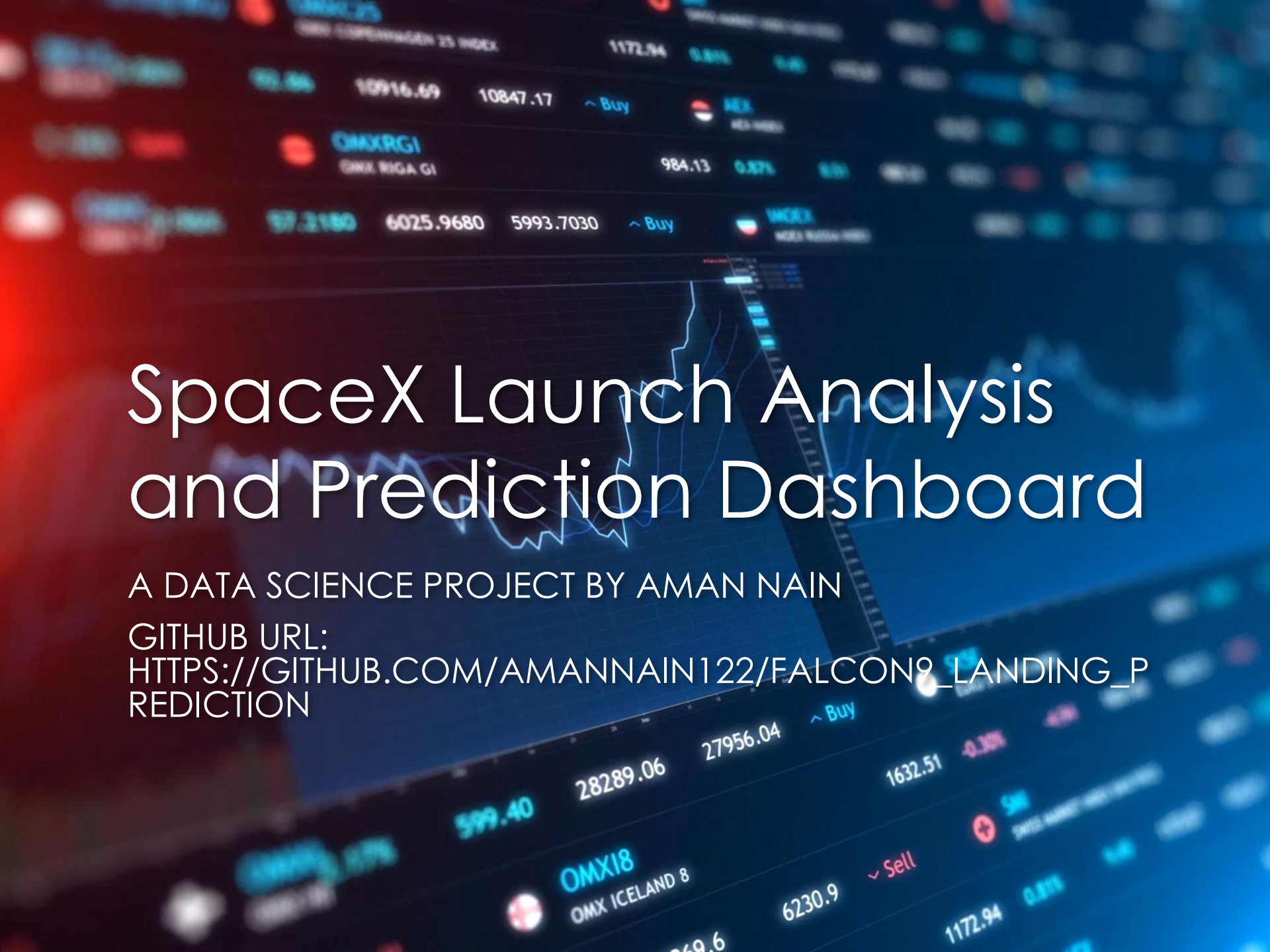


SpaceX Launch Analysis and Prediction Dashboard

A DATA SCIENCE PROJECT BY AMAN NAIN

GITHUB URL: HTTPS://GITHUB.COM/AMANNAIN122/FALCON9_LANDING_PREDICTION

Executive Summary

- Objective: Analyze SpaceX launch data to uncover insights and predict launch outcomes.
- Key Outcomes: Launch site success rates, payload success correlations, F9 Booster performance, and a predictive model for launch success.

Introduction

- Problem Statement: How can SpaceX optimize future launches based on historical data?
- Goals: Use data science techniques for analysis, build interactive dashboards, and predict launch outcomes.
- Data Sources: API, web scraping, and public datasets.

Data Collection and Wrangling Methodology

- Collection Methods: API access and web scraping to gather launch records and payload information.
- Wrangling Techniques: Handling missing values, data cleaning, and transformations to prepare for analysis.

EDA and Interactive Visual Analytics Methodology

- Techniques: Examining distributions, correlations, and launch site success rates.
- Interactive Elements: Dropdowns, sliders, and Folium maps to explore data interactively.

Predictive Analysis Methodology

- Model Selection: Classification models to predict success rates.
- Data Split: Training and testing data split with feature selection based on payload and booster characteristics.

EDA with Visualization Results

- Visualizations: Key charts and findings:
- - Launch success distribution by site.
- - Payload mass vs. success rate correlation.
- Key Insights: Sites with the highest success rates and significant payload ranges affecting success.

EDA with SQL Results

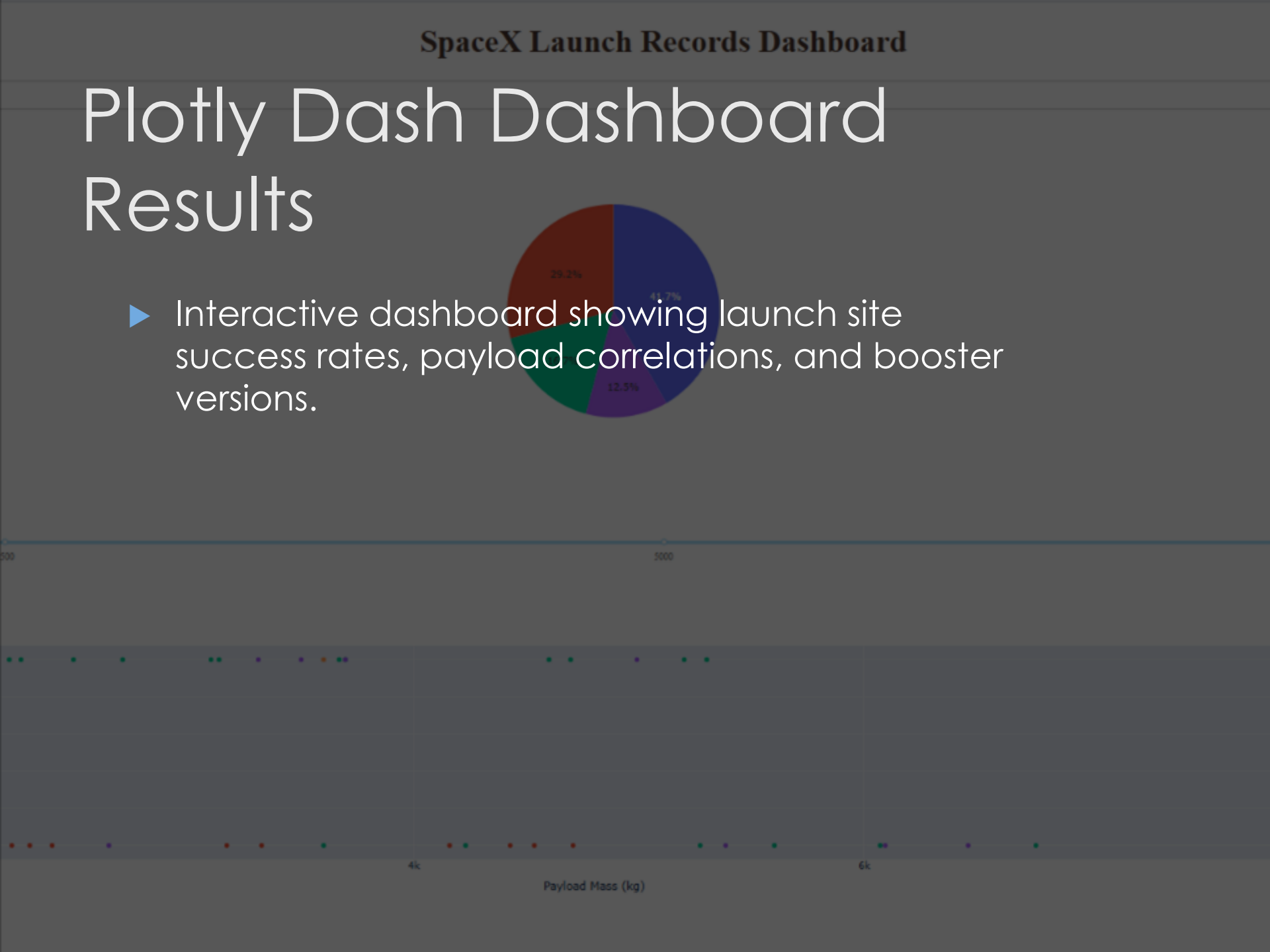
- SQL Analysis: Queries for site-specific success rates and booster performance.
- Findings: Insights into launch sites with the most successful launches and booster types with high performance.

Interactive Map with Folium Results

- Map Insights: Displayed launch sites with success and failure markers.
- Interpretation: Geographical trends or site performance differences.

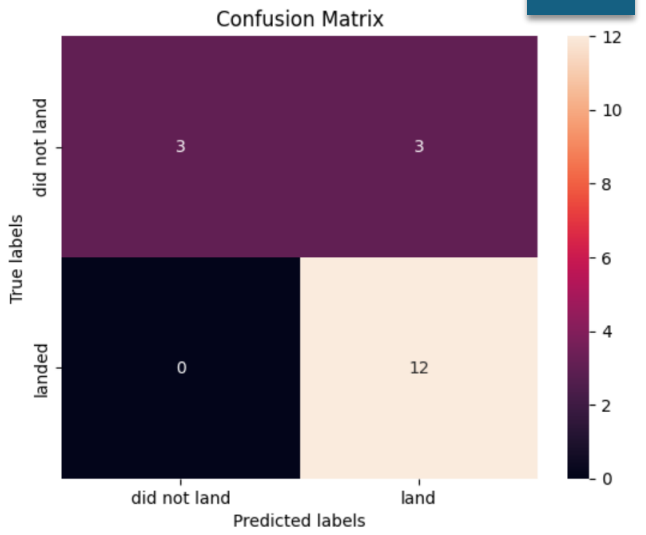
Plotly Dash Dashboard Results

- Interactive dashboard showing launch site success rates, payload correlations, and booster versions.



Predictive Analysis (Classification) Results

- Best Model: Decision Tree model with 88.89% accuracy.
- Confusion Matrix and model evaluation details.




```
results = {
    "Logistic Regression": logreg_test_accuracy,
    "SVM": svm_test_accuracy,
    "Decision Tree": tree_test_accuracy,
    "KNN": knn_test_accuracy
}

best_model = max(results, key=results.get)
print("Best performing model:", best_model, "with accuracy:", results[best_model])
```

```
Best performing model: Decision Tree with accuracy: 0.8888888888888888
```

Conclusion

With lesser data I was able to discover that with higher payload the landing is more sure.

Inclusion of more data and better predictive modelling will help me gain more insights about the data.