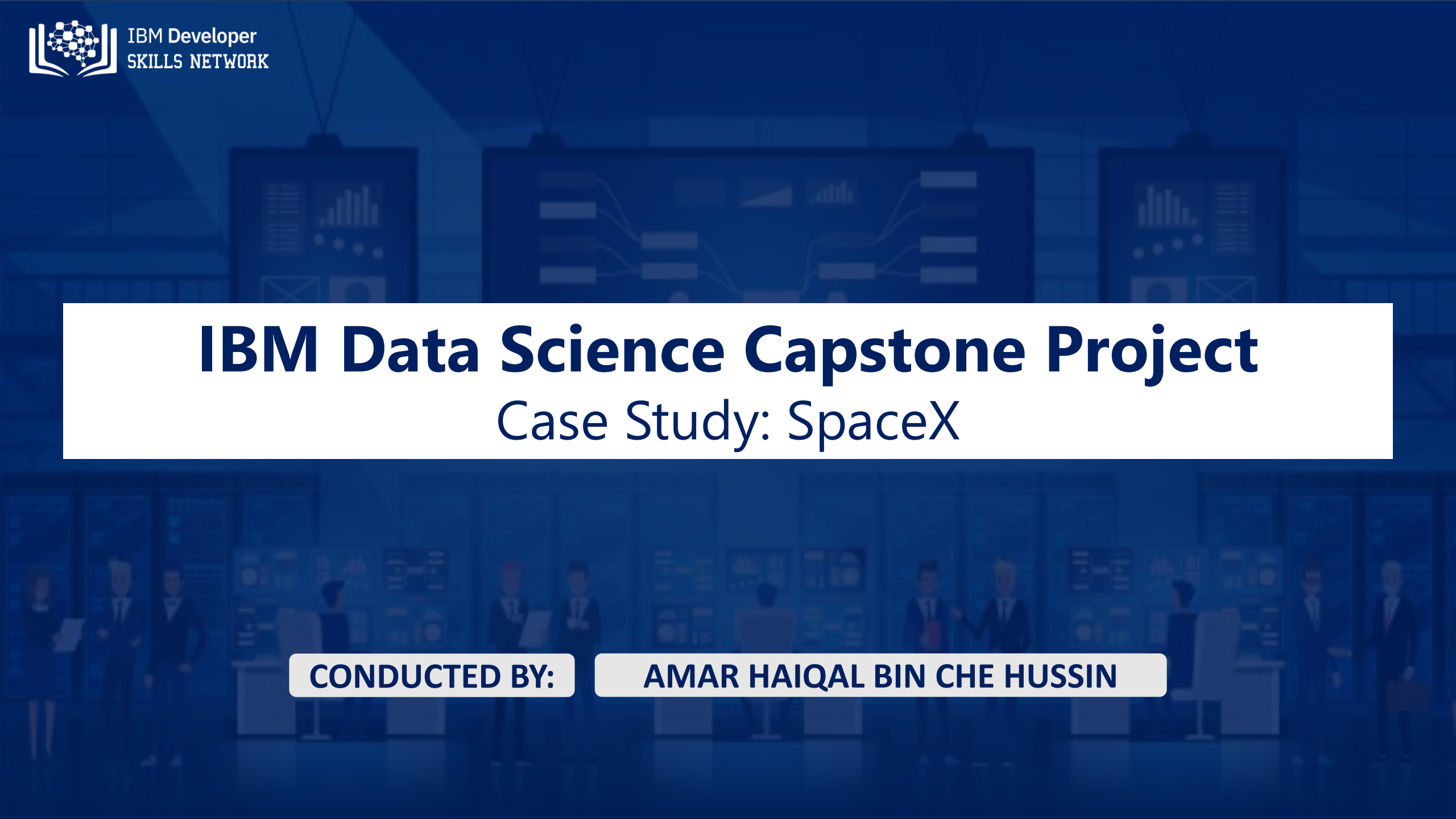


IBM Data Science Capstone Project

Case Study: SpaceX

CONDUCTED BY: AMAR HAIQAL BIN CHE HUSSIN

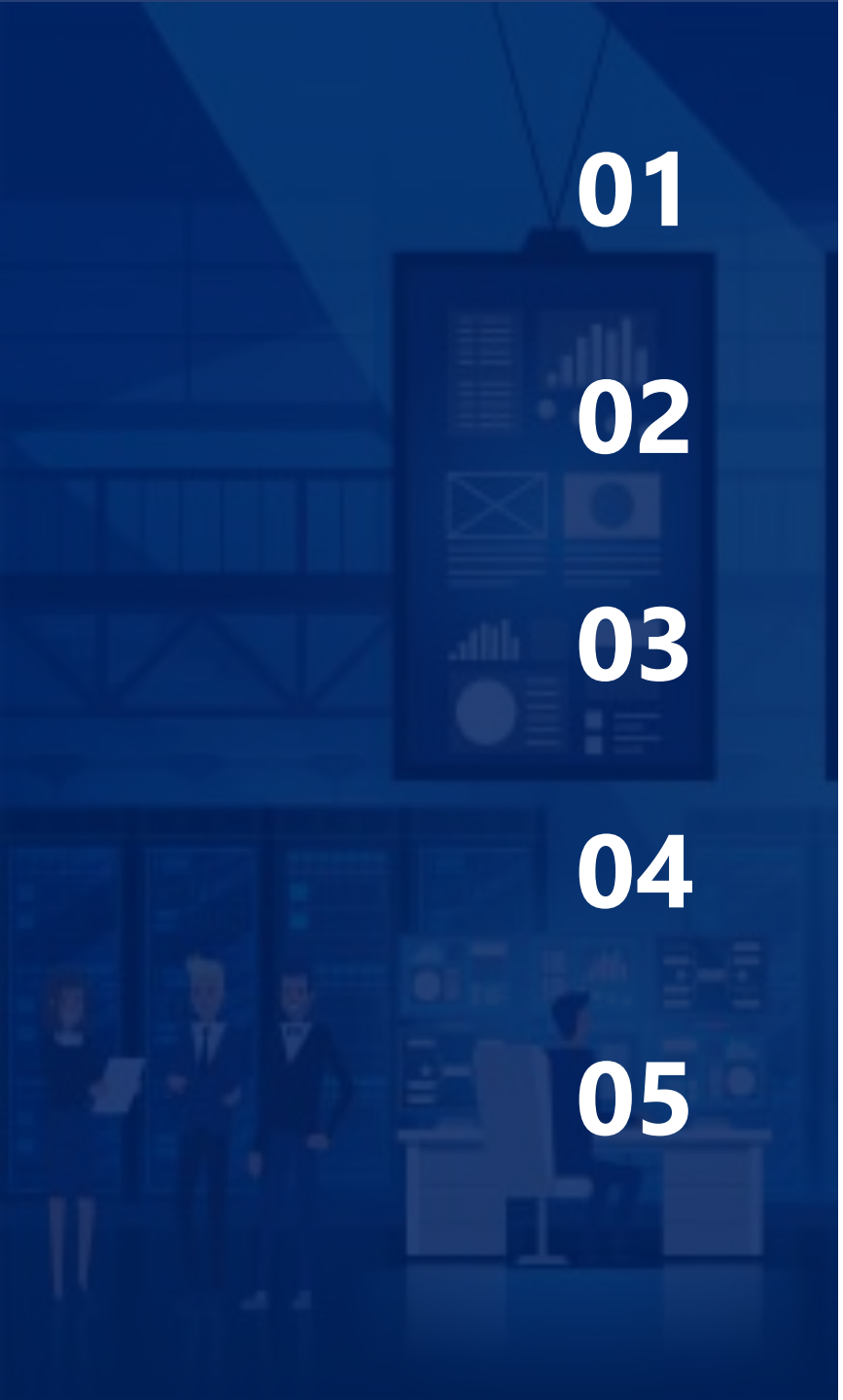


TABLE OF CONTENTS

01
Executive Summary

EXECUTIVE SUMMARY

This presentation outlines the methodology of this Capstone project which includes:

Data Collection & Wrangling

EDA

Interactive Map via Folium

Dashboard via Dash

Classification Predictive Analysis

EXECUTIVE SUMMARY

The results manifested that:

- higher payload will increase the chance of a successful landing
- The highest success rate is seen from the orbit ES-L1, SSO, HEO and GEO
- KSC-LC39A Has the highest Success Rate
- Decision Tree is the best model for classification in this project

02
Introduction

INTRODUCTION

PROJECT BACKGROUND

- This project aims **to predict** if the first stage of the SpaceX Falcon 9 rocket will **land successfully.**

- This will give us **insights on the cost** for launching a Falcon 9 and **supports that Falcon 9 is cheaper to operate** than its competitiors

USD62,000,000

VS.

CompanyZ

USD165,000,000

INTRODUCTION

PROBLEM STATEMENT

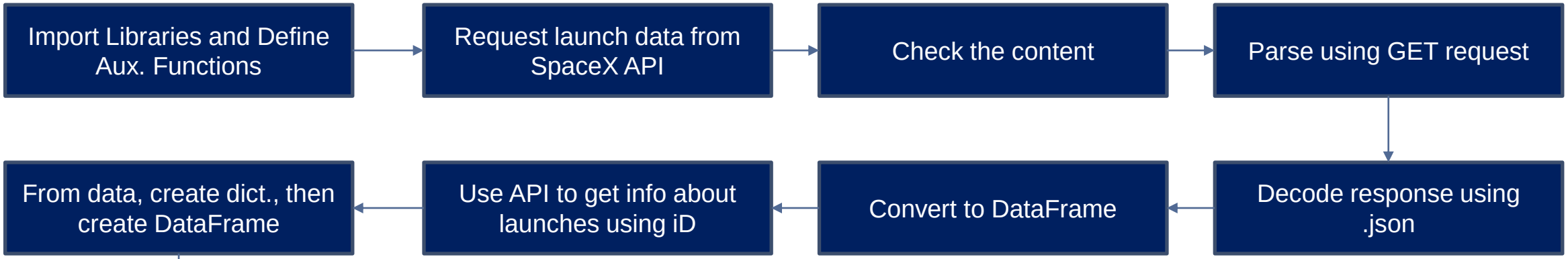
- The **parameters and variables** of the rocket **may affect the success rate** of the landing

- There is a need to **visualize and predict** the best condition **to maximize the rocket** success landing rate

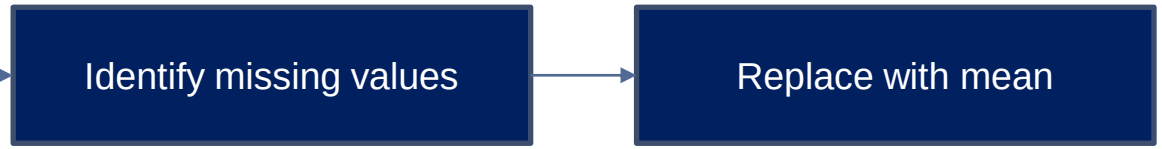
03
Methodology

1. Data Collection & Wrangling

Data Collection (API)

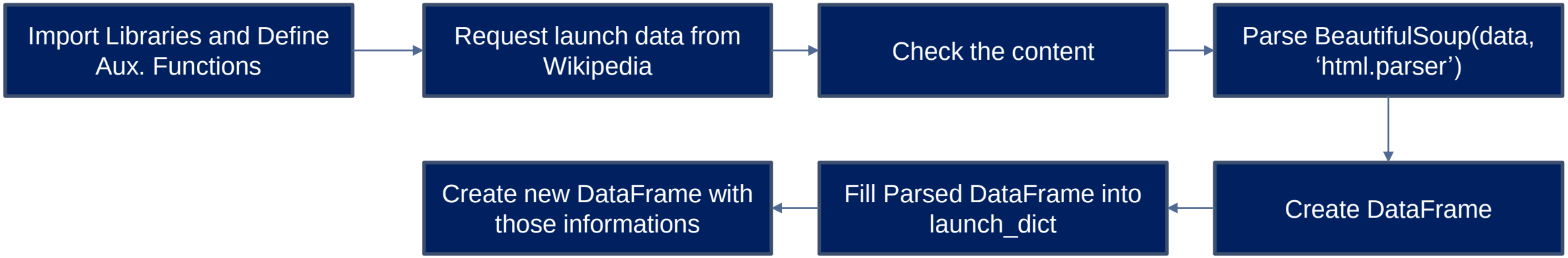


Data Wrangling



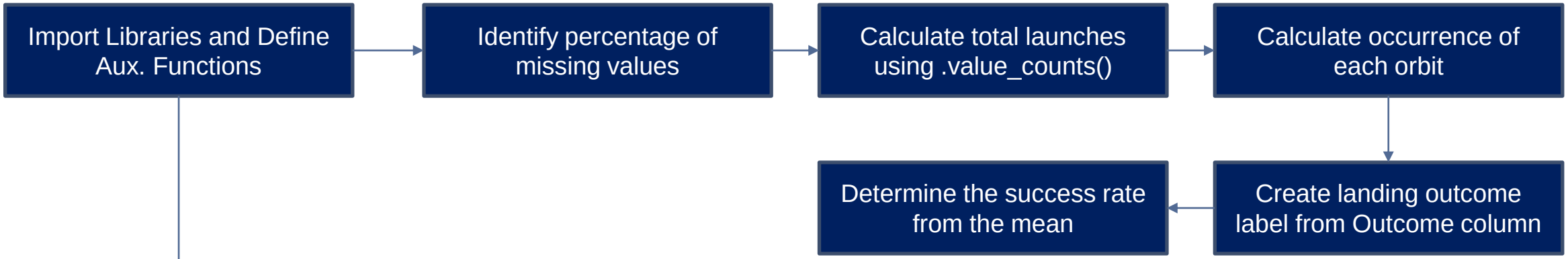
1. Data Collection & Wrangling

Data Collection (Wikipedia)



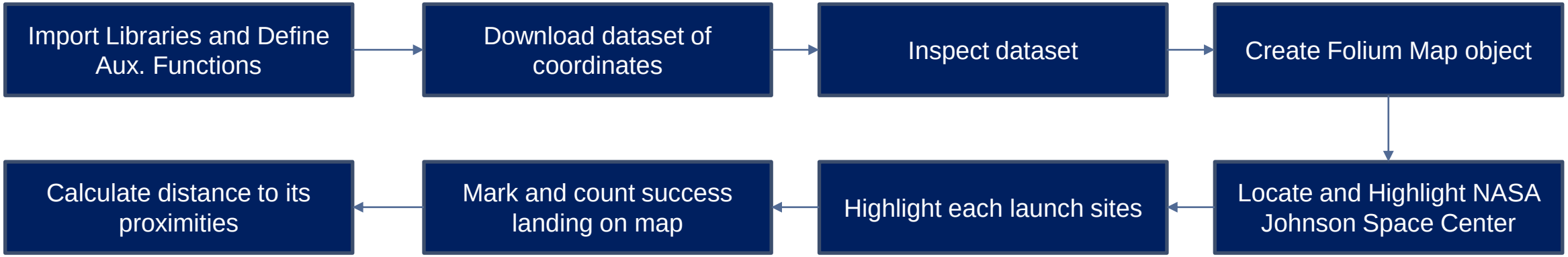
2. Interactive Visual Analytics & EDA

Data Wrangling

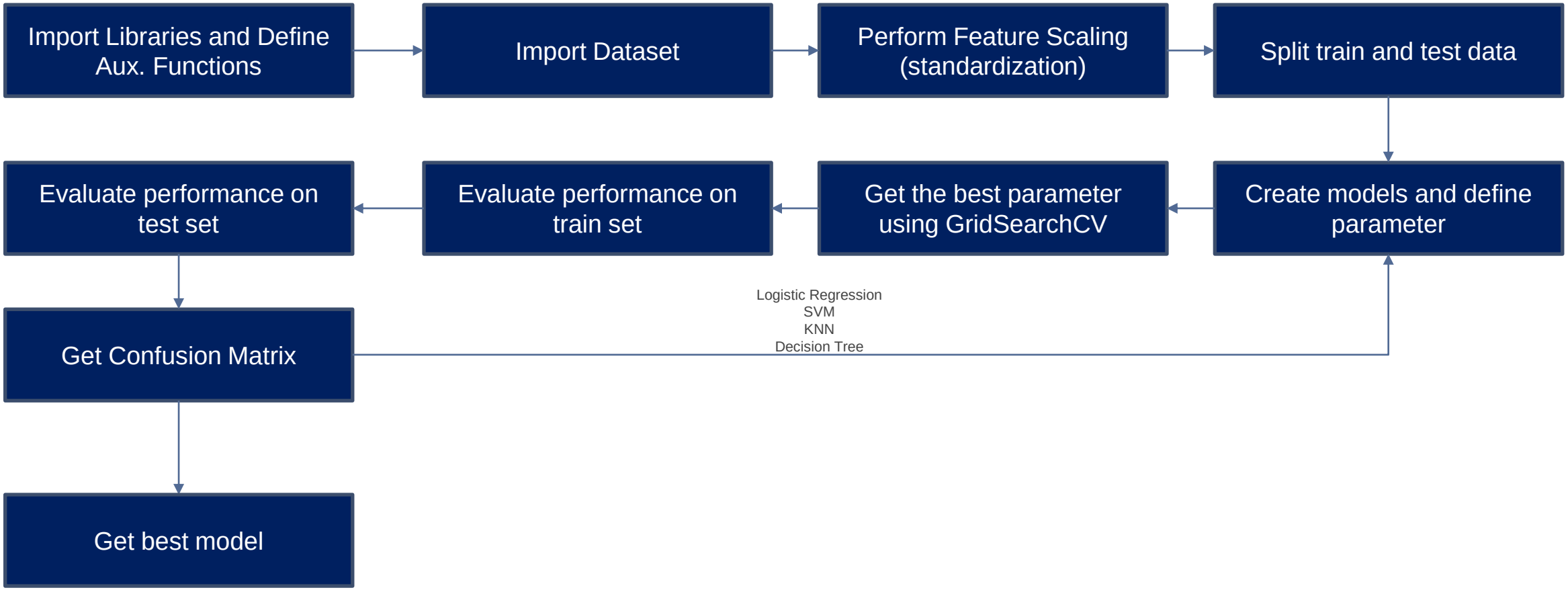


Data Visualization

3.Interactive Visual Analytics

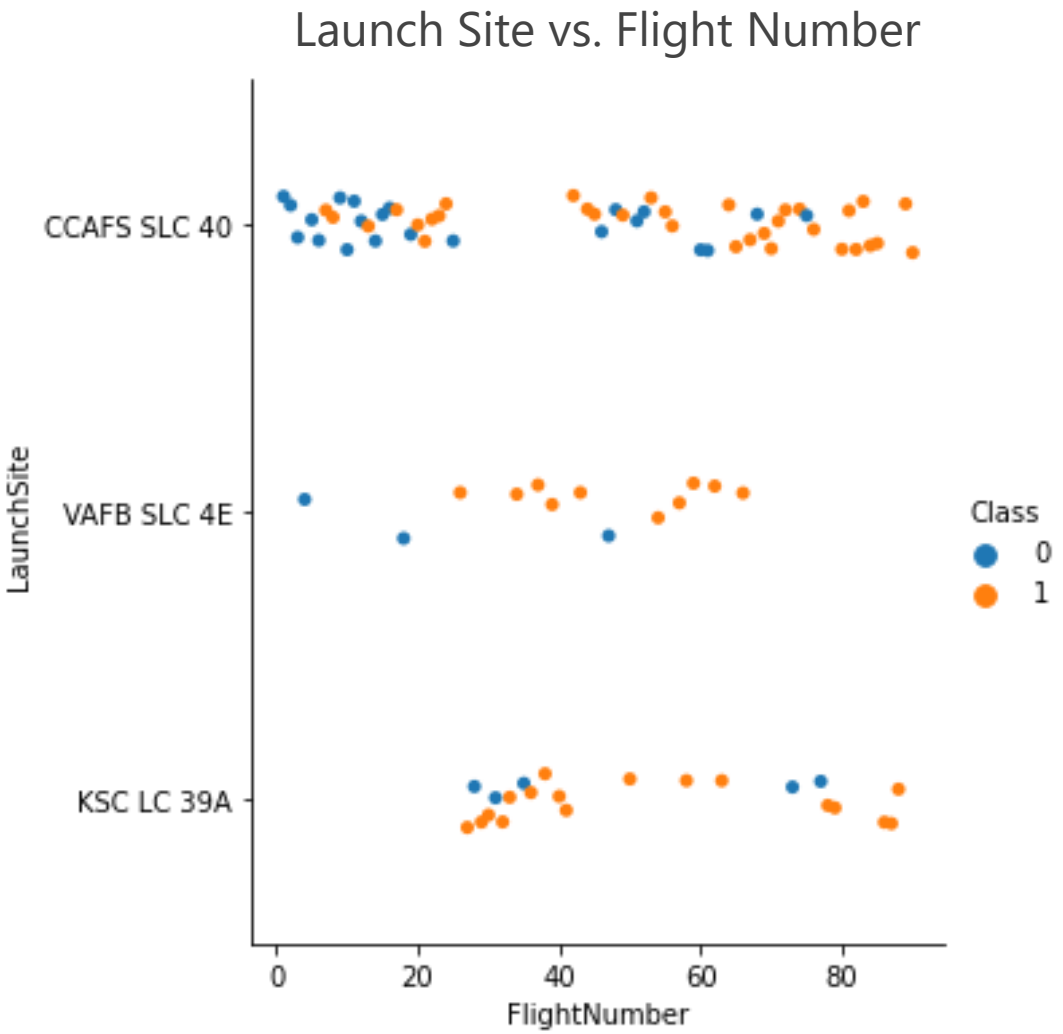


4.Predictive Analysis



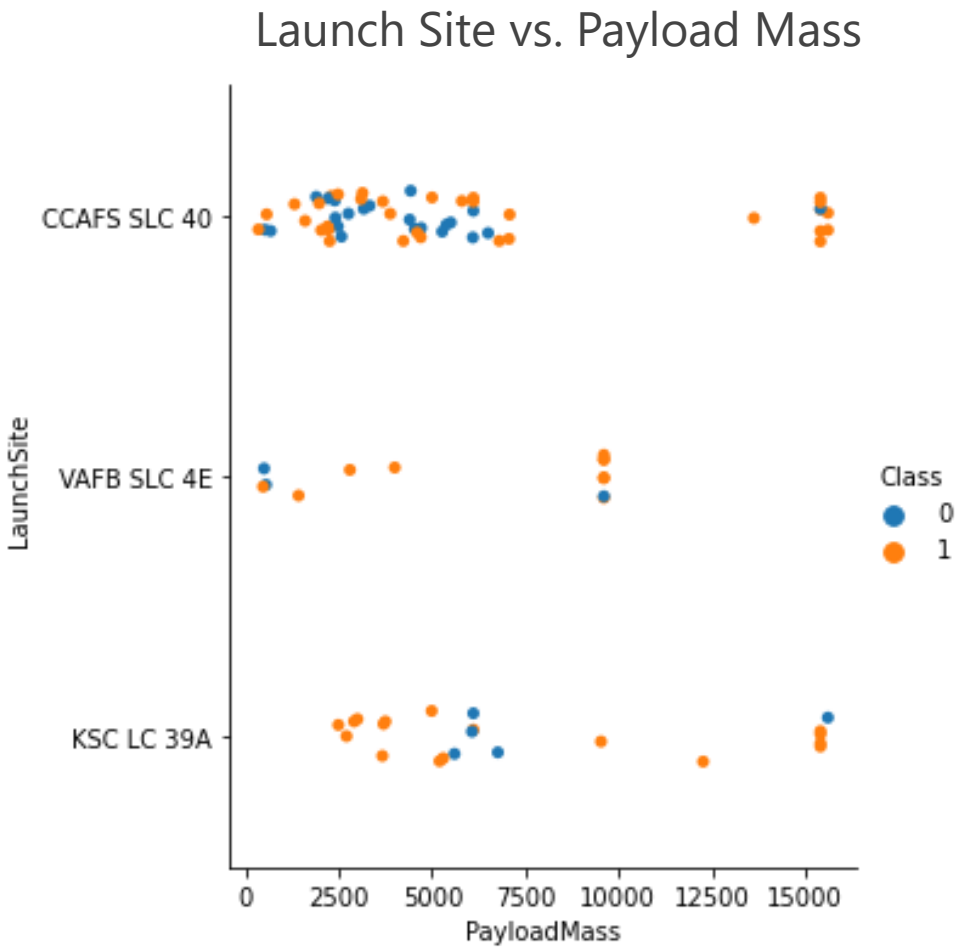
04
Results

EDA Results



It is deduced that from this chart, as **more successful landing** as the **number of flight increased**

EDA Results

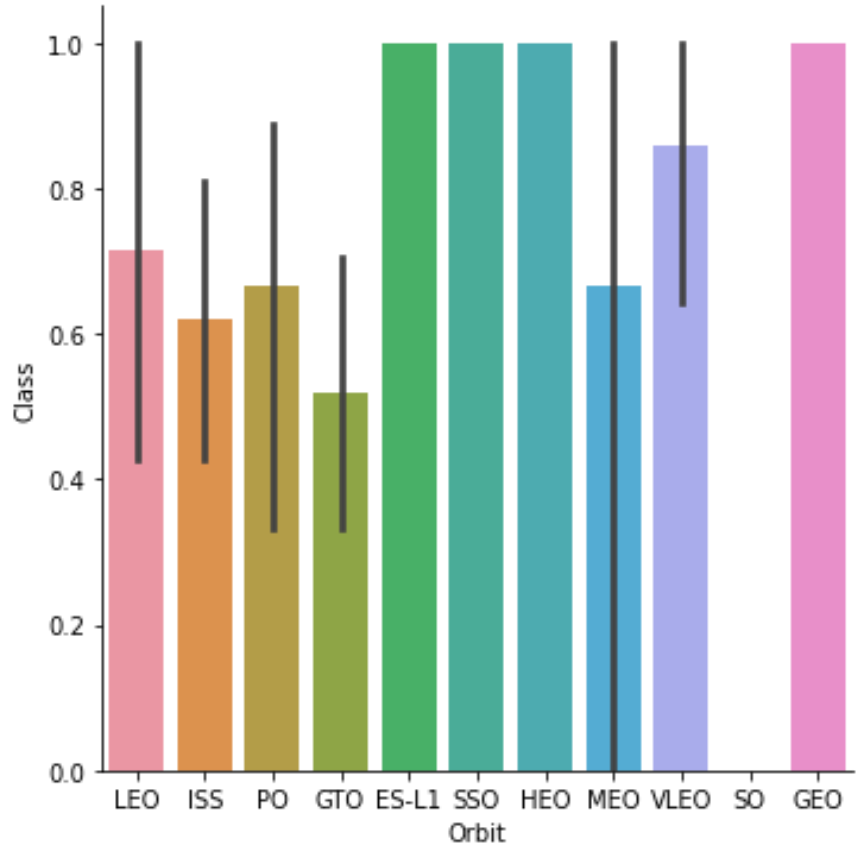


A direct intuition suggested that **higher payload will increase the chance of a successful landing**

However, it is hard to obtain more insights from this graph as the **trend is not visible** enough

EDA Results

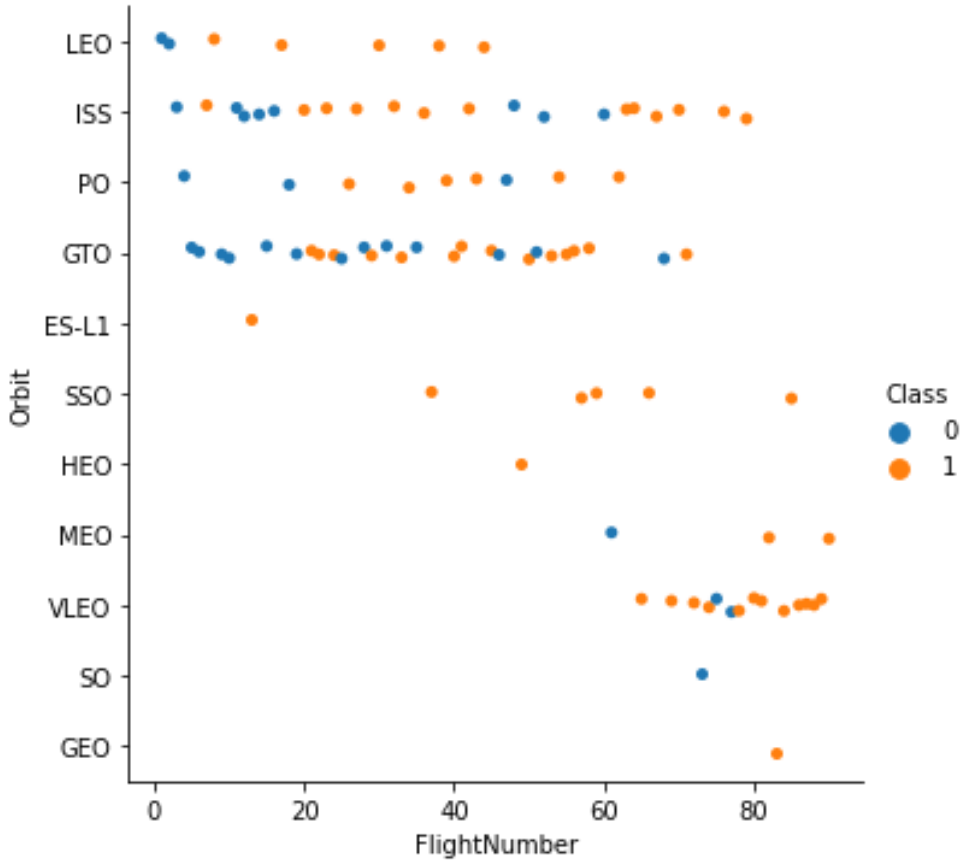
Success Rate vs. Orbit



The highest success rate is seen from the orbit **ES-L1, SSO, HEO and GEO**

EDA Results

Orbit vs. Flight Number

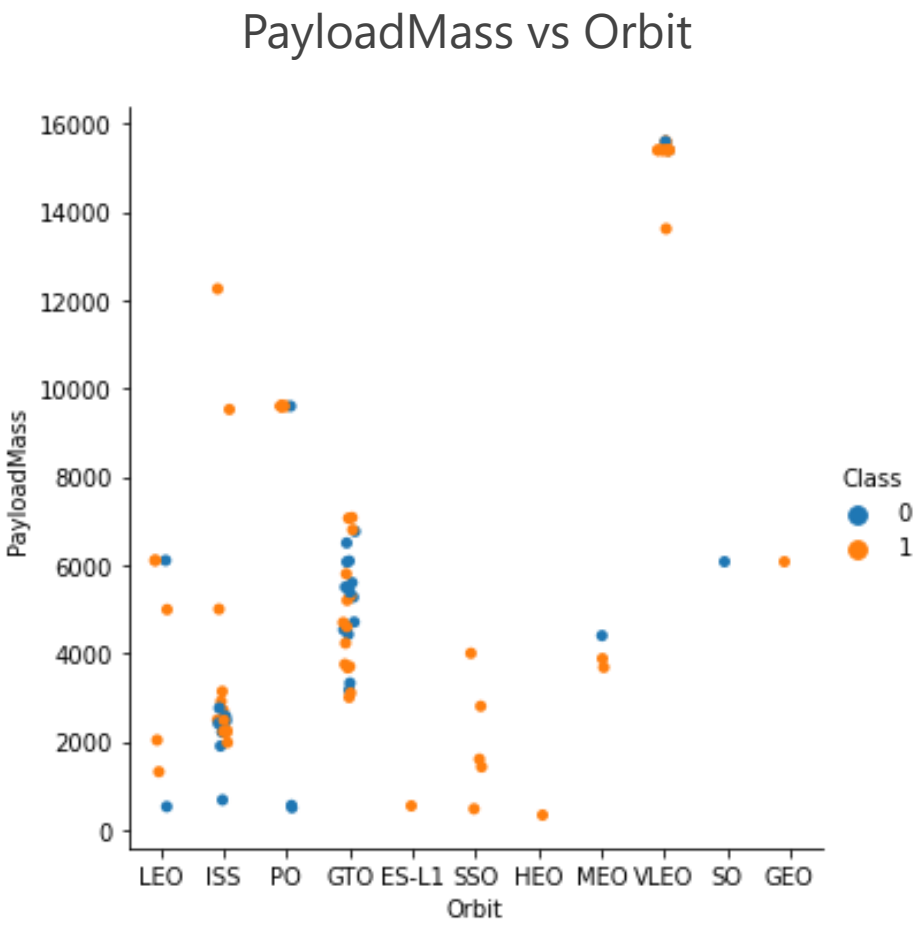


It is observed that the **LEO orbit** shows a trend of **high frequency of successful rate at increasing flight number**

However, the **rest of it shows no visible relationship**

ES-L1,SSO, HEO and GEO has the highest success rate, thus supports the previous statement and bar charts

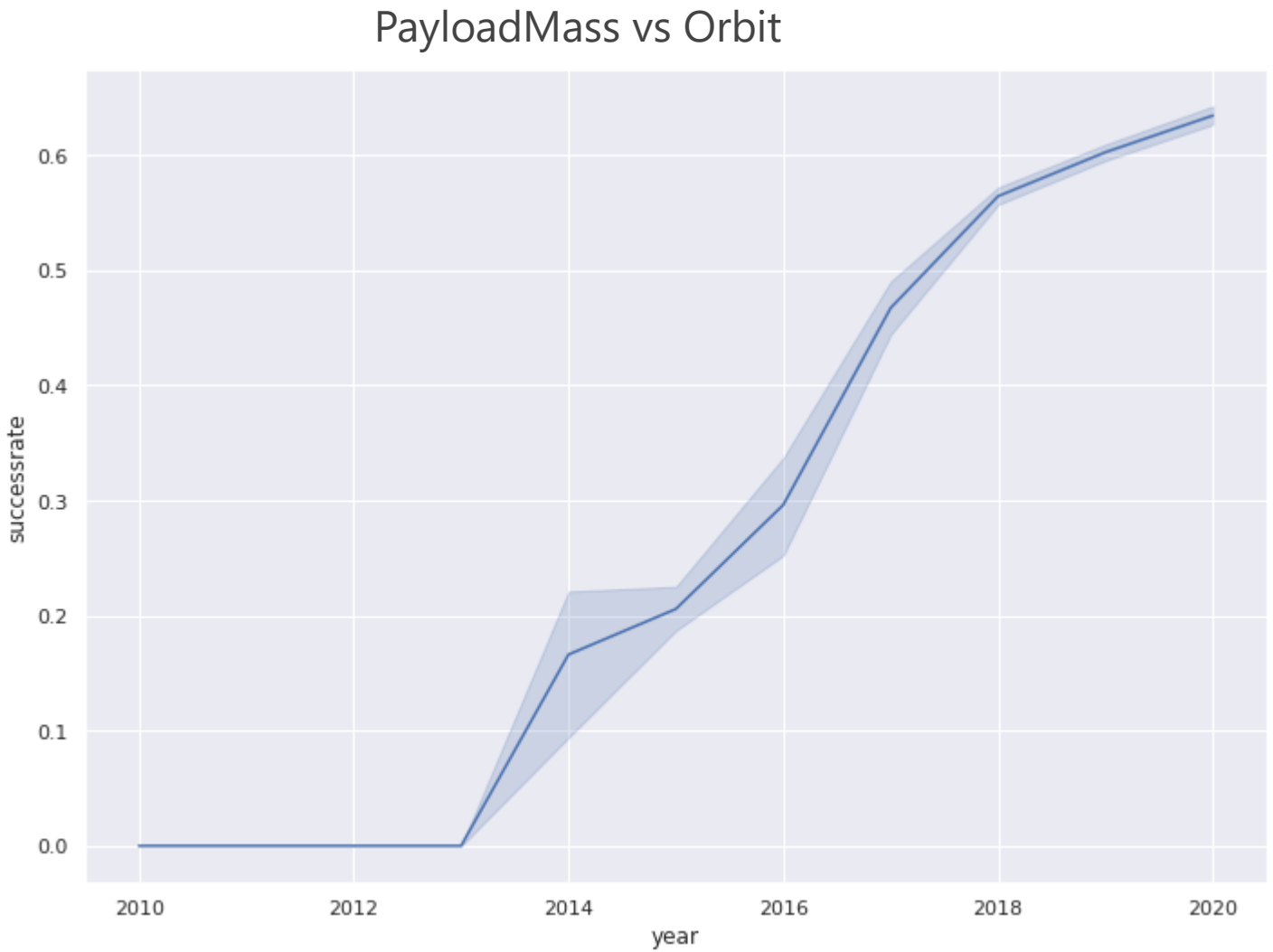
EDA Results



The highest success rate is seen from the orbit **ES-L1, SSO, HEO and GEO**

However, higher payload seems to affect the success rate of **GTO orbit**

EDA Results



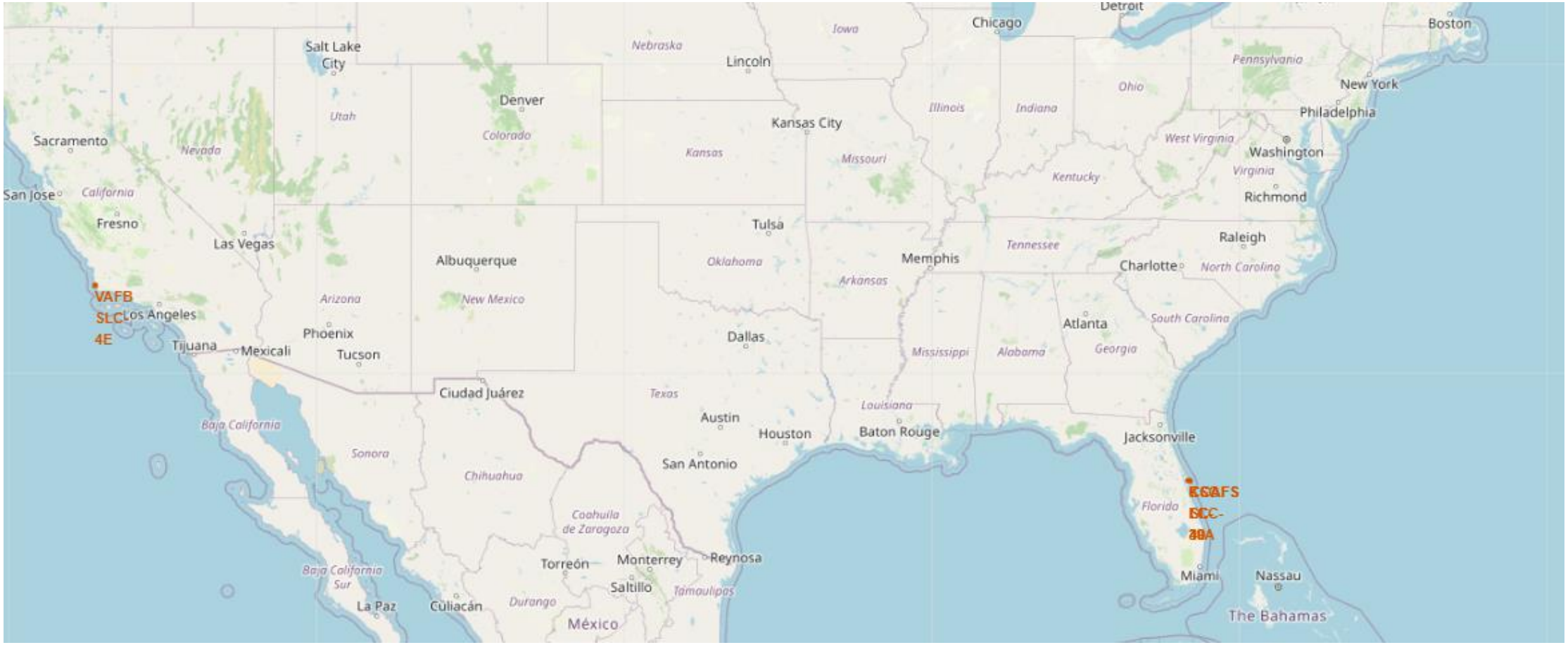
The highest success rate is seen from the orbit **ES-L1, SSO, HEO and GEO**

Interactive Map via Folium

All Launch Sites

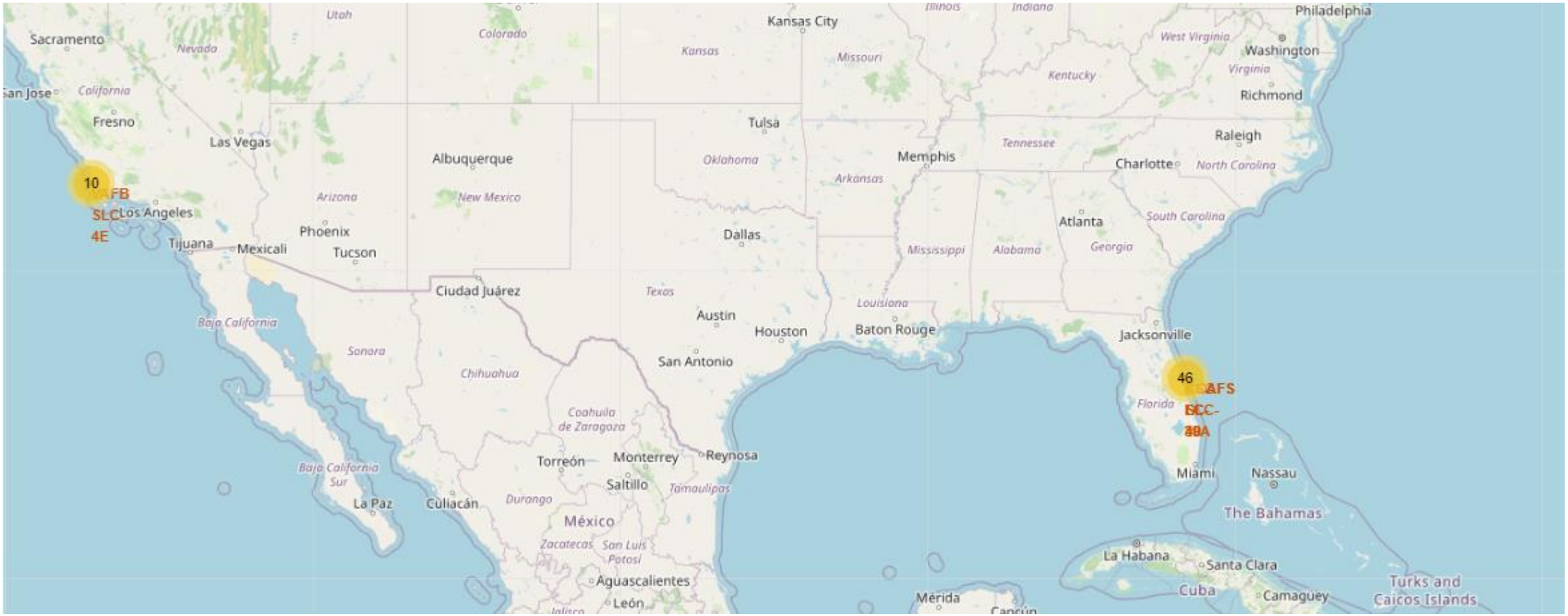
Interactive Map via Folium

All Launch Sites



Interactive Map via Folium

Launch Outcomes

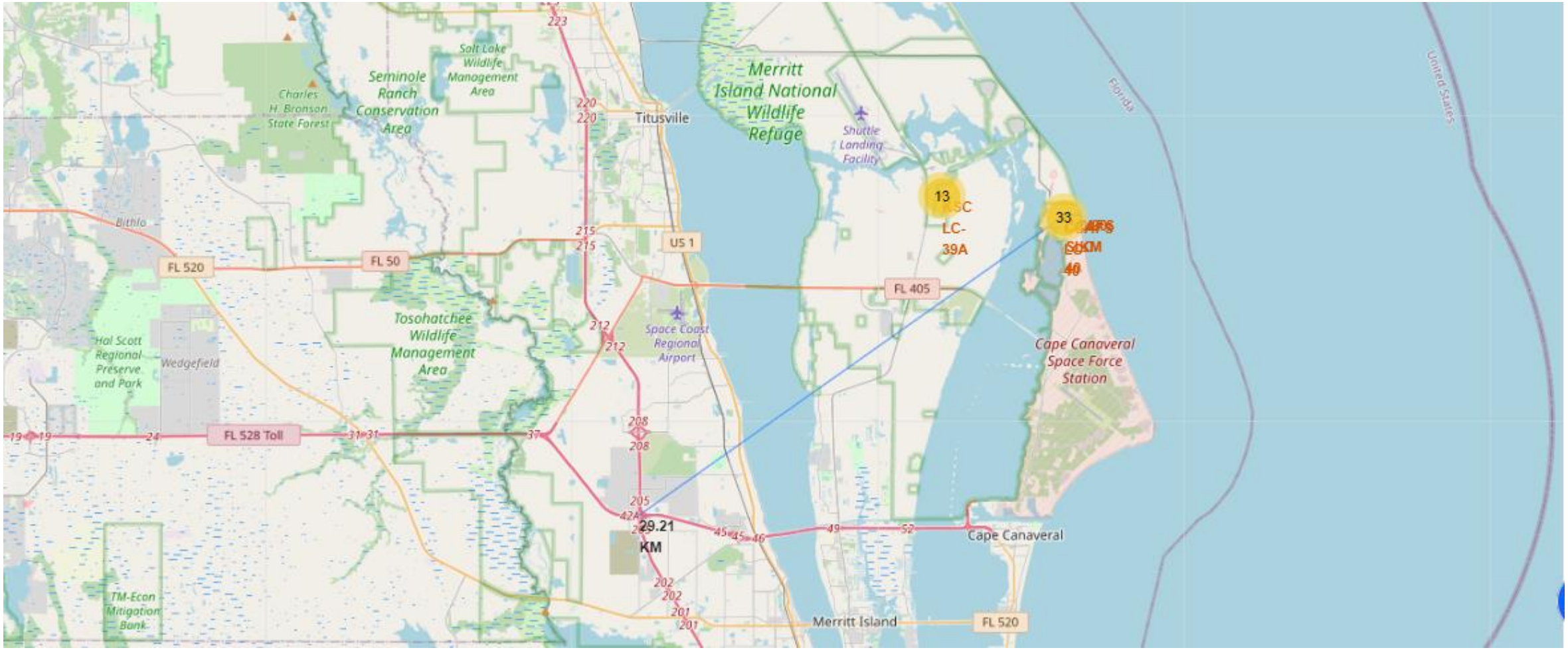


Interactive Map via Folium

Launch Site and its Proximites

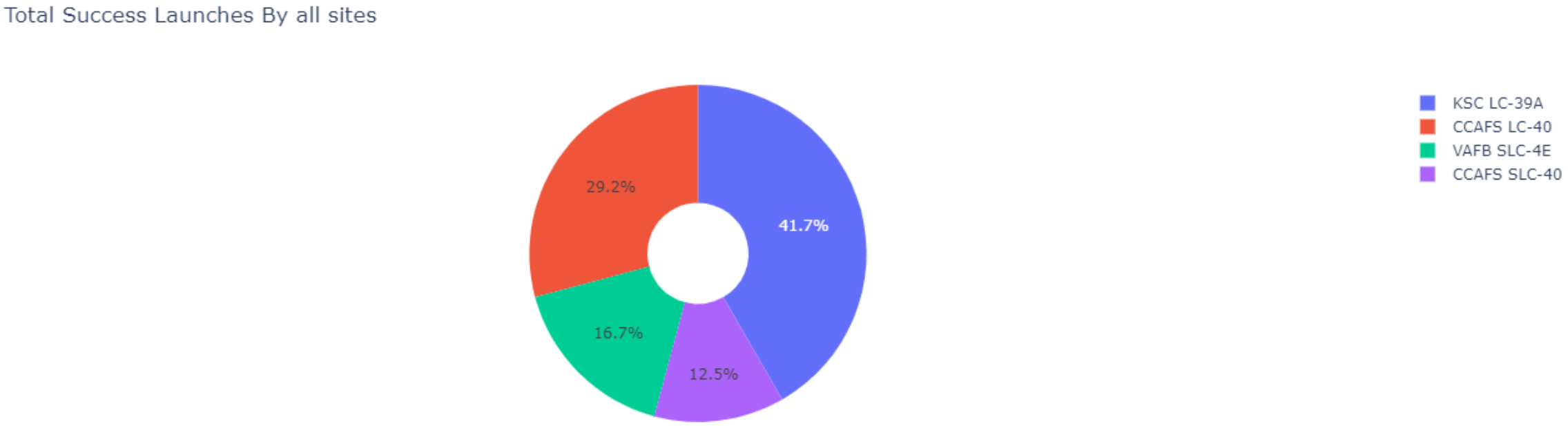
Interactive Map via Folium

Launch Site and its Proximites



Plotly Dash Dashboard

Pie Chart for Launch Site



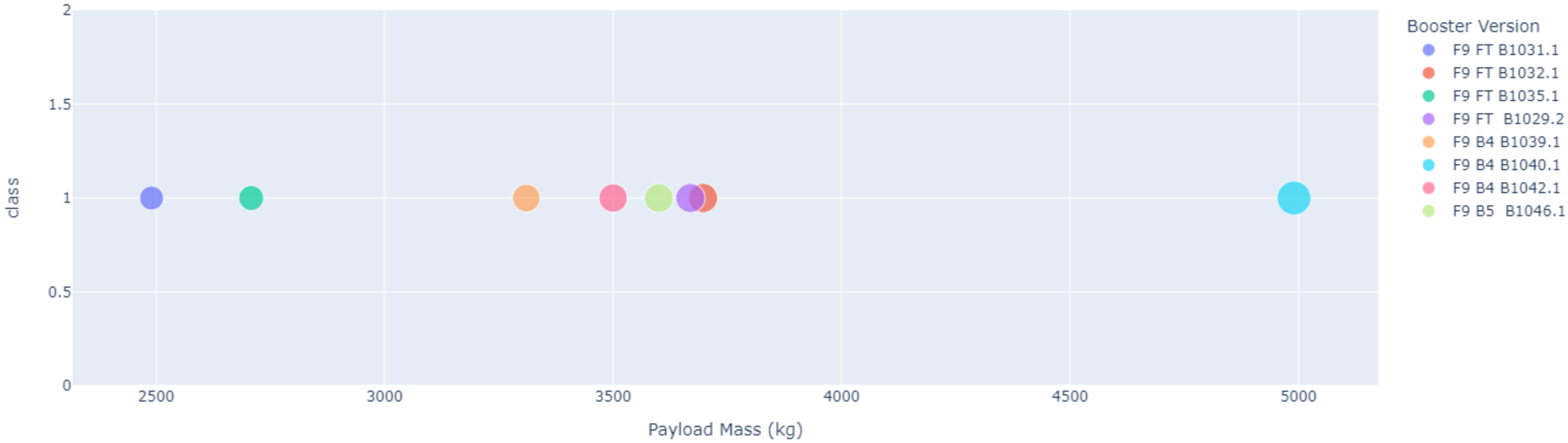
Plotly Dash Dashboard

Highest Launch Success

Total Success Launches for site KSC LC-39A

Plotly Dash Dashboard

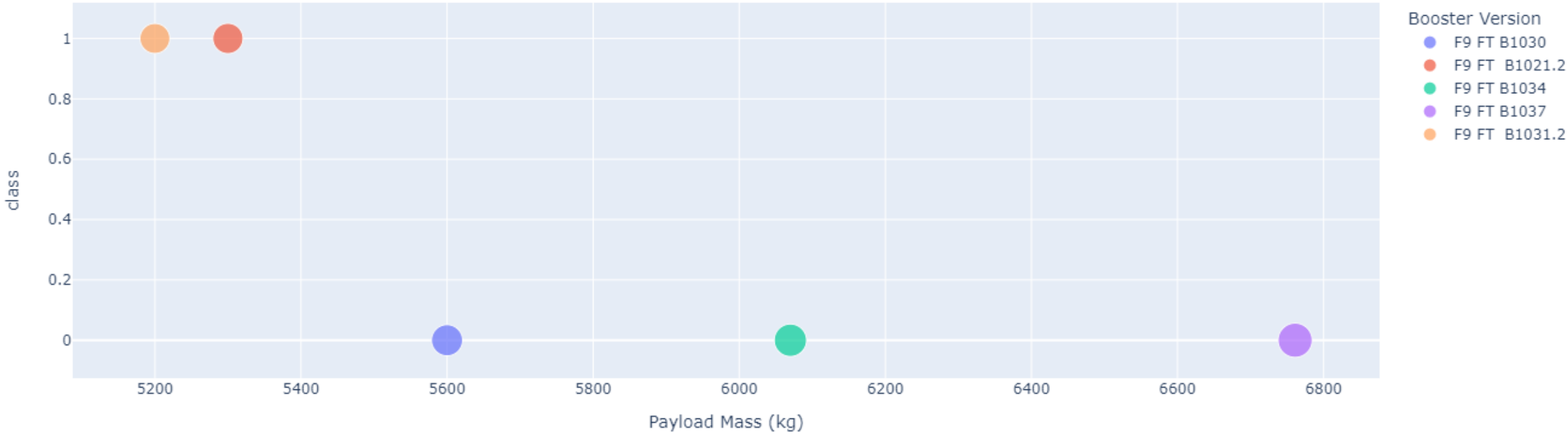
Payload vs. Launch Outcome



From 0 – 4000kg, all has Success Rate

Plotly Dash Dashboard

Payload vs. Launch Outcome



Above 5600kg, The Success Rate is 0

Predictive Analysis

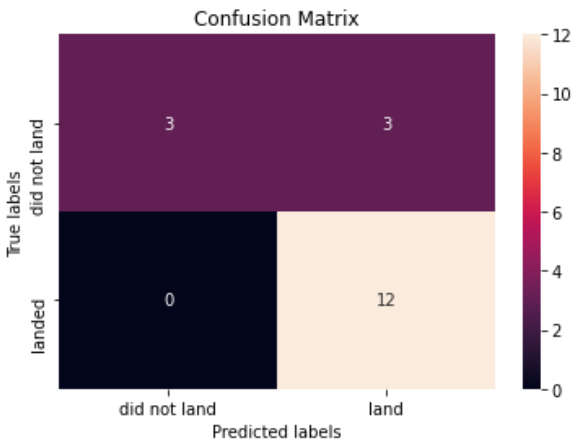
Accuracy (Train and Test)

Logistic Regression	KNN	SVM	Decision Tree
84.642% **83.33%**	**84.82%** **83.34%**	**84.821%** **83.34%**	**90.35%** **77.78%**

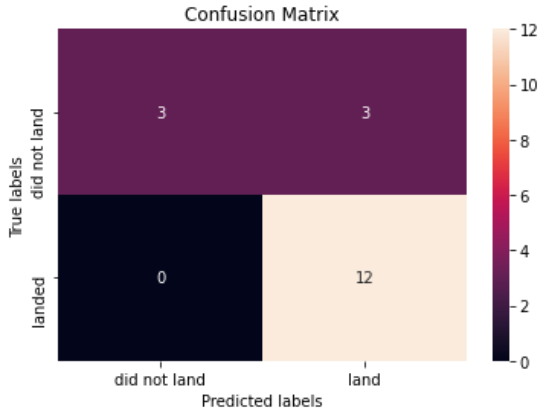
4.Predictive Analysis

Confusion Matrix

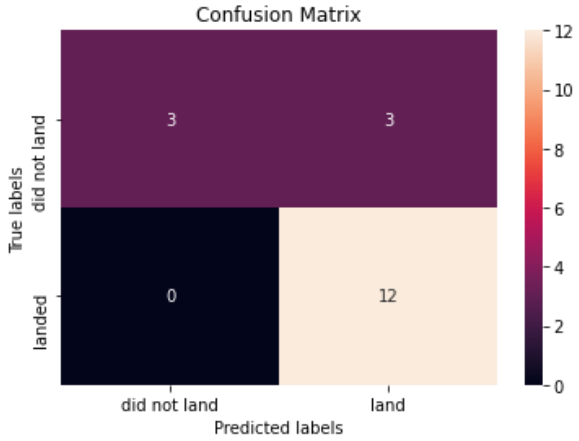
Logistic Regression



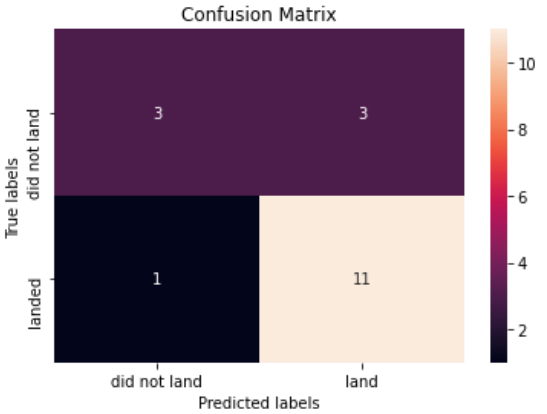
KNN



SVM



Decision Tree



Conclusion

- higher payload will increase the chance of a successful landing
- The highest success rate is seen from the orbit ES-L1, SSO, HEO and GEO
- KSC-LC39A Has the highest Success Rate
- Site in Florida has the highest success rate
- Decision Tree is the best model for classification in this project

GitHub Sources

1. Data Collection API
2. Data Collection with Web Scraping API
3. EDA
4. EDA with SQL
5. EDA with Visualization
6. Dashboard
7. Folium Lab
8. Predictive Analysis