Space Race with Falcon X

Phuc Thinh Nguyen

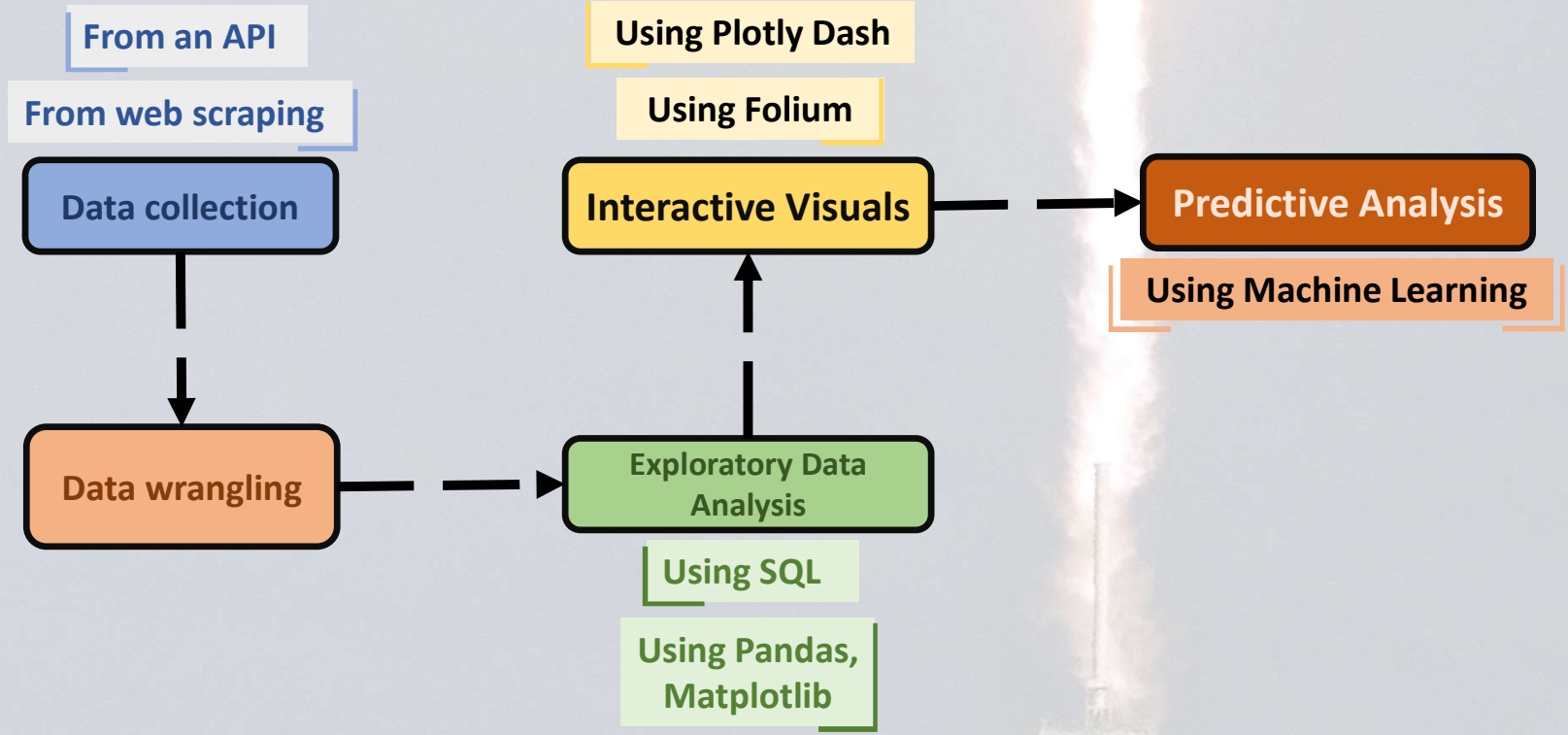
June 20th, 2024

Outline

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

Executive Summary

Summary of methodologies



Executive Summary

Summary of all results

Introduction

Project background and context

Predict if the Falcon 9 first stage will land successfully to determine the cost of a launch

Problems that need answers

Will payload mass and launch sites of the rocket have any relationships with its success rate?

Do the distances of the launch sites from its neighbors important?

What machine learning tools will we use to predict the success rate?

Methodology

Phuc Thinh Nguyen

June 20th, 2024

Methodology

Data collection using API

Get the URL of past SpaceX launches from SpaceX API

Using GET request

Using json.normalize()

Transform the data into a Python DataFrame

Filter the dataframe to needed columns

Using pre-defined helper functions

Get the needed data from the API to the filtered dataframe

Filter the dataframe to only need Falcon 9 launches

Replace missing values in "Payload Mass" data with the mean

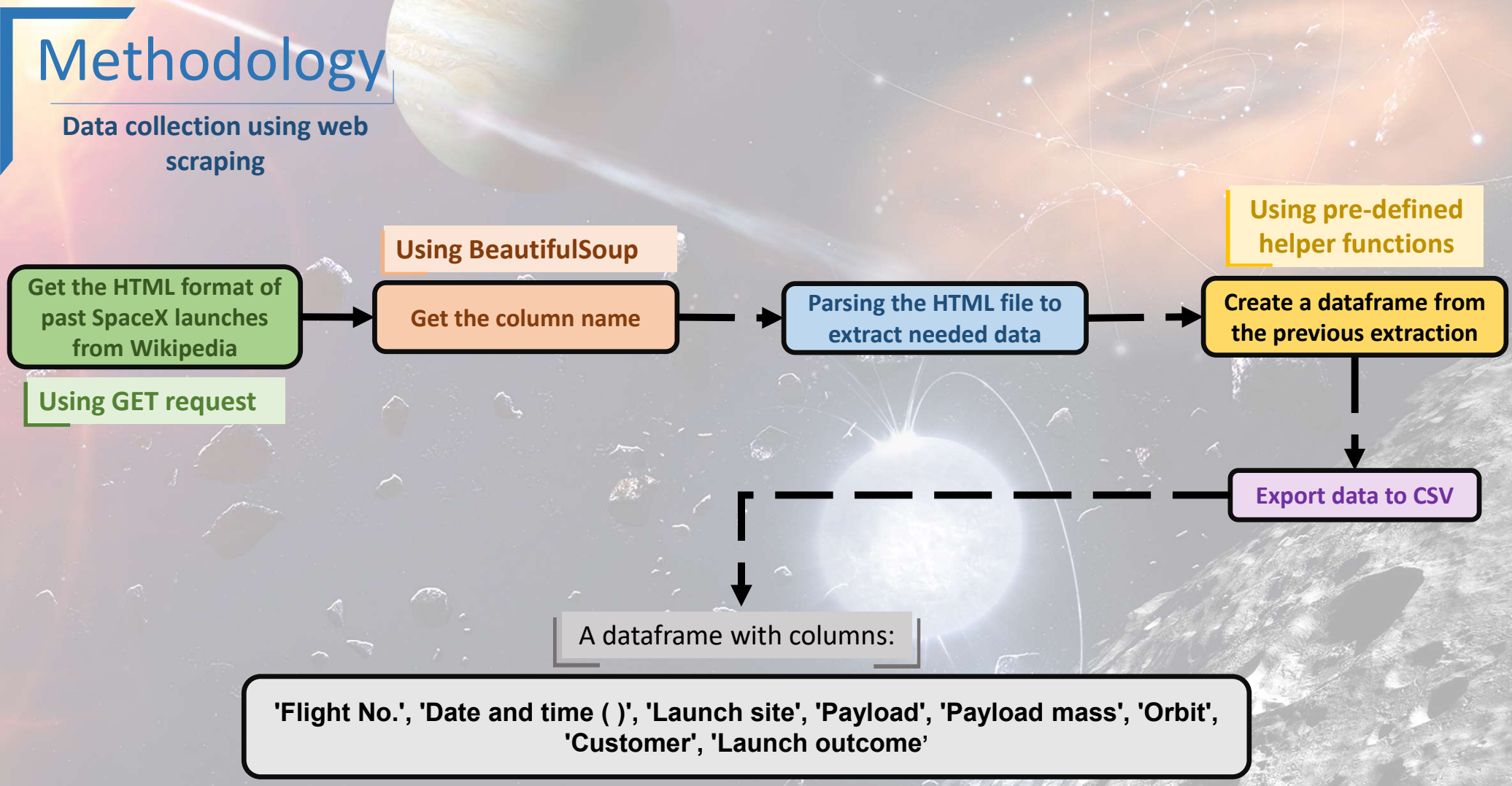
A dataframe with columns:

FlightNumber, Date, BoosterVersion, PayloadMass, Orbit, LaunchSite, Outcome, Flights, GridFins, Reused, Legs, LandingPad, Block, ReusedCount, Serial, Longitude, Latitude

GitHub: Data Collection API

Methodology

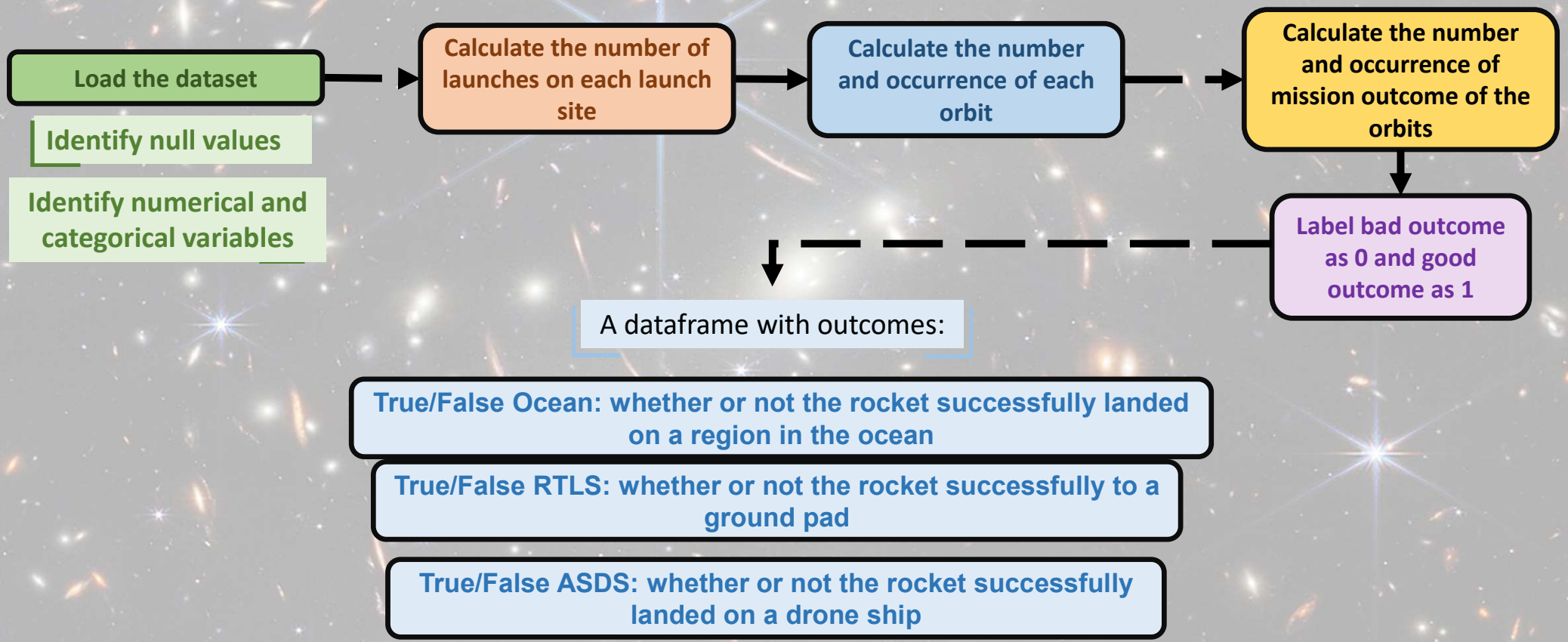
Data collection using web scraping



GitHub: Data Collection Web Scraping

Methodology

Data wrangling



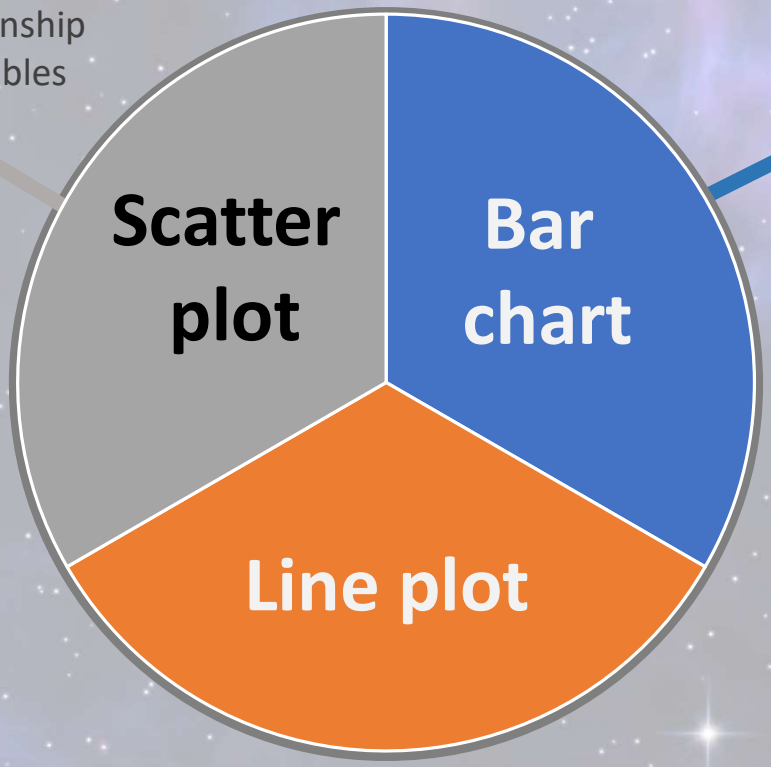
GitHub: Data Wrangling

Methodology

EDA with Visualizations

Visualize the relationship between two variables

Visualize the relationship between categorical variables through its frequency



Visualize the relationship of a single variable over time

GitHub: EDA with Data Visualizations

Methodology

EDA with SQL

- *Get unique launch sites*
- *Get records where launch sites begin with 'CCA'*
- *Get total payload mass carried by boosters launched by NASA (CRS)*
- *Get average payload mass carried by booster version F9 v1.1*
- *Get the first successful landing outcome date in ground pad*
- *Get 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 booster versions which have carried the maximum payload mass*
- *List the records with months, failure landing outcomes in drone ship, booster versions, launch site for the months in year 2015.*
- *Rank in descending order the count of landing outcomes between the date 2010-06-04 and 2017-03-20*

GitHub: EDA with SQL

Methodology

Interactive Map With Folium

GitHub: Interactive Map with Folium

*Marker: identify the launches with success rate (green means success, **red means failure**)*

Methodology

Dashboard with Plotly Dash

Pie plot: illustrate the total success launches for each sites

Drop-down: allows the user to choose the launch site they want information from

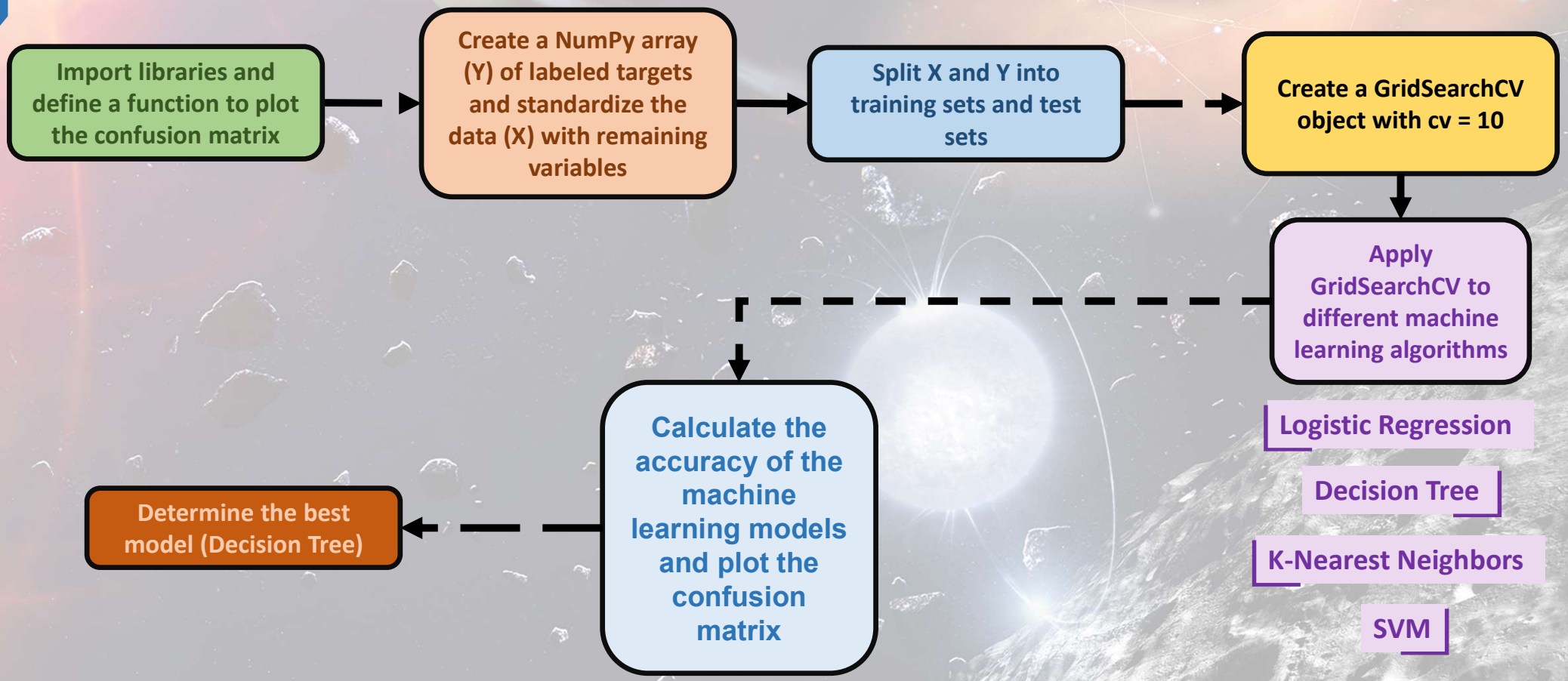
Slider: provides the information of the range of payload mass to be displayed

Scatterplot: gives information of the payload mass range of all launch sites, with its success status

GitHub: Plotly Dash Python Source Code

Methodology

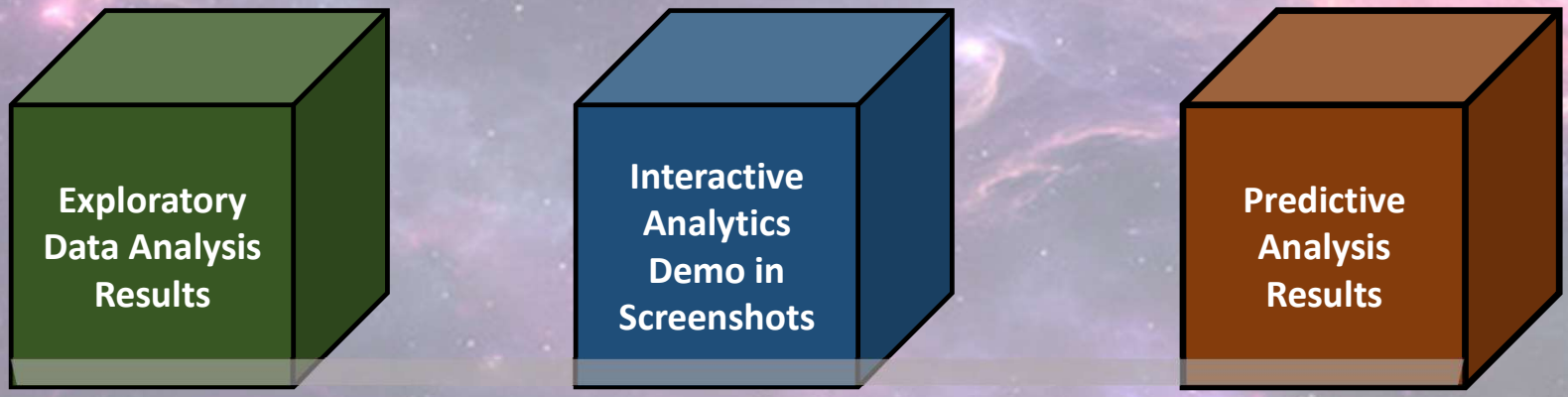
Predictive Analysis (Classification)



GitHub: Predictive Analysis (Classification)

Methodology

Results



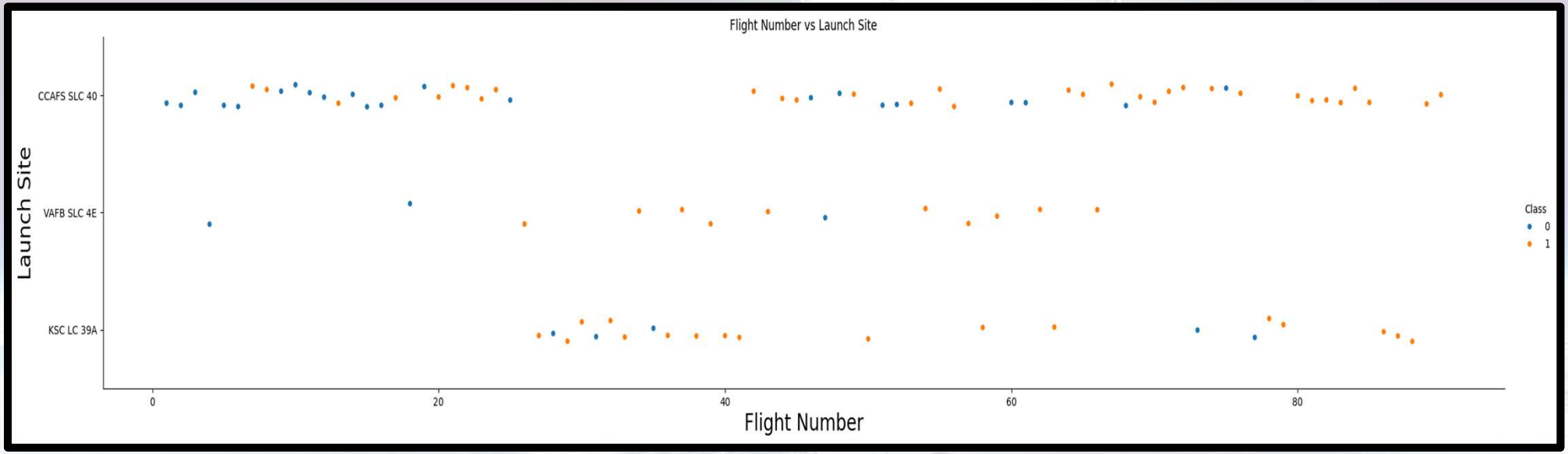
Insights Drawn From EDA

Phuc Thinh Nguyen

June 20th, 2024

Insights Drawn From EDA

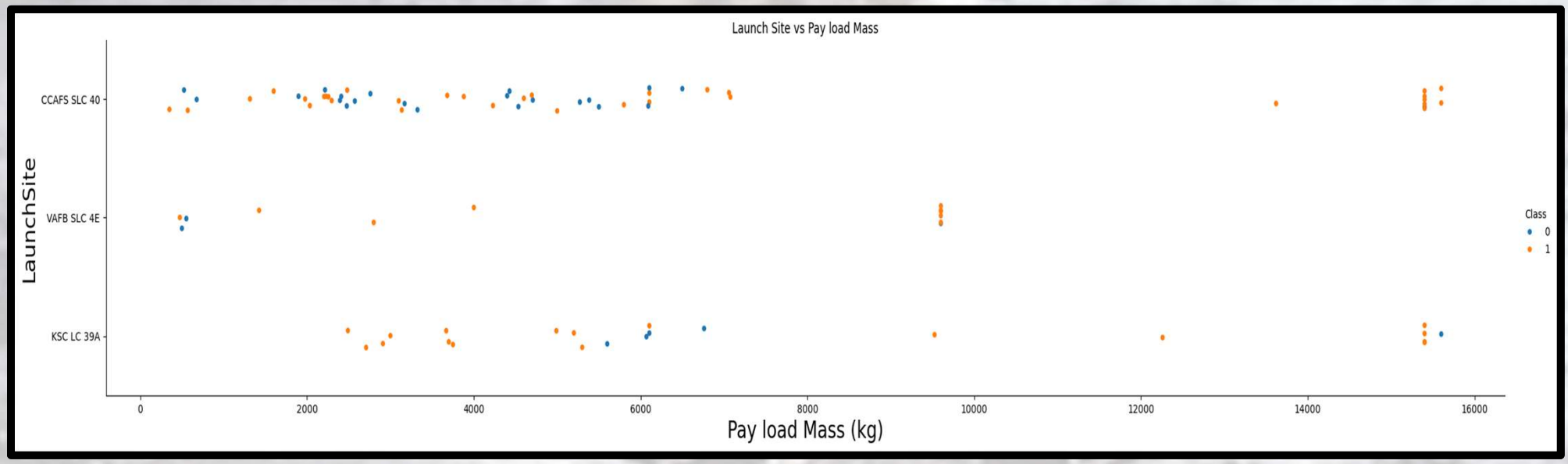
Flight Number vs. Launch Site



- ***Most of the flights are launched at CCAFS SLC 40.***
- ***Most of the first flights are unsuccessful (blue), while most of the latest flights are successful (orange)***
- ***VAFB SLC 4E and KSC LC 39A are launch sites with high success rate.***

Insights Drawn From EDA

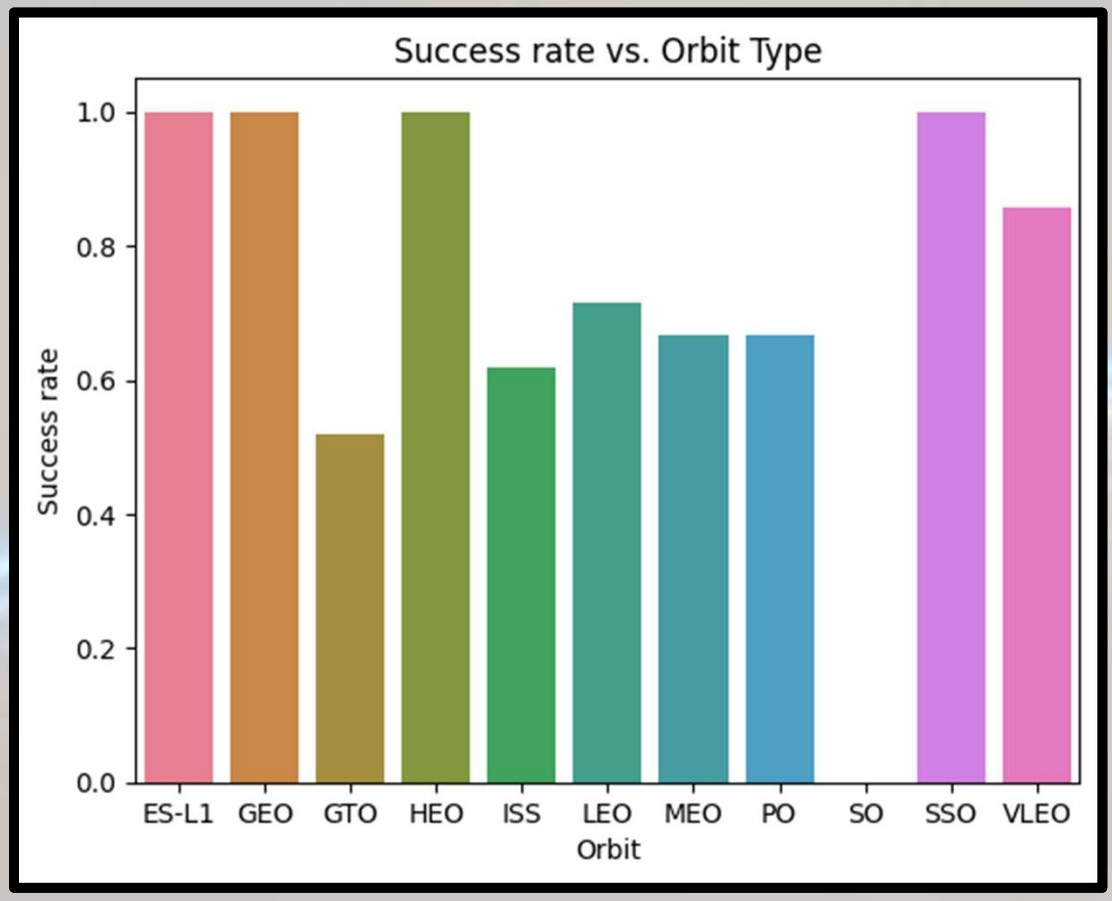
Payload Mass vs. Launch Site



- ***Most of the flights are launched at CCAFS SLC 40, with payload mass approximately ranging from 1000 – 7500kg.***
- ***Most successful launches have payload mass at least approximately 9750kg.***
- ***VAFB SLC 4E and KSC LC 39A are launch sites with high success rate.***

Insights Drawn From EDA

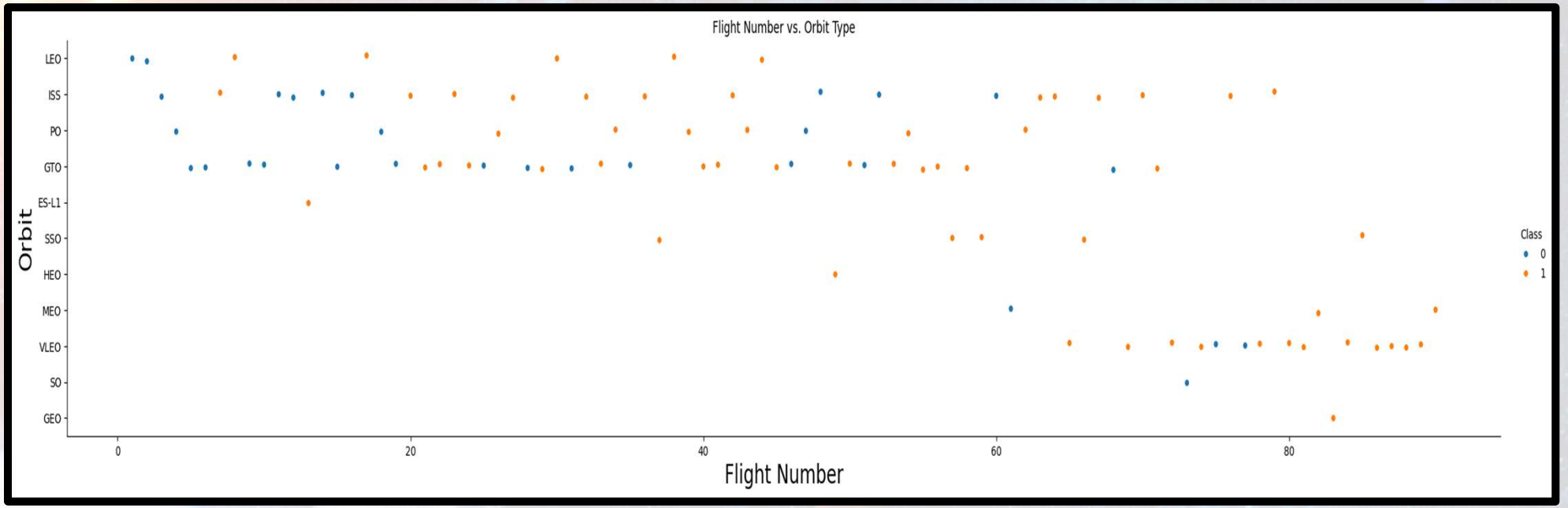
Success Rate vs. Orbit Type



- ***ES-L1, GEO, HEO, and SSO have 100% success rate.***
- ***SO has 0% chance of having successful launches***
- ***The remaining orbits have at least 50% chance of success.***

Insights Drawn From EDA

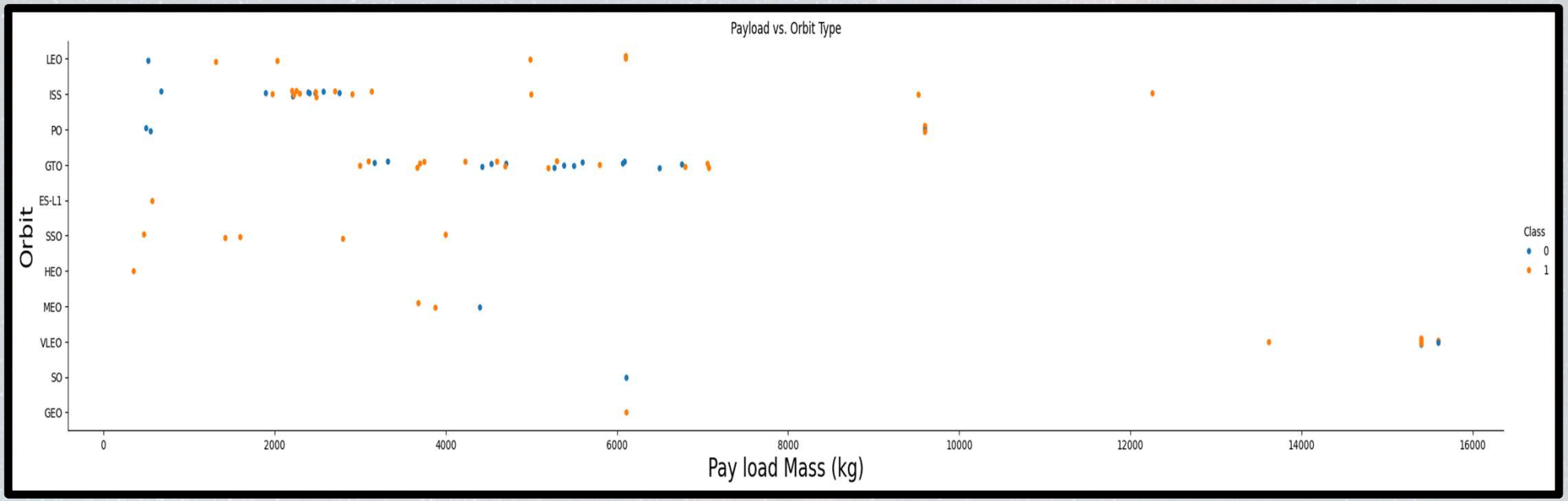
Flight Number vs. Orbit Type



- *The most recent launches had a successful outcome, mostly with VLEO orbit type.*
- *Most flights in LEO orbit are successful.*

Insights Drawn From EDA

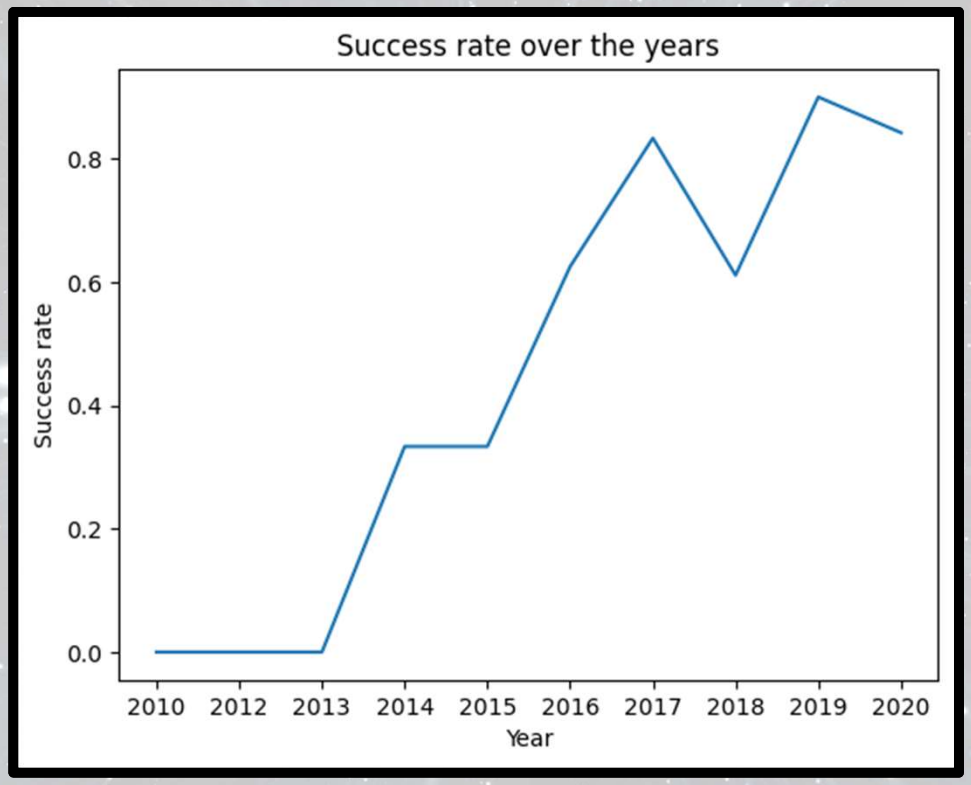
Payload vs. Orbit Type



- ***Heavy payload masses have more successful launches in LEO, PO, ISS, and VLEO orbits.***
- ***Low payload masses have more successful launches in ES-L1, SSO, and HEO.***

Insights Drawn From EDA

Launch Success Yearly Trend



- ***The success rate of the launches increases from 2013 until 2020, despite a decrease in 2018.***

Insights Drawn From EDA

All Launch Site Names

Unique Launch Sites
CCAFS LC-40
VAFB SLC-4E
KSC LC-39A
CCAFS SLC-40

- **CCAFS LC-40, VAFB SLC-4E, KSC LC-39A, and CCAFS SLC-40 are the names of the unique launch sites in the space mission.**

Insights Drawn From EDA

Launch Site Names Begin With ‘CCA’

Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer	Mission_Outcome	Landing_Outcome
2010-06-04	18:45:00	F9 v1.0 B0003	CCAFS LC-40	Dragon Spacecraft Qualification Unit	0	LEO	SpaceX	Success	Failure (parachute)
2010-12-08	15:43:00	F9 v1.0 B0004	CCAFS LC-40	Dragon demo flight C1, two CubeSats, barrel of Brouere cheese	0	LEO (ISS)	NASA (COTS) NRO	Success	Failure (parachute)
2012-05-22	7:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525	LEO (ISS)	NASA (COTS)	Success	No attempt
2012-10-08	0:35:00	F9 v1.0 B0006	CCAFS LC-40	SpaceX CRS-1	500	LEO (ISS)	NASA (CRS)	Success	No attempt
2013-03-01	15:10:00	F9 v1.0 B0007	CCAFS LC-40	SpaceX CRS-2	677	LEO (ISS)	NASA (CRS)	Success	No attempt

- **Above are 5 records where launch sites begin with the string 'CCA’**

Insights Drawn From EDA

Total Payload Mass

Total Payload Mass (kg)
45596

- The total payload mass carried by boosters launched by NASA (CRS) is 45596kg

Insights Drawn From EDA

Average Payload Mass by F9 v1.1

Average Payload Mass (kg)
2928.4

- The average payload mass carried by booster version F9 v1.1 is 2928.4kg

Insights Drawn From EDA

First Successful Ground Landing Date

First Successful Landing Date
2015-12-22

- **The date when the first successful landing outcome in ground pad was achieved is November 22nd, 2015.**

Insights Drawn From EDA

Successful Drone Ship Landing with Payload between 4000 and 6000

Booster_Version
F9 FT B1022
F9 FT B1026
F9 FT B1021.2
F9 FT B1031.2

- **F9 FT B1022, F9 FT B1026, F9 FT B1021.2, F9 FT B1031.2 are the names of the boosters which have success in drone ship and have payload mass greater than 4000 but less than 6000.**

Insights Drawn From EDA

Total Number of Successful and Failure Mission Outcomes

Successful Missions	Failure Missions
100	1

- There are a total of 100 successful mission outcomes and 1 failure mission outcomes.

Insights Drawn From EDA

Boosters Carried Maximum Payload

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

- **On the left are the names of the booster versions which have carried the maximum payload mass.**

Insights Drawn From EDA

2015 Launch Records

Month	Landing_Outcome	Booster_Version	Launch_Site
01	Failure (drone ship)	F9 v1.1 B1012	CCAFS LC-40
04	Failure (drone ship)	F9 v1.1 B1015	CCAFS LC-40

In 2015:

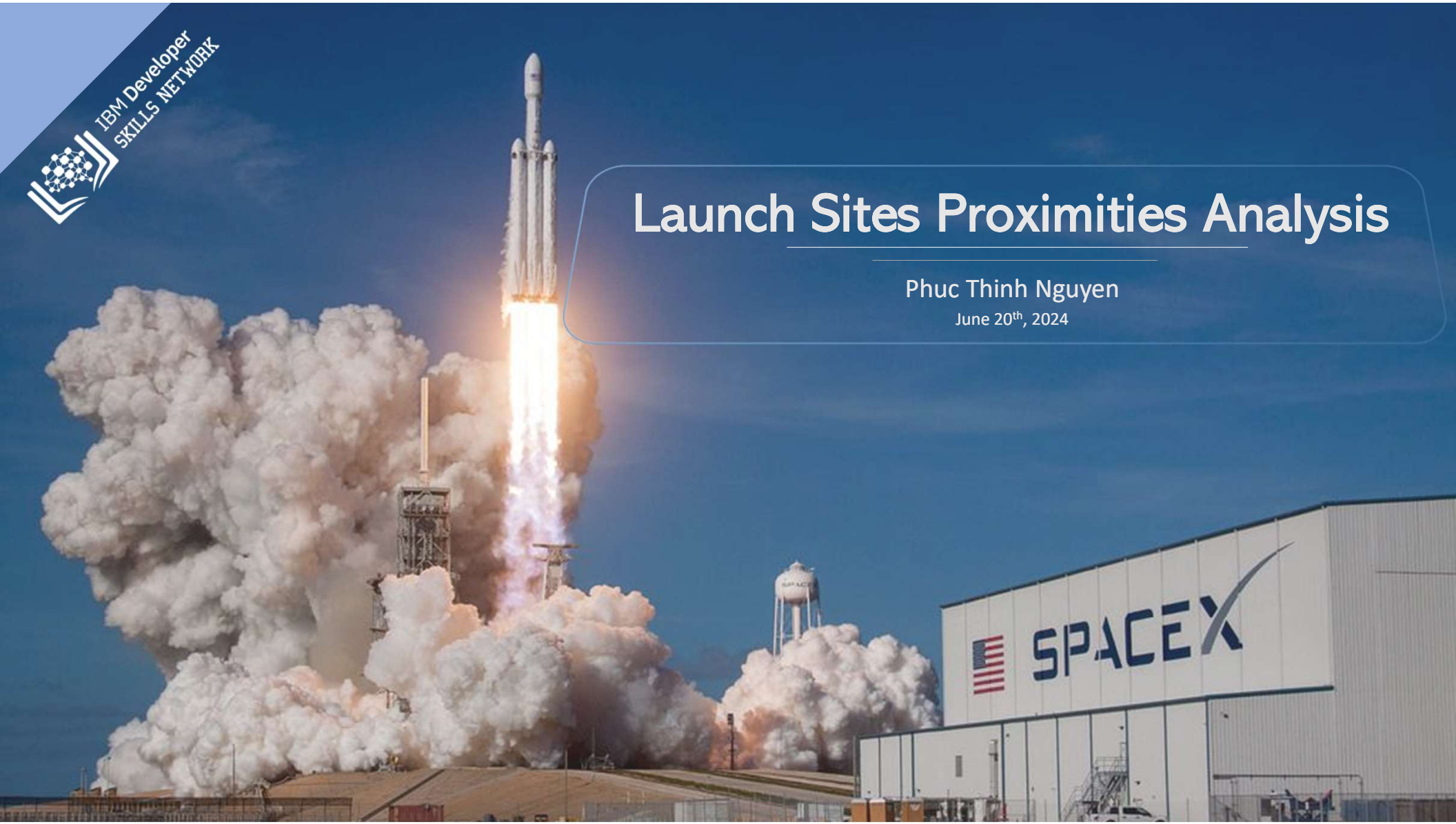
- **In January, a rocket launched from CCAFS LC-40 with booster version F9 v1.1 B1012 failed to land on a drone ship.**
- **In April, a rocket launched from CCAFS LC-40 with booster version F9 v1.1 B1015 failed to land on a drone ship.**

Insights Drawn From EDA

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

Landing Outcome	Number of Outcomes
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

- **On the left are the count of landing outcomes between April 6, 2010 and March 20th, 2017 in descending order. For example, there are 5 rocket launches that successfully landed on a drone ship.**



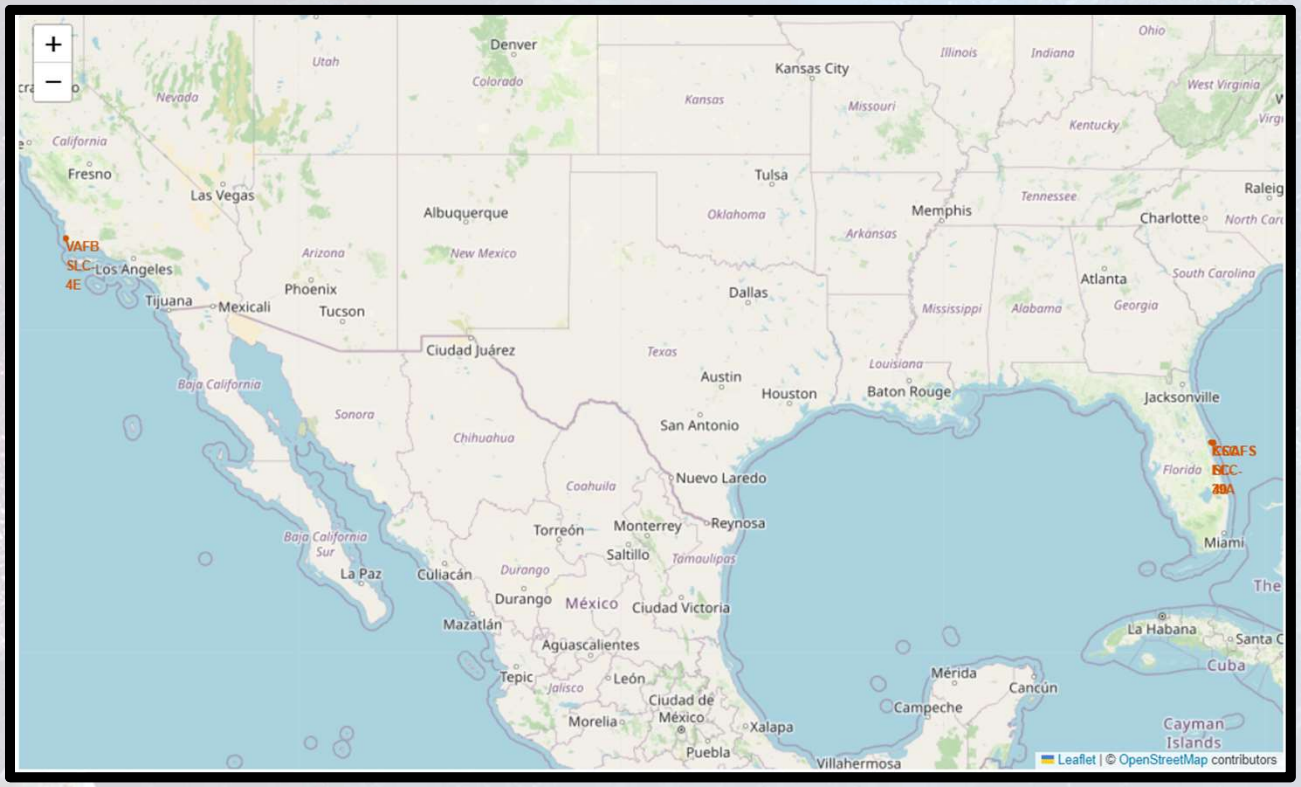
Launch Sites Proximities Analysis

Phuc Thinh Nguyen

June 20th, 2024

Insights Drawn From EDA

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

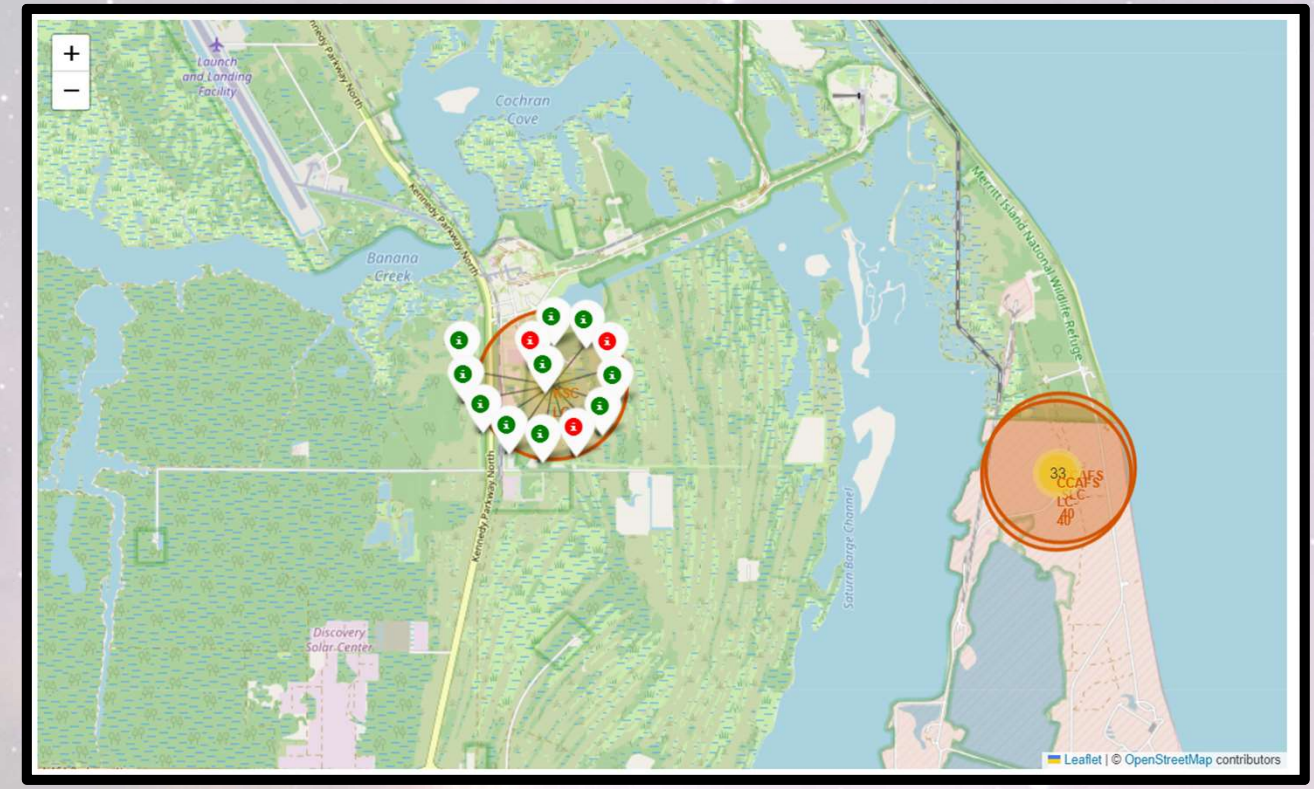


- ***All of the launches are in close proximity to the equator (0-degree orbit) so that less maneuver is needed, saving fuel and manufacturing cost.***

- ***All of the launches are in very close proximity to the coast so that it stays away from the cities, highways, and railroads.***

Insights Drawn From EDA

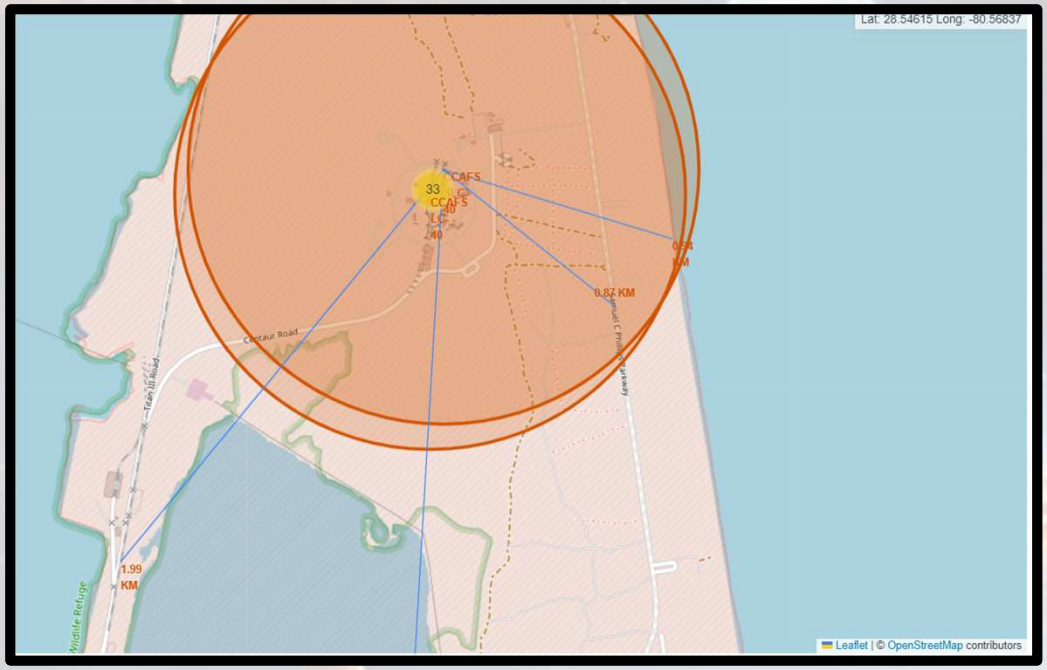
Folium Map of Color-Labeled Launch Outcomes

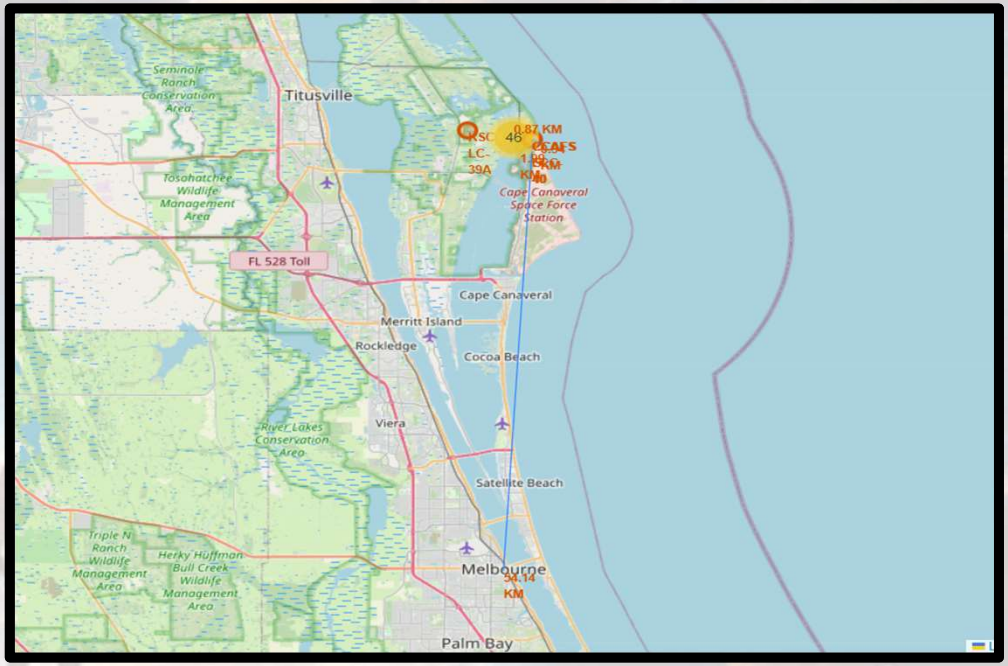


- *The green markers shows successful launches, while red markers indicate failure outcomes.*
- *The map indicates that KSC-LC 39A has a high success rate.*

Insights Drawn From EDA

Folium Map of Distances From A Launch Site





- *This map shows that CCAFS-LC 40 is:*

(+) 0.94km away from the closest coastline

(+) 0.87km away from the closest road

(+) 1.99km away from the closest railroad

- *This map shows that CCAFS-LC 40 is 54.14km from Melbourne, one of the closest cities, indicating that a failure outcome at the launch site would negatively affect nearby cities.*

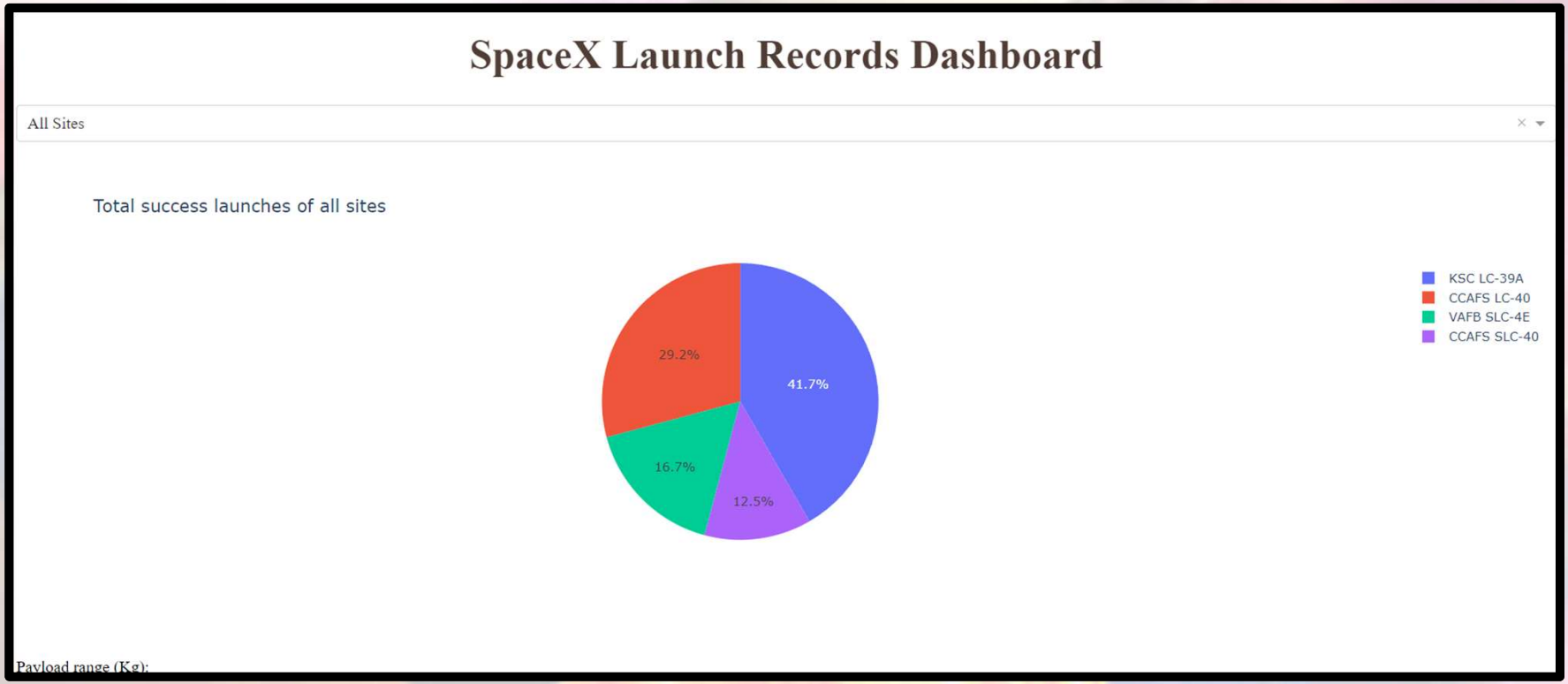
Dashboard with Plotly Dash

Phuc Thinh Nguyen

June 20th, 2024

Insights Drawn From EDA

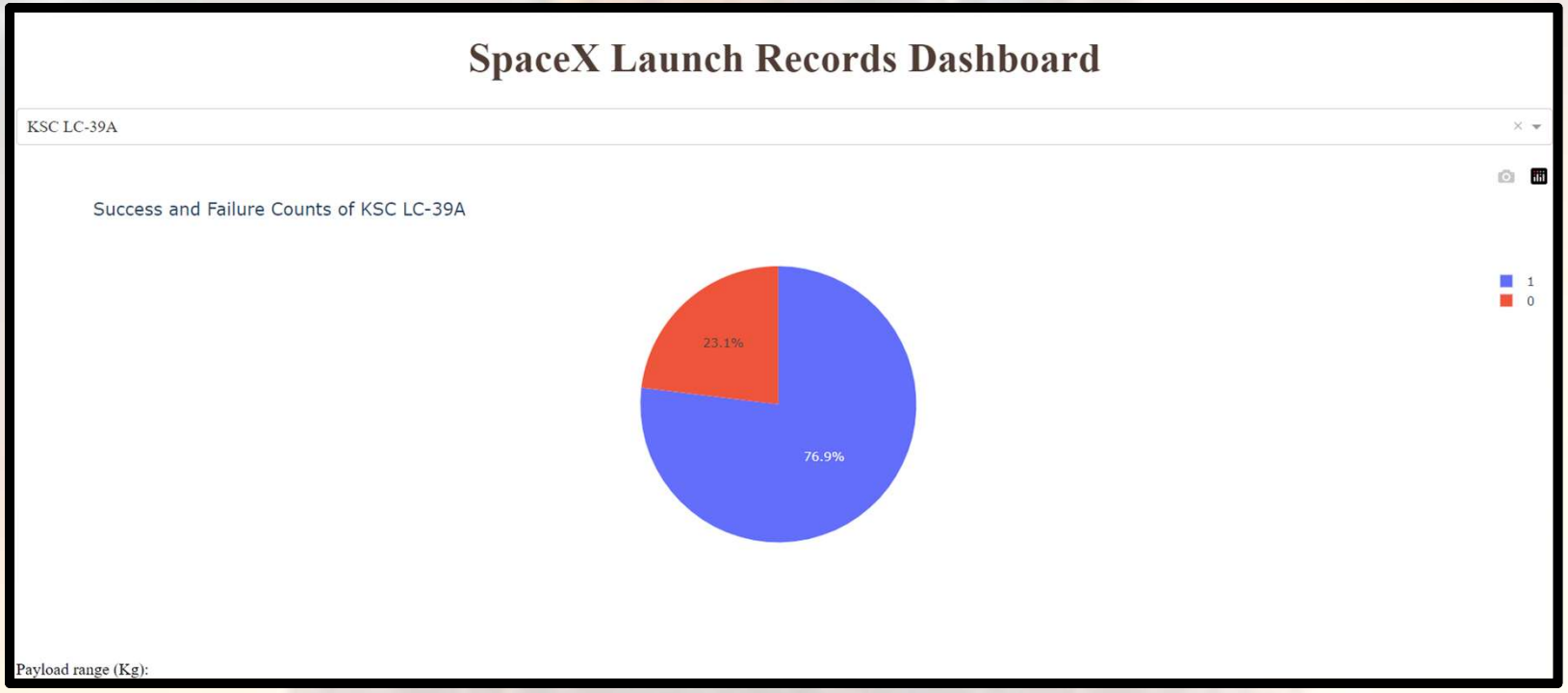
Plotly Pie Chart of Total Success Launches of All Sites



- *The pie chart indicates that KSC LC-39A has the highest success launches rate (41.7%), followed by CCAFS LC-40 (29.2%), VAFB SLC-4E (16.7%), and CCAFS SLC-40 (12.5%)*

Insights Drawn From EDA

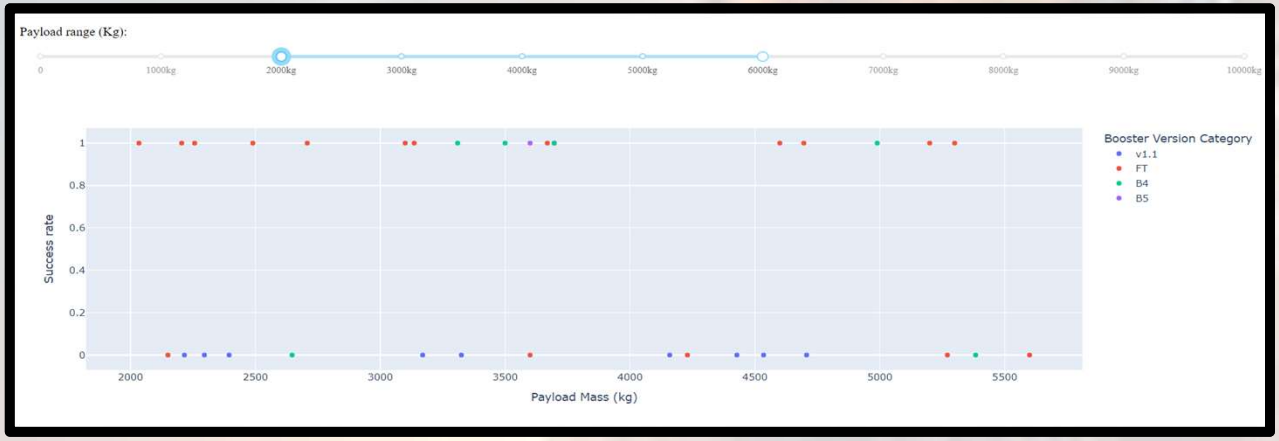
Plotly Pie Chart For The Launch Site with Highest Success Rate

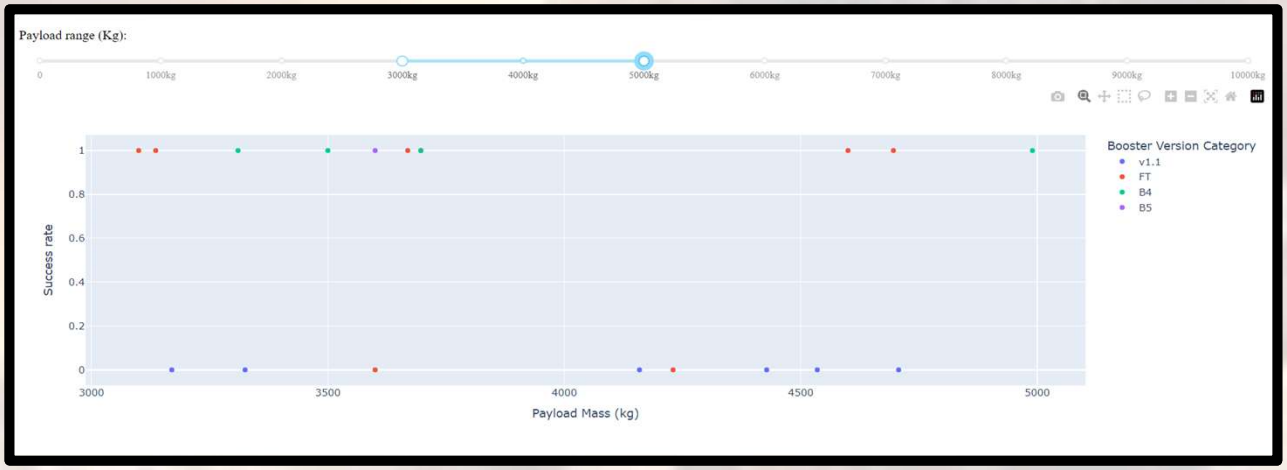


- ***The pie chart indicates that KSC LC-39A has the highest success rate (76.9%).***

Insights Drawn From EDA

Plotly Scatter Plot For Payload Mass vs. Success Rate





- ***The scatter plot shows that payload mass ranging between 2000kg – 6000kg has the highest success rate.***

- ***The scatter plot also shows that booster version FT has the highest success rate.***

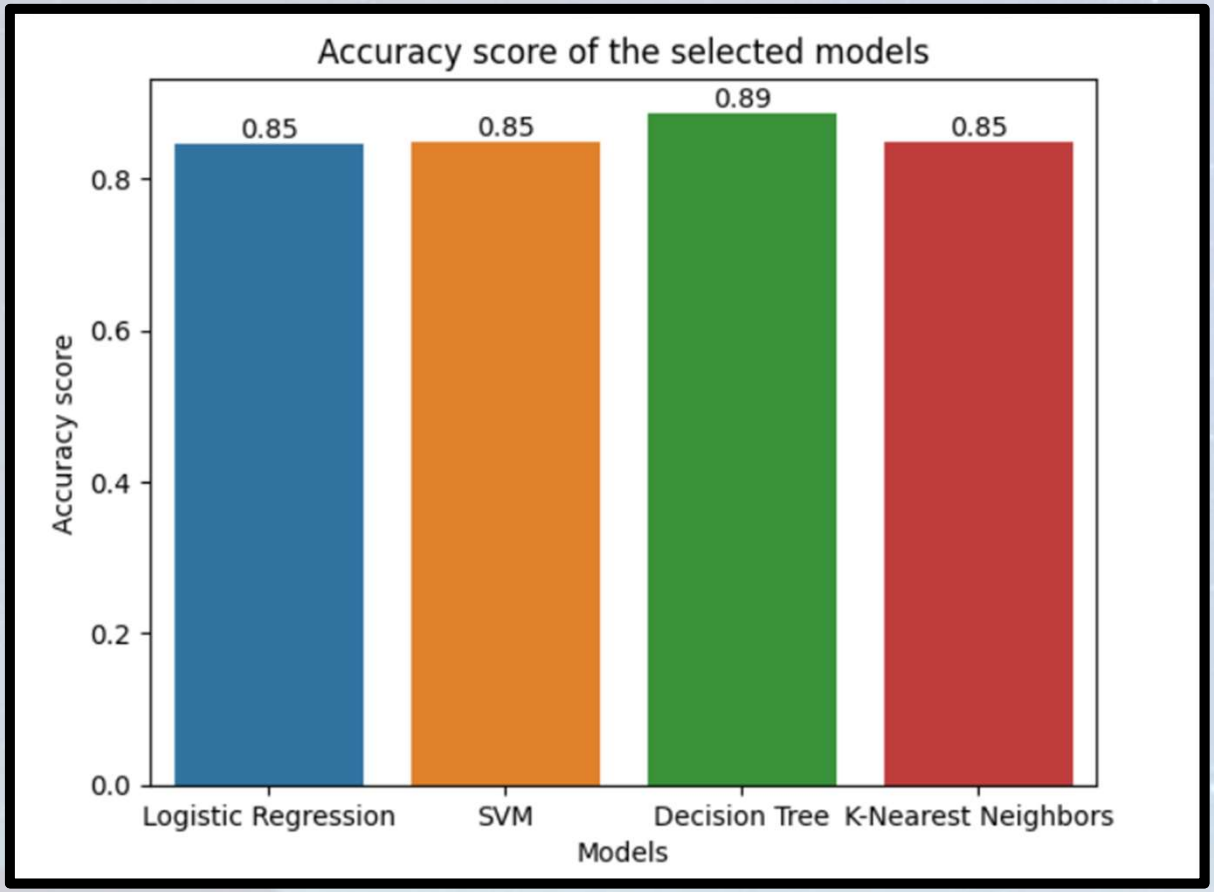
Predictive Analysis (Classification)

Phuc Thinh Nguyen

June 20th, 2024

Predictive Analysis

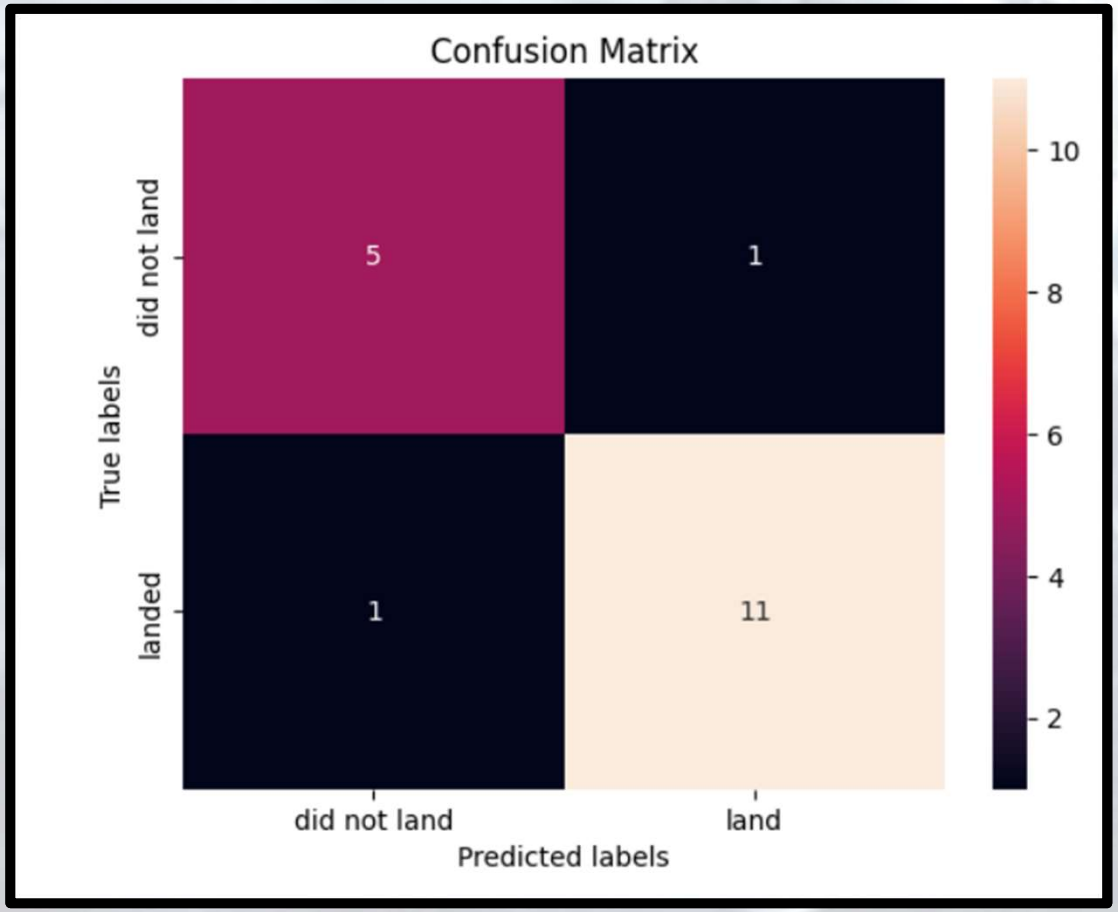
Classification Accuracy



- ***The bar chart clearly indicates that decision tree is the best model to predict, with accuracy score of 0.89***

Predictive Analysis

Confusion Matrix



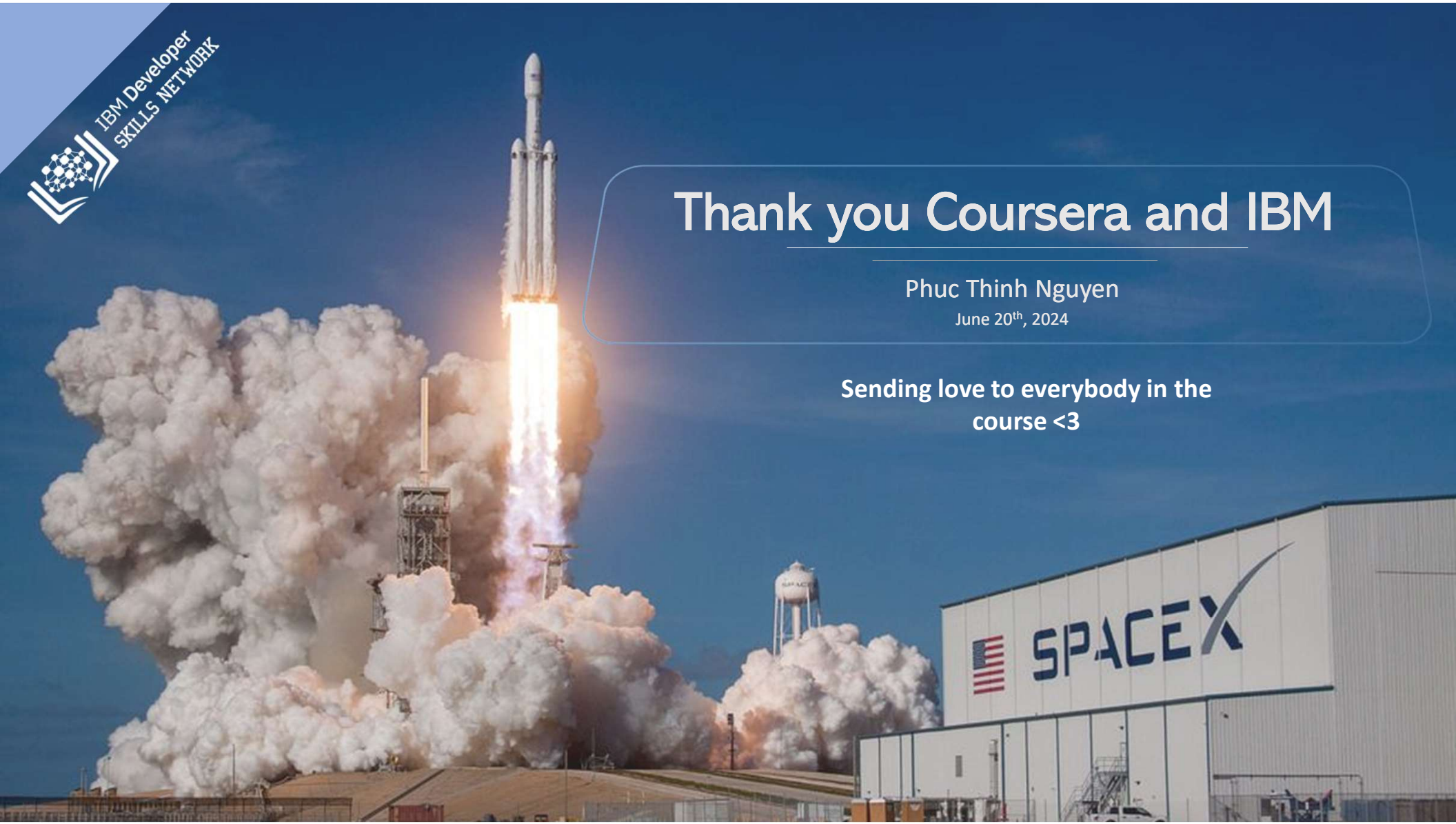
- *The confusion matrix correctly predicts 5 launches that did not land.*
- *However, it misclassified 1 launches that did not land as the launch that lands.*
- *The confusion matrix correctly predicts 11 launches that land.*
- *However, it misclassified 1 launches that lands as the launch that did not land.*

=> It can correctly predict $\frac{5}{5+1}$ *= 83% of launches that did not land, and* $\frac{11}{12}$ *= 91.6% of launches that lands.*

Predictive Analysis

Conclusion

- KSC LC-39A is the recommend launch site for the next launch, since it has the highest success rate.
- Low payload masses with ES-L1, SSO, and HEO orbits will have high success rate.
- The success rate of the launches increases from 2013 until 2020.
- All of the launches are in close proximity to the equator and extremely close proximity to the coastline.
- Decision Tree is the best machine learning algorithm in this classification scenario.



Thank you Coursera and IBM

Phuc Thinh Nguyen

June 20th, 2024

Sending love to everybody in the course <3