IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Brunel Hounza
<06/12/24>

Outline

- Executive Summary
- Introduction
- Methodology
- Results
- Conclusion
- Appendix

Executive Summary

- Summary of methodologies
- Summary of all results

Introduction

Section 1

Methodology

Methodology

Executive Summary

- Data collection methodology:
- Perform data wrangling
- Perform exploratory data analysis (EDA) using visualization and SQL
- Perform interactive visual analytics using Folium and Plotly Dash
- Perform predictive analysis using classification models

Data Collection

- The data was collected from the SpaceX API as well as its website in JSON, CSV formats, etc

Calculate below the mean for the PayloadMass using the .mean(). Then use the mean and the .replace() function to replace np.nan values in the data with the mean you calculated.

Data Collection – SpaceX API

Calculate below the mean for the PayloadMass using the .mean(). Then use the mean and the .replace() function to replace np.nan values in the data with the mean you calculated.

Filter the dataframe to only include Falcon 9 launches

Data Collection - Scraping

Filter the dataframe to only include Falcon 9 launches

Extract all column/variable names from the HTML table header

Create a data frame by parsing the launch HTML tables

Data Wrangling

- Calculate the number of launches on each site
- Calculate the number and occurrence of each orbital
- Calculate the number and occurence of mission outcome of the orbits

EDA with Data Visualization

- Visualize the relationship between Flight Number and Launch Site
- Visualize the relationship between Payload and Launch Site
- Visualize the relationship between Payload and Launch Site
- Visualize the relationship between FlightNumber and Orbit type
- Add the GitHub URL of your completed EDA with data visualization notebook, as an external reference and peer-review purpose

EDA with SQL

- List the date when the first succesful landing outcome in ground pad was achieve
- List the names of the boosters which have success in drone ship and have payload mass greater than 4000 but less than 6000
- List the total number of successful and failure mission outcomes
- List the records which will display the month names, failure landing_outcomes in drone ship ,booster versions, launch_site for the months in year 2015.
- Add the GitHub URL of your completed EDA with SQL notebook, as an exteral reference and peer-review purpose

Build an Interactive Map with Folium

- try to understand the change in average estimated fire area over time
- find if there is any correlation between mean estimated fire radiative power and mean confidence level
- See the percentage on the pie is not looking so good as it is overlaped for Region SA, TA, V
- find the portion of count of pixels for presumed vegetation fires vary across regions we will develop a pie chart for this

- Add the GitHub URL of your completed interactive map with Folium map, as an external reference and peer-review purpose

Build a Dashboard with Plotly Dash

- Add a callback function to render success-pie-chart based on selected site dropdown
- Add a callback function to render the success-payload-scatter-chart scatter plot
- Add the GitHub URL of your completed Plotly Dash lab, as an external reference and peer-review purpose

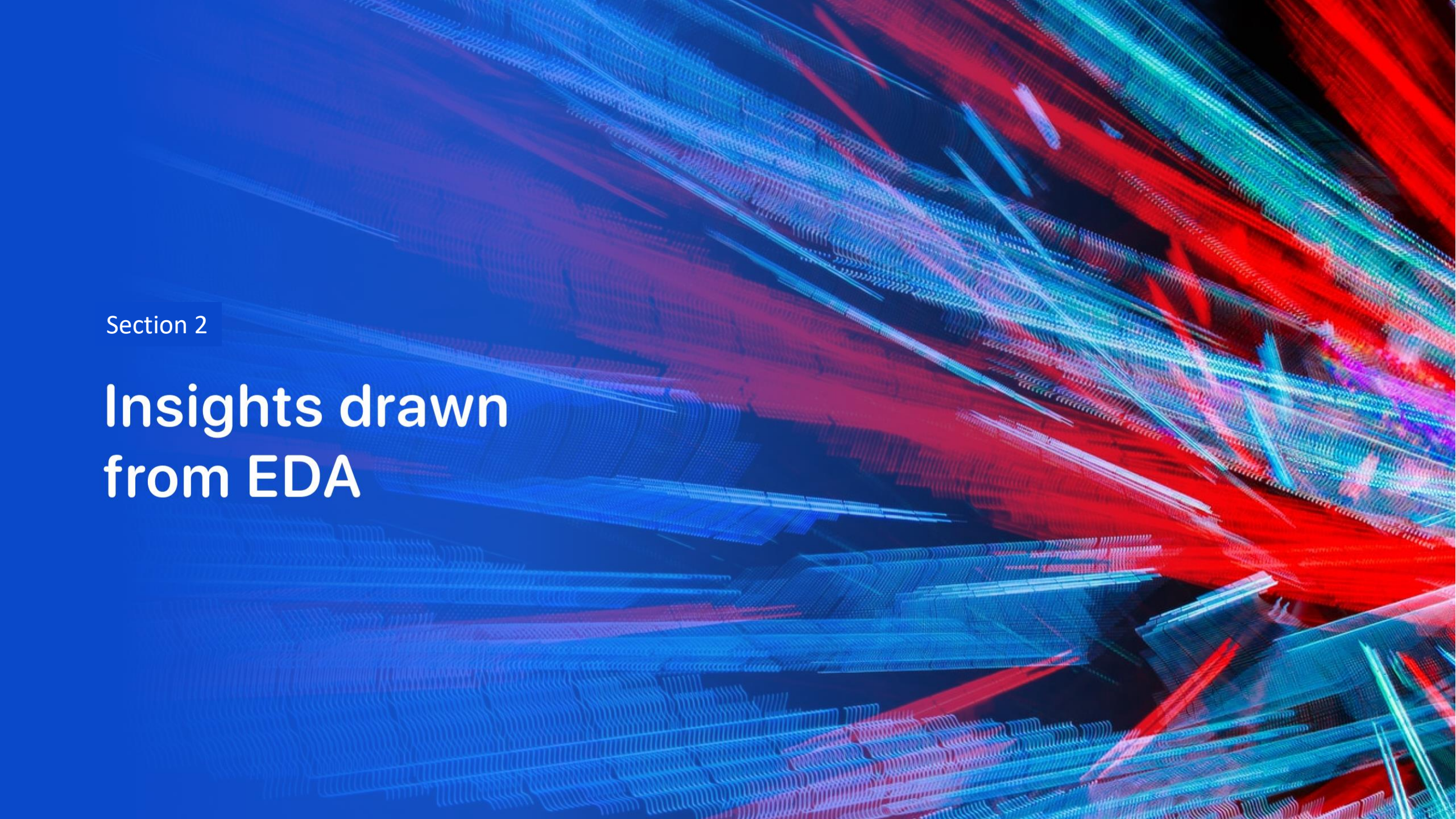
Predictive Analysis (Classification)

- Standardize the data in X then reassign it to the variable X using the transform provided below.
- Create a logistic regression,KNN,Tree decision object then create a GridSearchCV object
- WE use plot the confusion matrix
- Add the GitHub URL of your completed predictive analysis lab, as an external reference and peer-review purpose

Calculate the accuracy of knn_cv on the test data using the method score:

Results

- Calculate the accuracy of knn_cv on the test data using the method score:
- Find the method performs best


Section 2
Insights drawn from EDA

Flight Number vs. Launch Site

- Show a scatter plot of Flight Number vs. Launch Site

- Show the screenshot of the scatter plot with explanations

Payload vs. Launch Site

- Show a scatter plot of Payload vs. Launch Site

- Show the screenshot of the scatter plot with explanations

Success Rate vs. Orbit Type

- Show a bar chart for the success rate of each orbit type

- Show the screenshot of the scatter plot with explanations

Flight Number vs. Orbit Type

- Show a scatter point of Flight number vs. Orbit type

- Show the screenshot of the scatter plot with explanations

Payload vs. Orbit Type

- Show a scatter point of payload vs. orbit type

- Show the screenshot of the scatter plot with explanations

Launch Success Yearly Trend

- Show a line chart of yearly average success rate

- Show the screenshot of the scatter plot with explanations

All Launch Site Names

- Find the names of the unique launch sites
- Present your query result with a short explanation here

Launch Site Names Begin with 'CCA'

- Find 5 records where launch sites begin with `CCA`
- Present your query result with a short explanation here

Total Payload Mass

- Calculate the total payload carried by boosters from NASA
- Present your query result with a short explanation here

Average Payload Mass by F9 v1.1

- Calculate the average payload mass carried by booster version F9 v1.1
- Present your query result with a short explanation here

First Successful Ground Landing Date

- Find the dates of the first successful landing outcome on ground pad
- Present your query result with a short explanation here

Successful Drone Ship Landing with Payload between 4000 and 6000

- List the names of boosters which have successfully landed on drone ship and had payload mass greater than 4000 but less than 6000

- Present your query result with a short explanation here

Total Number of Successful and Failure Mission Outcomes

- Calculate the total number of successful and failure mission outcomes
- Present your query result with a short explanation here

Boosters Carried Maximum Payload

- List the names of the booster which have carried the maximum payload mass
- Present your query result with a short explanation here

2015 Launch Records

- List the failed landing_outcomes in drone ship, their booster versions, and launch site names for in year 2015

- Present your query result with a short explanation here

Rank Landing Outcomes Between 2010-06-04 and 2017-03-20

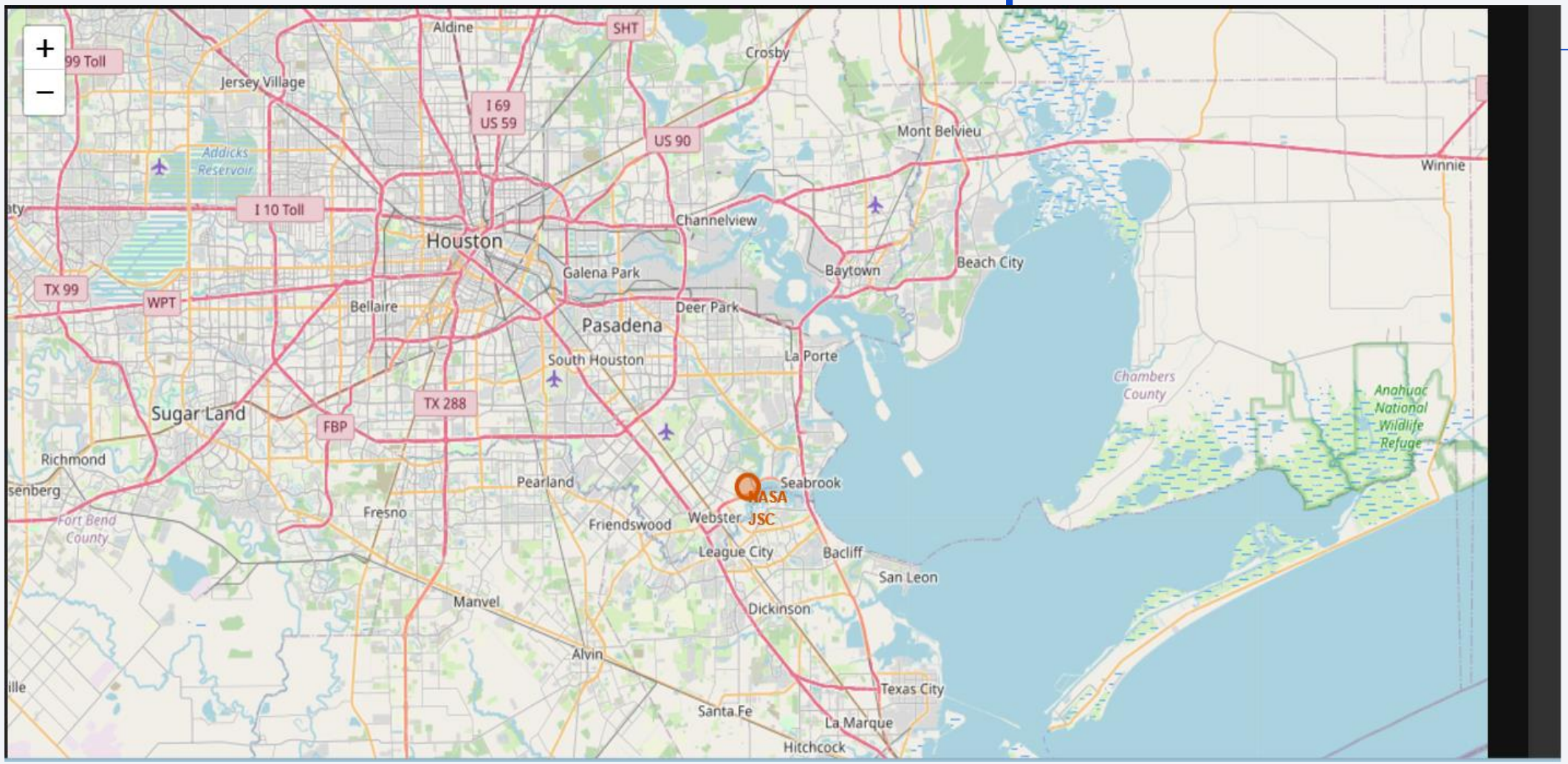
- Rank the count of landing outcomes (such as Failure (drone ship) or Success (ground pad)) between the date 2010-06-04 and 2017-03-20, in descending order

- Present your query result with a short explanation here

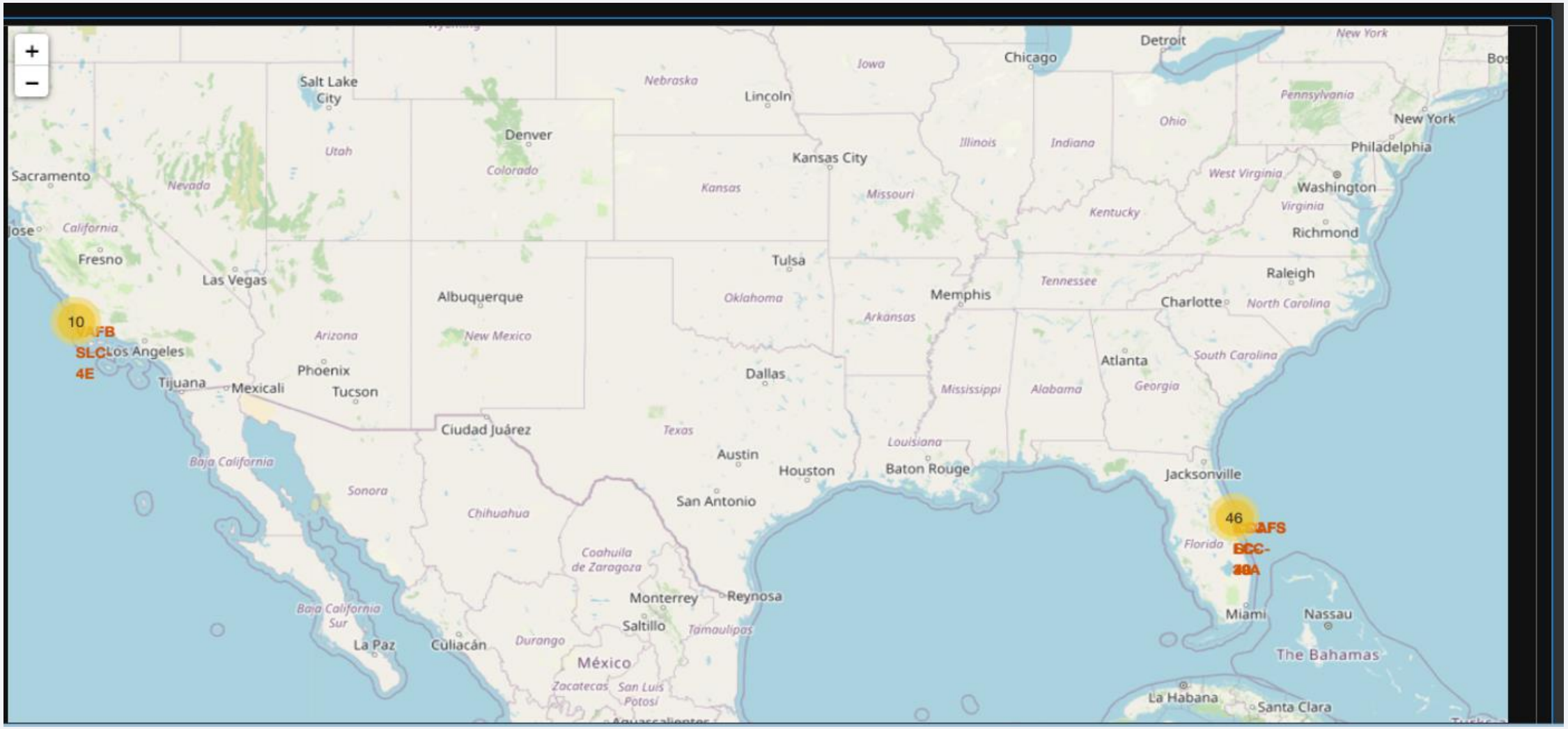
Section 3

Launch Sites Proximities Analysis

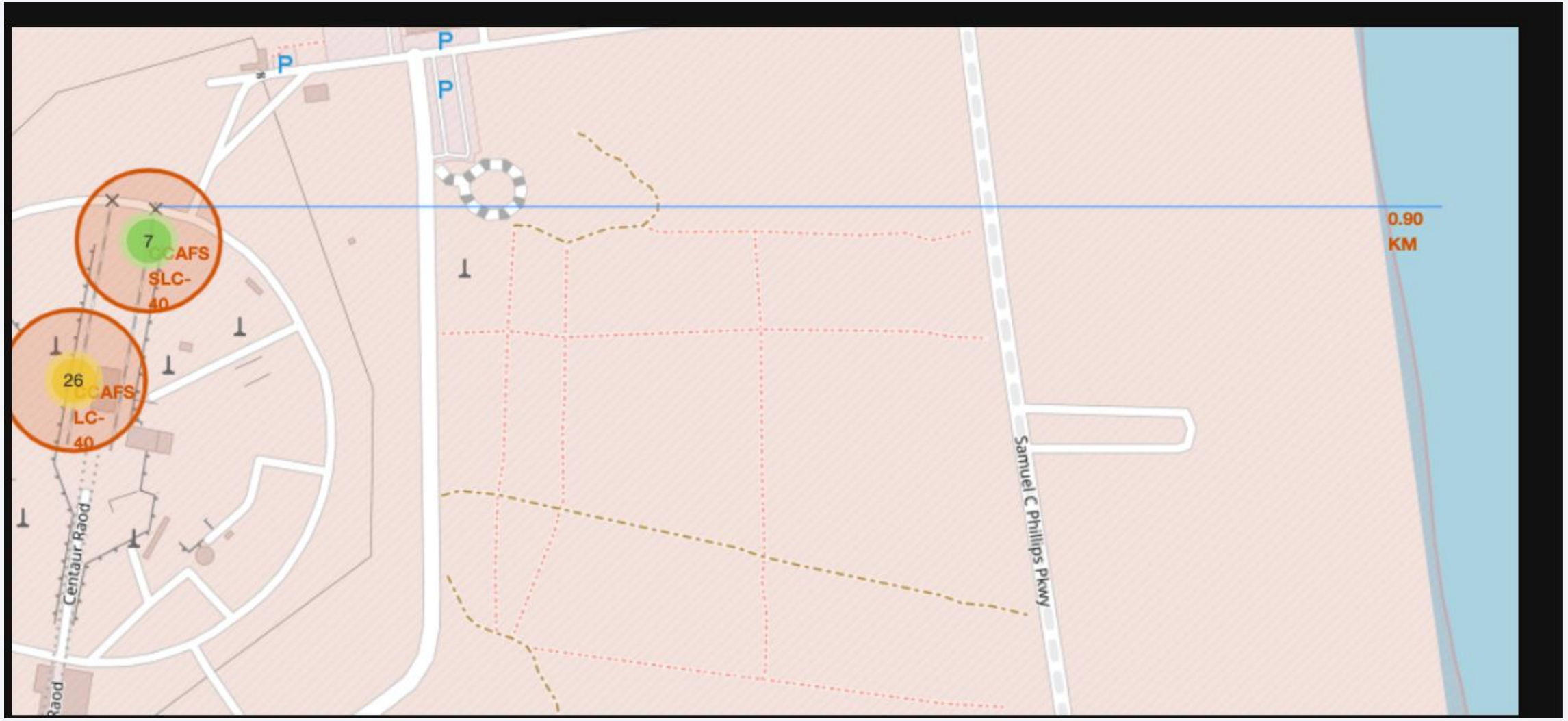
<Mark all launch sites on a map>

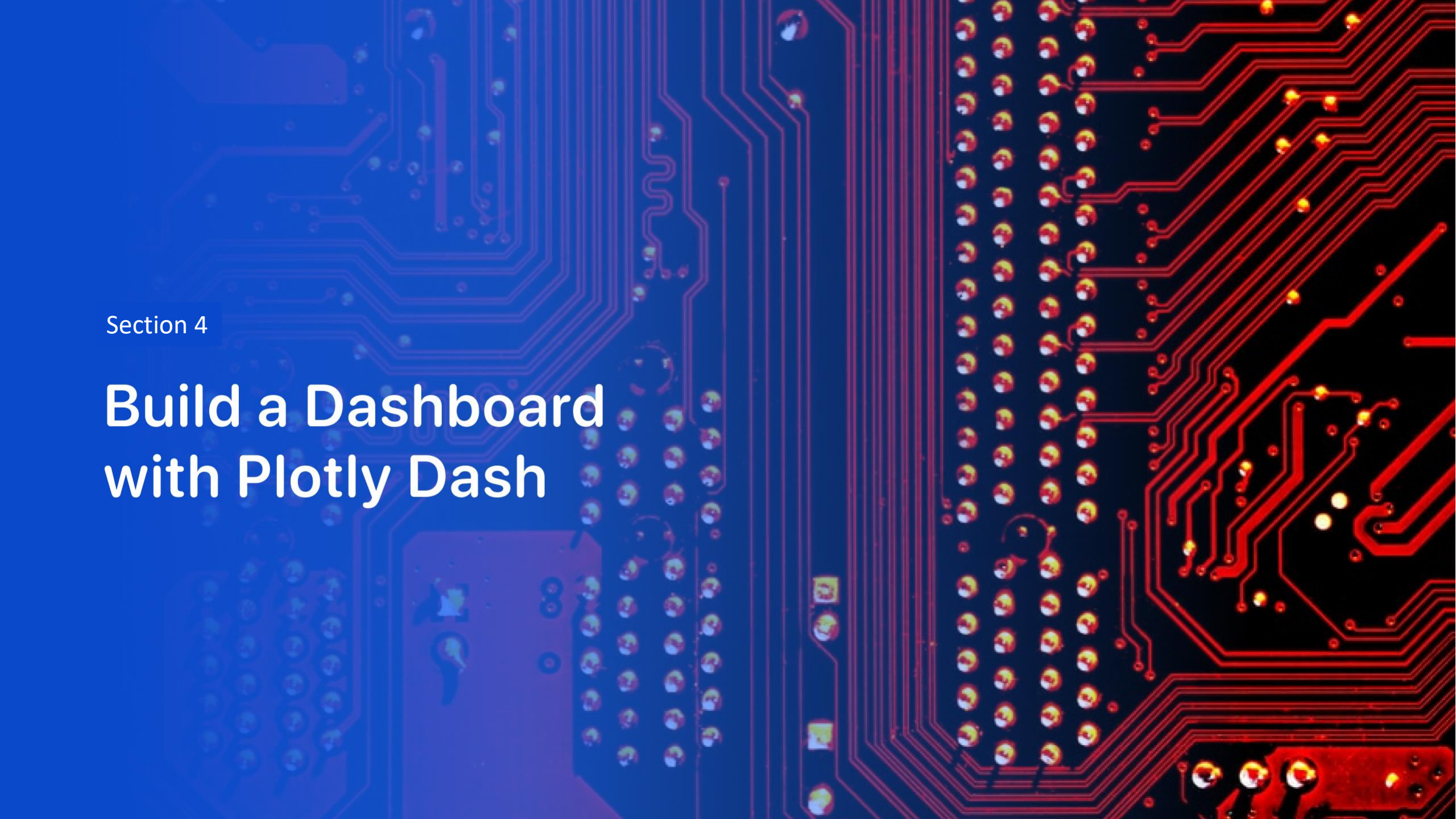


< the success/failed launches for each site on the map>

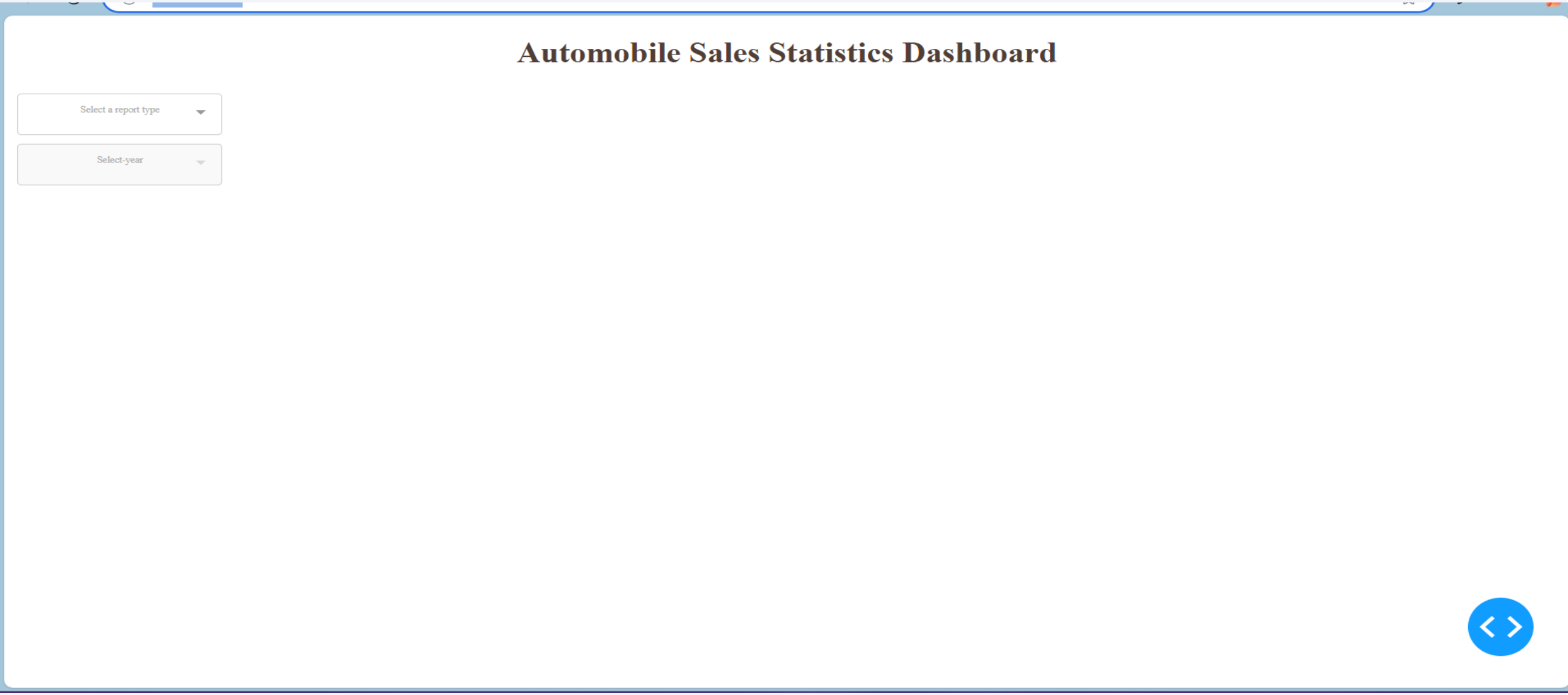


< the distances between a launch site to its proximities>

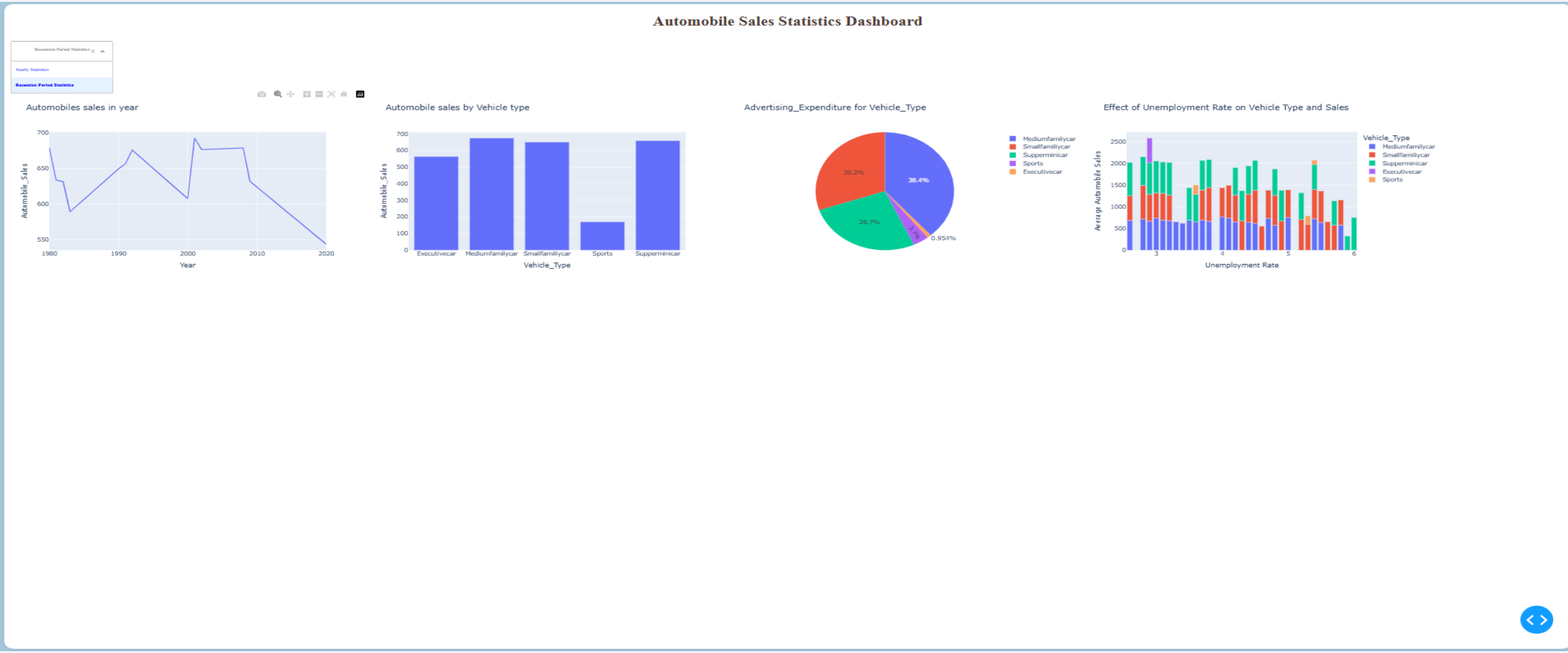



Section 4
Build a Dashboard with Plotly Dash

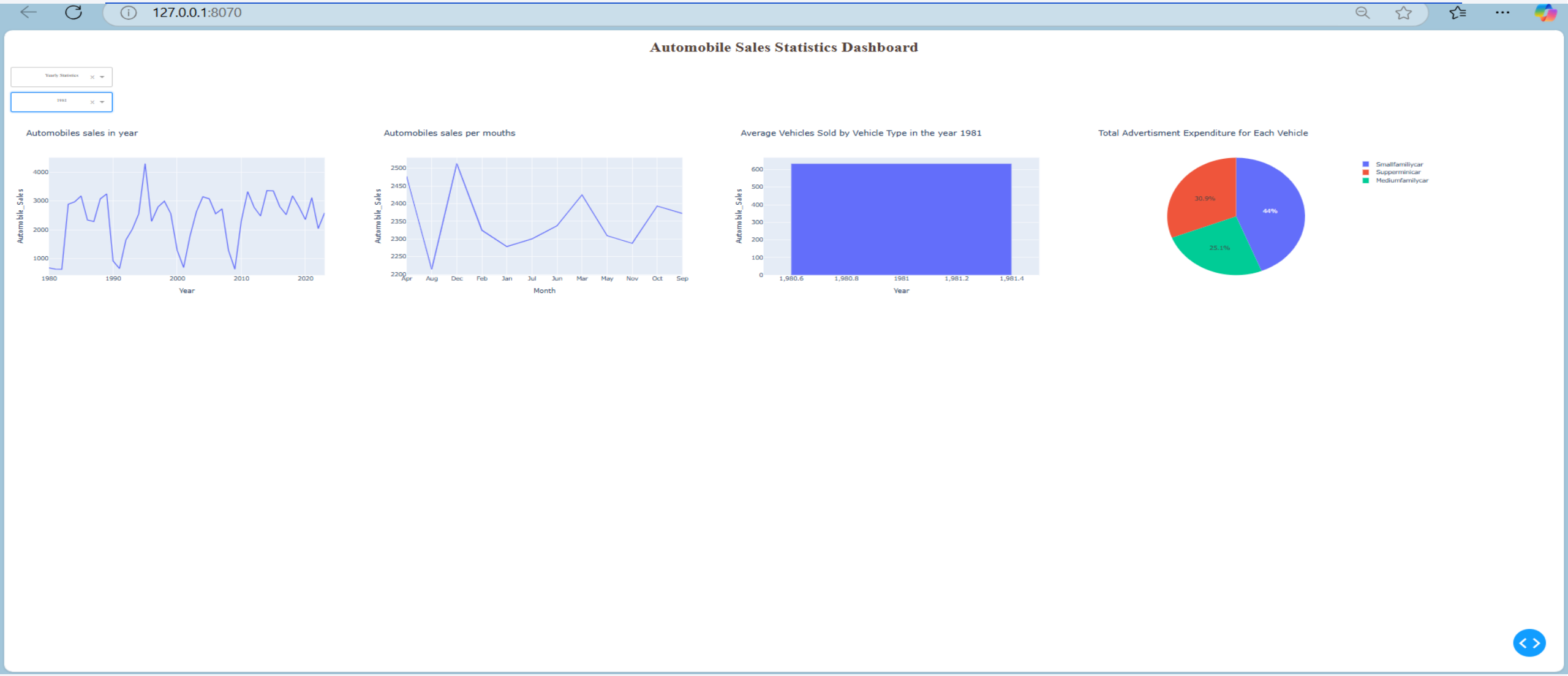
<Dashboard >


Automobile Sales Statistics Dashboard
Select a report type
Select-year

<RECESSION STATISTICS>



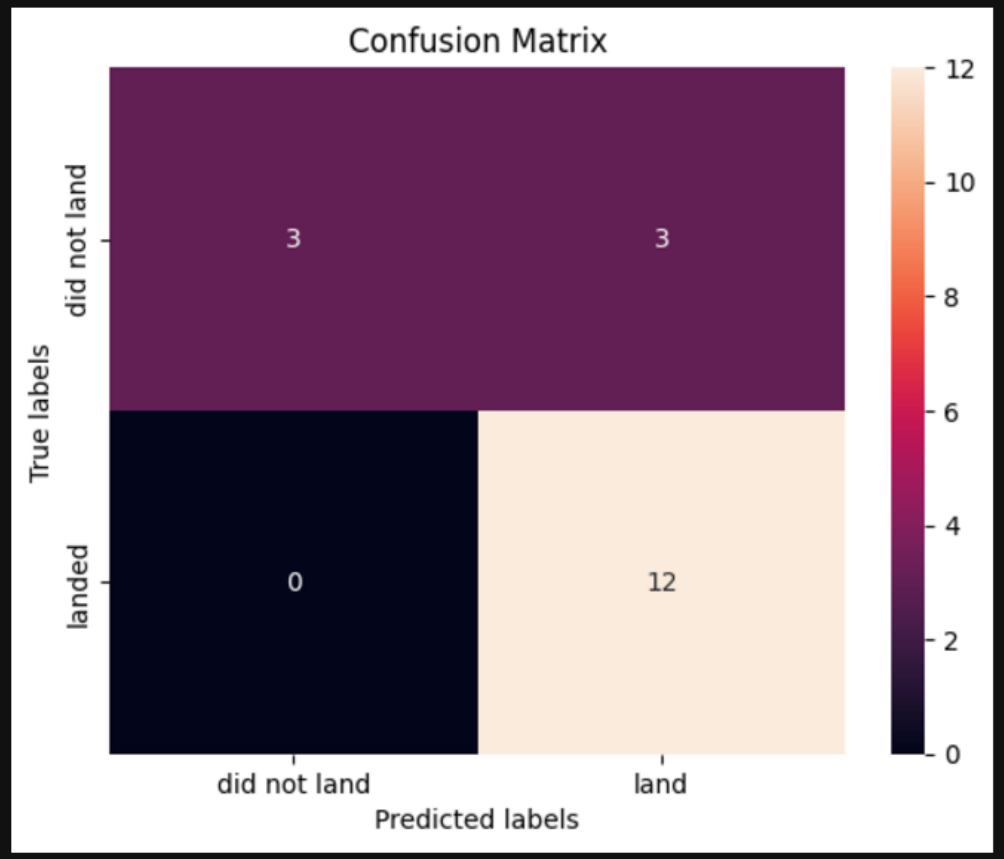
<YEARLY STATISTICS >



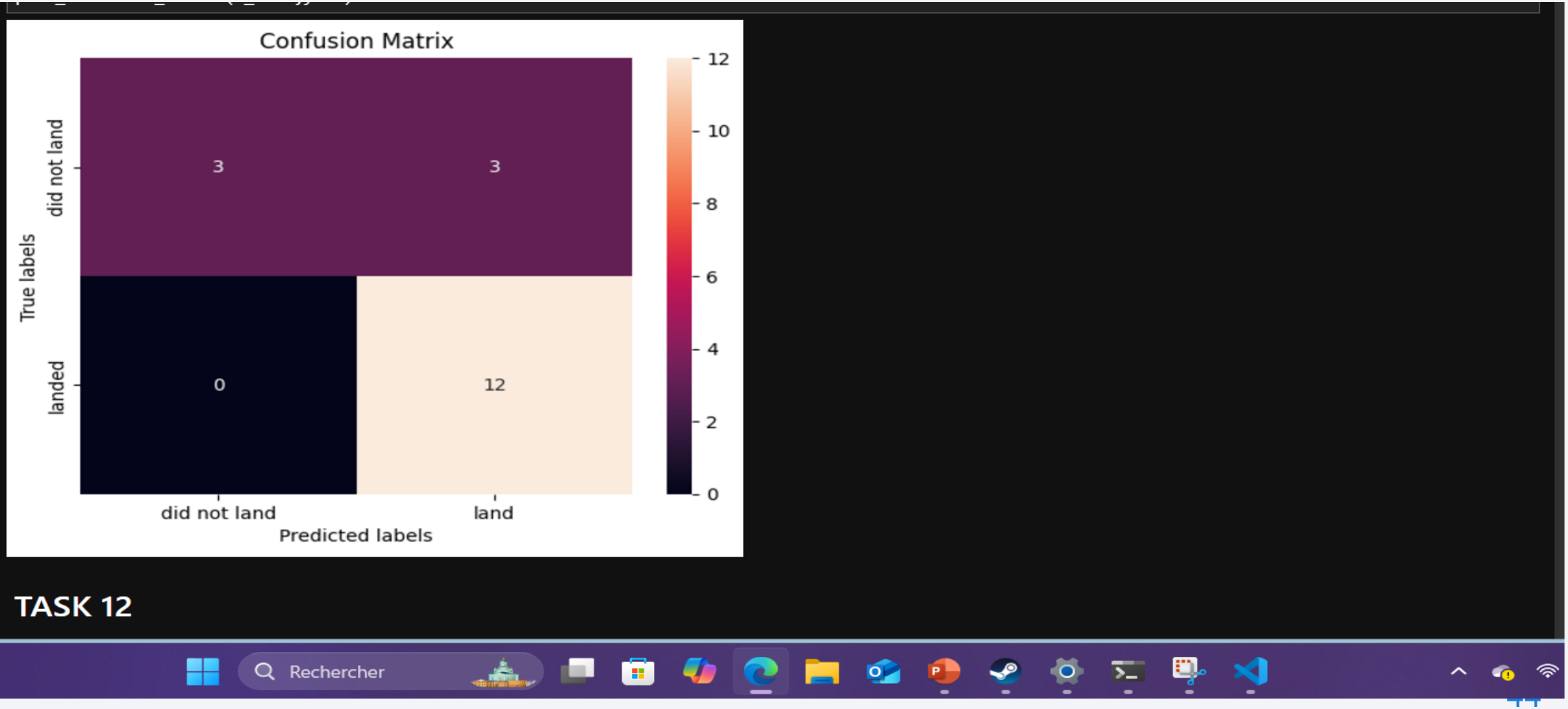
Section 5

Predictive Analysis (Classification)

Classification Accuracy



Confusion Matrix FOR BEST MODEL



Conclusions

Efficiency and Distribution of Launches: The analysis of SpaceX's launch site data reveals a strategic distribution of launches across various sites. This not only maximizes operational efficiency but also reduces turnaround times between launches. The data shows a trend of increasing successful launches, demonstrating the growing reliability of the chosen sites and the technologies employed. Reusability of Rocket Stages: One of SpaceX's key strengths is the reusability of its first rocket stages. The data analysis indicates a significant reduction in costs due to this reusability. Launches where the first stage is recovered show substantial savings compared to traditional launches, giving SpaceX a major competitive advantage in the market. Diversity of Orbits Achieved: The data on achieved orbits shows that SpaceX is capable of launching a variety of missions, covering a wide range of orbits from low Earth orbit (LEO) to geostationary orbit (GEO). This flexibility enables SpaceX to meet the diverse needs of its clients, whether for satellite deployment, space exploration missions, or manned flights.

Thank you!