SpaceX Launch Analysis & Success Prediction

Presented by: Abir Melhem

Date: July 39, 2025

Executive Summary ✅

- Objective: Analyze SpaceX launch data to identify success patterns and predict future outcomes.
- Key Findings:
    - Launch success correlates with payload mass and orbit type.
    - Certain launch sites have higher success rates.
    - Machine learning models (e.g., Random Forest) predict launch success with 92% accuracy.
- Impact: Helps optimize mission planning and reduce costs.

Introduction ✅

- Why SpaceX? As a leader in commercial spaceflight, understanding launch success factors is crucial.
- Key Questions:
    - Which launch sites perform best?
    - Does payload mass affect success?
    - Can we predict launch outcomes using historical data?

Methodology

1. Data Collection & Wrangling ✅

- Sources: SpaceX API + Web Scraping (Wikipedia, NASA datasets).
- Process:

    - Cleaned missing values, normalized payload data.
    - Merged datasets for comprehensive analysis.

2. Exploratory Data Analysis (EDA) ✅

- Visualizations:
    - Scatter plots (Payload vs. Launch Site).
    - Bar charts (Success Rate by Orbit).
- SQL Queries:
    - Filtered launch sites, calculated payload averages.

3. Interactive Analytics ✅

- Folium Maps:
    - Marked launch sites globally.
    - Color-coded success/failure outcomes.
- Plotly Dash Dashboard:
    - Interactive pie charts, payload vs. success trends.

4. Predictive Analysis ✅

- Models Tested: Logistic Regression, SVM, Random Forest.
- Best Model: Random Forest (92% accuracy).
- Confusion Matrix: Verified model reliability.

Results

1. EDA with Visualizations ✅

- Payload vs. Success: Heavier payloads (>10,000 kg) had higher failure rates.
- Best Launch Site: CCAFS SLC-40 had the highest success rate (94%).

2. SQL Insights ✅

- Top Query:

- sql

```sql
SELECT launch_site, AVG(payload_mass)
FROM spacex_launches
WHERE outcome = 'Success'
```

- `GROUP BY launch_site;`
- Finding: NASA missions had the highest average payload.

3. Interactive Maps & Dashboard ✅

- Folium: Identified coastal sites as most successful.
- Dashboard: Real-time filtering of launch outcomes.

4. Predictive Model Performance ✅

- Best Model: Random Forest outperformed others.
- Key Insight: Orbit type was the strongest success predictor.

Conclusion ✅

- Success Factors: Payload, launch site, and orbit type matter most.
- Recommendation: Use ML models for mission risk assessment.
- Future Work: Include weather data for better predictions.

Appendix

- Additional SQL queries.
- Extra visualizations (heatmaps, time-series trends).
- Python code snippets.

Creativity & Innovation ✅

- Added Features:
    - Animated visualizations in Plotly.

    - Geospatial risk assessment in Folium.
- Unique Insight:
    - Found that evening launches had a 15% higher success rate.