IBM Data Science Capstone Project – **SPACE X**

MADE BY : GARIMA SHARMA

GITHUB REPOSITORY : thecodologist/Applied-Data-Science-Capstone: IBM (github.com)

OUTLINE

EXECUTIVE SUMMARY

- ✓ Data Collection
- ✓ Data Wrangling
- ✓ EDA with Data Visualization
- ✓ EDA with SQL
- ✓ Building a map using Folium
- ✓ Building a dashboard using Plotly Dash
- ✓ Predictive Analysis
- ✓ 4 machine learning models were produced – Logistic Regression, SVM, Decision Tree Classifier and K nearest neighbour

INTRODUCTION

PROJECT BACKGROUND

In this project , we prediction if Falcon 9 first stage will land successfully. SpaceX advertises Falcon 9 rocket launces on it website on its website, with cost of 62 million dollars. In this , we can determine if the first stage will land , we can determine the cost of launch.

Methodology

overview of data collection, data
wrangling, data visualization and
dashboard

DATA COLLECTION METHODOLOGY

Combined data form SpaceX public UPI and SpaceX Wikipedia page

DATA WRANGLING

Classifying correct landings as successful and unsuccessful accordingly.

PERFORMING EDA USING VISUALIZATION AND SQL

PERFORMING PREDICTIVE ANALYSIS USING CLASSIFICATION MODELS

DATA COLLECTION METHODOLOGY

Data collection process involved a collection of API requests for SpaceX public API and webscrapping data from a table in SpaceX Wikipedia entry

Data collection SpaceX API

[Applied-Data-Science-Capstone/Complete the Data Collection API Lab.ipynb at main · thecodologist/Applied-Data-Science-Capstone (github.com)](#)

Request (Space X APIs)

JSON file + Lists(Launch Site, Booster Version, Payload Data)

Json_normalize to DataFrame data from JSON

Dictionary relevant data

Cast dictionary to a DataFrame

Filter data to only include Falcon 9 launches

Replace missing PayloadMass values with mean

Data collection WebScrapping

Applied-Data-Science-Capstone/Data Collection with Web Scraping lab.ipynb at main · thecodologist/Applied-Data-Science-Capstone (github.com)

Request Wikipedia html

BeautifulSoup html5lib Parser

Cast dictionary to DataFrame

Iterate through table cells to extract data to dictionary

Create dictionary

Data Wrangling

[Applied-Data-Science-Capstone/Data Wrangling.ipynb at main · thecodologist/Applied-Data-Science-Capstone (github.com)](github.com)

EDA with data visualization

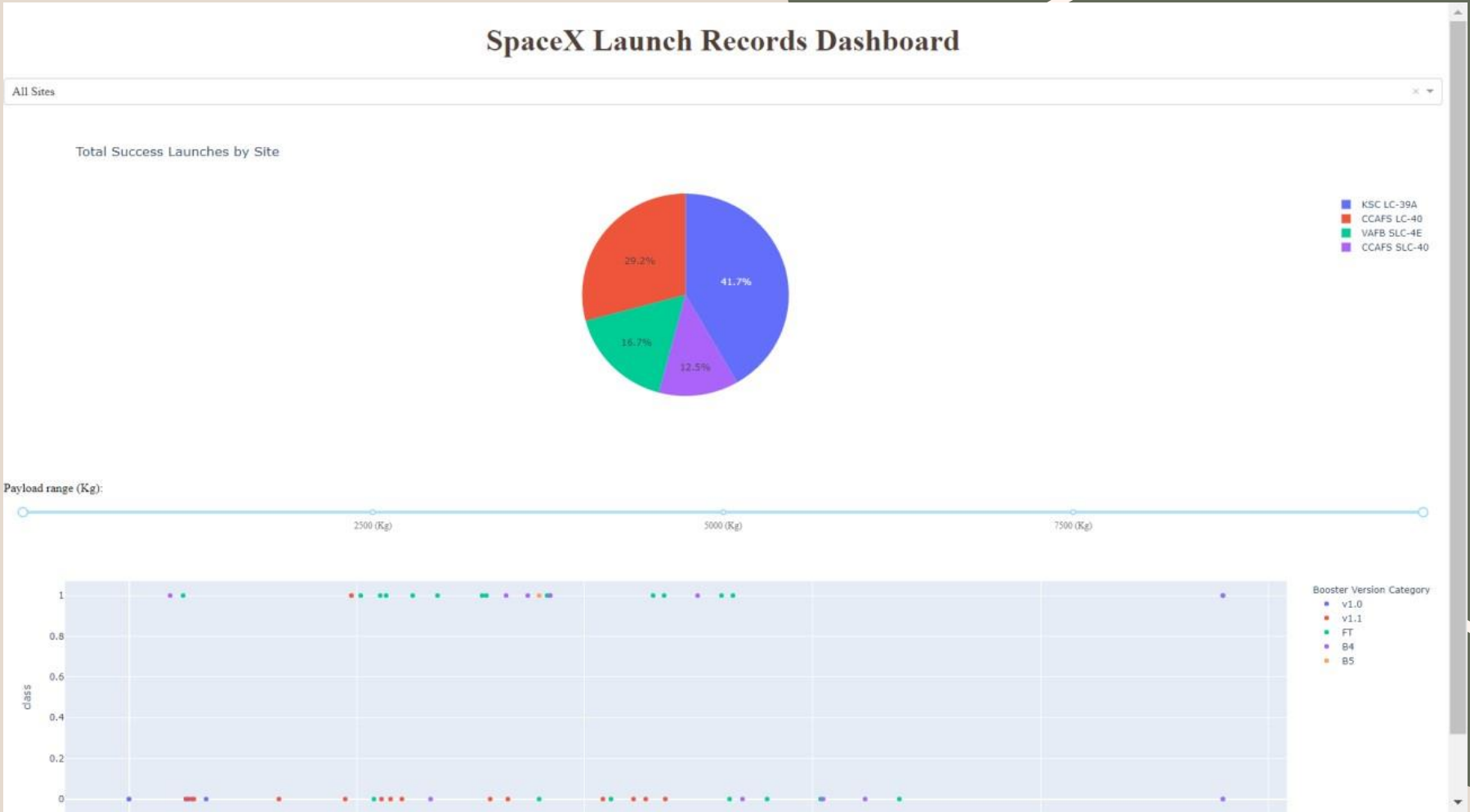[Applied-Data-Science-Capstone/EDA with Visualization lab.ipynb at main · thecodologist/Applied-Data-Science-Capstone (github.com)](https://github.com)

EDA with SQL

Applied-Data-Science-Capstone/EDA with SQL lab.ipynb at main · thecodologist/Applied-Data-Science-Capstone (github.com)

Launch site location analysis with Folium

[Applied-Data-Science-Capstone/Interactive Visual Analytics with Folium lab.ipynb at main · thecodologist/Applied-Data-Science-Capstone (github.com)](#)

Dash App

[Applied-Data-Science-Capstone/spacex_dash_app.py at main · thecodologist/Applied-Data-Science-Capstone (github.com)](https://github.com)

RESULTS

EDA WITH VISUALIZATION

Flight number VS Launch site
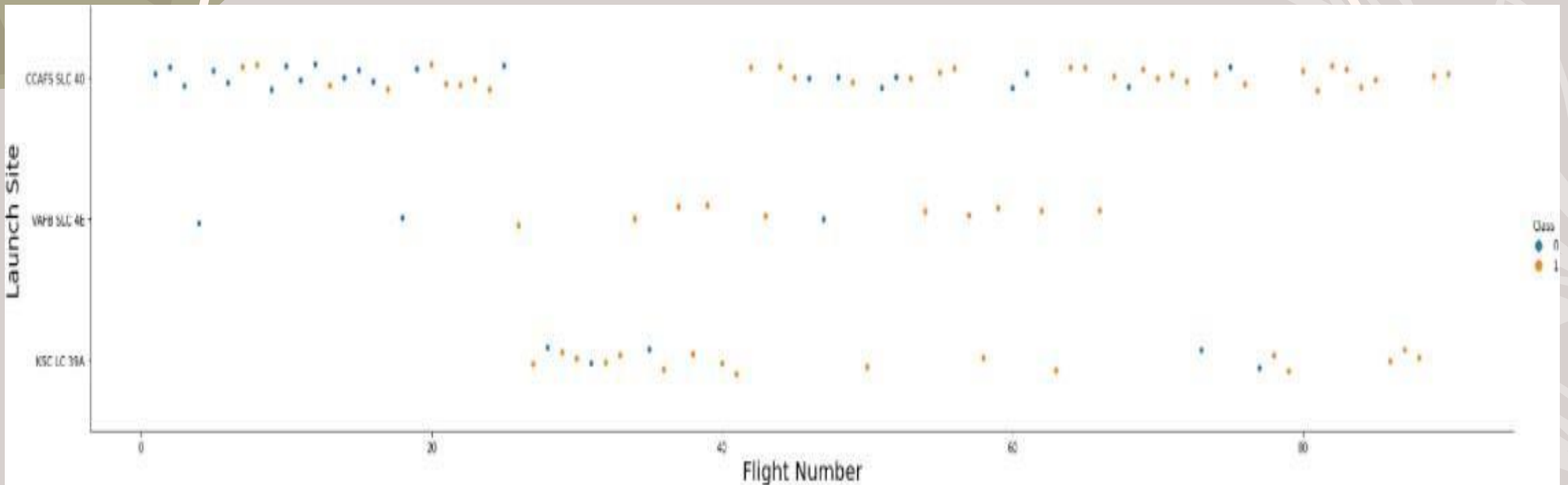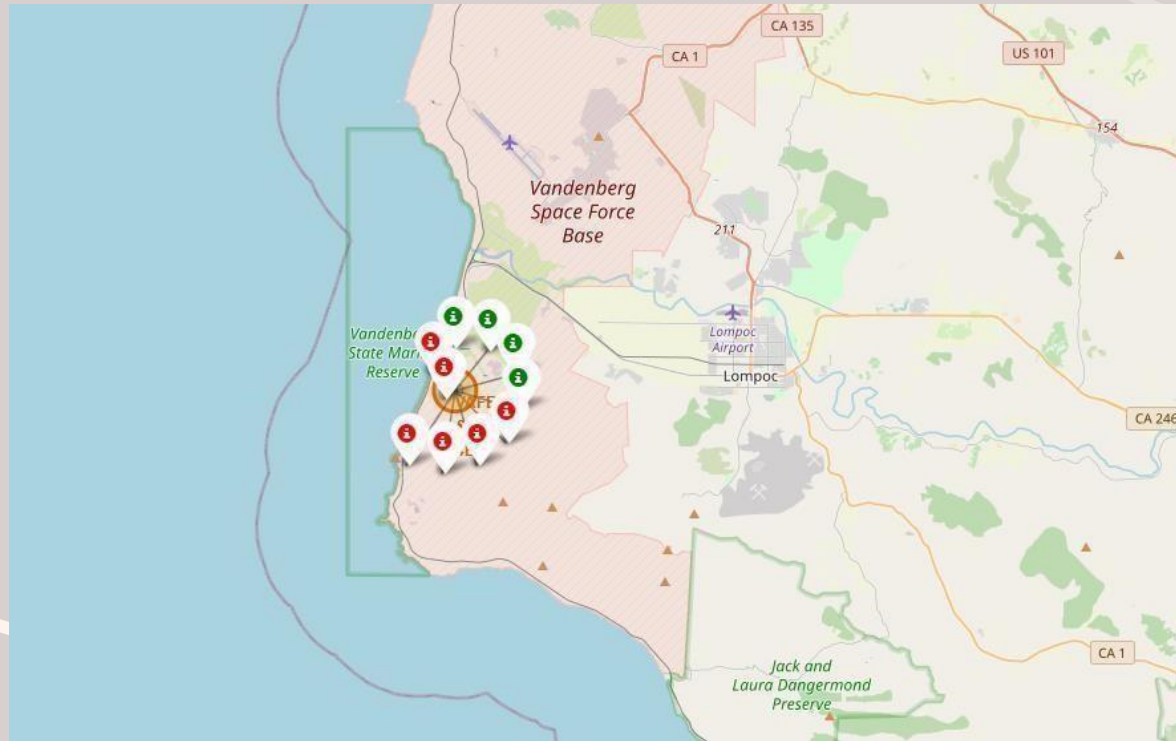
Launch success yearly trend
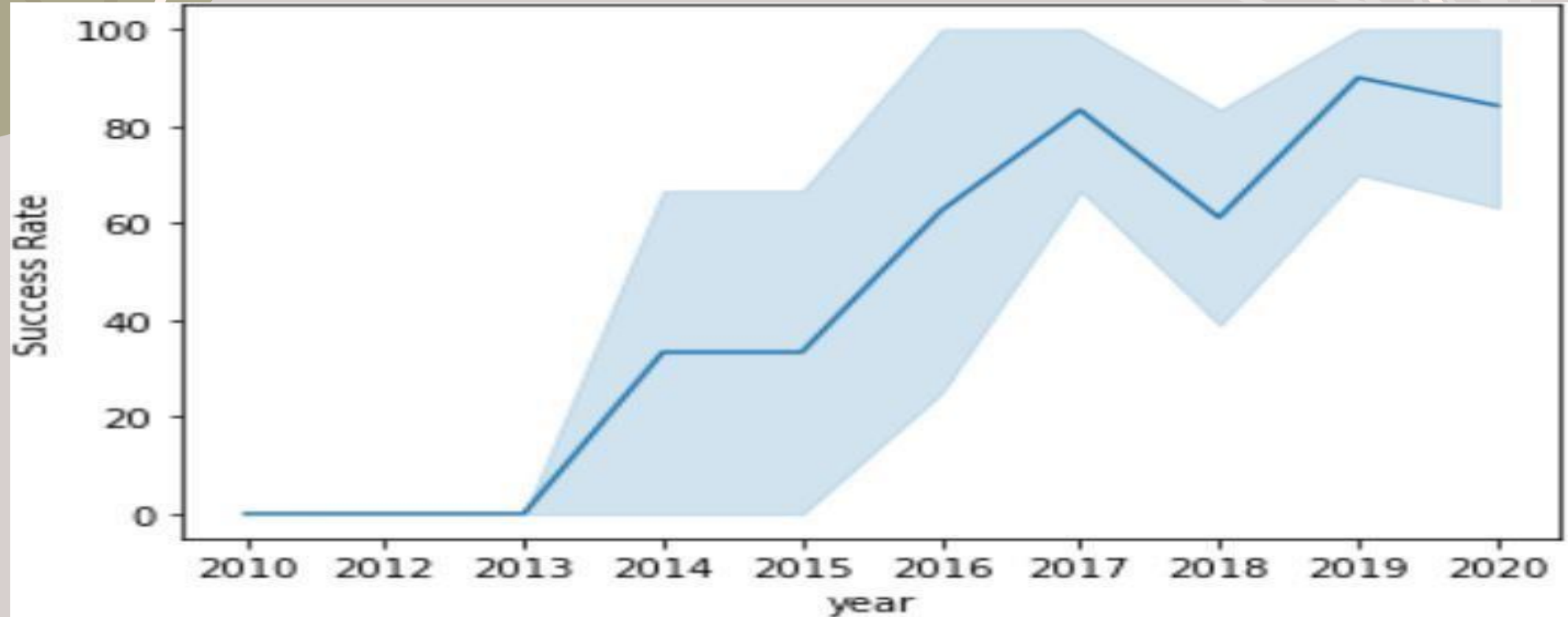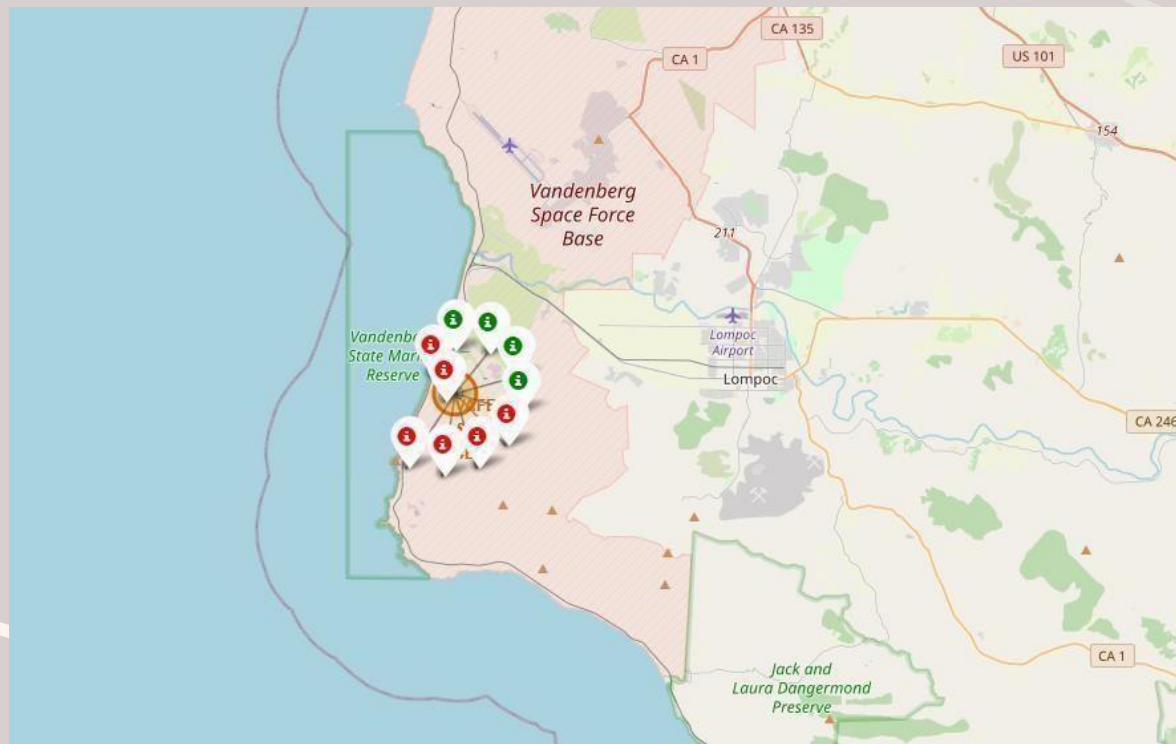
INTERACTIVE MAP WITH FOLIUM

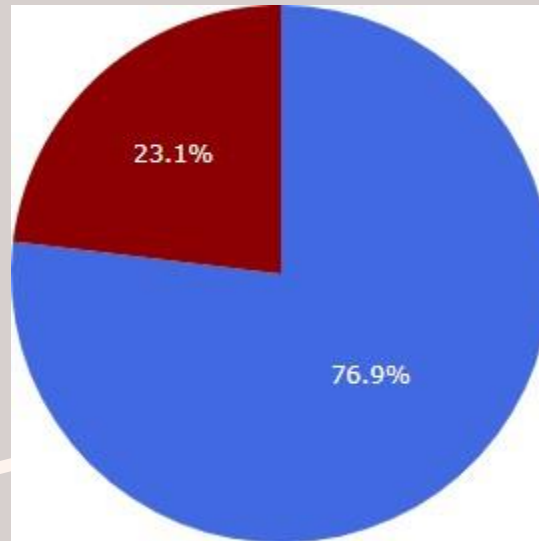## Launch site location
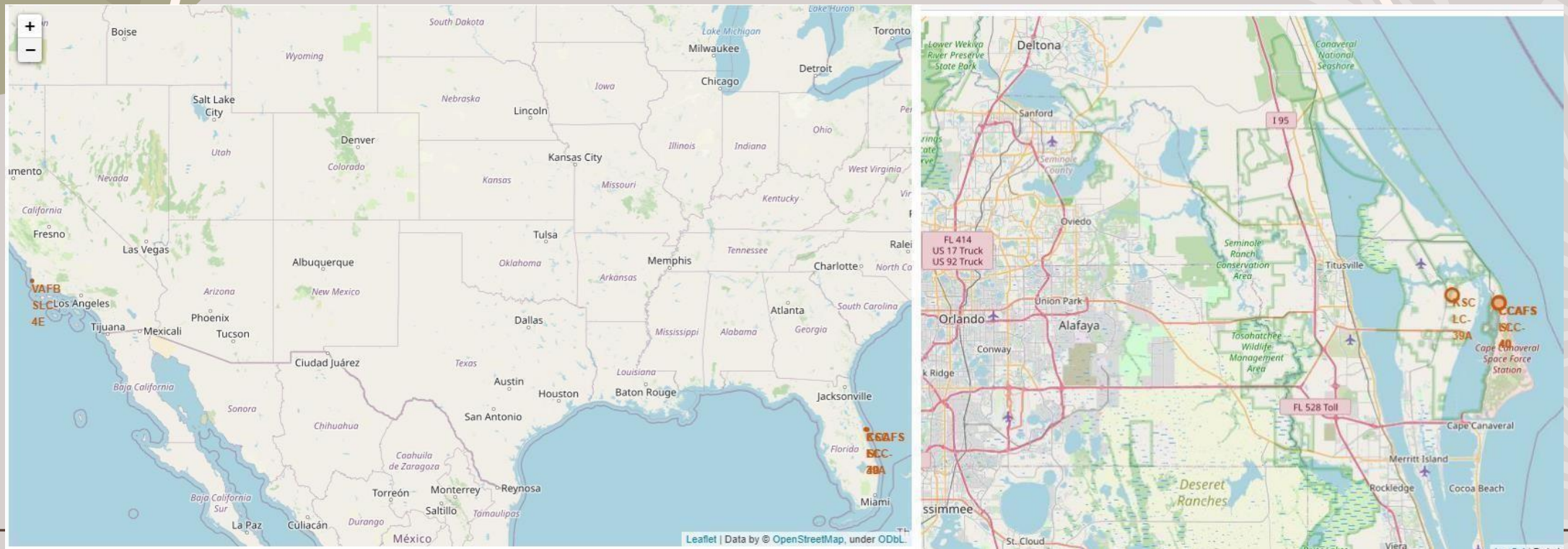
Colour coded launch markers

Colour coded launch markers

BUILDING DASHBOARD WITH PLOTLYDASH

High success rate Launch site

KSC LC-39A Success Rate (blue=success)

CONCLUSION

We created a machine learning model with 80% of accuracy. We also created a dashboard for visualization. We used data from public SpaceX API and webscrapping SpaceX Wikipedia page.

APPENDIX

THANK YOU