SpaceX Falcon 9 First Stage Landing Prediction

SON VUONG

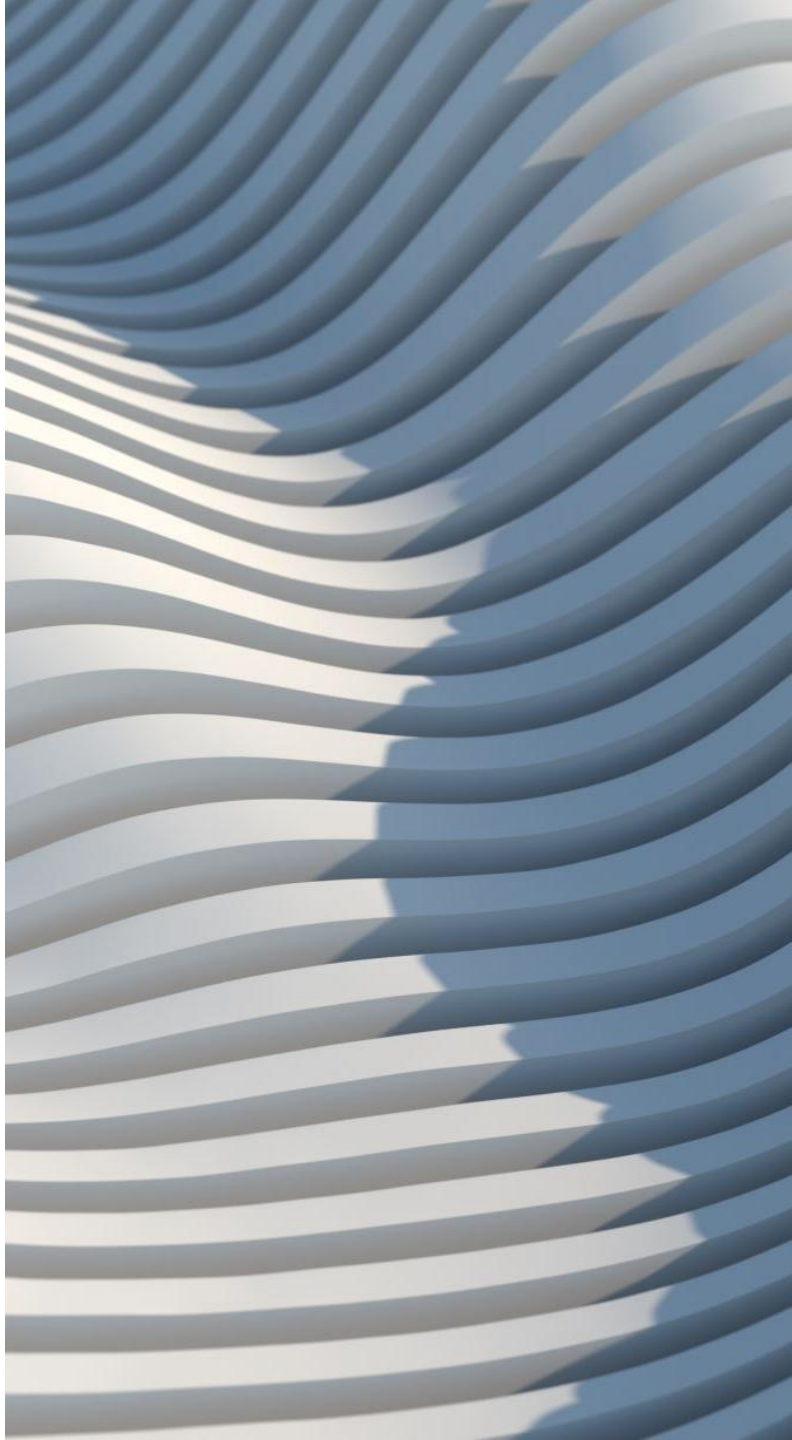
Data Science Capstone Project

Executive Summary

- GOAL: PREDICT FALCON 9 FIRST STAGE LANDING OUTCOME

- METHODS: DATA WRANGLING, EDA, MACHINE LEARNING

- RESULT: BEST MODEL ACHIEVED > 80% ACCURACY



Introduction

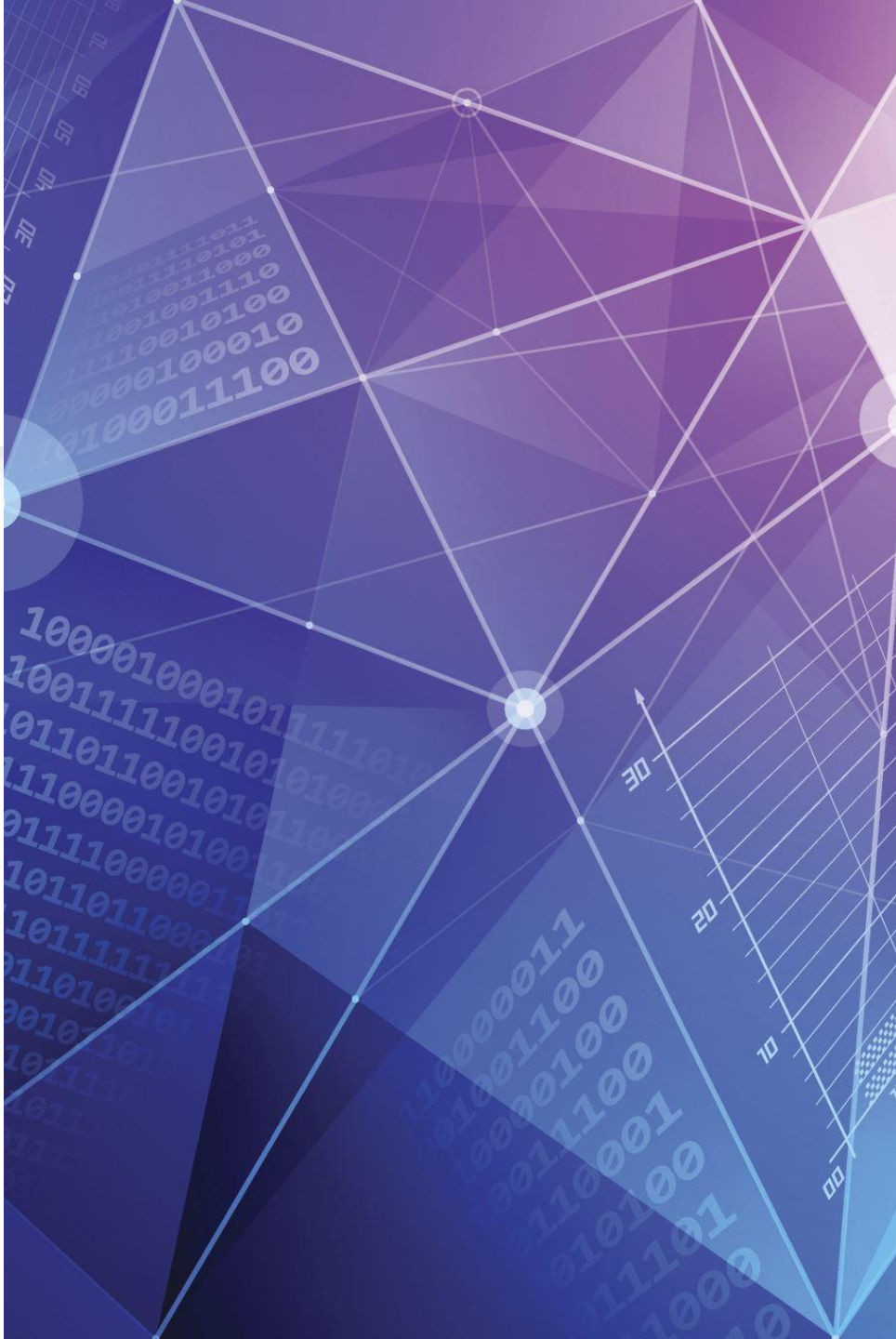
- SpaceX aims to reuse rockets to reduce costs

- Objective: Predict landing success using mission data

- Tools: Python, SQL, Folium, Plotly Dash

Data Collection & Wrangling

- • Data from API and Wikipedia tables
- • Cleaned missing values, engineered labels
- • One-hot encoding on categorical variables



EDA & Visual Analytics

- • Bar plots, boxplots, scatterplots
- • Interactive charts using Plotly
- • Correlation heatmap for feature analysis

Predictive Modeling Approach

- • Models: Logistic Regression, SVM, Decision Tree, KNN
- • Used GridSearchCV for hyperparameter tuning
- • Metrics: Accuracy, confusion matrix

EDA Visualization Results

- • Launch success rate by site and orbit
- • Payload impact on landing outcome
- • Time trend of mission successes

SQL Query Results

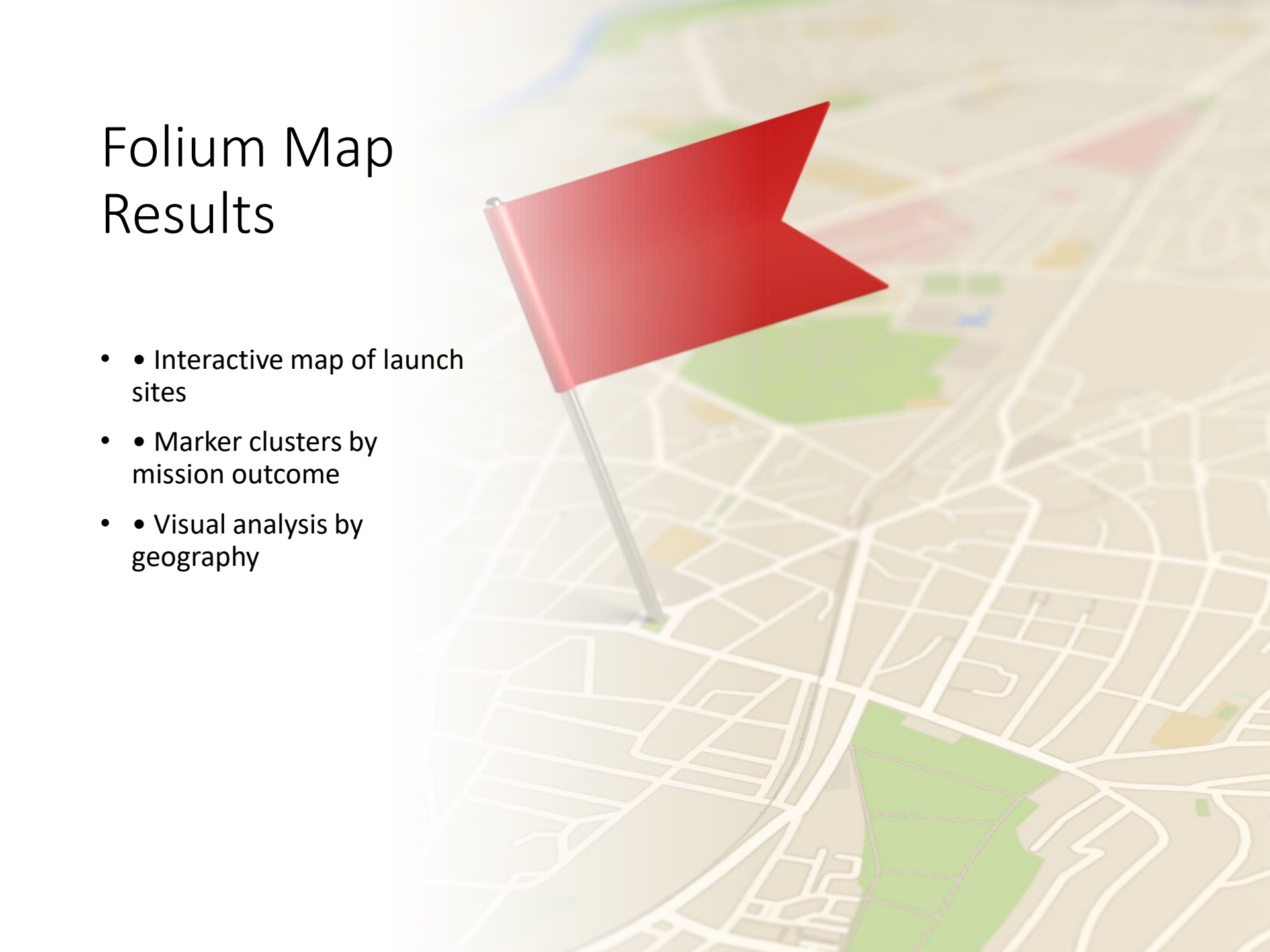
- Launches per site and orbit type

- Min/max payload analysis

- Joined mission outcome queries

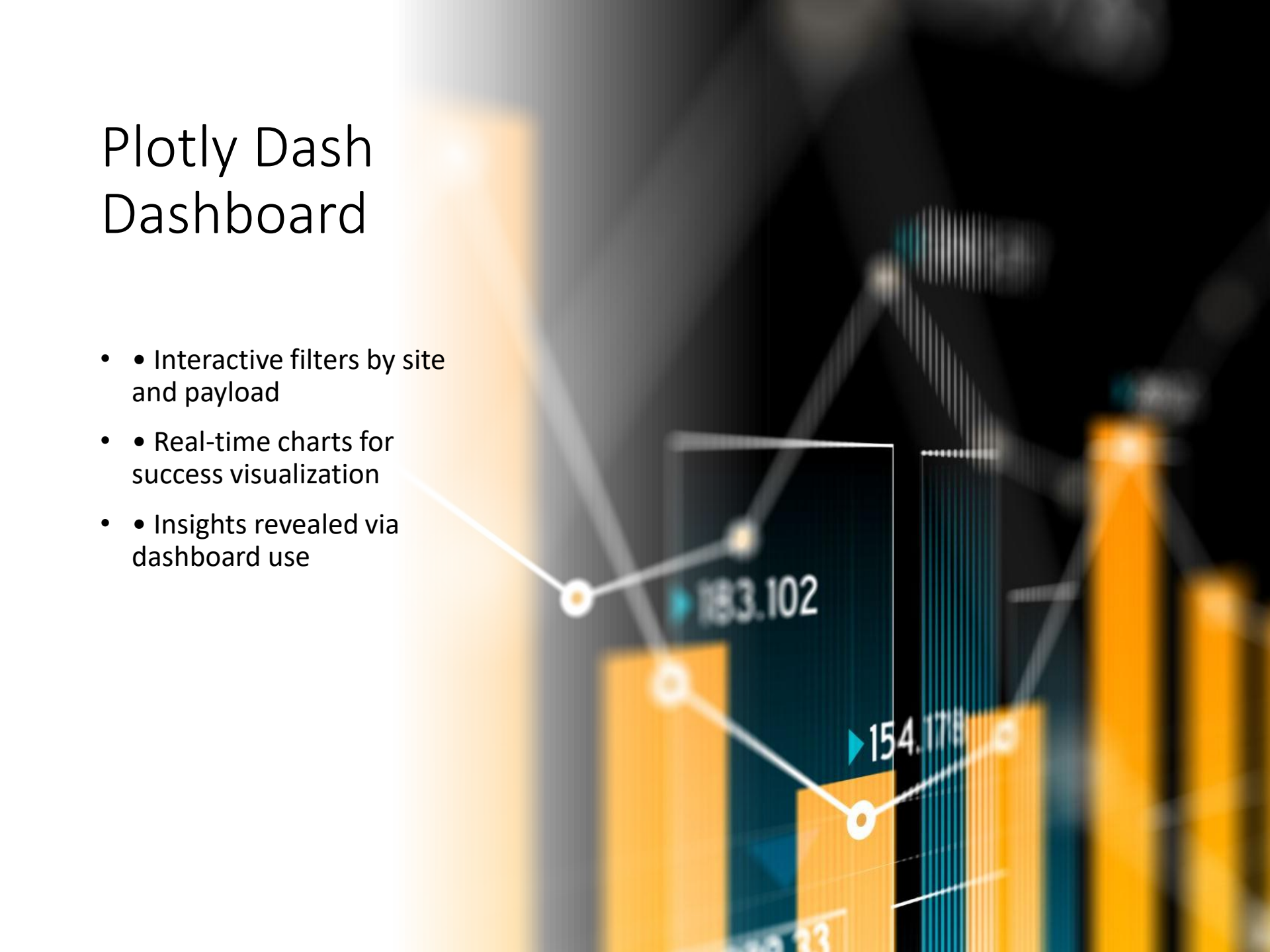
Folium Map Results

- • Interactive map of launch sites
- • Marker clusters by mission outcome
- • Visual analysis by geography



Plotly Dash Dashboard

- • Interactive filters by site and payload
- • Real-time charts for success visualization
- • Insights revealed via dashboard use



Classification Results

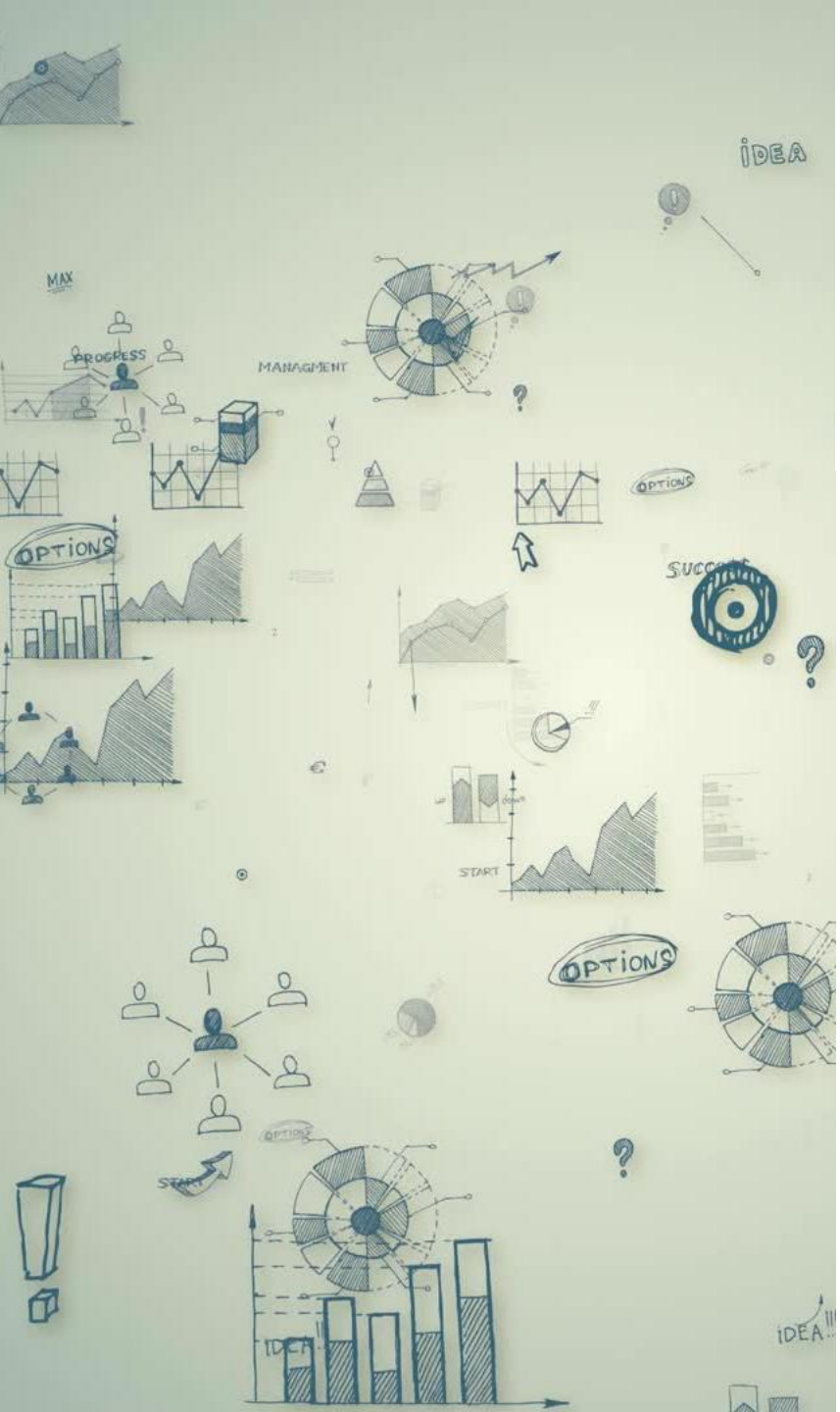
- • Model comparison: accuracy and confusion matrices
- • Best model: Logistic Regression (example)
- • Additional: ROC curve, feature importance

Conclusion

- Achieved reliable prediction of landing success

- Feature analysis supports SpaceX decision making

- Future work: Regression for cost prediction



Creativity & Innovation

- • Dashboard with interactive sliders
- • Folium map integration
- • Custom visual insights beyond template