IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Arreola Ramirez Edwin Daniel
Jan/14/25

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

- 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

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

- Linecharts, scatteplots and barcharts were made. In order to easily visualize the existing relationships between variables such as: number of flights and orbits, payload and orbit, time success rate, flights and payload, etc.

- ProjectM10/edadataviz.ipynb at main · daniuncoffee/ProjectM10

EDA with SQL

- Display records of the launchs, as names, launch sites, payload mass, booster version, etc.
- Display the firsts successful landing achieved
- Display the total number of successful and failure mission and information about them
- Display name of the boosters which have success in drone ship with a payload mass grater than 4000 and less than 6000.

- ProjectM10/jupyter-labs-eda-sql-coursera_sqllite.ipynb at main · daniuncoffee/ProjectM10

Build an Interactive Map with Folium

- Markers were created to identify the location of mission launch sites
- These markers were grouped into clusters (circles) for better visualization of the locations
- Added lines representing the distance from the launch site to important locations (such as train tracks, cities, and highway)

- ProjectM10/lab_jupyter_launch_site_location.ipynb at main · daniuncoffee/ProjectM10

Build a Dashboard with Plotly Dash

- To easily visualize the data, graphs such as piecharts of successful launches, and payload and successful launch plots were added to a dashboard.
- In addition, you can choose to display this information on all launch sites or for a specific site

- ProjectM10/Dashboard.txt at main · daniuncoffee/ProjectM10

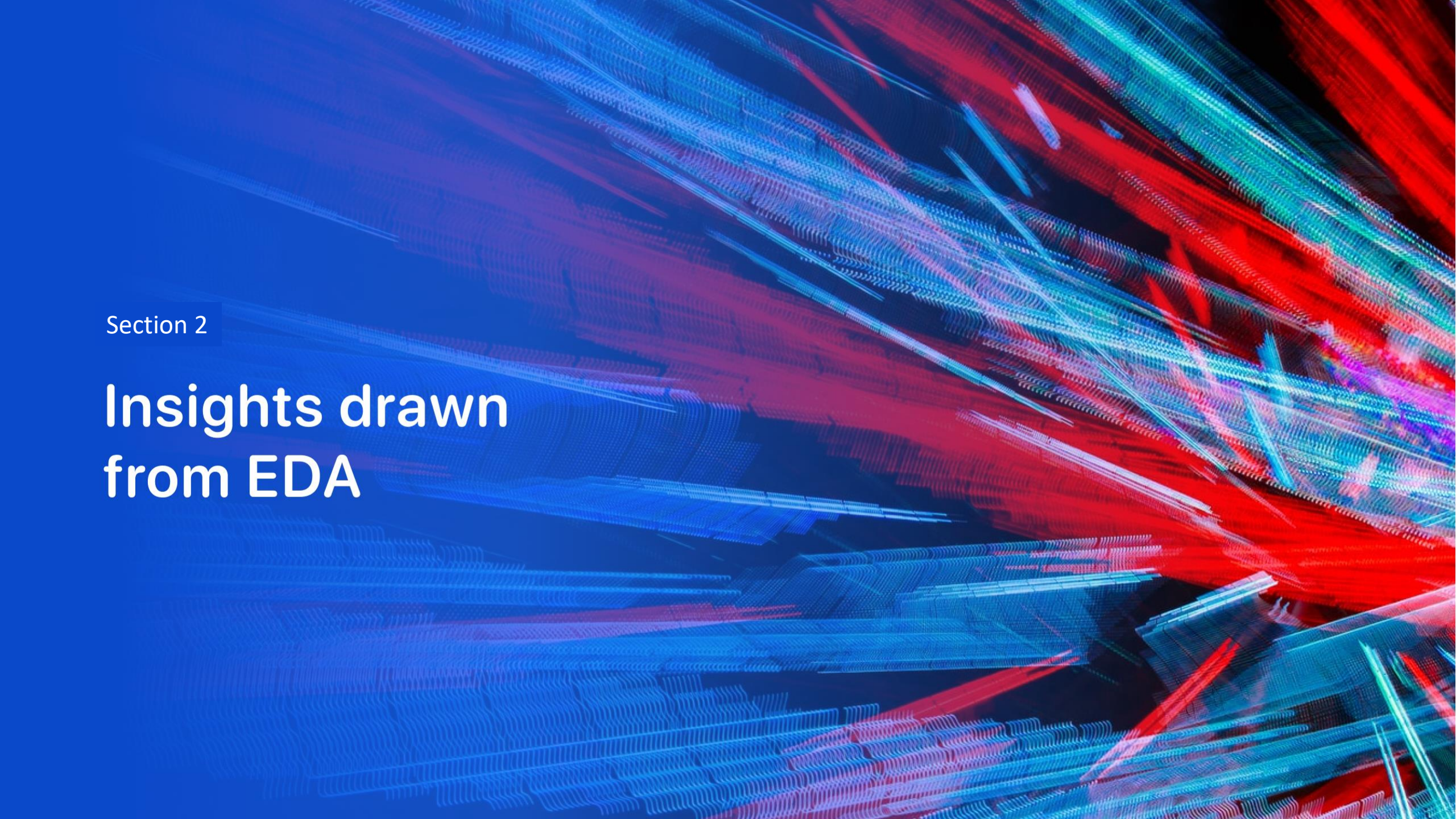
Predictive Analysis (Classification)

- With the 'test_train_test' function in Python, we obtained samples to train and test the models
- With the help of GridSearchCV, it was determined between several models such as: tree model, SVM, Regression, etc. The most accurate models.

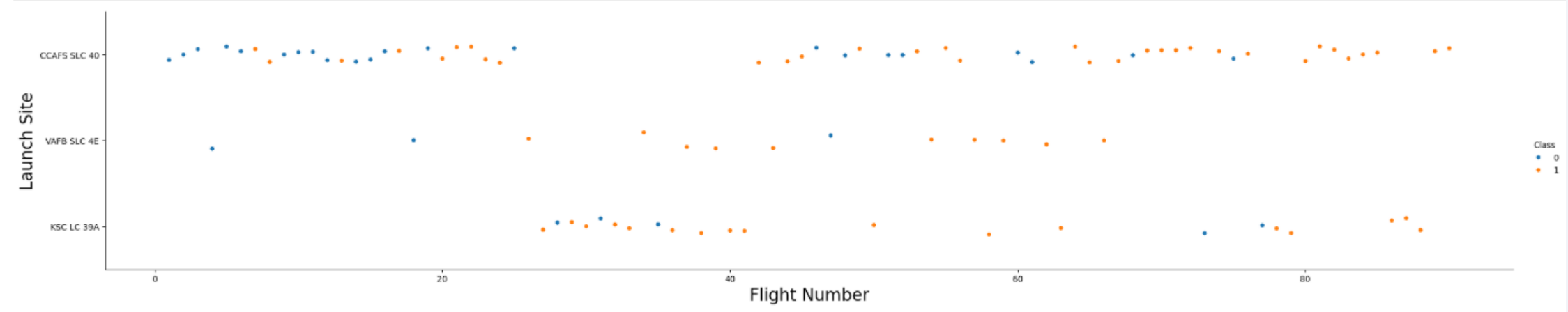
- ProjectM10/SpaceX_Machine Learning Prediction_Part_5.ipynb at main · daniuncoffee/ProjectM10

Results

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

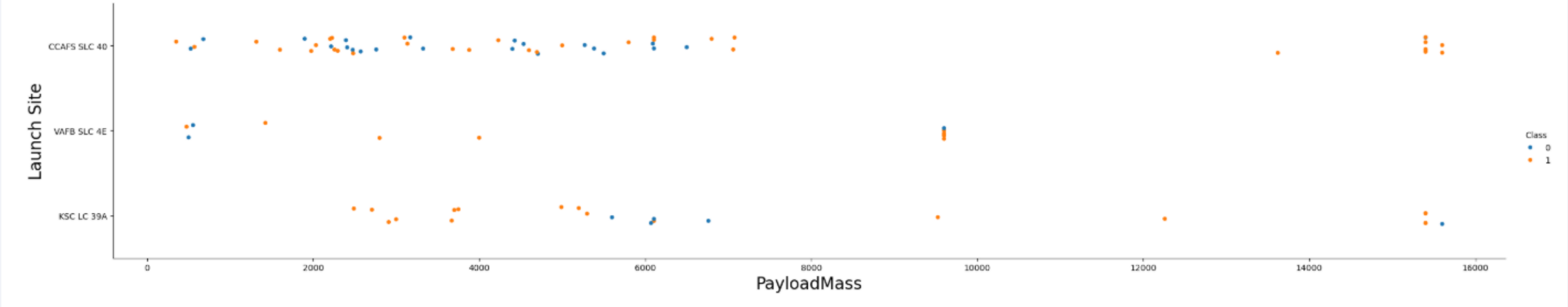

Section 2
Insights drawn from EDA

Flight Number vs. Launch Site



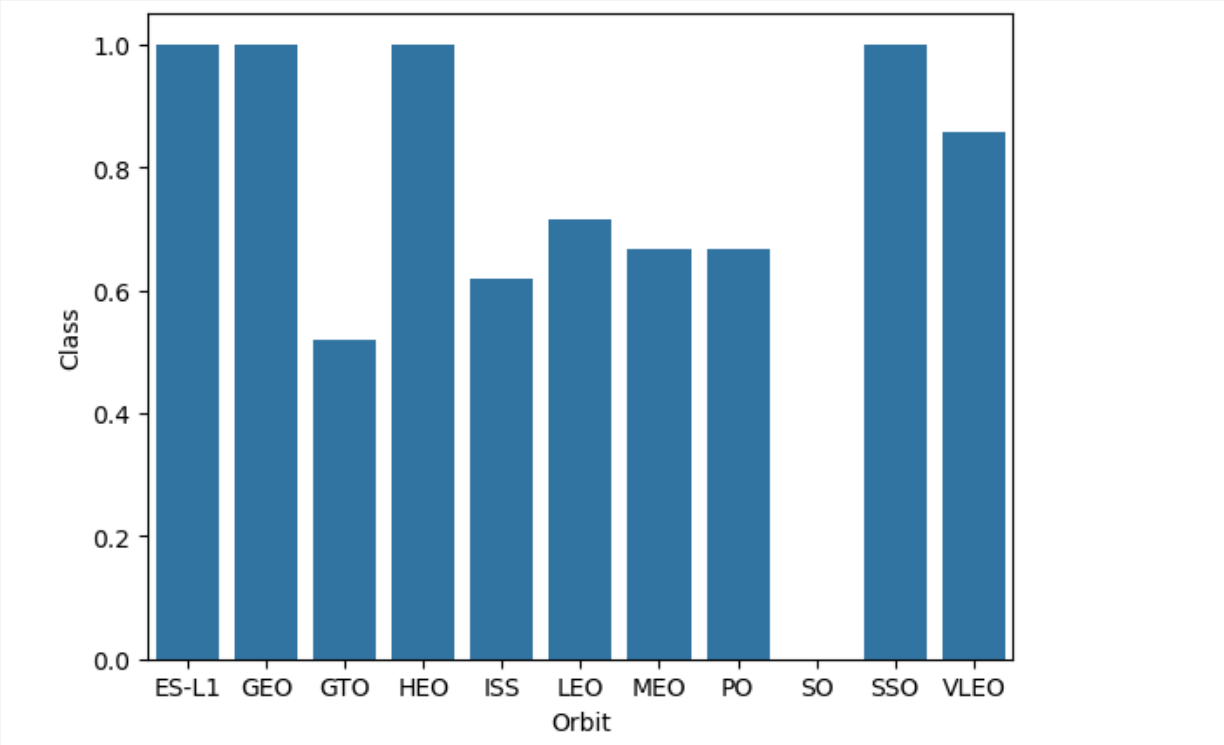
We can see that the majority of the Flights, are launch in the site ‘CCAFS SLC-40’

Payload vs. Launch Site



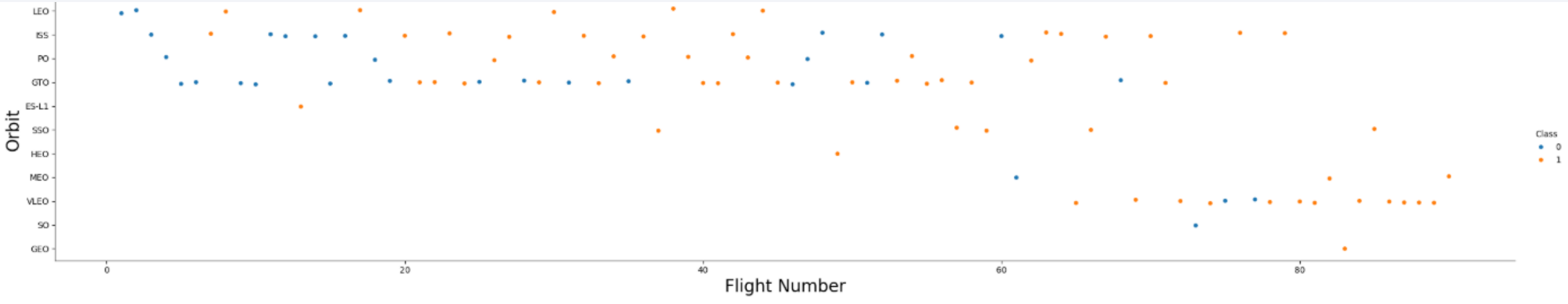
We can see that, in the site ‘KSC’, launch with low payload mass (<6000kg) are successful. While in site ‘CCAFS’, the launchs with maximum payload are successful

Success Rate vs. Orbit Type



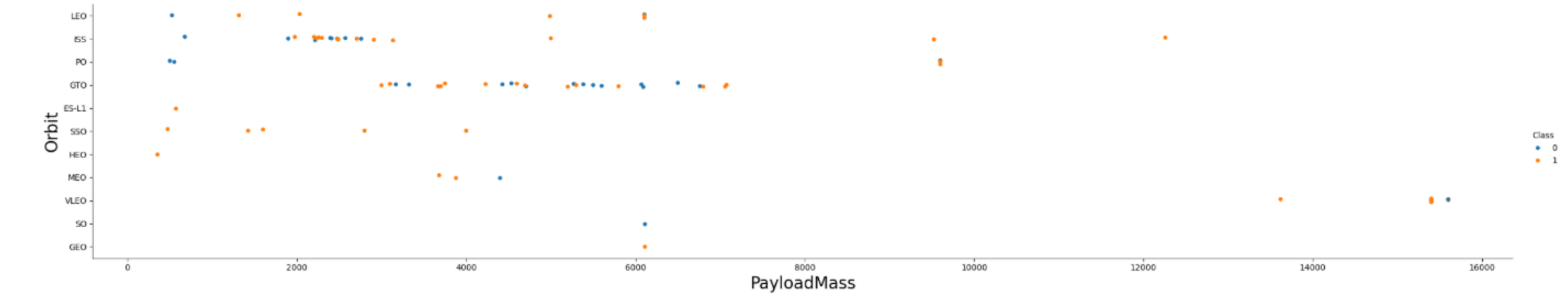
The orbits ES-L1, GEO, HEO, SSO and VLEO have a high successful rate, GTO, ISS, LEO, EO and PO a medium rate and last, SO is zero.

Flight Number vs. Orbit Type



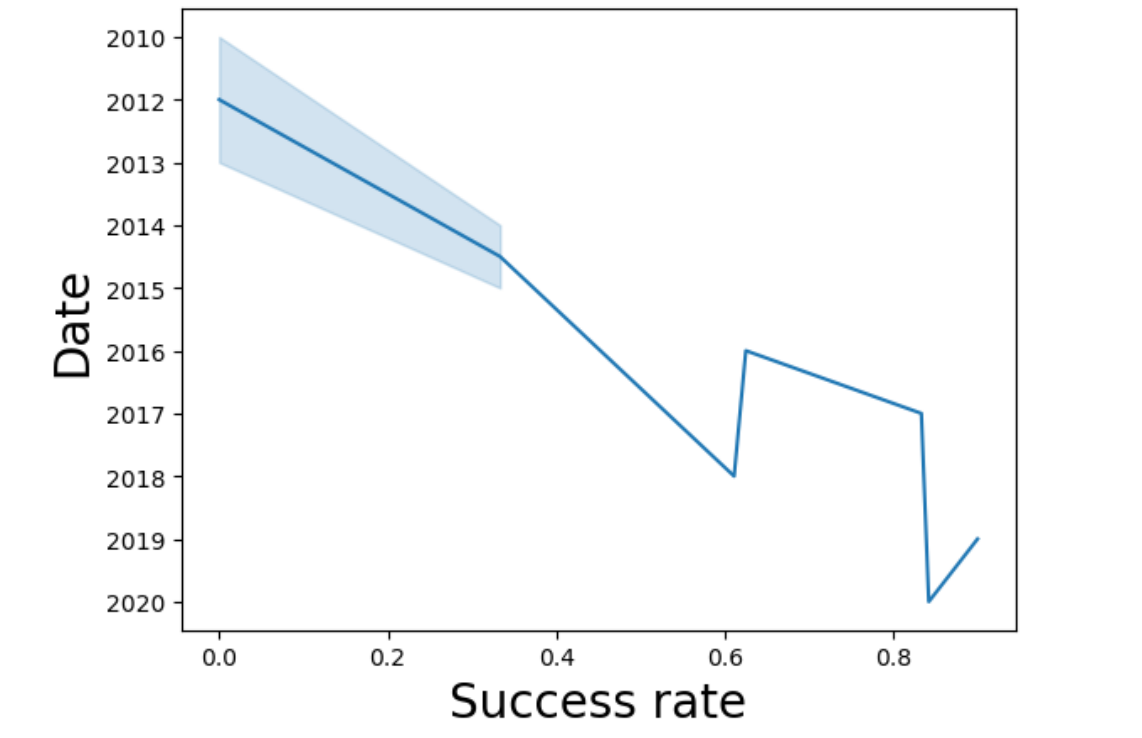
Initially, the orbits in the flight were LEO, ISS, PO and GTO, in majority, in recently flight is VLEO

Payload vs. Orbit Type



The flight with payload between 2000-3000 are in majority in the orbit ISS, and the range of payload between 3000-7000 are in majority in the orbit GTO

Launch Success Yearly Trend



The success rate has been increasing over the years

All Launch Site Names

The launch site are:

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

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

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

The table display records with Launch Location start with ‘CCA’

Total Payload Mass

Total_payload_mass_kg
45596

The average payload is 45596 kg

Average Payload Mass by F9 v1.1

avg(PAYLOAD_MASS__KG_)
2534.666666666665

The F9 v1.1 can charge an average payload of 2534.66kg

First Successful Ground Landing Date

Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer	Mission_Outcome	Landing_Outcome
2015-12-22	1:29:00	F9 FT B1019	CCAFS LC-40	OG2 Mission 2 11 Orbcomm-OG2 satellites	2034	LEO	Orbcomm	Success	Success (ground pad)
2016-04-08	20:43:00	F9 FT B1021.1	CCAFS LC-40	SpaceX CRS-8	3136	LEO (ISS)	NASA (CRS)	Success	Success (drone ship)
2016-05-06	5:21:00	F9 FT B1022	CCAFS LC-40	JCSAT-14	4696	GTO	SKY Perfect JSAT Group	Success	Success (drone ship)
2016-05-27	21:39:00	F9 FT B1023.1	CCAFS LC-40	Thaicom 8	3100	GTO	Thaicom	Success	Success (drone ship)
2016-07-18	4:45:00	F9 FT B1025.1	CCAFS LC-40	SpaceX CRS-9	2257	LEO (ISS)	NASA (CRS)	Success	Success (ground pad)

The first Success landing was in 2015-12-22, with booster version F9 FT, in Orbit Leo and a payload mass of 2034 kg, in the launch site CCAFS LC-40

Successful Drone Ship Landing with Payload between 4000 and 6000

Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer	Mission_Outcome	Landing_Outcome
2016-05-06	5:21:00	F9 FT B1022	CCAFS LC-40	JCSAT-14	4696	GTO	SKY Perfect JSAT Group	Success	Success (drone ship)
2016-08-14	5:26:00	F9 FT B1026	CCAFS LC-40	JCSAT-16	4600	GTO	SKY Perfect JSAT Group	Success	Success (drone ship)
2017-03-30	22:27:00	F9 FT B1021.2	KSC LC-39A	SES-10	5300	GTO	SES	Success	Success (drone ship)
2017-10-11	22:53:00	F9 FT B1031.2	KSC LC-39A	SES-11 / EchoStar 105	5200	GTO	SES EchoStar	Success	Success (drone ship)

4 success drone ship landing with a middle payload mass

Total Number of Successful and Failure Mission Outcomes

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

Allmost all the mission outcomes were success

Boosters Carried Maximum Payload

Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer	Mission_Outcome	Landing_Outcome
2019-11-11	14:56:00	F9 B5 B1048.4	CCAFS SLC-40	Starlink 1 v1.0, SpaceX CRS-19	15600	LEO	SpaceX	Success	Success
2020-01-07	2:33:00	F9 B5 B1049.4	CCAFS SLC-40	Starlink 2 v1.0, Crew Dragon in-flight abort test	15600	LEO	SpaceX	Success	Success
2020-01-29	14:07:00	F9 B5 B1051.3	CCAFS SLC-40	Starlink 3 v1.0, Starlink 4 v1.0	15600	LEO	SpaceX	Success	Success
2020-02-17	15:05:00	F9 B5 B1056.4	CCAFS SLC-40	Starlink 4 v1.0, SpaceX CRS-20	15600	LEO	SpaceX	Success	Failure
2020-03-18	12:16:00	F9 B5 B1048.5	KSC LC-39A	Starlink 5 v1.0, Starlink 6 v1.0	15600	LEO	SpaceX	Success	Failure

The maximum payload were in the orbit LEO, with the booster version F9 B5

2015 Launch Records

Date	Landing_Outcome	Booster_Version	Launch_Site
2015-01-10	Failure (drone ship)	F9 v1.1 B1012	CCAFS LC-40
2015-04-14	Failure (drone ship)	F9 v1.1 B1015	CCAFS LC-40

Launch records for landing in drone ship for 2015-2016

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

count(Landing_Outcome)	Landing_Outcome
31	Failure (parachute)

31 landing for the period 2010-06-04 and 2017-03-20, were Failure

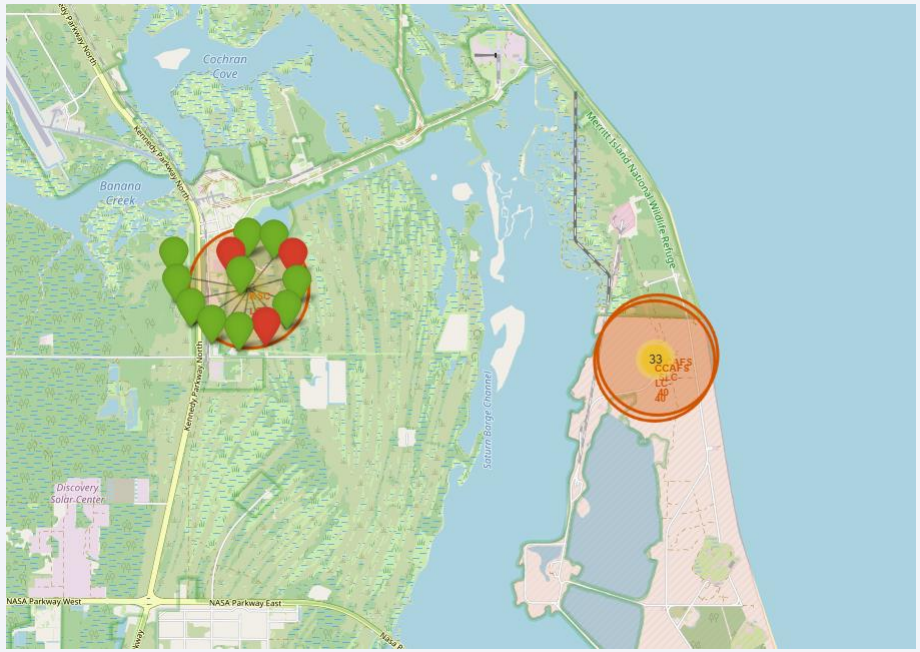
Section 3

Launch Sites Proximities Analysis

Locations of Launch sites

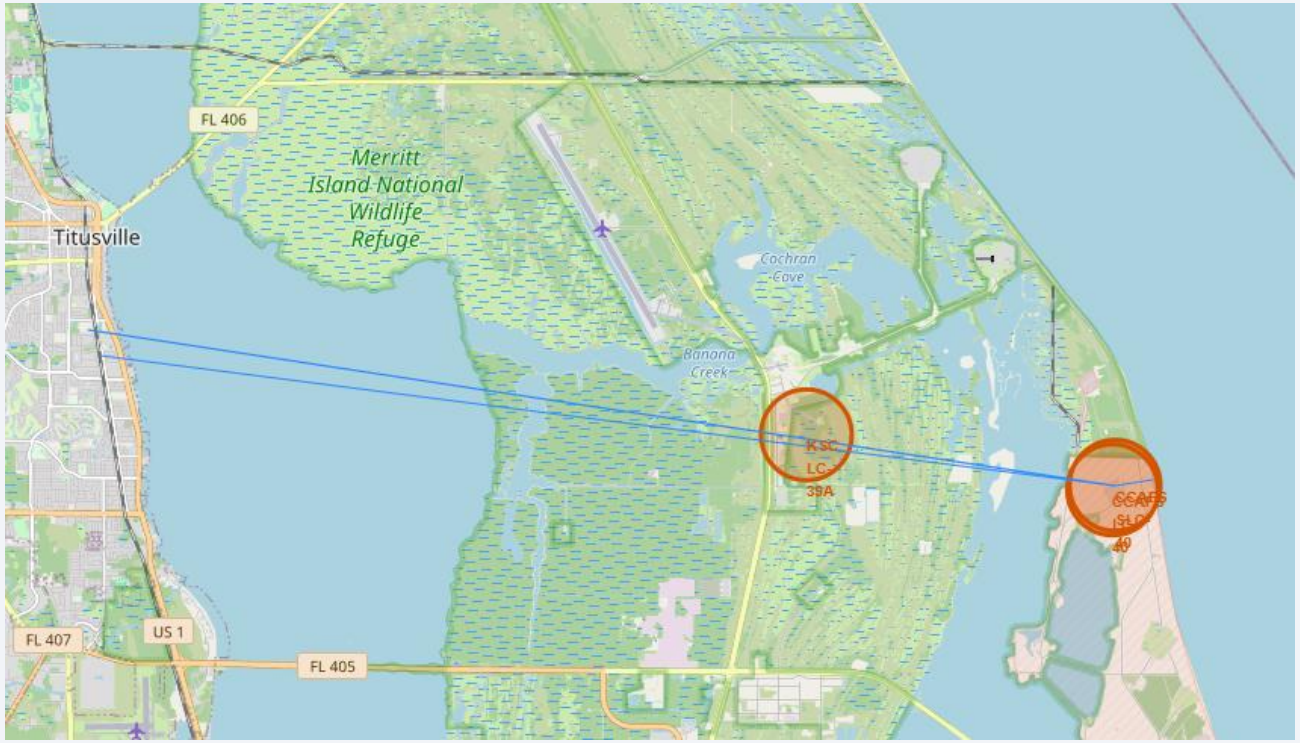
- Markers for the location of launch sites

Map of launch outcome

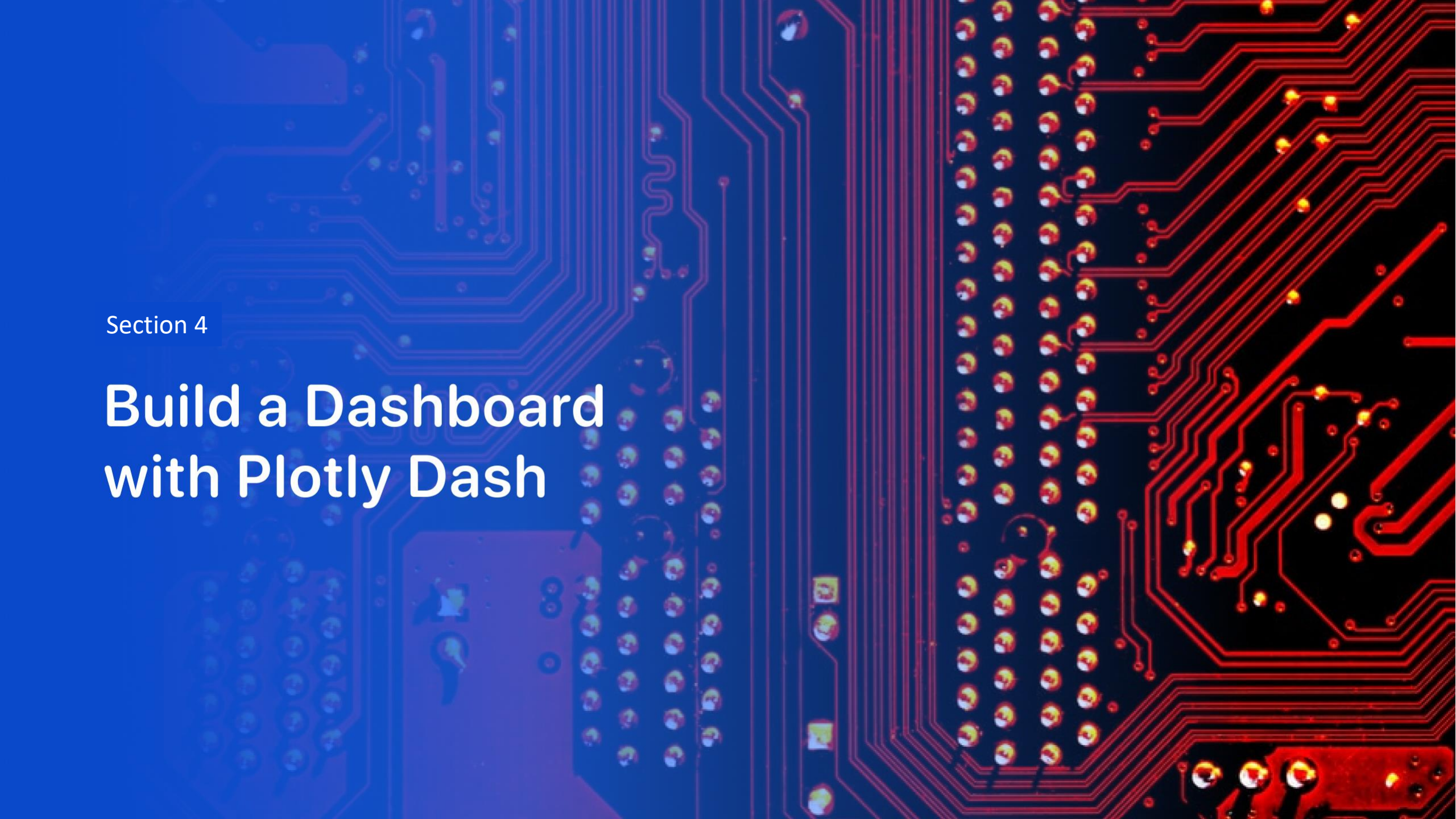


- Clusters of launch sites, markers in green were successful and markers in red were failure

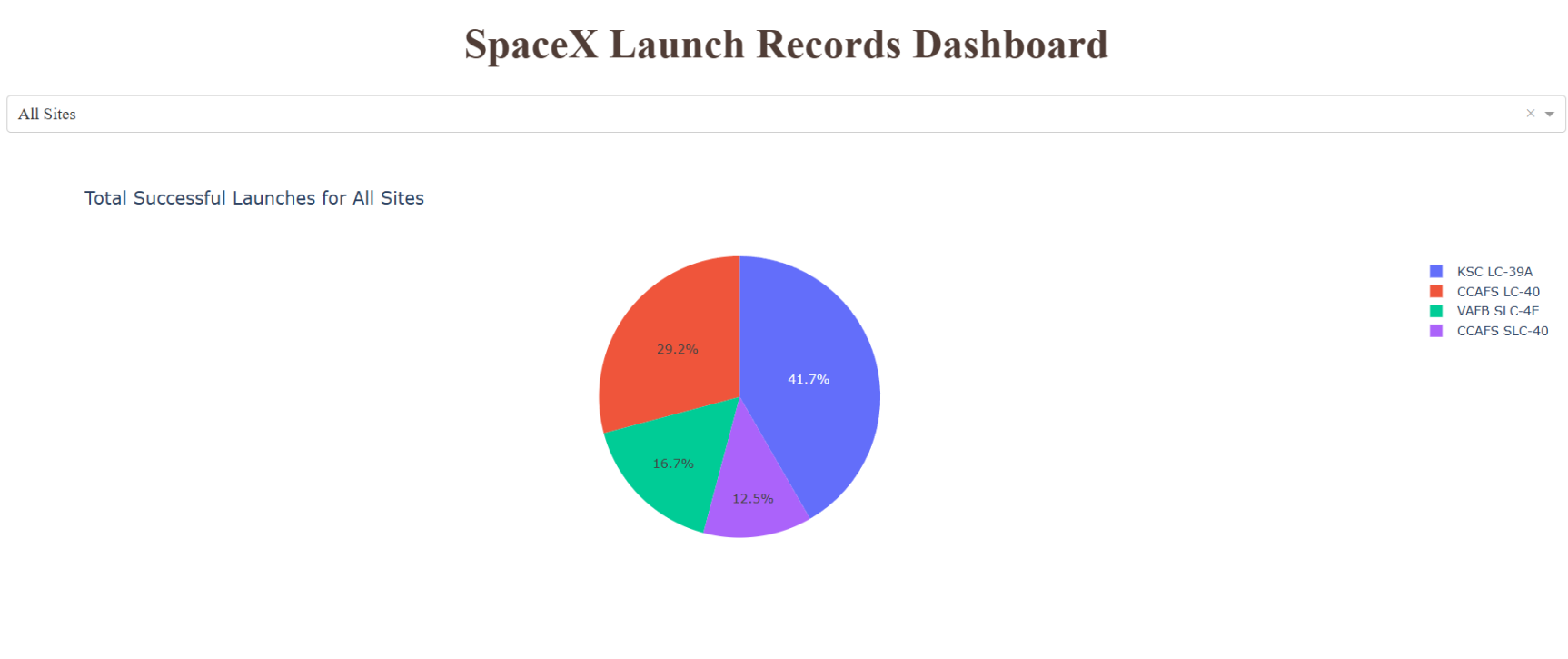
Locations near of the launch site



Distance between the launch site CCAFS SLC 40 with a city, a highway and a line coast

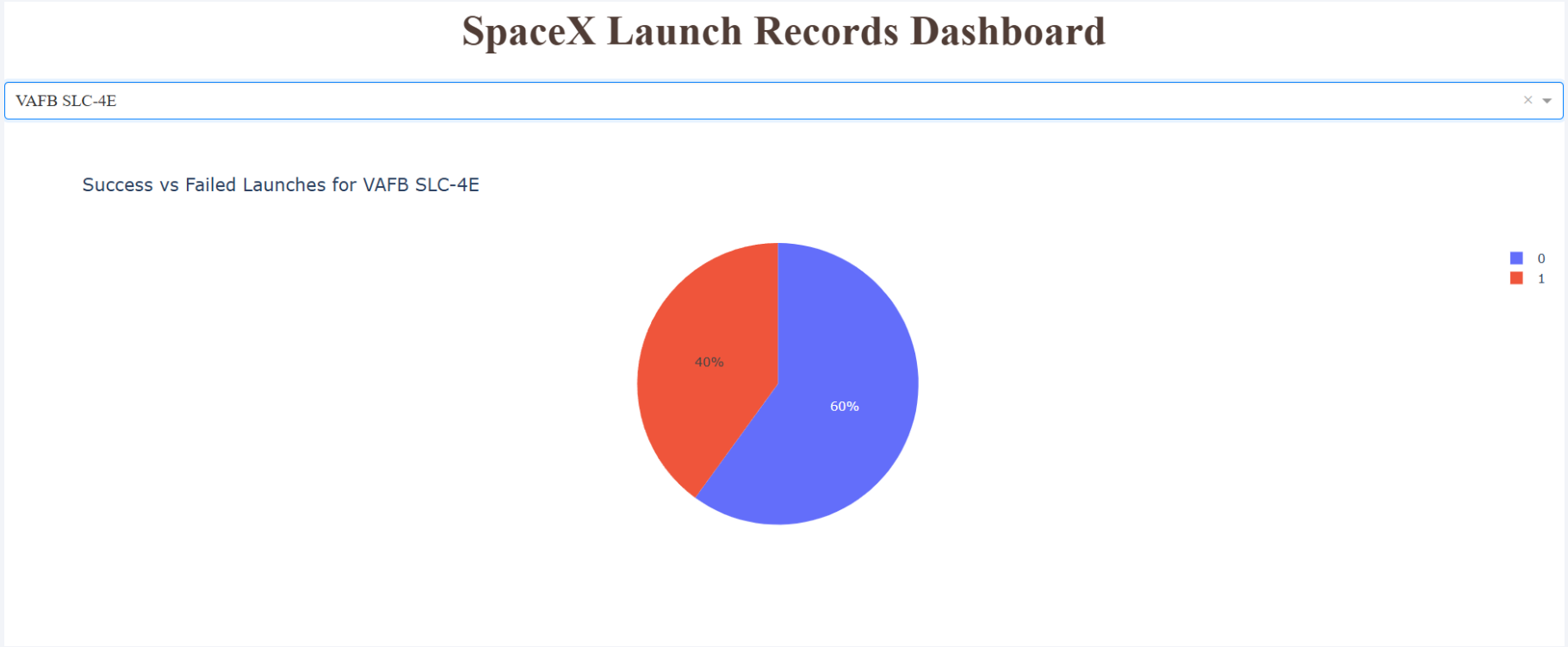

Section 4
Build a Dashboard with Plotly Dash

Successful Launches for All Sites



We can see that the majority of total successful launches were in the site KSC-LC-39A

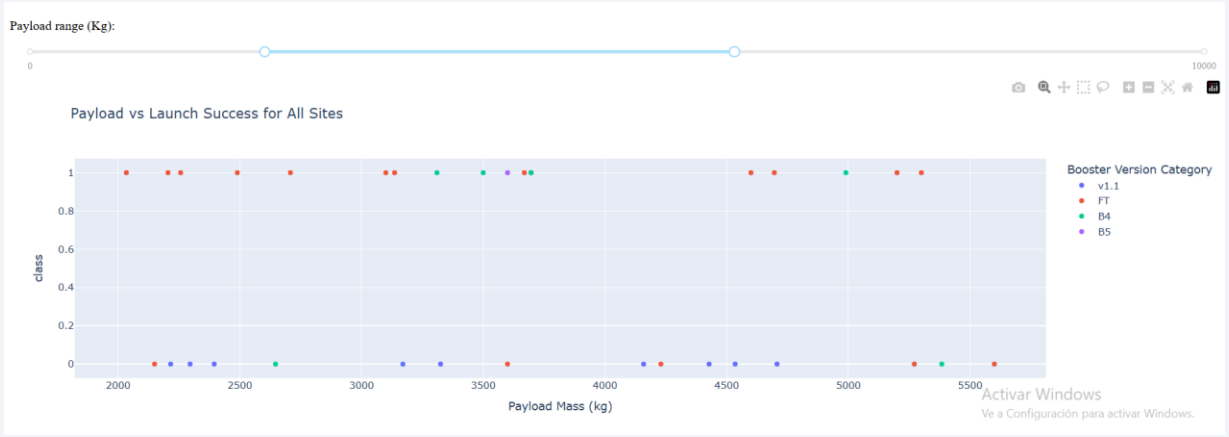
Highest Success ratio



The launch site with the highest success ratio is VAFB SLC-4E

Payload vs Launch Outcome

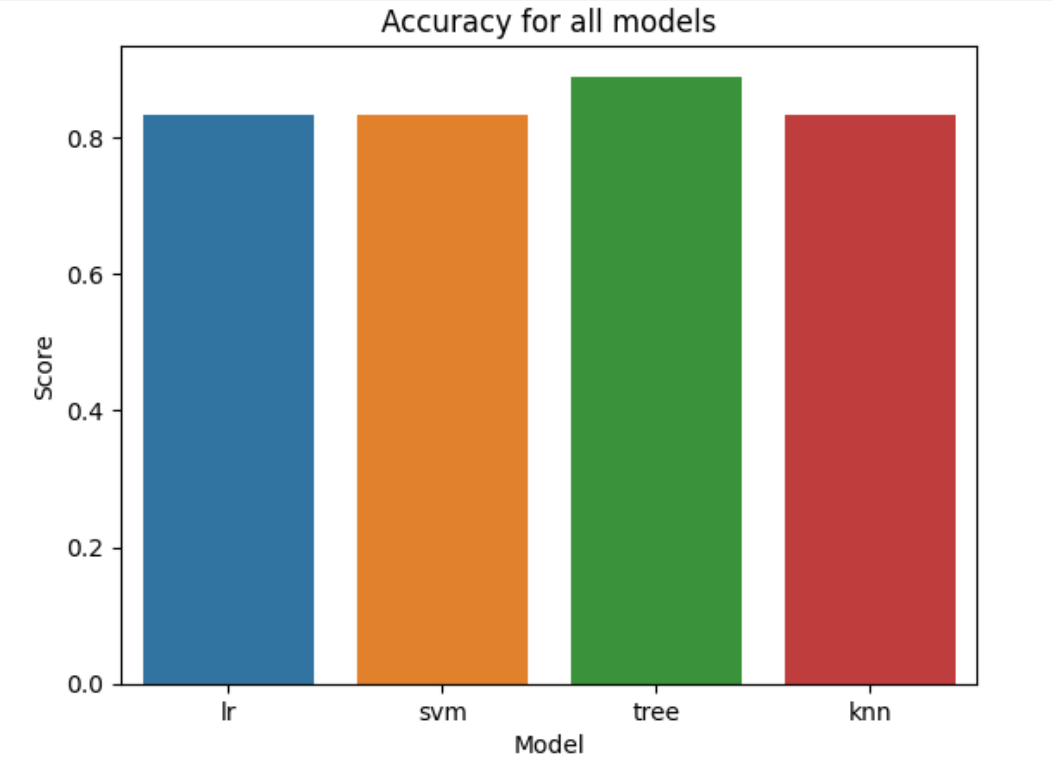
The majority of launch are in the range between 2000-6000kg, see this range, we can see that the majority of launchs have the booster version FT, with multiple success launch



Section 5

Predictive Analysis (Classification)

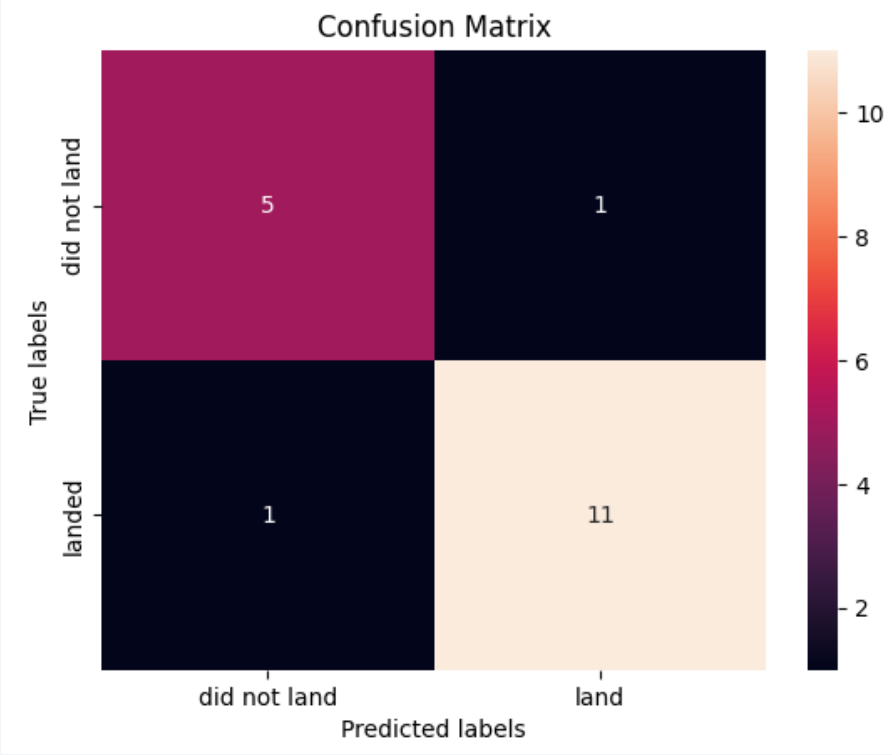
Classification Accuracy



The tree model has the highest accuracy

Confusion Matrix

- For a sample of 18, 5 didn't land was predicted correctly and 1 predicted incorrectly, and 11 land was predicted correctly while that 1 incorrectly



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!