Winning Space Race with Data Science

Bolu Amosu
9th January 2025

Outline

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

Executive Summary

Summary of methodologies

SpaceX Data collected using SpaceX API and Web Scraping

SpaceX Data wrangling

SpaceX EDA using SQL

SpaceX EDA using Pandas and Matplotlib

SpaceX Machine Learning landing prediction

Summary of all results

EDA results

Interactive Visual Analytics and Dashboards

Predictive Analysis for classification

Introduction

Project background and context

SpaceX advertises its Falcon 9 rocket launches as cheaper than its competitors as SpaceX can reuse the first stage. The rocket is designed to return to earth after launch, landing either on land - at one of SpaceX's landing zone – or on an autonomous drone ship in the ocean.

Due to this, if we can determine if the first stage will land, it is possible to determine the cost of a launch.

Problems you want to find answers

Due to this, if we can determine if the first stage will land, it is possible to determine the cost of a launch.

We will use data from Falcon 9 rocket launches advertised on its website to achieve this.

Section 1

Methodology

Methodology

Executive Summary

- Data collection methodology:
  - Data was gathered using two main sources - SpaceX API and web scraping on Wikipedia
- Data wrangling
- Exploratory data analysis (EDA) using visualization and SQL
- Interactive visual analytics using Folium and Plotly Dash
- Predictive analysis using classification models
  - How to build, tune, evaluate classification models

Data Collection

The data was first collected using SpaceX API by making a request.

Web scraping was also performed to collect Falcon 9's historical launch records from Wikipedia page of the launch records. By using BeautifulSoup, it was possible to extract the table of records from Wikipedia and convert it to a Pandas data frame.

EDA with Data Visualization

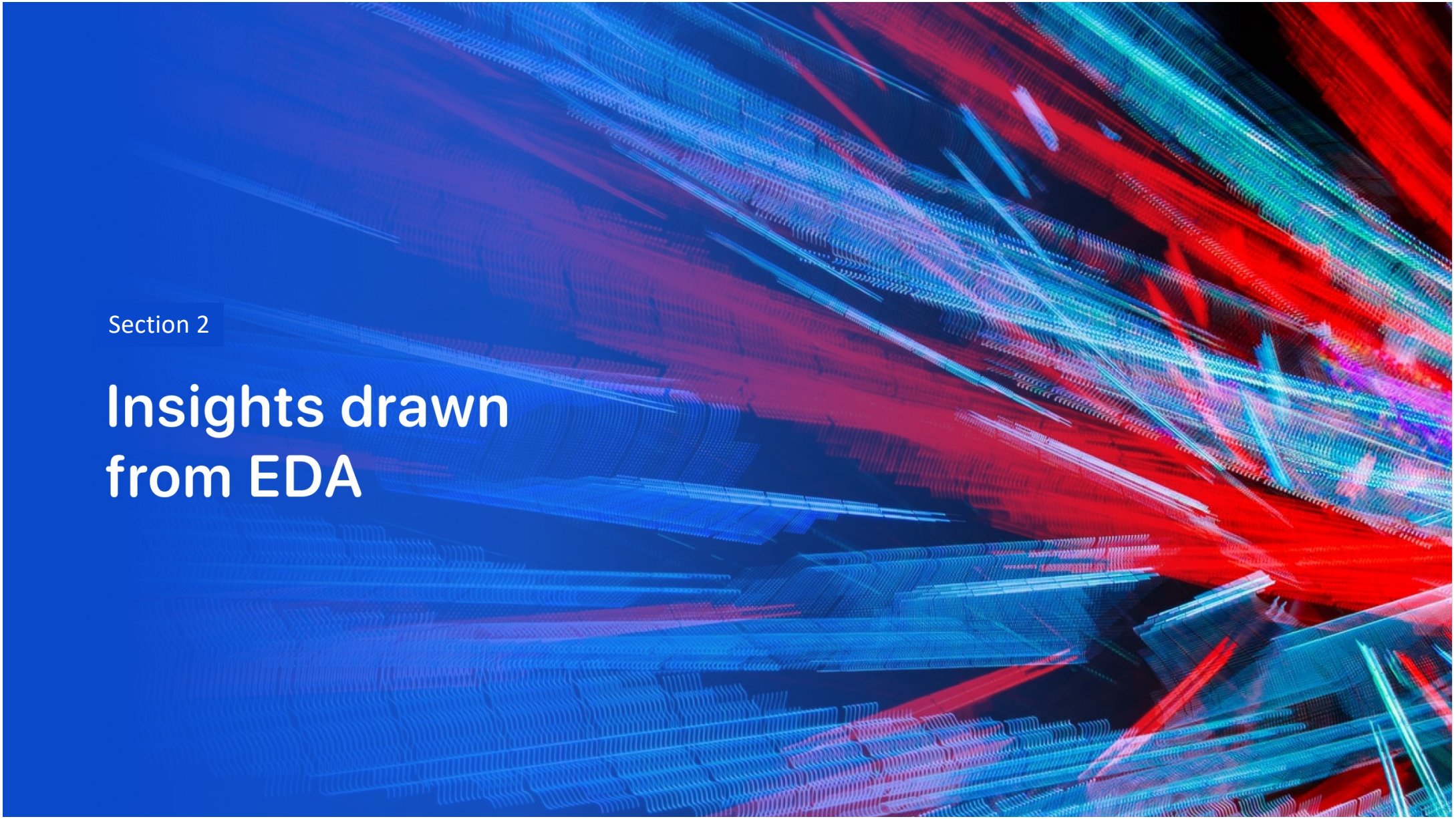
- Summarize what charts were plotted and why you used those charts
- Add the GitHub URL of your completed EDA with data visualization notebook, as an external reference and peer-review purpose

EDA with SQL

- Performed SQL Queries such as:
  - Displaying the names of Unique launch sites
  - Displaying records of launch sites beiging with KSC
  - Displaying total mass carried by boosters from NASA (CRS)
  - Ranking landing outcomes, etc.
- GitHub URL: https://github.com/boluamosu/CapstoneProject/blob/main/jupyter-labs-eda-sql-edx_sqllite.ipynb

Results

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


Section 2
Insights drawn from EDA

All Launch Site Names

- Find the names of the unique launch sites

- There are 4 unique launch sites, namely: CCAFS LC-40, VAFB SLC-4E, KSC LC-39A, and CCAFS SLC-40

```
 * sqlite:///my_data1.db
Done.
```

Launch_Site
CCAFS LC-40
VAFB SLC-4E
KSC LC-39A
CCAFS SLC-40

Launch Site Names Begin with 'KSC'

- Find 5 records where launch sites' names start with `KSC`

Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer	Mission
2017-02-19	14:39:00	F9 FT B1031.1	KSC LC-39A	SpaceX CRS-10	2490	LEO (ISS)	NASA (CRS)	
2017-03-16	6:00:00	F9 FT B1030	KSC LC-39A	EchoStar 23	5600	GTO	EchoStar	
2017-03-30	22:27:00	F9 FT B1021.2	KSC LC-39A	SES-10	5300	GTO	SES	
2017-05-01	11:15:00	F9 FT B1032.1	KSC LC-39A	NROL-76	5300	LEO	NRO	
2017-05-15	23:21:00	F9 FT B1034	KSC LC-39A	Inmarsat-5 F4	6070	GTO	Inmarsat	

Total Payload Mass

- Calculate the total payload carried by boosters from NASA
- The total payload is 45,596

```
 * sqlite:///my_data1.db
Done.
```

sum(PAYLOAD_MASS__KG_)
45596

Average Payload Mass by F9 v1.1

- Calculate the average payload mass carried by booster version F9 v1.1
- The average payload mass for F9 v1.1 is 2928.4

```
 * sqlite:///my_data1.db
Done.
```

.1]:

avg(PAYLOAD_MASS__KG_)
2928.4

First Successful Ground Landing Date

The date of the first successful landing outcome on drone ship was in April 2016

Done.

min("Date")
2016-04-08

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

- There are 3 successful boosters

Done.

Booster_Version
F9 FT B1032.1
F9 B4 B1040.1
F9 B4 B1043.1

Total Number of Successful and Failure Mission Outcomes

- Calculate the total number of successful and failure mission outcomes
- There have been 99 successful and failed mission outcomes

```
 * sqlite:///my_data1.db
Done.
```

count(Mission_Outcome)
99

Boosters Carried Maximum Payload

- List the names of the booster which have carried the maximum payload mass
- There are 12 boosters that have carried the maximum payload mass

```
 * sqlite:///my_data1.db
Done.
```

5]:

Booster_Version
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

substr(Date,6,2)	substr(Date, 9, 2)	Booster_Version	Launch_Site	Mission_Outcome	Landing_Outcome
02	19	F9 FT B1031.1	KSC LC-39A	Success	Success (ground pad)
05	01	F9 FT B1032.1	KSC LC-39A	Success	Success (ground pad)
06	03	F9 FT B1035.1	KSC LC-39A	Success	Success (ground pad)
08	14	F9 B4 B1039.1	KSC LC-39A	Success	Success (ground pad)
09	07	F9 B4 B1040.1	KSC LC-39A	Success	Success (ground pad)
12	15	F9 FT B1035.2	CCAFS SLC-40	Success	Success (ground pad)

2017 Launch Records

- List the records which will display the month names, succesful landing_outcomes in ground pad ,booster versions, launch_site for the months in year 2017

- The months in 2027 that had successful ground pad landings were: February, May, June, August, September and December

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

- The most coming outcome in this timeframe was no attempt. With the rockets having the same number of drone ship success and failure.

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

Thank you!