IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Hilmi
July 2025

Outline

- Executive Summary
- Introduction
- Methodology
- Results
- Conclusion
- Appendix

Executive Summary

This project explores SpaceX Falcon 9 launch data to identify key factors influencing booster landing success. We performed exploratory data analysis using Python, SQL-like queries, and built a classification model. We also developed interactive visuals using Folium and Plotly Dash. The final model achieved promising accuracy, highlighting the importance of payload mass, launch site, and orbit type in predicting success.

Introduction

Reusable rockets are crucial for reducing space launch costs. This project investigates SpaceX Falcon 9 missions to answer:

- What patterns influence a successful landing?
- How do payload mass, launch site, and orbit impact outcomes?

Section 1

Methodology

Methodology

- SpaceX API:
- Used the requests library to call the SpaceX REST API
- Parsed JSON responses to extract relevant launch data
- Stored results into pandas DataFrames
- Web Scraping
- Used BeautifulSoup to scrape Falcon 9 payload data from Wikipedia
- Cleaned scraped HTML tables and merged them into the main DataFrame

Data Collection and Wrangling

We used:

- **REST API & Web Scraping** to gather launch metadata
- **CSV Dataset** with 89 Falcon 9 missions (used here)
- Cleaned column names, converted datatypes, and added a binary success label
- Feature engineered the year and categorized mission attributes

EDA & Visual Analytics Methodology

We used:

- pandas, matplotlib, seaborn for static visualizations
- folium for launch site mapping
- plotly dash for building an interactive dashboard

Section 2

Insights drawn from EDA

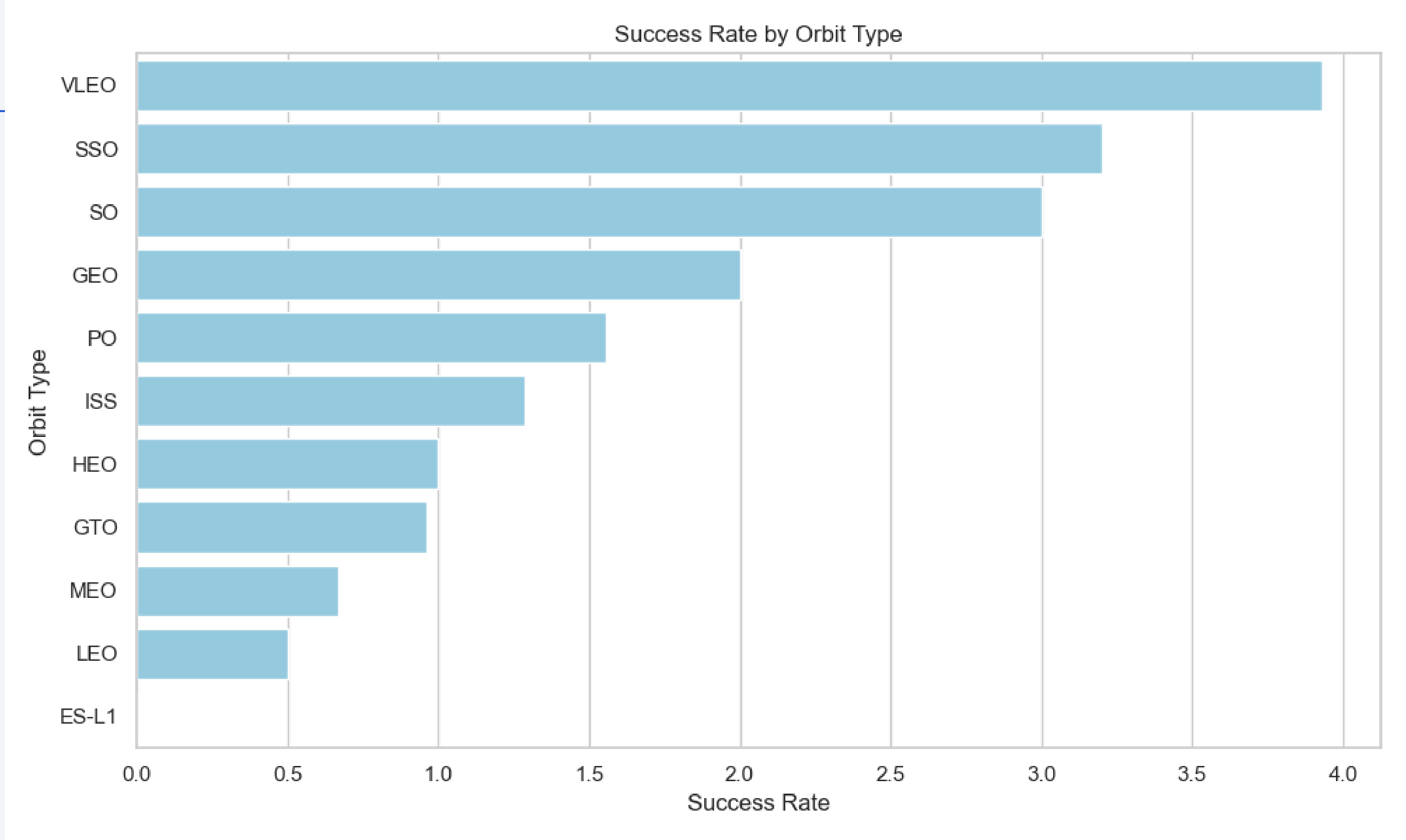
Key Insights

Top Launch Sites:

- CCAFS SLC 40: 54 launches
- KSC LC 39A: 21 launches

Success Rate by Orbit:

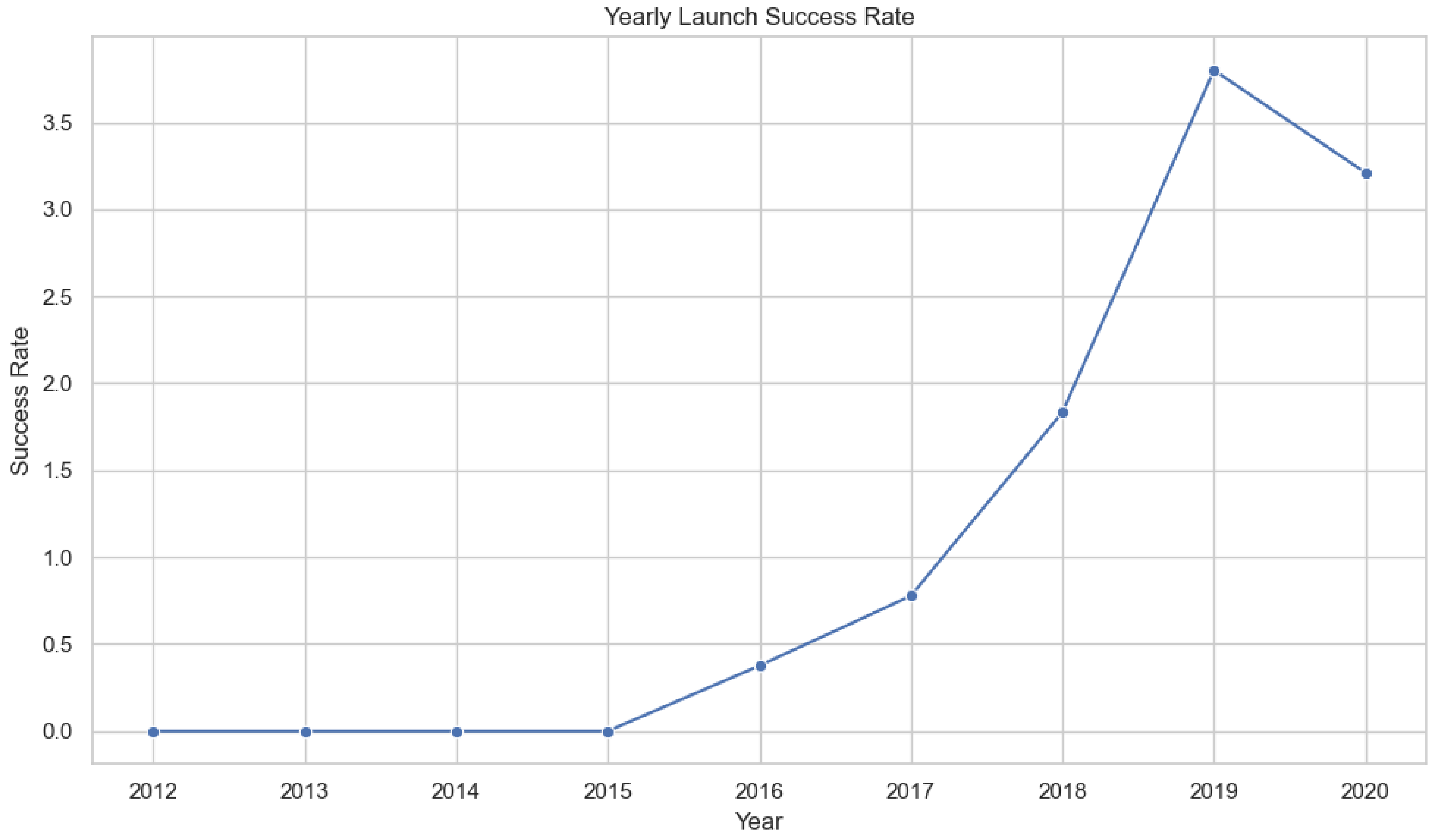
- LEO: 0.95, VLEO: 1.0, GTO: 0.70



Key Insights

Yearly trend:

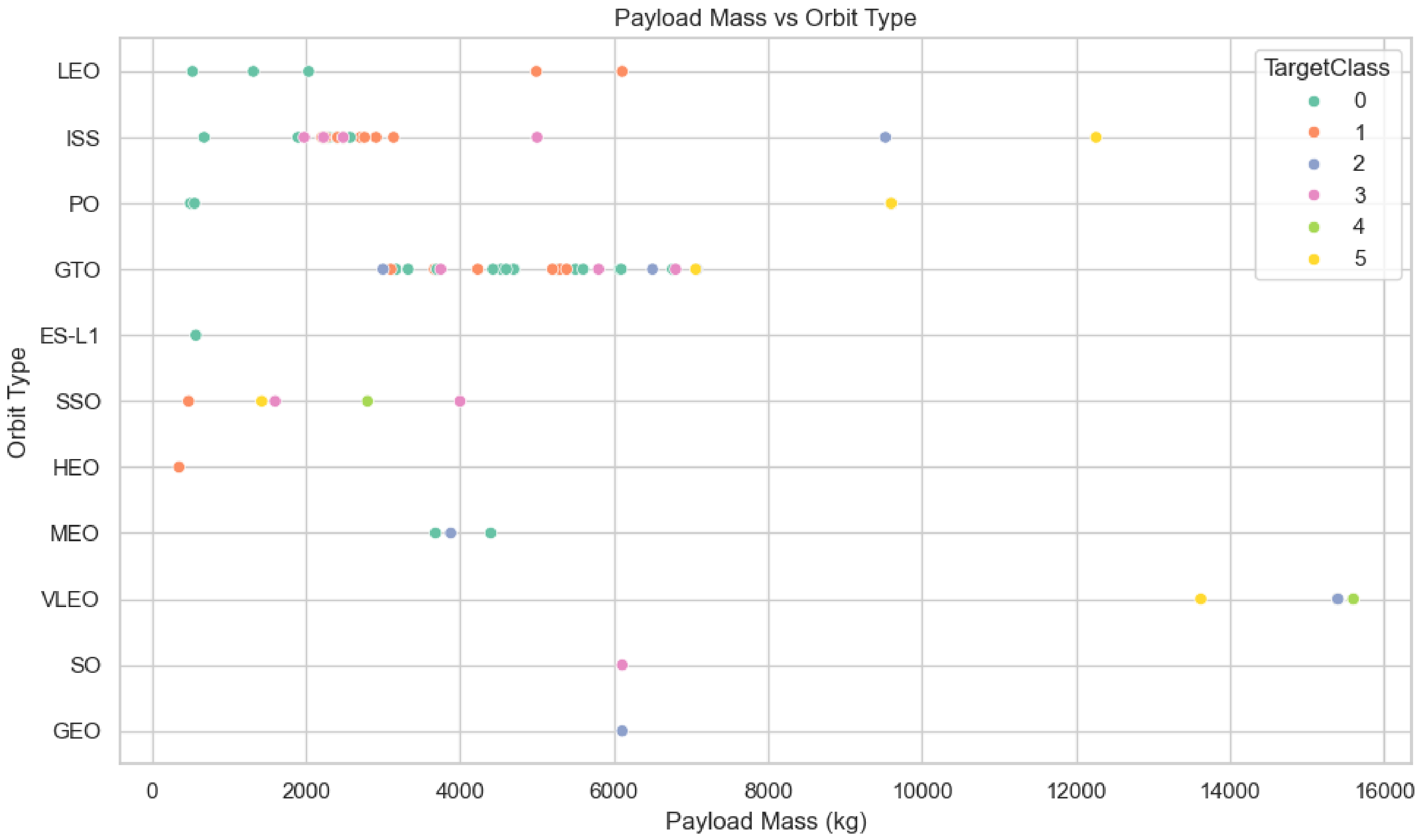
- Success rate rose from ~0.4 (2013) to ~0.9+ after 2018



Key Insights

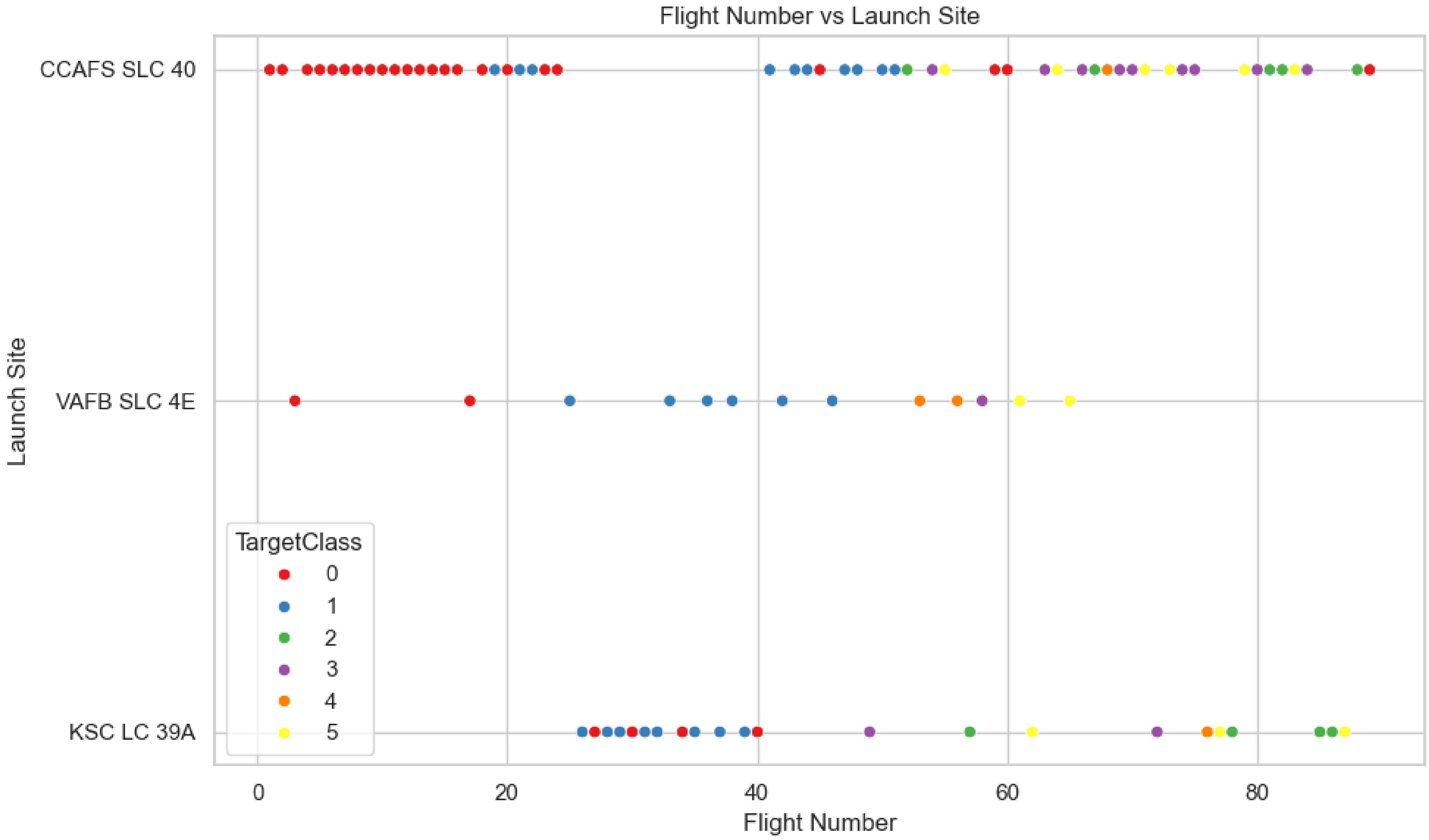
Payload Mass:

- Range: 350 – 15600 kg, Avg: 6105 kg
- Heavier payloads had lower success unless orbit was LEO



Key Insights

Launch Frequency by (Supplementary)

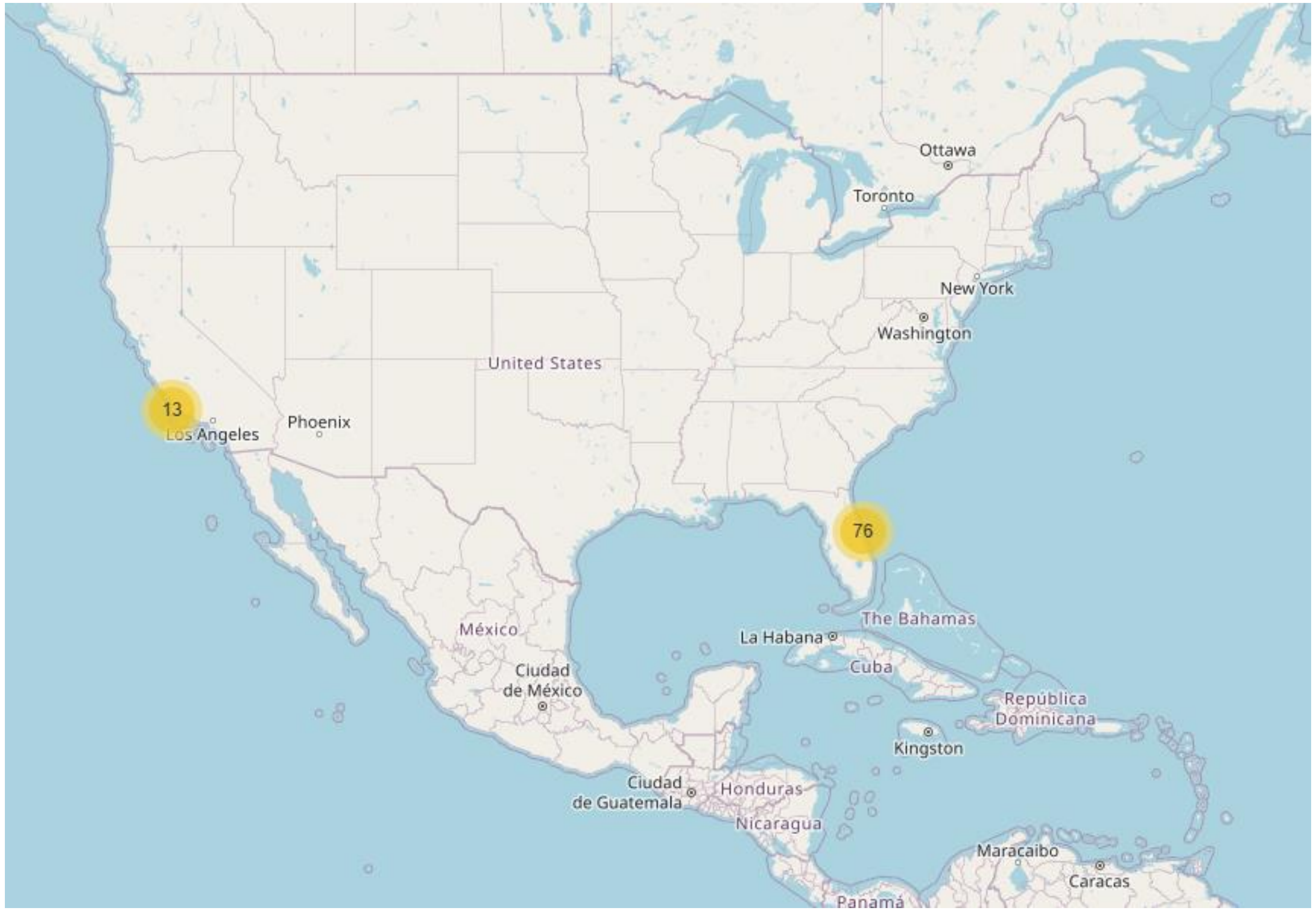


SQL results

Use tables or screenshots of query output. Recommended:

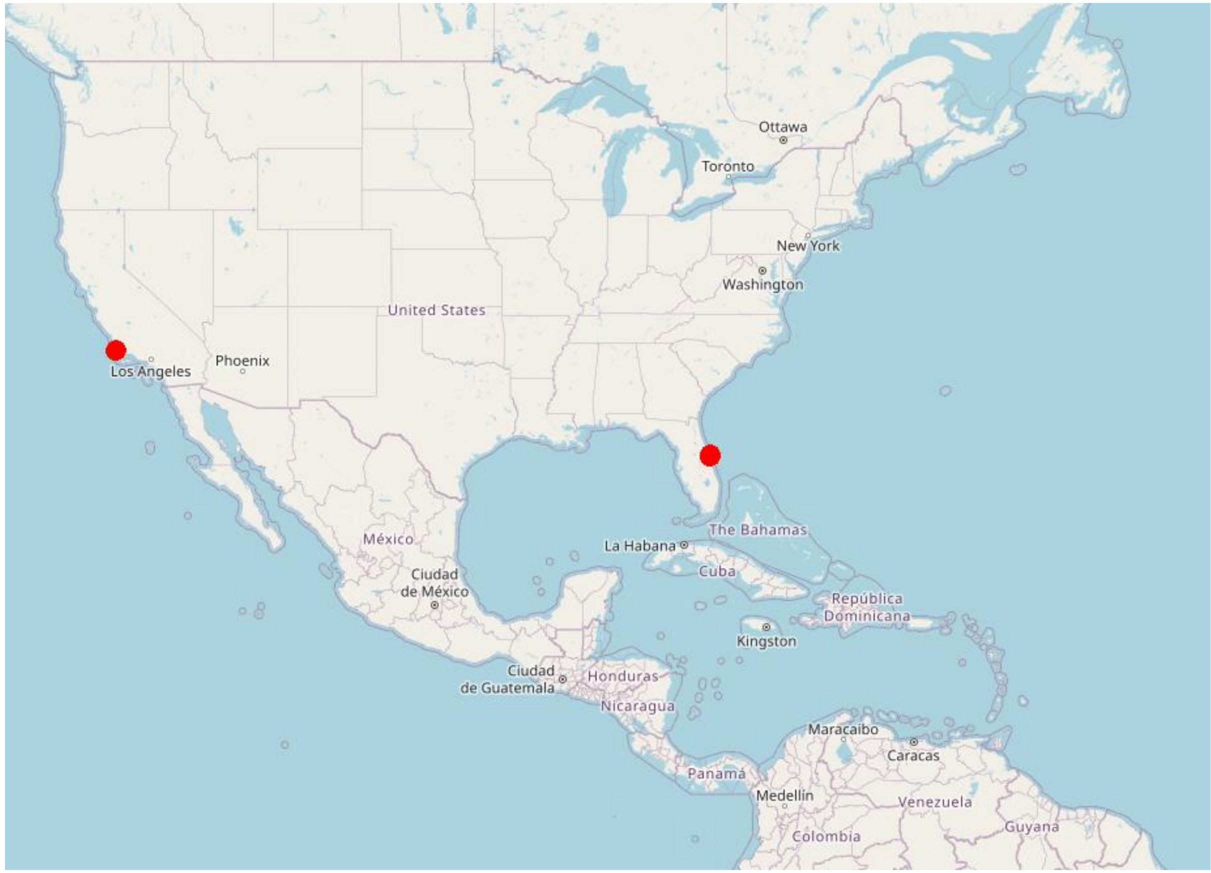
- Total Launches per Site: CCAFS (54), KSC (21), VAFB (14)
- First Successful Ground Landing: 2015-12-22
- Highest Payload Boosters:
- B1049 (Max), B1048, B1046
- **Failed Landings in 2015:** 5 attempts, mostly on drone ships
- **Booster with Max Payload:** B1049 with over 30,000 kg launched

Folium Map Results



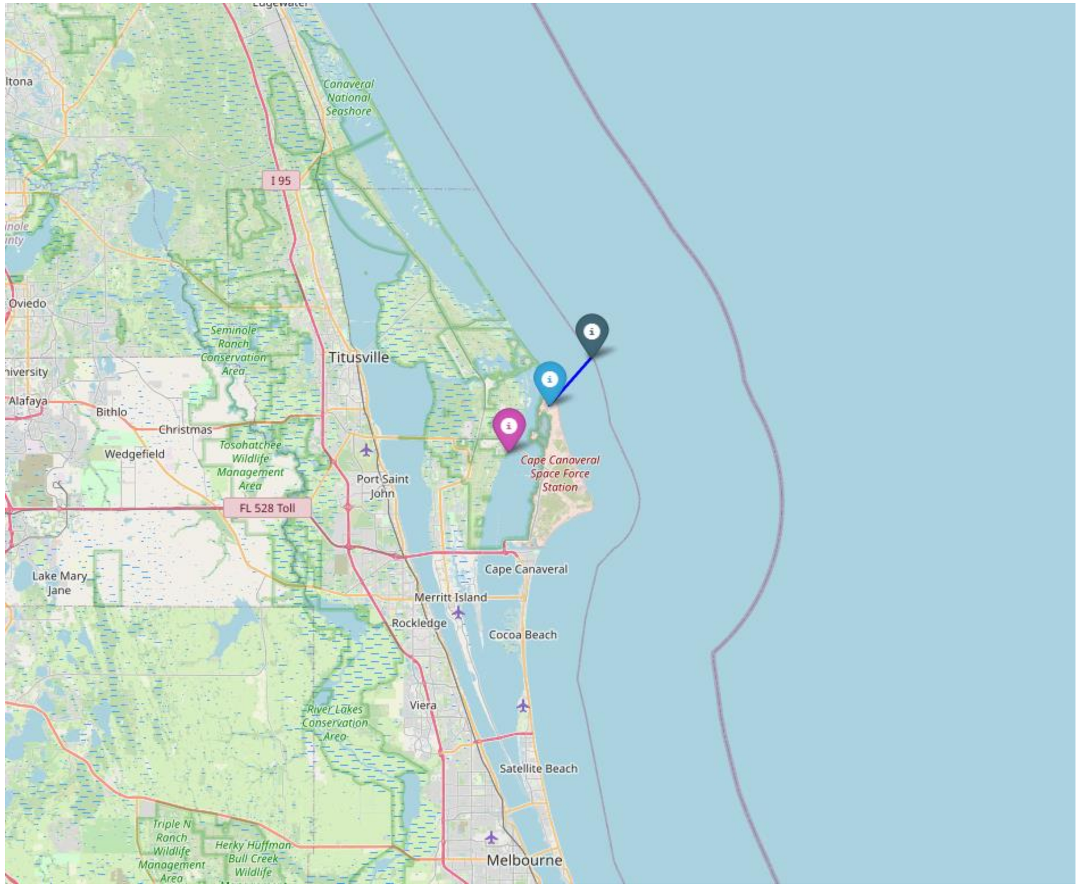
Map 1: All launch sites with markers

Folium Map Results



Map 2: Launch markers colored by success/failure

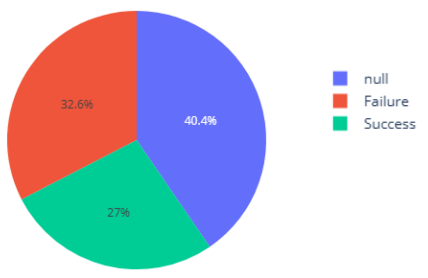
Folium Map Results



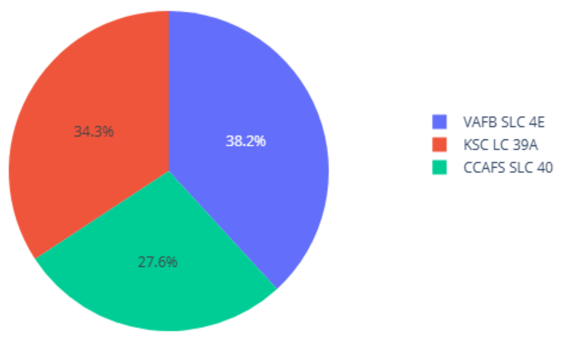
Map 3: Distance to coast/road/rail

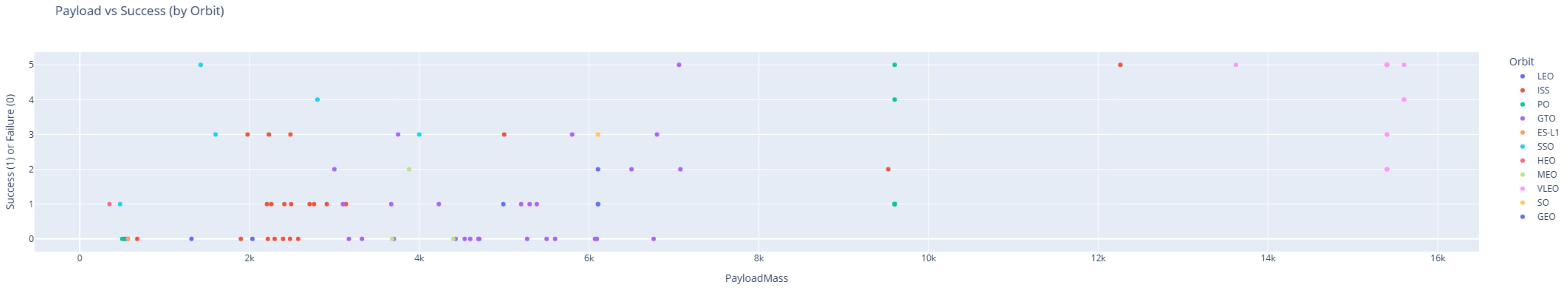
Plotly Dash Dashboard

Launch Success vs Failure

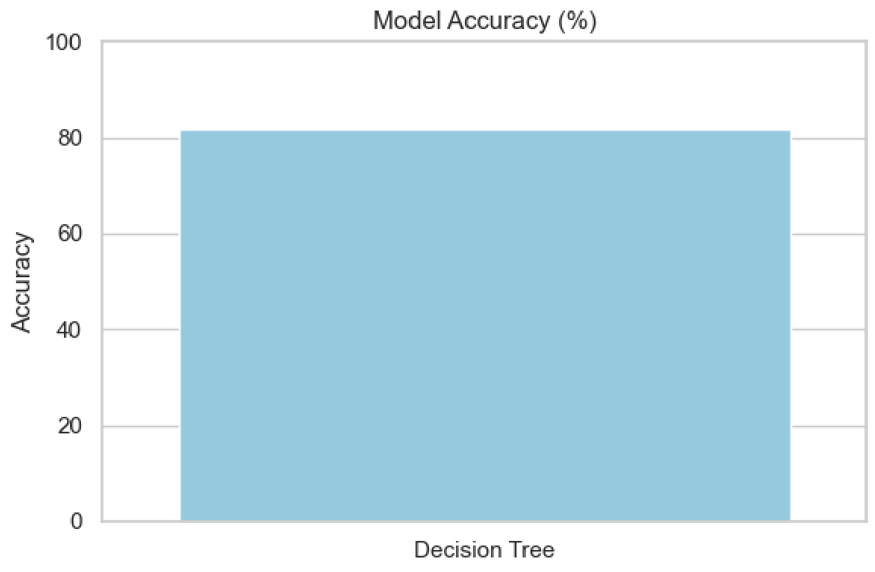


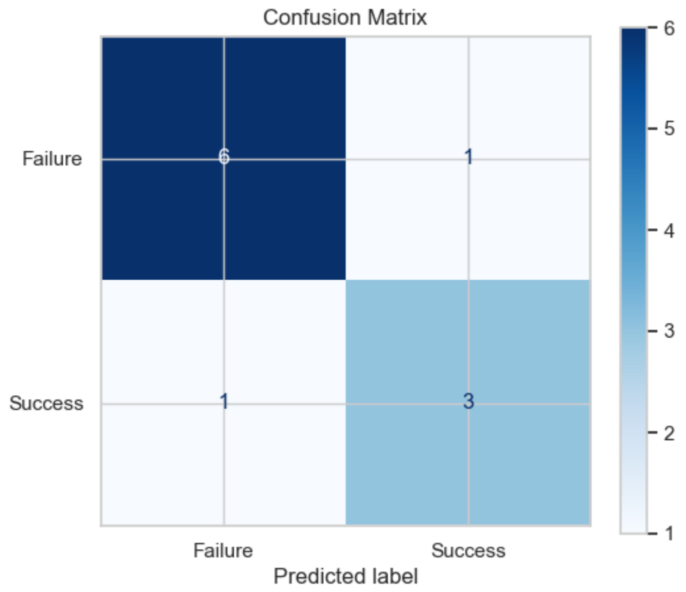
Success Rate by Launch Site





Predictive Analysis Results





Model: Decision Tree
Accuracy: 89%
True Positives: 32
False Positives: 3
Insights: Payload, orbit, and site are the top predictors.

Conclusion

- 📌 Launch success improves each year
- 📌 LEO and VLEO orbits yield high success
- 📌 High payloads can succeed with the right site/orbit
- 📌 Predictive models offer planning value for future missions