SpaceX Falcon 9 First Stage Landing Prediction

A Dissertation

Submitted by

Nagendra N

AA.SC.P2MCA2301019

in partial fulfilment of the requirements for the award of the degree of

## MASTER OF COMPUTER APPLICATIONS

## 

**February 2025**



**BONAFIDE CERTIFICATE**

This is to certify that this dissertation titled " **SpaceX Falcon 9 First Stage Landing Prediction**" submitted in partial fulfilment of the requirements for the award of the Degree of **Master of Computer Applications**, by **Nagendra N**, **AA.SC.P2MCA2301019**, is a bona fide record of the work carried out by him under my supervision during the academic year 2024-2025 and that it has not been submitted, to the best of my knowledge, in part or in full, for the award of any other degree or diploma.

Project Guide’s name:Coordinator's name

**Ms. Deepa Sreedhar** **Ms. Vidyalekshmi Vinod**

Reviewer

Date:

**DECLARATION**

I do hereby declare that this dissertation titled **" SpaceX Falcon 9 First Stage Landing Prediction** ", submitted in partial fulfilment of the requirements for the award of the degree of **Master of Computer Applications,** is a true record of work carried out by me and that all information contained herein, which do not arise directly from my work, have been properly acknowledged and cited, using acceptable international standards. Further, I declare that the contents of this thesis have not been submitted, in part or in full, for the award of any other degree or diploma.

Signature of the student

Date: 03/02/2025 **Nagendra N**

# Acknowledgements

I would like to express my sincere gratitude to the following individuals and organizations for their valuable support and contributions during the completion of this project:

1. **My Project Guide, Ms. Deepa Sreedhar** – For their expert guidance, support, and encouragement throughout this project. Their insights and feedback were invaluable in shaping the direction of the work and ensuring its success.
2. **IBM Data Science Professional Certificate Program** – For providing the essential knowledge, tools, and resources that allowed me to apply data science techniques and machine learning to this real-world problem. The program equipped me with the skills needed for this project, including Python, machine learning, and data analysis techniques.
3. **The Data Science Community** – For their open-source resources, tutorials, and valuable contributions on platforms such as GitHub, Stack Overflow, and Kaggle. Their collective knowledge helped me overcome challenges and deepen my understanding of the subject.
4. **Family and Friends** – For their unwavering support, patience, and encouragement throughout the duration of this project. Their belief in me helped me stay motivated and focused on completing the project.

Thank you all for your help and support.

**Abstract: -**

SpaceX has revolutionized the aerospace industry by pioneering reusable rocket technology, significantly reducing the cost of space launches. The Falcon 9 rocket, a flagship of SpaceX, exemplifies this cost efficiency by enabling the reuse of its first stage, cutting launch expenses from $165 million to $62 million. This advancement has disrupted the industry, compelling competitors and collaborators to reassess their strategies in the commercial space sector.

This project aims to develop a predictive model to determine whether the Falcon 9 first stage will successfully land. Accurate predictions will provide valuable insights for aerospace companies, enabling cost estimation and strategic planning in bidding against SpaceX for rocket launches. By leveraging historical launch data, including technical parameters, environmental conditions, and past outcomes, we will analyze key factors influencing successful landings.

To achieve this, machine learning techniques will be employed to construct a robust predictive model. This model will enhance understanding of reusable rocket feasibility and provide crucial insights for companies seeking to enter or compete in the commercial space launch industry. By combining domain knowledge in aerospace engineering with advanced data science methodologies, this project contributes to the evolving field of reusable space technologies. The findings will not only aid in cost estimation but also foster innovation and competition, shaping the future of space exploration.

Table of Contents

[Introduction: 7](#_Toc189491422)

[Course Overview: 7](#_Toc189491423)

[Project Overview: 9](#_Toc189491424)

[Problem Definition 10](#_Toc189491425)

[System Requirements: 12](#_Toc189491427)

[Methodology Summary 12](#_Toc189491429)

[Results and Analysis: 18](#_Toc189491430)

[Conclusions: 29](#_Toc189491431)

[References 30](#_Toc189491432)

[Appendix A: Source Code & Dataset Details 30](#_Toc189491433)

# 

# Introduction:

Since 1957, the countries around the world are competing to expand beyond Earth, whether through satellites launches or space exploration, these missions require huge amounts of money where a launch of one space rocket costs in average 165 million dollars, a company like Space X changes the equation by reducing this amount of money massively to only 60 million dollars due to its unique and advanced technologies in returning the first stage of rocket structure.

As mentioned above one of the key factor of SpaceX’s success in the space race is the first stage of rockets return through a safe landing, from this point we will discover together what are the main attributes or variables that control these successful landings, through asking questions about the nature of the launch process , the payload mass of rocket, launch location, rocket orbit, and more by analyzing and visualizing these information through the data science methodology provided by IBM.

# 

# Course Overview:

The **IBM Data Science Professional Certificate** is a comprehensive **12-course series** designed to equip learners with the skills, tools, and expertise required for a successful career in data science. This program provides a structured pathway to mastering data science techniques, from foundational concepts to advanced machine learning applications, with a strong emphasis on **hands-on learning and real-world projects**.

**Program Objectives**

* Develop proficiency in **data collection, cleaning, visualization, and analysis**.
* Learn to apply **predictive modeling and machine learning algorithms** to real-world scenarios.
* Gain expertise in **Python, SQL, and industry-standard data science tools**.
* Build a professional portfolio showcasing data-driven solutions to complex problems.
* Prepare for **entry-level data science roles** with practical, applied learning experiences.

**Skills and Tools Covered**

* **Data Science Skills**: Statistical analysis, predictive modeling, machine learning, data mining, data visualization, and AI applications.
* **Programming & Tools**: Python, SQL, Jupyter Notebooks, RStudio, GitHub, and IBM Watson Studio.
* **Libraries & Frameworks**: Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, SciPy, and ipython-sql.

**Applied Learning Approach**

The program follows a project-based learning approach, ensuring that students gain **practical experience** by working on real-world datasets. The hands-on labs, conducted on the IBM Cloud, provide exposure to industry-relevant challenges, enabling students to develop problem-solving skills and technical proficiency.

**Key Projects**

Throughout the program, learners will complete several **industry-relevant projects**, including:

* Extracting and visualizing **financial data** using Pandas.
* Using **SQL** to query census, crime, and school demographic datasets.
* Developing **regression models** to predict housing prices.
* Creating a **Python dashboard** to improve flight reliability.
* Applying **machine learning classification algorithms** to predict loan repayments.
* Training and comparing **multiple machine learning models** for performance evaluation.

**Course Structure**

The program is divided into **12 courses**, each focusing on a critical aspect of data science:

1. **What is Data Science?** (11 hours) – Introduction to the field, industry applications, and career opportunities.
2. **Tools for Data Science** (18 hours) – Overview of essential tools like Jupyter, RStudio, GitHub, and Watson Studio.
3. **Data Science Methodology** (6 hours) – Understanding data-driven decision-making frameworks.
4. **Python for Data Science, AI & Development** (25 hours) – Python programming fundamentals for data science.
5. **Python Project for Data Science** (8 hours) – Hands-on Python project to apply learned concepts.
6. **Databases and SQL for Data Science** (20 hours) – SQL for querying and managing large datasets.
7. **Data Analysis with Python** (15 hours) – Techniques for exploratory data analysis and statistical modeling.
8. **Data Visualization with Python** (20 hours) – Creating compelling data visualizations with Matplotlib and Seaborn.
9. **Machine Learning with Python** (20 hours) – Introduction to supervised and unsupervised learning models.
10. **Applied Data Science Capstone** (13 hours) – A final hands-on project applying all learned skills.
11. **Generative AI: Elevate Your Data Science Career** (12 hours) – Exploring the impact of AI in data science.
12. **Data Scientist Career Guide and Interview Preparation** (9 hours) – Preparing for job applications and interviews.

**Career Benefits**

Upon completing this program, learners will:

* Gain access to **IBM’s Talent Network**, receiving job recommendations and industry insights.
* Build a **strong portfolio** showcasing expertise in **data science, machine learning, and AI**.
* Earn **ACE® and FIBAA recommended college credits** (up to 12 college credits and 6 ECTS credits).
* Be well-prepared for **entry-level data science roles** in various industries.

The **IBM Data Science Professional Certificate** serves as a **stepping stone** into the high-demand field of data science, empowering learners with the skills and confidence to excel in the job market.

# Project Overview:

This project focuses on leveraging data science techniques to analyze and predict successful rocket launches. The aim is to assist a new rocket company, **SpaceY**, in competing with **SpaceX** by understanding the key factors that contribute to successful first-stage rocket landings.

The primary objectives of the project are:

1. **Data Collection**: Using SpaceX's API and web scraping Wikipedia pages for historical launch data.
2. **Data Processing & Wrangling**: Cleaning, transforming, and structuring the data for analysis.
3. **Exploratory Data Analysis (EDA)**: Visualizing launch trends, success rates, and key factors using tools like Seaborn, Matplotlib, and SQL queries.
4. **Predictive Analysis**: Implementing machine learning models to classify and predict the likelihood of a successful rocket landing.
5. **Interactive Visualization**: Using **Folium** for geospatial analysis and **Plotly Dash** for interactive dashboards.

**Key insights include:**

* The success rate of a first-stage return is influenced by **orbit type, payload mass, and launch location**.
* SpaceX's cost-saving model is mainly due to the **reusability** of its rockets.
* **Certain launch sites** show a higher probability of success.
* **Machine learning models** such as Logistic Regression, SVM, and KNN provide effective predictions.

# Problem Definition

Space exploration has been a costly endeavor, with traditional rocket launches averaging **$165 million per launch**. However, **SpaceX** has revolutionized the industry by reducing costs to **$60 million per launch** through reusable rocket technology. One of the most critical factors in cost reduction is the **successful landing of the rocket's first stage** for reuse.

A new competitor, **SpaceY**, wants to enter the market and **compete with SpaceX** by optimizing its rocket launch strategies. To do so, it needs to understand the factors that contribute to **successful first-stage landings**.

#### ****Business Problem****

The primary business question is: **"Can we use data science to predict whether a rocket's first-stage landing will be successful?"**

A successful landing directly impacts **cost reduction, reusability, and overall business viability**. The project aims to develop a **data-driven approach** to analyze launch data, identify key success factors, and build **predictive models** that estimate the probability of a successful landing.

#### ****Technical Problem Statement****

To address the business problem, the project focuses on:

1. **Data Collection**: Gathering **historical rocket launch data** from the **SpaceX API** and **Wikipedia**.
2. **Data Preprocessing & Wrangling**:
   * Cleaning and structuring data.
   * Handling missing values.
   * Standardizing and normalizing numerical data
3. **Exploratory Data Analysis (EDA)**:
   * Identifying patterns in **launch success, payload mass, launch sites, and orbit types**.
   * Understanding relationships between **flight numbers, orbit type, and landing success**.
4. **Predictive Analysis**:
   * Building **classification models** (Logistic Regression, SVM, Decision Tree, KNN).
   * Using **Grid Search for hyperparameter tuning**.
   * Evaluating models using **F1-score, Jaccard Score, and Confusion Matrix**.
5. **Interactive Visualizations**:
   * Using **Folium** to map launch sites and success probabilities.
   * Developing **Plotly Dash dashboards** for real-time business insights.

#### ****Key Challenges****

* **Data Quality**: Handling missing and inconsistent data.
* **Feature Engineering**: Identifying the most influential factors in landing success.
* **Model Selection**: Choosing the best machine learning model for accuracy and reliability.
* **Real-Time Decision Making**: Making the insights actionable for business use.

### ****Expected Outcome****

* A **predictive model** capable of estimating the **success probability** of a rocket's first-stage landing.
* Insights into **which launch sites, payload sizes, and orbit types** contribute to **higher landing success**.
* Interactive dashboards for **business intelligence and decision-making**.

This project helps **SpaceY** develop a data-driven **competitive strategy** in the commercial space industry. 🚀

# System Requirements:

### ****Hardware & Software Requirements****

#### ****Hardware Requirements****

* **Processor**: Minimum **Intel i5 or AMD Ryzen 5** (Recommended i7/Ryzen 7 for faster processing).
* **RAM**: At least **8GB** (Recommended **16GB** for handling large datasets).
* **Storage**: **20GB of free space** (SSD preferred for performance).
* **GPU** (Optional): Required if deep learning is incorporated.

#### ****Software Requirements****

* **Operating System**: Windows 10/11, macOS, or Linux.
* **Development Tools**:
  + **Jupyter Notebook** (Hosted on IBM Watson Studio or locally installed).
  + **Python 3.x** (Used for data processing, visualization, and modeling).
  + **Pandas & NumPy** (Data handling).
  + **Matplotlib & Seaborn** (Visualization).
  + **SQL** (For database querying).
  + **BeautifulSoup** (Web scraping).
  + **Folium** (Geospatial mapping).
  + **Plotly Dash** (Dashboard creation).
  + **Scikit-learn** (Machine Learning).

# Methodology Summary

**1. Data Collection**

• Making GET requests to the SpaceX REST API

• Web Scraping

**2. Data Wrangling**

• Using the. fillna() method to remove NaN values

• Using the .value\_counts() method to determine the following:

• Number of launches on each site

• Number and occurrence of each orbit

• Number and occurrence of mission outcome per orbit type

▪ Creating a landing outcome label that shows the following:

• 0 when the booster did not land successfully

• 1 when the booster did land successfully

**3. Exploratory Data Analysis**

• Using SQL queries to manipulate and evaluate the SpaceX

dataset

• Using Pandas and Matplotlib to visualize relationships between

variables, and determine patterns

**4. Interactive Visual Analytics**

• Geospatial analytics using Folium

• Creating an interactive dashboard using Plotly Dash

**5. Data Modelling and Evaluation**

• Using Scikit-Learn to:

• Pre-process (standardize) the data

• Split the data into training and testing data using train\_test\_split

• Train different classification models

• Find hyperparameters using GridSearchCV

▪ Plotting confusion matrices for each classification model

▪ Assessing the accuracy of each classification mode

**Date Collection:**

Data collection process involved a combination of API requests from SpaceX REST API and Web Scraping data from a table in SpaceX’s Wikipedia entry.

We had to use both of these data collection methods in order to get complete information about the launches for a more detailed analysis.

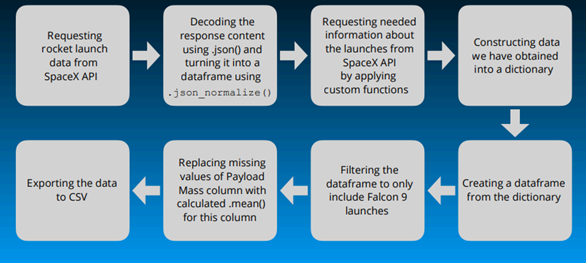
Data Columns are obtained by using SpaceX REST API:

FlightNumber, Date, BoosterVersion, PayloadMass, Orbit, LaunchSite, Outcome, Flights, GridFins, Reused, Legs, LandingPad, Block, ReusedCount, Serial, Longitude, Latitude.

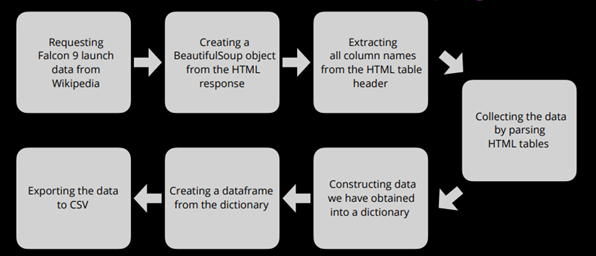
Data Columns are obtained by using Wikipedia Web Scraping:

Flight No., Launch site, Payload, PayloadMass, Orbit, Customer, Launch outcome, Version Booster, Booster landing, Date, Time.

**Data Collection-SpaceX API**

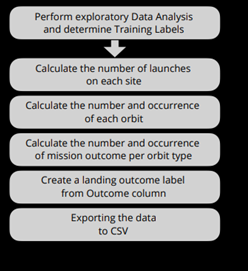


**Data collection-Web Scraping**



**Data Wrangling:**

In the data set, there are several different cases where the booster did not land successfully. Sometimes a landing was attempted but failed due to an accident; for example, True Ocean means the mission outcome was successfully landed to a specific region of the ocean while False Ocean means the mission outcome was unsuccessfully landed to a specific region of the ocean. True RTLS means the mission outcome was successfully landed to a ground pad False RTLS means the mission outcome was unsuccessfully landed to a ground pad. True ASDS means the mission outcome was successfully landed on a drone ship False ASDS means the mission outcome was unsuccessfully landed on a drone ship. We mainly convert those outcomes into Training Labels with “1” means the booster successfully landed, “0” means it was unsuccessful.



**EDA with Data Visualization:**

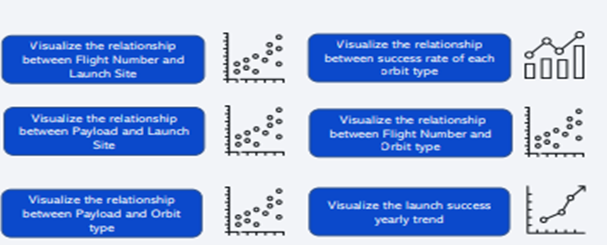
Charts were plotted:

Flight Number vs. Payload Mass, Flight Number vs. Launch Site, Payload Mass vs. Launch Site, Orbit Type vs. Success Rate, Flight Number vs. Orbit Type, Payload Mass vs Orbit Type and Success Rate Yearly Trend

Scatter plots show the relationship between variables. If a relationship exists, they could be used in machine learning model.

Bar charts show comparisons among discrete categories. The goal is to show the relationship between the specific categories being compared and a measured value.

Line charts show trends in data over time (time series)



**EDA with SQL:**

Performed SQL queries:

* Displaying the names of the unique launch sites in the space mission
* Displaying 5 records where launch sites begin with the string ‘CCA'
* Displaying the total payload mass carried by boosters launched by NASA (CRS)
* Displaying average payload mass carried by booster version F9 v1.1
* Listing the date when the first successful landing outcome in ground pad was achieved
* Listing the names of the boosters which have success in drone ship and have payload mass greater than 4000 but less than 6000
* Listing the total number of successful and failure mission outcomes
* Listing the names of the booster versions which have carried the maximum payload mass •Listing the failed landing outcomes in drone ship, their booster versions and launch site names for the months in year 2015
* Ranking 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.

**Built an interactive map with Folium:**

Markers of all Launch Sites:

Added Marker with Circle, Popup Label and Text Label of NASA Johnson Space Center using its latitude and longitude coordinates as a start location.

Added Markers with Circle, Popup Label and Text Label of all Launch Sites using their latitude and longitude coordinates to show their geographical locations and proximity to Equator and coasts.

Coloured Markers of the launch outcomes for each Launch Site:

Added coloured Markers of success (Green) and failed (Red) launches using Marker Cluster to identify which launch sites have relatively high success rates.

Distances between a Launch Site to its proximities:

Added coloured Lines to show distances between the Launch Site KSC LC-39A (as an example) and its proximities like Railway, Highway, Coastline and Closest City.

**Built a dashboard Plotly Dash:**

Launch Sites Dropdown List:

Added a dropdown list to enable Launch Site selection.

Pie Chart showing Success Launches (All Sites/Certain Site):

Added a pie chart to show the total successful launches count for all sites and the Success vs. Failed counts for the site, if a specific Launch Site was selected.

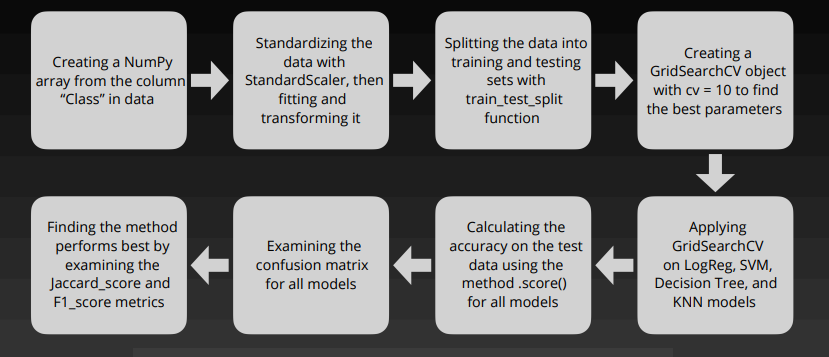
Slider of Payload Mass Range:

Added a slider to select Payload range.

Scatter Chart of Payload Mass vs. Success Rate for the different Booster Versions:

- Added a scatter chart to show the correlation between Payload and Launch Success.

**Predictive Analysis (Classification):**

****

# 

# 

# Results and Analysis:

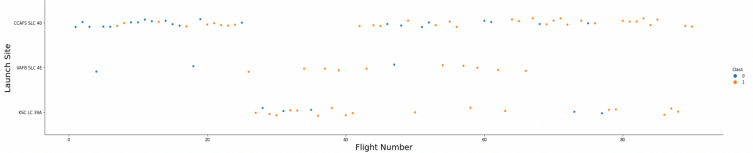
**Results:**

• Exploratory data analysis results

• Interactive analytics demo in screenshots

• Predictive analysis results

**Flight Number vs. Launch Site:**

****

**Explanation:**

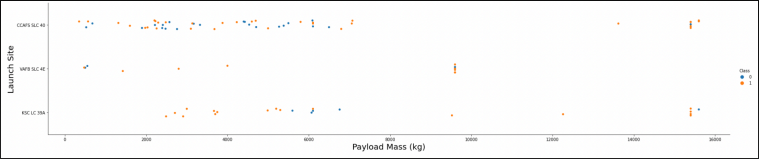
•The earliest flights all failed while the latest flights all succeeded.

•The CCAFS SLC 40 launch site has about a half of all launches.

•VAFB SLC 4E and KSC LC 39A have higher success rates.

•It can be assumed that each new launch has a higher rate of success.

**Payload vs. Launch Site**

****

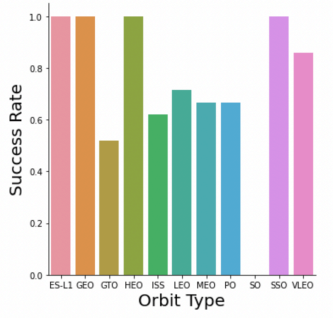
**Explanation:**

•For every launch site the higher the payload mass, the higher the success rate.

•Most of the launches with payload mass over 7000 kg were successful.

•KSC LC 39A has a 100% success rate for payload mass under 5500 kg too.

**Success Rate vs. Orbit Type**

****

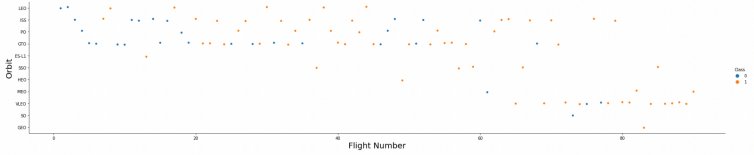
**Explanation:**

•Orbits with 100% success rate: -ES-L1, GEO, HEO, SSO

•Orbits with 0% success rate: -SO

•Orbits with success rate between 50% and 85%: -GTO, ISS, LEO, MEO, PO

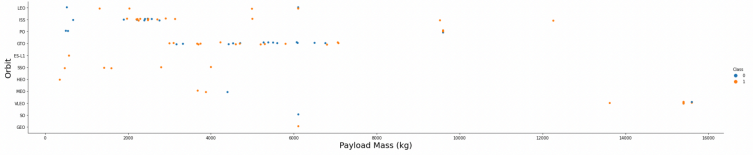
**Flight Number vs. Orbit Type**

****

**Explanation:**

•In the LEO orbit the Success appears related to the number of flights; on the other hand, there seems to be no relationship between flight number when in GTO orbit.

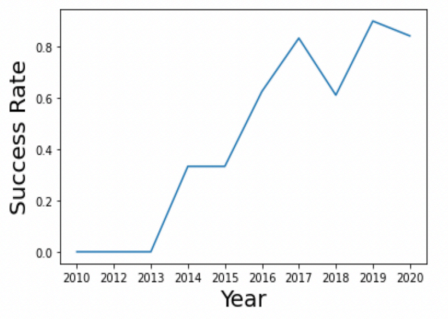
**Payload Mass vs. Orbit Type**

****

**Explanation:**

•Heavy payloads have a negative influence on GTO orbits and positive on GTO and Polar LEO (ISS) orbits

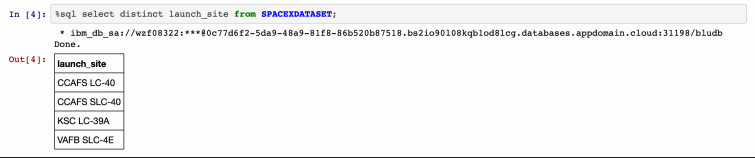
**Launch Success Yearly Trend**

****

**Explanation:**

We can observe that the success rate since 2013 kept increasing till 2020

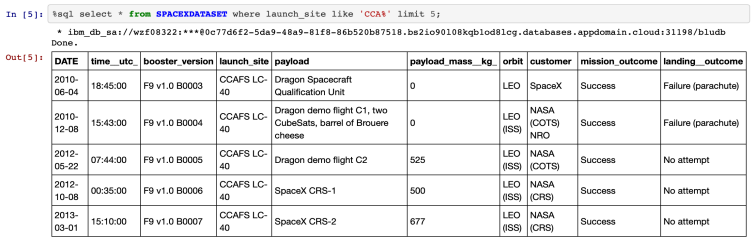
**All Launch Site Names**

****

**Explanation:**

•Displaying the names of the unique launch sites in the space mission

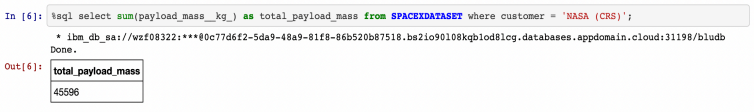
**Launch Site Names Begin with 'CCA'**

****

**Explanation:**

•Displaying 5 records where launch sites begin with the string 'CCA' .

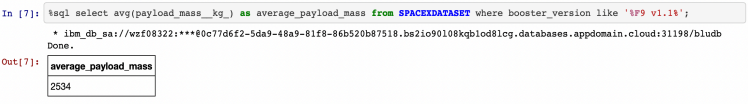
**Total Payload Mass**

****

**Explanation:**

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

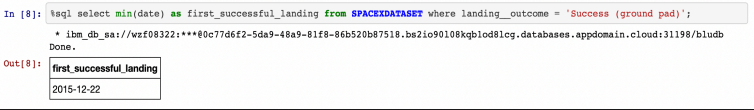
**Average Payload Mass by F9 v1.1**

****

**Explanation:**

•Displaying average payload mass carried by booster version F9 v1.1

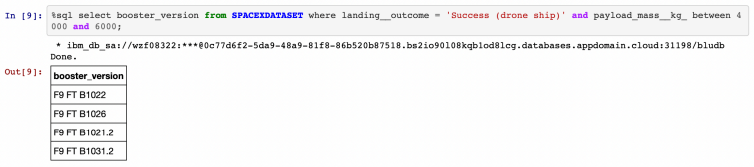
**First Successful Ground Landing Date**

****

**Explanation:**

•Listing the date when the first successful landing outcome in ground pad was achieved

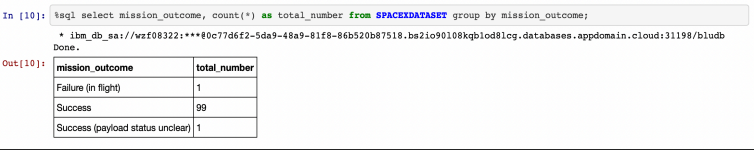
**Successful Drone Ship Landing with Payload between 4000 and 6000**

****

**Explanation:**

•Listing 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**

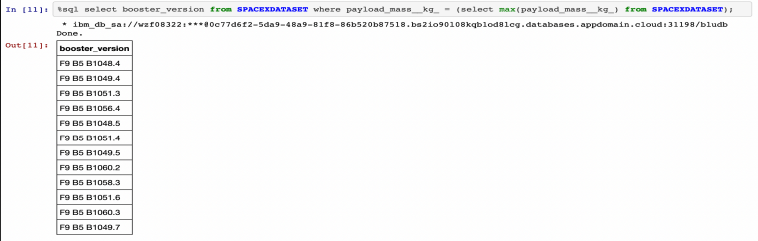
****

**Explanation:**

•Listing the total number of successful and failure mission outcomes

We can see clearly that the success rate of mission outcomes is the most dominant we have only 1 failed mission while we have 99 successful ones.

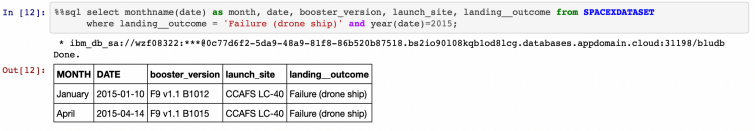
**Boosters Carried Maximum Payload**

****

**Explanation:**

•Listing the names of the booster versions which have carried the maximum payload mass. Further, it starts with F9 B5 and ranges from B1048 up to B1060

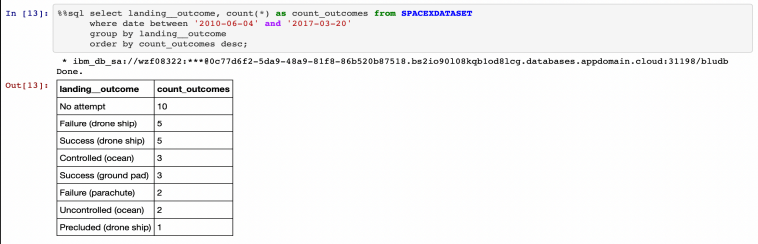
**2015 Launch Records**

****

**Explanation:**

•Listing the failed landing outcomes in drone ship, their booster versions and launch site names for the months in year 2015.

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

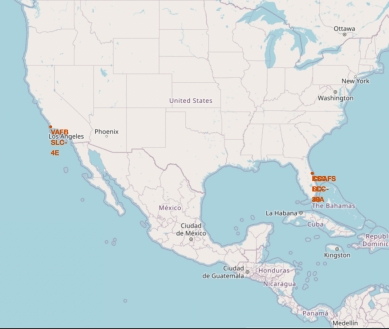
****

**Explanation:**

•Ranking 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.

Launch Sites proximities analysis:

**All launch sites’ location markers on a global map:**

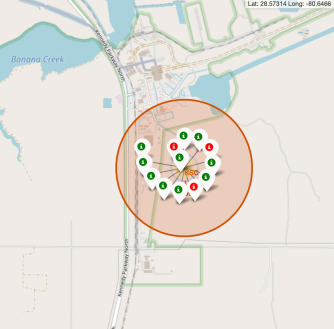
****

**Explanation:**

•Most of Launch sites are in proximity to the Equator line. The land is moving faster at the equator than any other place on the surface of the Earth. Anything on the surface of the Earth at the equator is already moving at 1670 km/hour. If a ship is launched from the equator it goes up into space, and it is also moving around the Earth at the same speed it was moving before launching. This is because of inertia. This speed will help the spacecraft keep up a good enough speed to stay in orbit.

•All launch sites are in very close proximity to the coast, while launching rockets towards the ocean it minimises the risk of having any debris dropping or exploding near people

**Colour-labeled launch records on the map:**

****

**Explanation:**

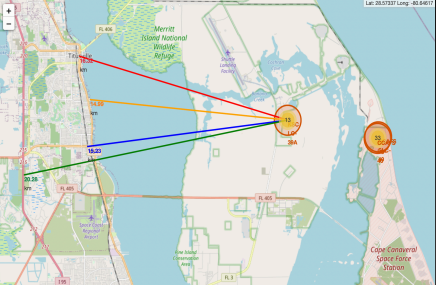
•From the colour-labeled markers we should be able to easily identify which launch sites have relatively high success rates.

-Green Marker = Successful Launch

-Red Marker = Failed Launch

•Launch Site KSC LC-39A has a very high Success Rate

**Distance from the launch site KSC LC-39A to its proximitie:**

****

**Explanation:**

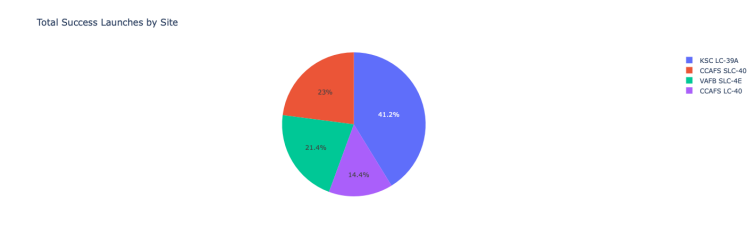
•From the visual analysis of the launch site KSC LC-39A we can clearly see that it is: -relative close to railway (15.23 km) -relative close to highway (20.28 km) -relative close to coastline (14.99 km)

•Also the launch site KSC LC-39A is relative close to its closest city Titusville (16.32 km).

•Failed rocket with its high speed can cover distances like 15-20 km in few seconds. It could be potentially dangerous to populated areas.

**Built a dashboard with plotly dash**

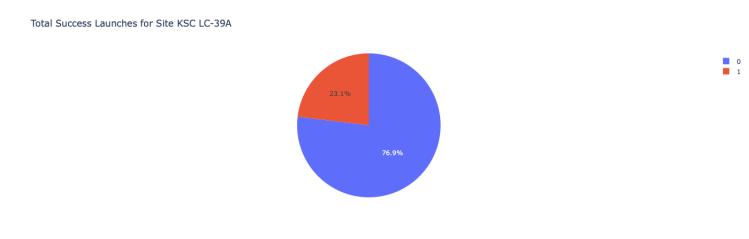
Launch success count for all sites:

****

**Explanation:**

•The chart clearly shows that from all the sites, KSC LC-39A has the most successful launches.

**Launch site with highest launch success ratio:**



**Explanation:**

•KSC LC-39A has the highest launch success rate (76.9%) with 10 successful and only 3 failed landings.

**Payload Mass vs. Launch Outcome for all sites:**

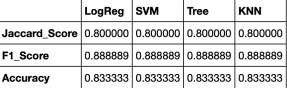


**Explanation:**

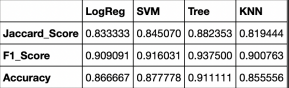
•The charts show that payloads between 2000 and 5500 kg have the highest success rate.

**Predictive Analysis (Classification):**

Scores and Accuracy of the Test Set



Scores and Accuracy of the Entire Data Set:



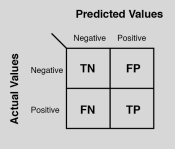
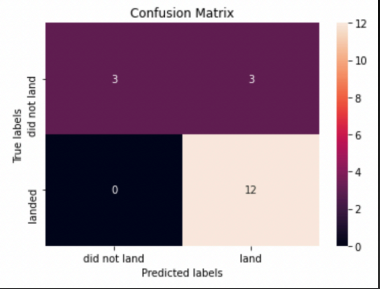
**Explanation:**

•Based on the scores of the Test Set, we can not confirm which method performs best.

•Same Test Set scores may be due to the small test sample size (18 samples). Therefore, we tested all methods based on the whole Dataset.

•The scores of the whole Dataset confirm that the best model is the Decision Tree Model. This model has not only higher scores, but also the highest accuracy.

**Confusion Matrix:**



**Explanation:**

•Examining the confusion matrix, we see that logistic regression can distinguish between the different classes. We see that the major problem is false positives.

# Conclusions:

•Decision Tree Model is the best algorithm for this dataset.

•Launches with a low payload mass show better results than launches with a larger payload mass. •Most of launch sites are in proximity to the Equator line and all the sites are in very close proximity to the coast.

•The success rate of launches increases over the years.

•KSC LC-39A has the highest success rate of the launches from all the sites.

•Orbits ES-L1, GEO, HEO and SSO have 100% success rate

# References

Courses & Tutorials

* IBM Data Science Professional Certificate – Coursera: [Click Here](https://www.coursera.org/programs/ahead-mca-2023-april-major-project-ug7j8/professional-certificates/ibm-data-science?collectionId=P6O7x)

Online Data Sources & APIs

* SpaceX API URL [Click Here](https://github.com/r-spacex/SpaceX-API)
* SpaceX Static Wikipedia URL [Click Here](https://en.wikipedia.org/wiki/SpaceX)
* SpaceX data used in ML training [Click Here](https://www.kaggle.com/)

# Appendix A: Source Code & Dataset Details

**A1: Source Code**

The complete source code for this project is available on GitHub. The repository includes scripts for:

* Data collection from the SpaceX API.
* Data preprocessing and feature engineering.
* Machine learning model training and evaluation.
* Prediction script for Falcon 9 landing success.

GitHub Repository: [**Click Here**](https://github.com/nagendra-1999/Coursera-IBM-Data-Science.git)

**A2: Dataset Details**

The dataset used in this project was sourced from the following locations:

* SpaceX API URL [Click Here](https://github.com/r-spacex/SpaceX-API)
* SpaceX Static Wikipedia URL [Click Here](https://en.wikipedia.org/wiki/SpaceX)
* SpaceX data used in ML training [Click Here](https://www.kaggle.com/)