



IBM Developer  
SKILLS NETWORK

# Winning Space Race with Data Science

Shaddai Lafleur  
16/10/2023



# Outline

---

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

# Executive Summary

---

- Summary of methodologies
- Summary of all results

# Introduction

---

- Predict if the Falcon 9 first stage will land successfully. SpaceX advertises Falcon 9 rocket launches on its website with a cost of 62 million dollars; other providers cost upward of 165 million dollars each, much of the savings is because SpaceX can reuse the first stage. Therefore if we can determine if the first stage will land, we can determine the cost of a launch



Section 1

# Methodology

# Methodology

---

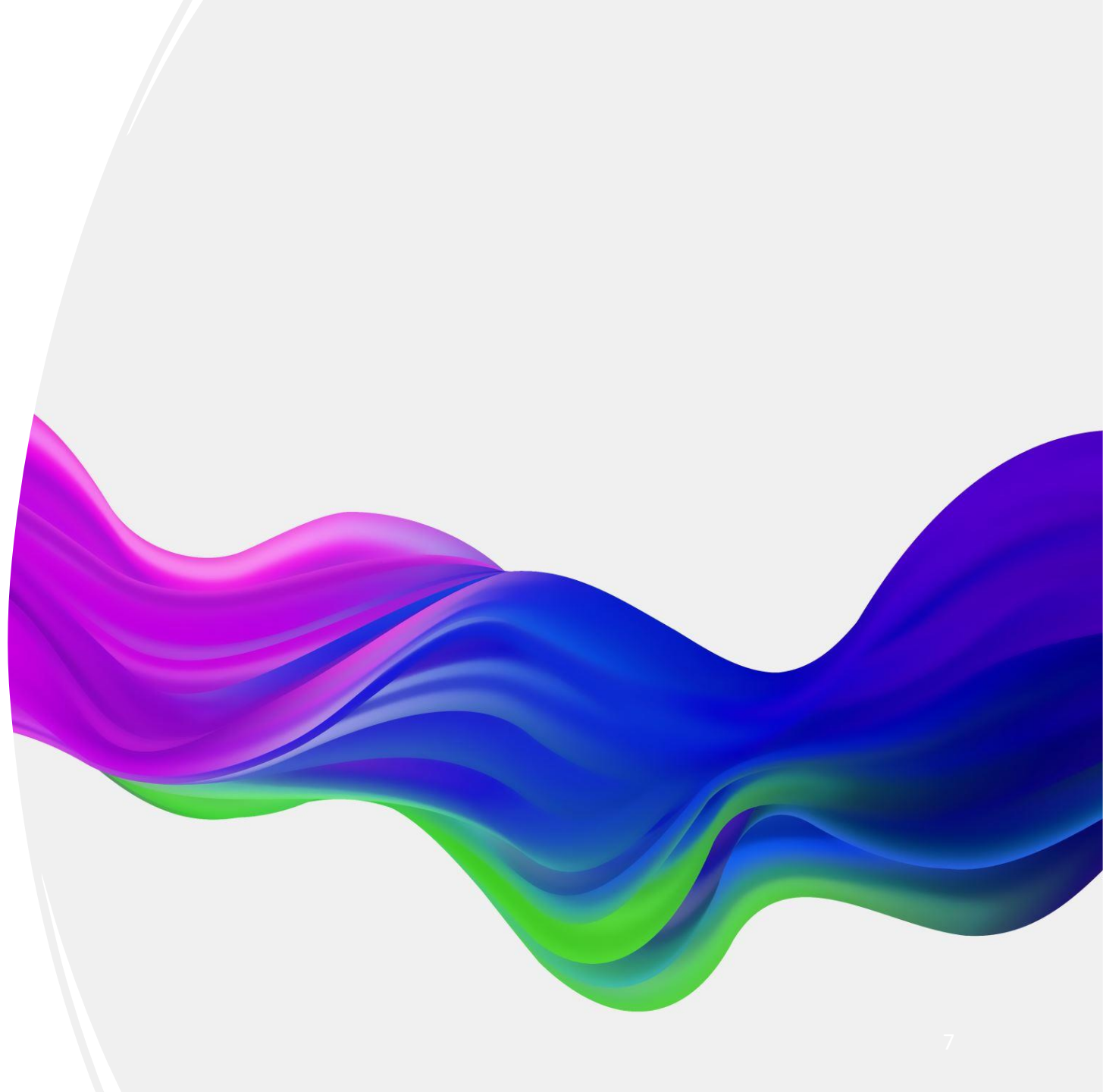
## Executive Summary

- Data collection methodology:
  - Describe how data was collected
- Perform data wrangling
  - Describe how data was processed
- Perform exploratory data analysis (EDA) using visualization and SQL
- Perform interactive visual analytics using Folium and Plotly Dash
- Perform predictive analysis using classification models
  - How to build, tune, evaluate classification models

# Data Collection

---

- Describe how data sets were collected.
- You need to present your data collection process use key phrases and flowcharts



# Data Collection – SpaceX API

- Request rocket launch data from SpaceX API with the following URL
- <https://github.com/Shady-90/dataScience/blob/main/jupyter-labs-spacex-data-collection-api.ipynb>

## SpaceX Falcon 9 first stage Landing Prediction

### Lab 1: Collecting the data

Estimated time needed: **45** minutes

In this capstone, we will predict if the Falcon 9 first stage will land successfully. SpaceX advertises Falcon 9 rocket launches on its website with a cost of 62 million dollars; other providers cost upward of 165 million dollars each, much of the savings is because SpaceX can reuse the first stage. Therefore if we can determine if the first stage will land, we can determine the cost of a launch. This information can be used if an alternate company wants to bid against SpaceX for a rocket launch. In this lab, you will collect and make sure the data is in the correct format from an API. The following is an example of a successful and launch.



Several examples of an unsuccessful landing are shown here:



# Data Collection - Scraping

- Request the Falcon9 Launch Wiki page from its URL
- Extract all column/variable names from the HTML table header
- Create a data frame by parsing the launch HTML tables
- <https://github.com/Shady-90/dataScience/blob/main/jupyter-labs-webscraping.ipynb>

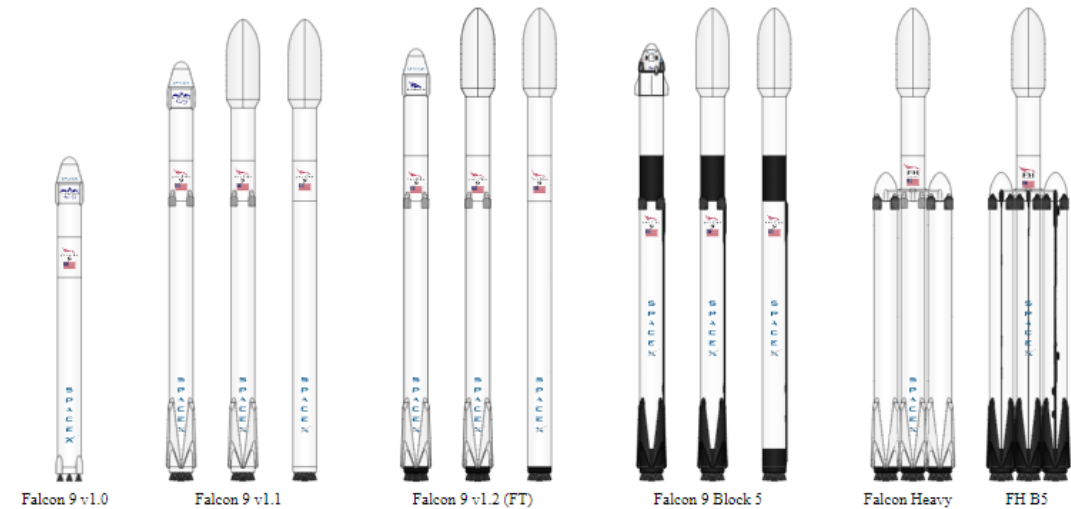
## Space X Falcon 9 First Stage Landing Prediction

### Web scraping Falcon 9 and Falcon Heavy Launches Records from Wikipedia

Estimated time needed: **40** minutes

In this lab, you will be performing web scraping to collect Falcon 9 historical launch records from a Wikipedia page titled `List of Falcon 9 and Falcon Heavy launches`

[https://en.wikipedia.org/wiki/List\\_of\\_Falcon\\_9\\_and\\_Falcon\\_Heavy\\_launches](https://en.wikipedia.org/wiki/List_of_Falcon_9_and_Falcon_Heavy_launches)



# Data Wrangling

---

- Describe how data were processed
- You need to present your data wrangling process using key phrases and flowcharts
- Add the GitHub URL of your completed data wrangling related notebooks, as an external reference and peer-review purpose

# EDA with Data Visualization

---

- Summarize what charts were plotted and why you used those charts
- Add the GitHub URL of your completed EDA with data visualization notebook, as an external reference and peer-review purpose

# EDA with SQL

---

- Using bullet point format, summarize the SQL queries you performed
- [https://github.com/Shady-90/dataScience/blob/main/jupyter-labs-eda-sql-coursera\\_sqlite.ipynb](https://github.com/Shady-90/dataScience/blob/main/jupyter-labs-eda-sql-coursera_sqlite.ipynb)

# Build an Interactive Map with Folium

---

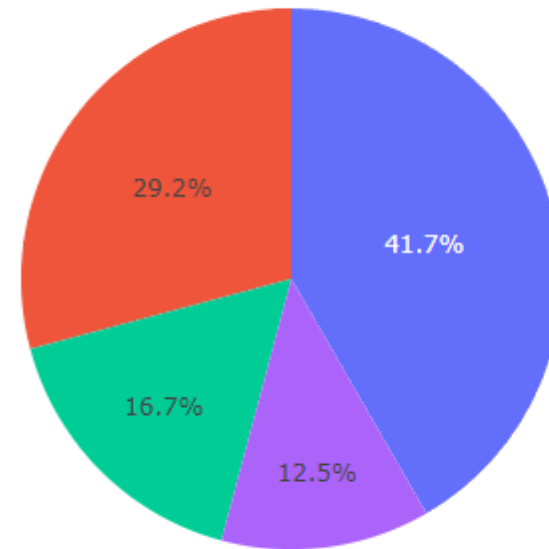
- Summarize what map objects such as markers, circles, lines, etc. you created and added to a folium map
- Explain why you added those objects
- Add the GitHub URL of your completed interactive map with Folium map, as an external reference and peer-review purpose



# Build a Dashboard with Plotly Dash

---

Success Count for all launch sites



# Predictive Analysis (Classification)

---

- Summarize how you built, evaluated, improved, and found the best performing classification model
- You need present your model development process using key phrases and flowchart
- Add the GitHub URL of your completed predictive analysis lab, as an external reference and peer-review purpose

# Results



EXPLORATORY DATA  
ANALYSIS RESULTS



INTERACTIVE ANALYTICS  
DEMO IN SCREENSHOTS



PREDICTIVE ANALYSIS  
RESULTS



The background of the slide is an abstract composition. It features a dark blue base color. Overlaid on this are numerous diagonal streaks in shades of red and cyan. A faint, light blue grid pattern is also visible, particularly in the lower half of the image. The overall effect is dynamic and technological.

Section 2

# Insights drawn from EDA



# Flight Number vs. Launch Site

---

- [https://github.com/Shady-90/dataScience/blob/main/Build%20an%20Interactive%20Dashboard%20with%20Plotly%20Dash%20-%20spacex\\_dash\\_app.py](https://github.com/Shady-90/dataScience/blob/main/Build%20an%20Interactive%20Dashboard%20with%20Plotly%20Dash%20-%20spacex_dash_app.py)



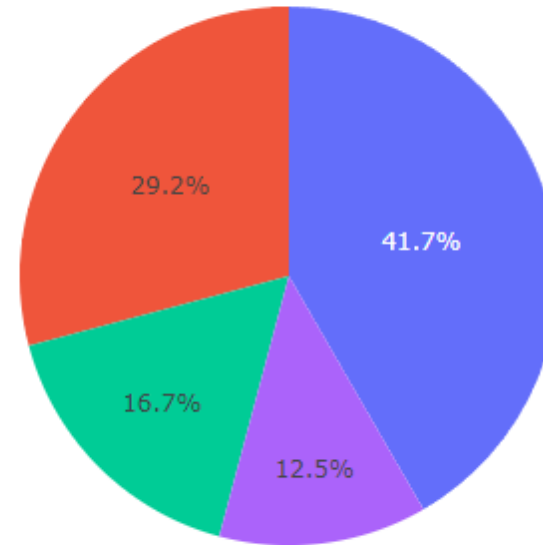
# Payload vs. Launch Site

---

- Show a scatter plot of Payload vs. Launch Site
- Show the screenshot of the scatter plot with explanations

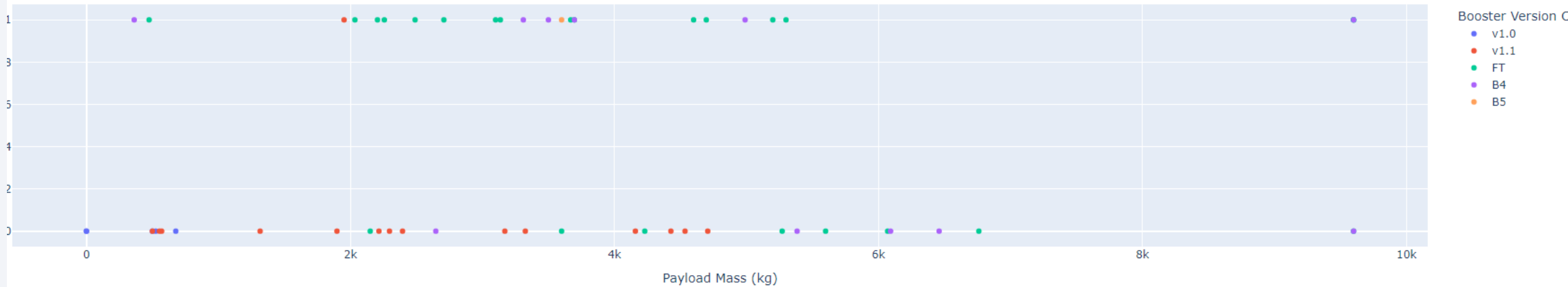
# Success Rate vs. Orbit Type

Success Count for all launch sites



# Flight Number vs. Orbit Type

Success count on Payload mass for all sites



# 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

---

**Launch\_Site**

---

CCAFS LC-40

VAFB SLC-4E

KSC LC-39A

CCAFS SLC-40

# Launch Site Names Begin with 'CCA'

---

**Launch\_Site**

CCAFS LC-40

CCAFS LC-40

CCAFS LC-40

CCAFS LC-40

CCAFS LC-40

# Total Payload Mass

---

- The total payload carried by boosters from NASA: 45596
- Present your query result with a short explanation here

# Average Payload Mass by F9 v1.1

---

- The average payload mass carried by booster version F9 v1.1: 340.4

# First Successful Ground Landing Date

- The dates of the first successful landing outcome on ground pad:

2015-12-22



# Successful Drone Ship Landing with Payload between 4000 and 6000

## Booster\_Version

F9 FT B1021.1

F9 FT B1022

F9 FT B1023.1

F9 FT B1026

F9 FT B1029.1

F9 FT B1021.2

F9 FT B1029.2

F9 FT B1036.1

F9 FT B1038.1

F9 B4 B1041.1

F9 FT B1031.2

F9 B4 B1042.1

F9 B4 B1045.1

F9 B5 B1046.1

# Total Number of Successful and Failure Mission Outcomes

- Calculate the total number of successful and failure mission outcomes

Mission_Outcome	TOTAL_NUMBER
Failure (in flight)	1
Success	98
Success	1
Success (payload status unclear)	1

# Boosters Carried Maximum Payload

- List the names of the booster which have carried the maximum payload mass

---

F9 B5 B1048.4

F9 B5 B1049.4

F9 B5 B1051.3

F9 B5 B1056.4

F9 B5 B1048.5

F9 B5 B1051.4

F9 B5 B1049.5

F9 B5 B1060.2

F9 B5 B1058.3

F9 B5 B1051.6

F9 B5 B1060.3

F9 B5 B1049.7

# 2015 Launch Records

---

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

# 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

Landing_Outcome	TOTAL_NUMBER
No attempt	10
Success (ground pad)	5
Success (drone ship)	5
Failure (drone ship)	5
Controlled (ocean)	3
Uncontrolled (ocean)	2
Precluded (drone ship)	1
Failure (parachute)	1

A satellite view of Earth from space, showing the curvature of the planet and city lights at night. The background is a deep blue gradient.

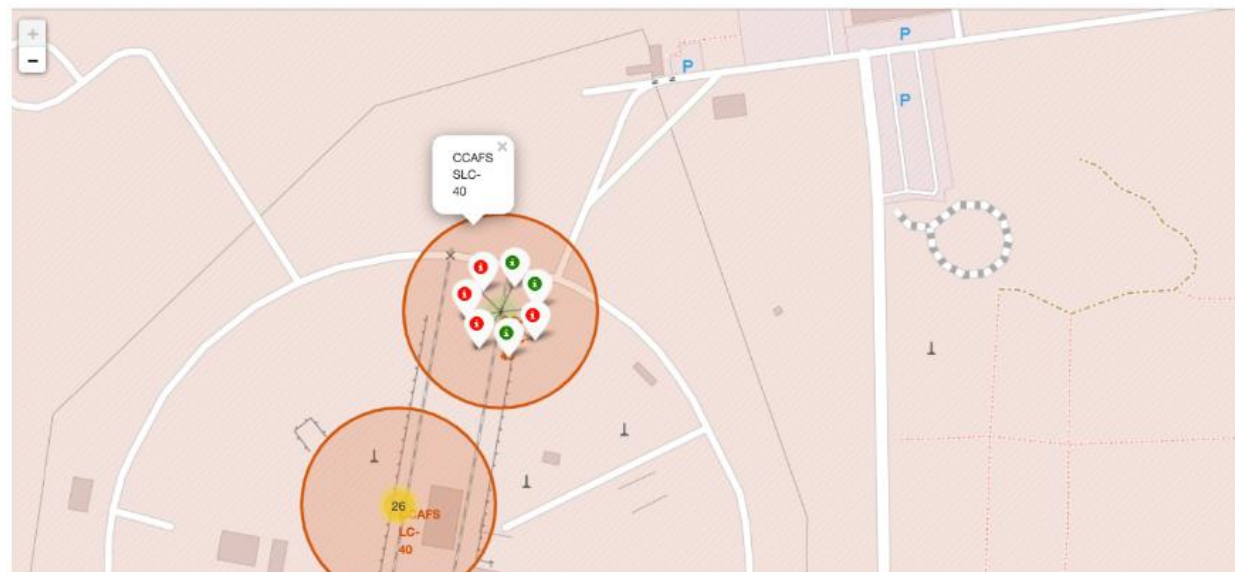
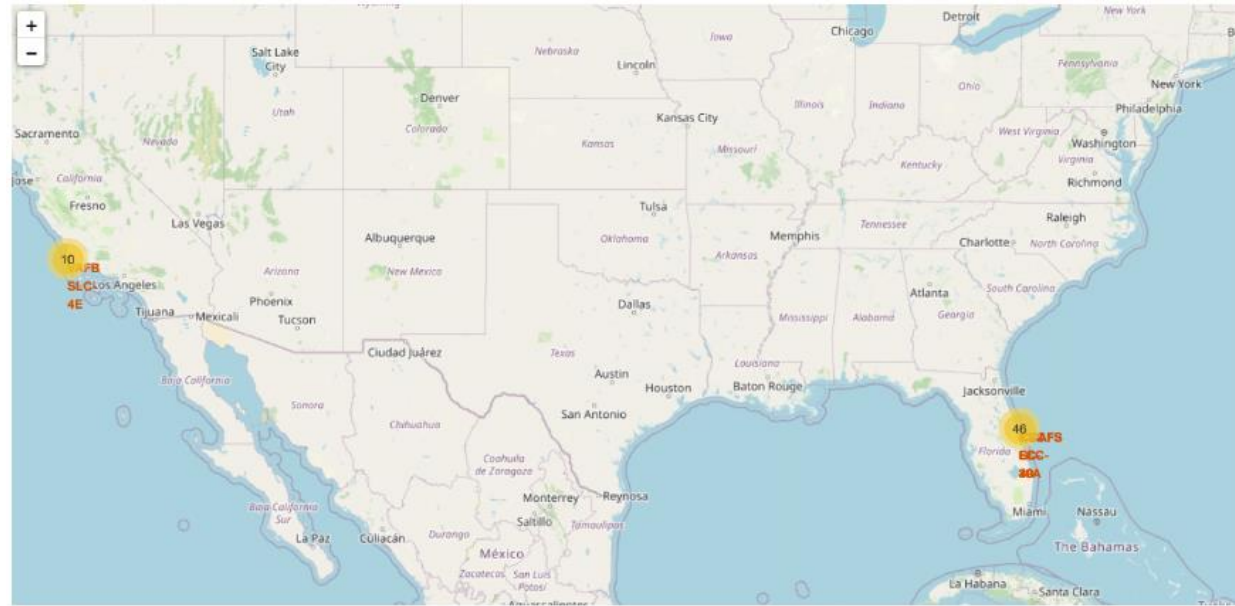
Section 3

# Launch Sites Proximities Analysis

# Launch Sites Locations Analysis with Folium











Section 4

# Build a Dashboard with Plotly Dash

# <Dashboard Screenshot 1>

---

- Replace <Dashboard screenshot 1> title with an appropriate title
- Show the screenshot of launch success count for all sites, in a piechart
- Explain the important elements and findings on the screenshot

## <Dashboard Screenshot 2>

---

- Replace <Dashboard screenshot 2> title with an appropriate title
- Show the screenshot of the piechart for the launch site with highest launch success ratio
- Explain the important elements and findings on the screenshot

## <Dashboard Screenshot 3>

---

- Replace <Dashboard screenshot 3> title with an appropriate title
- Show screenshots of Payload vs. Launch Outcome scatter plot for all sites, with different payload selected in the range slider
- Explain the important elements and findings on the screenshot, such as which payload range or booster version have the largest success rate, etc.

Section 5

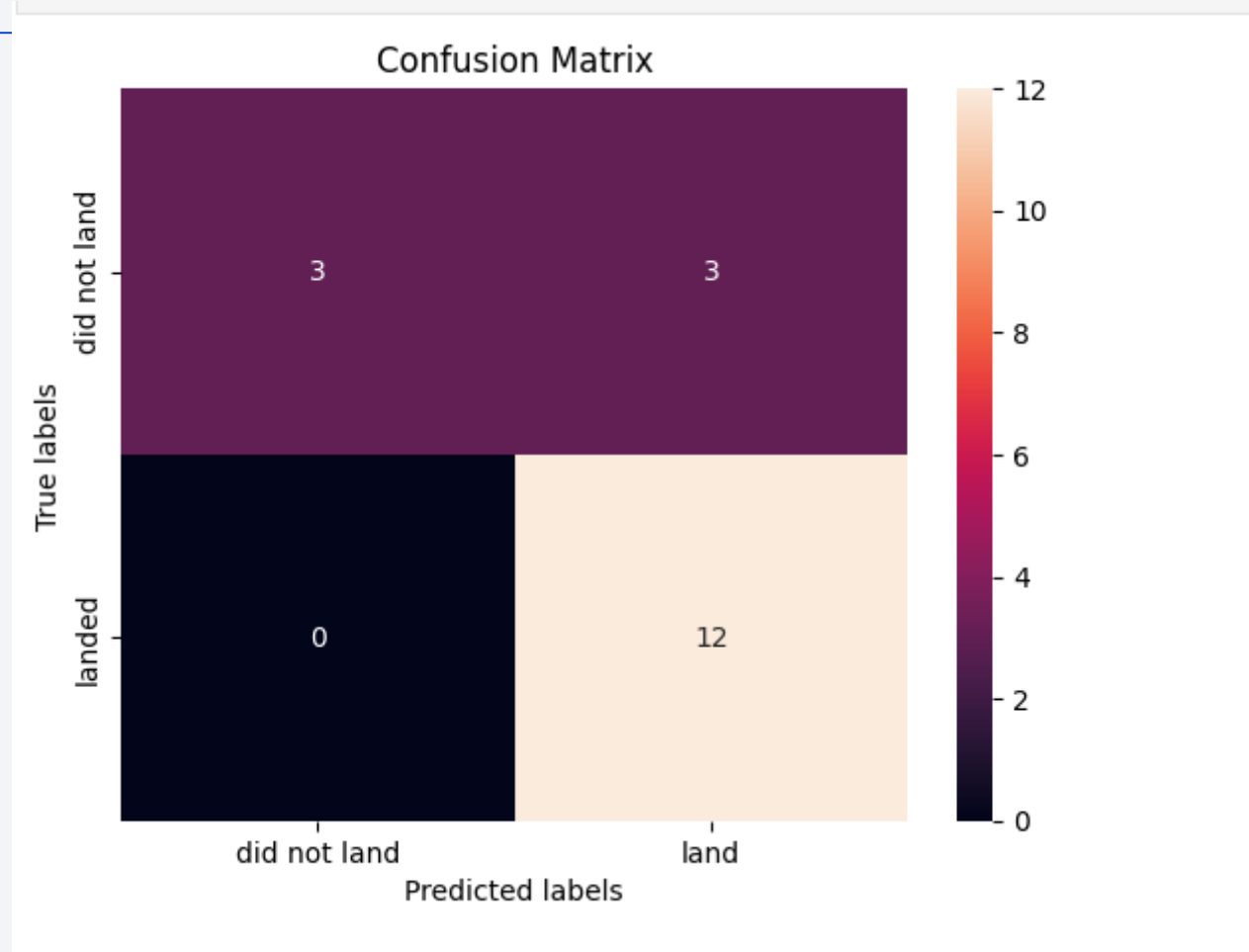
# Predictive Analysis (Classification)

# Classification Accuracy

---

- Visualize the built model accuracy for all built classification models, in a bar chart
- Find which model has the highest classification accuracy

# Confusion Matrix



# Conclusions

---

- **KSC LC – 39 A : 41.7 %**
- **NASA (CRS): 45596**



# Appendix

---

- Include any relevant assets like Python code snippets, SQL queries, charts, Notebook outputs, or data sets that you may have created during this project

Thank you!

