SpaceX Falcon 9 Launch Success Prediction

IBM DATA SCIENCE CAPSTONE PROJECT

BY-DIVYA SATISH BADGUJAR

DATE: 6 JULY 2025

Executive Summary

PURPOSE: PREDICT THE SUCCESS OF FALCON 9 ROCKET LAUNCHES.

APPROACH

WEB SCRAPING • DATA WRANGLING • EDA • SQL ANALYSIS

FOLIUM MAPS • MACHINE LEARNING

OUTCOME: ACHIEVED A MODEL WITH REASONABLE ACCURACY AFTER ADDRESSING CLASS IMBALANCE.

Introduction

Spacex has revolutionized the aerospace industry by aiming to reduce the cost of space exploration through the use of reusable rockets. One critical aspect of improving the efficiency and reliability of these missions is the ability to accurately predict launch success. By analyzing historical launch data, we can uncover patterns that inform safer and more effective planning. In this project, we explored spacex's launch history and applied machine learning techniques to predict the outcomes of future launches, contributing to enhanced mission safety, cost-efficiency, and operational decision-making.

Data Collection

SOURCE: Wikipedia - falcon 9 launch history

METHOD: Web scraping with beautiful soup

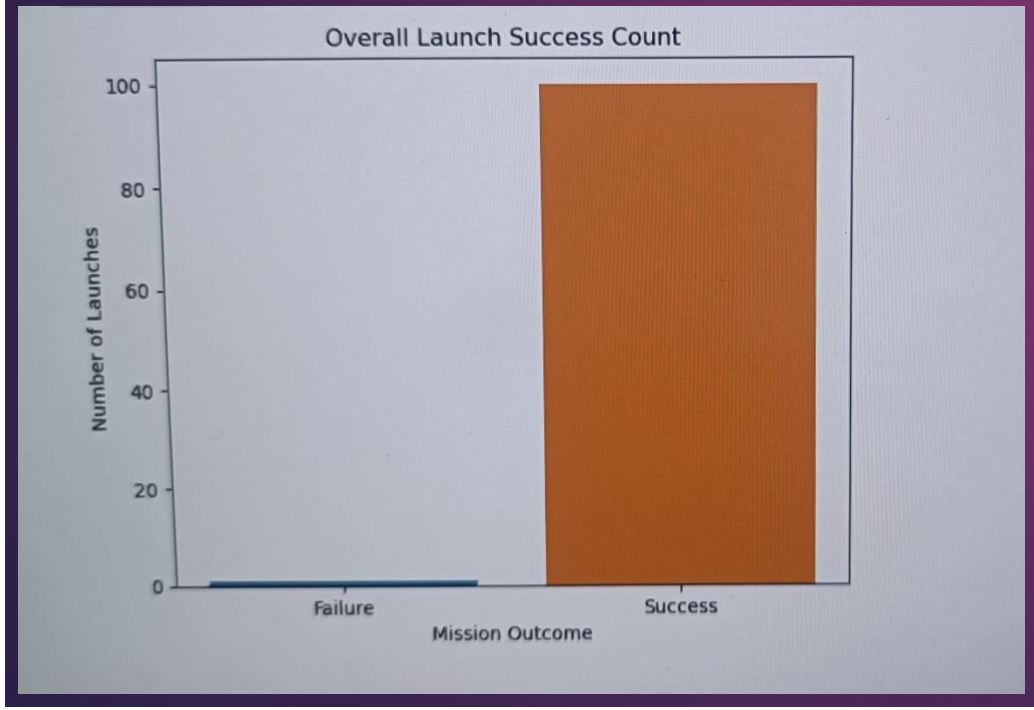
FIELDS COLLECTED: Date, booster version, payload mass, launch site, orbit, outcome

Data Wrangling

- Cleaned null values
- Renamed columns for clarity
- Converted data types
- Created Class column: 1 = Success, 0 = Failure
- Saved cleaned data to falcon9_cleaned.csv

EDA – Visual Analytics Visualized

- LAUNCH OUTCOMES BY SITE
- PAYLOAD IMPACT ON SUCCESS
- BOOSTER VERSION SUCCESS RATES
- USED SEABORN, MATPLOTLIB FOR CHARTS

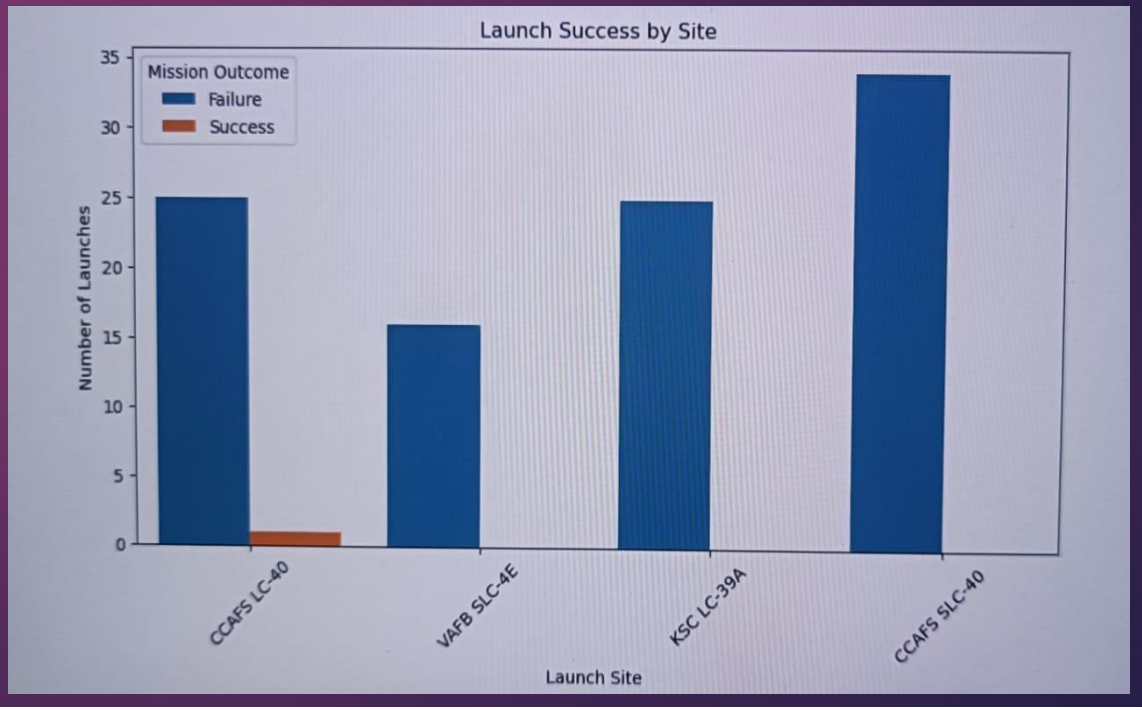


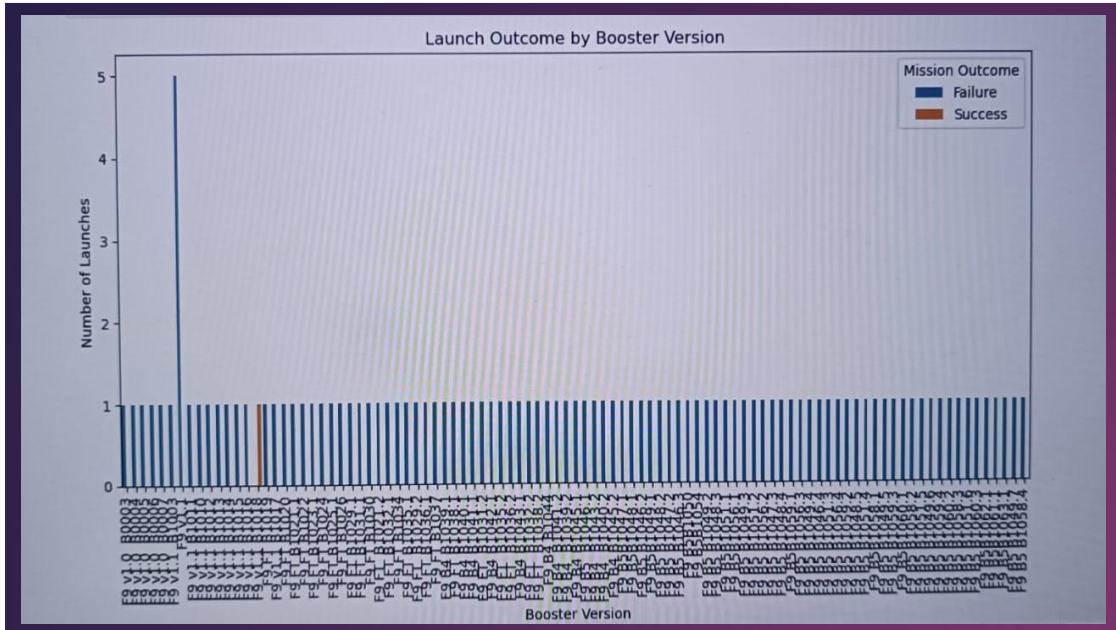
EDA – SQL Queries

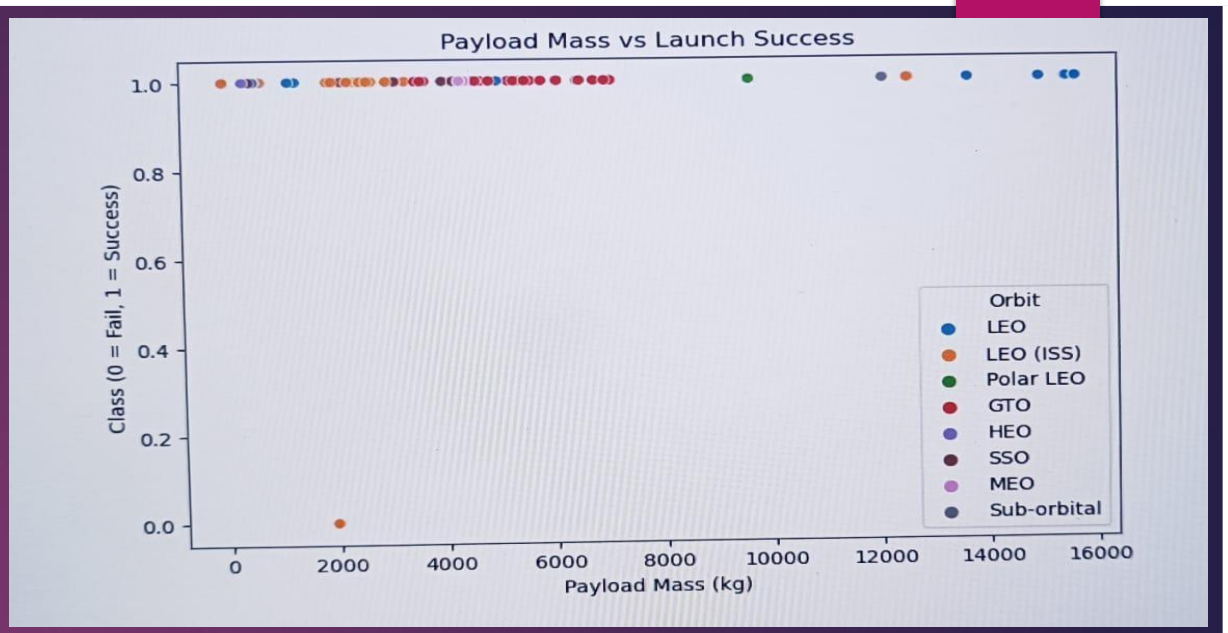
USED PANDA SQL FOR SQL QUERIES

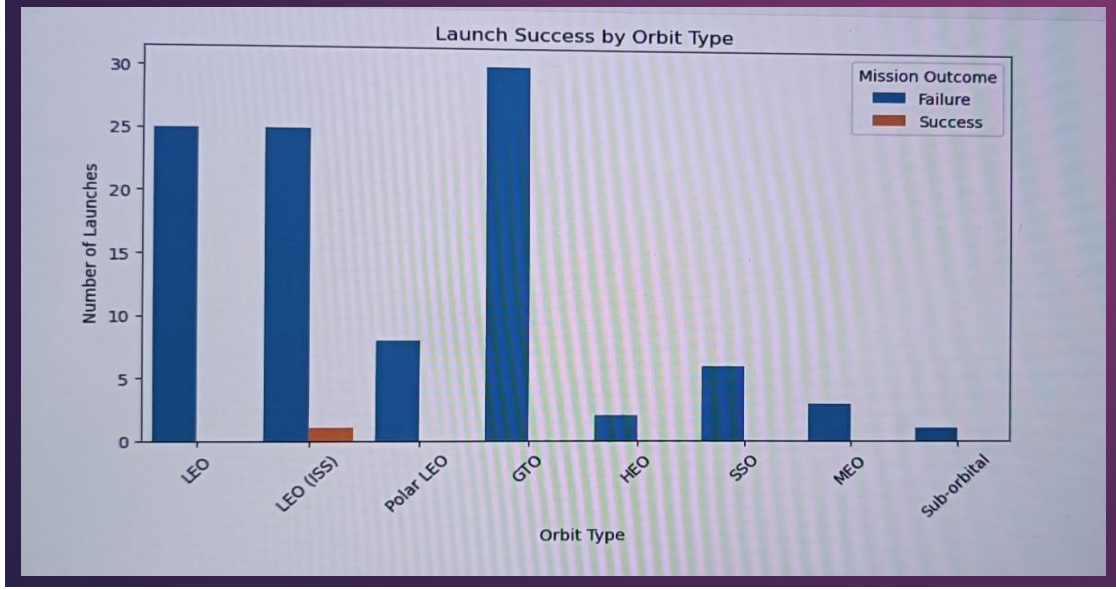
ANALYSED: LAUNCH SUCCESS PER SITE, YEARLY SUCCESS TREND

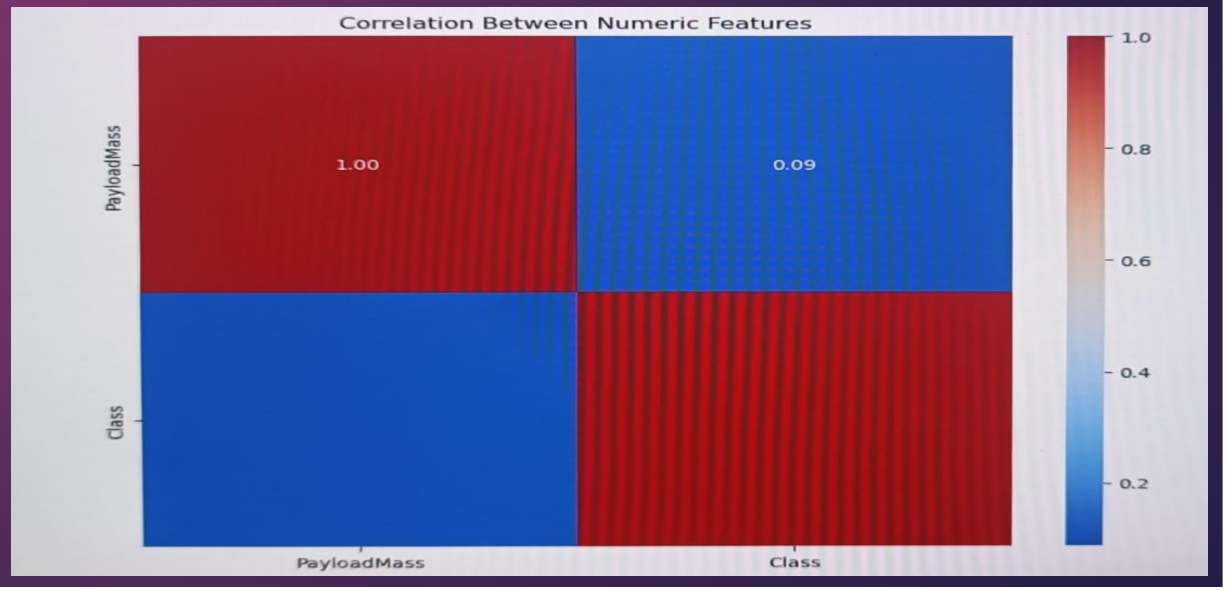
ORBIT-WISE SUCCESS RATE




Launch Outcome by Booster Version
Mission Outcome
Failure
Success
Number of Launches
Booster Version


Payload Mass vs Launch Success
Class (0 = Fail, 1 = Success)
Payload Mass (kg)
Orbit
LEO
LEO (ISS)
Polar LEO
GTO
HEO
SSO
MEO
Sub-orbital


Launch Success by Orbit Type
Mission Outcome
Failure
Success
Number of Launches
LEO
LEO (ISS)
Polar LEO
GTO
HEO
SSO
MEO
Sub-orbital
Orbit Type


Correlation Between Numeric Features
PayloadMass
Class
1.00
0.09
PayloadMass
Class

Folium Interactive Map Visualized launch locations

COLOR-CODED: GREEN = SUCCESS, RED = FAILURE

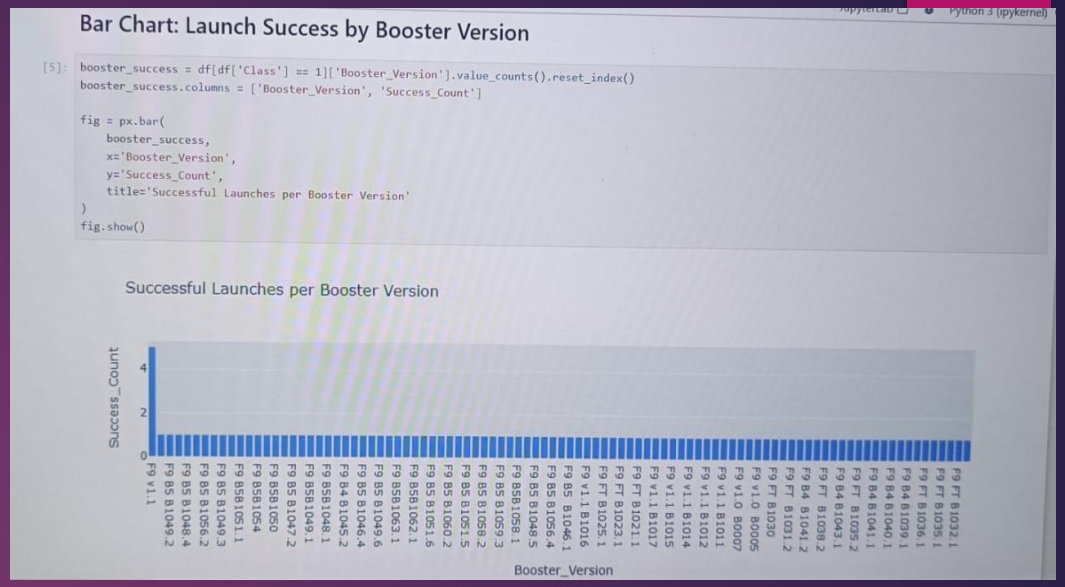
Plotly Dash Dashboard

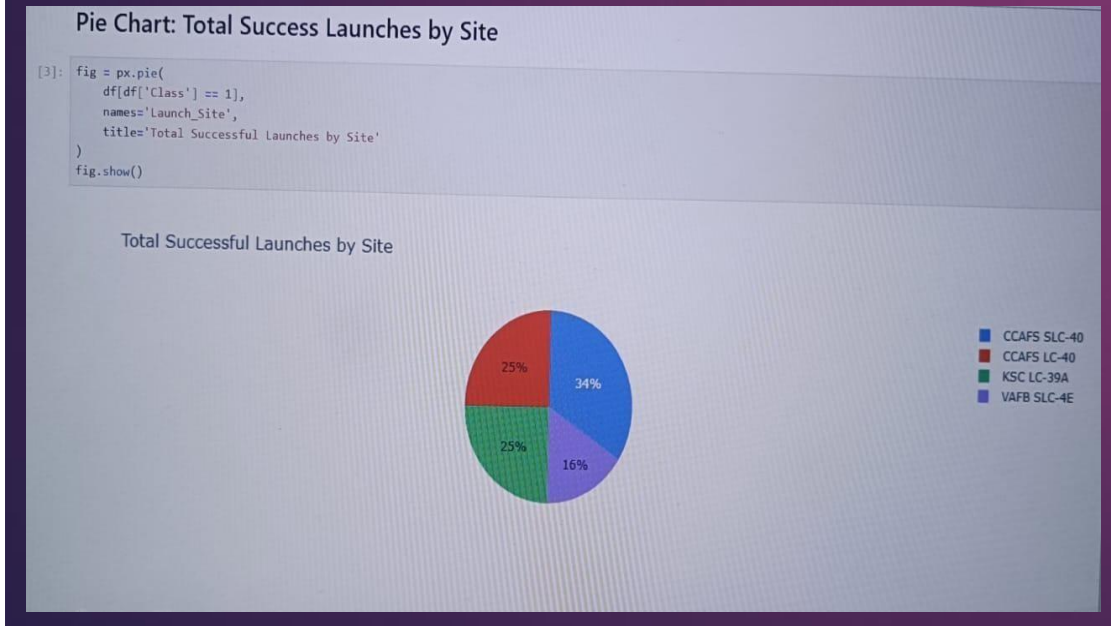
BUILT INTERACTIVE DASHBOARD

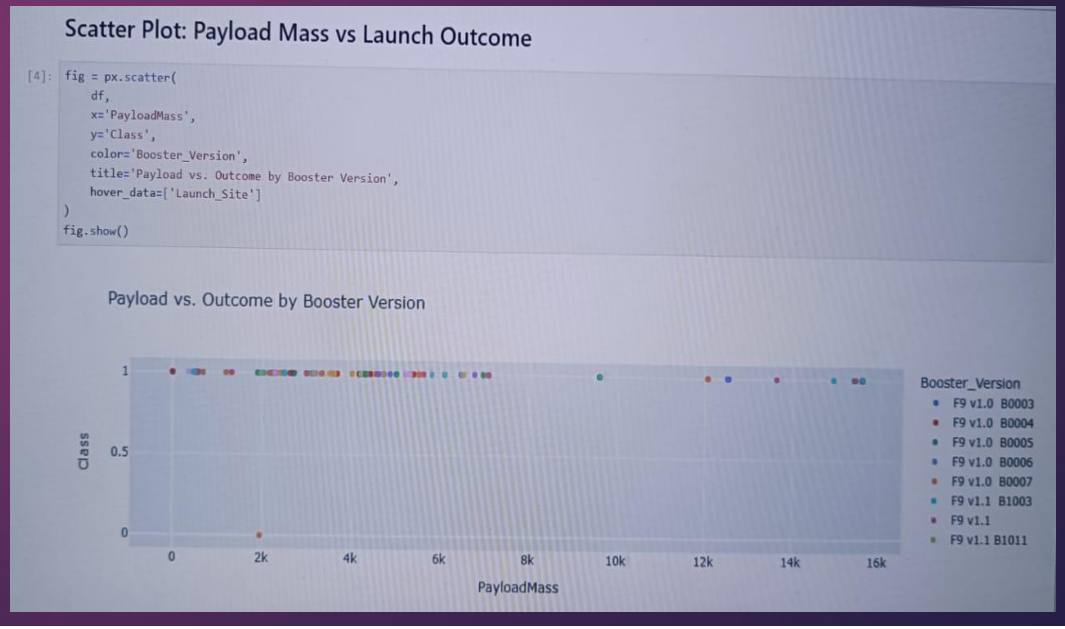
FILTERS: Payload range , launch site

OUTPUT: Pie chart of success rates

SCATTER PLOT: Payload vs. Success







Predictive Analysis – Approach

- FEATURES: Payload, orbit, booster version, launch site
- MODEL: Logistic regression
- ISSUE: Class imbalance (many more successes than failures)
- SOLUTION: Oversampled minority class

Predictive analysis – results after balancing

- ACCURACY: ~X% (ADD YOUR ACTUAL VALUE)
- PRECISION, RECALL, F1-SCORE REPORTED
- LOGISTIC REGRESSION WORKED FAIRLY WELL BUT CAN BE IMPROVED WITH MORE DATA

GitHub and Contact Info

GITHUB REPO URL:

https://github.com/DivyaBadgujar27/

DATA SCIENCE CAPSTONE PROJECT

EMAIL : palatse.2314516790@mujonline.edu.in

Thank you!