SPACEX FALCON 9 LANDING PREDICTION

Dayron García Pérez

17/06/2025

Skills Network

OUTLINE

1. Executive Summary
2. Introduction
3. Methodology
4. Results
    - SQL queries in database
    - Visualizations – Charts
5. Interactive Charts
    - Interactive Folium maps
    - Dashboard
6. Predictive Analysis (Classification)
7. Conclusion

EXECUTIVE SUMMARY

- Methodologies
  - Data collection from different sources
  - Data wrangling
  - Exploratory Data Analysis – Visualizations
  - Exploratory Data Analysis – SQL queries
  - Interactive Folium maps
  - Interactive Dashboard with Dash
  - Predictive Analysis – Classification Machine Learning Models
- Results
  - Classification Models performance tables
  - Confusion Matrix.

INTRODUCTION

SECTION II

- Falcon 9 is the first reusable orbital class rocket, significantly reducing space access costs
- With a cost of 62 million dollars; other providers cost upward of 165 million dollars each
- Main Characteristics:
  - **Reusability**: It is the first reusable orbital class rocket, allowing its first stage to be recovered and reused on multiple missions, reducing costs and increasing efficiency.
  - **Power and payload capacity**: With a height of 70 meters and a thrust of more than 7,600 kN, it can carry up to 22,800 kg to low-Earth orbit and 8,300 kg to geosynchronous transfer orbit.
  - **Advanced technology**: It uses nine Merlin engines in the first stage and a Merlin Vacuum engine in the second stage, with RP-1 kerosene propellants and liquid oxygen, optimizing its performance and reliability.
- Based on the data provided by the different sources, we propose to predict whether the landing will be successful.

METHODOLOGY

SECTION III

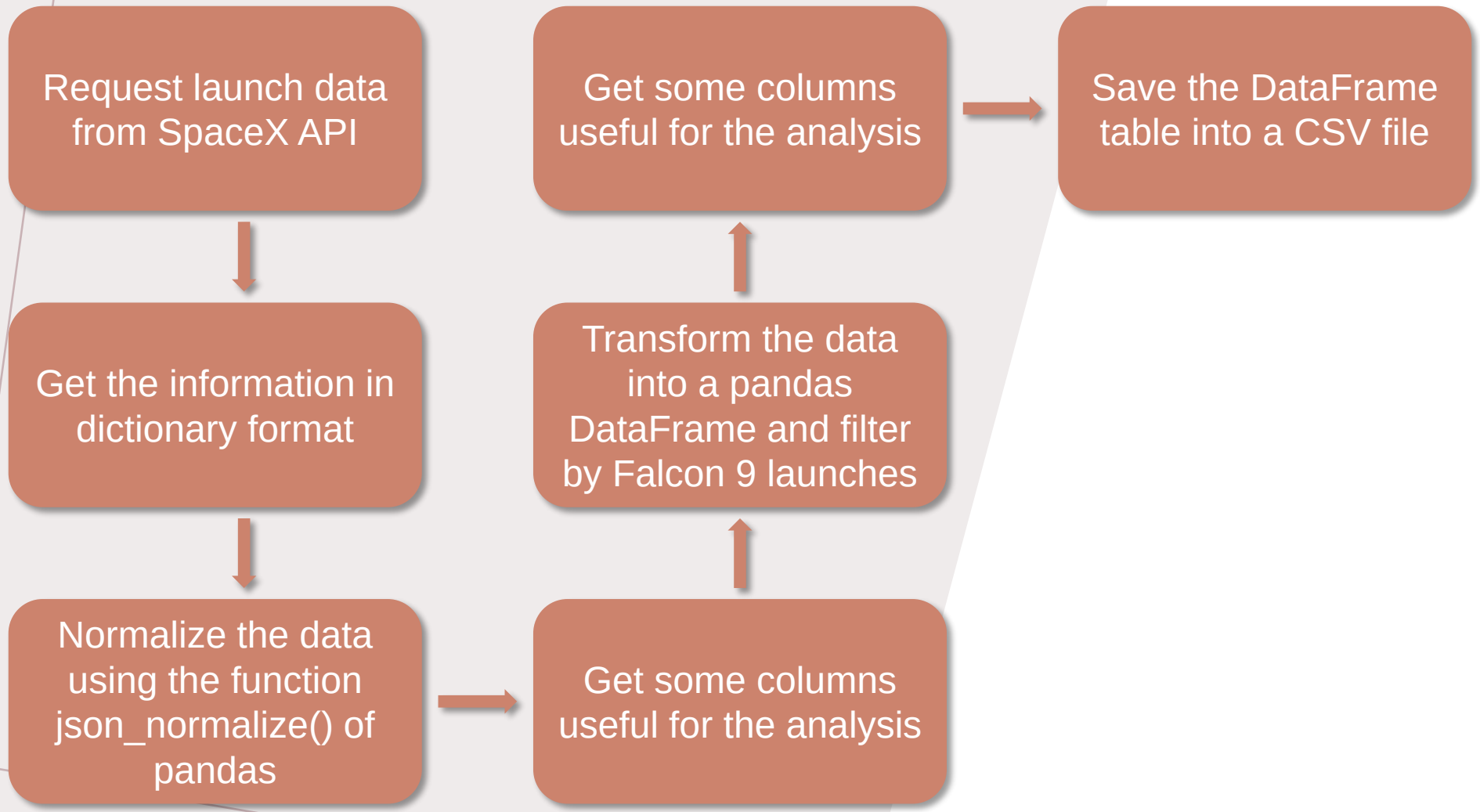
DATA COLLECTION

Data Sources: there are two main data sources in this project.

1. SpaceX
   - URL: https://api.spacexdata.com/v4/launches/past
   - Extract the query data as a json file and then manipulate the data to create a table of related data as a DataFrame.
   - Select the features that are useful to us and keep only the Falcon 9 related data.
2. Wikipedia
   - URL: https://en.wikipedia.org/w/index.php?title=List_of_Falcon_9_and_Falcon_Heavy_launches&oldid=1027686922
   - Extract the data from a web page using the BeautifulSoup library in this web page extract all tables using the HTML <table> tag.
   - Create a related data table (pandas DataFrame)

Save the data as CSV file.

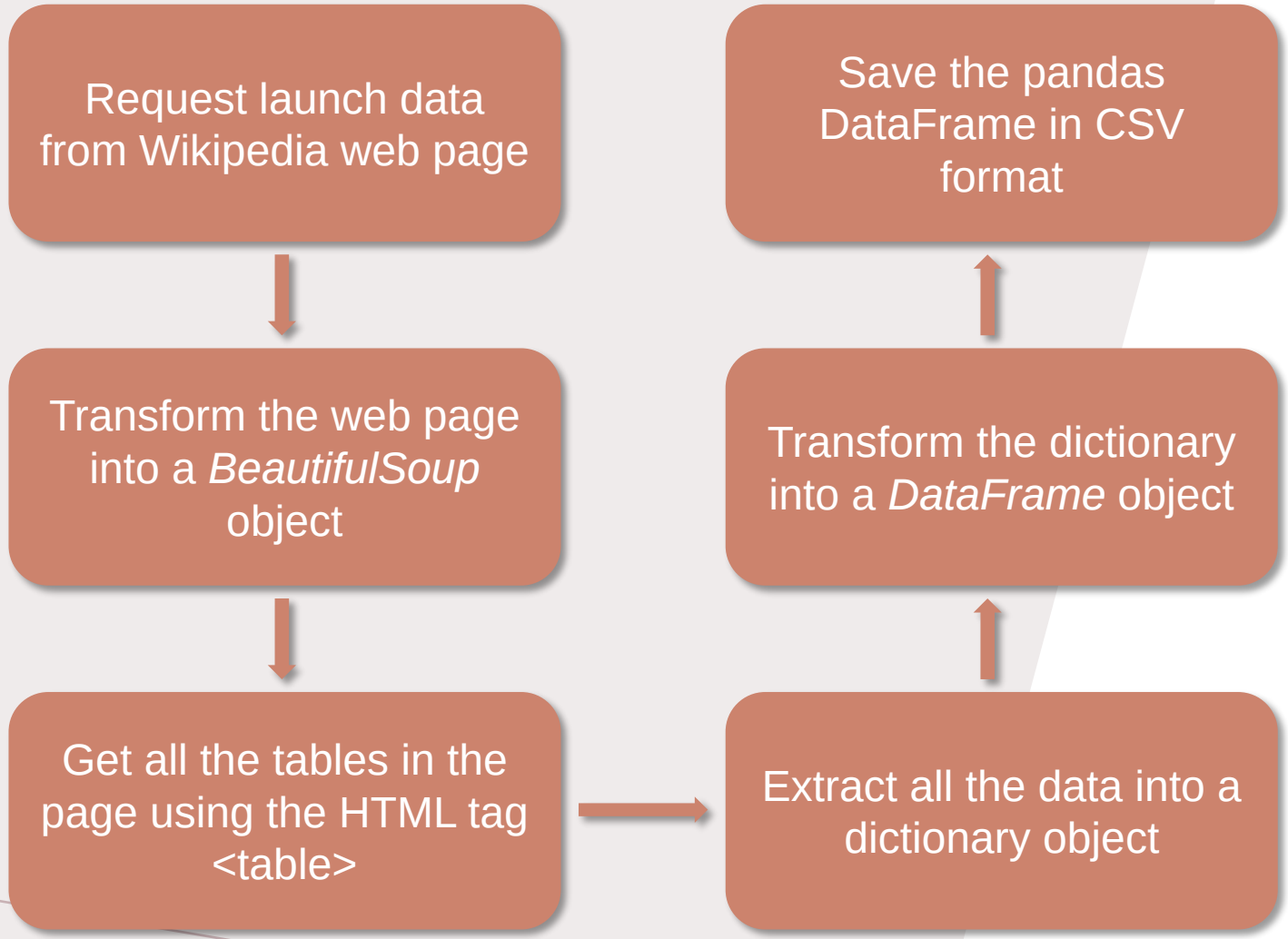
SPACEX API – DATA WRANGLING



COLUMNS

- FLIGHTNUMBER
- DATE
- BOOSTERVERSION
- PAYLOADMASS
- ORBIT
- LAUNCHSITE
- OUTCOME
- FLIGHTS
- GRIDFINS
- REUSED
- LEGS
- LANDINGPAD
- BLOCK
- REUSEDCOUNT
- SERIAL
- LONGITUDE
- LATITUDE

WIKIPEDIA PAGE – DATA WRANGLING



COLUMNS

- FLIGHT NO.
- LAUNCH SITE
- PAYLOAD
- PAYLOAD MASS
- ORBIT
- CUSTOMER
- LAUNCH OUTCOME
- VERSION BOOSTER
- BOOSTER LANDING
- DATE
- TIME

EXPLORATORY DATA ANALYSIS

Scatter plots:

- Relationship between Flight Number and Launch Site
- Relationship between Payload Mass and Launch Site
- Relationship between FlightNumber and Orbit type
- Relationship between Payload Mass and Orbit type

Bar plot:

- Relationship between success rate of each orbit type

Line plot:

- Visualize the launch success yearly trend

EXPLORATORY DATA ANALYSIS

SQL queries:

- Task 1: Display the names of the unique launch sites in the space mission.
- Task 2: Display 5 records where launch sites begin with the string 'CCA’
- Task 3: Display the total payload mass carried by boosters launched by NASA (CRS)
- Task 4: Display average payload mass carried by booster version F9 v1.1
- Task 5: List the date when the first successful landing outcome in ground pad was archived.
- Task 6: List the names of the boosters which have success in drone ship and have payload mass greater than 4000 but less than 6000.
- Task 7: List the total number of successful and failure mission outcomes
- Task 8: List all the booster_versions that have carried the maximum payload mass.
- Task 9: List the records which will display the month names, failure landing_outcomes in drone ship ,booster versions, launch_site for the months in year 2015.
- Task 10: 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.

INTERACTIVE MAP WITH FOLIUM

Marker of all Launch Sites:

- Added Marker with circle and text label of NASA Space Center Jon F. Kennedy and Vandenberg Space Launch Complex.
- Colored Markers of the launch outcomes for each Launch Site
- Calculate the distances between the Launch Sites and its proximities to the coast.

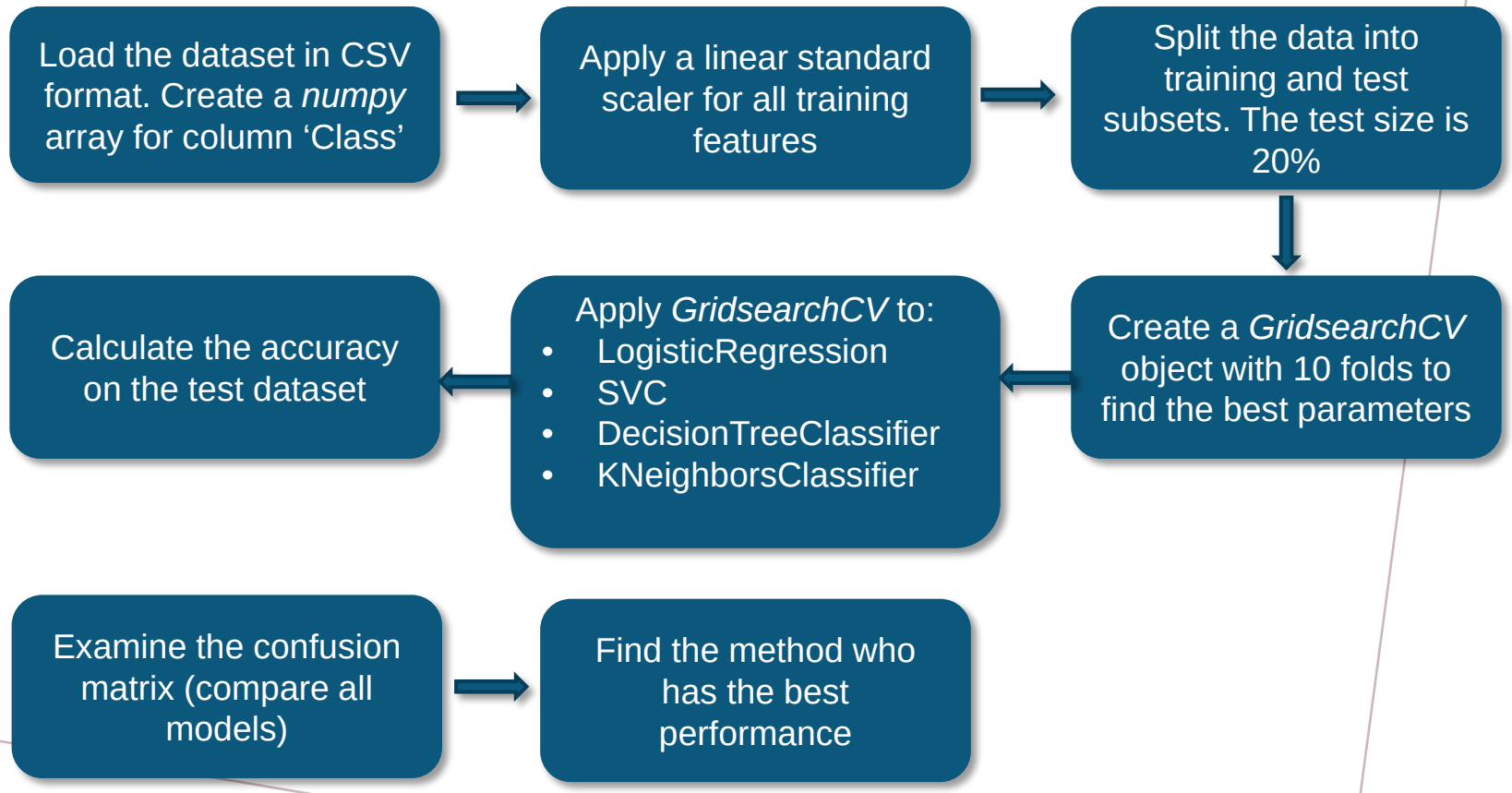
BUILD A DASHBOARD

The built Dashboard has two interactive menus:

- Dropdown list for all Launch Sites.
    - This dropdown menu filters a pie chart where is shown the total successful and fail launches count.
- Slicer for select the payload mass range.
    - This slicer is a filter for a scatter plot with a relationship between Payload Mass vs Success Rate for the different booster versions.

CLASSIFICATION MACHINE LEARNING ALGORITHM

The goal of this Project is to build a machine learning algorithm which can predict if a Falcon 9 launch will success or not based on the first stage data of the rocket.



RESULTS

SECTION IV

SQL QUERIES IN DATABASE

TASK 1: DISPLAY THE NAMES OF THE UNIQUE LAUNCH SITES IN THE SPACE MISSION.

```
%sql select distinct(Launch_Site) from SPACEXTABLE
```

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

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

SQL QUERIES IN DATABASE

TASK 2: DISPLAY 5 RECORDS WHERE LAUNCH SITES BEGIN WITH THE STRING 'CCA'

```
%sql select * from SPACEXTABLE where Launch_Site like 'CCA%' limit 5
```

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

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

SQL QUERIES IN DATABASE

TASK 3: DISPLAY THE TOTAL PAYLOAD MASS CARRIED BY BOOSTERS LAUNCHED BY NASA (CRS)

```
%sql select sum(PAYLOAD_MASS__KG_) as 'TOTAL PAYLOAD MASS BY NASA (CRS)' from SPACEXTABLE where Customer = 'NASA (CRS)'
```

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

TOTAL PAYLOAD MASS BY NASA (CRS)
45596

SQL QUERIES IN DATABASE

TASK 4: DISPLAY AVERAGE PAYLOAD MASS CARRIED BY BOOSTER VERSION F9 V1.1

```
%sql select avg(PAYLOAD_MASS__KG_) from SPACEXTABLE where Booster_Version like 'F9 v1.1'
```

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

avg(PAYLOAD_MASS__KG_)
2928.4

SQL QUERIES IN DATABASE

TASK 5: LIST THE DATE WHEN THE FIRST SUCCESSFUL LANDING OUTCOME IN GROUND PAD WAS ACHIEVED.

```
%sql select min(Date) from SPACEXTABLE where Landing_Outcome = 'Success (ground pad)'
```

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

min(Date)
2015-12-22

SQL QUERIES IN DATABASE

TASK 6: LIST THE NAMES OF THE BOOSTERS WHICH HAVE SUCCESS IN DRONE SHIP AND HAVE PAYLOAD MASS GREATER THAN 4000 BUT LESS THAN 6000

```
%sql select Booster_Version from SPACEXTABLE where Landing_Outcome = 'Success (drone ship)' and PAYLOAD_MASS__KG_ between 4000 and 6000
```

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

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

SQL QUERIES IN DATABASE

TASK 7: LIST THE TOTAL NUMBER OF SUCCESSFUL AND FAILURE MISSION OUTCOMES

```
%%sql select Landing_Outcome, count(Landing_Outcome) from SPACEXTABLE group by Landing_Outcome
having Landing_Outcome like 'Success%' or Landing_Outcome like 'Failure%'
```

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

Landing_Outcome	count(Landing_Outcome)
Failure	3
Failure (drone ship)	5
Failure (parachute)	2
Success	38
Success (drone ship)	14
Success (ground pad)	9

```
%%sql select Landing_Outcome, count(Landing_Outcome) from SPACEXTABLE group by Landing_Outcome
having Landing_Outcome like 'Success' or Landing_Outcome like 'Failure'
```

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

Landing_Outcome	count(Landing_Outcome)
Failure	3
Success	38

SQL QUERIES IN DATABASE

TASK 8: LIST ALL THE BOOSTER_VERSIONS THAT HAVE CARRIED THE MAXIMUM PAYLOAD MASS.

```
%sql select Booster_Version from SPACEXTABLE where PAYLOAD_MASS__KG_ = (select max(PAYLOAD_MASS__KG_) from SPACEXTABLE)
```

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

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

SQL QUERIES IN DATABASE

TASK 9: LIST THE RECORDS WHICH WILL DISPLAY THE MONTH NAMES, FAILURE LANDING_OUTCOMES IN DRONE SHIP, BOOSTER VERSIONS, LAUNCH_SITE FOR THE MONTHS IN YEAR 2015.

```
%%sql select substr(Date, 6,2), Landing_Outcome, Booster_Version, Launch_Site from SPACEXTABLE
where Landing_Outcome like 'Failure (drone ship)' and substr(Date,0,5) = '2015'
```

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

substr(Date, 6,2)	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

SQL QUERIES IN DATABASE

TASK 10: 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.

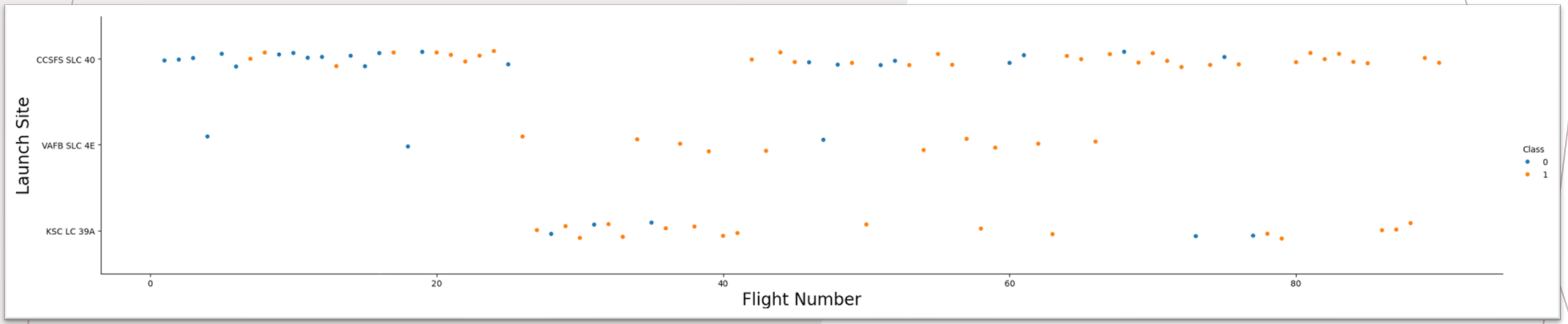
```
%%sql select Landing_Outcome, count(Landing_Outcome) from SPACEXTABLE
group by Landing_Outcome having Date between '2010-06-04' and '2017-03-20' order by count(Landing_Outcome) desc
```

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

Landing_Outcome	count(Landing_Outcome)
No attempt	21
Success (drone ship)	14
Success (ground pad)	9
Failure (drone ship)	5
Controlled (ocean)	5
Uncontrolled (ocean)	2
Failure (parachute)	2
Precluded (drone ship)	1

VISUALIZATIONS – CHARTS

TASK 1: VISUALIZE THE RELATIONSHIP BETWEEN FLIGHT NUMBER AND LAUNCH SITE

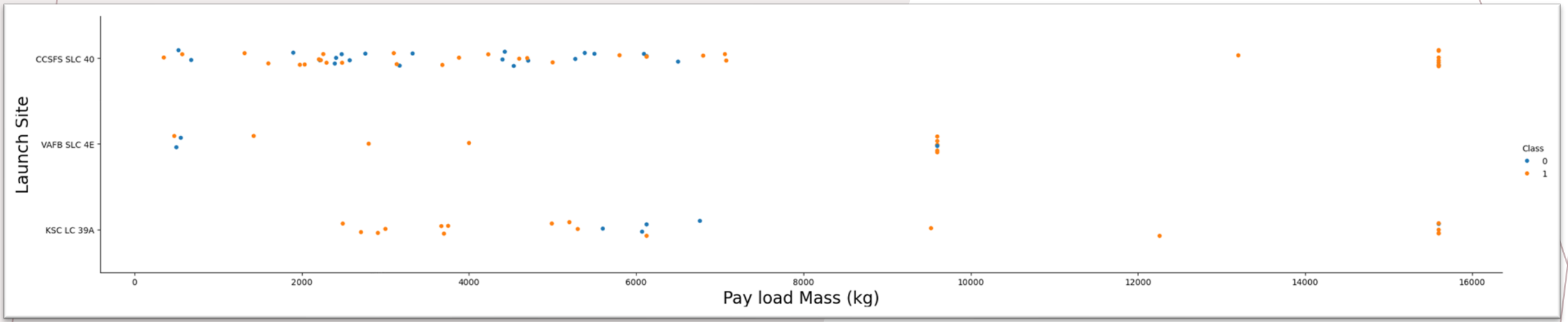


Remarks:

- The first launches were in CCSFS SLC 40 and they were a failure
- Most of the rockets were launched from location CCSFS SLC 40
- The last launches are a success
- The launch site VAFB SLC 4E has the minimum value of launches

VISUALIZATIONS – CHARTS

TASK 2: VISUALIZE THE RELATIONSHIP BETWEEN PAYLOAD MASS AND LAUNCH SITE

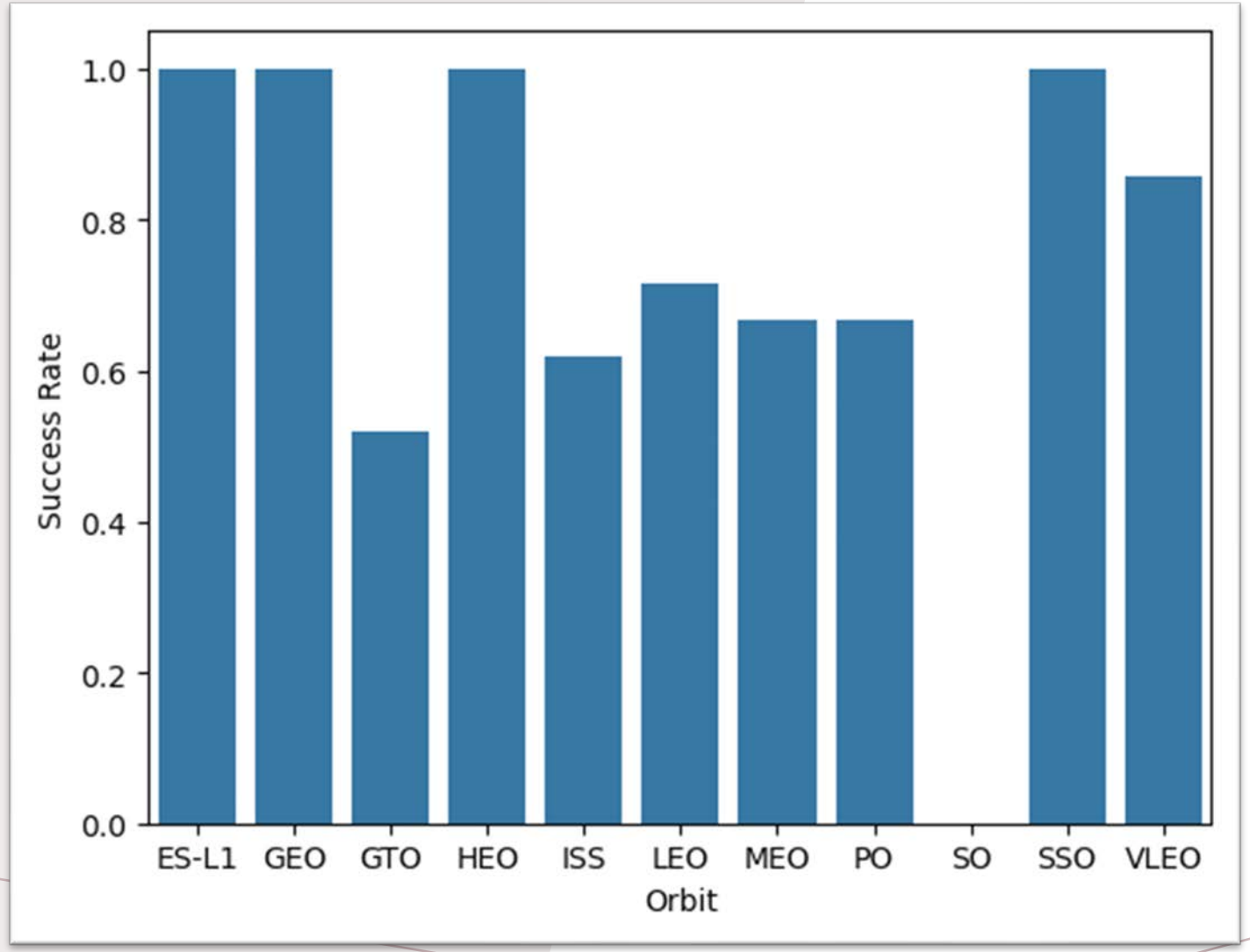


Remarks:

- Launches with Payload Mass less than 10 000 kg were done in CCSFS SLC 40 and KSC LC 39A
- The max of Payload Mass in VAFB SLC 4E is 9 600 kg
- Launches with the Payload max were in CCSFS SLC 40 and KSC LC 39A and it is 15 600 kg. The most of them were successful.

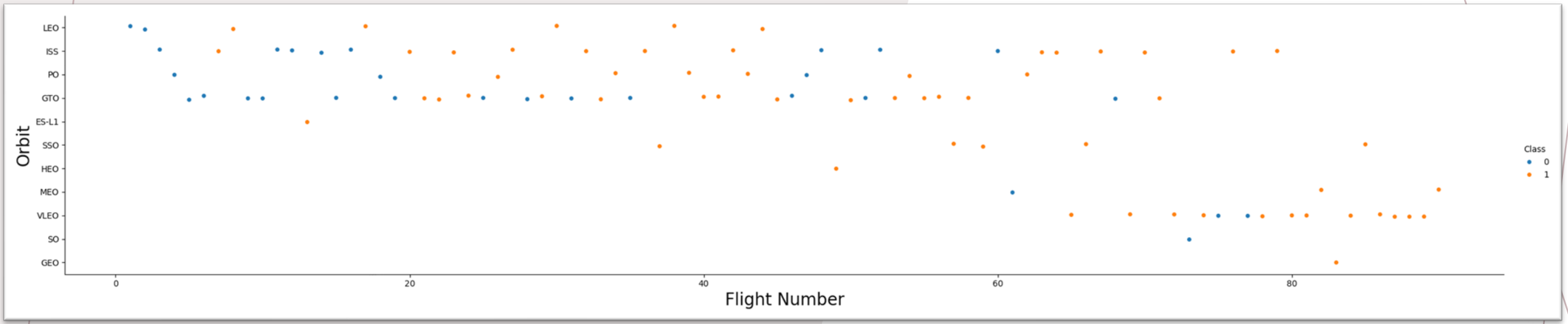
VISUALIZATIONS – CHARTS

TASK 3: VISUALIZE THE RELATIONSHIP BETWEEN SUCCESS RATE OF EACH ORBIT TYPE



VISUALIZATIONS – CHARTS

TASK 4: VISUALIZE THE RELATIONSHIP BETWEEN FLIGHT NUMBER AND ORBIT TYPE

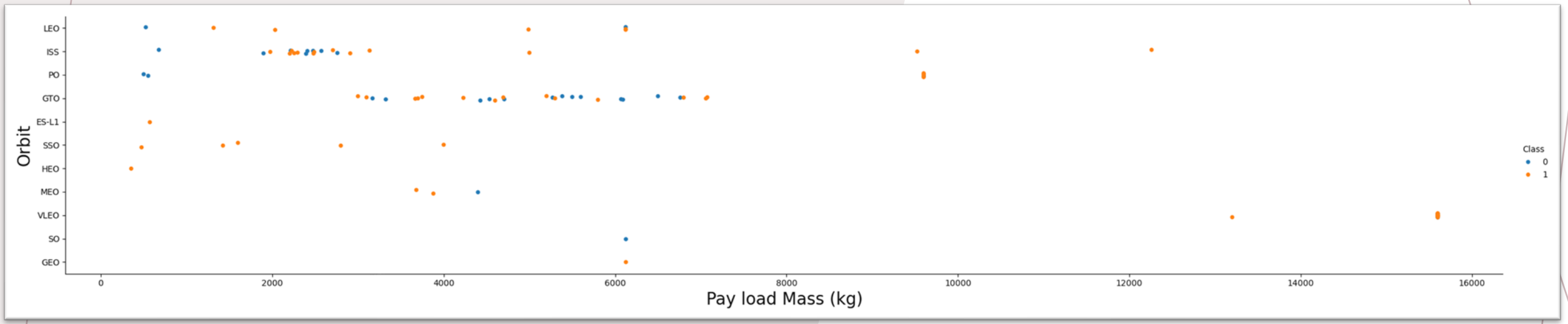


Remarks:

- The LEO and VLEO orbits success seems to be related to the number of flights.
- In the GTO orbit there appears to be no relationship between flight number and success.
- In the beginning the most popular orbits were LEO, ISS and GTO.
- The most popular orbit in the last launches is VLEO.

VISUALIZATIONS – CHARTS

TASK 5: VISUALIZE THE RELATIONSHIP BETWEEN PAYLOAD MASS AND ORBIT TYPE



Remarks:

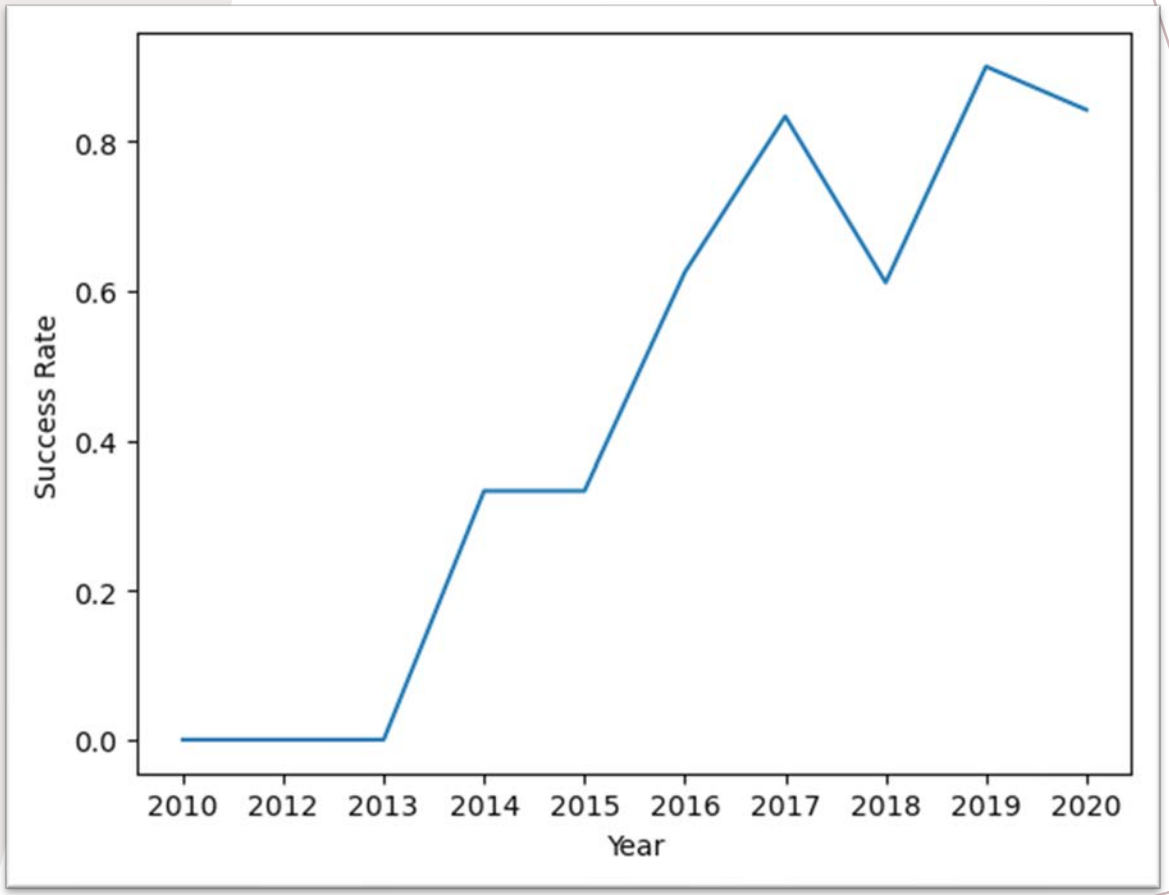
- With heavy Payloads the success ratio is high.
- In the case of GTO, it is difficult to distinguish the success and fail launch depending on the Payload mass.

VISUALIZATIONS – CHARTS

TASK 6: VISUALIZE THE LAUNCH SUCCESS YEARLY TREND

Remarks:

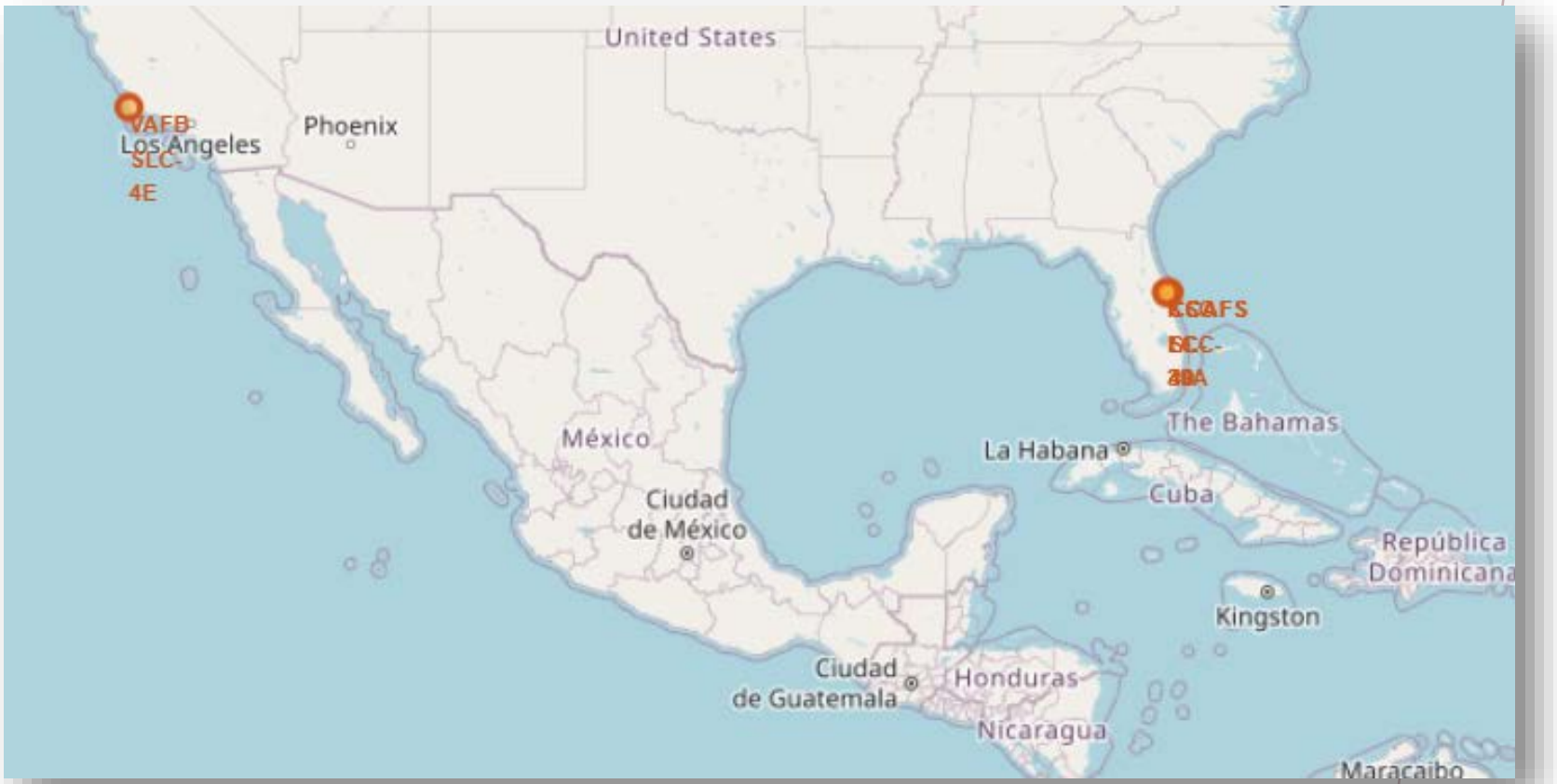
- The Success Rate increases over years.



INTERACTIVE CHARTS

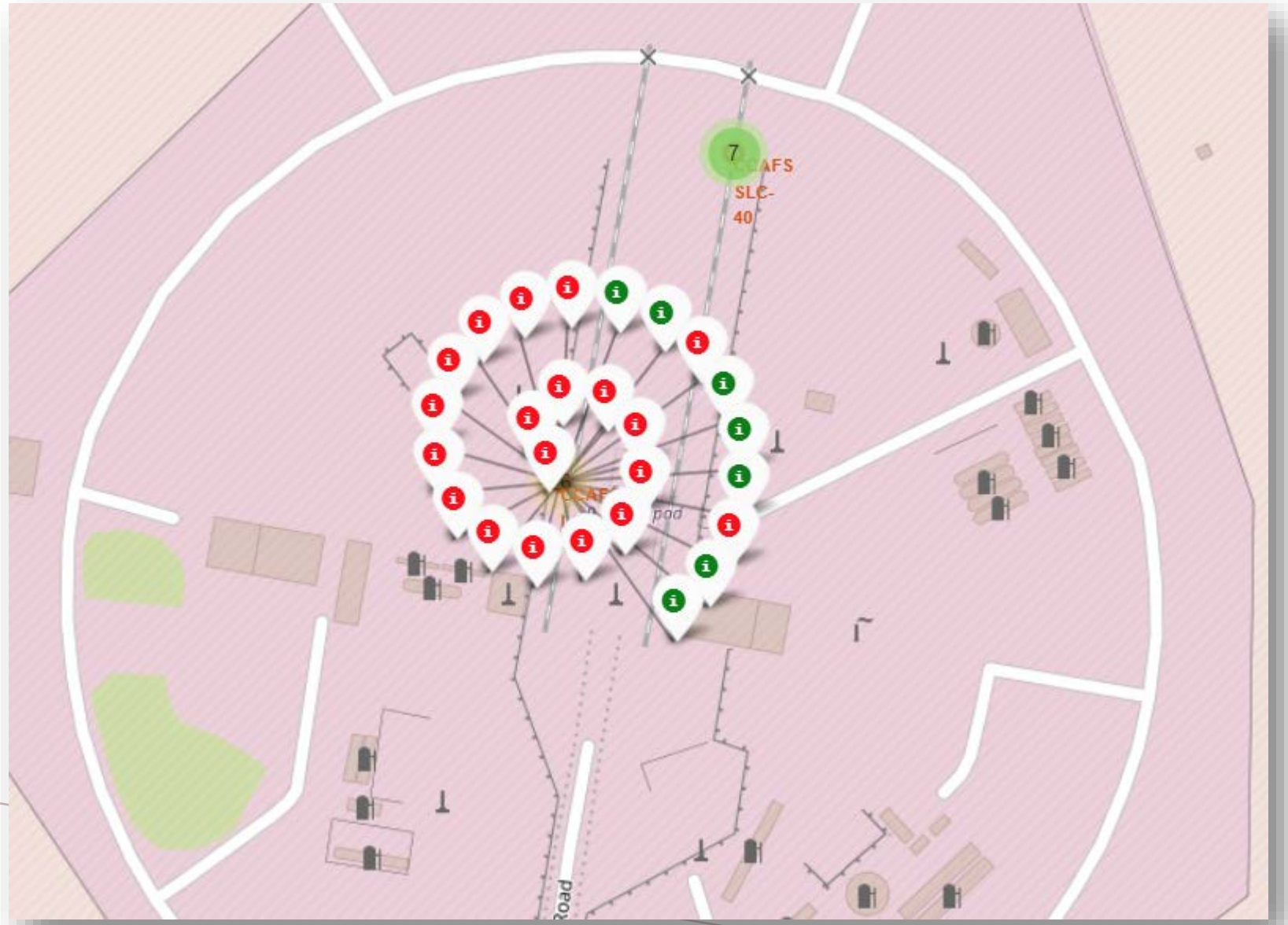
SECTION V

INTERACTIVE MAP WITH FOLIUM



All Launch Sites on the map

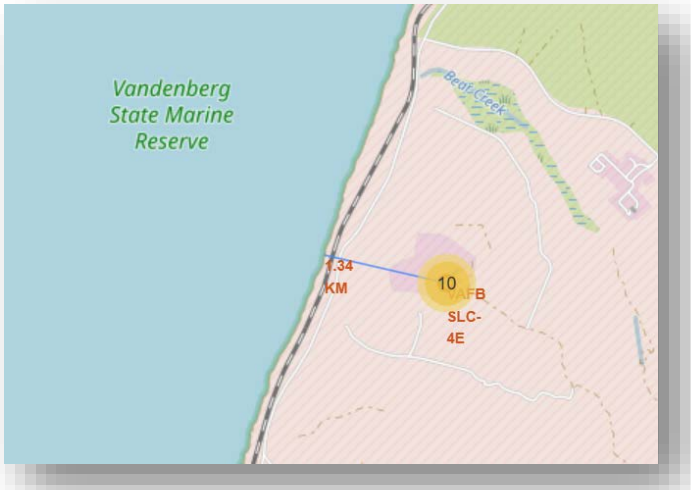
INTERACTIVE MAP WITH FOLIUM



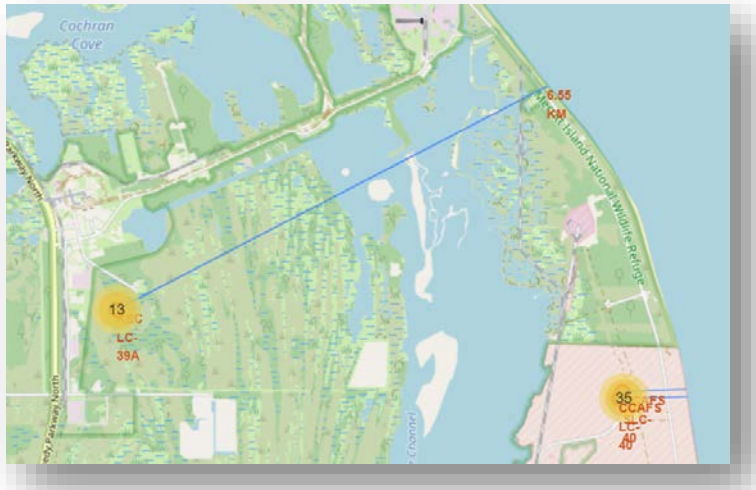
Total of launches in CCAFS LC – 40

- Red label indicates a fail launch.
- Green label indicates a success.

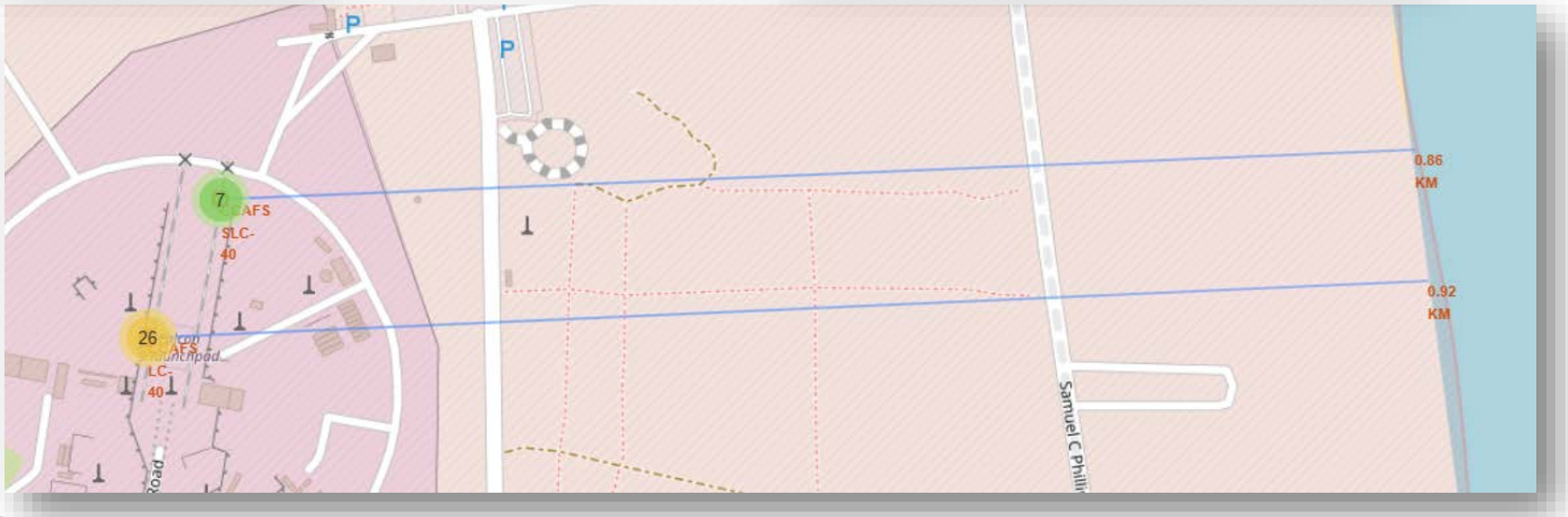
INTERACTIVE MAP WITH FOLIUM



VAFB SLC – 4E

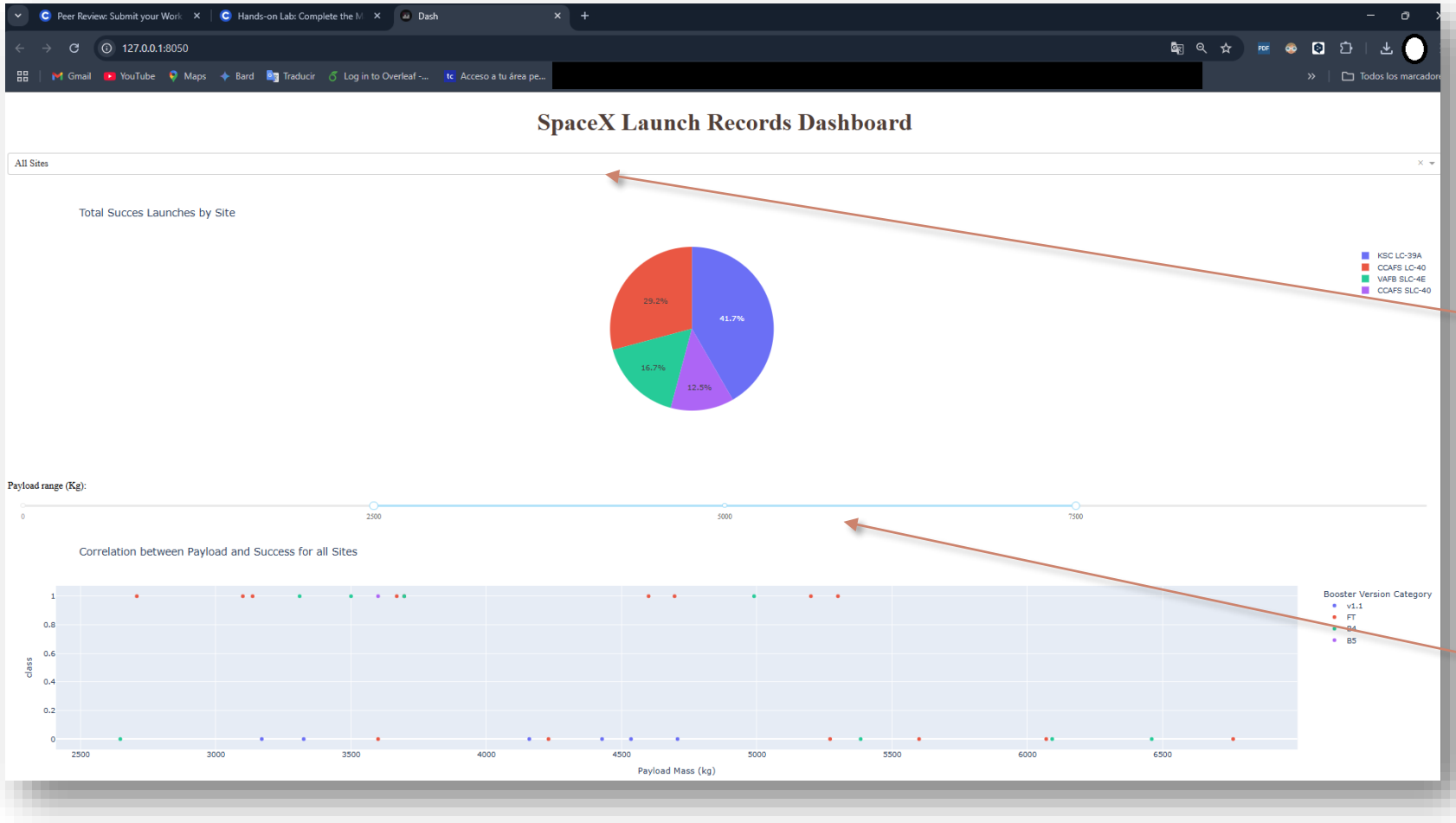


KSC LC – 39A



Distance to coast of Launch Sites CCAFS LC – 40 and CCAFS SLC – 40

DASHBOARD

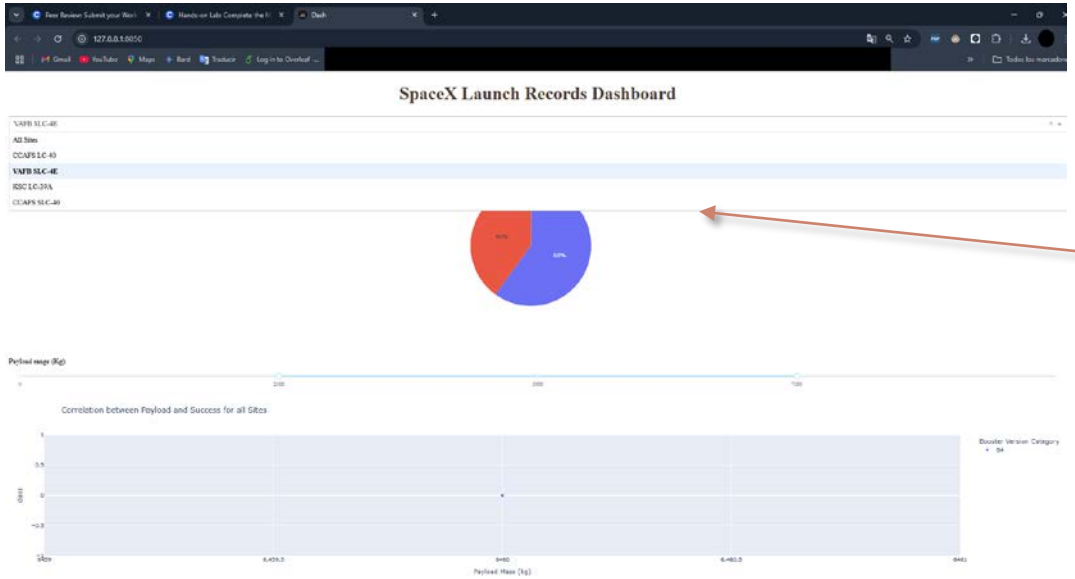


Dropdown
Launch Site

Slicer
Payload Mass
kg

https://github.com/dayrongp93/couse10_assigments/blob/main/Module_03/spacex_dash_app.py

DASHBOARD

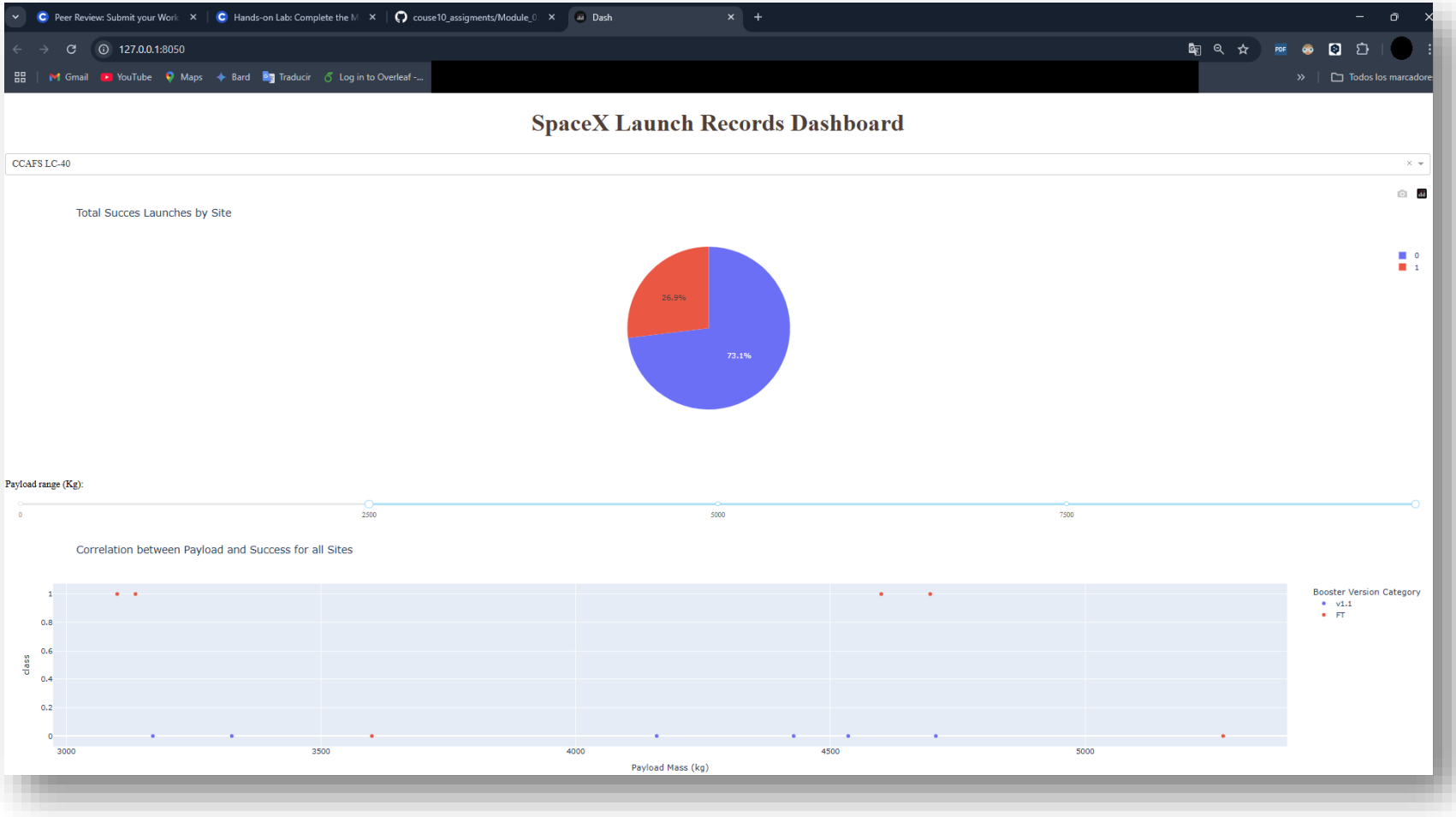


Dropdown
Launch Site

Slicer
Payload Mass
kg

https://github.com/dayrongp93/couse10_assigments/blob/main/Module_03/spacex_dash_app.py

DASHBOARD



Remarks:

- Launch Site dropdown for site CCAFS LC -40
- Payload Mass range between 2500 and 10000 kg

https://github.com/dayrongp93/couse10_assigments/blob/main/Module_03/spacex_dash_app.py

PREDICTIVE ANALYSIS (CLASSIFICATION)

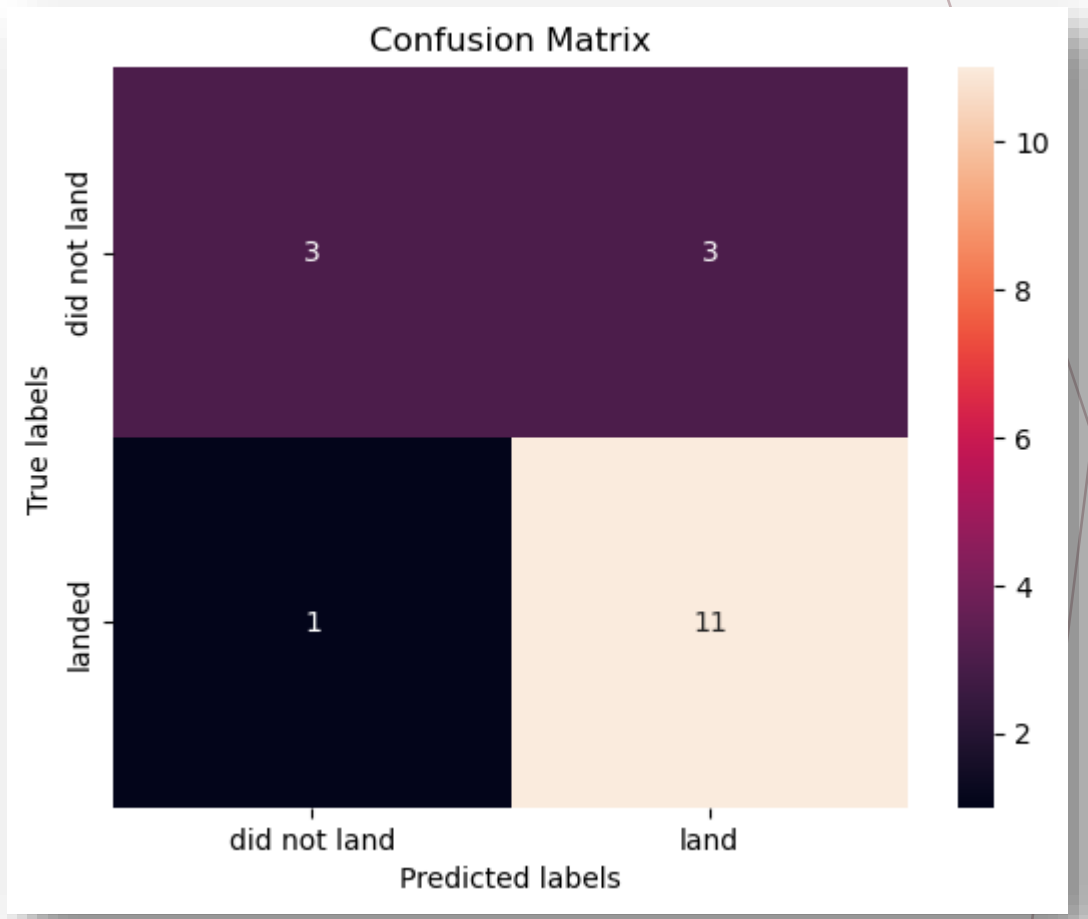
SECTION VI

MACHINE LEARNING MODELS RESULTS

Algorithm Logistic Regression

Score on Test Data: 78%

	Precision	Recall	F1-score	Support
0	0,75	0,50	0,60	6
1	0,79	0,92	0,85	12
Accuracy			0,78	18
Macro Avg	0,77	0,71	0,72	18
Weighted Avg	0,77	0,78	0,76	18

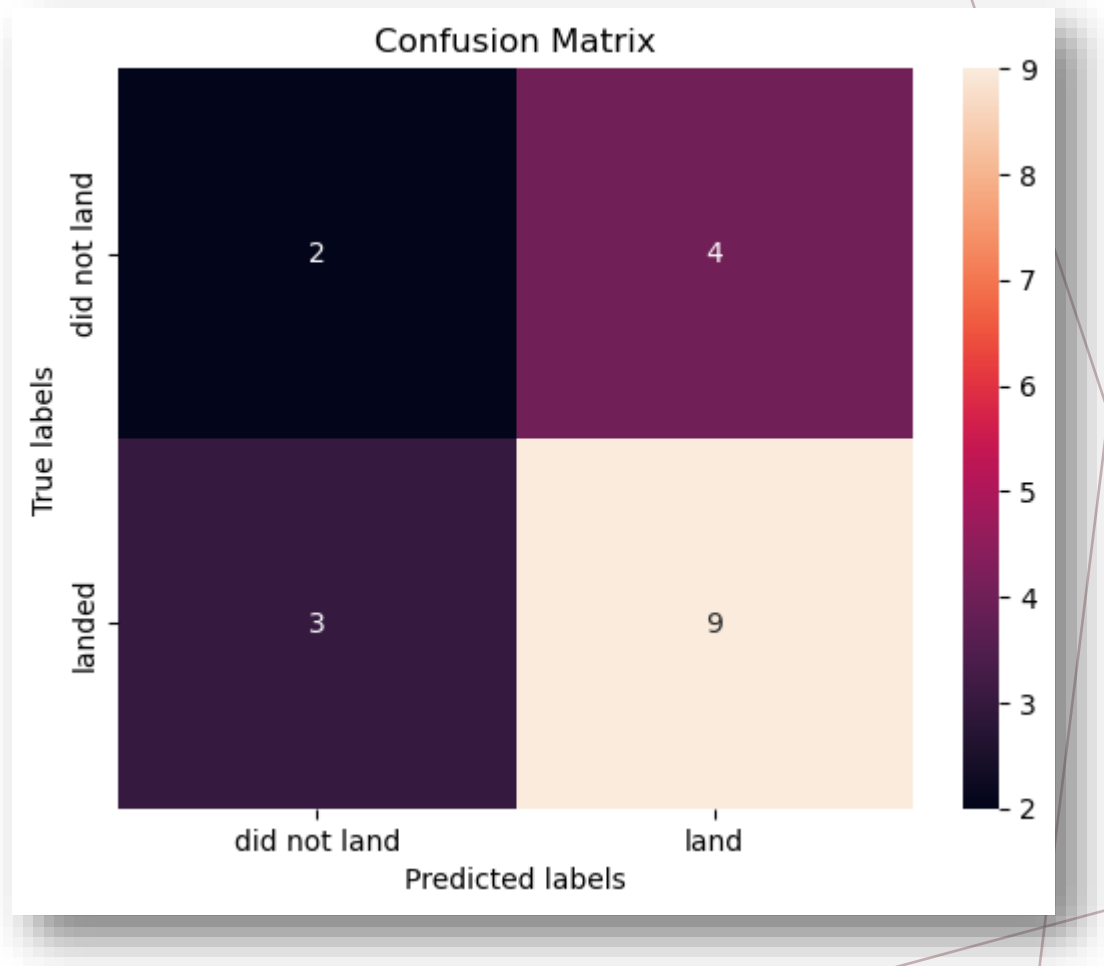


MACHINE LEARNING MODELS RESULTS

Algorithm Support Vector Machine (SVM)

Score on Test Data: 61%

	Precision	Recall	F1-score	Support
0	0,40	0,33	0,36	6
1	0,69	0,75	0,72	12
Accuracy			0,61	18
Macro Avg	0,55	0,54	0,54	18
Weighted Avg	0,59	0,61	0,60	18

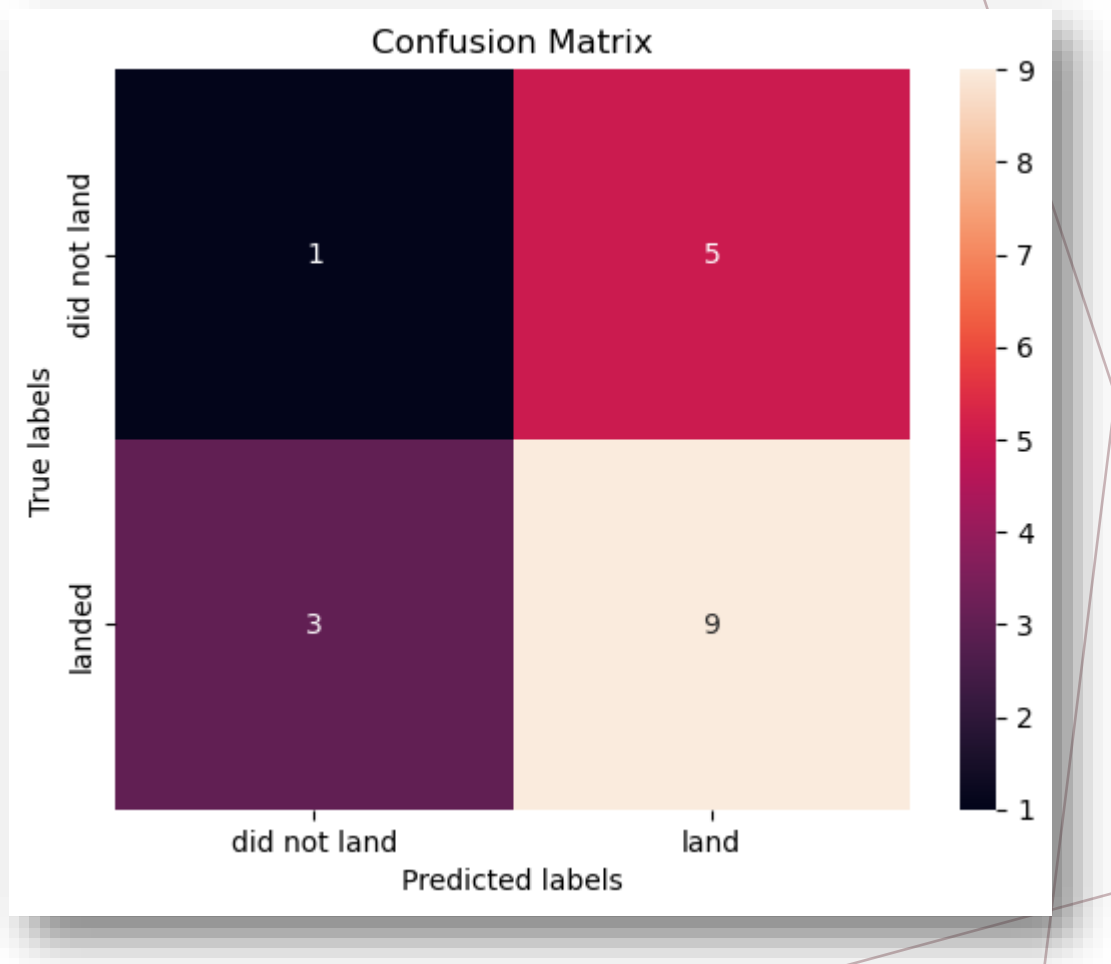


MACHINE LEARNING MODELS RESULTS

Algorithm Decision Tree Classifier

Score on Test Data: 56%

	Precision	Recall	F1-score	Support
0	0,25	0,17	0,20	6
1	0,64	0,75	0,69	12
Accuracy			0,56	18
Macro Avg	0,45	0,46	0,45	18
Weighted Avg	0,51	0,56	0,53	18

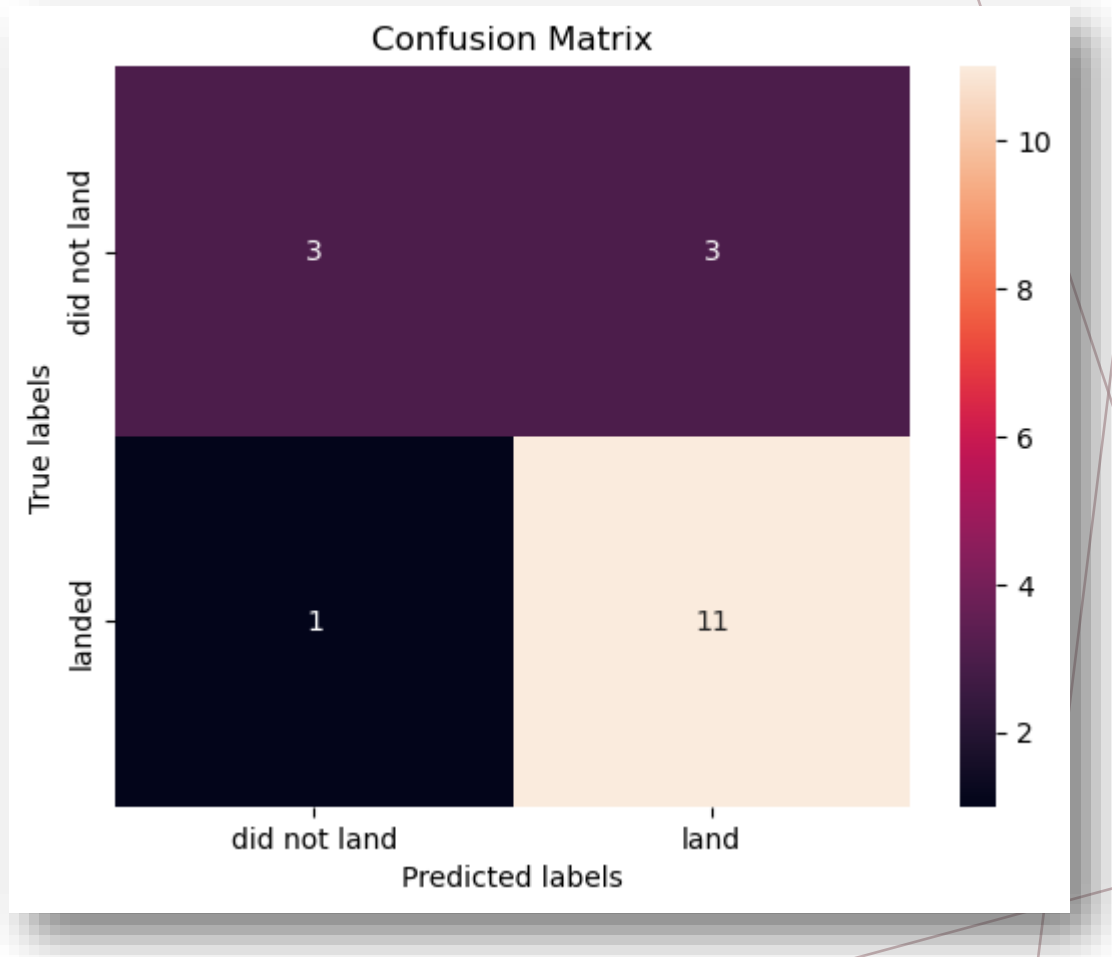


MACHINE LEARNING MODELS RESULTS

Algorithm K Nearest Neighbors (KNN)

Score on Test Data: 78%

	Precision	Recall	F1-score	Support
0	0,75	0,50	0,60	6
1	0,79	0,92	0,85	12
Accuracy			0,78	18
Macro Avg	0,77	0,71	0,72	18
Weighted Avg	0,77	0,78	0,76	18



CONCLUSIONS

SECTION VII

CONCLUSIONS

- The test data is only 18 samples. It is small.
- The classification algorithm who better performs is Logistic Regression and K Nearest Neighbors.
- The worst algorithm is Decision Tree Classifier.
- Launches with high Payload Mass have better results. They have a higher success ratio.
- The success rate of launches increases over the years.
- The launch site KSC LC – 39A has the highest success rate.

THANK YOU!