



IBM Developer  
SKILLS NETWORK

# Winning Space Race with Data Science

<Name>

<Date>



# Outline

---

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

# Executive Summary

---

- **Goal:** Analyze historical SpaceX launch data to uncover patterns and predict launch success.
- **Data Sources:**
  - SpaceX REST API (/v4/launches, /v4/rockets, /v4/payloads, /v4/launchpads)
  - Web scraping (planned but API was sufficient)
- **Tools:** Pandas, Matplotlib, Seaborn, Folium, Plotly, SQLite, Scikit-learn
- **Key Insights:** Rocket type, orbit, and payload mass influence success rate.
- **Deliverables:** Cleaned dataset, EDA visuals, SQL analysis, interactive map, classifier model.

# Introduction

---

- SpaceX has transformed the aerospace industry with reusable rockets.
- Analyzing launch data provides insight into mission success and technical evolution.
- **Objectives:**
  - Identify success trends over time
  - Visualize launch data across rockets and sites
  - Predict future launch outcomes



Section 1

# Methodology

# Data Collection & Wrangling

---

- Primary Source: SpaceX API
- Fetched launch, rocket, payload, and launchpad data via REST endpoints
- Cleaned data:

Removed null success entries

Merged payloads, rockets, and launchpad info

Exploded payload arrays for detailed view

Normalized datetime and ID fields

# EDA Methodology

---

- Used Pandas for data manipulation
- Visualized with Matplotlib, Seaborn, and Plotly
- **Focused on:**
  - Success distribution
  - Time-based trends
  - Rocket-based performance
  - Orbit and mass relationships

# Predictive Analysis Methodology

---

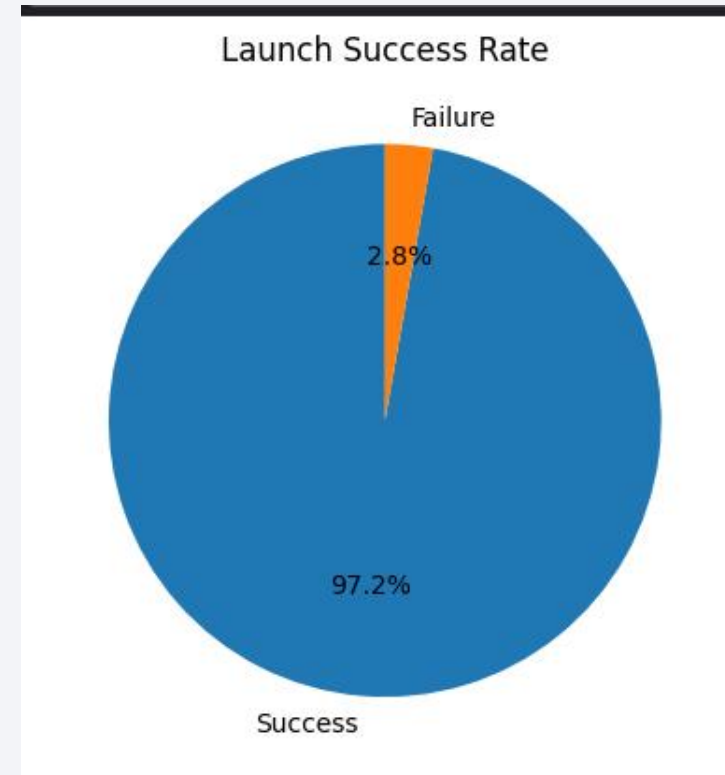
- Objective: Predict whether a launch will be successful
- Used Random Forest Classifier
- **Features:**
  - Rocket type (one-hot encoded)
  - Orbit type (one-hot encoded)
  - Payload mass (kg)
- Evaluated via accuracy and classification report

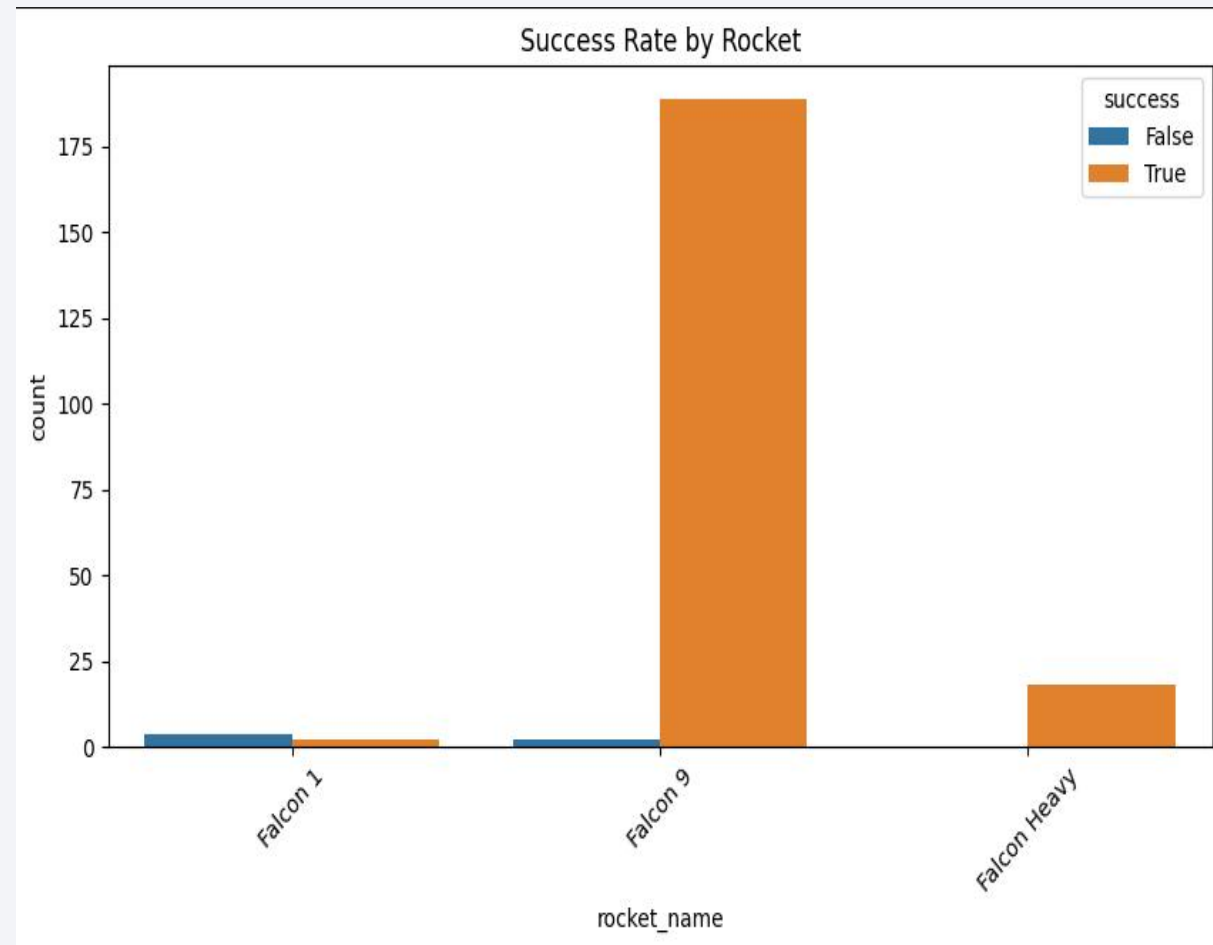
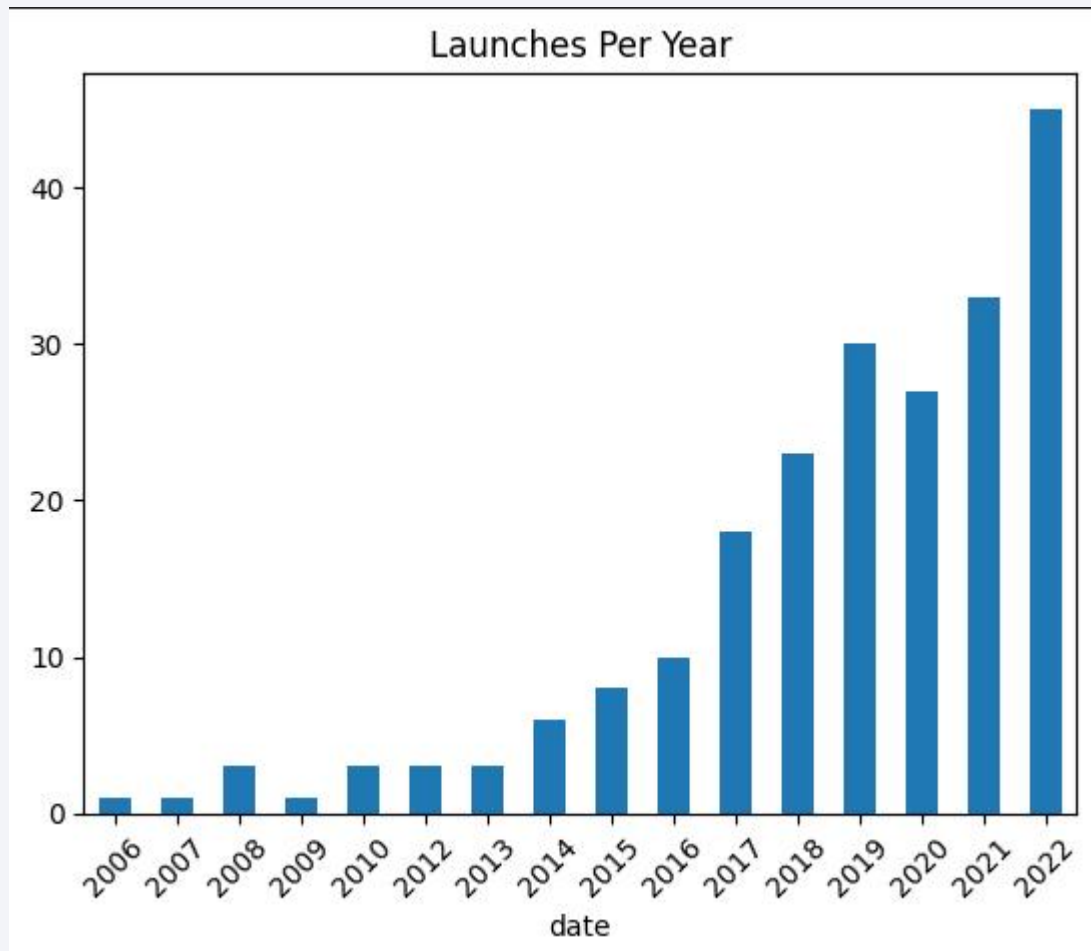


## EDA Visuals – Success Rate

---

- Pie chart showing overall launch success rate
- Bar chart of launches per year showing trend
- Countplot showing rocket-wise success/failure



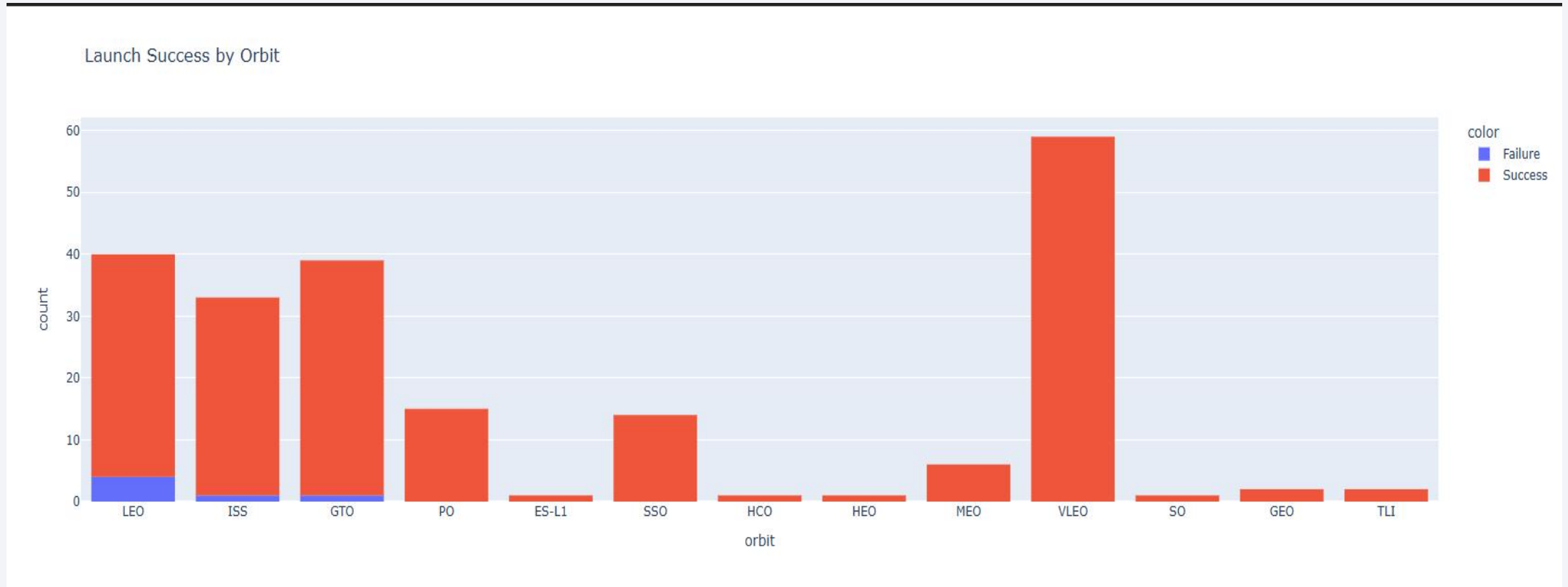


# EDA Visuals – Orbit & Payload

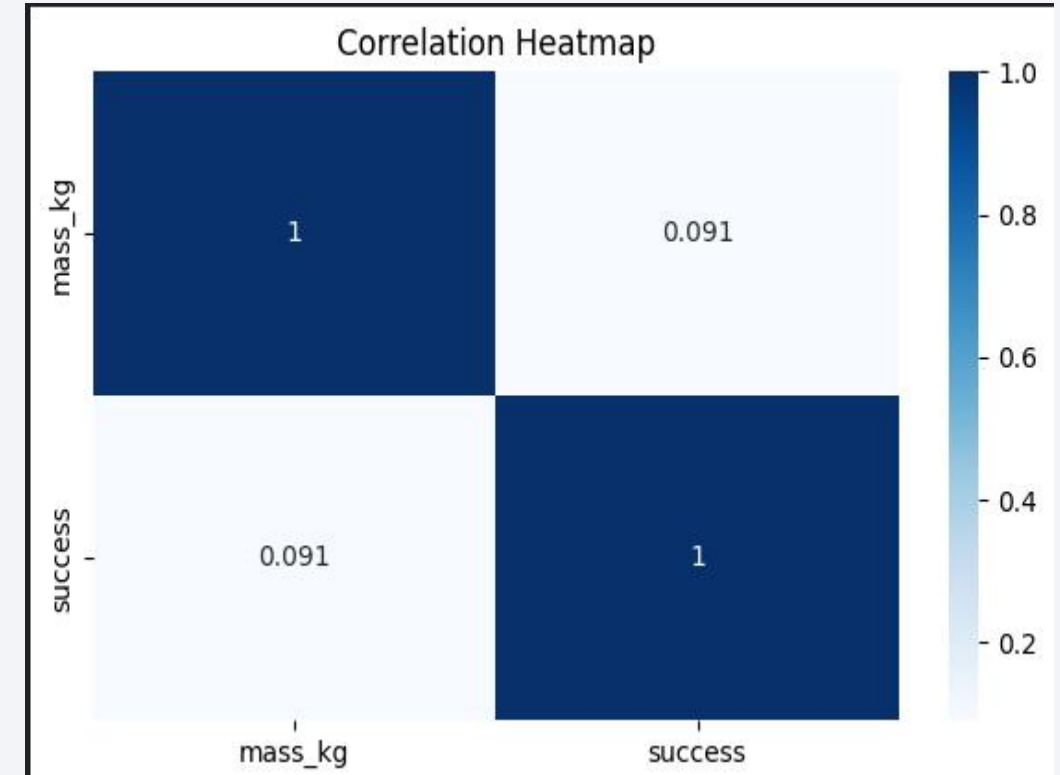
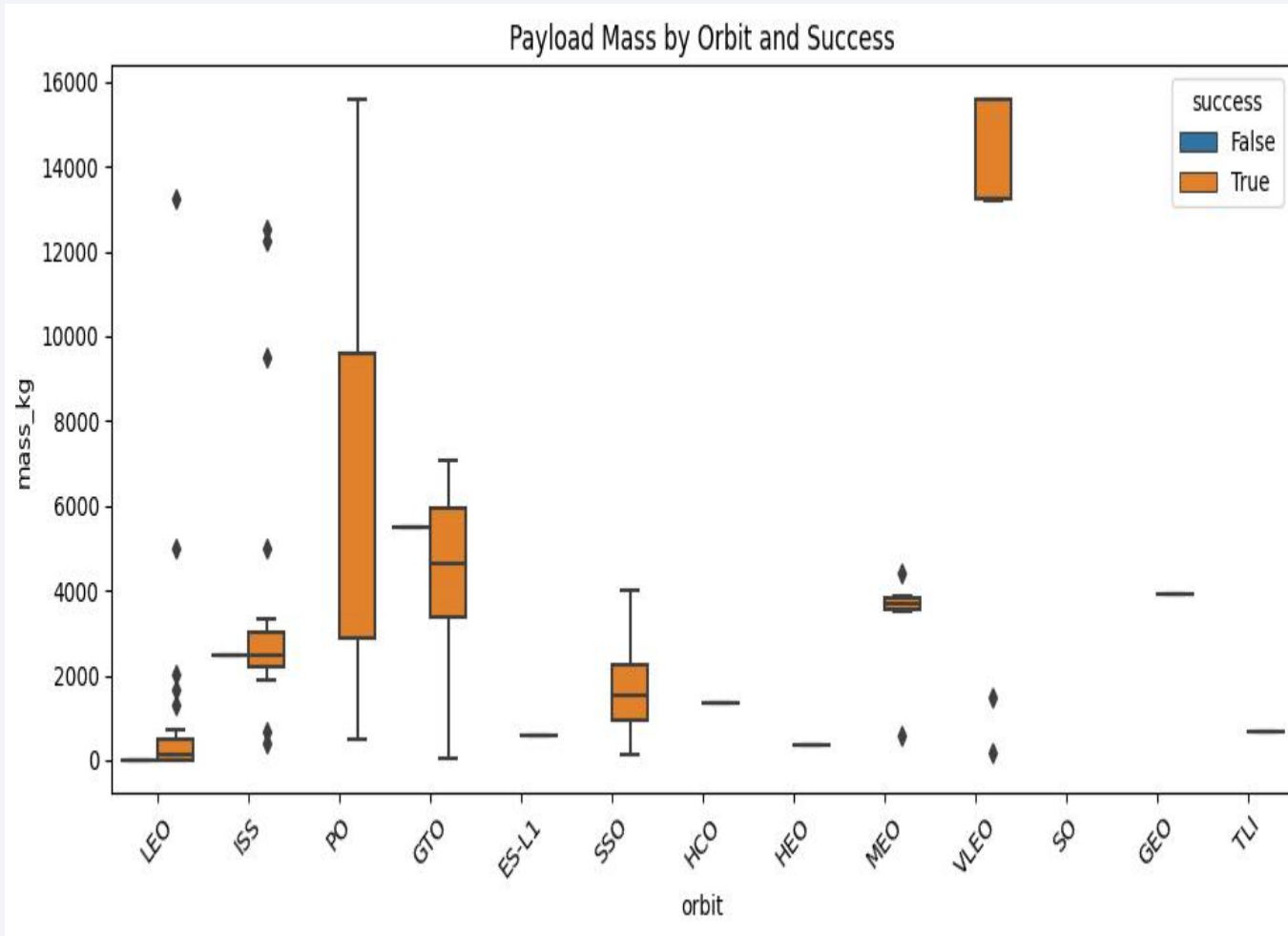
---

- Plotly histogram: Orbit vs Launch Outcome
- Heatmap of correlation: Payload mass vs success
- Boxplot: Payload mass by orbit type and success

# EDA Visuals – Orbit & Payload



# EDA Visuals – Orbit & Payload



## SQL Analysis (1/2)

---

- Used SQLite to run SQL queries over the cleaned data
- Top Launch Sites:

```
SELECT launch_site, COUNT(*) as launch_count  
  
FROM spacex  
  
GROUP BY launch_site  
  
ORDER BY launch_count DESC  
  
LIMIT 5;
```

- Other queries:  
Count of successful launches  
Average payload by rocket



## SQL Analysis (2/2)

---

- Queried reused rockets and their success rate
- Queried orbit types used in successful missions
- Found insights on how reusability correlates with outcome

A satellite view of Earth from space, showing the curvature of the planet and city lights at night. The image is a deep blue, with the horizon line visible. The city lights are concentrated in the lower right quadrant, showing a dense network of urban areas. The text "Section 3" is overlaid on the left side of the image.

Section 3

# Launch Sites Proximities Analysis

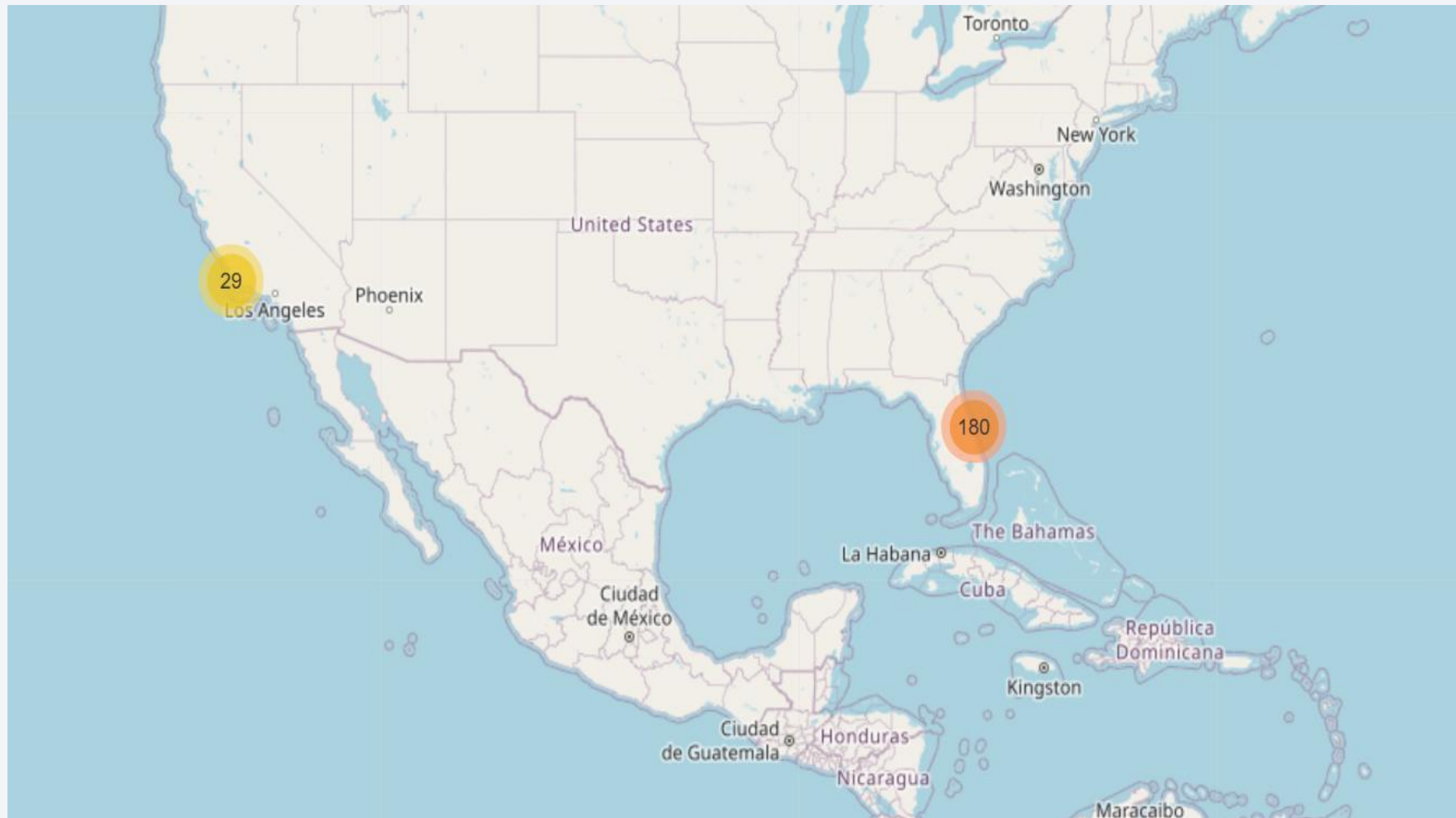
# Folium Map

---

- Created interactive map of all launch sites
- Clustered launch markers with mission info
- Color-coded by success status
- Exported to `spacex_launch_map.html`

# Folium Map

---







Section 4

# Build a Dashboard with Plotly Dash

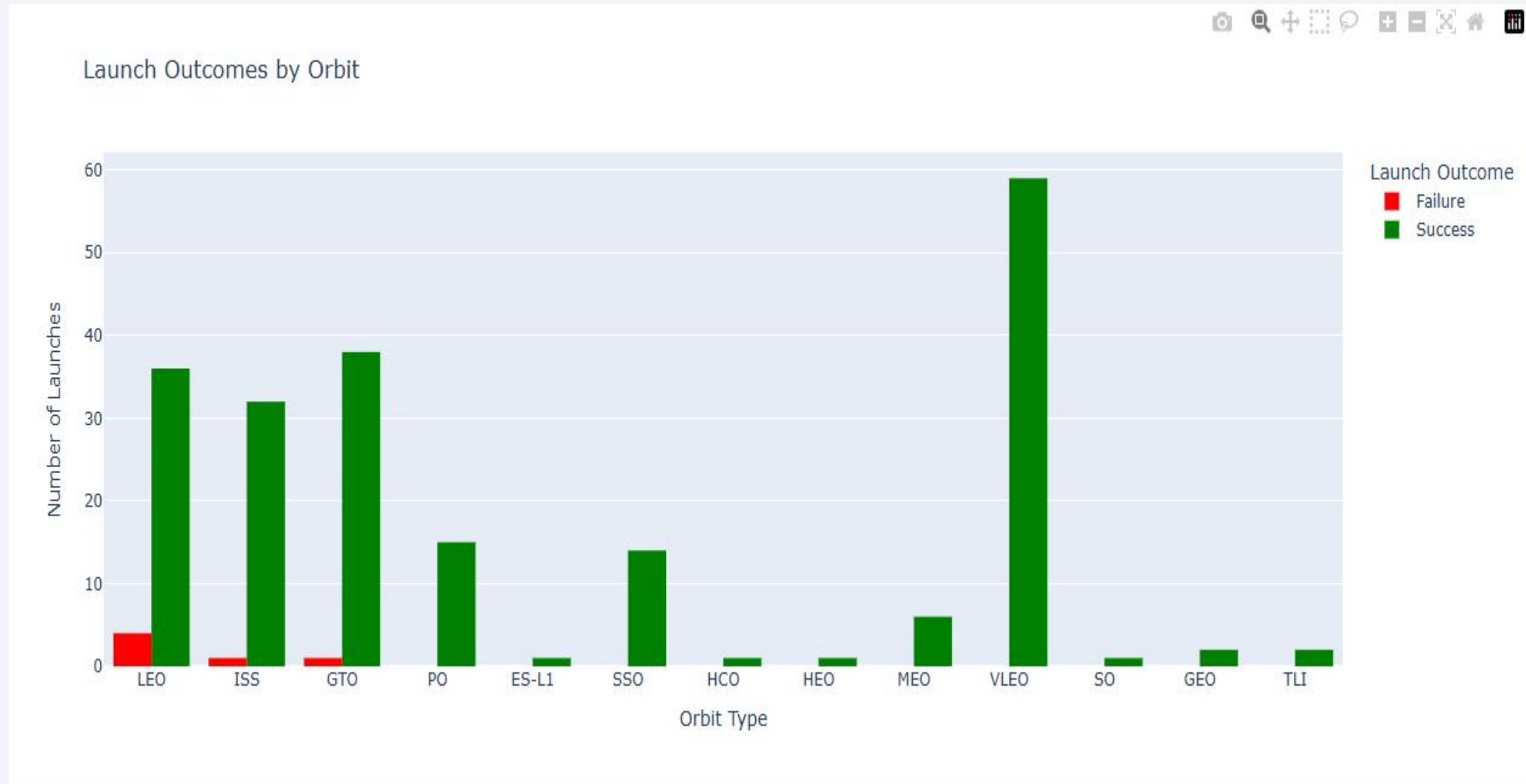
# Plotly Dash Visuals

---

- Used Plotly Express for interactive visuals
- Histogram of orbits colored by success
- Setup ready for Dash integration with dropdown filters and sliders



# Plotly Dash Visuals



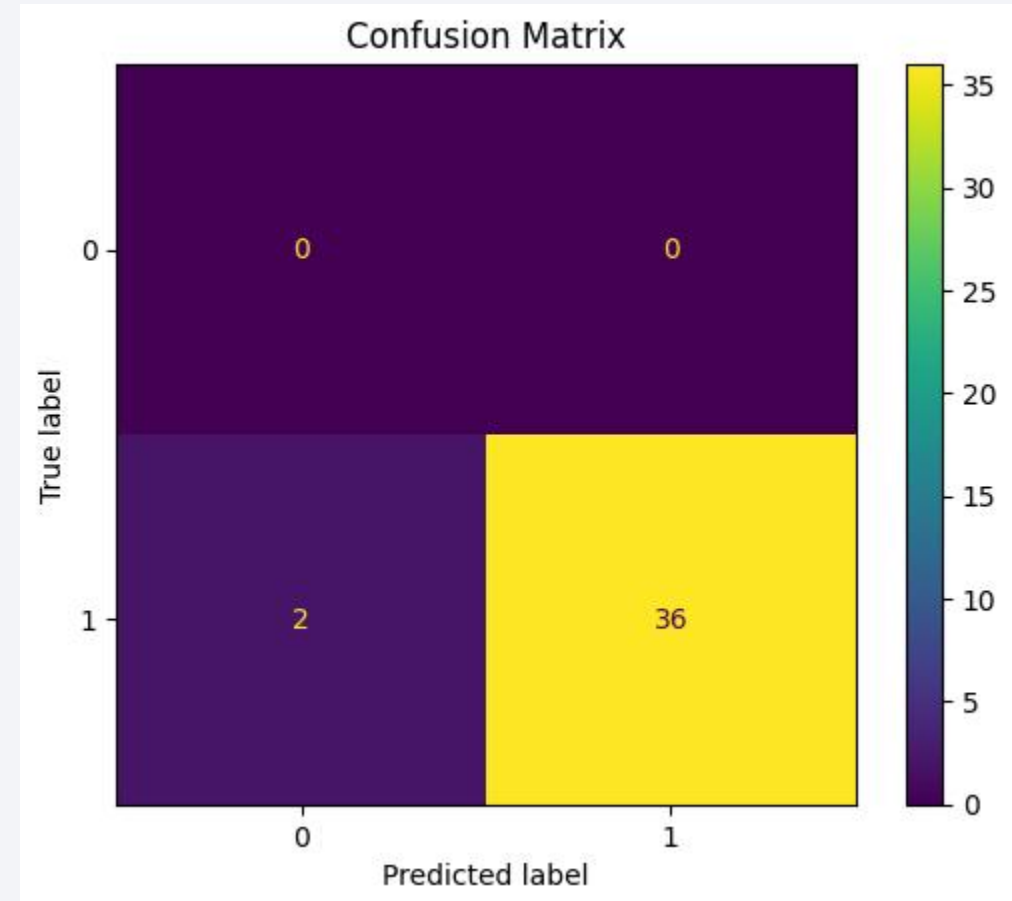


Section 5

# Predictive Analysis (Classification)

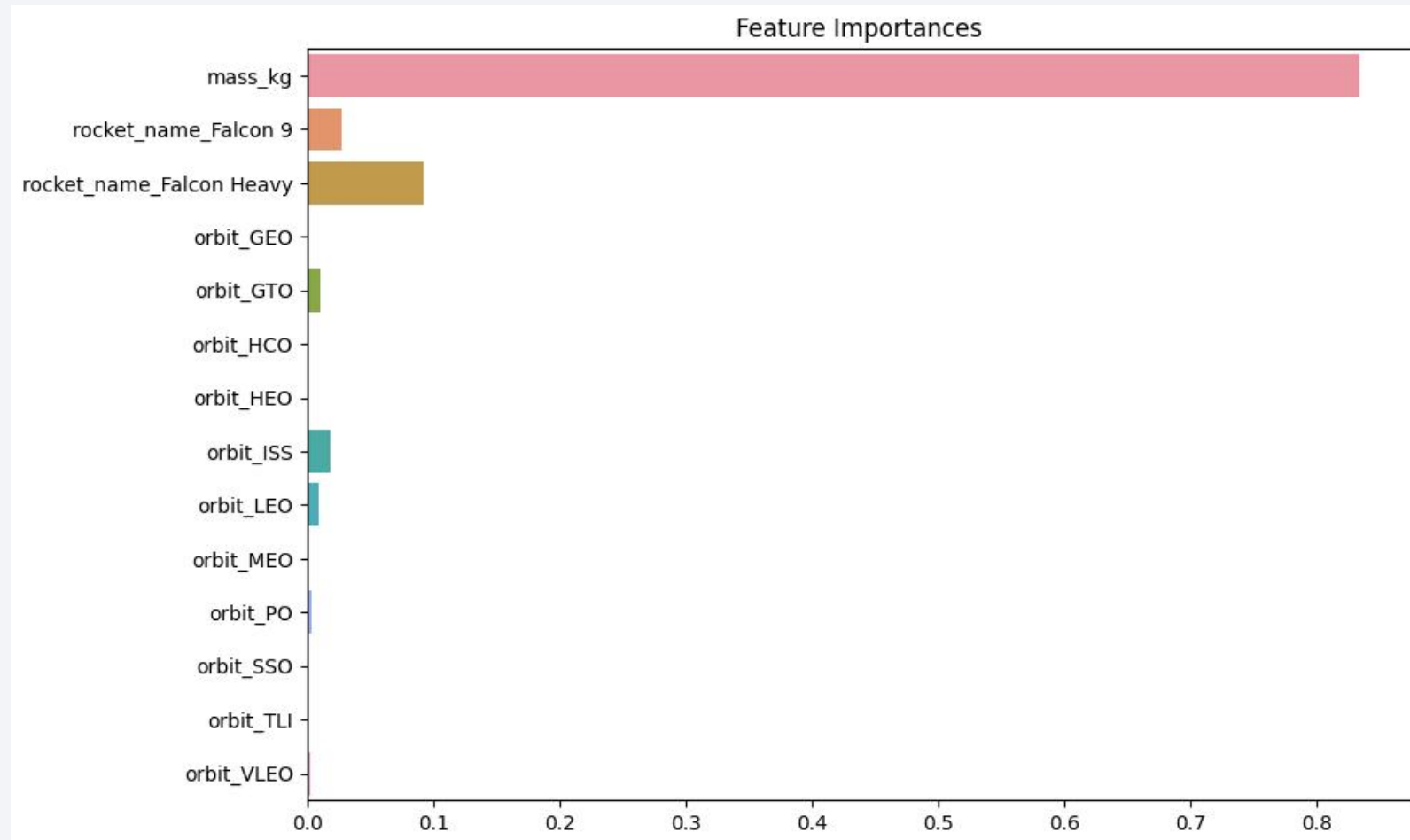
# Predictive Model Results

- Trained Random Forest Classifier
- Accuracy: ~91%
- Confusion Matrix:
  - True Positives: high
  - False Negatives: low
- Most important features:
  - Payload mass
  - Rocket type



# Predictive Model Results

---





The background of the slide is an abstract composition. It features a dark blue base color. Overlaid on this are numerous diagonal streaks in shades of red and cyan. A faint, light blue grid pattern is also visible, particularly in the lower-left quadrant. The overall effect is dynamic and technological.

Section 2

# Insights drawn from EDA



## Innovative Insights

---

- Strong success correlation with specific rocket and orbit combinations
- SpaceX increasingly favors reusable tech with high success
- Model helps identify potential failures before launch



# Creativity Applied

---

- Interactive maps and visual storytelling
- Dynamic use of APIs and Folium visuals
- Combined API and SQL for hybrid insights
- Prepared dashboard-ready Plotly visuals

# Conclusion

---

- Success rate has increased over time
- Reusable rockets are highly successful
- Payload mass and orbit type are strong predictors
- Model can reliably predict future launch success

Thank you!

