



APPLIED DATA SCIENCE CAPSTONE
IBM DATA SCIENCE PROFESSIONAL
CERTIFICATE (COURSERA)

FALCON 9

FIRST STAGE LANDING PREDICTION

KAUNG HTET HEIN

EXECUTIVE SUMMARY



This is an applied data science capstone project of IBM data science professional certificate. The aim of this project is to perform predictive analysis on SpaceX's Falcon 9 launches' first stage landing. Data collection, EDA, interactive visualization, & predictive analysis will be performed using machine learning models.

TABLE OF CONTENTS

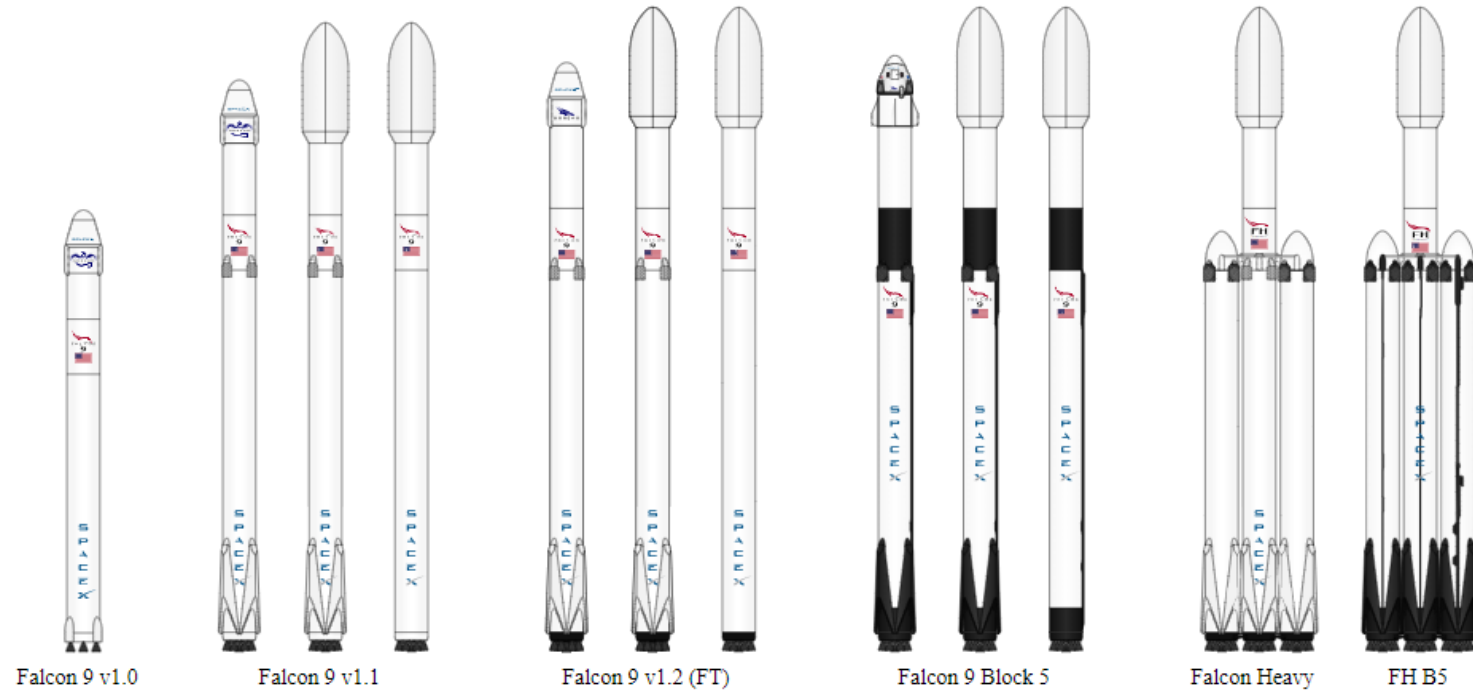
- Introduction
- Methodologies
- Results
- Conclusion



INTRODUCTION

- To predict if the Falcon 9 first stage will land successfully.
- SpaceX advertises Falcon 9 rocket launches cost 62 million dollars.
- Other providers cost upward of 165 million dollars each.
- Much of the savings is because SpaceX can reuse the first stage.
- If we can determine if the first stage will land, we can determine the cost of a launch.

FALCON 9 & FALCON HEAVY LAUNCHES



METHODOLOGIES

- Data collection
- Data wrangling

- EDA
- Interactive visual analytics

- Predictive analysis

DATA COLLECTION & DATA WRANGLING

Methods Utilized:

- RESTful API
- Web scrapping

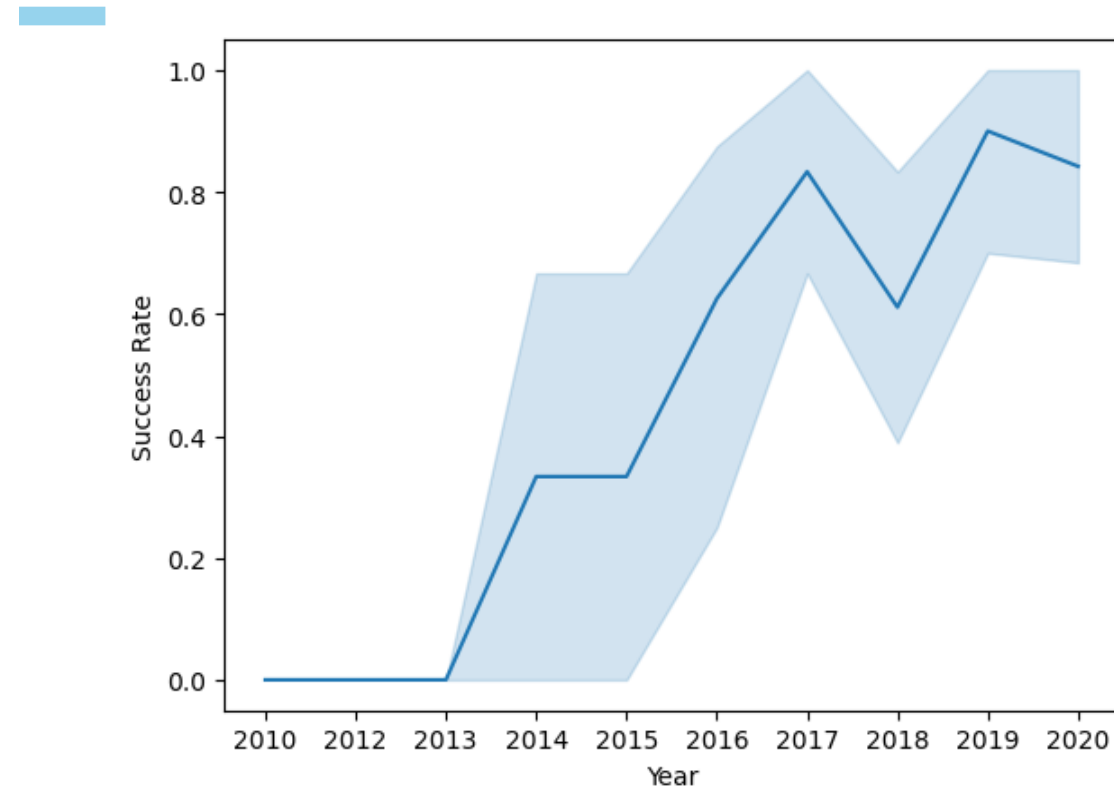
EDA & INTERACTIVE VISUAL ANALYTICS

Exploratory Data Analysis
Libraries:

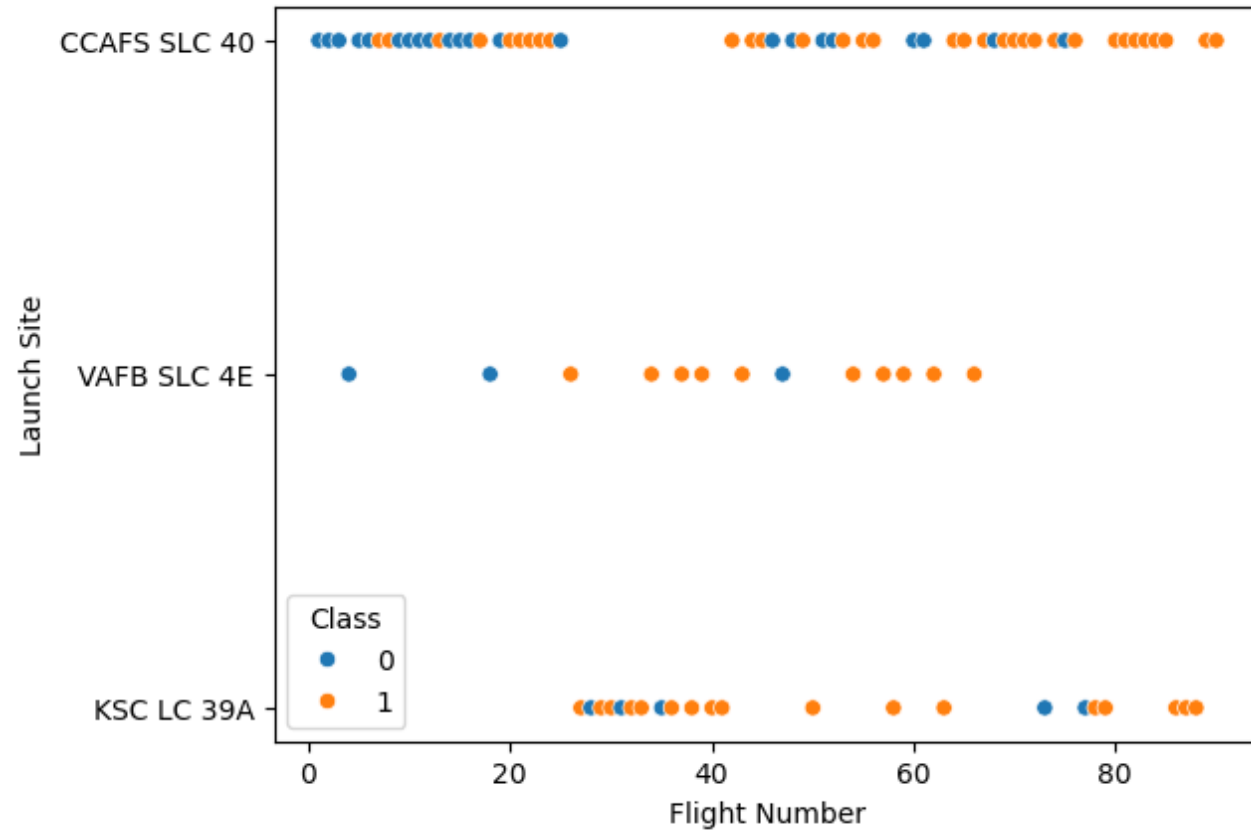
- SQLite
- Pandas
- Matplotlib
- Seaborn

Interactive Visual Analytics
Libraries:

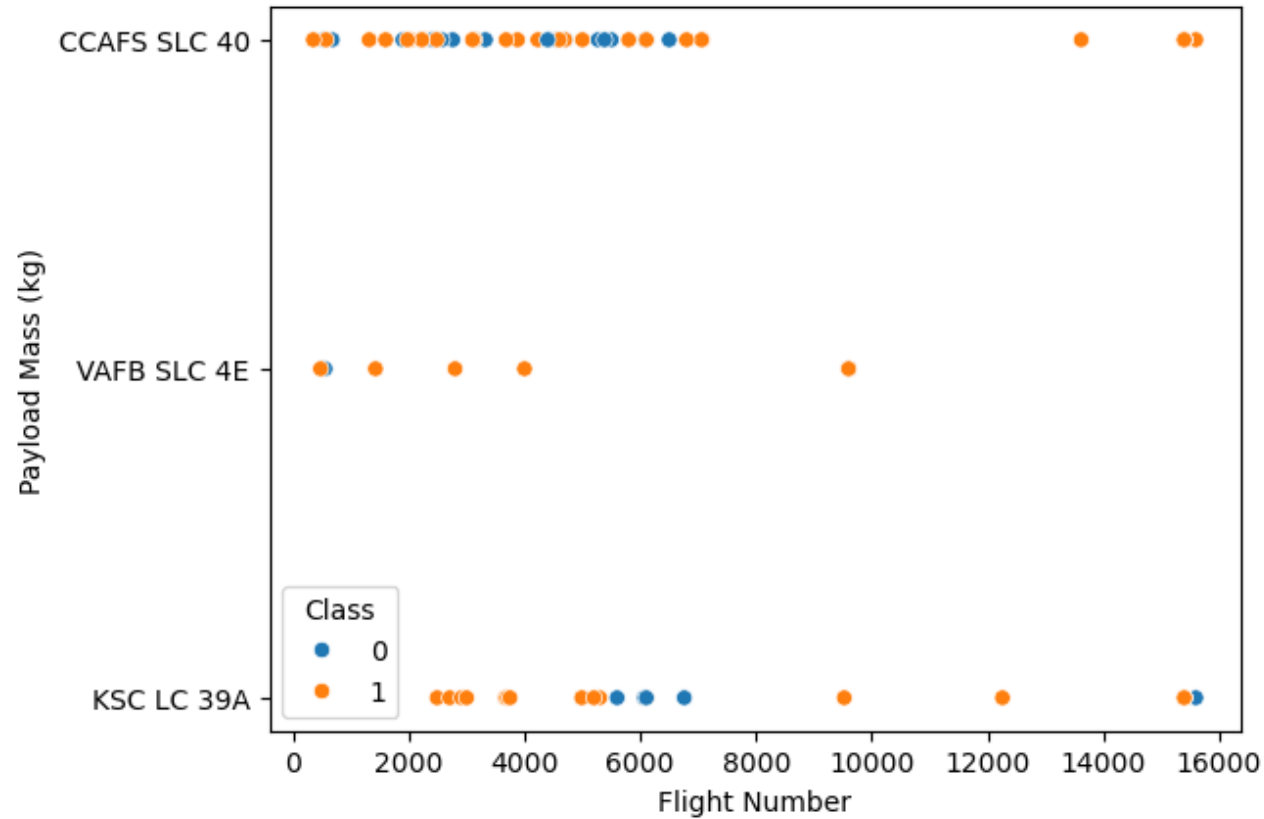
- Folium
- Plotly

EDA
VISUALISATION
RESULTSSUCCESS RATE OF FIRST
STAGE LANDING OVER YEARS

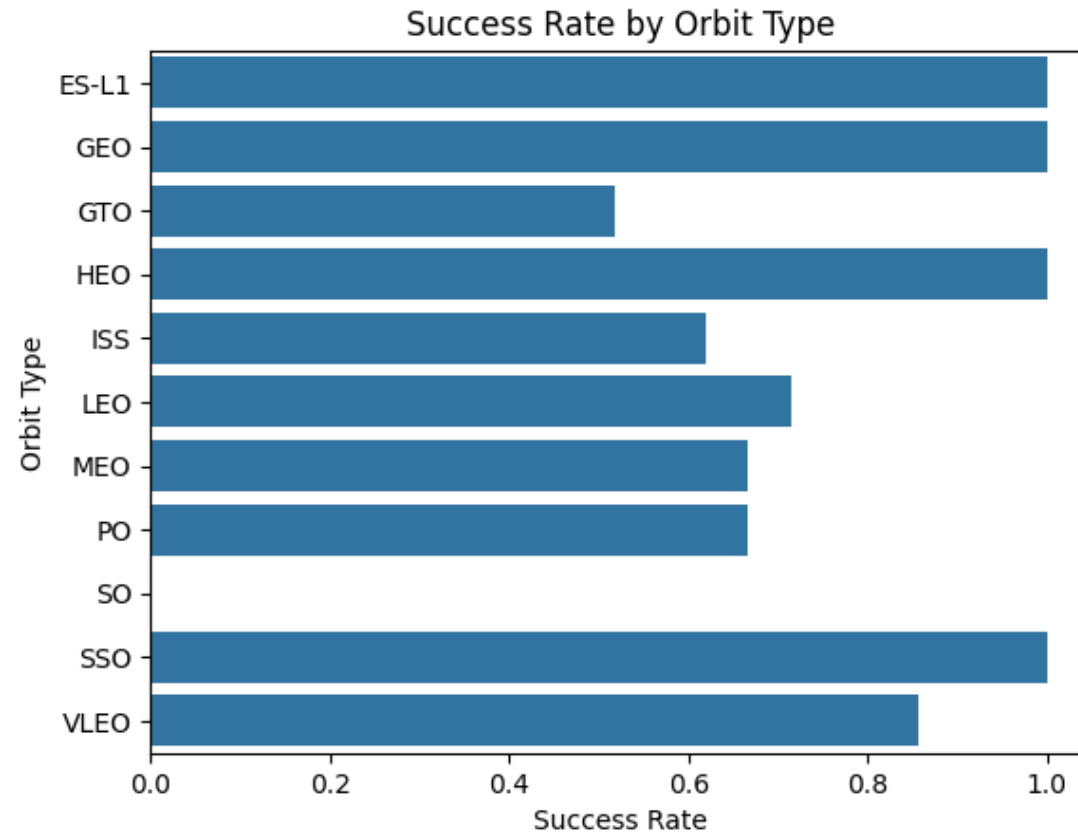
FLIGHT NUMBER VS LAUNCH SITE



FLIGHT NUMBER VS PAYLOAD MASS(KG)



SUCCESS RATE VS ORBIT TYPE



SQL ANALYTICS RESULTS

Unique Launch Sites

Launch_Site

CCAFS LC-40

VAFB SLC-4E

KSC LC-39A

CCAFS SLC-40

Total Payload Mass (KG)

total_payload_mass

45596

First Successful Landing

MIN("Date")

2015-12-22

SQL ANALYTICS RESULTS

Successful Drone Ship
Landings

Booster_Version

F9 FT B1022

F9 FT B1026

F9 FT B1021.2

F9 FT B1031.2

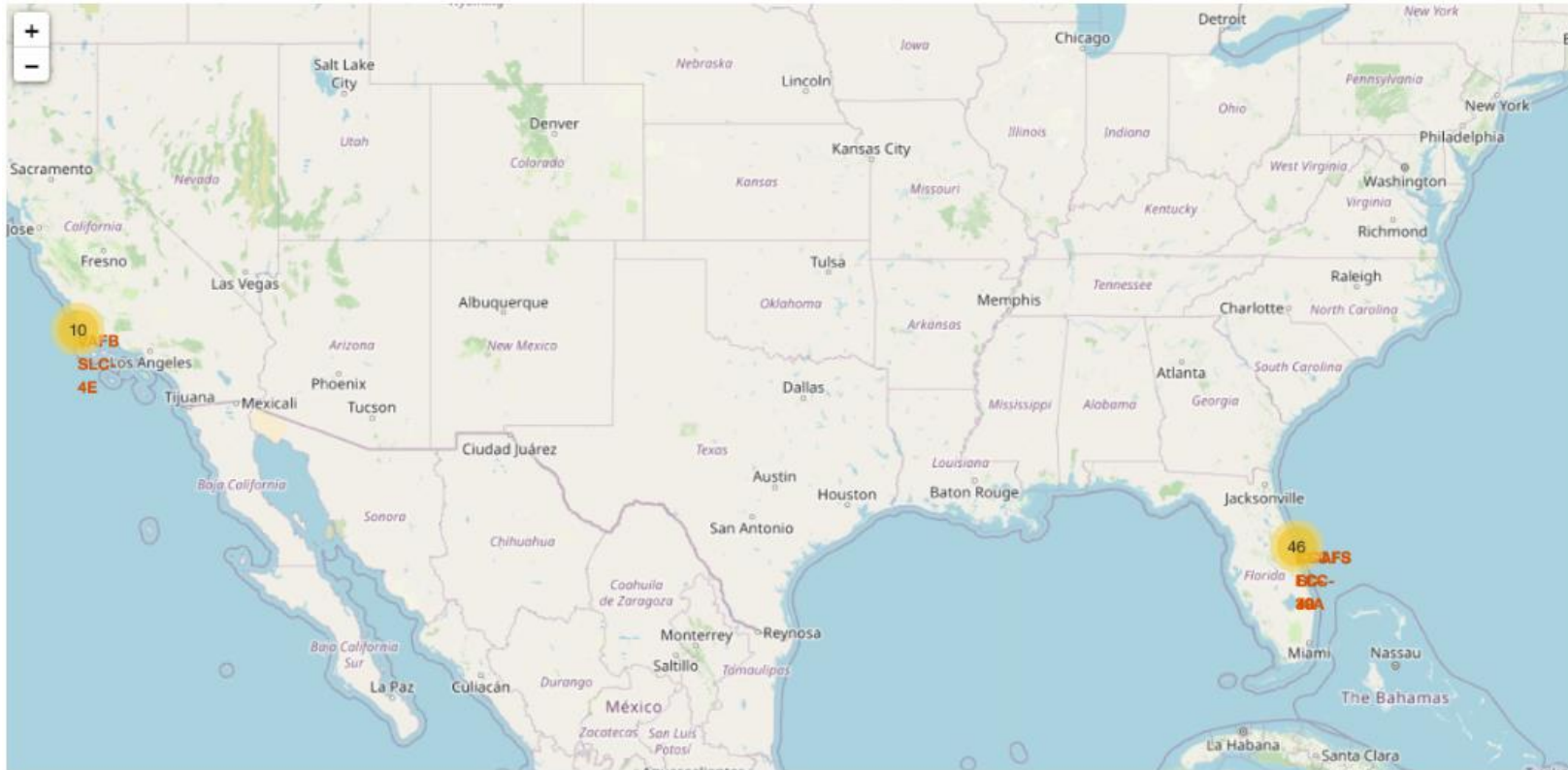
Total Successful
Missions

Mission_Outcome **total_count**

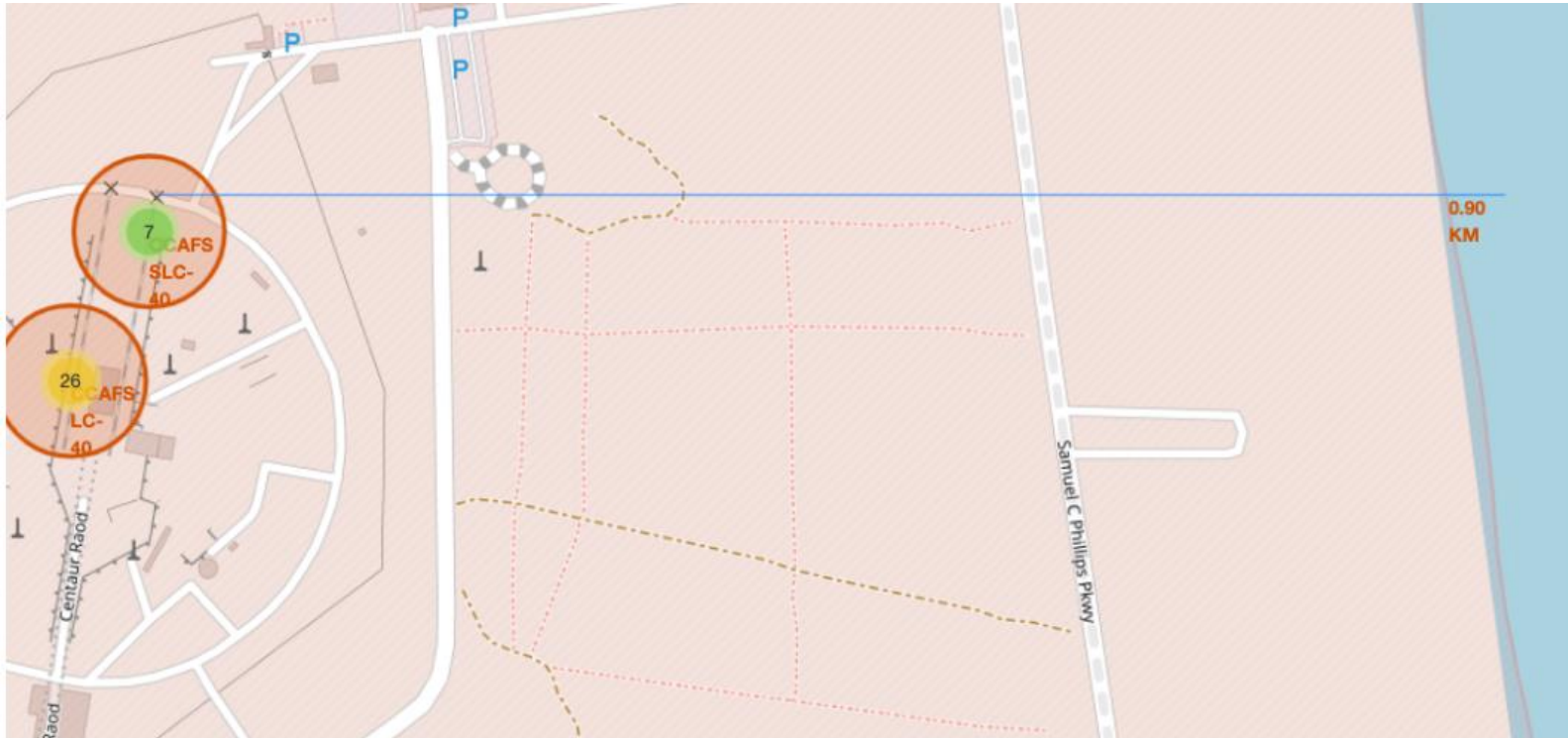
Success

98

LAUNCH SITE LOCATIONS IN USA



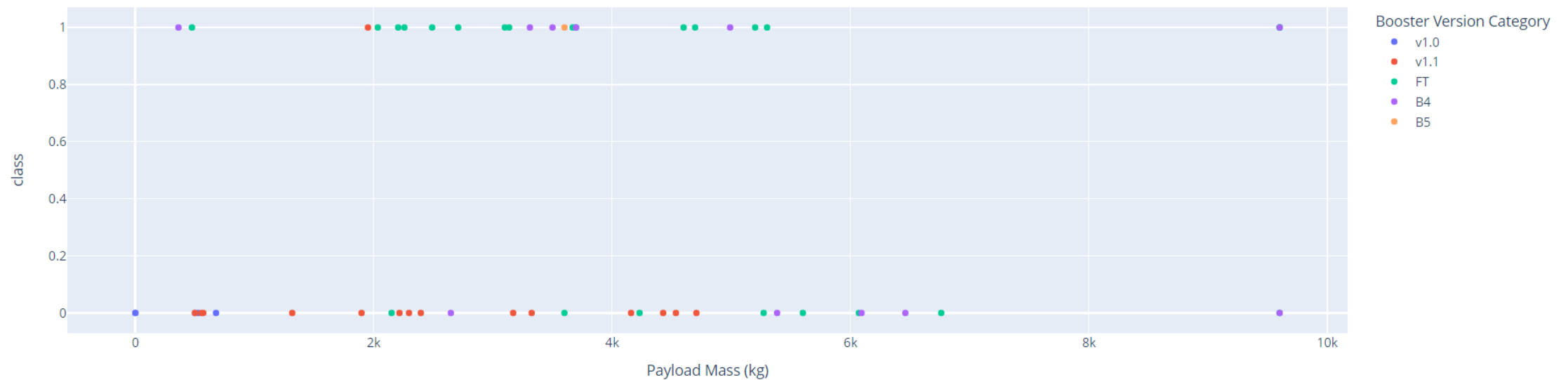
NEAREST COASTAL LINE FROM A SELECTED LAUNCH SITE



TOTAL LAUNCHES FOR ALL SITES



BOOSTER VERSIONS, PAYLOAD MASS & THEIR SUCCESS RATE



PREDICTIVE ANALYSIS

Machine Learning Models:

- Logistic regression
- Support Vector Machine (SVM)
- Decision Tree Classifier
- K-Nearest Neighbours (KNN)

Hyperparameters Tuning:

- Grid Search CV

Evaluation Method:

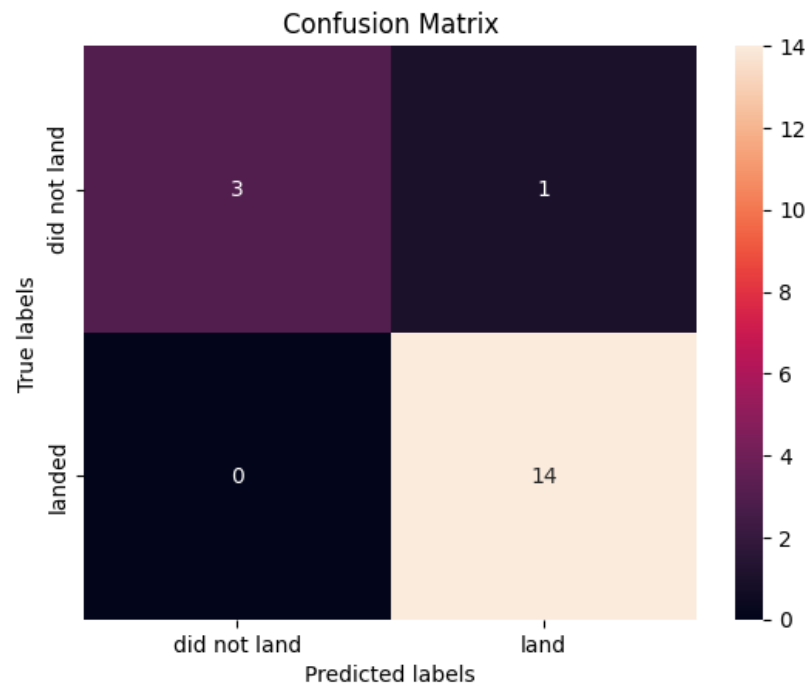
- Confusion Matrix

PREDICTIVE ANALYSIS (CLASSIFICATION) RESULTS

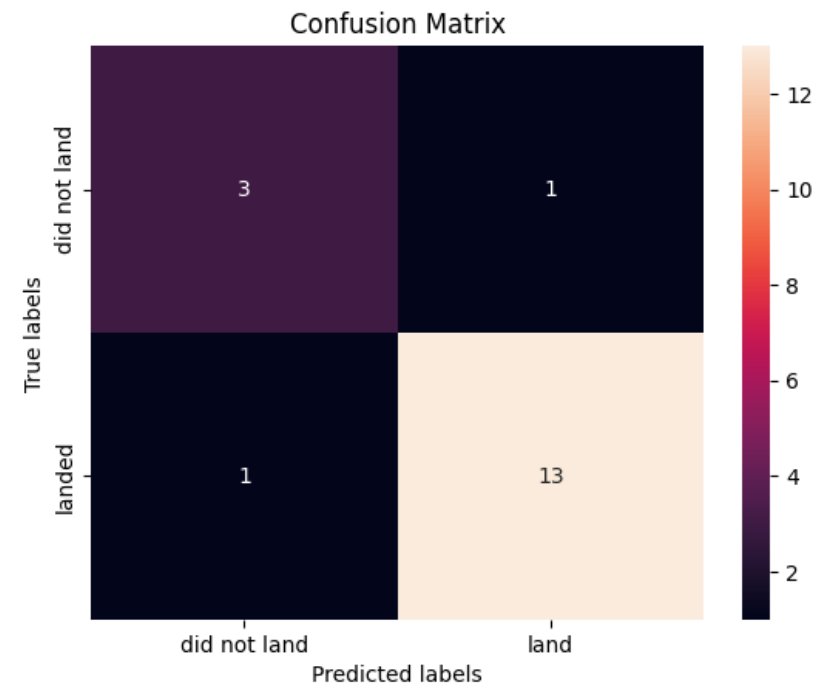
Model No.	Machine Learning Models	Accuracy on Test data
1	Logistic Regression	0.9444
2	Support Vector Machine (SVM)	0.8888
3	Decision Tree Classifier	0.9444
4	K-Nearest Neighbours (KNN)	0.9444

CONFUSION MATRICES

Logistic Regression

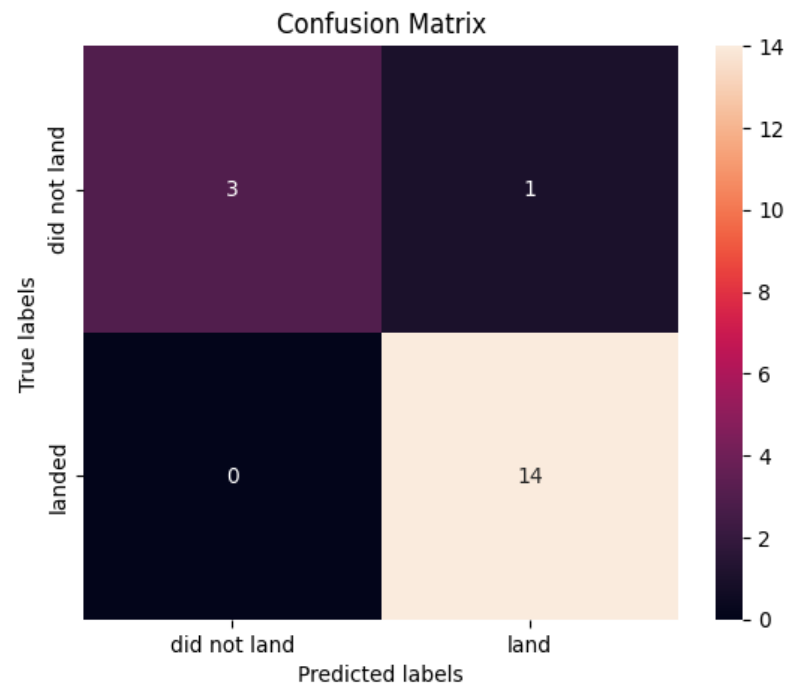


Support Vector Machine

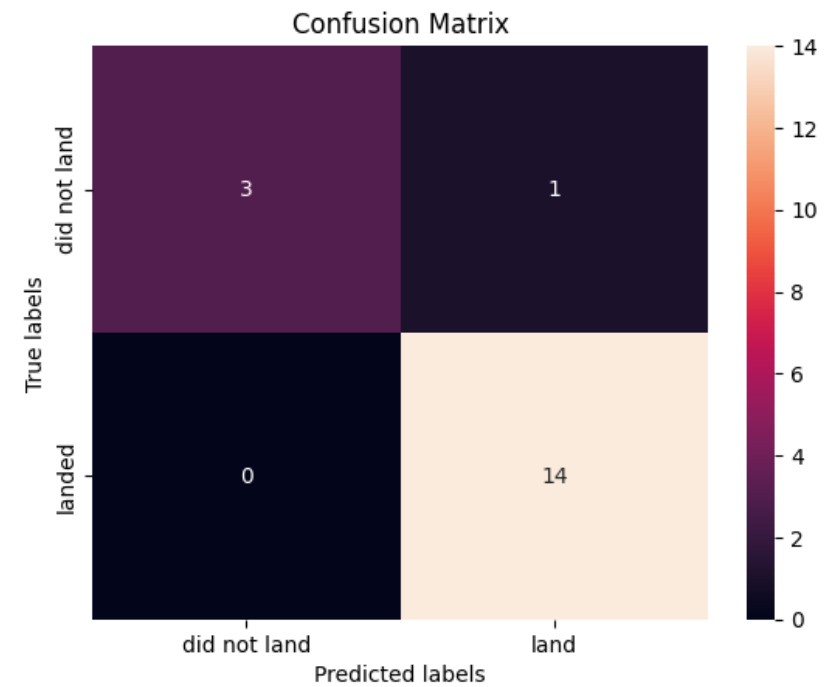


CONFUSION MATRICES

Decision Tree Classifier



K-Nearest Neighbours



CONCLUSION

Main Factors that contributed successful landings:

- Location
- Payload Mass
- Orbit Type



“To revolutionize space technology, with the ultimate goal of enabling people to live on other planets.” ~ SpaceX

THANK YOU 

APPLIED DATA SCIENCE CAPSTONE PROJECT (IBM DATA SCIENCE PROFESSIONAL CERTIFICATE)
BY KAUNG HTET HEIN