

# Winning Space Race with Data Science

Rafael da Silva Figueiredo  
29/11/2025



# Outline

---

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

# Executive Summary

---

- Summary of methodologies
- Summary of all results

# Introduction

---

- Project background and context
- Problems you want to find answers

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

---

Data from SpaceX's Falcon 9 rocket launches were collected from two different sources:

- SpaceX API (<https://api.spacexdata.com/v4>)
- Web scraping 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 Collection – SpaceX API

---

Specific requests were made in order to access different information related to each launch:

- **/rockets** to get the booster version;
- **/launchpads** to get the launch site and its longitude and latitude;
- **/payloads** to get the payloads attached to the rocket
- **/cores** to get information like outcome, number of flights, landing pad, among others.

GitHub URL of the completed SpaceX API calls notebook (must include completed code cell and outcome cell)

Place your flowchart of SpaceX API calls here

# 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

Place your flowchart of web scraping here

# 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

Add the GitHub URL of your completed EDA with SQL notebook, as an external reference and peer-review purpose

# 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

---

Summarize what plots/graphs and interactions you have added to a dashboard

Explain why you added those plots and interactions

Add the GitHub URL of your completed Plotly Dash lab, as an external reference and peer-review purpose

# Predictive Analysis (Classification)

---

Summarize how you built, evaluated, improved, and found the best performing classification model

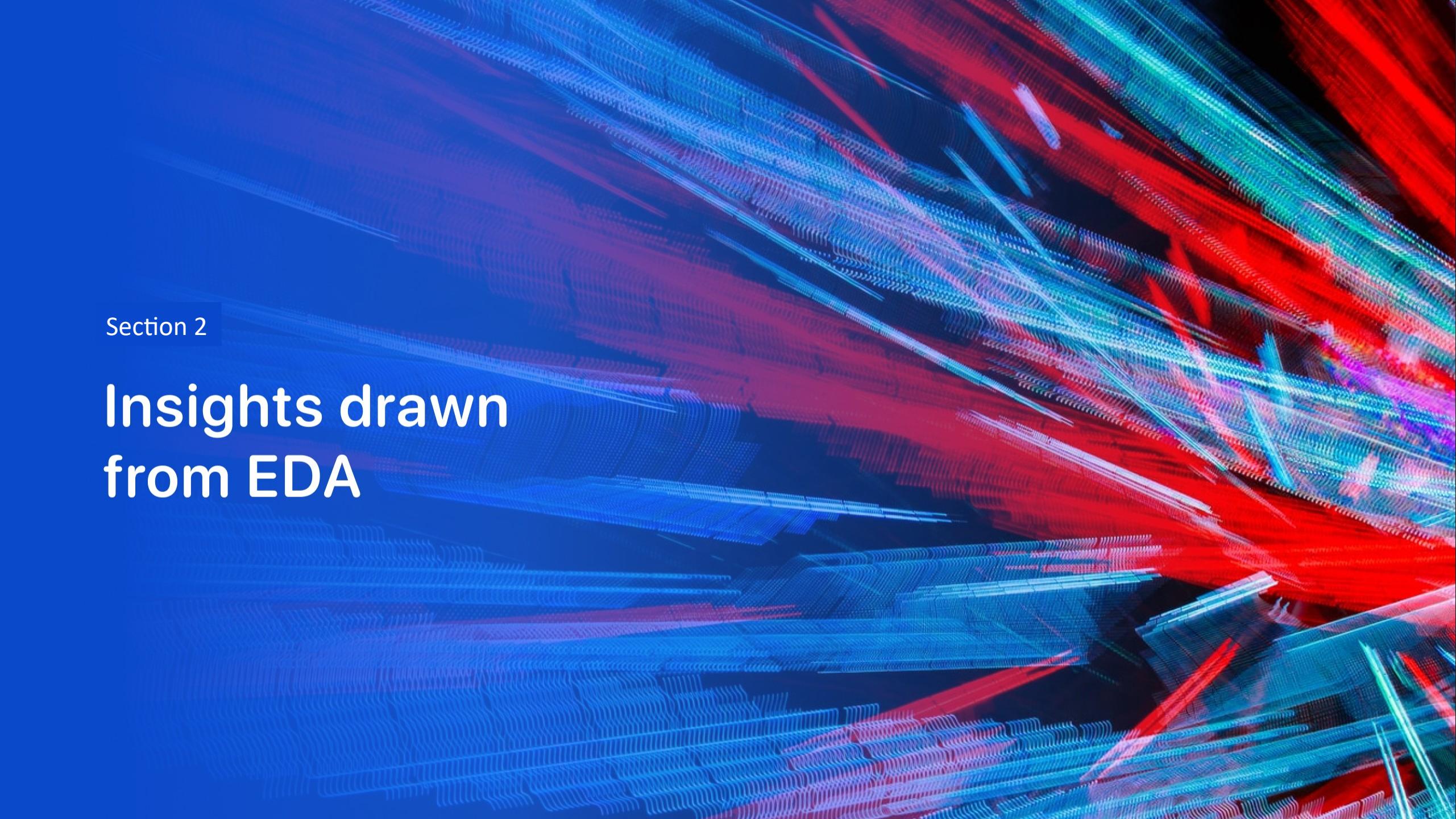
You need present your model development process using key phrases and flowchart

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 features a complex, abstract pattern of glowing lines. These lines are primarily blue and red, creating a sense of depth and motion. They appear to be composed of many small, individual particles or segments, giving them a textured, almost organic appearance. The lines converge and diverge, forming various shapes and directions across the dark, solid-colored background.

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's horizon against a dark blue sky. City lights are visible as small white dots, with larger clusters of lights indicating major urban areas. In the upper right quadrant, there is a bright green and yellow glow, likely representing the Aurora Borealis or a similar natural light display.

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>

---

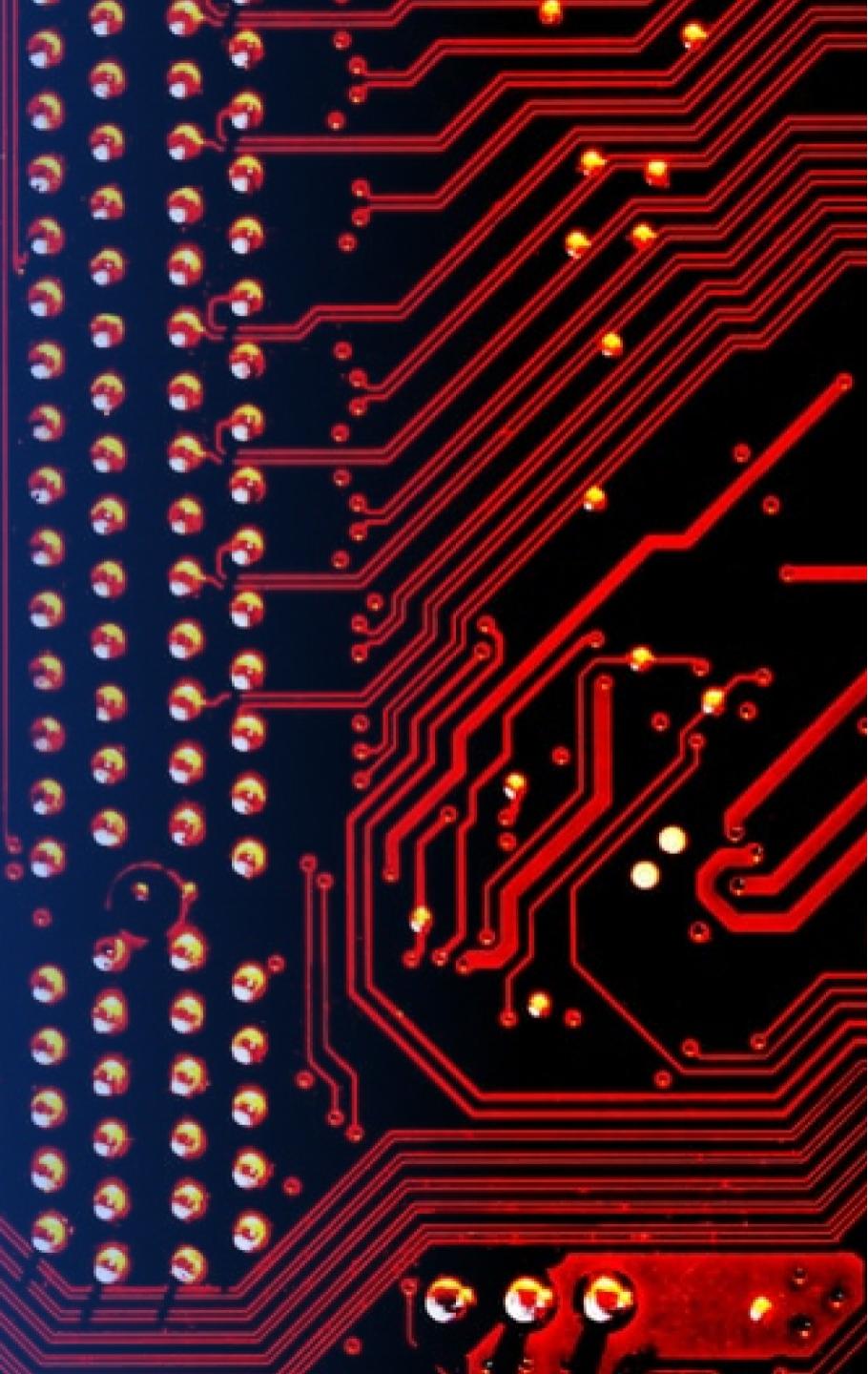
Replace <Folium map screenshot 3> title with an appropriate title

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!

