IBM Developer
SKILLS NETWORK

Catching Up in the Space Race: The SpaceX Model

Alin Airinei

01/14/2022

Outline

- Abstract
- Introduction
- Methodology
- Results
- Conclusion
- References
- Appendix

Abstract

- **Summary of methodologies**
  - In this study we analyze the launching records of the most used rocket in US space exploration history: the SpaceX Falcon 9
  - Starting with the booster version v1.2 (Blocks 1 to 5) the Falcon 9's first stage is partially reusable[1]
  - According to Elon Musk the first stage accounts for 70% of the launching cost[2]
  - The purpose of this study is to provide insights that help model a low cost, reliable, and competitive service for payload delivery to orbit using reusable rockets.
- **Summary of all results**
  - Using Exploratory Data Analysis on the Falcon 9 data we extracted valuable insights into the elements that set the standard for a reliable low cost space launching service.
  - Using Machine Learning we built and trained a model that predicts whether or not a Falcon 9 first stage booster will land successfully.

Introduction

- The Hawthorne based Space Exploration Technologies Corp. (SpaceX) sits at the forefront of the currently booming space industry by offering on-schedule payload deliveries to orbit through its Falcon 9 rockets[3]
- It’s reliable service depends in part on it’s first stage rocket reusability which, in addition to on-time service, also significantly lowers the cost of launching payloads into space (the current record is of 27 days between subsequent launches of the same booster after refurbishment)[4]
- Bank of America, tracking 23 aerospace companies, projects a space economy that will hit $2.7 trillion by 2045[5]
- Hiccups in development and high costs keep competition’s customers on edge as time sensitive launching schedules for their payloads and astronauts are met with frustrating delays[6]
- Many of the major players in the current space race are developing their own reusable launchers(see Blue Origin, Boeing, etc.).[7] By using the tools of Data Science we try to map some of the key aspects of the business model that a potential competitor might consider in the endeavor of matching SpaceX’s lead in the industry.

Section 1

Methodology

Methodology

Executive Summary

- *Data collection methodology:*
  - *The data for this study was collected using both the SpaceX REST API and by web scraping Falcon 9 records from Wikipedia.*
- *Data Processing*
  - *The data was prepared by processing the missing values, selecting features, creating labels and encoding.*
- *Exploratory data analysis (EDA) using visualization and SQL*
  - *Insights were drawn from data by using graphs and performing SQL queries*
- *Interactive visual analytics using Folium and Plotly Dash*
  - *To provide easy access to the results, interactive maps and dashboards were created using folium and Plotly Dash.*
- *Predictive analysis using classification models*
  - *Different classification models were tested and the results compared to provide predictive insights into the current development of reusable rockets.*

Data Collection

- To collect the information needed for this study two methods were used:

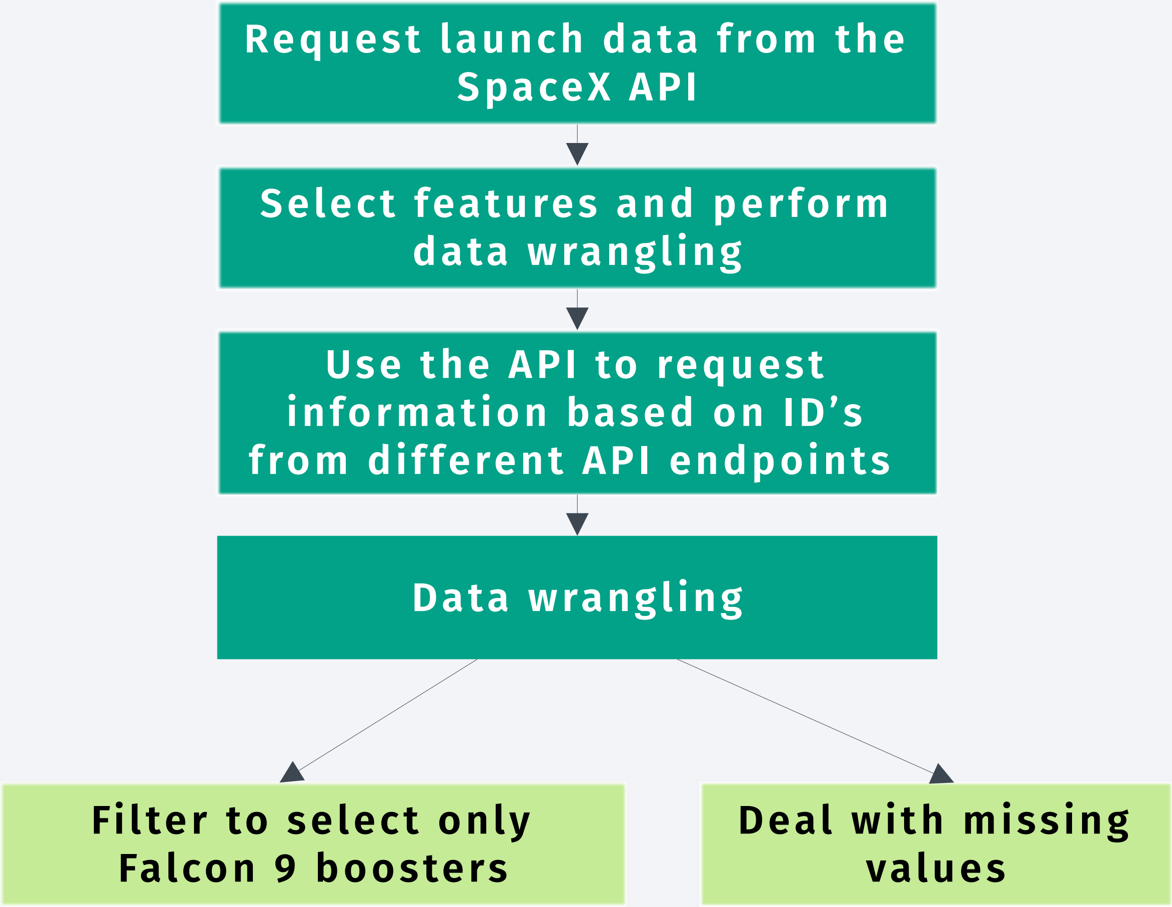
  1) Using get requests to receive records from the SpaceX REST API

  2) Pulling the information from the Falcon 9 flights Wiki page using the BeautifulSoup library

  Next, the two methods will be discussed in detail.

Data Collection – SpaceX API

- Using the SpaceX REST API data related to past launches was collected using get requests.
- After selecting the features, the final data set was made complete by calling different API endpoints to retrieve booster, payload, and launch site information.
- Data points irrelevant to the current study were removed and missing payload values were replaced by their means.



The GitHub link of the API Data Collection notebook:
https://github.com/Al-1n/IBM_SpaceX_Capstone/blob/master/Data%20Collection%20API.ipynb

Data Collection - Scraping

- After using the requests.get() method to pull the text of the HTML from Wikipedia, the BeautifulSoup library with the html5lib parser were used to create a BeautifulSoup object.
- Using the newly created object the find_all() method was applied to find all the table type elements.
- After identifying the table of interest the table was pulled by index and a predefined function was used to extract the column names from the header elements.
- Adding the column names as keys to a launch dictionary, the corresponding lists of values were subsequently added by iterating through each row.
- The final result was converted into a pandas DataFrame and exported to a CSV file.

Retrieve the Falcon 9 Wiki page by using requests.get()

↓

Find and retrieve the table of interest

↓

Extract the column names from the table header elements

↓

Populate the launch dictionary with the corresponding values

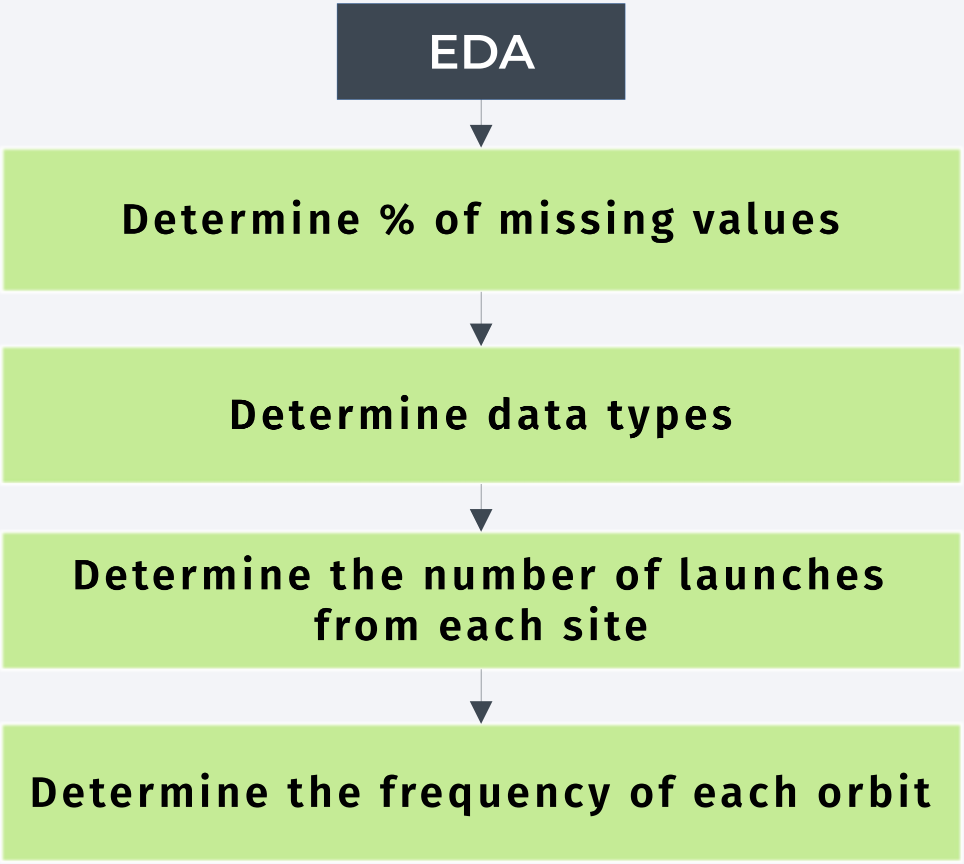
↓

Convert the launch dictionary into a pandas DataFrame
&
Export the DataFrame to a CSV file

The Data Web scraping notebook can be found at:
https://github.com/Al-1n/IBM_SpaceX_Capstone/blob/master/Data%20Collection%20with%20Web%20Scraping%20lab.ipynb

Data Wrangling I

- The Data Wrangling process involves two objectives:
  1) Perform EDA to determine how the data needs to be transformed and manipulated.
  2) Use the insights of EDA to determine the classification labels.
- By determining the percentage of missing values for each column we determine if the results will be significantly affected by the missing items. This will prompt whether or not further action is needed in dealing with the missing values.
- Determining the data types will establish the need for encoding.
- Weighing the results according to the number of launches from each site and the frequency of each orbit will provide more general results.

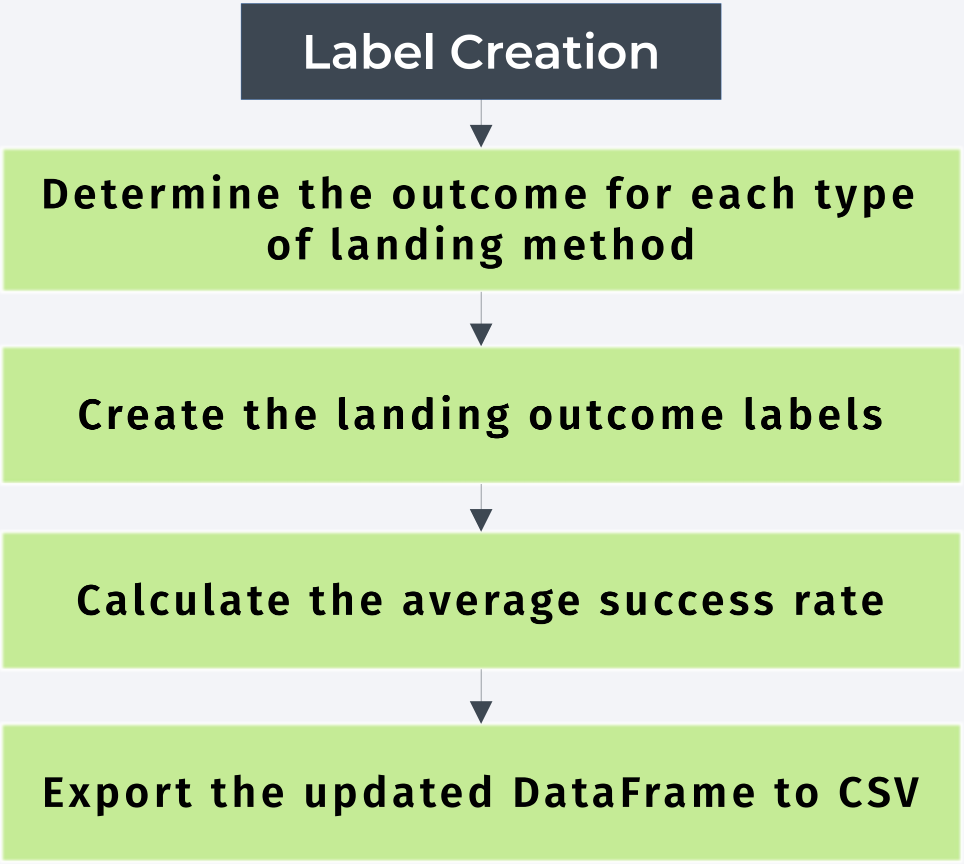


Data Wrangling II

- We first extract the number of outcomes to retrieve the labels for each landing location.
- Using the location labels from the previous step a set of bad outcomes was created which will be used to filter and split the the total number of outcomes into two categories:

  1) Class 0 for failed landings.
  2) Class 1 for successful landings

- (**Note: at this step “True Ocean” should have been included in the bad_outcomes set or removed, as this type of landings are only relevant for testing purposes and do not yield reusable boosters. For the consistency of the assignment the results were kept as they are.**)
- The binary encoding of the classification labels allows for the calculation of the average success rate.



The GitHub URL of the completed data wrangling notebook is:
https://github.com/Al-1n/IBM_SpaceX_Capstone/blob/master/labs-jupyter-spacex-Data%20wrangling.ipynb

EDA with Data Visualization I

Following is a list of charts that were plotted for the visual analysis of the data:

1) Flight number vs launch site
   - The purpose of this graph is to help identify variations in location with time and any possible correlation to the landing success rate
2) Payload mass vs launch site
   - This chart helps determine if the payload range is restricted by location
3) Landing success rate vs orbit type
   - This will help establish if certain destinations are more likely to yield successful landings.
4) Flight number vs orbit type
   - The graph will help determine if the stage of development of the Falcon 9 booster is related to the destination orbits and the landing success rate.
   - This could also indicate if the industry demand for a certain orbit type has changed with time.

EDA with Data Visualization II

5) Payload mass vs orbit type
   - The graph will help visualize which orbit types are more successful with heavy payloads.
6) Launch success rate – yearly trend
   - The purpose of this plot is to help visualize whether the success rate has been increasing or decreasing with time.

- The GitHub URL of the completed EDA with data visualization notebook: https://github.com/Al-1n/IBM_SpaceX_Capstone/blob/master/jupyter-labs-eda-dataviz.ipynb

EDA with SQL I

Following is a list of the SQL queries applied to perform EDA for this study:

- Requesting the display of unique site names
- Retrieve the records of Cape Canaveral locations.
- Calculate the total payload carried for NASA CRS missions.
- Determine the average payload mass carried by F9 v1.1 boosters.
- List the date of the first successful ground pad landing.
- Retrieve the booster names that carried between 4000 and 6000 Kg, and landed on drone ships.
- Calculate the total number of mission outcomes

EDA with SQL II

SQL query list continued:

- List the names of the booster versions which carried the maximum payload mass.
- List the failed drone ship landing outcomes and launch sites for the year 2015.
- Rank the count of landing outcomes for drone ship and ground pad between 6/4/2010 and 3/20/2017.

- GitHub URL of the EDA with SQL notebook: https://github.com/Al-1n/IBM_SpaceX_Capstone/blob/master/jupyter-labs-eda-sql-coursera.ipynb

Build an Interactive Map with Folium

The following graphic elements were used to allow users to interact with the launching site maps:

- Circles
  - To mark the location of each launching center, circles with popup location names were added to the map.
- Markers
  - Associated with each circled location, a marker containing an icon element for displaying the static name of the site was added to the map at each location.
- Lines
  - To display distances, lines with endpoint markers displaying the distance have been used.

- GitHub URL of the interactive map with folium notebook: https://github.com/Al-1n/IBM_SpaceX_Capstone/blob/master/lab_jupyter_launch_site_location.ipynb

Build a Dashboard with Plotly Dash

- The interactive dashboard created for this study includes two types of charts:

1) **Success-failure pie-charts** which allow the comparison of both the ratio of successful landings for all the launching sites as well as the individual rate of success vs failure for each individual site. The navigation between each chart is made possible via a drop-down menu. Also pop-ups on mouse hover offer information about the individual elements of the charts.

2) **Success-payload scatter plots** allow the user to explore the connection between features such as booster version, payload and success rate. The payload range is controlled via a slider and on mouse hover each data point offers information on booster version, payload mass, and landing outcome via pop-ups.

- GitHub URL of the Plotly Dash Lab: https://github.com/Al-1n/IBM_SpaceX_Capstone/blob/master/spacex_dash_app.py

Predictive Analysis (Classification)

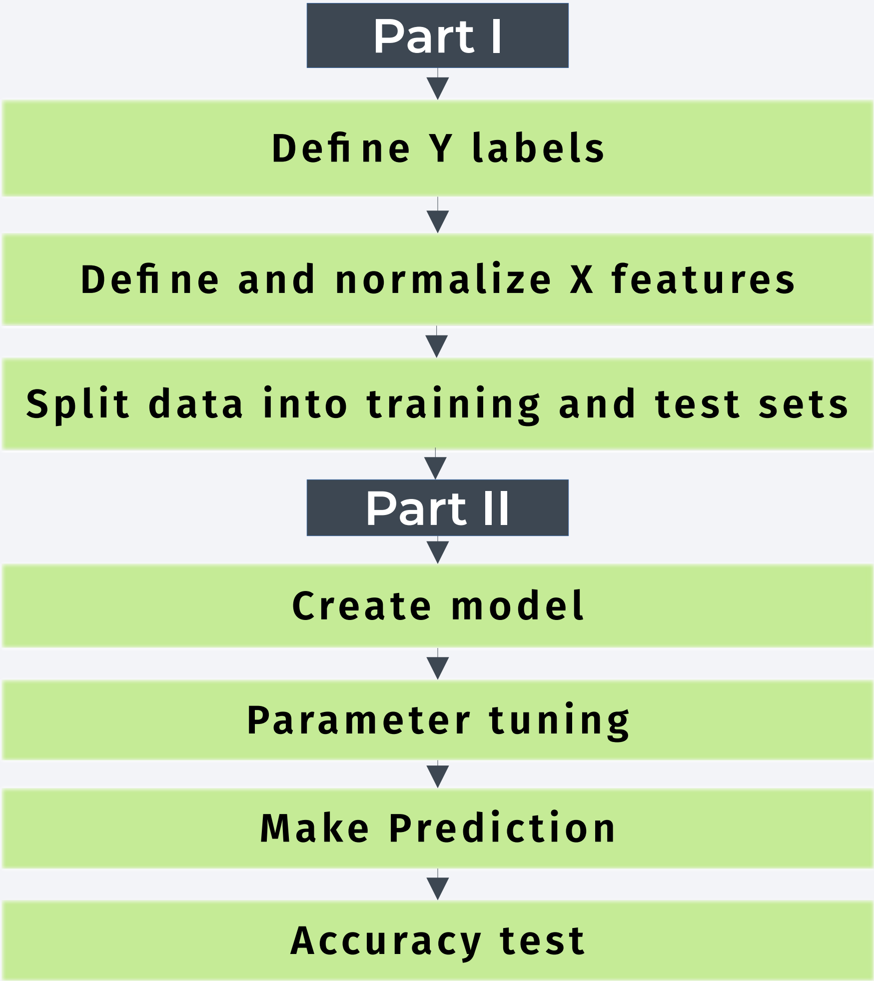
The classification process involves two parts:

Part I: Creating labels and splitting the data

- load the outcome data and assign labels to the Y variable
- load, normalize, and assign the features to the X variable
- split the data into training and test sets

Part II: Building and evaluating the models

- create a model object and a parameter dictionary
- optimize the parameters by using the GridSearchCV tuning function
- make a prediction using the test data
- estimate the prediction accuracy using the accuracy_score function
- plot the confusion matrix



- GitHub URL of the Machine Learning Notebook: https://github.com/Al-1n/IBM_SpaceX_Capstone/blob/master/SpaceX_Machine%20Learning%20Prediction_Part_5.ipynb

Results

Next, the results are listed according to the following categories:

- **Exploratory data analysis results**
  - A discussion of the graphs and the results of the SQL database queries performed on the Falcon 9 data
- **Interactive analytics demo in screenshots**
  - A demonstration of the interactive tools generated during this project
- **Predictive analysis results**
  - An evaluation of the results and the accuracy of the predictive models employed in our study

Section 2

Insights drawn from EDA

Flight Number vs. Launch Site

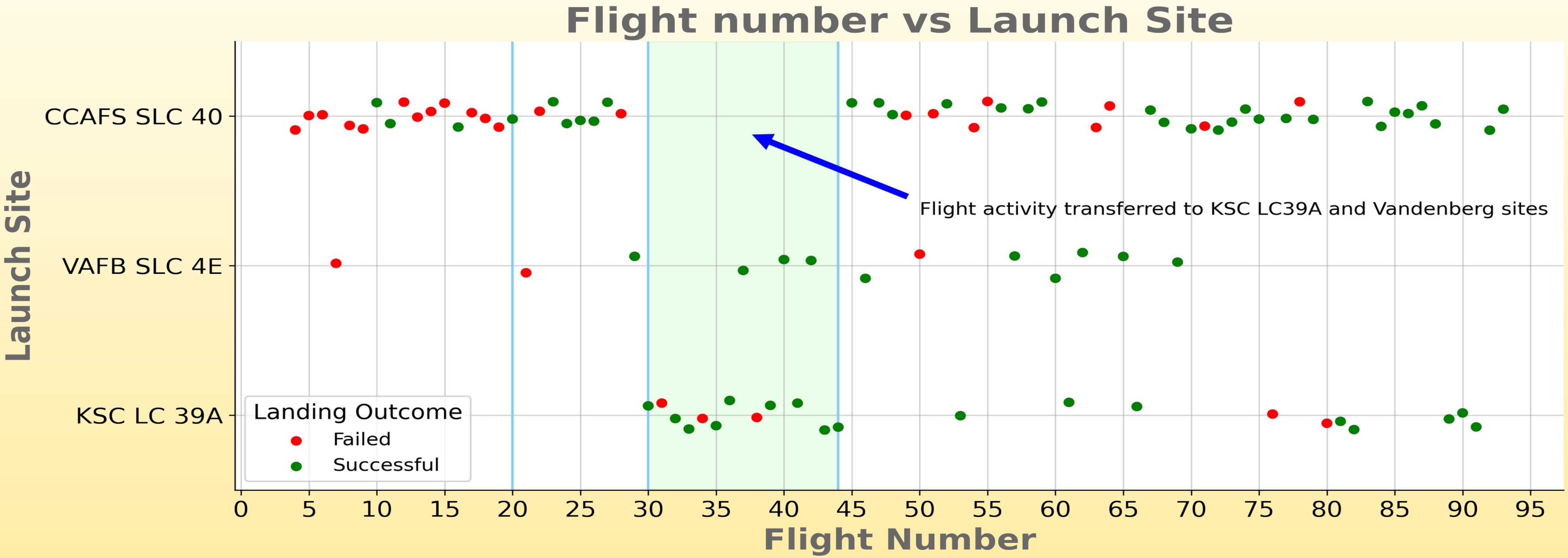


Fig 1. Scatter plot showing the distribution of flights for the three main SpaceX launch sites. The markers are color-coded to indicate the landing outcome.

Flight Number vs. Launch Site (analysis)

Note: In the figure above, ocean landings (which are not drone ship landings) are marked as Class 1, i.e. successful, as per the assignment . That might not be accurate as SpaceX and the literature in general do not refer to controlled soft ocean landings as successful other than for test and proof of concept purposes. In ocean landings rockets tip over and usually get destroyed. Next we will list the most important findings.

- All initial flights and test flights were launched from Cape Canaveral SLC-40 except for one, unsuccessful mission, launched from VAFB SLC-4E. There are no initial launches from KSC LC-39A.
- Therefore, overall individual success rates are not very relevant in determining the dependence of the landing outcome on location.
- After the historic Flight 20, which marks the first successful landing on a land pad, all three sites have comparable success rates proportional to the number of flights from each site.
- From the graph above we can also see that the bulk of the SpaceX launches are located at Florida sites (Cape Canaveral in particular), in relative proximity to each other. All the Launch sites are on lease from NASA or the US Air Force.
- There is a complete shift in density between flights 30 and 44 between CCAFB SLC-40 and KSC LC-39A. This might have been caused by a static test explosion at CCAFB SLC-40.[8]

Payload vs. Launch Site

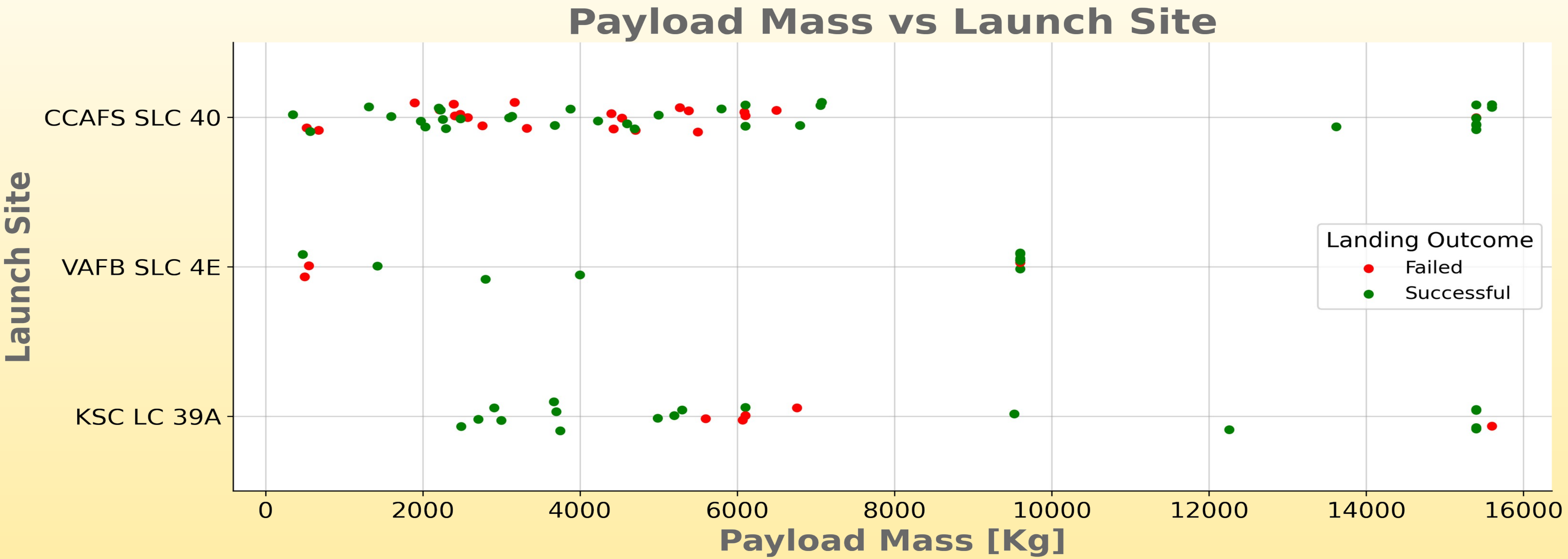


Fig 2. Scatter plot showing the distribution of payload per mission for the three main SpaceX launch sites. The markers are color-coded to indicate the landing outcome.

Payload vs. Launch Site (analysis)

- From the Payload vs Launch Site plot above we can see that both Cape Canaveral and Kennedy Space Complex share a similar payload range.

- The Vandenberg location does not launch payloads larger than 10000 Kg and the Kennedy complex did not launch payloads that were less than 2000 Kg.

Success Rate vs. Orbit Type

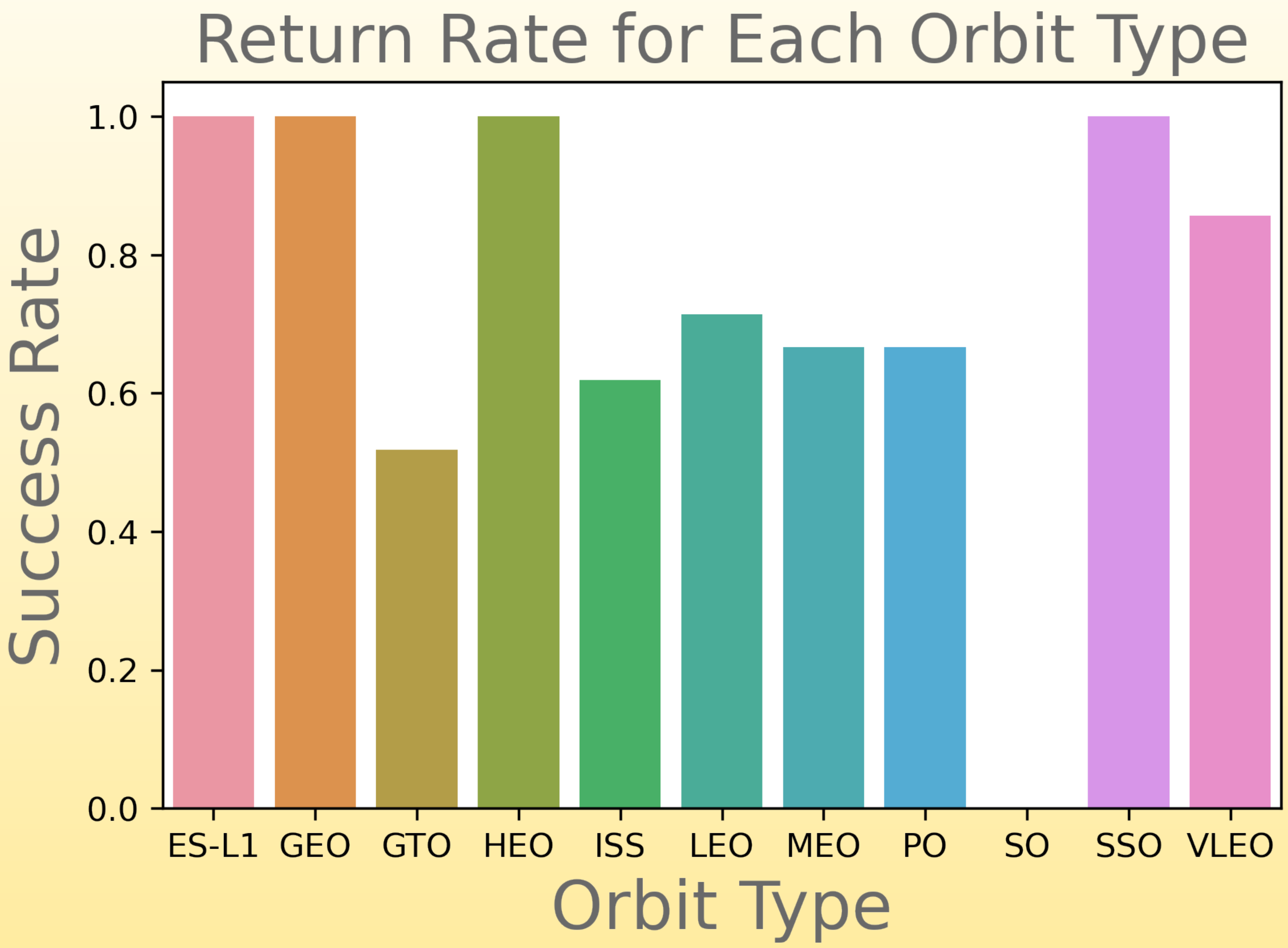


Fig 3. Bar plot showing the mission success rate for each orbit type.

Success Rate vs. Orbit Type (analysis)

- From the bar chart in Fig. 3 it can be deduced that high Earth orbits (HEO, L1, GEO) have the highest success rates while most low orbits have less positive results with the exception of SSO.

- This results can be deceiving as it will become more visible in Fig. 5 bellow.

- To list a few of the reasons this results are not suggestive of any causal connection between success rate and orbit type:
  - Firstly, the number of flights to high Earth orbits is much smaller than the flights to low Earth orbits, and thus, there are not sufficient events to provide conclusive data.
  - The two orbits with the highest number of flights (see Fig. 5 bellow) actually have the lowest success rate. It merely shows that the more you fly, the more likely it is to have failures.
  - For high orbits the specifications for the first stage (currently the only reusable part that lands itself on a pad) do not change as long as the payload and trajectory are within its capabilities. It is the job of the second stage to actually deliver the payload to it's final orbit. Therefore, no real change in circumstances that might affect the likelihood of either failure or success.[9]

- The most relevant aspect of this graph is <u>the high success rate for the SSO orbit missions</u>. This might be related to the specific trajectory of a polar launch. The effect of the polar trajectory on landing success rate indicates the need for further investigation.[10]

Flight Number vs. Orbit Type

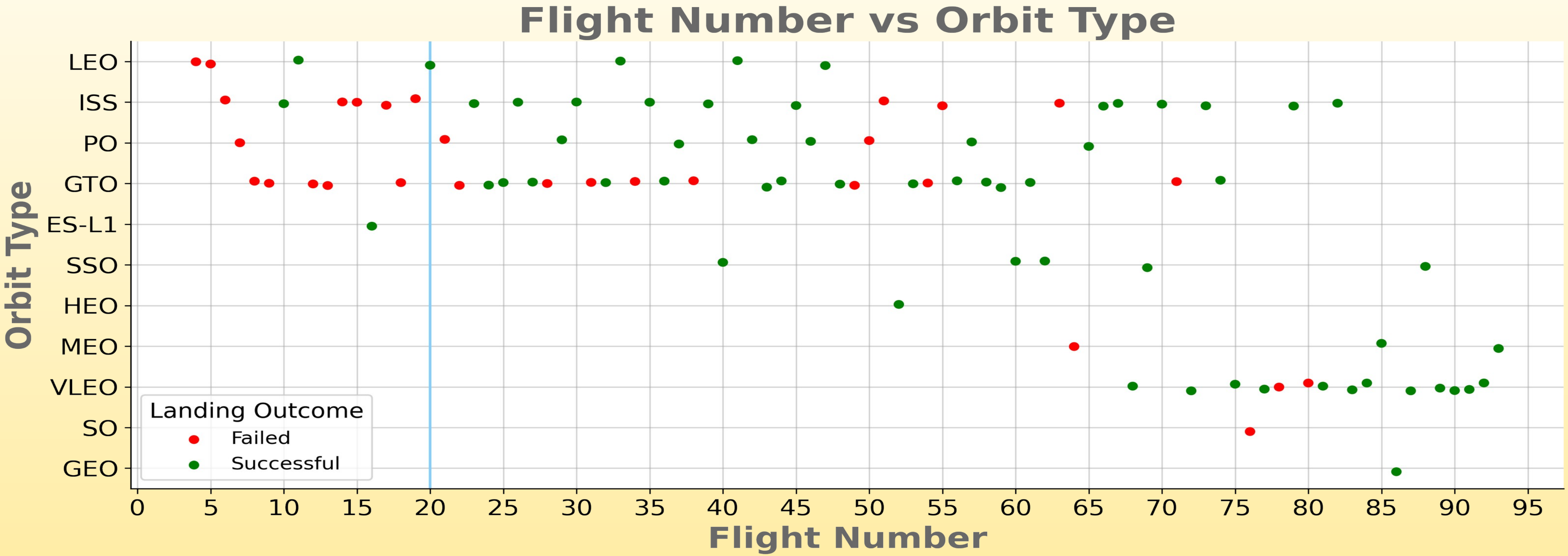


Fig 4. Scatter plot showing the relation between flight number and orbit type. The markers are color-coded to indicate the landing outcome.

Flight Number vs. Orbit Type (analysis)

- From the graph above we can see that the most frequent orbit types are: LEO, ISS, PO, and GTO.
- There was only one flight to ES-L1.
- Flight 40 marks the first ascension to a sun-synchronous orbit (SSO). This might have been preceded by engine improvements on the FT boosters.[14]
- Flight 47 marks the last low Earth orbit (LEO) mission and Flight 65 the last polar LEO (PO).
- The emergence of the B5 boosters as the new standard boosters seems to correlate with a transition in focus on very low orbits (VLEO) and ISS LEO (ISS), as well as first attempts at SO and GEO. This is shown on the graph between Flights 65 and 93. This transition is preceded by first attempts at SSO, HEO, and MEO between Flights 40 and 65 which mark the era of the B4 boosters.

Payload vs. Orbit Type

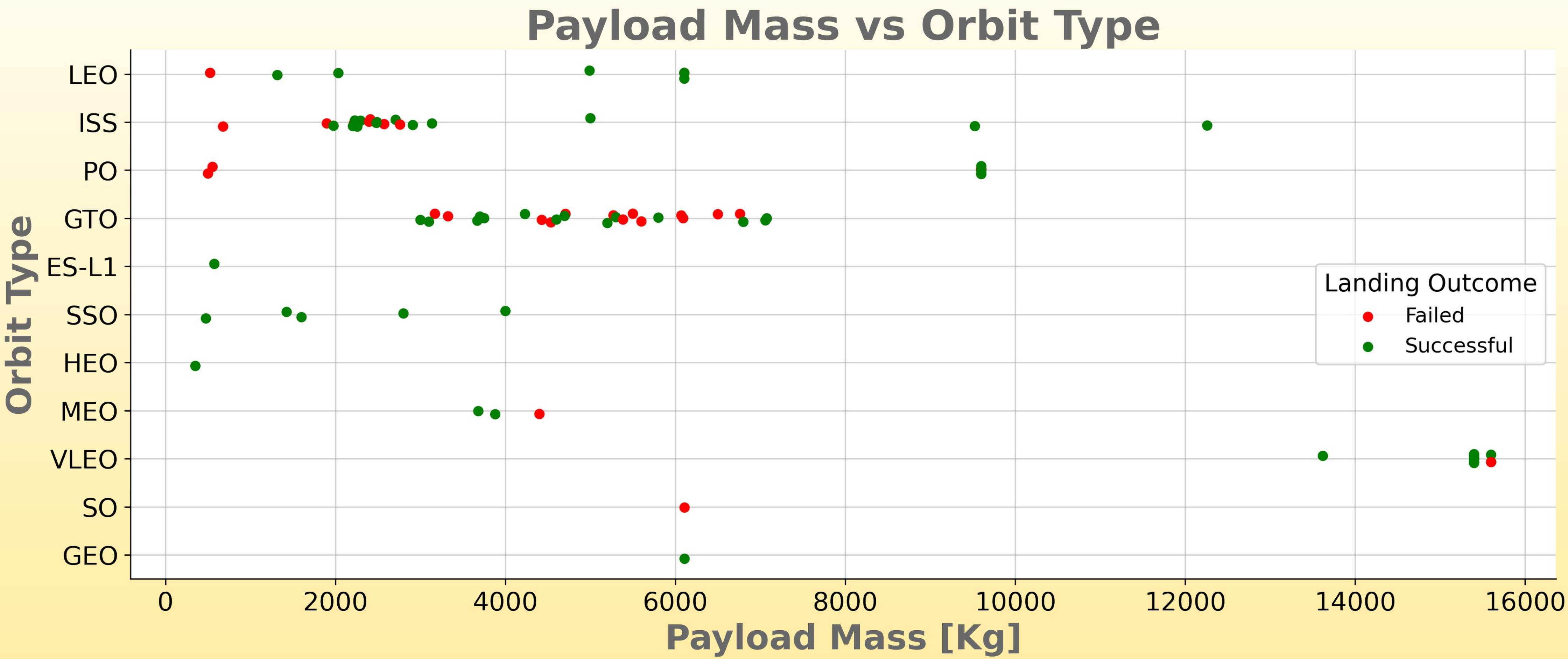


Fig 5. Scatter plot showing the relation between flight payload, orbit type and landing outcome.

Payload vs. Orbit Type (analysis)

- From the graph in Fig. 5 we observe that satellites destined for a geostationary orbit (GEO) are mostly delivered to the transfer orbit (GTO) and have a payload ranging from 3000 to 7000Kg with an even success/failure ratio.
- High orbits are failure free but have an extremely small number of missions with one mission for each of HEO, GEO and Lagrange (L1). HEO and L1 have very small payloads with GEO having the largest payload at just above 6000Kg.
- Low Earth orbits (LEO, ISS, VLEO, SSO, PO) have the highest rate of success for heavy payloads with the heaviest payloads going to VLEO.

Launch Success Yearly Trend

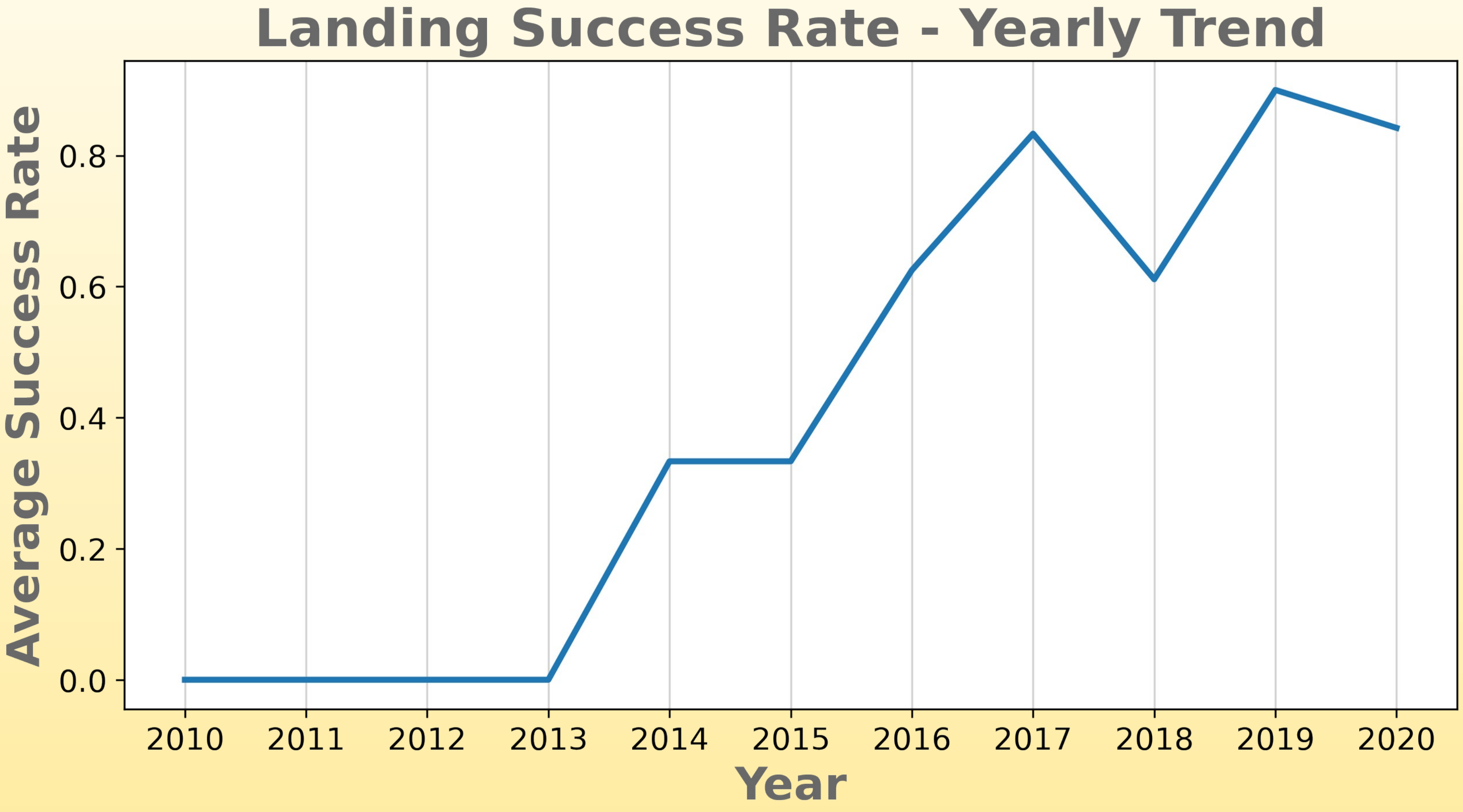


Fig 6. Trend line showing the average landing success rate

Landing Success Yearly Trend (analysis)

- The graph in Fig. 6 shows the yearly increase in landing success rate with no landings prior to 2013
- The increase between the years 2013 and 2015 is due to the test missions (no true landings)[14]
- After 2015 the success rate has been virtually on a continuous rise without dropping bellow 60%, and with a maximum rate of 84%.

All Launch Site Names

- The SQL query to extract the unique site name returns four locations:

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

- The initialisms stand for:[11]
  - Cape Canaveral Air Force Station Launch Complex 40
  - Cape Canaveral Air Force Station Space Launch Complex 40
  - Kennedy Space Center Launch Complex 39A
  - Vandenberg Air Force Base Space Launch Complex 4E
- The initial division between launching sites was meant to serve different markets (Air Force/Military from Cape Canaveral, NASA CRS from Kennedy, Polar from Vandenberg, and an additional Texas location for commercial launches)[12]
- The first two sites in the list are one and the same location as LC-40 has been renamed to SLC-40[13]
- The initial division has changed, nowadays SpaceX using all three sites interchangeably, including restoring the abandoned polar launch capability to Cape Canaveral with a new trajectory and new safety measures.[14]

Launch Site Names Begin with 'CCA'

- Since the Cape Canaveral LC-40 and SLC-40 refer to the same location, an SQL query meant to retrieve the records of the sites that include the 'CCA' string was made to include both denominations
- The first five records are shown below:

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

- From the results we can see that the first flights from Cape Canaveral were demo and qualification missions for the NASA CRS program aimed at delivering supplies to the ISS.
- We notice that SpaceX considered and studied booster recovery from the very beginning by the two failed attempts by parachute at the top of the table.
- The first actual supply deliveries to ISS are also listed as the CRS-1 and CRS-2 missions.

Total Payload Mass

- Next we queried the database for the total payload carried to ISS for the NASA Commercial Resupply Services (NASA CRS). The result is shown below:

NASA(CRS) total payload[kg]
45596

- The result offers a good confirmation of the Falcon 9's continuous success as NASA is one of the most important space mission and contract-awarding entities for both SpaceX and its competitors.[15]

Average Payload Mass by F9 v1.1

- Using SQL to calculate the average payload mass carried by the F9 v1.1 boosters we obtain the following result:

AVG Mass Carried By F9 v1.1 Boosters
2534

- The result suggests an average mass that is significantly bellow the rockets specifications for both LEO or GTO[16]
- The reason for the low average can be traced to under-loading (in particular the LEO missions have been set on course with loads way bellow the rocket's specifications; that is not the case for the GTO missions though).[14]

First Successful Ground Landing Date

- Below is the result of the SQL query made to retrieve the date of the first successful ground pad landing:

First Successful Landing Date
2015-12-22

- The date of the first successful landing is of particular historical importance for SpaceX and for the development of the Falcon 9. It marks the first flight of the F9 v1.2 – Full Thrust booster (at it's current Block 5 generation), and the beginning of the reusable rocket industry.[17]

Successful Drone Ship Landing with Payload between 4000 and 6000

- The list of names of boosters which have successfully landed on drone ship and had payload mass greater than 4000 but less than 6000 is given below:

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

- The first booster listed is the first to land after a GTO mission. The significance comes from the higher difficulty in recovery due to the higher speed needed for GTO missions [18]
- The second booster in the list is the first to execute a single-engine landing burn[19]
- The third listing marks the first payload to fly on a re-used booster and the first recovery of the payload fairing[20]
- The last item marks the first triple reuse and recovery of the same booster.[21]

Total Number of Successful and Failed Mission Outcomes

- Using SQL to calculate the total landing attempts for the sample data studied returned the following result:

Total Landing Attempts
71

- Comparing the result with the total flights in our sample data we observe that for 70% of the Falcon 9 missions there was a landing attempt.

Boosters that Carried the Maximum Payload

- The list of boosters that carried the maximum payload is displayed below:

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

- It can easily be spotted that the maximum payload was carried by boosters from the Block 5 family.
- The Block 5 boosters have an 8% increase in thrust from Block 4 for the first stage and 5% increase for the second stage. It also has twice the payload capacity of v1.0.[22]

2015 Launch Records

- Listed below are the failed drone ship landings for the 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

- The two records presented here represent the first drone ship landing attempts for SpaceX and Falcon 9. Both failed due to technical difficulties. [14]

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

- Following is the ranking of successful landing outcomes for drone ship and ground pad landings between the dates 2010-06-04 and 2017-03-20:

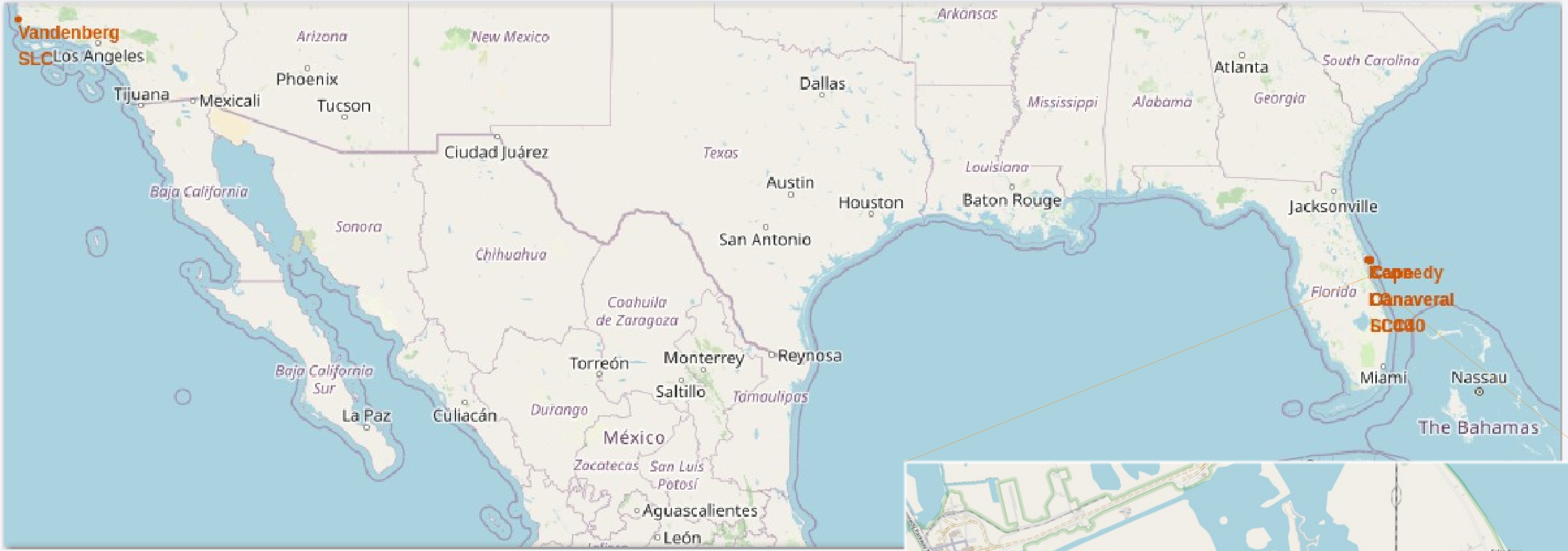
landing__outcome	outcome_count
Success (drone ship)	5
Success (ground pad)	3

- The result does not suggest a higher likelihood of success for drone ship landings as the result could be determined by different factors such as the difficulty or delay in obtaining ground landing authorizations, experimental permits etc.[23]

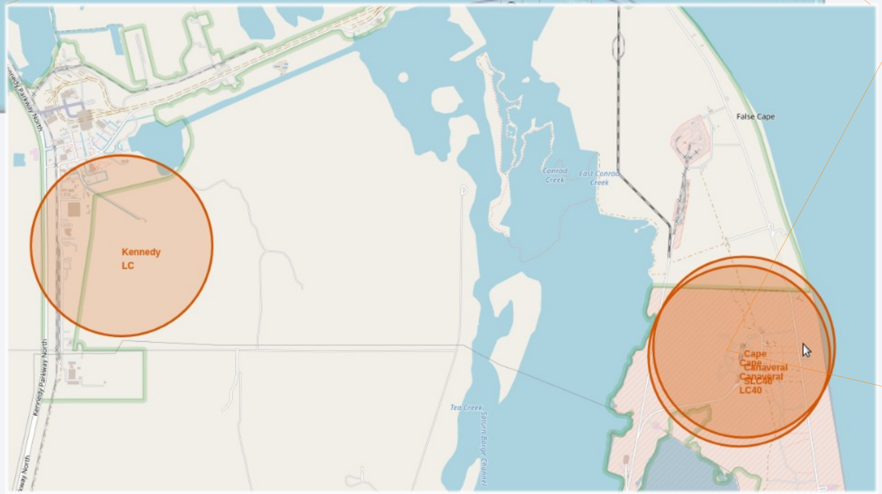
Section 4

Proximity Analysis

Global View of the Falcon 9 Launching Sites I



Global View of the SpaceX launching sites showing both East coast and West coast locations.



Expanded view of the East coast locations.

Expanded view of the Cape Canaveral sites.

Global View of the Falcon 9 Launching Sites II

Detail showing the markers for the Florida sites.

Detail showing the marker for the Vandenberg SLC.

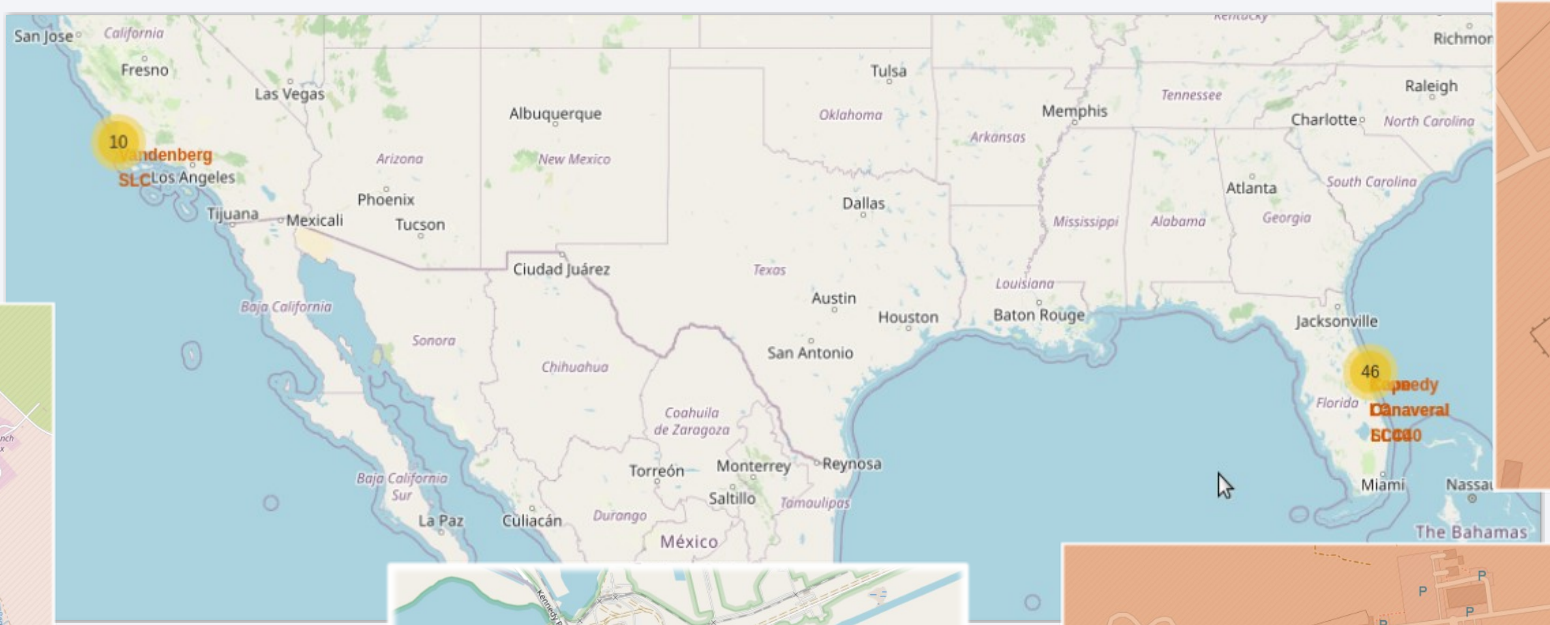
Note: to increase legibility the map above was tweaked to show only SLC-40 for the Cape Canaveral launching pads as both LC-40 and SLC-40 virtually share the same location.

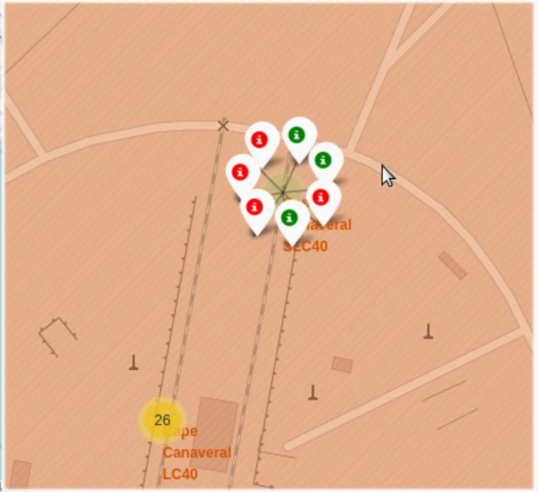
Global View of the Falcon 9 Launch Sites (analysis)

- From the map above it can be seen that all the launch sites are located at some of the lowest latitudes of the United States.
- The reason for their location in proximity to the Equator is due to the fact that at those latitudes the Earth acts as a slingshot as a result of its equatorial bulge. A location at the Equator rotates faster with respect to the Earth's axis than a location closer to the poles.[24]
- The slingshot effect allows for extra savings in fuel and less booster power needed.
- All sites are also in close proximity to the water. This allows for convenient splash-landing and capsule recovery, as well as drone ship landing.
- On the global map the sites in Florida appear superimposed due to their proximity. The blown up details show the individual sites with the Cape Canaveral sites virtually sharing the same location.

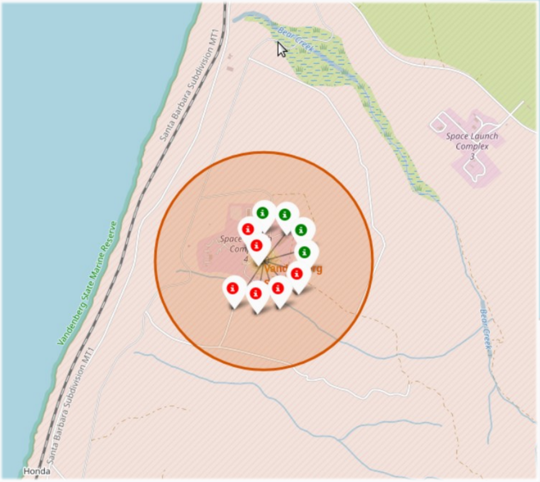
Landing Outcomes

Global view showing East coast and West coast launch clusters.

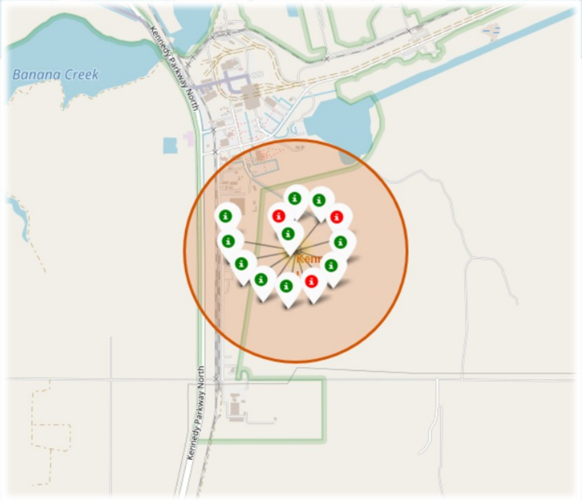




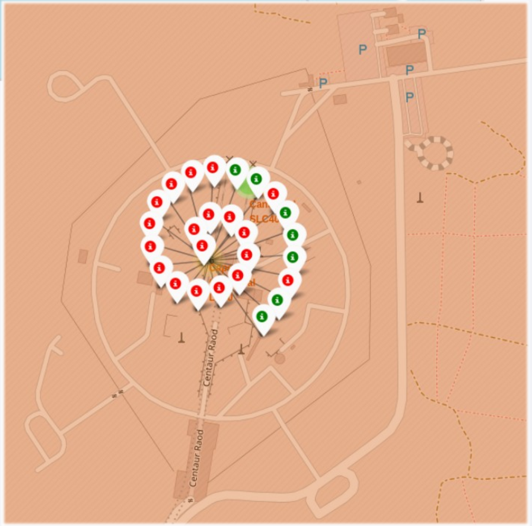
Landing outcomes for the Cape Canaveral SLC-40



Landing outcomes for the Vandenberg SLC



Landing outcomes for the Kennedy LC-39A



Landing outcomes for the Cape Canaveral LC-40

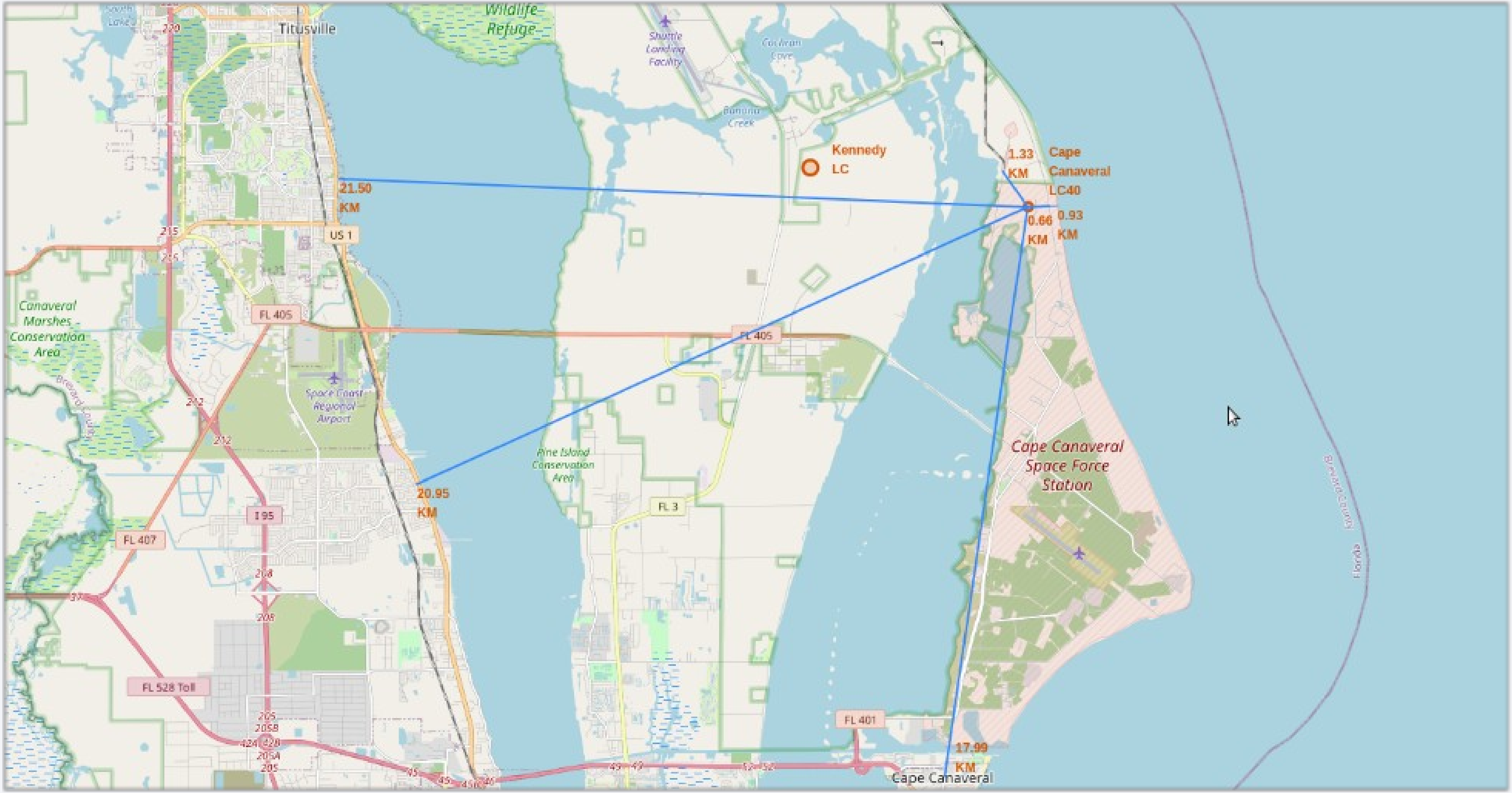
Failed landing

Successful landing

Landing Outcomes

- The map on the previous slide shows upon zoom-in on the individual locations color-labeled marker clusters which help identify the success rate of each site.
- Based on the blown-up details we can observe that the highest success rate occurs at the Kennedy Space Center.
- The highest number of landing attempts are made for launches from the Cape Canaveral location.

Proximity Map



Distance to the nearest cities, highway, coastline, and railroad from the Cape Canaveral LC-40.

Proximity Map (analysis)

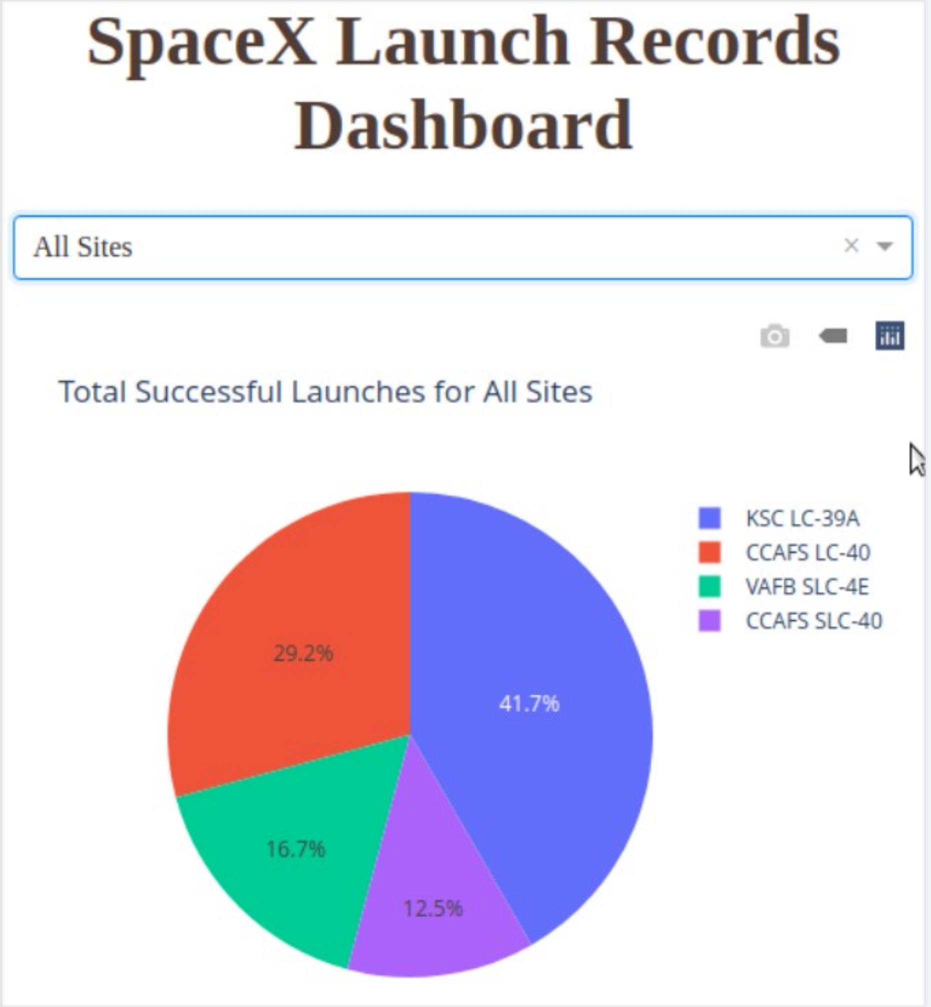
- From the proximity map above we can see that the average distance from Cape Canaveral to the nearest cities is approximately 20 Km
- The average distance from nearby populated areas does not indicate a pattern in launch pad location or a preference for a minimum distance related to the calculated average as other sites have varying distances from populated areas.
- Most launches (except polar) from Cape Canaveral follow an eastern direction due to Earth's direction of rotation[25]
- The launch corridors are designed to avoid falling debris over populated areas in case of rocket failure. Self destructing mechanisms are also implemented for the same purpose[26]
- For range safety, a safe distance of just over 2 miles is the most cited (see Kennedy SC Visitor Complex)[27]
- Launches are not always free of incidents where debris fall over populated areas.[28] Thus, the most significant insight resulted from the proximity analysis is that Cape Canaveral has a particular advantage in its location given by its proximity to a large body of water on the eastern side (i.e. about 0.93 Km from the Atlantic Ocean as shown on the map).
- Rather than seen from a safety perspective, the proximity to the nearest railroad (NASA Railroad, 1.33 Km) and the nearest highway (Samuel C. Philips Pkwy, 0.66 Km) should instead be seen as a possible logistic advantage for facility operations.

Section 5

Build a Dashboard with Plotly Dash

Total Successful Launches

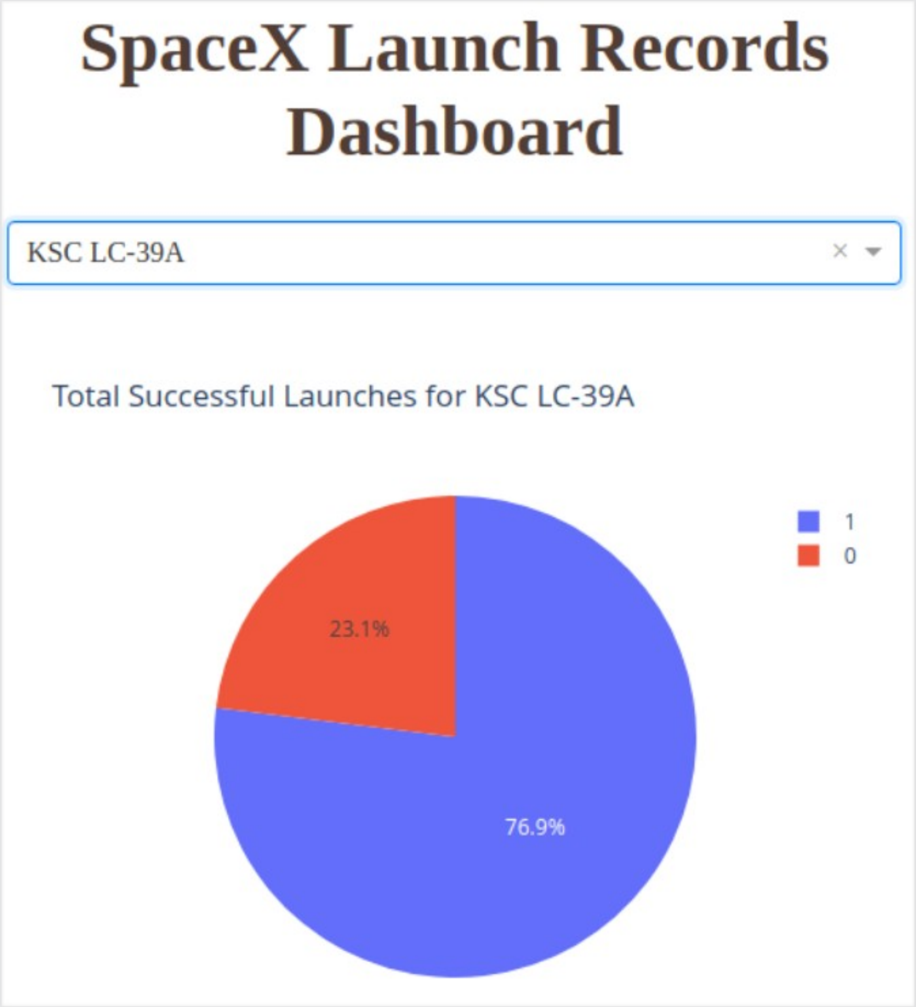
- Based on the data sample chosen for this dashboard demo the most successful launching site is the Kennedy Space Center LC-39A.
- Comparing this result with Section 2 above we might see that this result is not exactly accurate.



Pie chart of the total successful launches for all sites. (Plotly dash demo)

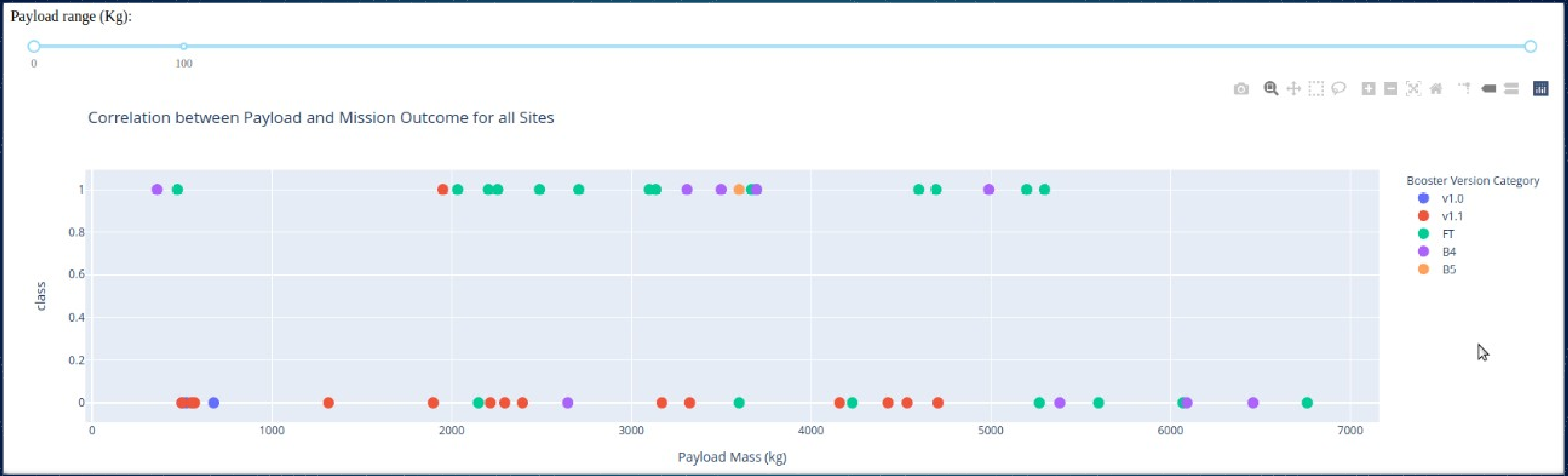
The Most Successful Site

- Focusing on the Kennedy Space Center LC-39A launching pad we observe a success rate of 76.9%.
- The result seems to be in agreement with the results discussed in Section 2 above.
- On the next slide we will focus on the interactive scatter plot section of the dashboard.



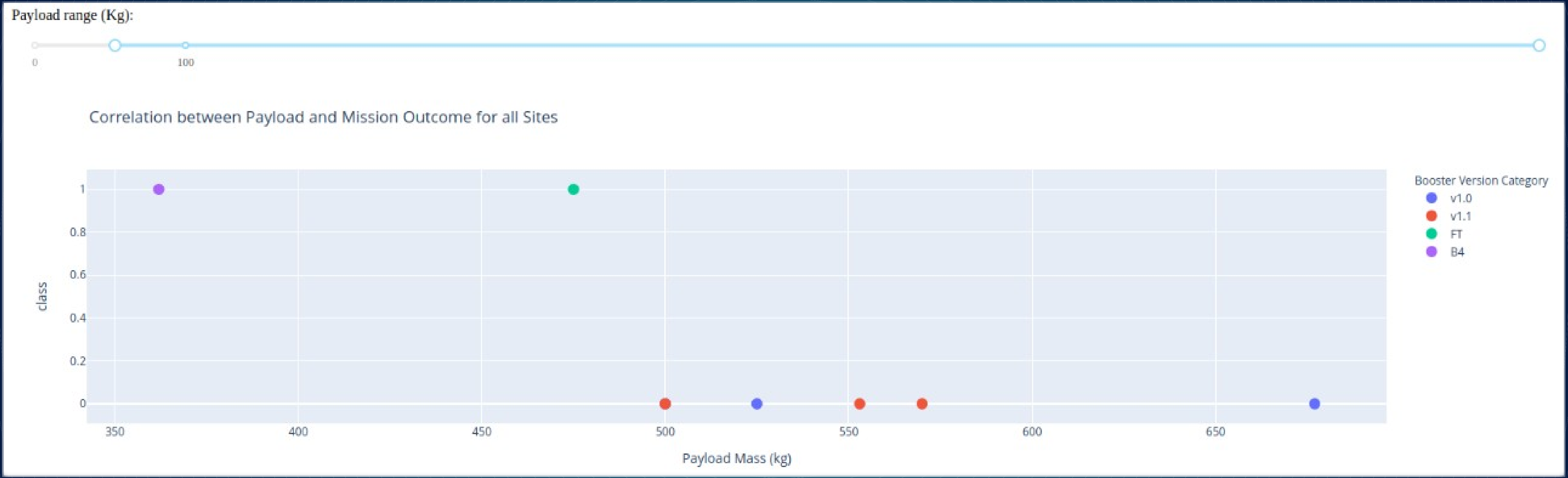
Pie chart showing the success rate of the most successful site. (Plotly dash demo)

Payload vs Outcome I (dashboard demo)



Screenshot 1. Demo of the interactive dashboard scatter plot. The plot shows the entire payload range with the outcomes color-coded by booster version. In this screenshot the payload slider is not yet initialized.

Payload vs Outcome II (dashboard demo)



Screenshot 2. Once the slider is moved, the outcomes displayed are limited by the range of the slider (i.e. 0-1000 Kg).

Payload vs Outcome III (dashboard demo)

Screenshot 3. The slider is moved to show only outcomes in the range 500 – 1000 Kg.

Payload vs Outcome (analysis)

- From the first screenshot of the interactive scatter plot we observe that the most successful booster is FT (Full Thrust) which denotes version v1.2, blocks 1 to 3.

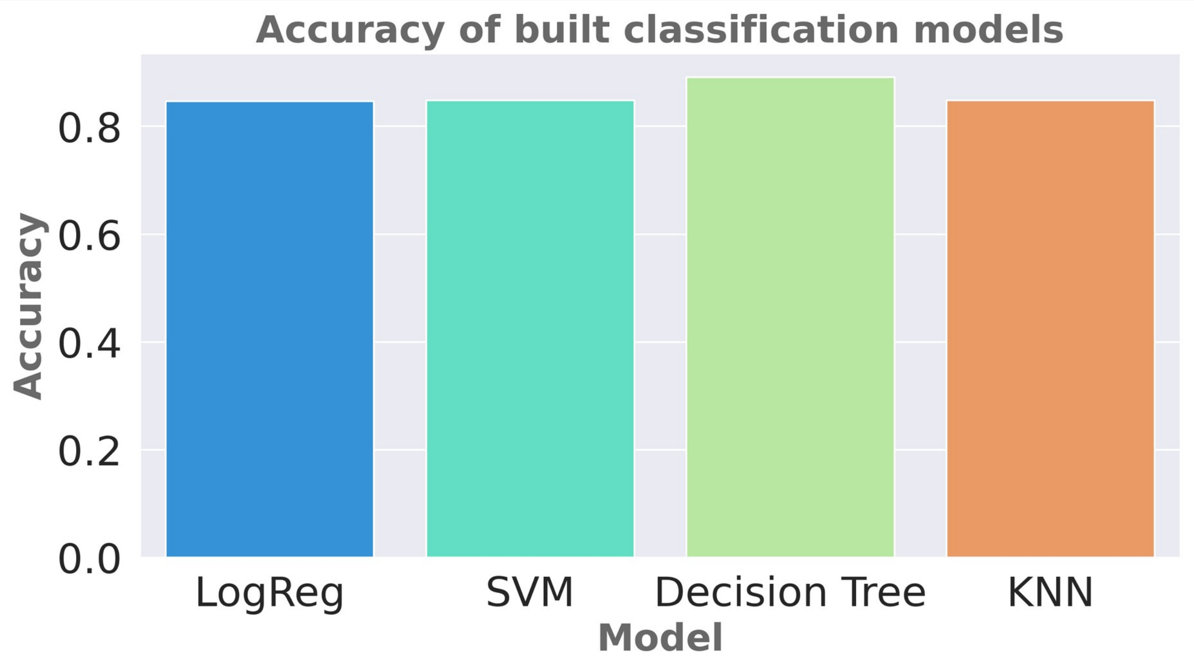
- Two other boosters with successful landings are B4 (Block 4) and B5 (Block 5) which are also part of the v1.2 full thrust family.

- The only other booster with successful landing is v1.1 with only one successful mission carrying a payload just under 2000 Kg. **This particular result cannot be verified against flight records as the two flights that seem to qualify (Flight 17 and Flight 19) both failed to land.** [14]

Section 6

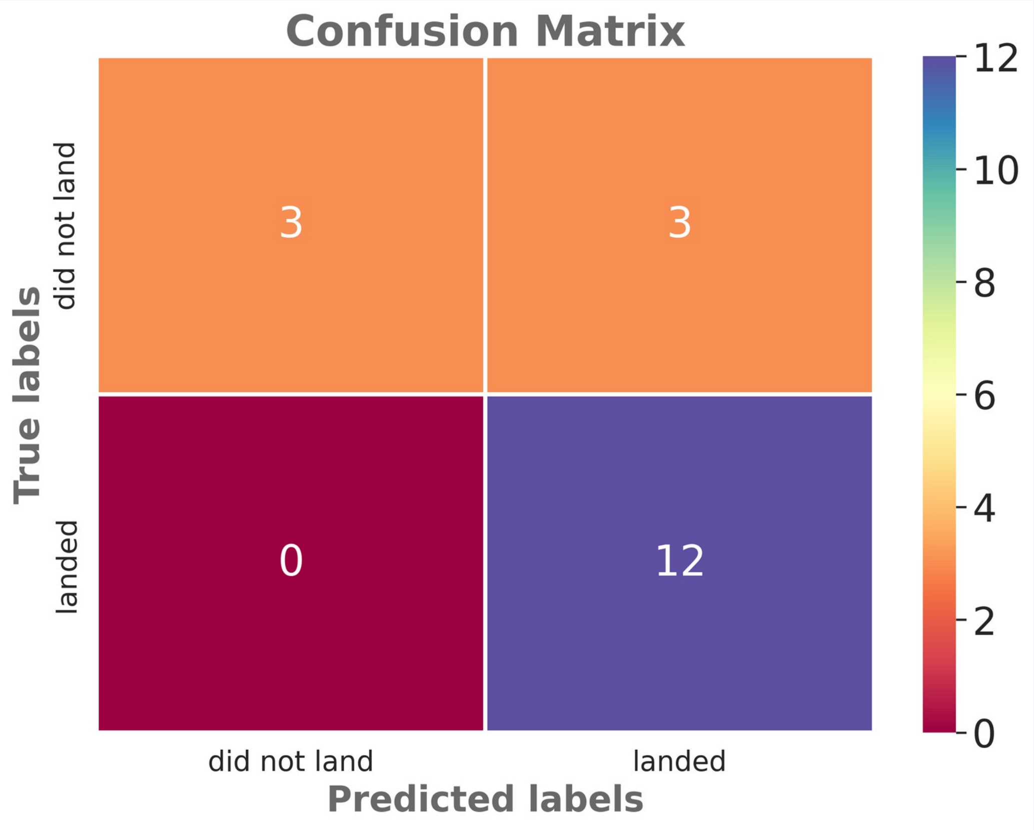
Predictive Analysis (Classification)

Classification Accuracy



- The graph above shows the accuracy of all the classification models tested
- The highest classification accuracy was obtained with the Decision Tree model
- Next, we will take a closer look at the Decision Tree model performance

Confusion Matrix



- The confusion matrix above shows that the model performed best in predicting the successful landing outcomes
- The model incorrectly predicted three out of six negative outcomes as positive.

Conclusions I

From the analysis of the flight density for each launch site it was established that the bulk of the SpaceX missions are handled at Florida sites with the Cape Canaveral SLC-40 site at the lead followed by the Kennedy Space Center.

Building on the above insight we looked at the payload range for each location and confirmed that CCAFS SLC-40 and KSC LC-39A share similar payload ranges while the Vandenberg location on the West coast does not handle payloads larger than 10000 Kg.

By comparing the success rates of different orbit types it became noticeable that launches following a polar trajectory have a strikingly high rate of successful landings. Future investigations need to take a more in dept look at the relation between polar trajectories and landing outcomes.

Looking at the frequency of each orbit type we observed that low Earth orbits are by far the bulk of the business at the current time, with GTO missions at the higher altitude limit. We also observed a shift in frequency between GTO and VLEO with the emergence of the B5 boosters.

By analyzing the payload mass range specific to each orbit we deduced that GTO has a well defined range (between 3000 and 7000 Kg) while ISS has the widest range of payloads. The highest payloads, corresponding to VLEO orbits might also be explained by the higher capacity of the latest boosters.

Conclusions II

Looking at the yearly trend for the landing success rate of the Falcon 9 boosters we observed a rising trend between the years 2013 and 2020. The trend reflects ongoing technology improvements.

Querying the SpaceX data set for unique launch locations returned four site names. Based on external sources[13] it was concluded that two of the denominations refer to the same location i.e. Cape Canaveral LC-40 and Cape Canaveral SLC-40.

The next query merged the results for the two denominations mentioned above and we learned that NASA has been the main SpaceX customer during early test flights and missions. We learned that test landings have been performed early on in the rocket's development confirming that reusability has been an ongoing goal for SpaceX.

Another interesting result suggests that for early LEO missions the F9 v1.1 boosters carried payloads significantly less heavy than the rocket's specifications.

Querying the database for the first successful ground pad landing correctly returned the date of the historical Flight 20 which marks the first true landing of a first stage Falcon 9 booster.

Conclusions III

By retrieving the booster names for boosters that carried payloads between 4000 and 6000 Kg we identified a series of breakthrough missions which mark the constant improvement in technology that defines the increasing reliability and cost efficiency of the Falcon 9 rockets.

Next, it was demonstrated that SpaceX attempted to land only 70% of its missions. A future study needs to further consider the factors that lead to the dismissal of a landing opportunity. Improvements in dealing with such factors might hint at the possibility of developing a competitive edge for a prospective SpaceX competitor.

We also determined that Block 5 boosters are the state of the art in terms of maximum payload capacity.

It was also proved that, while more difficult to master, drone ship landings are more frequent in early flights. This might be due to safety concerns and possible delays in obtaining ground landing authorizations.

The proximity analysis of the launch sites underlines both the importance of site location near the Equator to ensure maximum launch efficiency, and the importance of proximity to a large body of water to ensure launch corridor safety in relation to nearby populated areas.

Conclusions IV

To enforce our results, the interactive dashboard accompanying the current study reveals insights into the success rate of each launch site location as well the considerable success rate of boosters in the Full Thrust family.

Upon testing various classification models it was established that the Decision Tree model yields the highest predictive accuracy. Future model developments will benefit from the accumulation of new data.

References I

1. Falcon 9 full thrust, Wikipedia. (2021). https://en.wikipedia.org/wiki/Falcon_9_Full_Thrust (accessed January 12, 2022).

2. M. Wall, SpaceX rocket could be 100-percent reusable by 2018, Elon Musk says, Space.com. (2017). https://www.space.com/36412-spacex-completely-reusable-rocket-elon-musk.html (accessed January 12, 2022).

3. SpaceX company profile, Fortune. (2020). https://fortune.com/company/spacex/ (accessed January 12, 2022).

4. S. Clark, SpaceX reaches Rocket Reuse Milestone on starlink launch, Spaceflight Now. (2021). https://spaceflightnow.com/2021/05/09/spacex-reaches-rocket-reuse-milestone-on-starlink-launch/ (accessed January 12, 2022).

5. The space industry will be worth nearly $3 trillion in 30 years, Bank of America predicts, CNBC. (2017). https://www.cnbc.com/2017/10/31/the-space-industry-will-be-worth-nearly-3-trillion-in-30-years-bank-of-america-predicts.html (accessed January 12, 2022).

6. J. Wattles, Boeing set to retry test flight in bid to take on SpaceX, CNN. (2021). https://www.cnn.com/2021/08/02/tech/boeing-nasa-starliner-flight-test-scn/index.html (accessed January 12, 2022).

7. M. Mayfield, Ready For Reuse, Launch Program Leveraging Recycled Rockets. (n.d.). https://www.nationaldefensemagazine.org/articles/2021/8/6/launch-program-leveraging-recycled-rockets (accessed January 12, 2022).

8. T. Malik, Launchpad explosion destroys SpaceX Falcon 9 rocket, satellite in Florida, Space.com. (2016). https://www.space.com/33929-spacex-falcon-9-rocket-explodes-on-launch-pad.html (accessed January 12, 2022).

9. Falcon 9 fact sheet, Spaceline. (n.d.). https://www.spaceline.org/cape-canaveral-rocket-missile-program/falcon-9-fact-sheet/ (accessed January 13, 2022).

10.W. William, SpaceX launches satellites in Rare Polar Orbit Mission, CBS News. (2021). https://www.cbsnews.com/news/spacex-launches-satellites-starlinks-polar-orbit/ (accessed January 13, 2022).

11. SpaceX launch facilities, SpaceX Launch Facilities. (2021). https://en.wikipedia.org/wiki/SpaceX_launch_facilities (accessed January 9, 2022).

12.Spaceflight now: Breaking news: SpaceX says robust market can support four launch pads. (n.d.). https://spaceflightnow.com/news/n1310/18spacex (accessed January 9, 2022).

13.Cape Canaveral Space Launch Complex 40, Wikipedia. (2021). https://en.wikipedia.org/wiki/Cape_Canaveral_Space_Launch_Complex_40 (accessed January 9, 2022).

14.List of falcon 9 and Falcon Heavy launches, Wikipedia. (2022). https://en.wikipedia.org/wiki/List_of_Falcon_9_and_Falcon_Heavy_launches#Past_launches (accessed January 9, 2022).

References II

15.G. Sage, Latest contract releases, NASA. (2018). https://www.nasa.gov/office/procurement/latest-contract-releases (accessed January 13, 2022).

16.Capabilities and Services, SpaceX. (2012). https://web.archive.org/web/20131007205105/http://www.spacex.com/about/capabilities (accessed January 10, 2022).

17.Falcon 9 Flight 20, Wikipedia. (2021). https://en.wikipedia.org/wiki/Falcon_9_flight_20 (accessed January 10, 2022).

18.J. Amos, SpaceX records another rocket landing, BBC News. (2016). https://www.bbc.com/news/science-environment-36223745 (accessed January 10, 2022).

19.P. Blau, Falcon 9 rocket lifts Japanese communications satellite, aces high-energy ocean landing, Falcon 9 JCSat16 Spaceflight101. (2017). https://spaceflight101.com/falcon-9-jcsat-16/falcon-9-rocket-lifts-japanese-communications-satellite-aces-high-energy-ocean-landing/ (accessed January 10, 2022).

20.W. Graham, SpaceX conducts historic Falcon 9 re-flight with SES-10 - lands booster again, NASASpaceFlight.com. (2017). https://www.nasaspaceflight.com/2017/03/spacex-historic-falcon-9-re-flight-ses-10/ (accessed January 10, 2022).

21.S. Clark, SES agrees to launch another satellite on previously-flown Falcon 9 booster, Spaceflight Now. (2017). https://spaceflightnow.com/2017/08/04/ses-agrees-to-launch-another-satellite-on-a-previously-flown-falcon-9-booster/ (accessed January 10, 2022).

22.H. Caleb , C. Henry, Musk details block 5 improvements to Falcon 9, SpaceNews. (2018). https://spacenews.com/musk-details-block-5-improvements-to-falcon-9/ (accessed January 11, 2022).

23.B. Evans, C.C. says: A. says: D. says: J. says: T.A. says: et al., USAF confirms SpaceX awaiting FAA approval for Booster Landing at cape for return to Flight, AmericaSpace. (2015). https://www.americaspace.com/2015/12/02/usaf-confirms-spacex-awaiting-faa-approval-for-booster-landing-at-cape-for-return-to-flight/ (accessed January 11, 2022).

24.Basics of space flight - solar system exploration: NASA science, NASA. (n.d.). https://solarsystem.nasa.gov/basics/chapter14-1/ (accessed January 11, 2022).

25.Launch a rocket from a spinning planet, NASA. (n.d.). https://spaceplace.nasa.gov/launch-windows/en/ (accessed January 12, 2022).

26.Range safety, Wikipedia. (2021). https://en.wikipedia.org/wiki/Range_safety#Launch_corridor (accessed January 12, 2022).

27.Rocket Launch Viewing Guide: Where to Watch Launches from Cape Canaveral. (n.d.). http://www.launchphotography.com/Launch_Viewing_Guide.html (accessed January 12, 2022).

28.J. Roulette, SpaceX rocket debris lands on Man's Farm in Washington, The Verge. (2021). https://www.theverge.com/2021/4/2/22364582/spacex-rocket-debris-falls-farm-washington (accessed January 12, 2022).

Appendix

GitHub Repo Link:

https://github.com/Al-1n/IBM_SpaceX_Capstone/tree/master


Thank you!
SPACEX