

# Winning Space Race with Data Science

Antonio Luque  
August, 2025



# Outline

---

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

# Executive Summary

---

- Summary of methodologies
  - Data Collection though API
  - Data Collection with Web Scraping
  - Data Wrangling
  - Exploratory Data Analysis with SQL
  - Exploratory Data Analysis with Data Visualization
  - Interactive Visual Analytics with Folium
  - Machine Learning Prediction
- Summary of all results
  - Exploratory Data Analysis result
  - Interactive analytics in screenshots
  - Predictive Analytics result

# Introduction

---

- Project background and context

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 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.

- Problem you want to find answers

Predicting if the Falcon 9 first stage will land successfully.

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

---

- Present your data collection with SpaceX REST calls using key phrases and flowcharts
- Add the GitHub URL of the completed SpaceX API calls notebook (must include completed code cell and outcome cell), as an external reference and peer-review purpose

[https://github.com/antoniyo6lt/  
SpaceX\\_Falcon9/blob/main/jupyter-labs-  
spacex-data-collection-api.ipynb](https://github.com/antoniyo6lt/SpaceX_Falcon9/blob/main/jupyter-labs-spacex-data-collection-api.ipynb)

# Data Collection - Scraping

---

- Present your web scraping process using key phrases and flowcharts
- Add the GitHub URL of the completed web scraping notebook, as an external reference and peer-review purpose

[https://github.com/antoniyo6lt/  
SpaceX\\_Falcon9/blob/main/jupyter-labs-  
webscraping.ipynb](https://github.com/antoniyo6lt/SpaceX_Falcon9/blob/main/jupyter-labs-webscraping.ipynb)

# Data Wrangling

---

- Describe how data were processed
- You need to present your data wrangling process using key phrases and flowcharts
- [https://github.com/antoniyo6lt/SpaceX\\_Falcon9/blob/main/labs-jupyter-spacex-Data wrangling.ipynb](https://github.com/antoniyo6lt/SpaceX_Falcon9/blob/main/labs-jupyter-spacex-Data%20wrangling.ipynb)

# EDA with Data Visualization

---

- Summarize what charts were plotted and why you used those charts
- [https://github.com/antoniyo6lt/SpaceX\\_Falcon9/blob/main/edadataviz.ipynb](https://github.com/antoniyo6lt/SpaceX_Falcon9/blob/main/edadataviz.ipynb)

# EDA with SQL

---

- Using bullet point format, summarize the SQL queries you performed
- [https://github.com/antoniyo6lt/SpaceX\\_Falcon9/blob/main/jupyter-labs-eda-sql-coursera\\_sqlite.ipynb](https://github.com/antoniyo6lt/SpaceX_Falcon9/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
- [https://github.com/antoniyo6lt/SpaceX\\_Falcon9/blob/main/  
lab\\_jupyter\\_launch\\_site\\_location.ipynb](https://github.com/antoniyo6lt/SpaceX_Falcon9/blob/main/lab_jupyter_launch_site_location.ipynb)

# Build a Dashboard with Plotly Dash

---

- Summarize what plots/graphs and interactions you have added to a dashboard
- Explain why you added those plots and interactions

# Predictive Analysis (Classification)

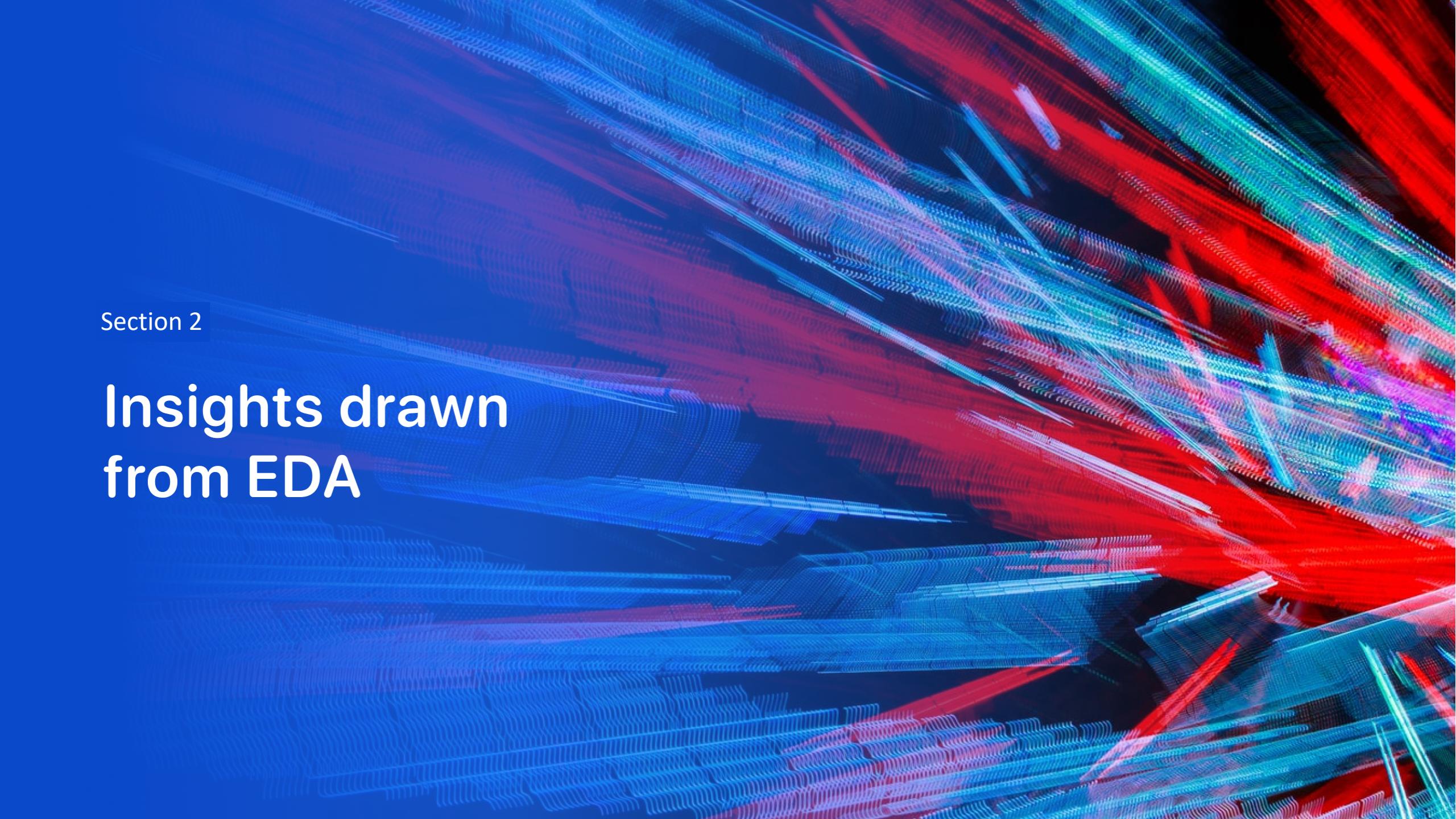
---

- 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
- [https://github.com/antoniyo6lt/SpaceX\\_Falcon9/blob/main/SpaceX\\_Machine\\_Learning\\_Prediction\\_Part\\_5.ipynb](https://github.com/antoniyo6lt/SpaceX_Falcon9/blob/main/SpaceX_Machine_Learning_Prediction_Part_5.ipynb)

# Results

---

- Exploratory data analysis results
- Interactive analytics demo in screenshots
- Predictive analysis results

The background of the slide features a complex, abstract digital visualization. It consists of numerous thin, glowing lines that create a sense of depth and motion. The lines are primarily blue and red, with some green and white highlights. They form a grid-like structure that is more dense and vibrant towards the right side of the frame, while appearing more sparse and blue-tinted on the left. The overall effect is reminiscent of a high-energy particle simulation or a futuristic circuit board.

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

The background of the slide is a photograph taken from space at night. It shows the curvature of the Earth against a dark blue and black void of space. City lights are visible as small white dots and larger clusters of light, primarily concentrated in the lower right quadrant where the United States and Mexico would be. In the upper right, there is a bright green and yellow aurora borealis or aurora australis visible in the atmosphere.

Section 3

# Launch Sites Proximities Analysis

## <Folium Map Screenshot 1>

---

- Replace <Folium map screenshot 1> title with an appropriate title
- Explore the generated folium map and make a proper screenshot to include all launch sites' location markers on a global map
- Explain the important elements and findings on the screenshot

## <Folium Map Screenshot 2>

---

- Replace <Folium map screenshot 2> title with an appropriate title
- Explore the folium map and make a proper screenshot to show the color-labeled launch outcomes on the map
- Explain the important elements and findings on the screenshot

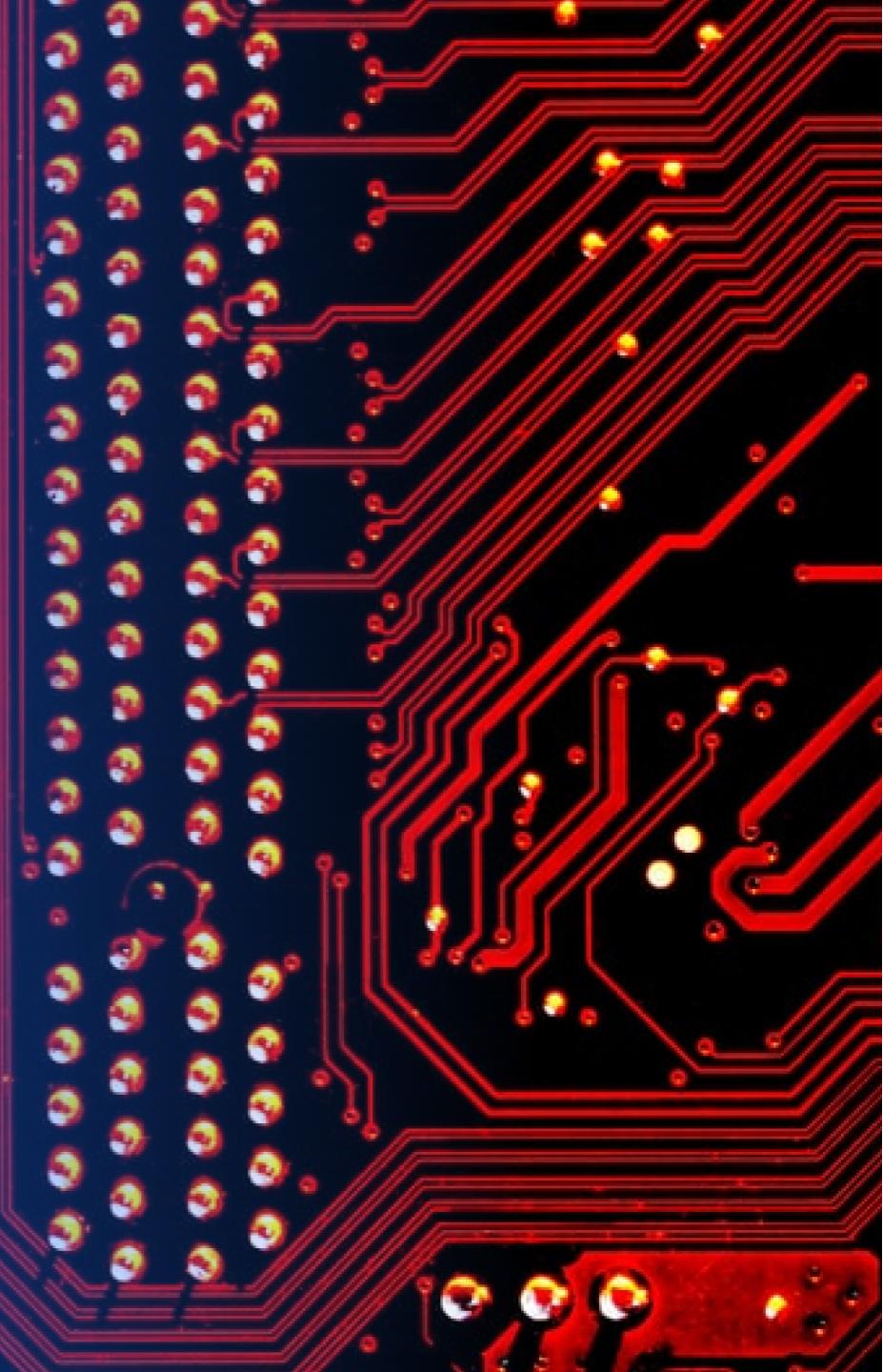
## <Folium Map Screenshot 3>

---

- Replace <Folium map screenshot 3> title with an appropriate title
- Explore the generated folium map and show the screenshot of a selected launch site to its proximities such as railway, highway, coastline, with distance calculated and displayed
- Explain the important elements and findings on the screenshot

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

---

- Show the confusion matrix of the best performing model with an explanation

# Conclusions

---

- Point 1
- Point 2
- Point 3
- Point 4
- ...

# 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!

