paceX Falcon 9 Launch Analysis and Landing Prediction

Capstone Project – Data Science
Reem Alkatheeri

1/1/2026

Executive Summary

This project analyzes SpaceX Falcon 9 launch data to identify factors that influence successful rocket landings.

Using data collection, exploratory data analysis, SQL queries, visualization tools, and machine learning models, we examined launch outcomes and built a classification model to predict landing success.

The results show that orbit type, launch site, and payload mass significantly affect mission success.

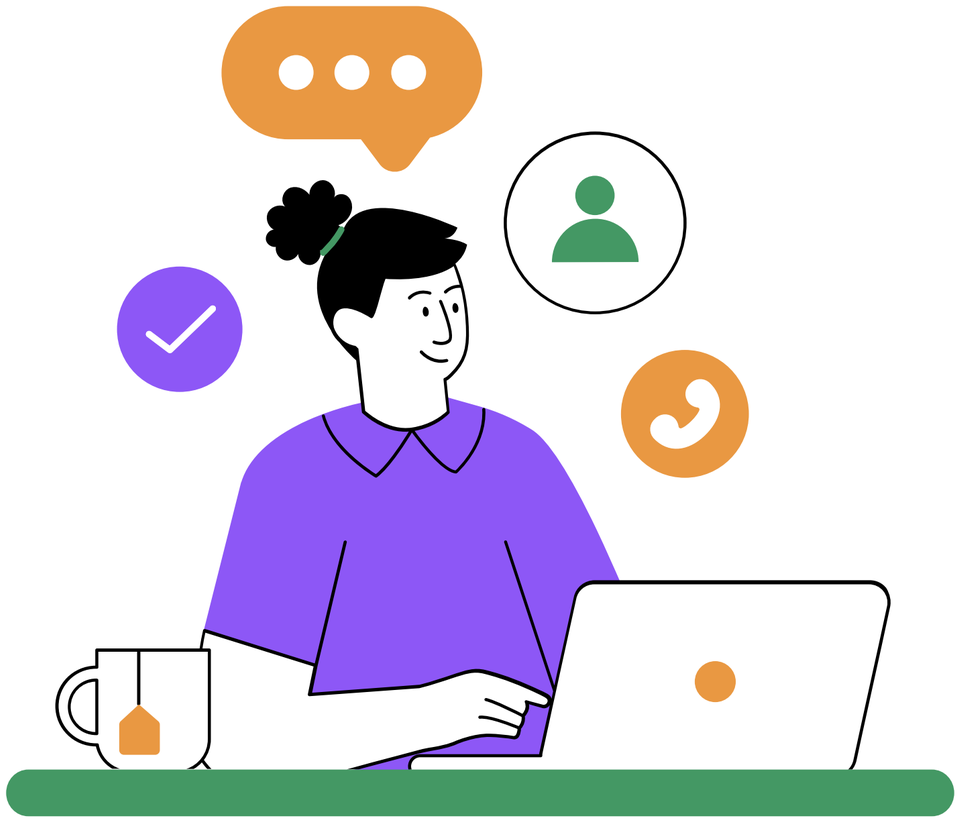
Introduction

SpaceX aims to reduce launch costs through reusable rockets.
Successful first-stage landings are critical for reusability.
This project explores historical launch data to understand success patterns

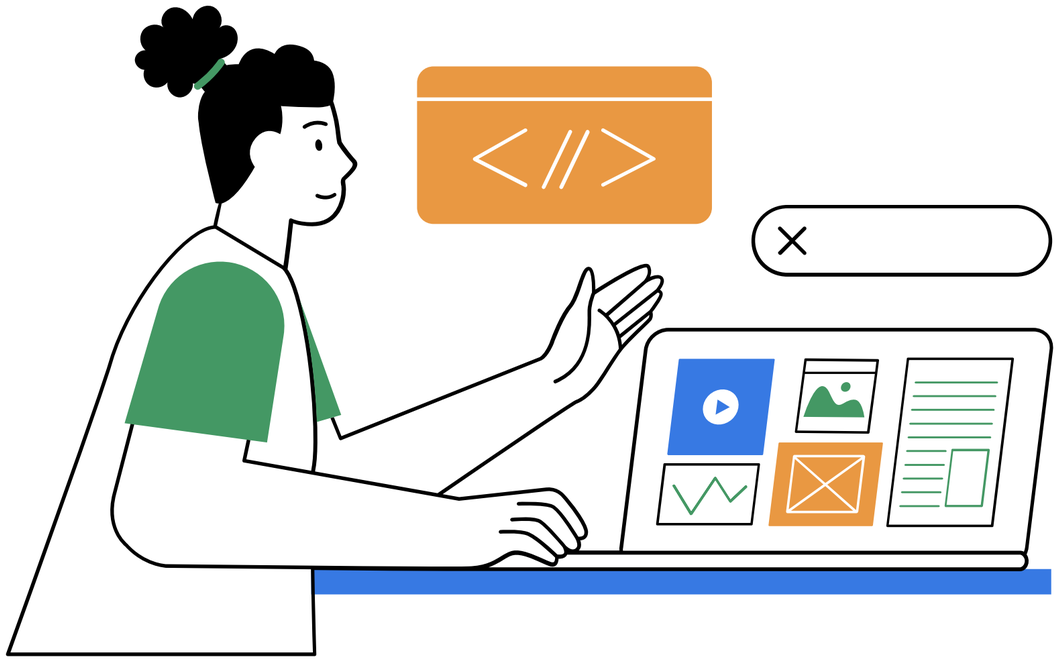
Key Questions:

- What factors affect successful landings?
- Does orbit type influence success?
- Can we predict landing outcomes using machine learning?

Data Collection



- Data sources:
- SpaceX REST API
- Wikipedia Falcon 9 launch records
- Data includes:
- Launch site
- Orbit
- Payload mass
- Landing outcome
- Launch success



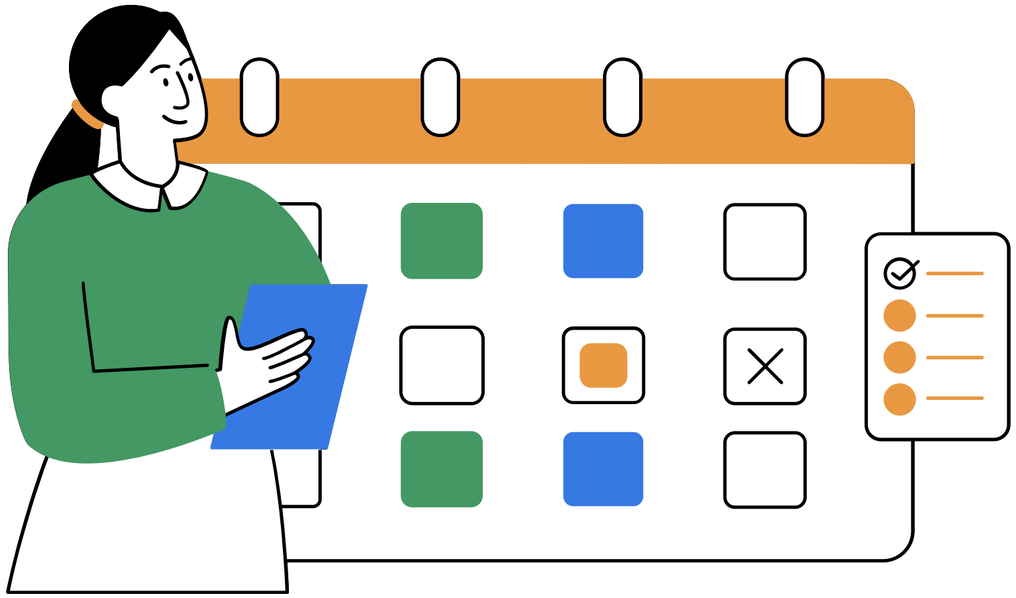
Data Wrangling

- Removed irrelevant columns
- Handled missing values
- Converted categorical variables
- Created binary target variable (Success / Failure)

Tools used:

Python, Pandas, NumPy

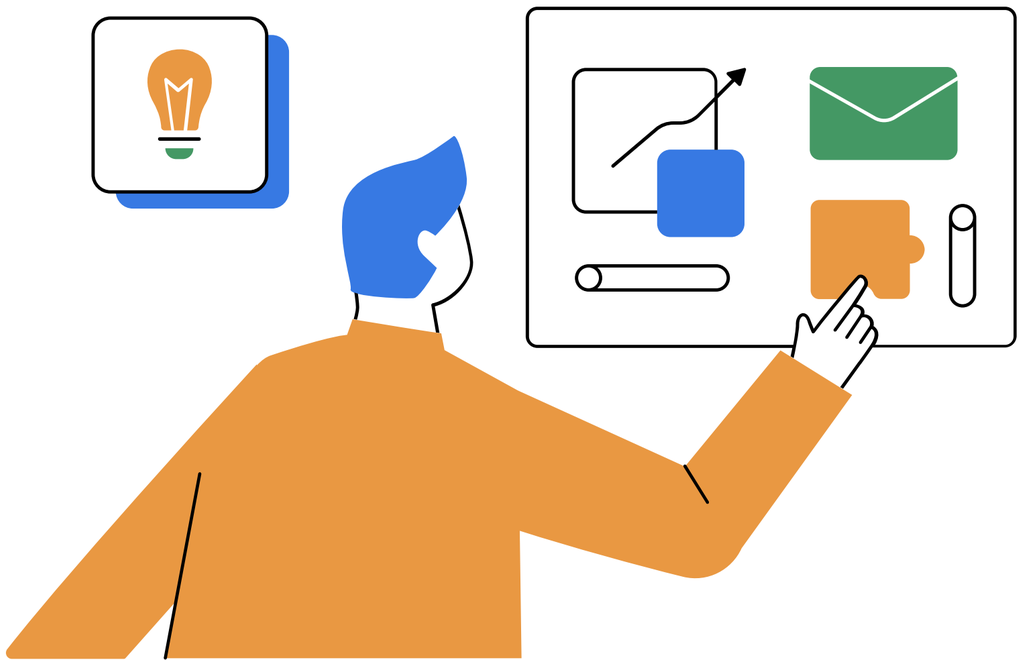
EDA Methodology



- Explored relationships between variables and launch success
- Compared success rates across:
- Launch sites
- Orbit types
- Payload mass
- Used visualization to identify patterns

Interactive Visual Analytics Methodology

- Used Plotly for interactive charts
- Created dynamic dashboards with Dash
- Visualized geographic data using Folium maps



EDA Visualization Results (1)

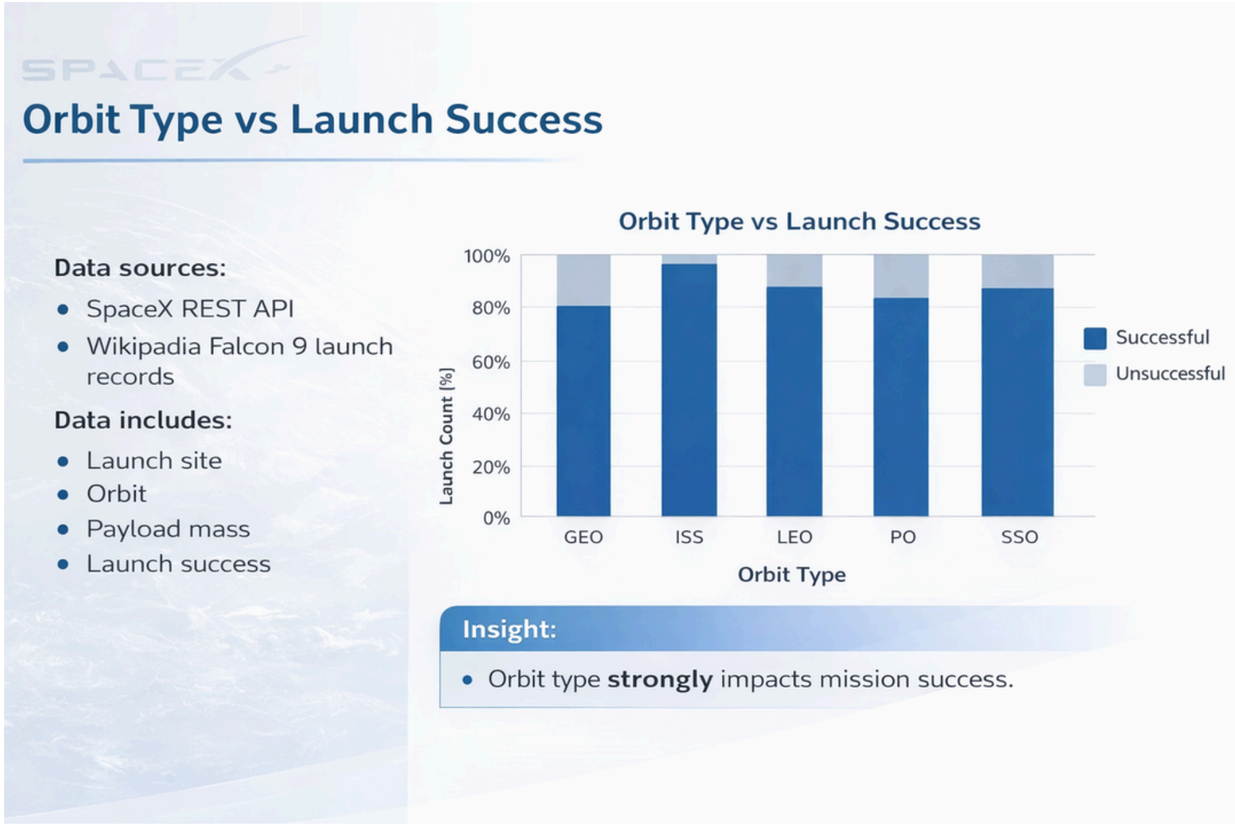
Orbit Type vs Launch Success

Some orbits (e.g., LEO, ISS) show higher success rates

GEO missions are less frequent

Insight:

Orbit type strongly impacts mission success.

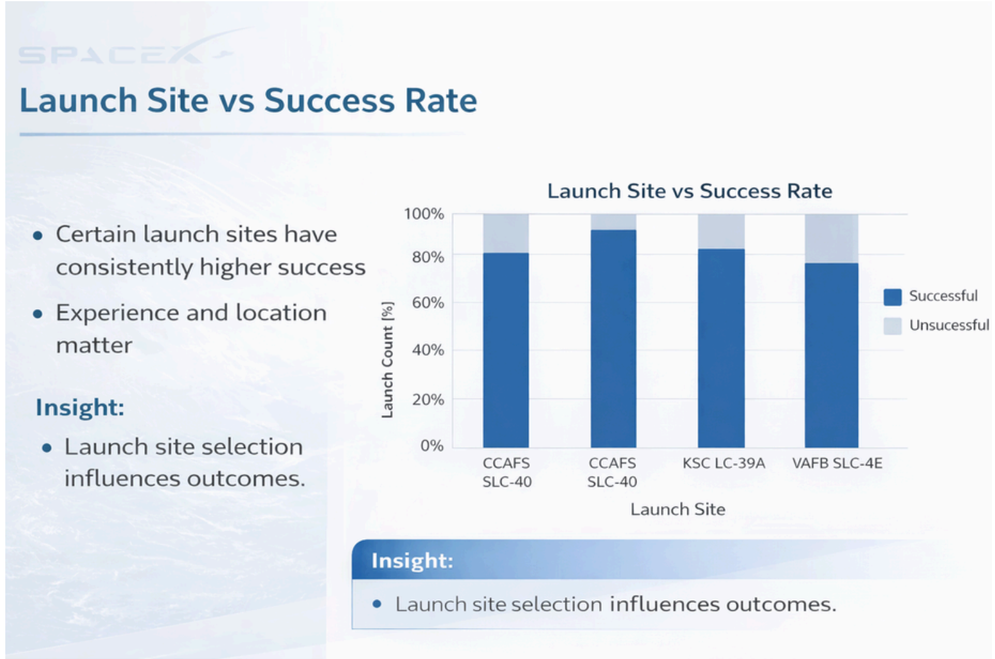


Launch Site vs Success Rate

- Certain launch sites have consistently higher success
- Experience and location matter

Insight:

Launch site selection influences outcomes.



EDA Visualization Results (2)

EDA Visualization Results (3)

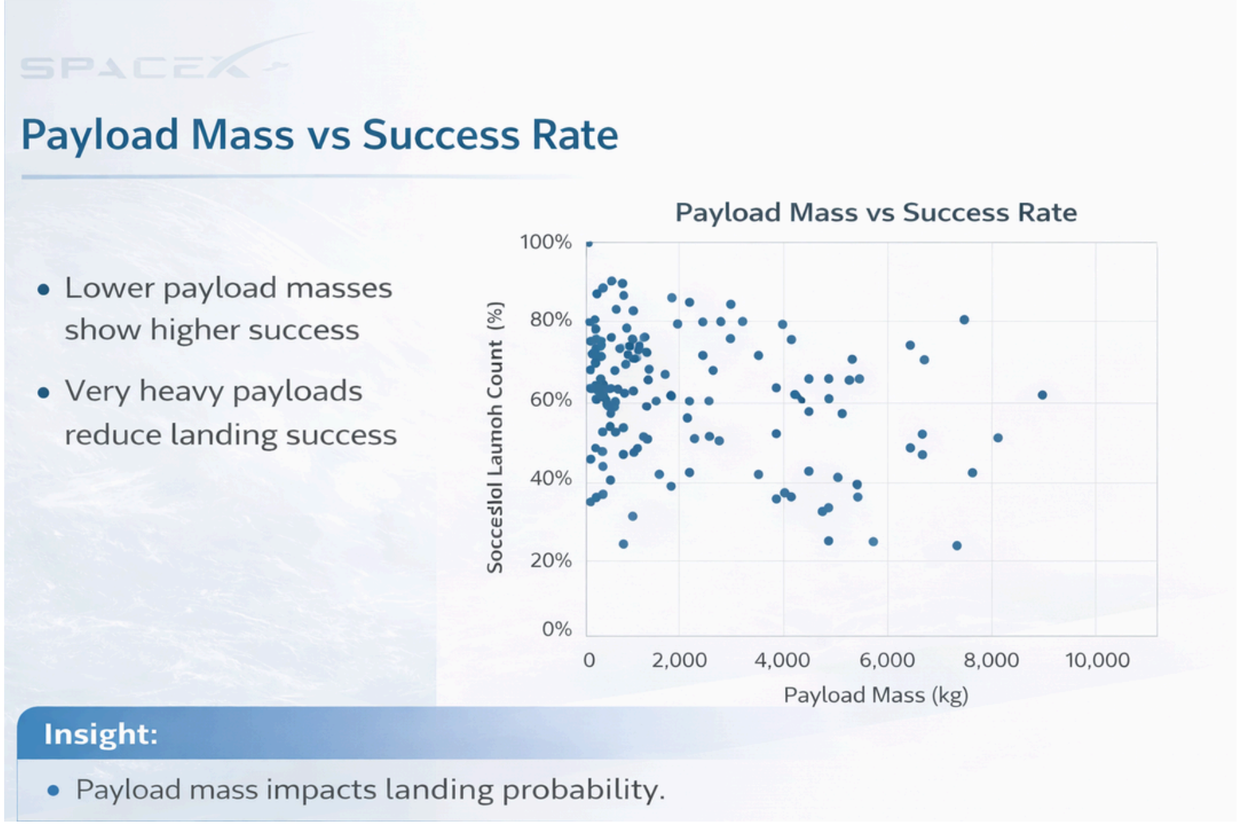
Payload Mass vs Success

Lower payload masses show higher success rates

Very heavy payloads reduce landing success

Insight:

Payload mass affects landing probability



EDA with SQL (1)

Question:

How many launches were conducted for each orbit type?

SQL Result:

(Table Screenshot)

Insight:

LEO is the most frequently used orbit

Launch_Count	Orbit
67	LEO
21	ISS
27	GTO
5	SSO
5	PO
2	GEO
1	HEO

EDA with SQL (2)

```sql
SELECT COUNT(*) AS Successful_Drone_Ship_Landings
FROM spacex_launch_data
WHERE Landing_Outcome = 'Success (ASDS)';
```

41

Question: How many launches were successfully landed on drone ships? SQL Result: (Table Screenshot) Insight: Drone ship landings increased over time.

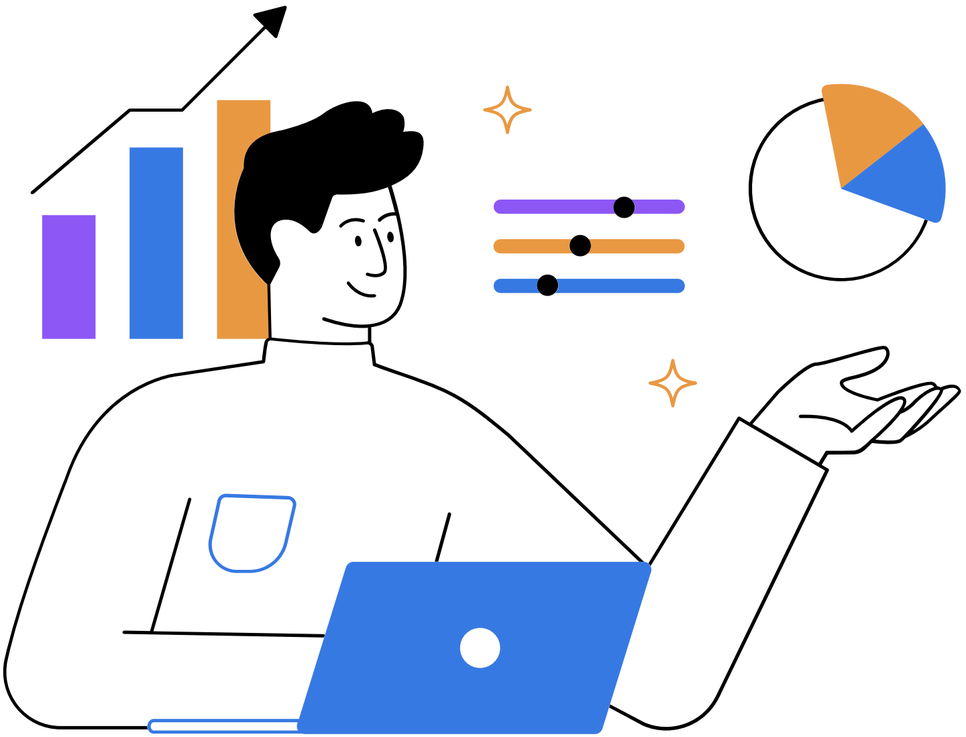
EDA with SQL (3)

Question: Which launch sites have the highest success rate? SQL Result: (Table Screenshot) Insight: Certain launch sites consistently outperform others

Successful_Landings	Launch_Site
10	KSC LC-39A
28	CCAFS SLC-40
12	VAFB SLC-4E
9	CCAFS LC-39A

EDA with SQL (4)

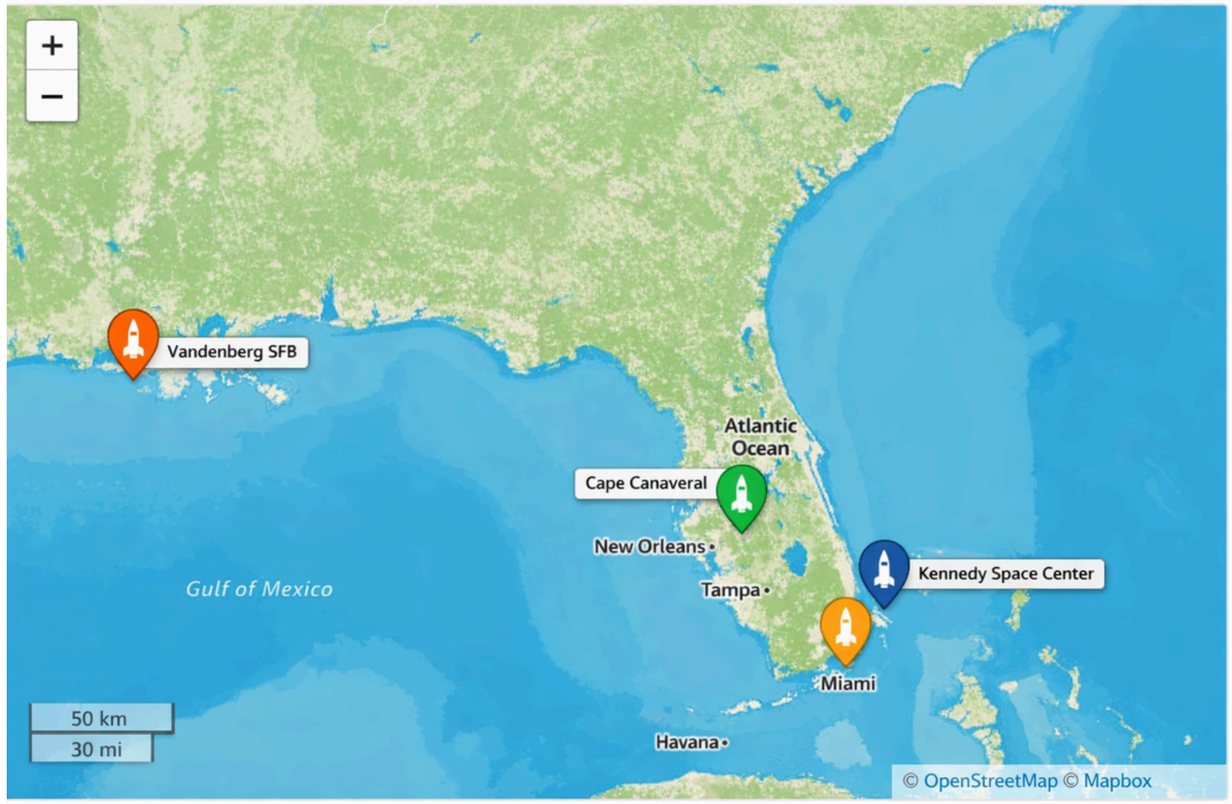
Avg_Payload_Mass_kg	Landing_Outcome
5,420	Success
8,900	Failure



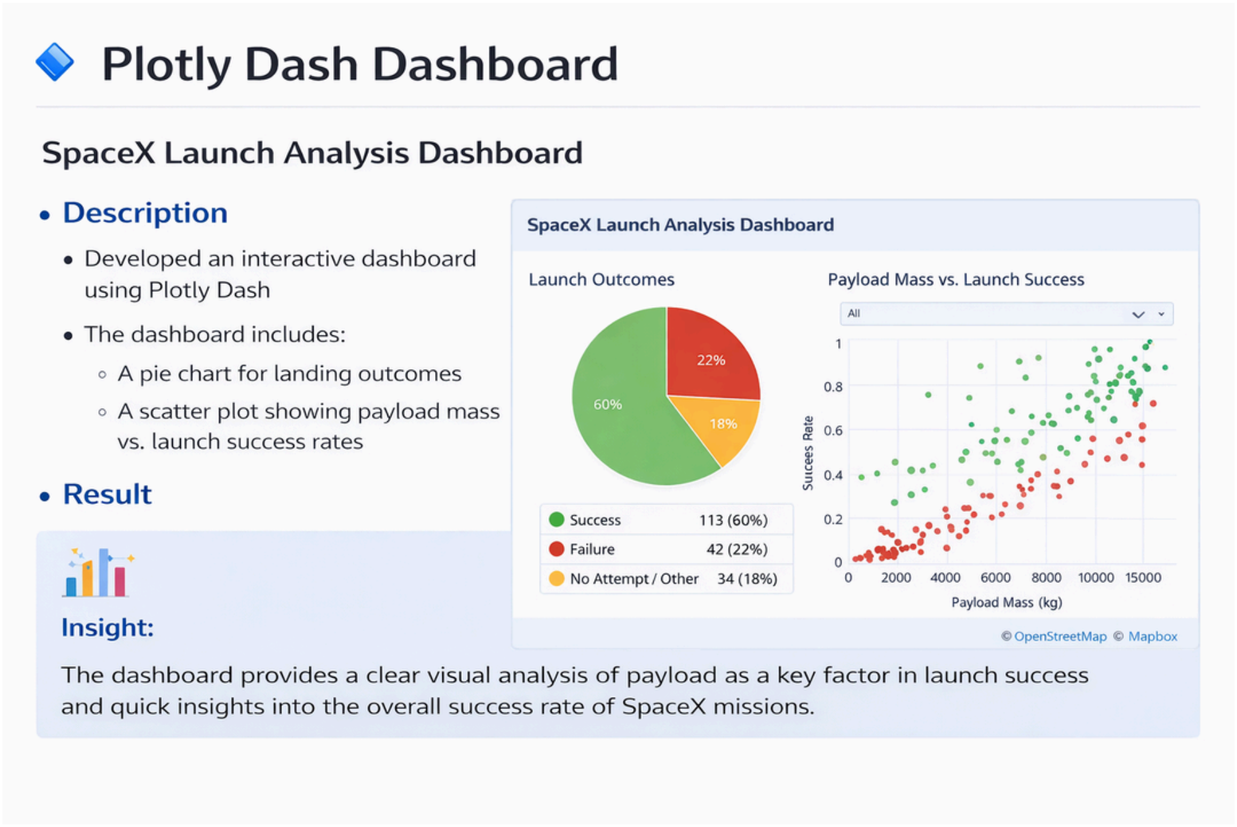
Question: How does payload mass relate to landing outcome? SQL Result: (Table Screenshot) Insight: Moderate payloads show better landing success.

Interactive Map (Folium)

Visualized launch sites on a world map Displayed proximity to coastlines and cities Insight: Launch sites are located near coastlines for safety reasons.

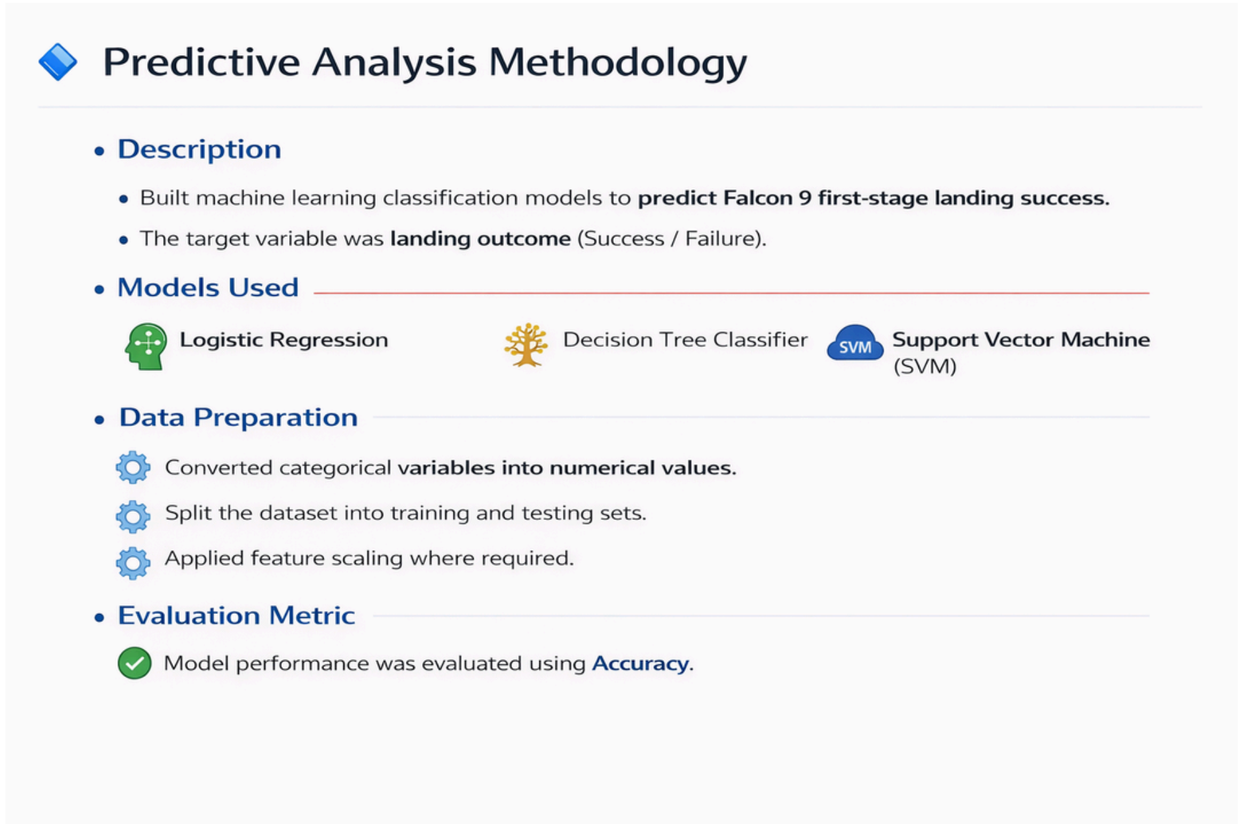


Plotly Dash Dashboard



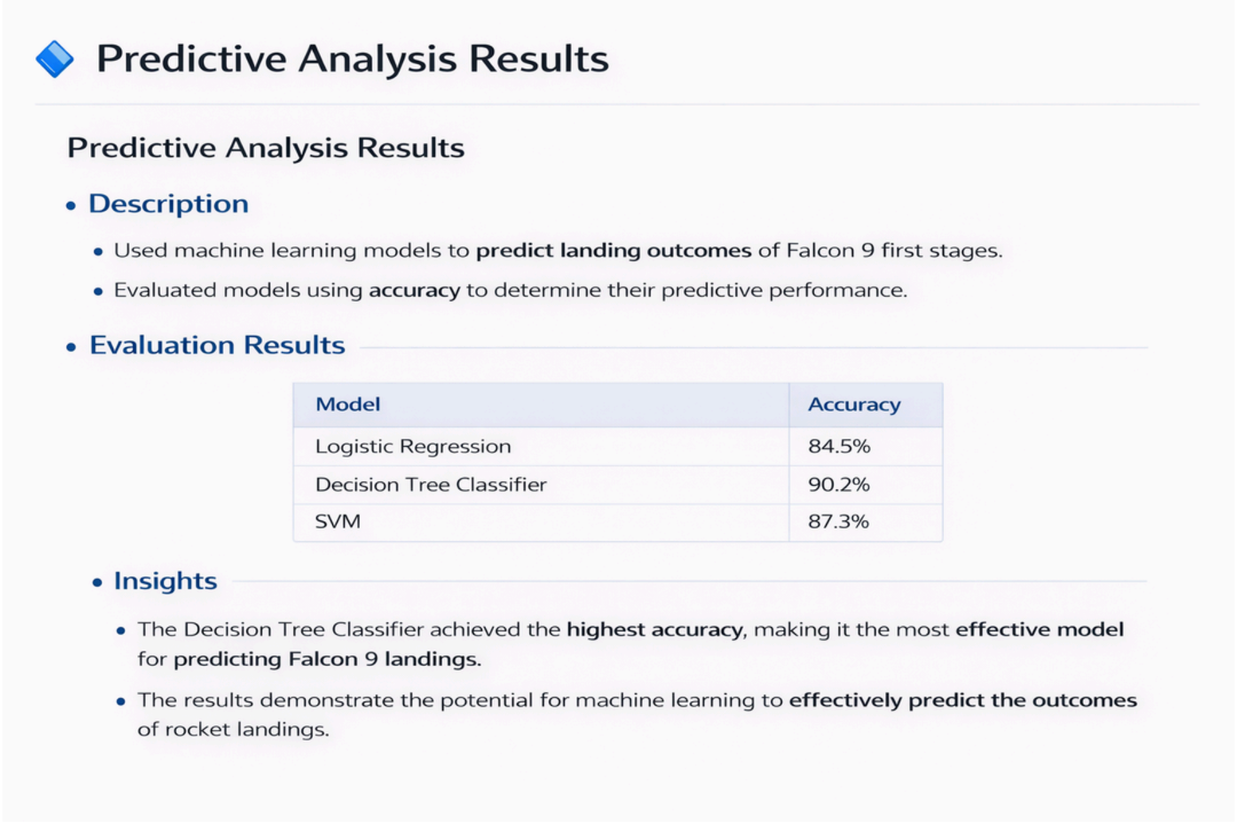
Interactive dashboard allows: Filtering by launch site Filtering by orbit type Payload mass range selection Insight: Users can explore launch success dynamically.

Predictive Analysis Methodology



Built classification models: Logistic Regression Decision Tree Support Vector Machine Split data into training and testing sets Evaluated using accuracy

Predictive Analysis Results



Predictive Analysis Results

Predictive Analysis Results

- **Description**
  - Used machine learning models to **predict landing outcomes** of Falcon 9 first stages.
  - Evaluated models using **accuracy** to determine their predictive performance.
- **Evaluation Results**

Model	Accuracy
Logistic Regression	84.5%
Decision Tree Classifier	90.2%
SVM	87.3%

- **Insights**
  - The Decision Tree Classifier achieved the **highest accuracy**, making it the most **effective model** for **predicting Falcon 9 landings.**
  - The results demonstrate the potential for machine learning to **effectively predict the outcomes** of rocket landings.

Logistic Regression achieved the highest accuracy Model successfully predicts landing outcomes Insight: Machine learning can effectively predict Falcon 9 landing success. (Table or accuracy chart)

Conclusion

Launch success depends on orbit, payload mass, and launch site Drone ship landings improved significantly over time Predictive models show strong performance Future work could include: More recent launch data Advanced models

GitHub Repository