of the booster_versions which have carried the maximum payload mass
  - SELECT Booster_Version, MAX(PAYLOAD_MASS__KG_) FROM SPACEXDATASET GROUP BY Booster_Version
- **Task 9:** Show the records where launch sites begin with the string 'KSC'
  - SELECT * FROM SPACEXDATASET WHERE Launch_Site LIKE 'KSC%'
- **Task 10:** Display the total number of successful landing_outcomes at each landing_site
  - SELECT Landing_Site, COUNT(*) FROM SPACEXDATASET WHERE Landing_Outcome = 'Success' GROUP BY Landing_Site

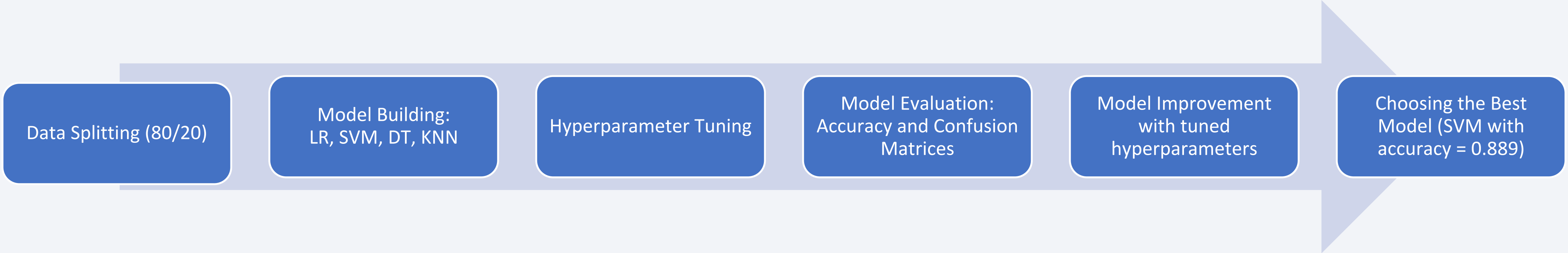
Build an Interactive Map with Folium

- **Markers:** Used to mark the location of each launch site (CCAFS LC-40, CCAFS SLC-40, KSC LC-39A, VAFB SLC-4E) using their latitude and longitude coordinates. This visualised the exact geographical location of each launch site.
- **CircleMarkers:** Each launch site was encircled with a CircleMarker. This enhanced the visual clarity, distinctly highlighting the launch sites on the map.
- **Marker Clusters:** Used to group markers of the launch results. It helps in handling multiple records with the same coordinates, providing a clean, uncluttered representation of successful (green) and failed (red) launches at sites.
- **PolyLines:** Created to represent distances between various points of interest such as the launch site, closest city, coastline, railway, and highway. This line representation provides a clear idea of proximity to important infrastructure and geographic features.
- **Popups:** These were attached to the CircleMarkers to display the name of the launch site when clicked, providing an interactive element to the map.
- **Reasoning**:
    - To represent launch site locations visually for easier geographical interpretation.
    - To distinguish between successful and failed launches, helping to identify trends or patterns.
    - To understand the proximities to different infrastructure and geographic features, which are crucial for logistical and environmental considerations.

Build a Dashboard with Plotly Dash

- The Plotly Dash web application performs real-time visual analytics on the SpaceX launch data. The following interactive features and visual plots were added and improved our ability to extract and understand the underlying data:
- **Launch Site Drop-down Input:** This feature allows users to filter the dashboard visuals based on either all launch sites or a specific one, enhancing the user's ability to view data from multiple perspectives.
- **Pie Chart:** This graphical representation displays the total successful launches when 'All Sites' is selected and distinguishes between successful and failed launches when a specific site is chosen. It simplifies complex data into a visually comprehensive format.
- **Payload Range Slider:** This interactive element allows users to conveniently select different payload ranges and identify visual patterns, providing a more intuitive data exploration experience.
- **Scatter Chart:** It illustrates the potential correlation between payload and mission outcomes for selected site(s). With the color-label Booster version on each scatter point, it visually highlights mission outcomes with different boosters, aiding in better data interpretation.
- These enhancements have substantially enriched the dashboard's interactivity and analytical capabilities. Here are some questions the updated dashboard can answer more effectively:
    - Which site has the highest number of successful launches? (Answer: KSC LC-39A with 10 launches)
    - Which site boasts the highest launch success rate? (Answer: KSC LC-39A with a 76.9% success rate)
    - Which payload range(s) sees the highest launch success rate? (Answer: 2000 – 5000 kg)
    - Which payload range(s) experience the lowest launch success rate? (Answer: 0-2000 and 5500 - 7000 kg)
    - Which F9 Booster version (v1.0, v1.1, FT, B4, B5, etc.) has the highest launch success rate? (Answer: FT)
    - The rationale behind these additions was to boost user engagement, simplify complex data visualization, and extract more meaningful insights from the SpaceX launch data.

Predictive Analysis (Classification)



1. **Data Splitting:** The dataset is split into 80% for training and 20% for testing, ensuring separate subsets for model development and evaluation.
2. **Model Building**: Several algorithms are employed to build predictive models, including Logistic Regression, Support Vector Machine, Decision Tree, and K-Nearest Neighbors. Each algorithm offers unique strengths and characteristics.
3. **Hyperparameter Tuning**: Cross Validation is utilized to optimize model performance by systematically searching for the best combination of hyperparameters. This process ensures robust parameter selection.
4. **Model Evaluation: The** accuracy metric is employed to assess model performance on the test data. Additionally, confusion matrices are plotted to gain insights into the model's predictive performance across different classes.
5. **Model Improvement:** The top-performing models are selected based on their hyperparameter settings and performance evaluation. These models are refined and considered for further analysis.
6. **Choosing the Best Model:** The Support Vector Machine (SVM) algorithm is identified as the best model with an accuracy of 0.889. This model demonstrates the highest predictive capability among the tested algorithms.

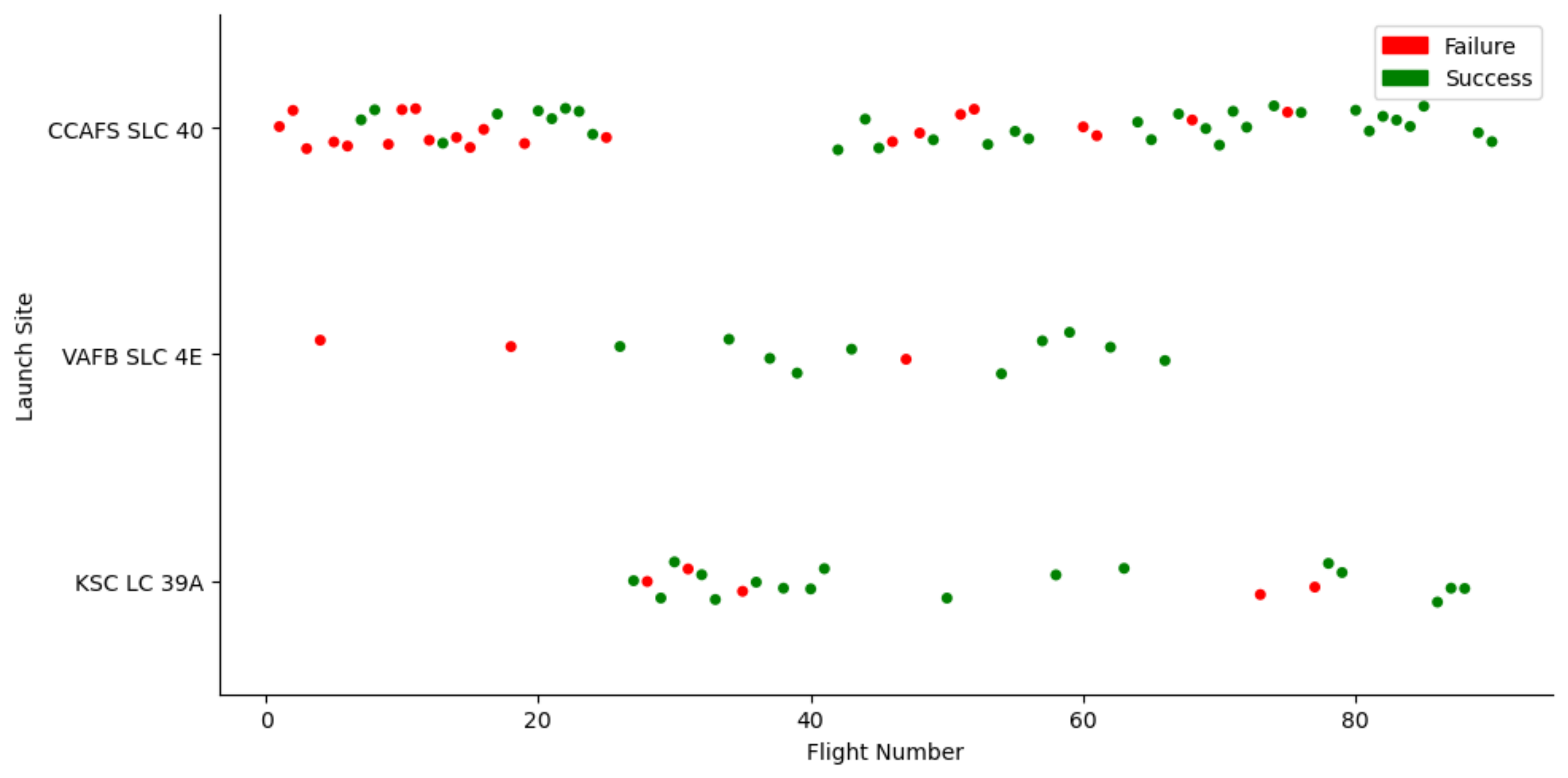
Results

- **Exploratory Data Analysis (EDA) Results:** The EDA revealed key insights about SpaceX's Falcon 9 launches.
  - For instance, it was found that the success rate of launches increased over time, certain launch sites were preferred for heavier payloads, and specific orbit types had higher success rates. These findings were visualised using scatter plots, bar charts, and line charts.
- **Interactive Analytics Demo:** The interactive analytics demo provided dynamic insights into the data.
  - Using Folium, an interactive map was created that visualised the location of each launch site and the results of each launch.
  - With Plotly Dash, a dashboard was developed that allowed real-time analysis of the data. Screenshots of these interactive elements are included to demonstrate their functionality and the insights they provided.
- **Predictive Analysis Results:**
  - The predictive analysis involved building, tuning, and evaluating various classification models.
  - The Support Vector Machine (SVM) model was identified as the best model with an accuracy of 0.889. This model demonstrated the highest predictive capability among the tested algorithms, suggesting it could accurately predict the outcome of future SpaceX Falcon 9 launches.
- **Conclusion**:
  - The results of the EDA, interactive analytics, and predictive analysis provided valuable insights into SpaceX's Falcon 9 launches. These insights could be used to inform future launch strategies and improve the success rate of launches.

Section 2

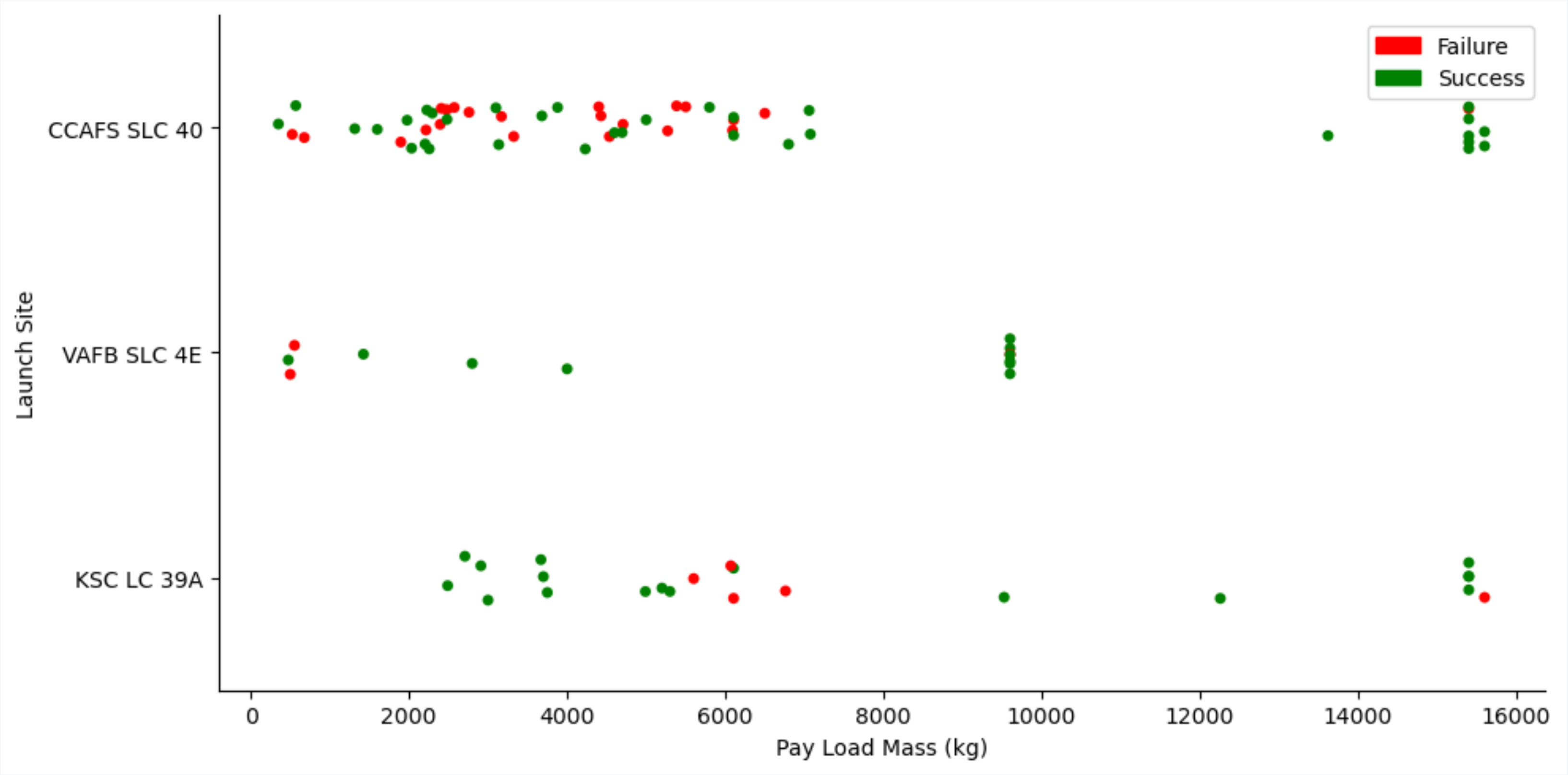
Insights drawn from EDA

Flight Number vs. Launch Site



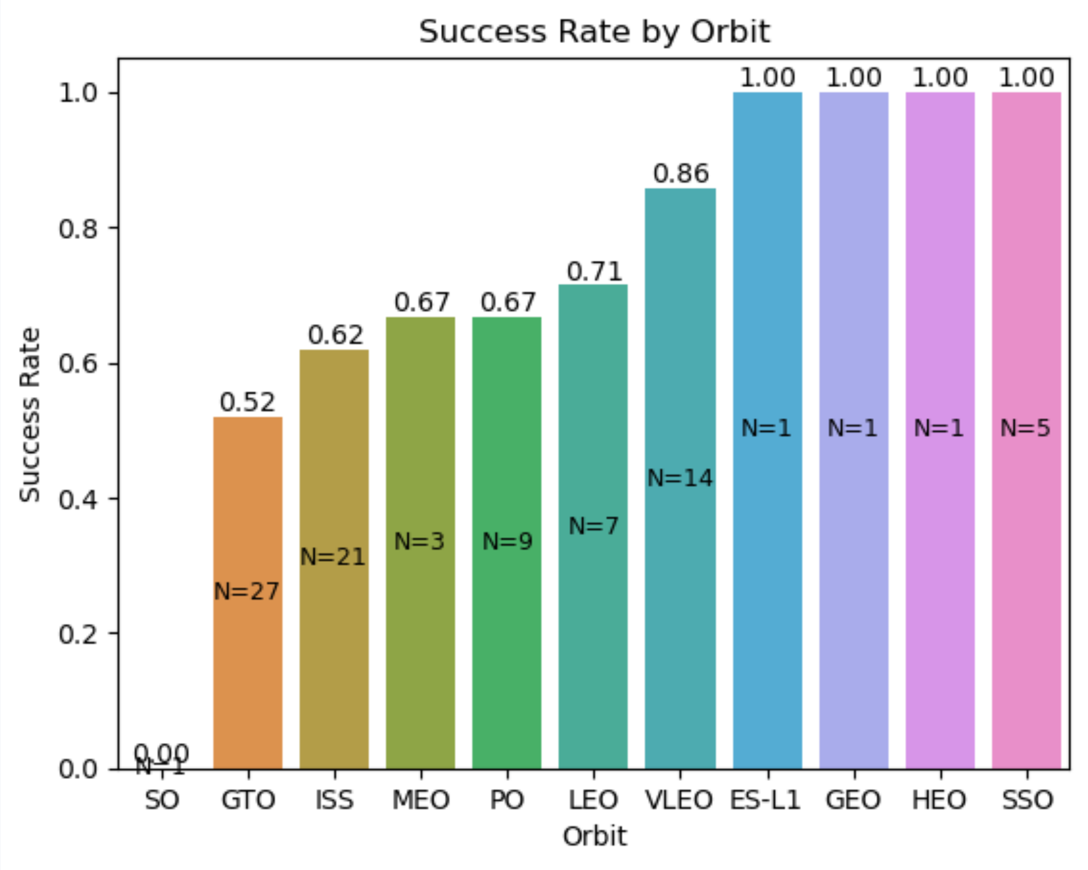
- The success proportion increases as the number of flights grow across all launch sites.
- CCAFS SLC 40 was the preferred launch site.
- KSC LC39A emerged as the most popular launch site between the 25th and 40th flights.

Payload vs. Launch Site



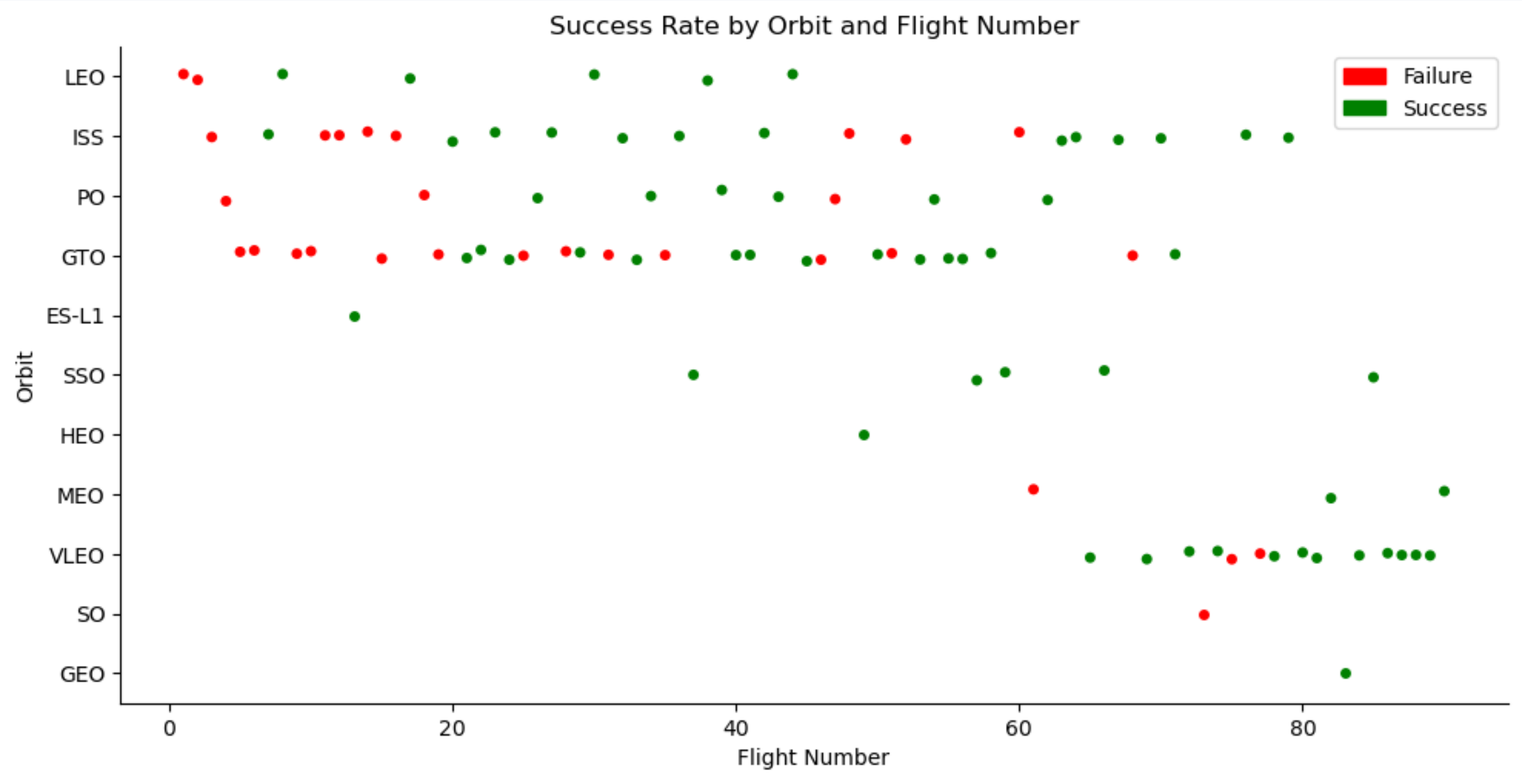
- No rockets were launched at 'VAFB SLC 4E' for payloads exceeding 10,000 kg.
- The proportion of successes increase with payload mass.
  -> May be attributable to better preparation.
- There is a relative gap of payload masses between 10T and 15T with just 2 datapoints

Success Rate vs. Orbit Type



- Orbits ES-L1, GEO, HEO and SSO have the highest success rate.
- However, when only those with at least five launches are considered:
  - 1. SSO: 100% success
  - 2. VLEO: 86%
  - 3. LEO: 71%
- GTO has by far the worst success rate (52%) when SO with just one launch is excluded.

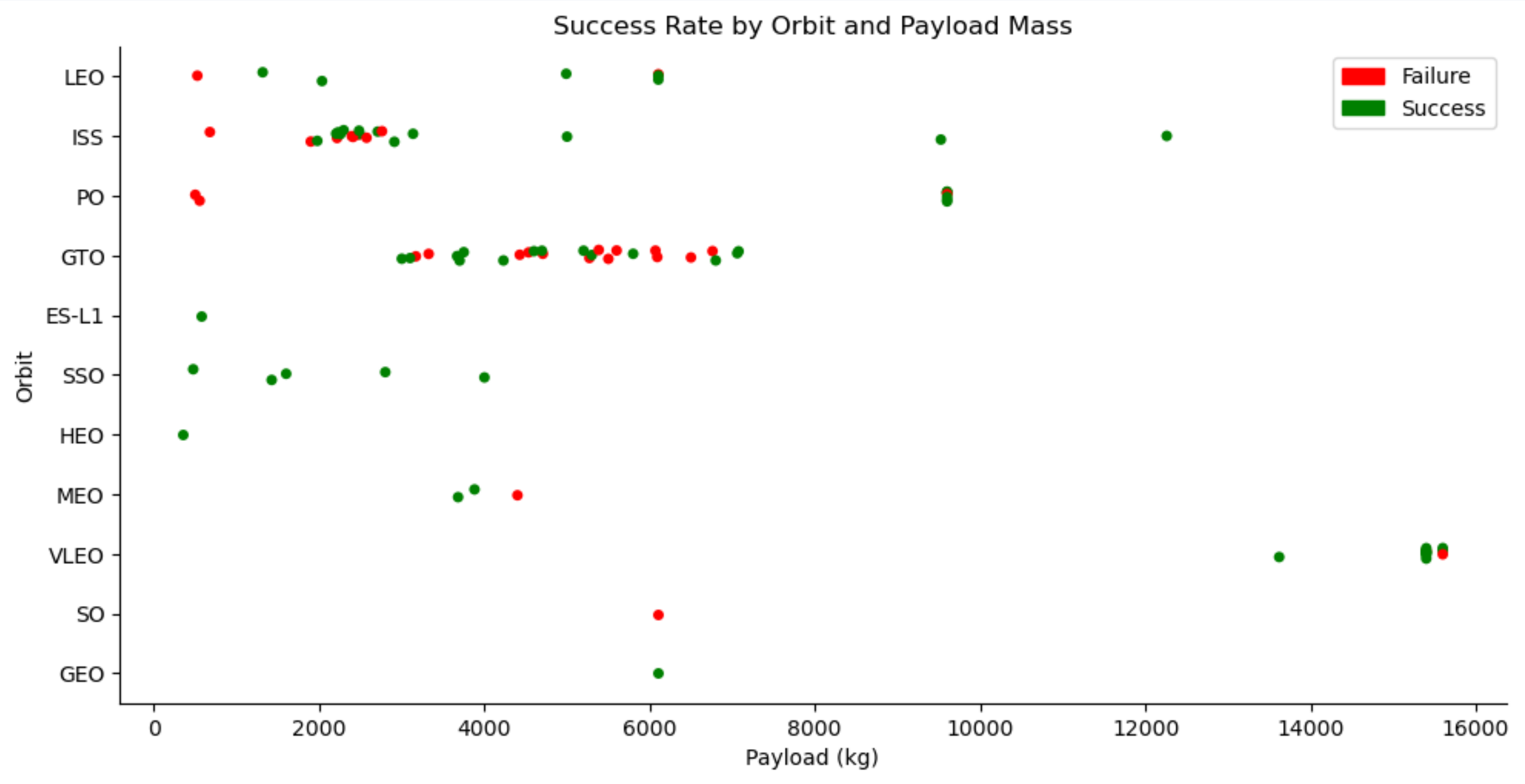
Flight Number vs. Orbit Type



Key Points:

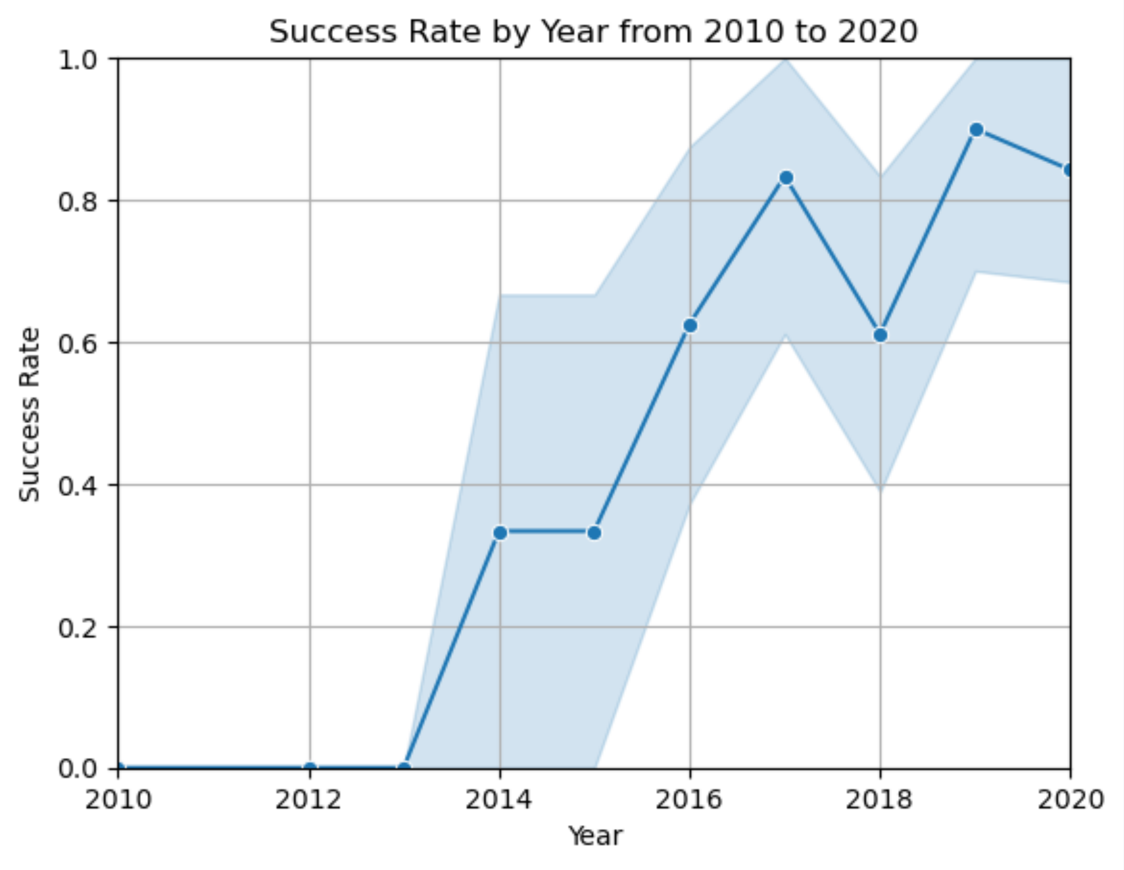
- For the first 50 launches most were in the LEO, ISS, PO and GTO orbit.
- The first launches in each orbit were often unsuccessful.
- The success rate seems to increase for each orbit across flights.

Payload vs. Orbit Type



- VLEO orbit launches were the heaviest and the only orbit with 13T+ launches
- There were no launches over 5T for ES-L1, SSO, HEO and MEO
- ISS had the widest range with 1T to >12T

Launch Success Yearly Trend



Key Points:

- Success Rate increased from 0% in 2012 to >80% in 2020.
- Launches started in 2012.
- Since 2016 the success rate has been over 60% every year.
- The first successful launch was in 2014
- 2017-2018 and 2019-2020 are the only two years where success rate decreased relative to the previous year

All Launch Site Names

Main Points:
- There were four distinct launch sites.
- CCAFS is the only launch site with two sub-sites (LC-40 and SLC-40)

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

Launch Site Names Begin with 'CCA'

Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer	Mission_Outcome	Landing _Outcome
04-06-2010	18:45:00	F9 v1.0 B0003	CCAFS LC-40	Dragon Spacecraft Qualification Unit	0	LEO	SpaceX	Success	Failure (parachute)
08-12-2010	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)
22-05-2012	07:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525	LEO (ISS)	NASA (COTS)	Success	No attempt
08-10-2012	00:35:00	F9 v1.0 B0006	CCAFS LC-40	SpaceX CRS-1	500	LEO (ISS)	NASA (CRS)	Success	No attempt
01-03-2013	15:10:00	F9 v1.0 B0007	CCAFS LC-40	SpaceX CRS-2	677	LEO (ISS)	NASA (CRS)	Success	No attempt

- The launches were conducted exclusively at CCAFS LC-40 (Cape Canaveral Air Force Station, Launch Complex 40).
- The initial five launches took place between 2010 and 2013.
- The first three launches carried the Dragon payload, while the 2012 and 2013 launches featured the SpaceX payload.
- All launches employed F9 v1.0 Boosters, which achieved successful mission outcomes but did not successfully land (two failures and three no attempts).
- SpaceX was the first customer, with NASA being the customer for the remaining four launches.

Total Payload Mass for NASA boosters

- 45.596T was the total payload mass for all NASA boosters launched by SpaceX

45596

Average Payload Mass by F9 v1.1

- The average Payload Mass carried by F9 v1.1 Boosters was 2.928T

2928

First Successful Ground Landing Date

- The first Successful Ground Landing was on the 22nd December 2015

min_date
2015-12-22

Successful Drone Ship Landing with Payload between 4000 and 6000

Booster_Version	PAYLOAD_MASS__KG_
F9 B5 B1046.2	5800
F9 B5 B1047.2	5300
F9 B5 B1046.3	4000
F9 B5 B1048.3	4850
F9 B5 B1051.2	4200
F9 B5B1060.1	4311
F9 B5 B1058.2	5500
F9 B5B1062.1	4311

- Range: 4200kg to 5800kg
- All used versions of the F9 Booster
- A slightly different version was used each time, suggesting that improvements were made between every successful drone ship landing

Total Number of Successful and Failure Mission Outcomes

mission_outcome	counts
Failure (in flight)	1
Success	99
Success (payload status unclear)	1

- Only one of 101 flights failed
- 99 mission outcomes were successful
- 1 was successful, but the status of the payload was unclear

Boosters Carried Maximum Payload

Booster Version	Maximum Payload
F9 B5 B1048.4	15600
F9 B5 B1049.4	15600
F9 B5 B1051.3	15600
F9 B5 B1056.4	15600
F9 B5 B1048.5	15600
F9 B5 B1051.4	15600
F9 B5 B1049.5	15600
F9 B5 B1060.2	15600
F9 B5 B1058.3	15600
F9 B5 B1051.6	15600
F9 B5 B1060.3	15600
F9 B5 B1049.7	15600

Key Points:
1. The maximum payload carried for boosters was 15.6T (metric tons).
2. A total of 12 F9 boosters successfully carried the maximum payload.
3. Each booster represented a distinct version, suggesting ongoing improvements between subsequent maximum payload launches.

2015 Launch Records – Failed Landings

- There were two failed F9 Booster landings in 2015 (January and April)
- Both were launched from CCAFS LC-40
- Both failures were attributed to the drone ship

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

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

- No attempt was made 10 times, compared to 22 attempts.
- 15 attempts were successful, 7 failed
- There were 5 successful landings on the drone ship and ground pad, with 5 ocean landings.
- 6 failures happened on the drone ship and1 during the parachute

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

Section 3

Launch Sites Proximities Analysis

Map of All Space X Falcon9 Launch Sites

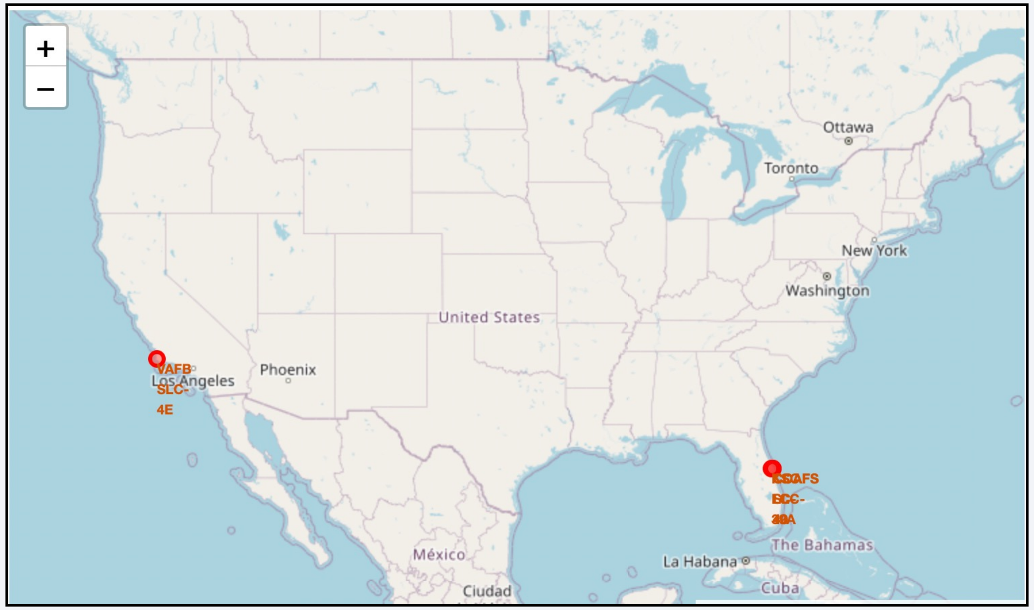


Figure 1: Global Map with Marked Launch Sites

Figure 3: Zoom Map with VAFP SLC-4E Launch Site

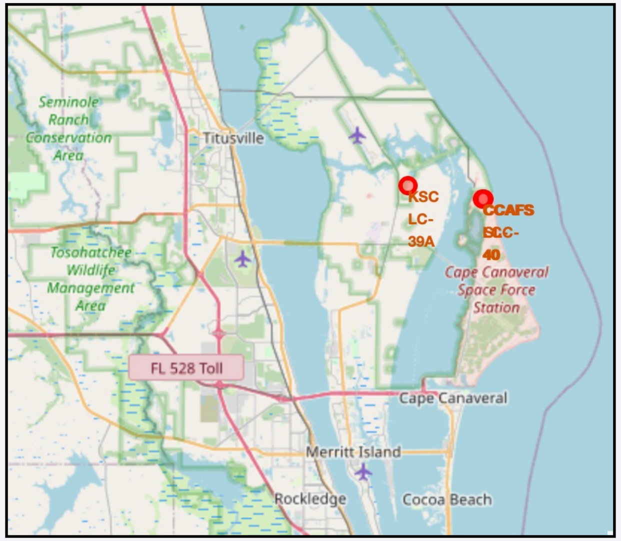


Figure 4: Zoom Map showing:

i) KSC LC-39A
ii) CCAFS LC-40
iii) CCAFS SLC40

- Each launch site is indicated by a labelled circle and a popup for location details.
- Figure 1 shows all Falcon 9 launch sites. All four are located in the US, one in California and three in Florida.
- Figure 2 shows the VAFP SLC-4E launch site in California and its surroundings, including airports and coast.
- Figure 3 shows the remaining three launch sites and their surroundings.

Map showing successful and failed SpaceX F9 launches

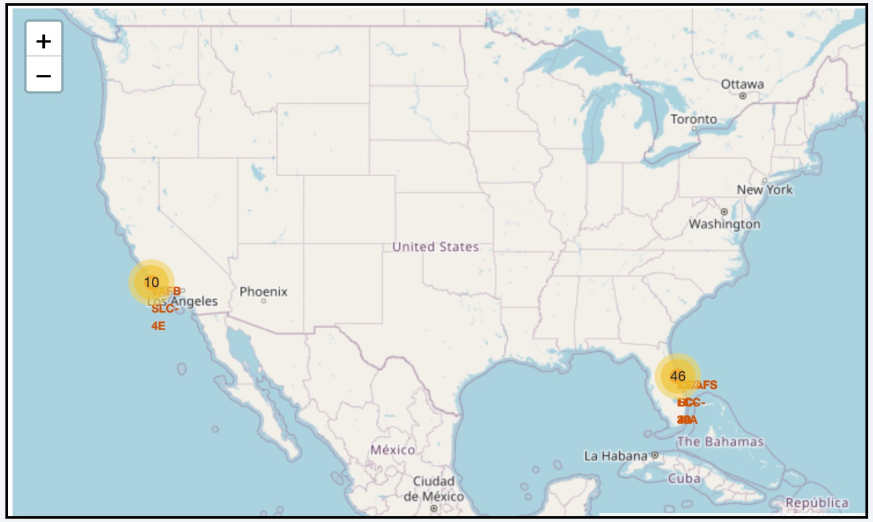


Figure 1: US Map showing the Falcon 9 Launch Sites.

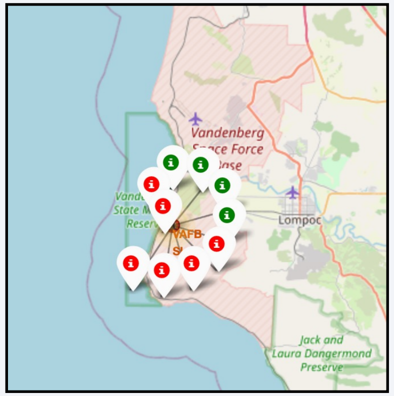


Figure 2: VAFB-SLC-4E Launch Site

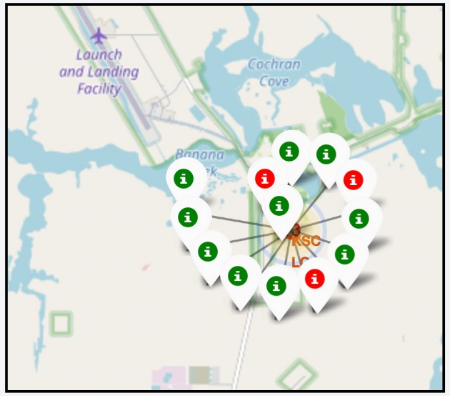


Figure 3: KSC LC-39A Launch Site

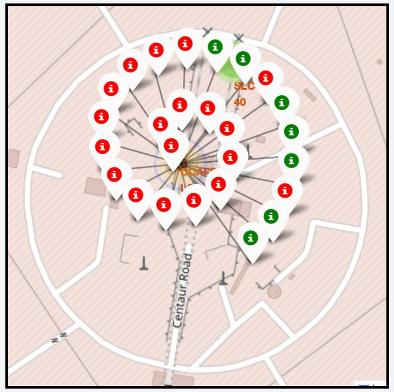


Figure 4: CCAFS LC-40 Launch Site

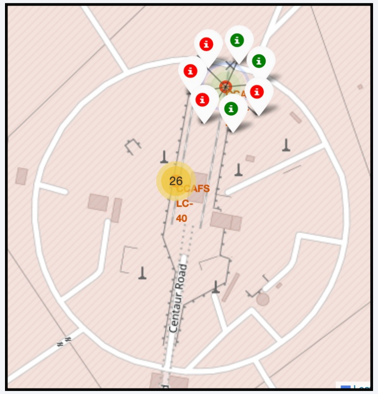


Figure 5: CCAFS SLC-40 Launch Site

- Figure 1 displays a US map with Falcon 9 launch sites, with numbers indicating the cumulative count of successful (green) and failed (red) launches.
- Figures 2, 3, 4, and 5 provide closer views of each launch site, showcasing success/fail markers (green for success, red for failure).
- Based on the site maps, the KSC LC-39A Launch Site stands out with the highest number of successful launches.

VAFB SLC-4E Launch Site Proximity Map

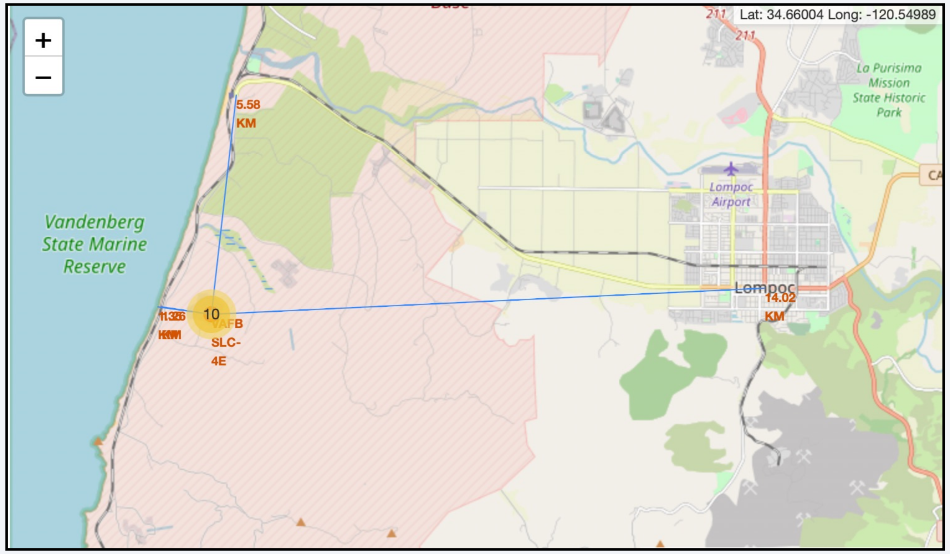


Figure 1: VAFB SLC-4E Distance from the city Lompoc and its airport, highway, railroad and coast.

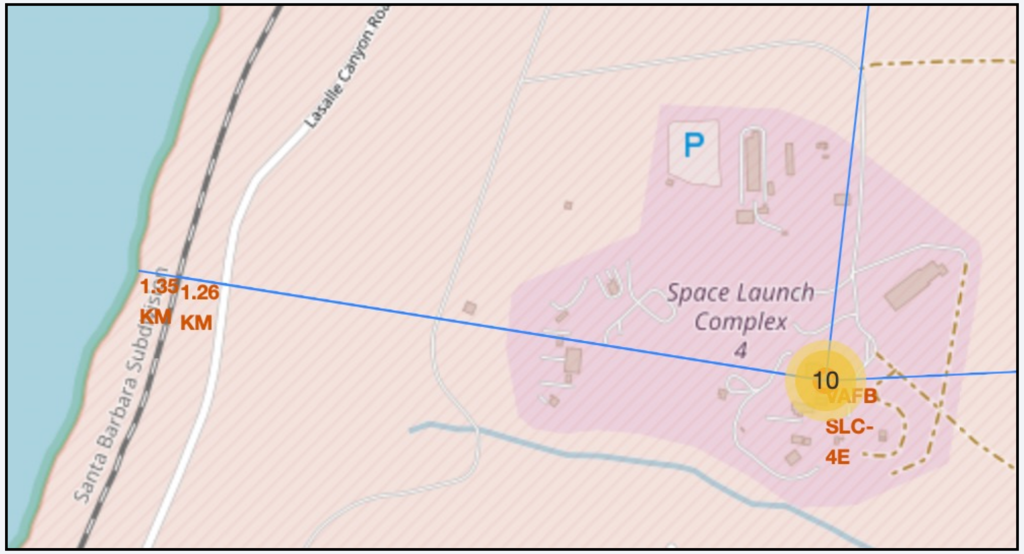


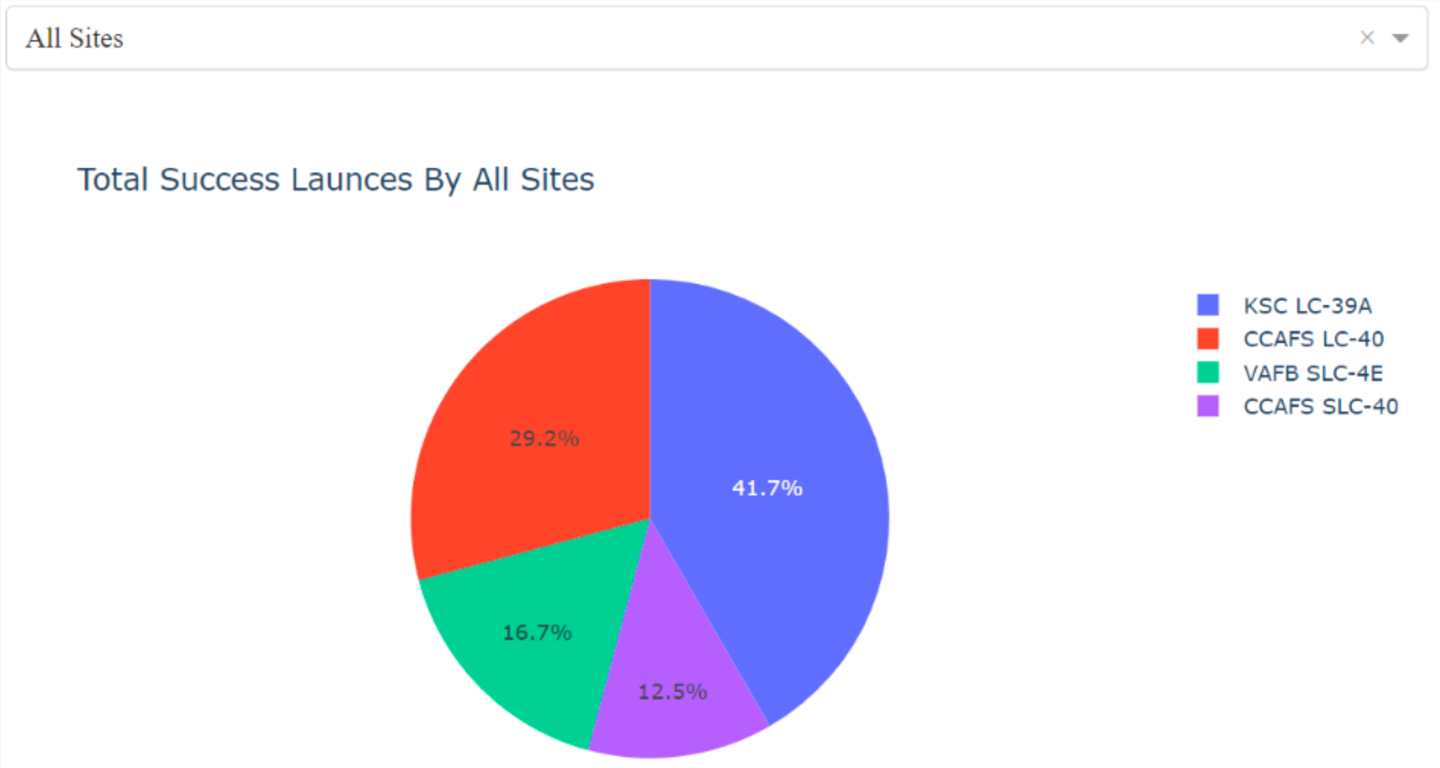
Figure 2: Zoom of VAFB SLC-4E Launch Site and distance to railroad and coastline.

- Are launch sites in close proximity to railways?
  - Yes, the VAFB SLC-4E launch site is in close proximity to railways, with a distance of 1.26 KM to the closest railway.
- Are launch sites in close proximity to highways?
  - Yes, the VAFB SLC-4E launch site is in close proximity to highways, with a distance of 5.58 KM to the closest highway.
- Are launch sites in close proximity to the coastline?
  - Yes, the VAFB SLC-4E launch site is in close proximity to the coastline, with a distance of 1.35 KM to the closest coastline.
- Do launch sites keep a certain distance away from cities?
  - Yes, the VAFB SLC-4E launch site is at a distance of 14.02 KM from the closest city.
- Brief Interpretation:
  - The close proximity to highways, railways and the coastline as well as an airport provide logistical advantages for transportation and infrastructure needs associated with launching rockets.
  - The 14 KM proximity to a small city provides some safety net when launches fail.

Section 4

Build a Dashboard with Plotly Dash

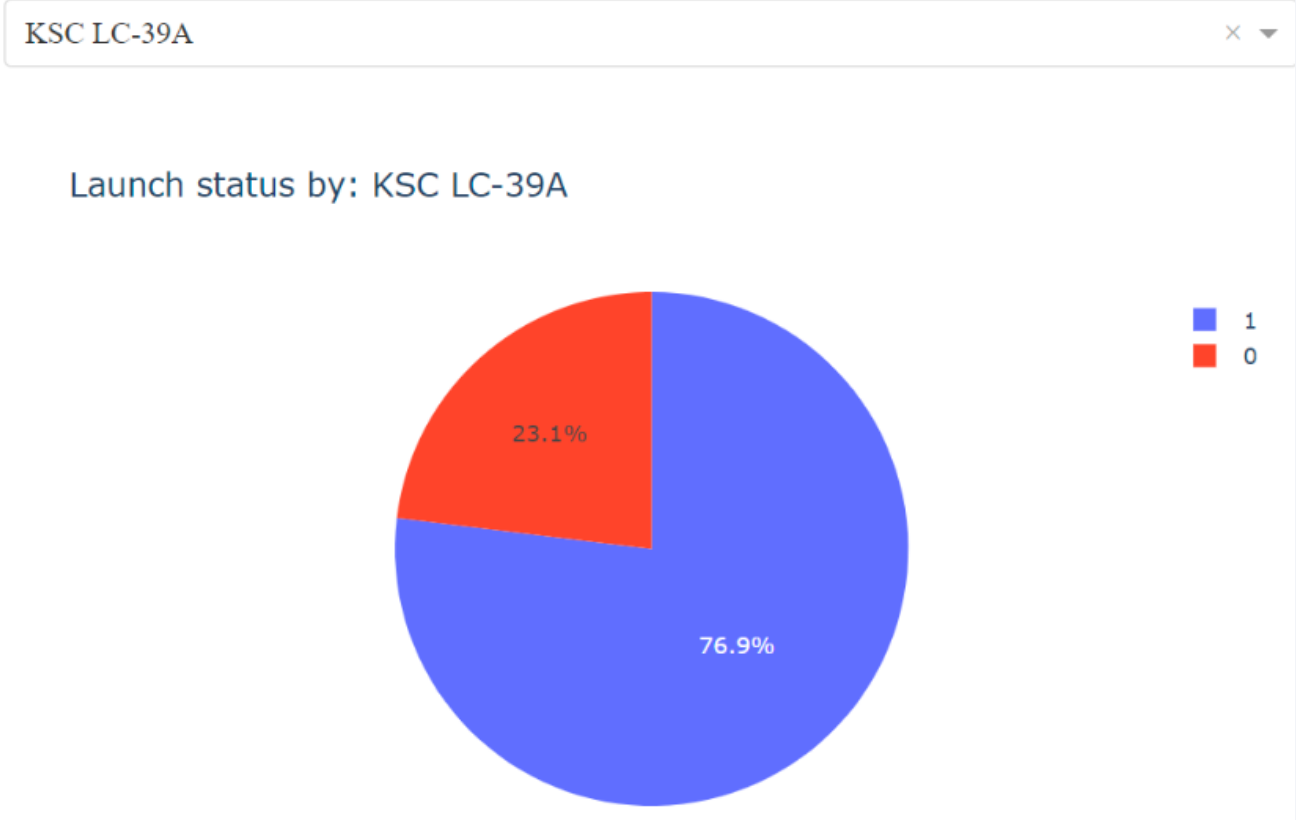
Pie Chart 1 - Proportion of Total Successful Launches by Launch Site



Key Points:

- The KSC LC-39A launch site had the highest share of successful launches, representing 41.7% of the total.
- The CCAFS SLC-40 launch site had a comparatively lower share, accounting for only 12.5% of successful launches.

Pie Chart 2 - Launch Site with the Highest Success Ratio



Key Points:

- The KSC LC-39A launch site had the highest share of successful launches, representing 41.7% of the total.
- 76.9% of launches succeeded
- 23.1% of launches failed

Scatter Plot -

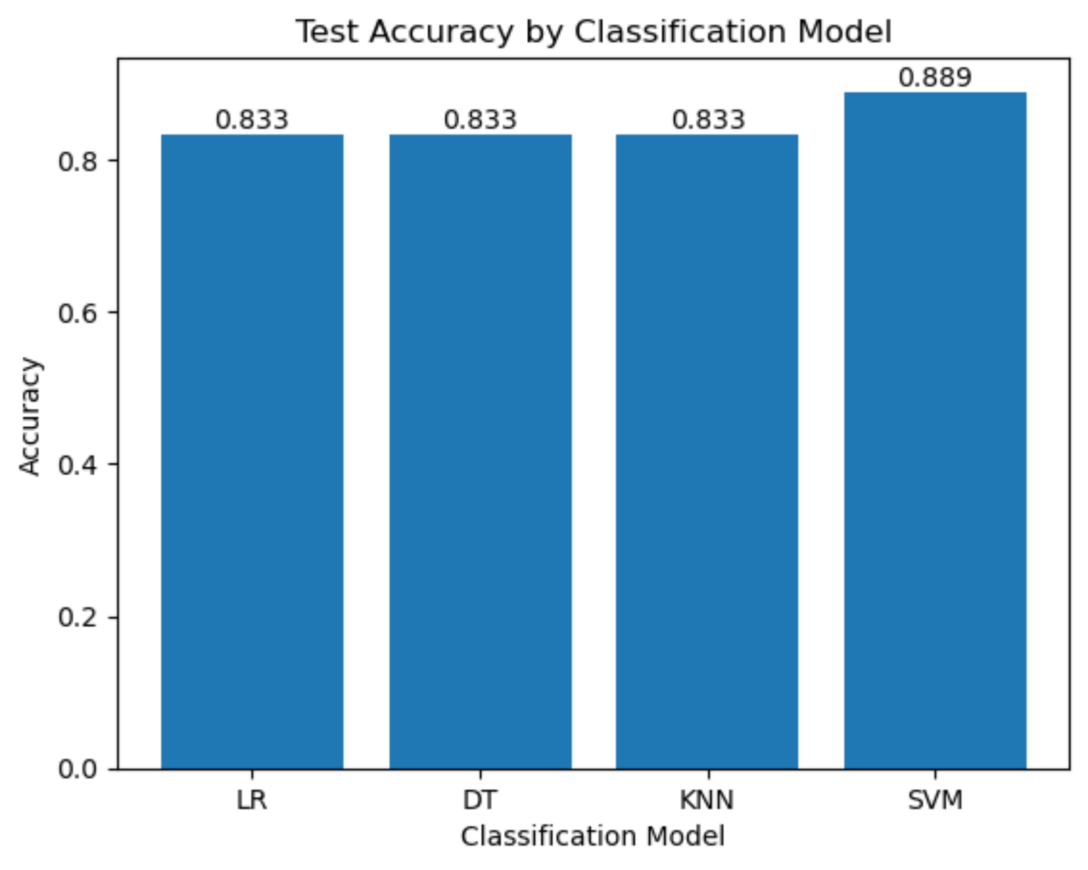
Key Points:

- Most successful launches occur with payloads ranging from 2000 to around 5500.
- The 'FT' booster version had the highest number of successful launches.
- The 'B4' booster category is the only category that achieved a successful launch with a payload greater than 6000.

Section 5

Predictive Analysis (Classification)

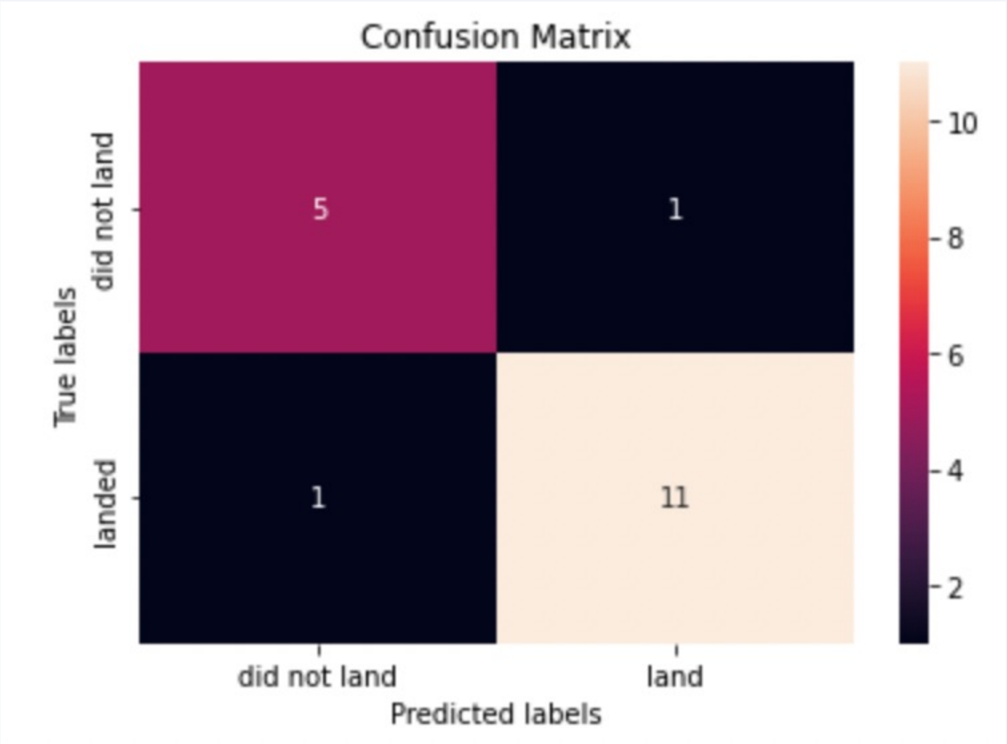
Classification Accuracy By Model on Test Data



Key Points:

- The SVM model had the highest classification accuracy with 88.9%
- The remaining models all had an accuracy of 83.3%
- The large accuracy suggests the models generalized beyond the training data to the test data.
- -> Considering the similarity in accuracy scores and identical test scores, a larger sample size might be necessary to refine and improve the models further.

Confusion Matrix



Key Points:

- The classifier made a total of 18 predictions.
- Out of the predictions, 11 were correctly classified as successful landings (True positive).
- Five predictions were correctly classified as unsuccessful landings (True negative).
- There was one false positive and false negative respectively.
- Overall, the classifier achieves an accuracy rate of around 89.9% ((TP + TN) / Total) and an error rate of 21.1% ((FP + FN) / Total).

Conclusions

- **Comprehensive Analysis:** The combination of data collection, data wrangling, exploratory data analysis, interactive visual analytics, and predictive analysis provided a comprehensive understanding of SpaceX's Falcon 9 launches.
- **Key Insights:** The analysis revealed key insights such as the increasing success rate of launches over time, the preference for certain launch sites for heavier payloads, and the higher success rates of specific orbit types.
- **Interactive Visualisations:** The use of interactive visualisations and dashboards significantly enhanced the ability to understand and interpret the data. These tools allowed for dynamic exploration of the data and provided insights that would be more difficult with static visualisations.
- **Predictive Modelling:** The predictive analysis demonstrated that the Support Vector Machine (SVM) model was the most accurate in predicting the outcome of future SpaceX Falcon 9 launches. This suggests that the model could be a valuable tool for planning future launches.
- **Future Work:** While the current analysis provided valuable insights, there is potential for further analysis. This could include exploring other predictive models, incorporating additional data, or focusing on specific aspects of the launches in more detail.
- **Impact:** The insights gained from this analysis could inform future launch strategies, contribute to improvements in launch success rates, and ultimately support SpaceX's mission to advance space exploration.

Thank you!