IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Jie Min
Feb 2025

Outline

Executive Summary

- In this analysis, historical Falcon 9 launch data were used to gain insights into predicting successful landings. The data were initially retrieved from the SpaceX API and Wikipedia. Next, exploratory data analysis (EDA) was performed using Python, SQL, and various visualization tools, including graphs, charts, maps(Folium), and interactive visualizations(Dash), to identify patterns and relationships within the data. Following EDA, a range of classification algorithms were applied, including Logistic Regression, Support Vector Machine, Decision Trees, and K-Nearest Neighbors. Additionally, GridSearchCV was used to optimize model performance by selecting the best hyperparameters.

- Landing success rate has improved significantly, rising from 0% in 2013 to 84.2% in 2020; Cape Canaveral SLC-40 has handled the majority of launches and has improved to be the site with the highest landing success rate in recent years compared to the others

- The primary orbit type has shifted from GTO and ISS in earlier flights to VLEO in later flights. VLEO missions typically carry heavier payloads and have a significantly higher landing success rate

- Booster FT has a significantly higher landing success rate in the 2,000 - 4,000 kg payload range compared to other ranges, while booster B4 has significantly higher success rate (83%) for payloads under 5,000 kg

- Decision tree model is the best performed machine model comparing to other algorisms that have been used in this analysis. It achieved 94.4% accuracy score in test data

Introduction

- The commercial space industry has been rapidly expanding. Through rocket services, companies have successfully sent spacecraft to the International Space Station, developed satellite-based internet access, and launched manned missions into space.
- Our competitor, SpaceX, has been able to offer highly competitive pricing in the rocket services market by significantly reducing costs through first-stage recovery, allowing the rocket to land safely back on Earth after returning from space.
- This analysis utilizes SpaceX's historical launch data and data science techniques, including machine learning, to gather information and extract insights. The goal is to improve our ability to predict whether the first stage will land successfully, which in turn helps determine the pricing we can offer to the market.

Section 1

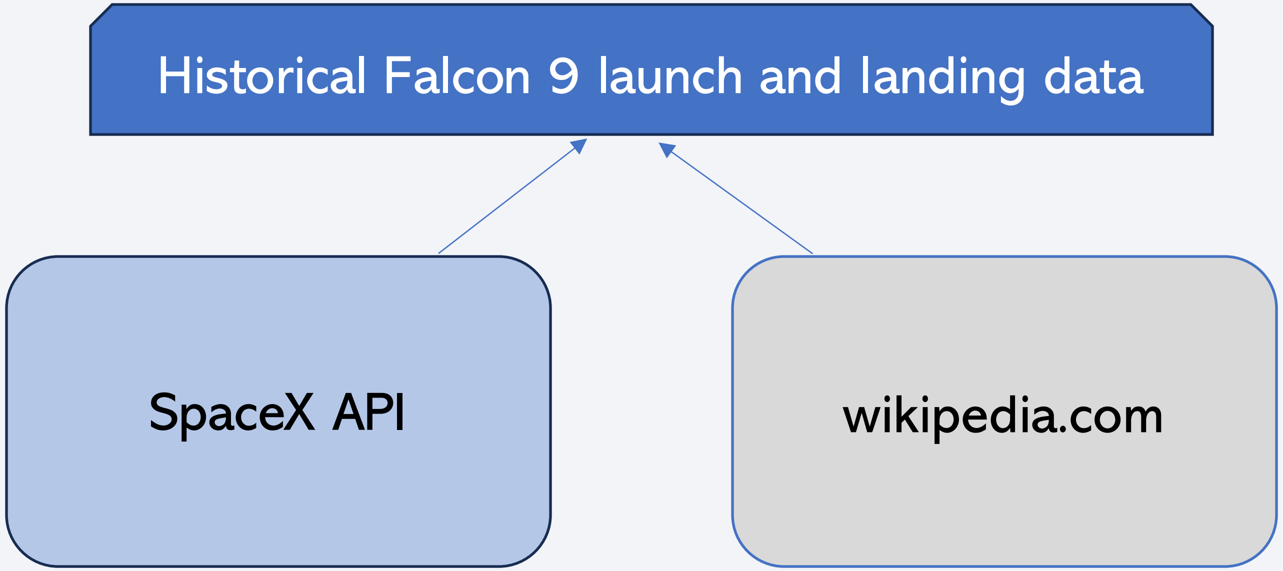
Methodology

Methodology

Executive Summary

- Data collection methodology:
  - Historical Falcon 9 launch and landing data at SpaceX API and Wikipedia page
  - Perform data wrangling, using Python libraries requests, BeautifulSoup, Pandas etc.
- Perform exploratory data analysis (EDA) using visualization and SQL
- Perform interactive visual analytics using Folium and Plotly Dash
- Perform predictive analysis using classification models
  - Logistic Regression, SVM, Decision Trees and K nearest Neighbors
  - Standardization feature scaling; Train/test split
  - GridSearchCV() to tune hyperparameters and cross validation
  - Evaluated with accuracy score and confusion matrix

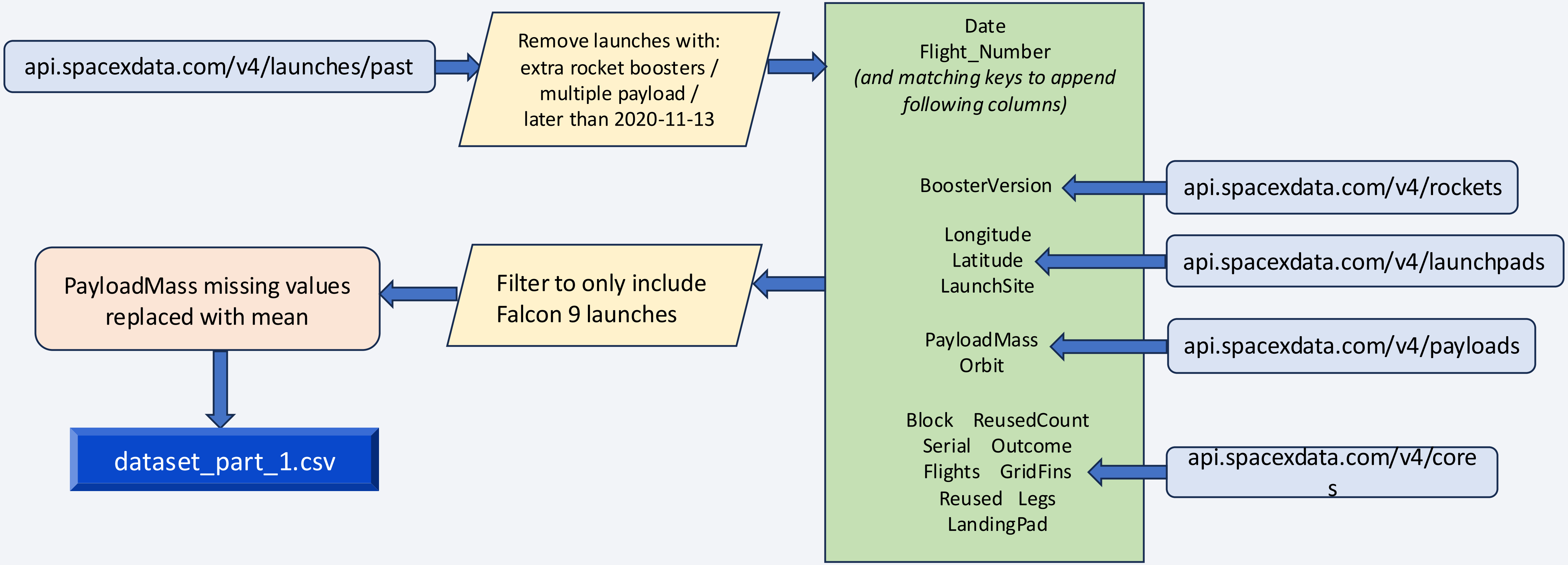
Data Collection



api.spacexdata.com/v4/launches/past

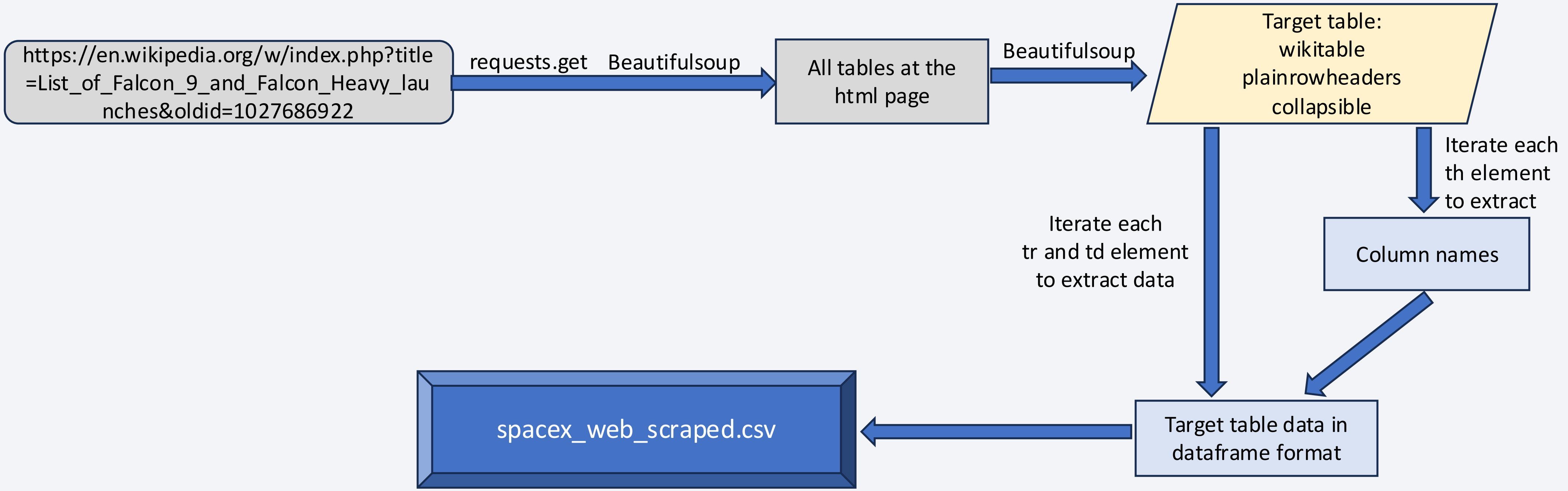
https://en.wikipedia.org/w/index.php?title=List_of_Falcon_9_and_Falcon_Heavy_launches&oldid=1027686922

Data Collection – SpaceX API



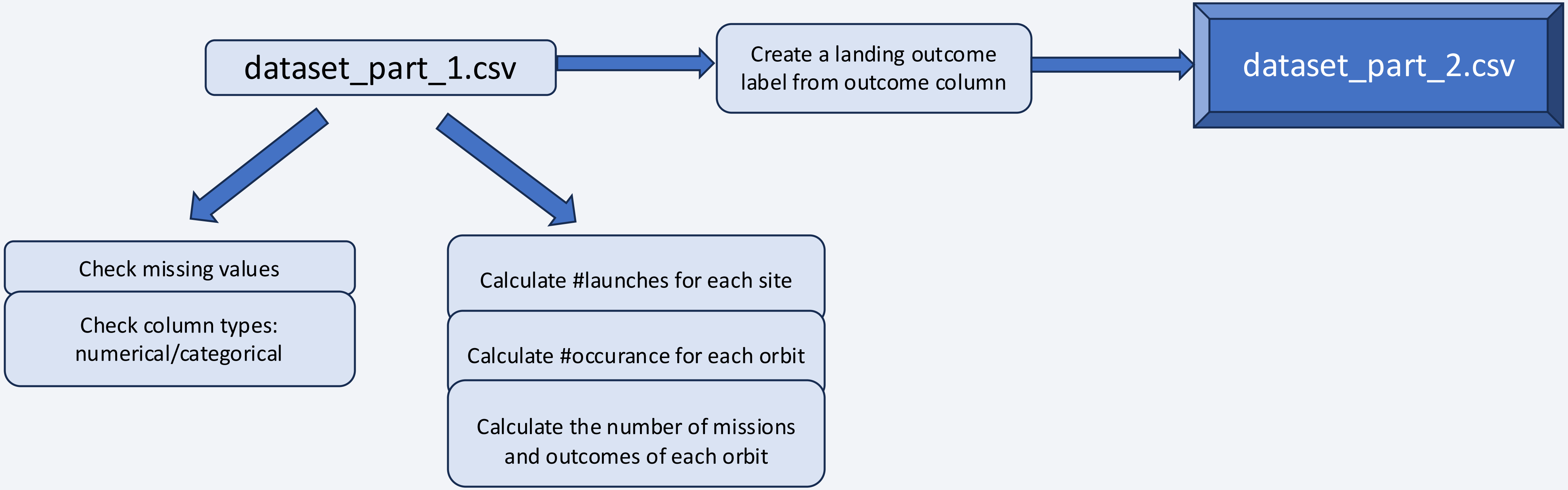
Github url: https://github.com/jmin2024/IBM_capstone/blob/main/jupyter-labs-spacex-data-collection-api.ipynb

Data Collection – Web Scraping



Github url: https://github.com/jmin2024/IBM_capstone/blob/main/jupyter-labs-webscraping.ipynb

Data Wrangling



Github url: https://github.com/jmin2024/IBM_capstone/blob/main/labs-jupyter-spacex-Data%20wrangling.ipynb

EDA with Data Visualization

- Visualization charts:
  - Scatter plot to show the relationship between FlightNumber and PayloadMass
  - Scatter plot to show the relationship between Flight Number and Launch Site
  - Scatter plot to show the relationship between Payload Mass and Launch Site
  - Bar chart to show the relationship between success rate of each orbit type
  - Scatter plot to show the relationship between FlightNumber and Orbit type
  - Scatter plot to show the relationship between Payload Mass and Orbit type
  - Line chart to show launch success yearly trend

- Feature engineering to create dummy variables for categorical columns
- Github url: https://github.com/jmin2024/IBM_capstone/blob/main/edadataviz.ipynb

EDA with SQL

- Information extracted in this section include:
    - Names of the unique launch sites in the space mission
    - Total payload mass carried by boosters launched by NASA (CRS)
    - Average payload mass carried by booster version F9 v1.1¶
    - The date when the first succesful landing outcome in ground pad was acheived.
    - The names of the boosters which have success in drone ship and have payload mass greater than 4000 but less than 6000
    - Total number of successful and failure mission outcomes
    - The names of the booster_versions which have carried the maximum payload mass
    - The records which will display the month names, failure landing_outcomes in drone ship ,booster versions, launch_site for the months in year 2015
    - Rank the count of landing outcomes (such as Failure (drone ship) or Success (ground pad)) between the date 2010-06-04 and 2017-03-20, in descending order.

- Github url: https://github.com/jmin2024/IBM_capstone/blob/main/jupyter-labs-eda-sql-coursera_sqllite.ipynb

Build an Interactive Map with Folium

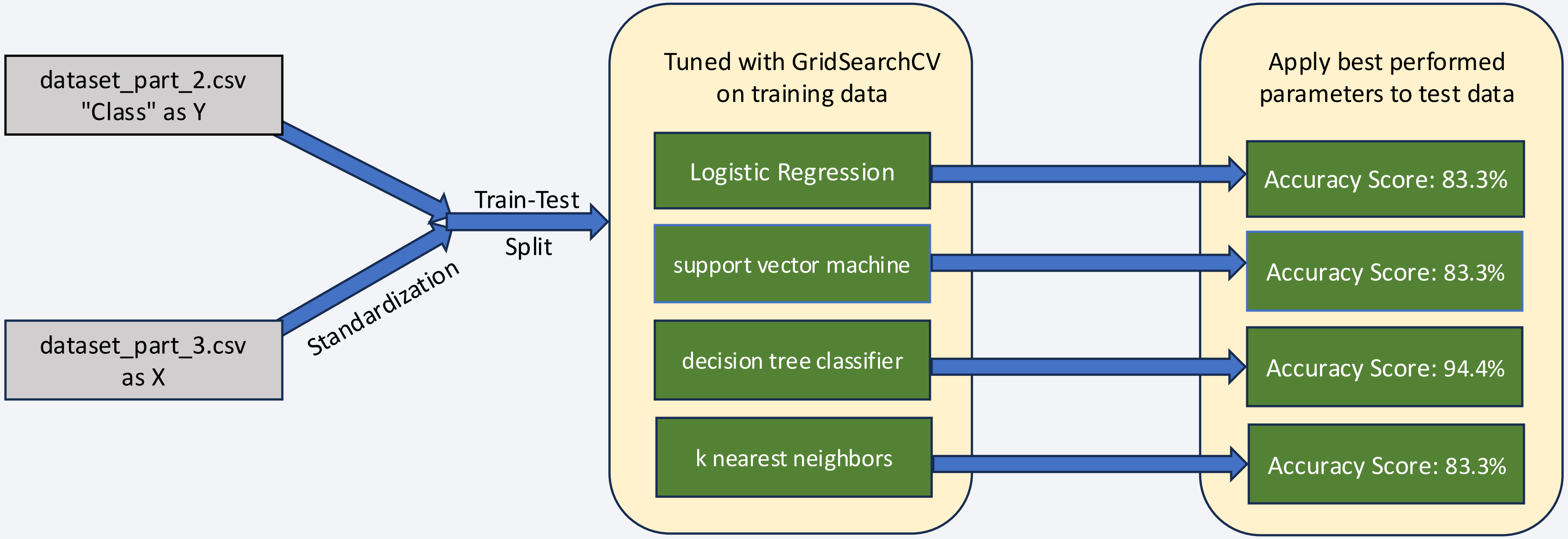
- All launch sites were displayed on a map through folium.Circle and Marker
- Launches for each site were shown on the map through folium.MarkerCluster, with success/failed differentiated through folium.Icon color
- Distances between a launch site to its proximities were calculated using predefined function and visualized with folium.PolyLine
- These procedures help us understand the location of each site, the outcomes of launches at these sites, and the distances between each site and nearby locations.
- GitHub URL: https://github.com/jmin2024/IBM_capstone/blob/main/lab_jupyter_launch_site_location.ipynb

Build a Dashboard with Plotly Dash

- Success rates at each site and the overall success distribution across all sites were visualized using pie charts, so that performance across different locations could be easily compared and analyzed.
- A scatter plot of payload versus success/failure by booster version was presented, featuring a payload range slider, so that the relationship between payload weight and launch outcomes could be explored by different booster versions.
- GitHub URL:https://github.com/jmin2024/IBM_capstone/blob/main/Dash_project.ipynb

Predictive Analysis (Classification)

- A total of 83 independent variables, including Payload Mass, Orbit, Launch Site, Number of Flights, and Reusability, are used to predict landing outcomes
- Four machine learning algorithms—Logistic Regression, Support Vector Machine (SVM), Decision Tree, and K-Nearest Neighbors (KNN)—are utilized for prediction. They are further tuned by GridSearchCV to optimize their performance.
- GitHub URL: https://github.com/jmin2024/IBM_capstone/blob/main/SpaceX_Machine%20Learning%20Prediction_Part_5.ipynb



Results

- Landing success rate has improved significantly, rising from 0% in 2013 to 84.2% in 2020;
- All the three launch sites are near coast line. Cape Canaveral SLC-40 has handled 61% of all launches and has improved to be the site with the highest landing success rate in recent years compared to the others
- Geosynchronous Transfer Orbit (GTO) and ISS are the primary orbit type before Flight 70, mostly carrying payloads under 8,000 kg, with 50% landing success rate. However, Very Low Earth Orbit (VLEO), which handles heavier payloads above 13,000 kg, gradually became the dominant orbit type afterward, achieving an 85.7% success rate; Most of the orbit types show an increasing landing success rate over time
- Boost FT, B4, and B5 are progressively newer booster versions over time. Within the timeframe of each type, certain payload ranges show higher landing success rates than others. For example, Booster FT has a significantly higher success rate in the 2,000 - 4,000 kg payload range, and Booster B4 has a much higher success rate (83%) for payloads under 5,000 kg; The first successful landing on ground pad happened on Dec 22, 2015; The heaviest payload Falcon 9 launched is 15,600 kg with F9 B5 booster
- Among the four machine learning algorithms (Logistic Regression, Support Vector Machine, Decision Trees, and K-Nearest Neighbors), the Decision Tree model performed the best compared to the other algorithms used in this analysis. It achieved an accuracy score of 94.4% on the test data

Interactive Visualization Demo Video

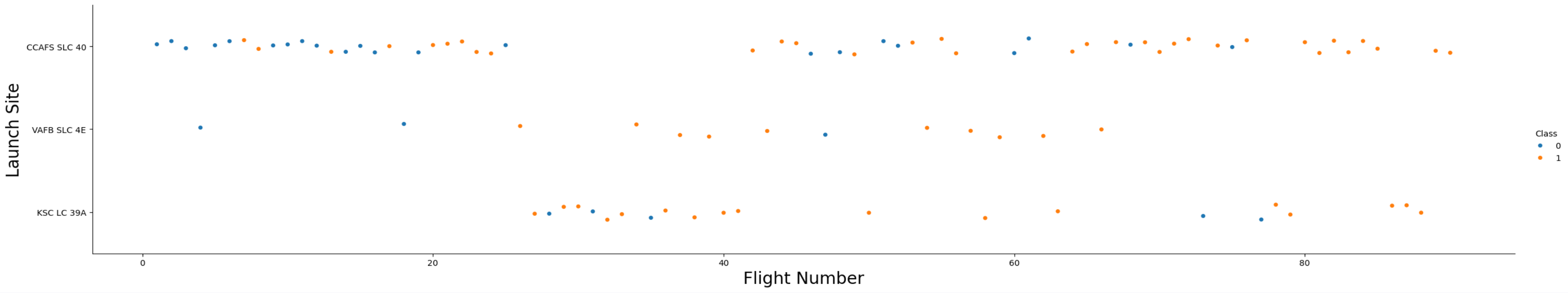
Section 2

Insights drawn from EDA

Flight Number vs. Launch Site

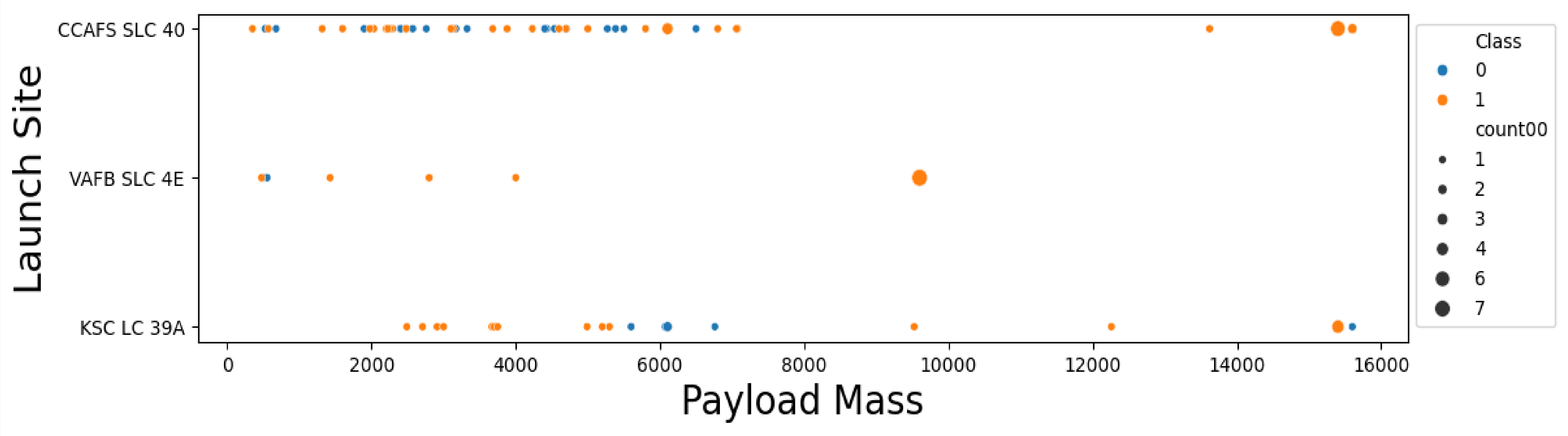
- Cape Canaveral SLC-40 site accounted for 61% of all Falcon 9 launches, with a success rate of 60%. The other two sites handled 24% and 14% of the total launches, with a success rate of 77%
- Cape Canaveral SLC-40 site success rate improved significantly after flight number 70
- Cape Canaveral SLC-40 typically handles the majority of Falcon 9 launches, however during flights 25 to 41, most launches appeared to shift to Kennedy LC-39A
- Vandenberg SLC-4E site did not launch any Falcon 9 after flight number 69, while Kennedy LC-39A did not launch any Falcon 9 before flight number 27

Launch Site	# Launch	% Total	Success Rate
Cape Canaveral, SLC-40	55	61%	60%
Vandenberg, SLC-4E	13	14%	77%
Kennedy, LC-39A	22	24%	77%
Total	90	100%	67%



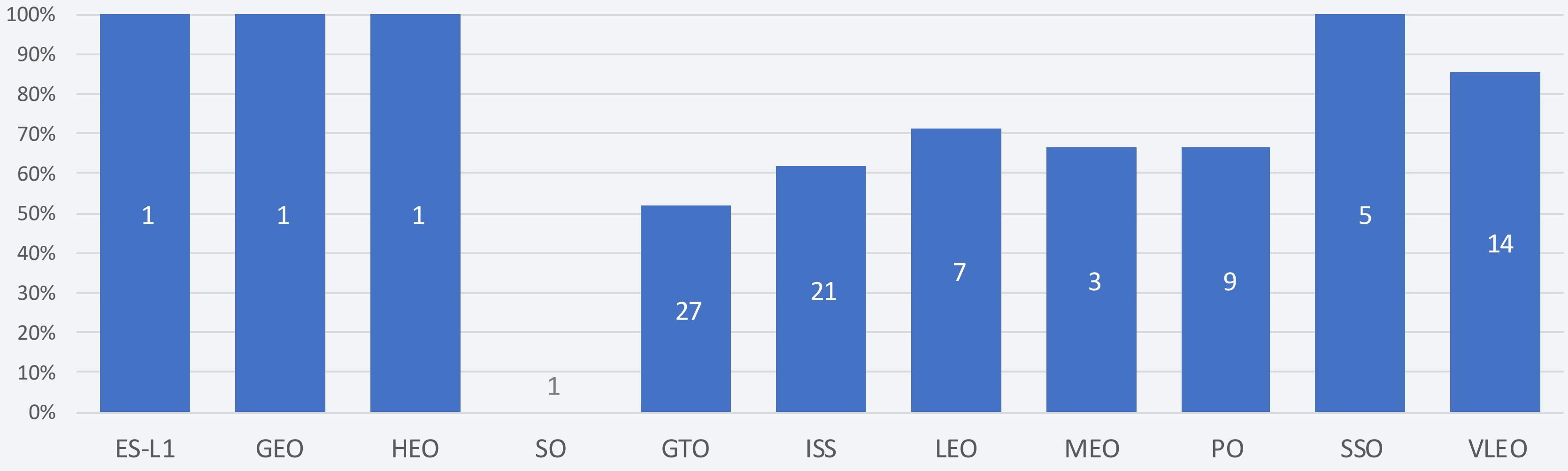
Payload vs. Launch Site

- The majority of flights carry payloads under 8,000 kg. Flights with payloads exceeding 8,000 kg have a significantly higher l anding success rate (87%), with most concentrated around 9,600 kg and 15,000 kg
- Cape Canaveral SLC-40 has a higher landing success rate for heavy payloads above 8,000 kg
- Most of the failed landings at the Kennedy LC-39A site involved payloads around 6,000 kg
- Vandenberg SLC-4E site did not handle any flight with payload higher than 9,600 kg



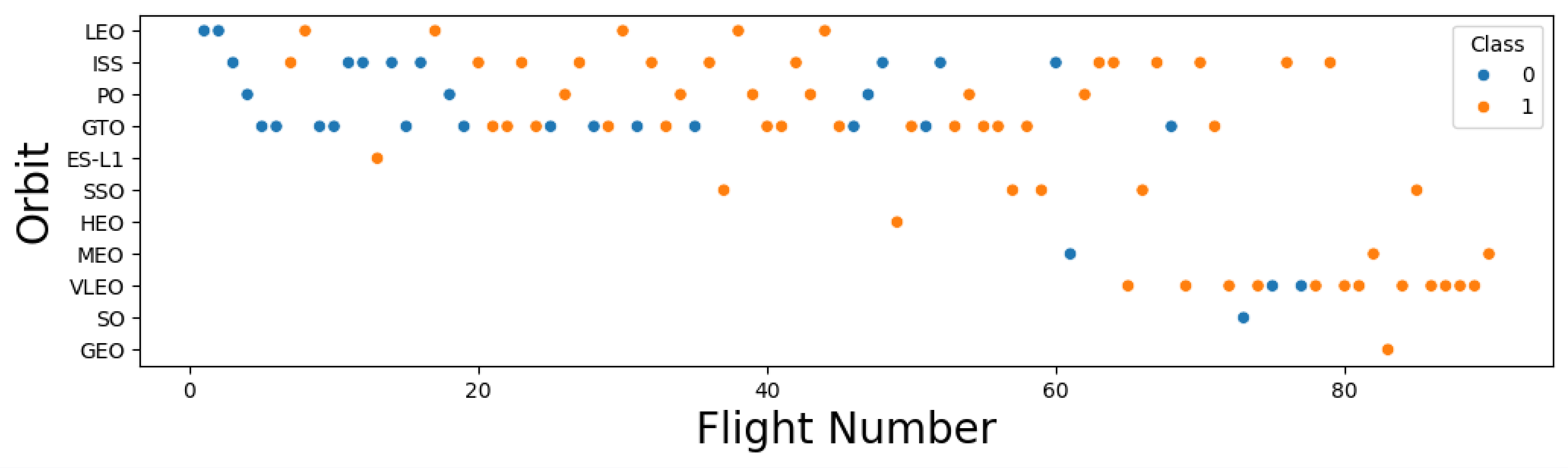
Success Rate vs. Orbit Type

- The orbits with a 100% success rate are all low-frequency orbits: SSO had five launches, while ES-L1, GEO, and HEO each had only one launch
- With 27 launches, GTO is the most frequently launched orbit type. However, it also has the lowest success rate, at just 52%
- VLEO, with 14 launches, has success rate of 86%
- ISS (21 launches), LEO (7 launches), MEO (3 launches) and PO (9 launches) have success rate of around 65%
- The definitions of orbit types can be found in the appendix on page 40



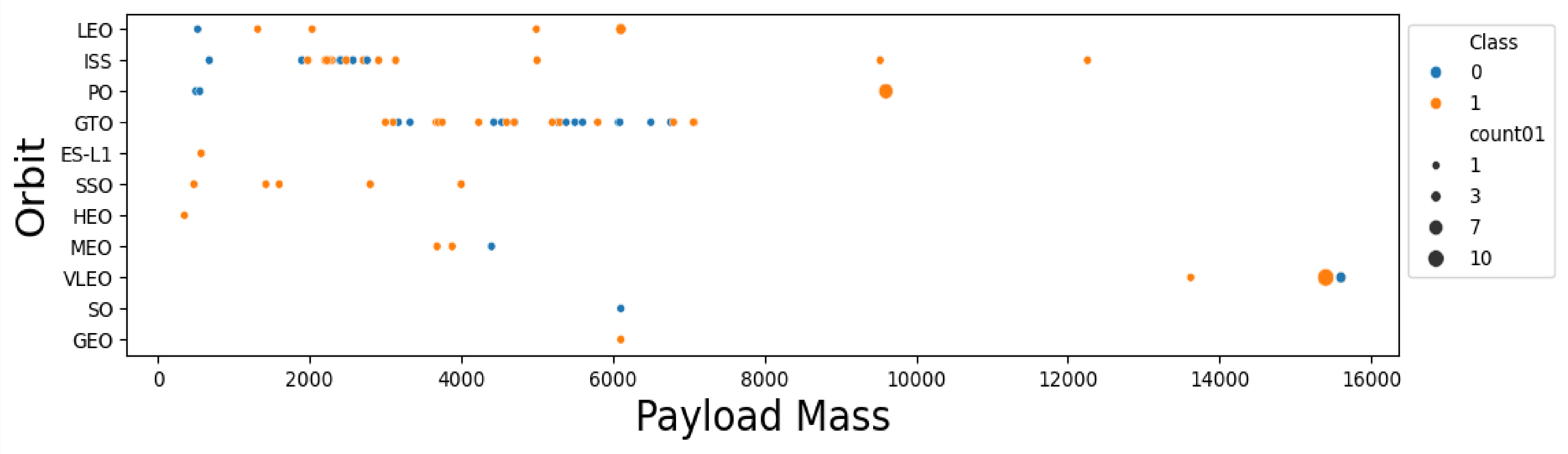
Flight Number vs. Orbit Type

- Launches before flight number 60 were mostly GTO, ISS, PO, and LEO
- VLEO gradually became the most frequent orbit after its first occurrence at at flight number 65
- Most orbits demonstrate improving success rate over time



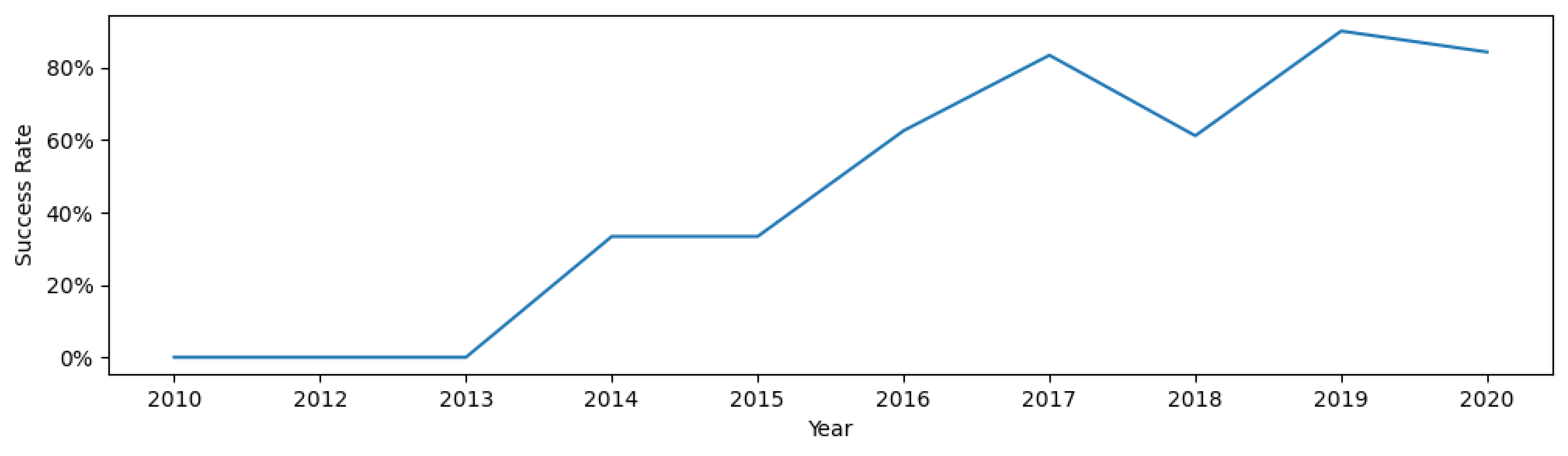
Payload vs. Orbit Type

- Overall, launches carrying payloads above 8,000 kg have a significantly higher landing success rate compared to those with payloads below 8,000 kg. However, higher payloads are also associated with later launches, which generally have higher success rates
- LEO, ISS, and PO orbits show a higher success rate at higher payload ranges than at lower ones, GTO payloads are distributed between 3,000 and 7,000 kg, with success rates appearing unrelated to payload mass
- ISS payloads are mostly concentrated between 2,000 and 3,000 kg, with all payloads above 3,000 kg achieving a 100% landing success rate
- SSO payloads are all below 4,000 kg, and every launch has been successful, regardless of payload mass.
- VLEO launches carry heavy payloads exceeding 13,000 kg, with a high success rate of 86%.
- MEO, HEO, SO, and GEO have too few launches to identify any clear patterns



Launch Success Yearly Trend

- Success rate of Falcon 9 shows upward trend, rise from 0% in 2010 to 84.2% in 2020
- The highest landing success rate occurred in 2019 at 90%
- There was a dip in 2018 following two years of rapid growth, but it recovered and achieved an even higher success rate afterward



All Launch Site Names

There are three main unique launch sites for SpaceX

Cape Canaveral Space Force Station (CCAFS SLC-40/ CCAFS LC-40) - Florida
One of the busiest launch sites for Falcon 9 missions, handling a wide range of payloads.

Kennedy Space Center (KSC LC-39A) - Florida
A historic NASA launch pad now leased by SpaceX

Vandenberg Space Force Base VAFB SLC-4E - California
Primarily used for polar(PO) and sun-synchronous(SSO) orbit launches.

```
%sql select distinct Launch_Site from SPACEXTABLE;
```

```
 * sqlite:///my_data1.db
Done.
```

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

Launch Site Names Begin with 'CCA'

- The LIKE clause in SQL is used to search for strings that match a specified pattern

```
%sql select * from SPACEXTABLE where Launch_Site like 'CCA%' limit 5;
```

```
 * sqlite:///my_data1.db
Done.
```

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

Total Payload Mass

- Total payload carried by boosters from NASA is 45,596 kg
- The code retrieves customer column and calculates the total PAYLOAD_MASS_KG from SPACEXTABLE, only on rows where customer column has a value of ‘NASA (CRS)’

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

```
%sql select customer, sum(PAYLOAD_MASS__KG_) from SPACEXTABLE where customer = 'NASA (CRS)';
```

```
 * sqlite:///my_data1.db
Done.
```

Customer	sum(PAYLOAD_MASS__KG_)
NASA (CRS)	45596

Average Payload Mass by F9 v1.1

- The average payload mass carried by booster version F9 v1.1 is 2928.4 kg
- The code retrieves Boost_Version column and calculates the average of PAYLOAD_MASS_KG from SPACEXTABLE, only on rows where Boost_Version column has a value of ‘F9 v1.1’

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

```
%sql select Booster_Version, avg(PAYLOAD_MASS__KG_) from SPACEXTABLE where Booster_Version='F9 v1.1';
```

```
 * sqlite:///my_data1.db
Done.
```

Booster_Version	avg(PAYLOAD_MASS__KG_)
F9 v1.1	2928.4

First Successful Ground Landing Date

- The first successful landing outcome on ground pad happened on Dec 22, 2015
- The code retrieves the minimum date value from SPACEXTABLE where Landing_Outcome has a value of 'Success (ground pad)'

```
%sql select min(date) from SPACEXTABLE where Landing_Outcome='Success (ground pad)';
```

```
 * sqlite:///my_data1.db
Done.
```

min(date)
2015-12-22

Successful Drone Ship Landing with Payload between 4000 and 6000

- The boosters that have successfully landed on drone ship and had payload mass greater than 4000 but less than 6000 are F9 FT B1022, B1026, B1021.2 and B1031.2
- The query retrivess the Boost_Version, PAYLOAD_MASS_KG_ and Landing_Outcome field from SPACEXTABLE where the Landing_Outcome colume has value of 'Success (drone ship)' and PAYLOAD_MASS_KG_ value is greater than 4000 and less than 6000

```
%sql select Booster_Version, PAYLOAD_MASS__KG_, Landing_Outcome from SPACEXTABLE where Landing_Outcome='Success (drone ship)'
    and PAYLOAD_MASS__KG_>4000 and PAYLOAD_MASS__KG_<6000;
```

```
 * sqlite:///my_data1.db
Done.
```

Booster_Version	PAYLOAD_MASS__KG_	Landing_Outcome
F9 FT B1022	4696	Success (drone ship)
F9 FT B1026	4600	Success (drone ship)
F9 FT B1021.2	5300	Success (drone ship)
F9 FT B1031.2	5200	Success (drone ship)

Total Number of Successful and Failure Mission Outcomes

- The total number of successful and failed mission outcomes are 100 and 1, respectively.
- Please note that the dataset used in this section (Pages 25-33) contains 12 more records than the dataset used on page 18-23. Additionally, the Mission_Outcome in this section differs from the Landing Outcome Success column ('Class') referenced elsewhere.
- The code retrieves all rows from SPACEXTABLE, groups them by the Mission_Outcome column, and counts the number of rows for each unique outcome value

List the total number of successful and failure mission outcomes

```
%sql select Mission_Outcome, count(*) from SPACEXTABLE group by Mission_Outcome;
```

```
 * sqlite:///my_data1.db
Done.
```

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

Boosters Carried Maximum Payload

- The boosters that have carried the maximum payload mass are listed in the following output.
- The query retrieved the maximum PAYLOAD_MASS_KG_ value and then returned it to the upper level query where booster_version and PAYLOAD_MASS_KG_ are pulled where PAYLOAD_MASS_KG_ equals to the returned value from the nested query

```
%sql select booster_version, "PAYLOAD_MASS__KG_" from SPACEXTABLE where "PAYLOAD_MASS__KG_"= (select max("PAYLOAD_MASS__KG_") from SPACEXTABLE)
```

```
 * sqlite:///my_data1.db
Done.
```

Booster_Version	PAYLOAD_MASS__KG_
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

2015 Launch Records

- The launches with failed landings in 2015 occurred in January and April, involving boosters F9 v1.1 B1012 and B1015
- The code retrieved all rows where the Landing_Outcome column is equal to ‘Failure (drone ship)’ and the year extracted from the Date column is ‘2015’

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

Note: SQLLite does not support monthnames. So you need to use substr(Date, 6,2) as month to get the months and substr(Date,0,5)='2015' for year.

```
%sql select substr(Date,0,5) as year, substr(Date, 6,2) as month, landing_outcome, Booster_Version, Launch_Site from SPACEXTABLE
    where substr(Date,0,5)='2015' and Landing_Outcome = 'Failure (drone ship)';
```

```
 * sqlite:///my_data1.db
Done.
```

year	month	Landing_Outcome	Booster_Version	Launch_Site
2015	01	Failure (drone ship)	F9 v1.1 B1012	CCAFS LC-40
2015	04	Failure (drone ship)	F9 v1.1 B1015	CCAFS LC-40

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

- Rank the count of landing outcomes (such as Failure (drone ship) or Success (ground pad)) between the date 2010-06-04 and 2017-03-20, in descending order

- The code retrieves all rows from SPACEXTABLE where the Date falls between '2010-06-04' and '2017-03-20'. It then groups the results by landing_outcome, counts the number of rows for each outcome, and displays the results in descending order based on the count.

```
%sql select landing_outcome, count(*) as count1 from SPACEXTABLE where Date between '2010-06-04' and '2017-03-20'
    group by landing_outcome order by count1 desc;
```

```
 * sqlite:///my_data1.db
Done.
```

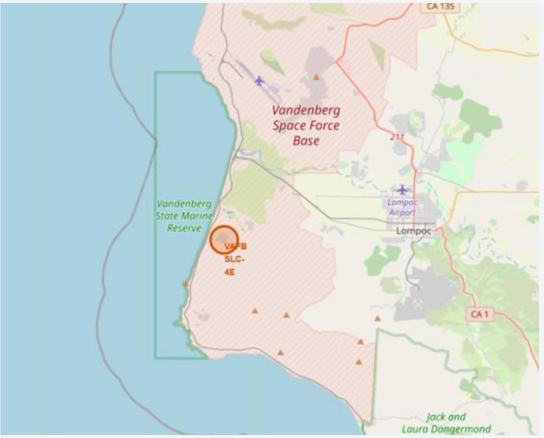
Landing_Outcome	count1
No attempt	10
Success (drone ship)	5
Failure (drone ship)	5
Success (ground pad)	3
Controlled (ocean)	3
Uncontrolled (ocean)	2
Failure (parachute)	2
Precluded (drone ship)	1

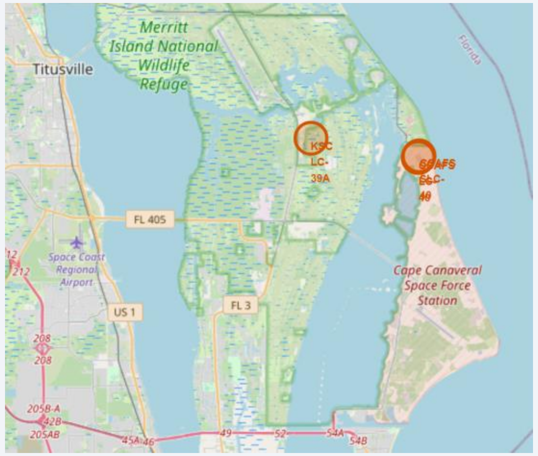
Section 3

Launch Sites Proximities Analysis

Falcon 9 Launch Site Map

- The left map provides an overview of all four sites in a single view. The middle map zooms in on the location of Vandenberg Space Force Base (VAFB SLC-4E), while the right map highlights the locations of Cape Canaveral Space Force Station (CCAFS SLC-40/CCAFS LC-40) and Kennedy Space Center (KSC LC-39A).
- All launch sites are in proximity to the Equator line, and in very close proximity to the coast

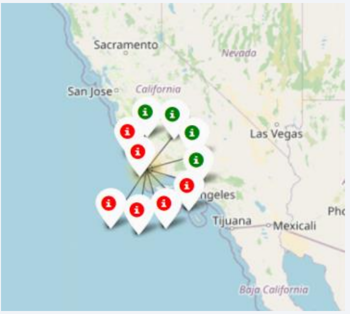




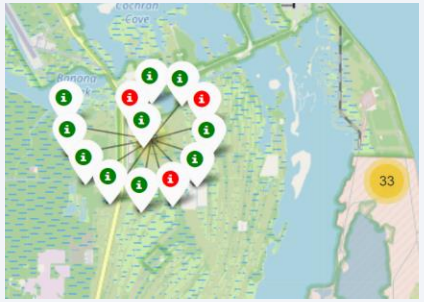
Landing Outcome by Launch Site

- The landing outcomes for each site are displayed on the following map. A red marker indicates failure, while a green marker indicates success.
- A Folium.Icon is added to folium.Marker and marker_cluster using a color identifier to differentiate success and failure.

Vandenberg Space Force Base
(VAFB SLC-4E)



Kennedy Space Center
(KSC LC-39A)



Cape Canaveral Space Force Station
CCAFS SLC-40

Cape Canaveral Space Force Station
CCAFS LC-40

Distance to Proximity

- Cape Canaveral Space Force Station is 0.86 km from the closest coastline , 20.2km from the center of city Cape Canaveral

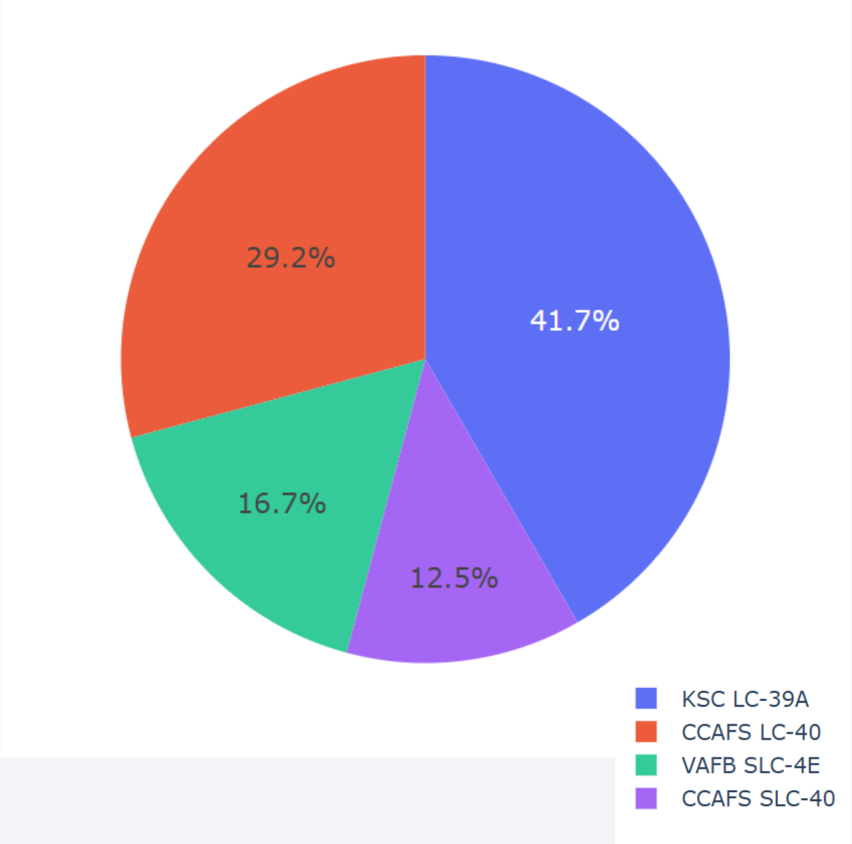
- A location extractor function is used to determine the longitude and latitude of the location where the mouse is pointing. A predefined function calculates the distance between the launch site and the point of interest (coastline) based on their coordinates. folium.PolyLine is used to draw a line between the two locations on the map, and the calculated distance is displayed using DivIcon, which allows HTML text to appear at the specified location

Section 4

Build a Dashboard with Plotly Dash

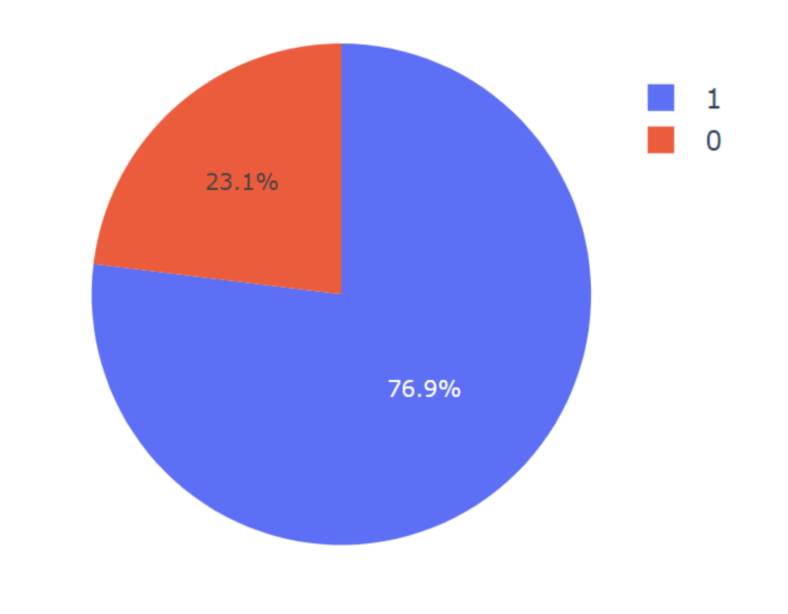
Launch Success Count for All Sites

- dcc.Dropdown is used to create a drop-down list of all sites. When "All Sites" is selected, a pie chart displays the composition of success counts across all sites. If one of the four specific site names is selected, the pie chart shows the comparison between success and failure counts for that site
- KSC LC-39A accounts for 41.7% of all successful launches, followed by CCAFS LC-40 at 29.2%, CCAFS SLC-40 at 12.5%, and VAFB SLC-4E at 16.7%.



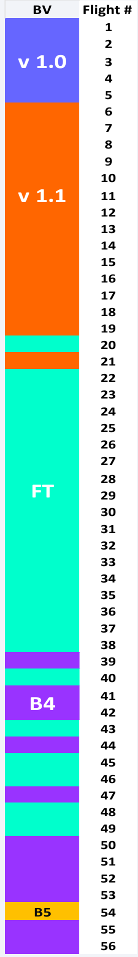
Highest Launch Success Ratio

- KSC LC-39A has the highest launch success ratio among all sites at 76.9%



Launch Outcome by Payload Range and Booster Version

- Class 1 represents Success, while class 0 represents Failure
- Booster FT has a significantly higher success rate in the 2,000–4,000 kg payload range compared to other ranges
- Booster B4 are mainly used in later launches and has significantly higher success rate (83%) in under 5000 kg payload range
- Booster version v1.1 and v1.0 were used in early flights(1-21) and consistently showed a low success rate across all payload ranges
- There are few launches with payloads in the 5,500–9,000 kg range, and they have a very low success rate too

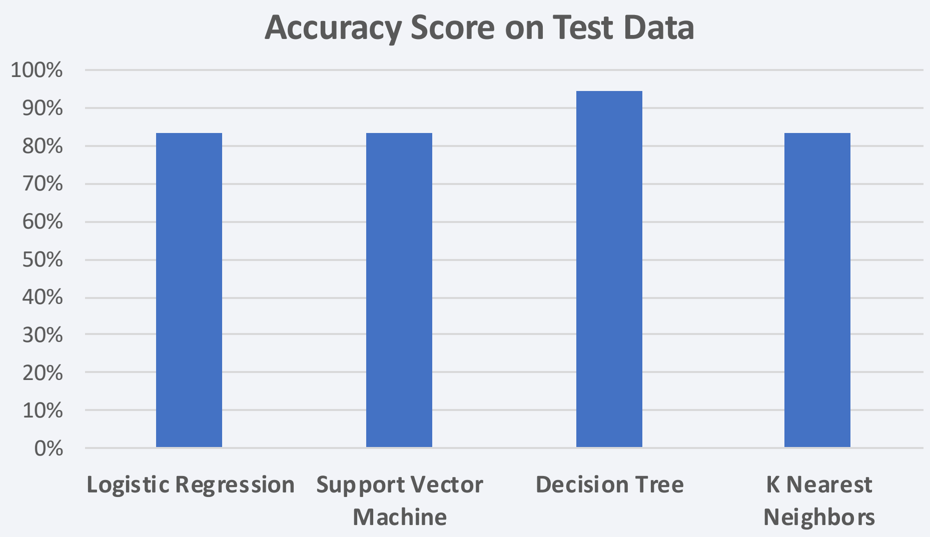


Section 5

Predictive Analysis (Classification)

Classification Accuracy

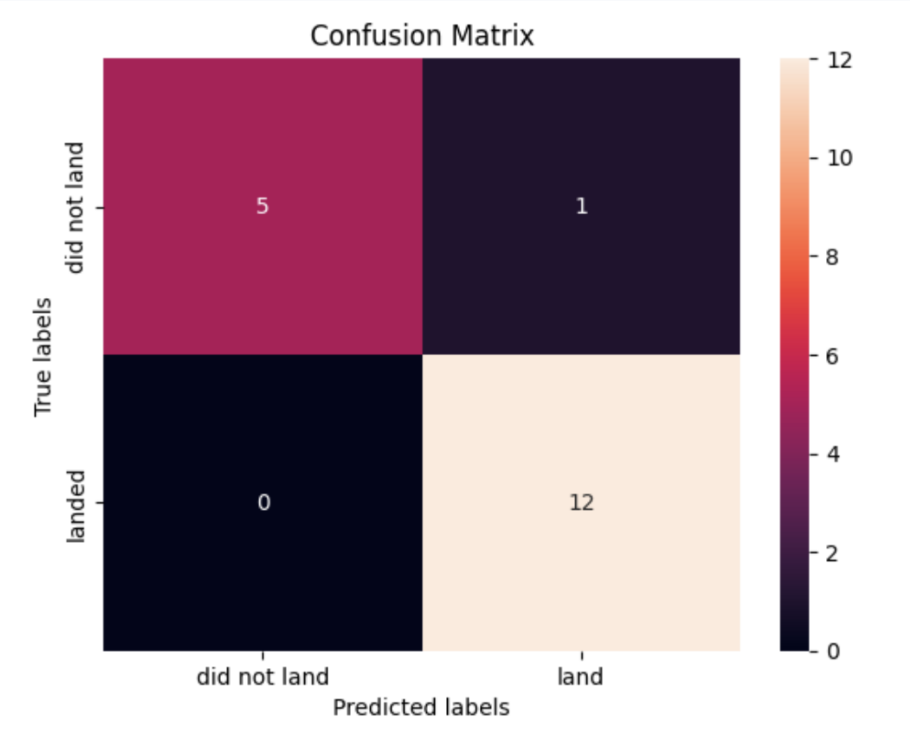
- The independent variables include: 'FlightNumber', 'PayloadMass', 'Orbit', 'LaunchSite', 'Flights', 'GridFins', 'Reused', 'Legs', 'LandingPad', 'Block', 'ReusedCount', and 'Serial'. Among these, categorical variables such as 'Orbit', 'LaunchSite', 'LandingPad', and 'Serial' were converted into dummy variables for analysis
- There are 90 launches in total. A random training and test split results in 72 launches for training and 18 launches for testing.
- All four models were tuned using GridSearchCV to identify the optimal parameters
- Logistic Regression, Support Vector Machine, and K-Nearest Neighbors achieved a similar level of accuracy on the test data
- The Decision Tree model outperformed the others, achieving the highest accuracy score of 94.4% on the test data



{

Confusion Matrix

- The best parameter found by GridSearchCV is : 'criterion': 'gini', 'max_depth': 16, 'max_features': 'sqrt', 'min_samples_leaf': 4, 'min_samples_split': 2, 'splitter': 'random’}
- Applying the best parameters to the test data, it achieved 94.4% accuracy score, with 92.3% precision and 100% recall

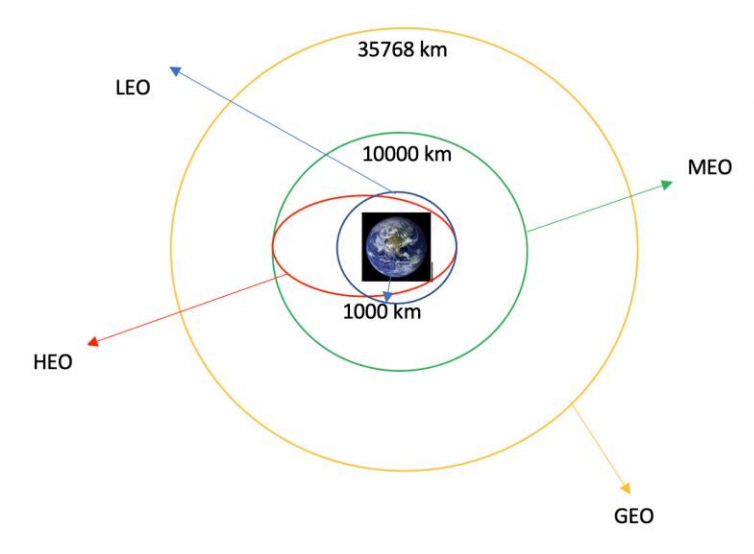


Conclusions

- The landing success rate has improved significantly over the years, primarily due to advancements at Cape Canaveral SLC-40, which handles the majority of launches. This improvement is also evident across most orbit types
- The primary orbit type has shifted from GTO and ISS in earlier flights to VLEO in later flights. VLEO missions typically carry heavier payloads and have a significantly higher landing success rate
- Heavier payloads show a higher landing success rate. However, these launches also occurred in later flights, which had an overall improved landing rate. Therefore, it makes more sense to attribute the higher success rate to time and overall advancements rather than just the payload weight
- Boost FT, B4, and B5 are progressively newer booster versions over time. Within the timeframe of each type, certain payload ranges show higher landing success rates than others. For example, Booster FT has a significantly higher success rate in the 2,000 - 4,000 kg payload range, and Booster B4 has a much higher success rate (83%) for payloads under 5,000 kg
- Among the four machine learning algorithms (Logistic Regression, Support Vector Machine, Decision Trees, and K-Nearest Neighbors), the Decision Tree model performed the best compared to the other algorithms used in this analysis. It achieved an accuracy score of 94.4% on the test data
- The Decision Tree model demonstrated better predictive performance on the test data compared to the other three algorithms used in this analysis. However, due to the relatively small dataset, there is a risk of overfitting

Appendix

- **LEO**: Low Earth orbit (LEO)is an Earth-centred orbit with an altitude of 2,000 km (1,200 mi) or less (approximately one-third of the radius of Earth)
- **VLEO**: Very Low Earth Orbits (VLEO) can be defined as the orbits with a mean altitude below 450 km.
- **GTO** A geosynchronous orbit is a high Earth orbit that allows satellites to match Earth's rotation.
- **SSO (or SO)**: It is a Sun-synchronous orbit also called a heliosynchronous orbit is a nearly polar orbit around a planet, in which the satellite passes over any given point of the planet's surface at the same local mean solar time
- **ES-L1** :At the Lagrange points the gravitational forces of the two large bodies cancel out in such a way that a small object placed in orbit there is in equilibrium relative to the center of mass of the large bodies. L1 is one such point between the sun and the earth
- **HEO** A highly elliptical orbit, is an elliptic orbit with high eccentricity, usually referring to one around Earth
- **ISS** A modular space station (habitable artificial satellite) in low Earth orbit. It is a multinational collaborative project between five participating space agencies: NASA (United States), Roscosmos (Russia), JAXA (Japan), ESA (Europe), and CSA (Canada)
- **MEO** Geocentric orbits ranging in altitude from 2,000 km (1,200 mi) to just below geosynchronous orbit at 35,786 kilometers (22,236 mi). Also known as an intermediate circular orbit.
- **HEO** Geocentric orbits above the altitude of geosynchronous orbit (35,786 km or 22,236 mi)
- **GEO** It is a circular geosynchronous orbit 35,786 kilometres (22,236 miles) above Earth's equator and following the direction of Earth's rotation
- **PO** It is one type of satellites in which a satellite passes above or nearly above both poles of the body being orbited (usually a planet such as the Earth
- From: https://labs.cognitiveclass.ai/v2/tools/jupyterlab?ulid=ulid-3ee2ae0b10b4321e8e614ad076b2bf3ed4892296



Thank you!