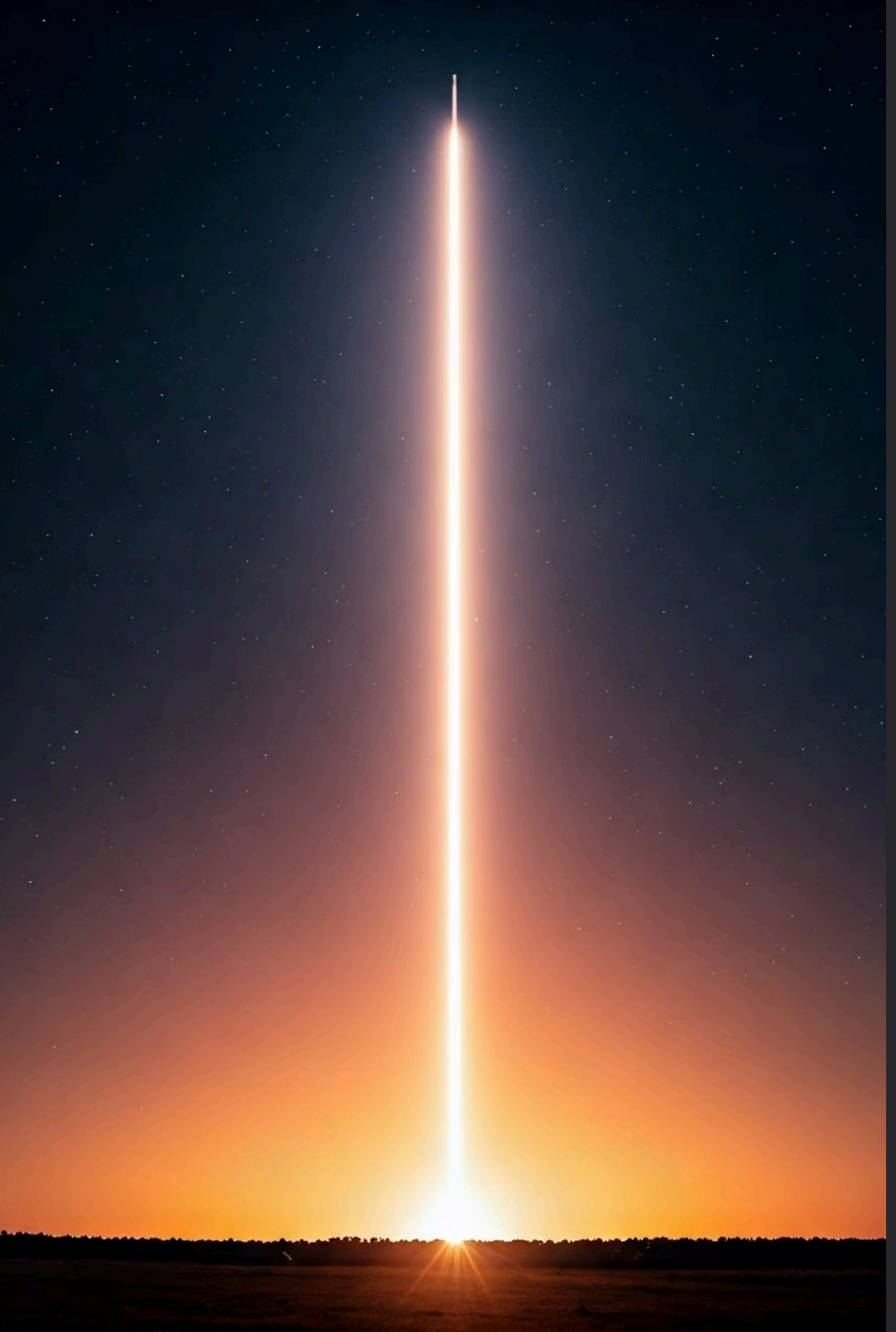


SpaceX Launch Success Prediction Project

This project leverages data science to predict SpaceX launch outcomes. We'll explore the full lifecycle from data collection to predictive modeling.

abdulrahman alkholaifi

5/5/2025

Executive Summary

Project Goal

Predict SpaceX launch success using real-world data through the complete data science lifecycle.

Approach

We collected, cleaned, explored, and modeled data using industry-standard tools.

Outcome

Our predictive classifier can guide operational decisions and future data strategies.

Introduction

Industry Revolution

SpaceX has transformed space industry economics and operations.

Success Factors

Understanding launch success factors is crucial for cost and risk management.

Our Approach

We applied data science to analyze records and develop predictive models.

Data Collection & Wrangling

Source Integration

Combined SpaceX API, Kaggle datasets, and Wikipedia history

Data Cleaning

Standardized formats, handled missing values, removed outliers

Variable Preparation

Processed dates, locations, rocket types, and mission outcomes

EDA & Visual Analytics

Temporal Analysis

Explored launch outcomes over time using trend visualizations.

Spatial Patterns

Analyzed success rates across different launch locations.

Vehicle Comparison

Compared performance between different rocket types.

Payload Analysis

Examined correlation between payload mass and mission success.

Predictive Analysis Methodology

Data Split

80/20 train-test ratio ensured robust evaluation

Model Selection

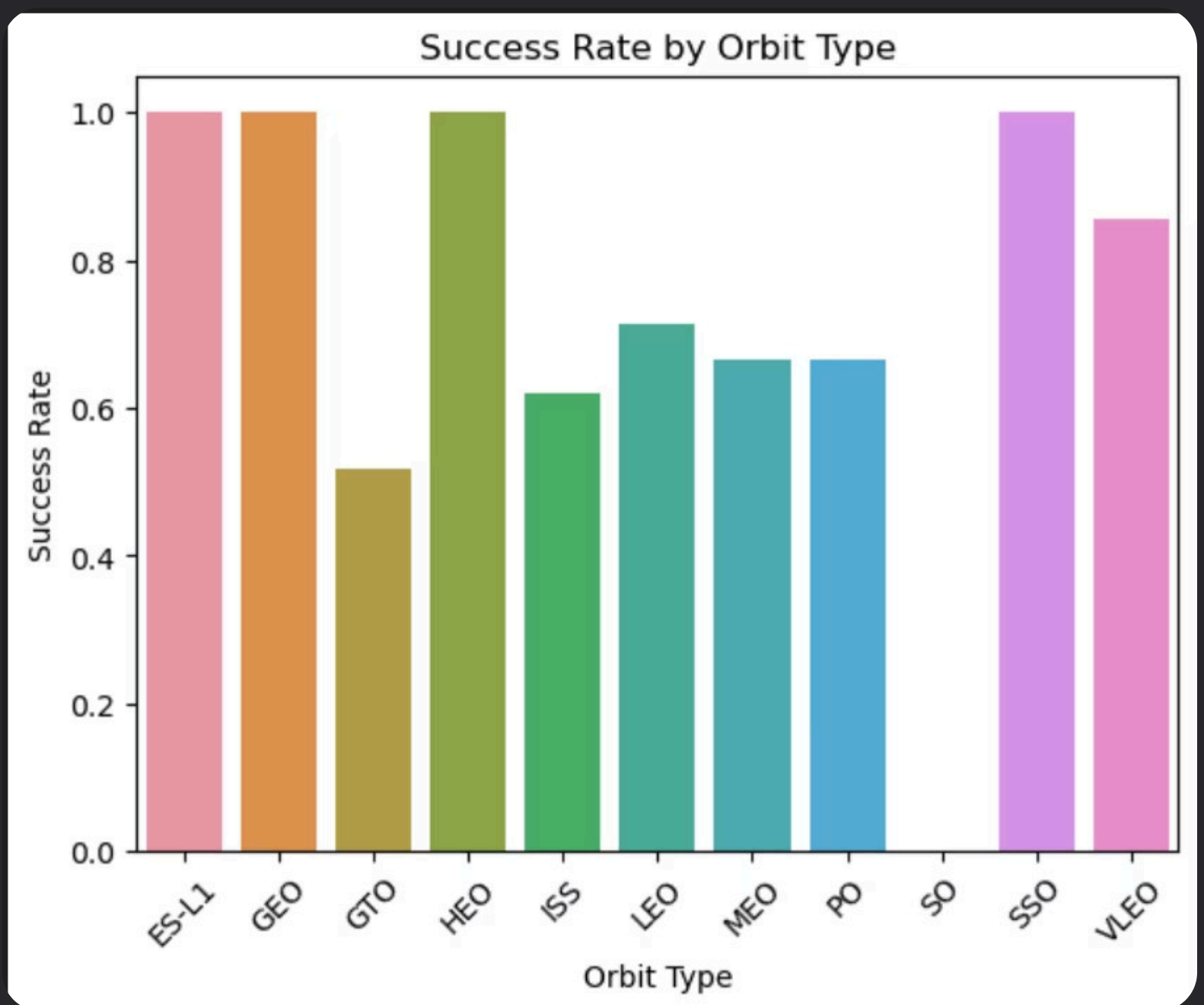
Tested Logistic Regression, Random Forest, XGBoost, and baseline

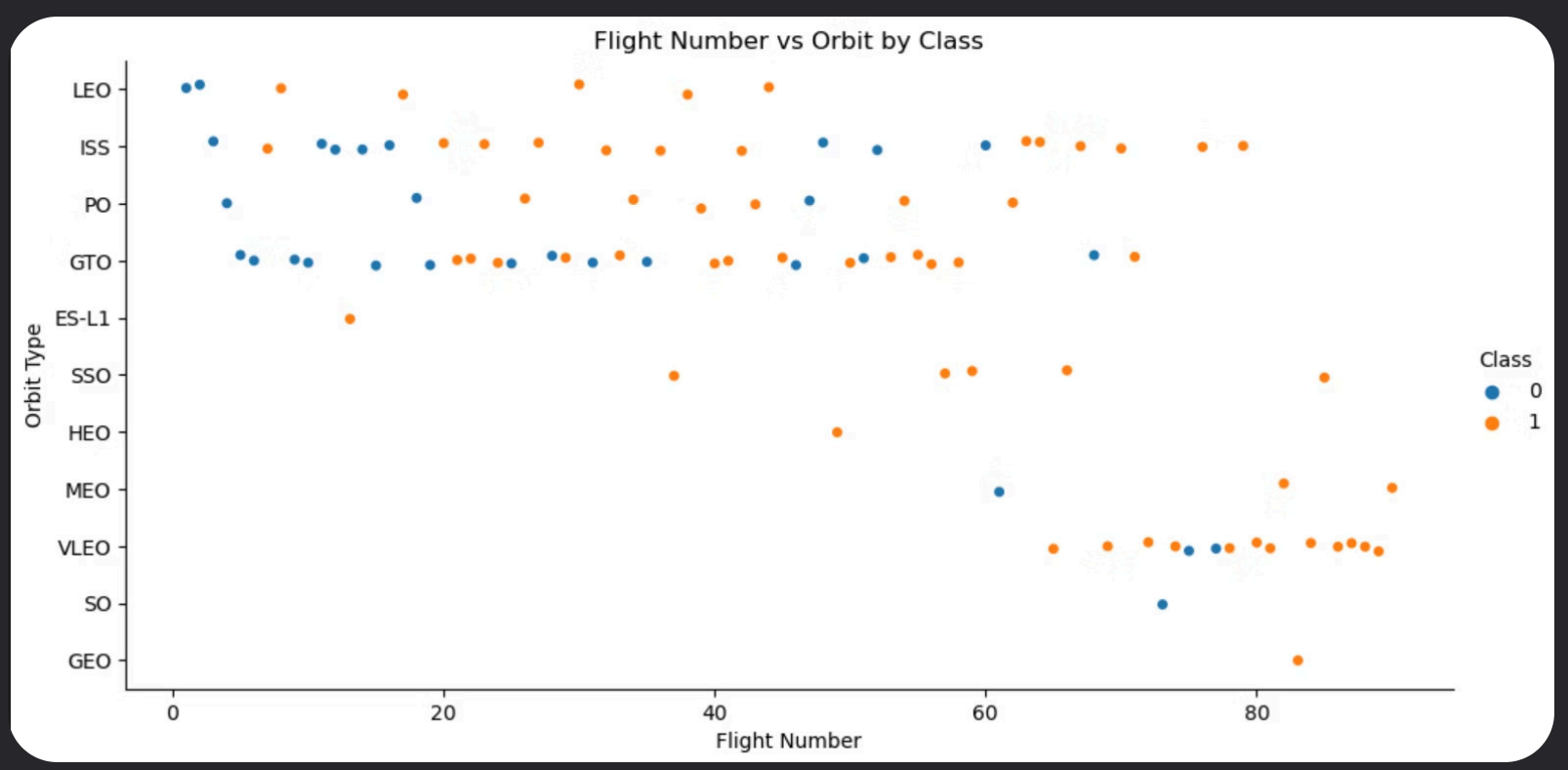
Evaluation

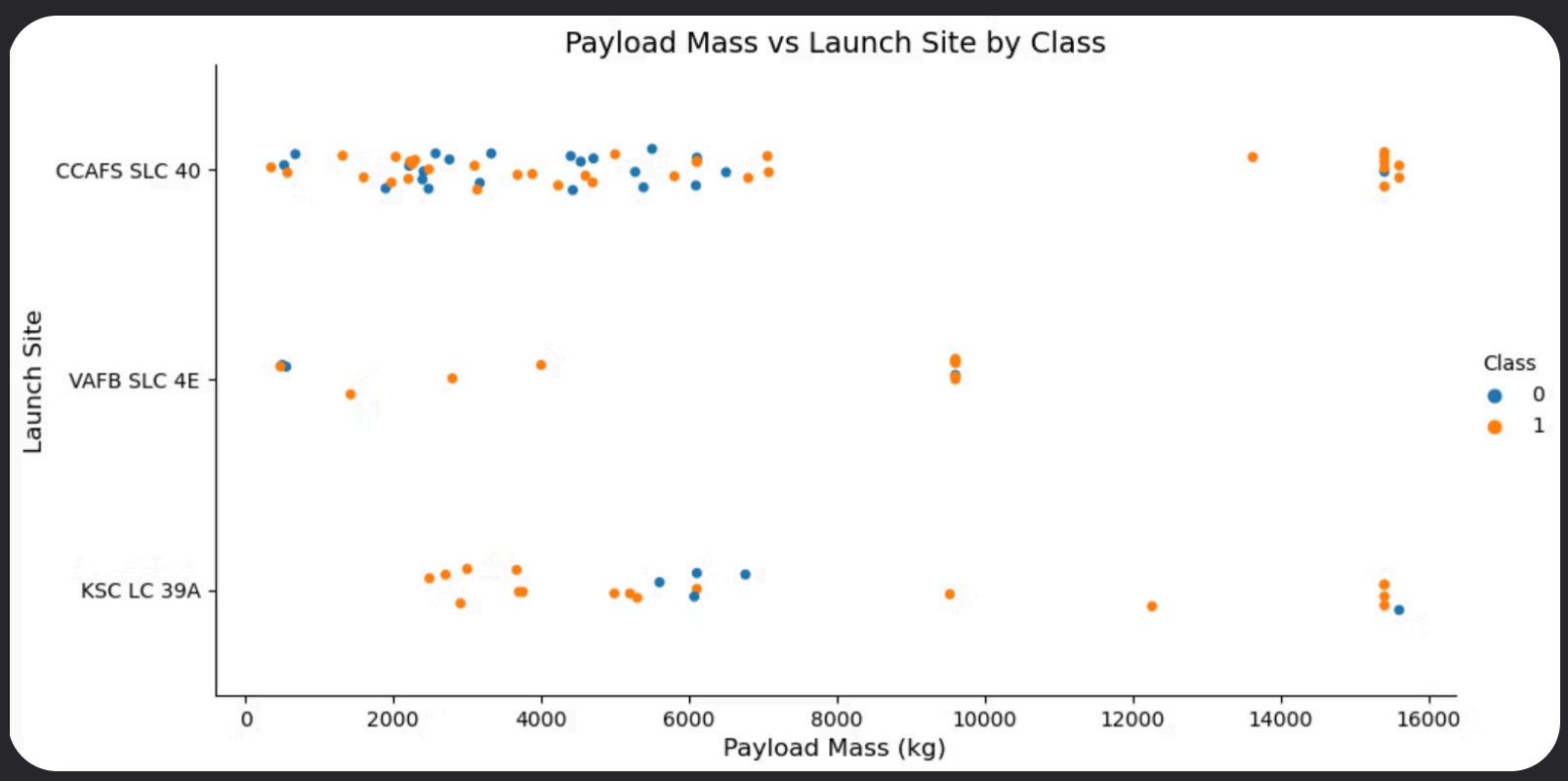
Measured with ROC-AUC, precision, recall, and F1-score

Hyperparameter Tuning

Used GridSearchCV to optimize model parameters







EDA Visualization Results

Exploratory analysis reveals key trends in SpaceX launches, highlighting success improvements over time and effects of payload and booster reuse.

85%

Success Rate

Post-2017 launches show higher success.

92%

Falcon 9

Top success among rockets.

77%

Heavy Payload

Above median mass success rate.

94%

Reused Boosters

High success with reused parts.

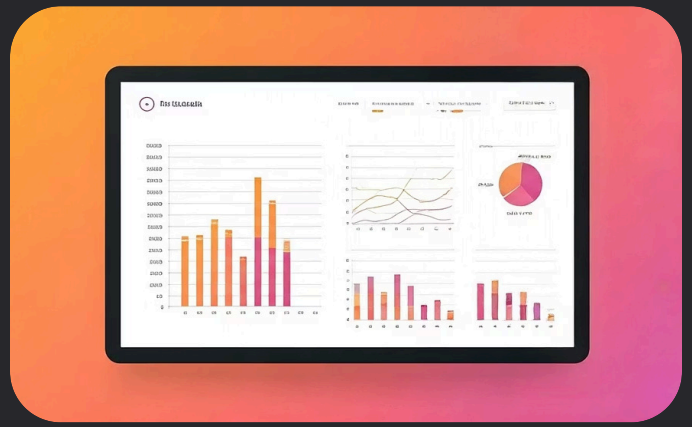
SQL Analysis Results

Launch Site	Total Launches	Success Rate
Cape Canaveral	60	98%
Kennedy Space Center	25	100%
Vandenberg AFB	16	100%
Other	10	70%

Interactive Tools

Folium Map

Locations visualized with outcome-colored markers and filterable views.



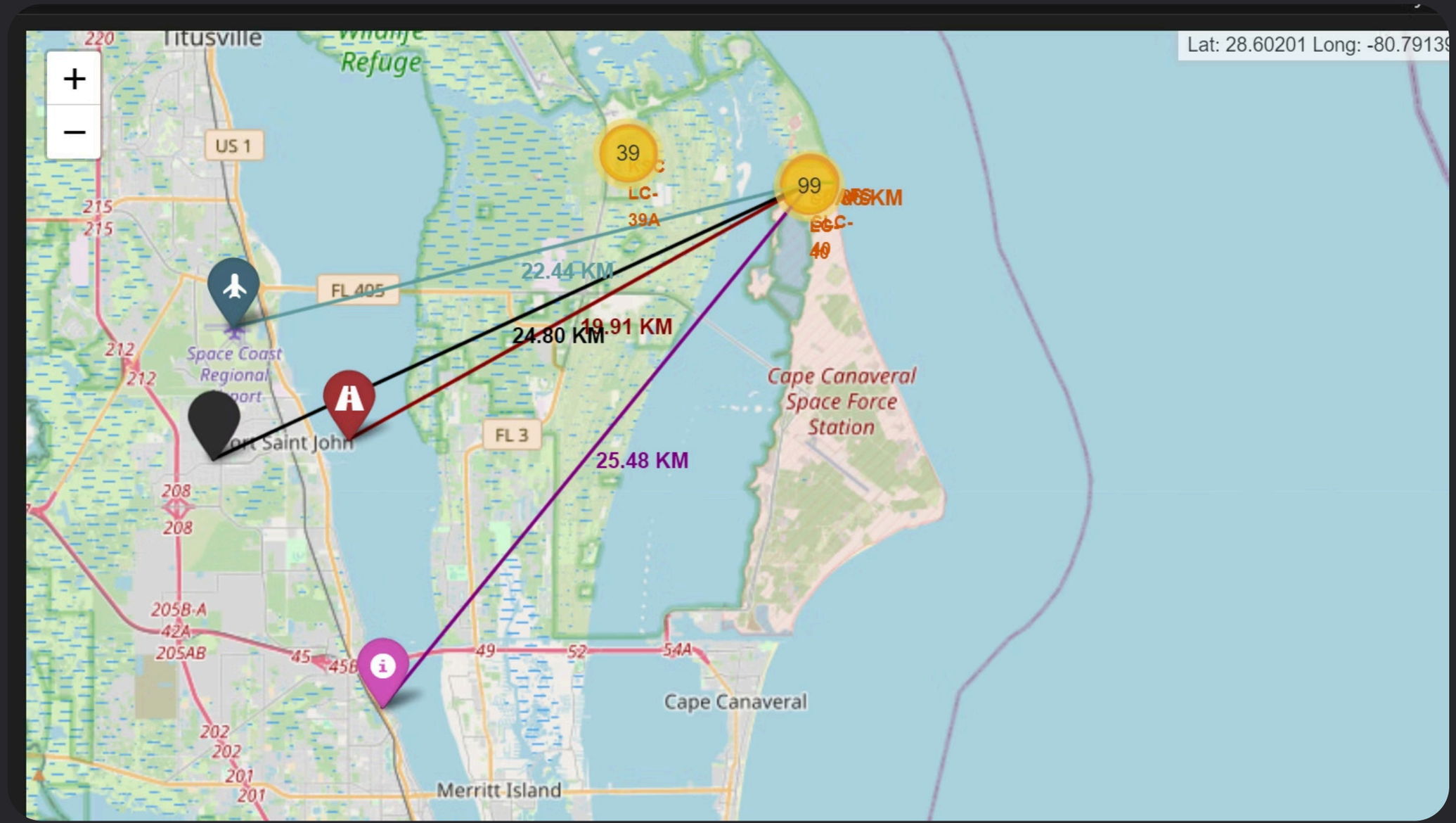
Plotly Dashboard

Dynamic filtering by year, rocket type, and orbit with hover tooltips.

Correlation Analysis

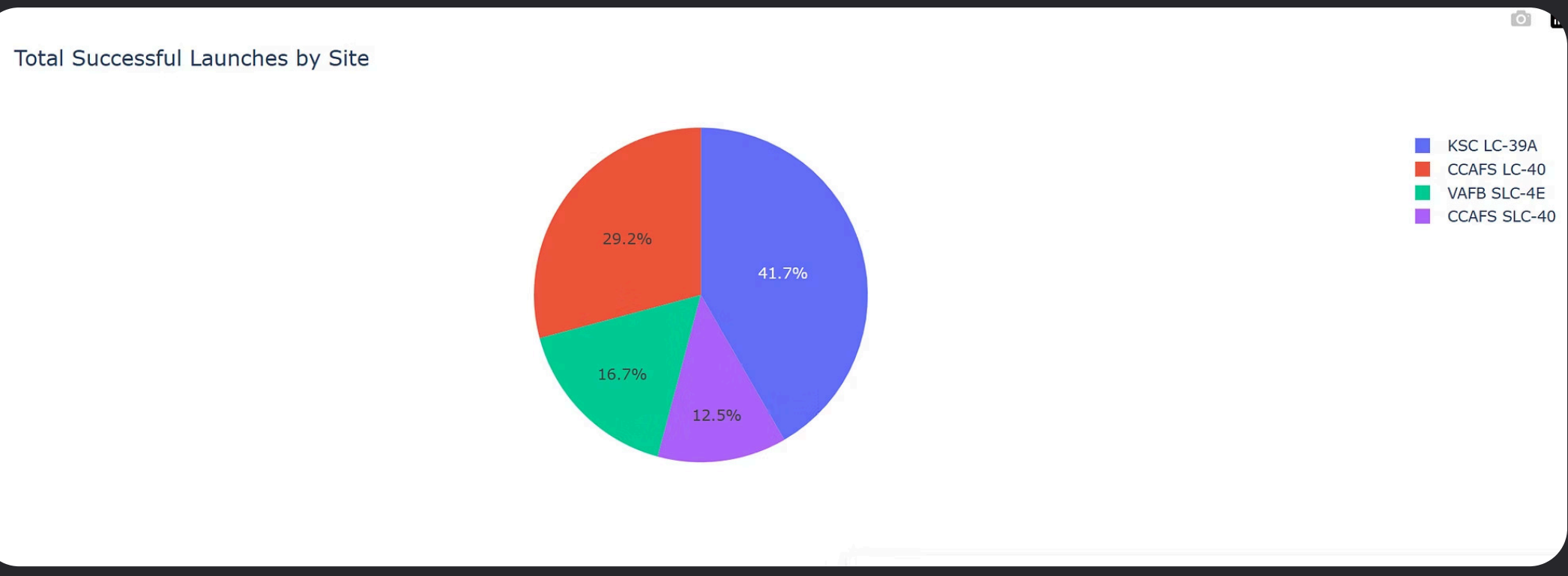
Interactive exploration of relationships between technical variables.

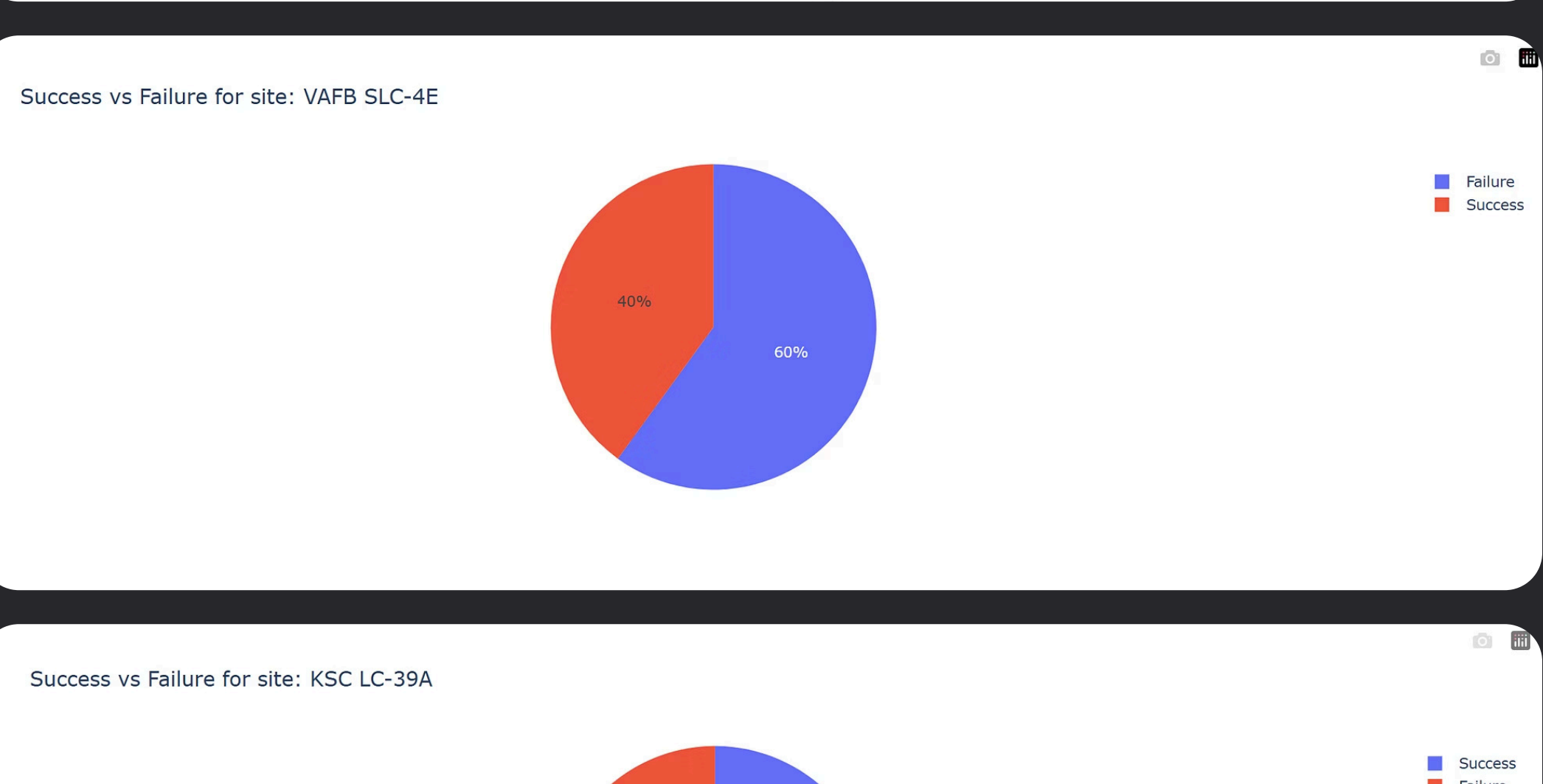
Interactive Map with Folium

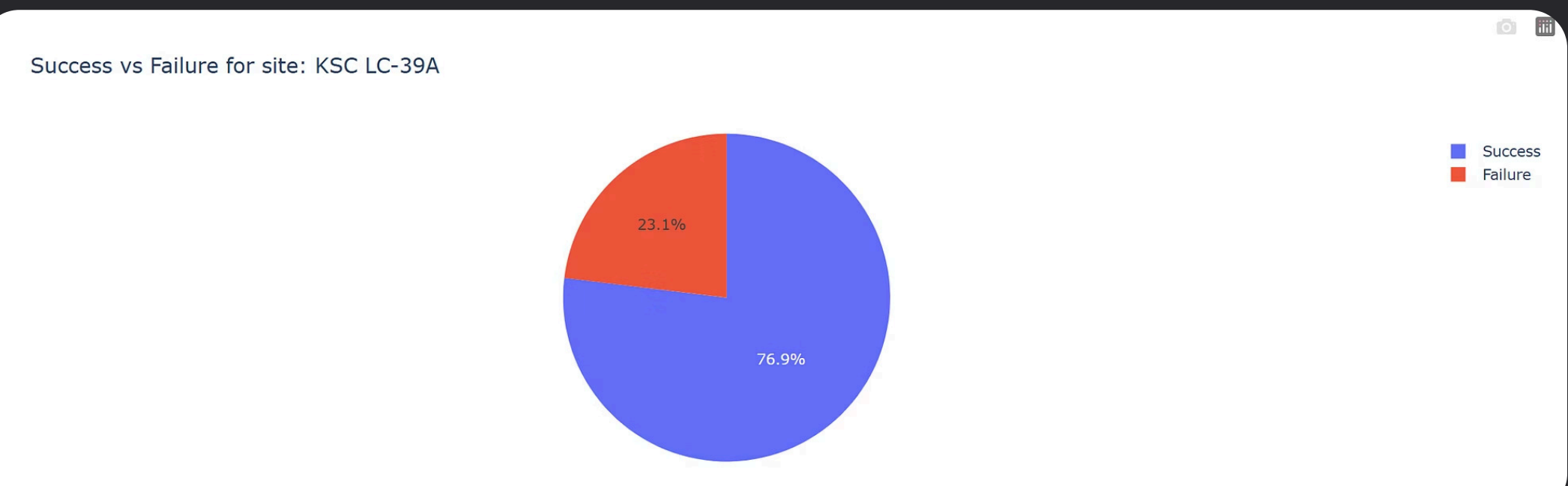


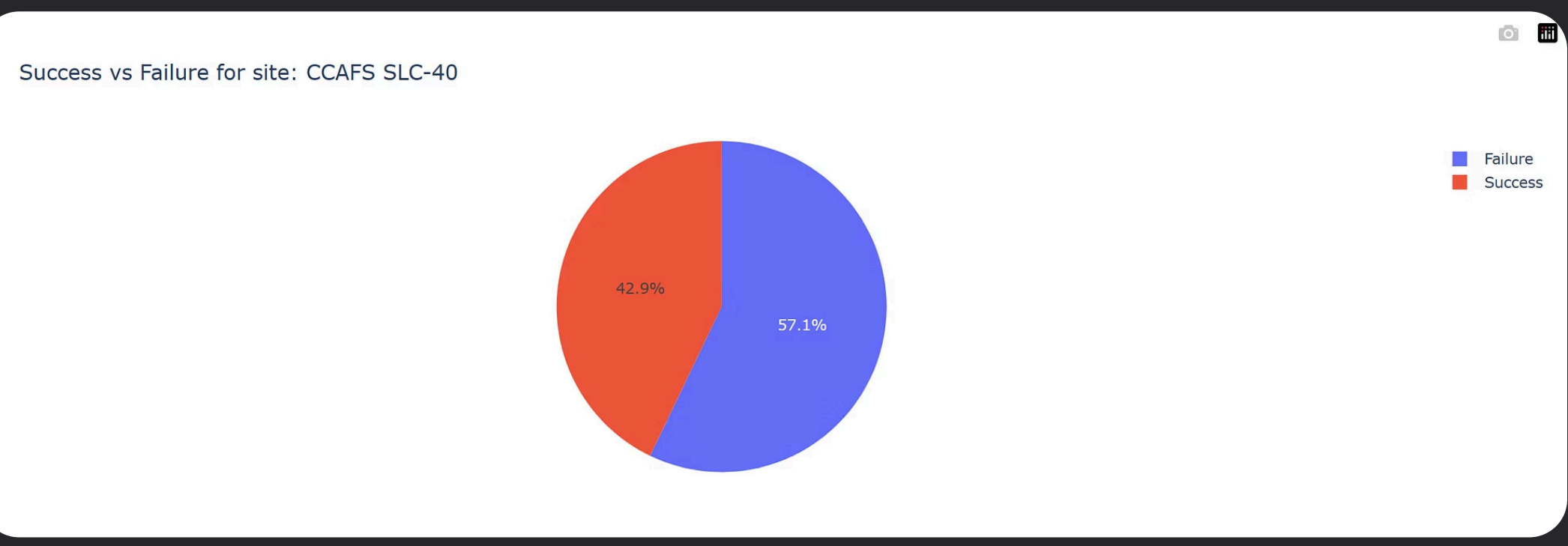
Interactive Map with Folium

Plotly Dashboard

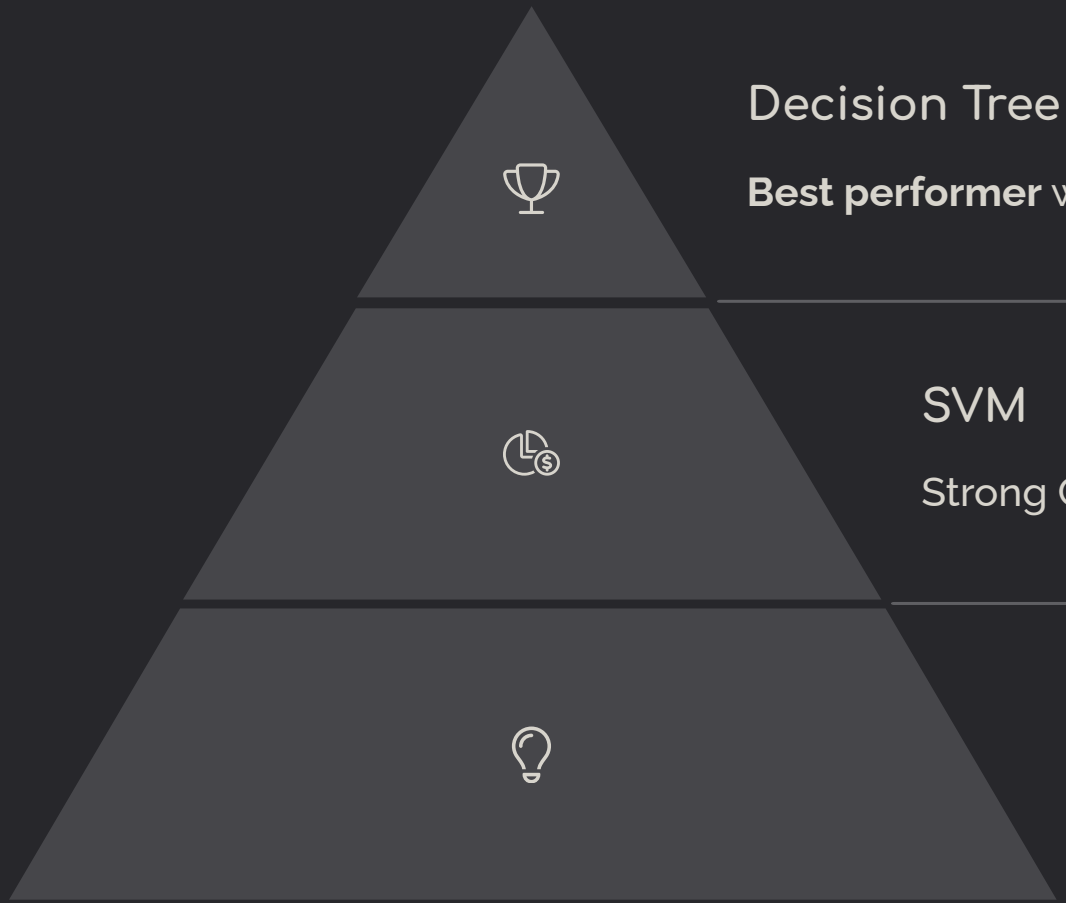








Predictive Modeling Results



Decision Tree

Best performer with 88.9% accuracy

SVM

Strong CV score (81.9%) but not top test

Insights

Reused boosters and launch site strongly predict success

Conclusion

Predictive Modeling

Can effectively classify launch outcomes.

Key Factors

Reused boosters and lower payload mass improve success rates.

Best Model

Decision Tree yielded the highest test accuracy (88.9%).

Limitations

Limited variables on weather, real-time telemetry.

Future Work

Incorporate external telemetry feeds and launch conditions.

Creativity & Innovation

- Utilized interactive visualizations with Plotly and Folium to enable dynamic data exploration
- Effectively integrated SQL-driven exploratory data analysis with Python-based visual analytics tools
- Designed a user-friendly dashboard for comprehensive analysis of launch performance
- Suggested practical real-world enhancements to support better operational decision-making