DATA SCIENCE

Project Proposal

Space X Falcon 9 First Stage Landing Prediction

TABLE OF CONTENTS

01

Summary

PROJECT Summary

- Data Collection
- Data Wrangling
- EDA with Data Visualisation
- EDA with SQL
- Interactive Map using Folium
- Dashboard with Plotly Dash
- Predictive Models
- EDA Results and Conclusions

02

Introduction & methodology

INTRODUCTION

In this Analysis, we will predict if the Falcon 9 first stage will land successfully. SpaceX advertises Falcon 9 rocket launches on its website, with a cost of 62 million dollars; other providers cost upward of 165 million dollars each, much of the savings is because SpaceX can reuse the first stage.

ABOUT THE PROJECT

The commercial space age is here, companies are making space travel affordable for everyone. Virgin Galactic is providing suborbital spaceflights. Rocket Lab is a small satellite provider. Blue Origin manufactures sub-orbital and orbital reusable rockets. Perhaps the most successful is SpaceX. SpaceX's accomplishments include: Sending spacecraft to the International Space Station. Starlink, a satellite internet constellation providing satellite Internet access. Sending manned missions to Space. One reason SpaceX can do this is the rocket launches are relatively inexpensive. SpaceX advertises Falcon 9 rocket launches on its website with a cost of 62 million dollars; other providers cost upwards of 165 million dollars each, much of the savings is because SpaceX can reuse the first stage. **Therefore, if we can determine if the first stage will land, we can determine the cost of a launch.**

Methodology- Data Collection

We collected data from 2 sources

- SpaceX REST API – This dataset gave us data about rocket used in launch, payload delivery, landing outcome etc.
    - https://github.com/Shubham2376G/trainingIBM/blob/main/jupyter_labs_spacex_data_collection_api.ipynb
- Wikipedia – Another popular data source we used to scrap Falcon 9 launch data using BeautifulSoup.
    - https://github.com/Shubham2376G/trainingIBM/blob/main/jupyter_labs_webscraping.ipynb

Methodology- Data Wrangling

https://github.com/Shubham2376G/trainingIBM/blob/main/labs_jupyter_spacex_Data_wrangling.ipynb

1st step

Check null values

2nd step

Calculates no. of launches on each site

3rd step

Calculate no. of occurrance of each orbit

Handled Null values

4th step

Create a landing outcome label from “Outcome” column

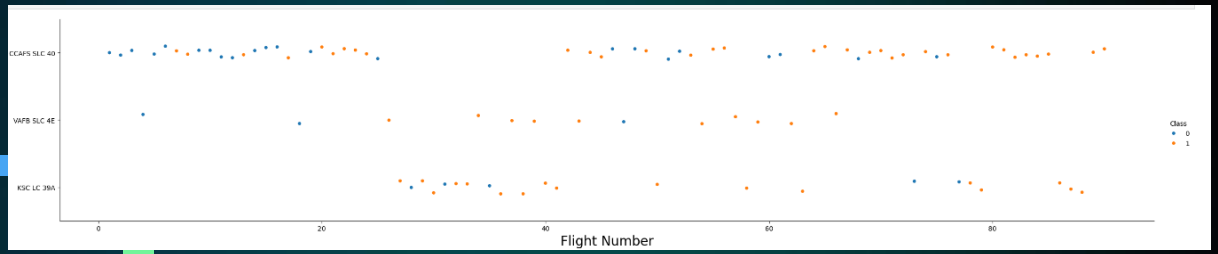
Methodology-EDA with SQL

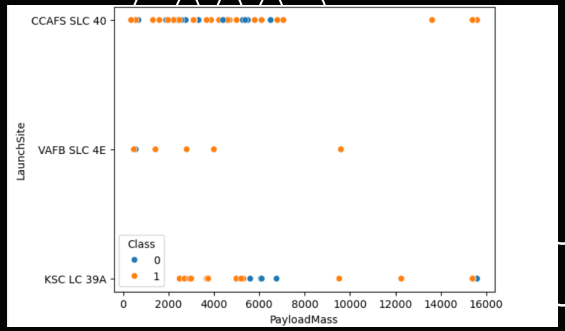
SQL queries performed includes

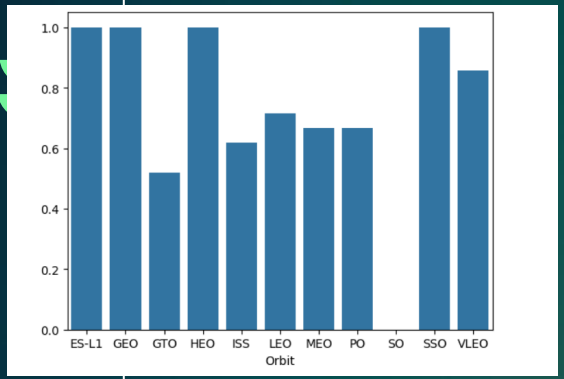
- Display the names of the unique launch sites in the space mission
- Display 5 records where launch sites begin with the string 'CCA'
- Display the total payload mass carried by boosters launched by NASA (CRS)
- Display average payload mass carried by booster version F9 v1.1
- List the date when the first succesful landing outcome in ground pad was acheived.
- List the names of the 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 the names of the booster_versions which have carried the maximum payload mass
- 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.

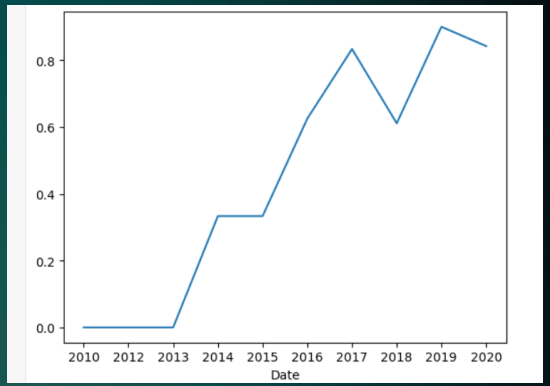
https://github.com/Shubham2376G/trainingIBM/blob/main/jupyter_labs_eda_sql_coursera_sqllite.ipynb

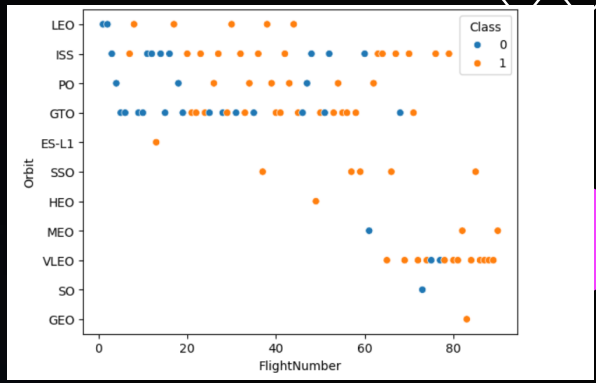
Methodology-EDA with Visualisation



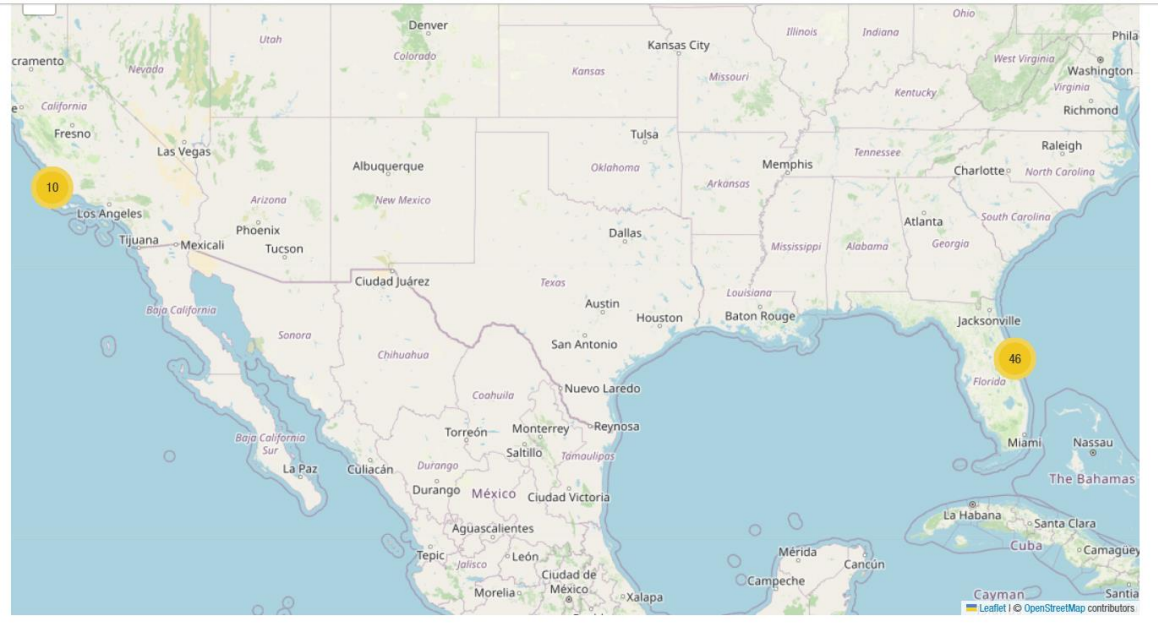






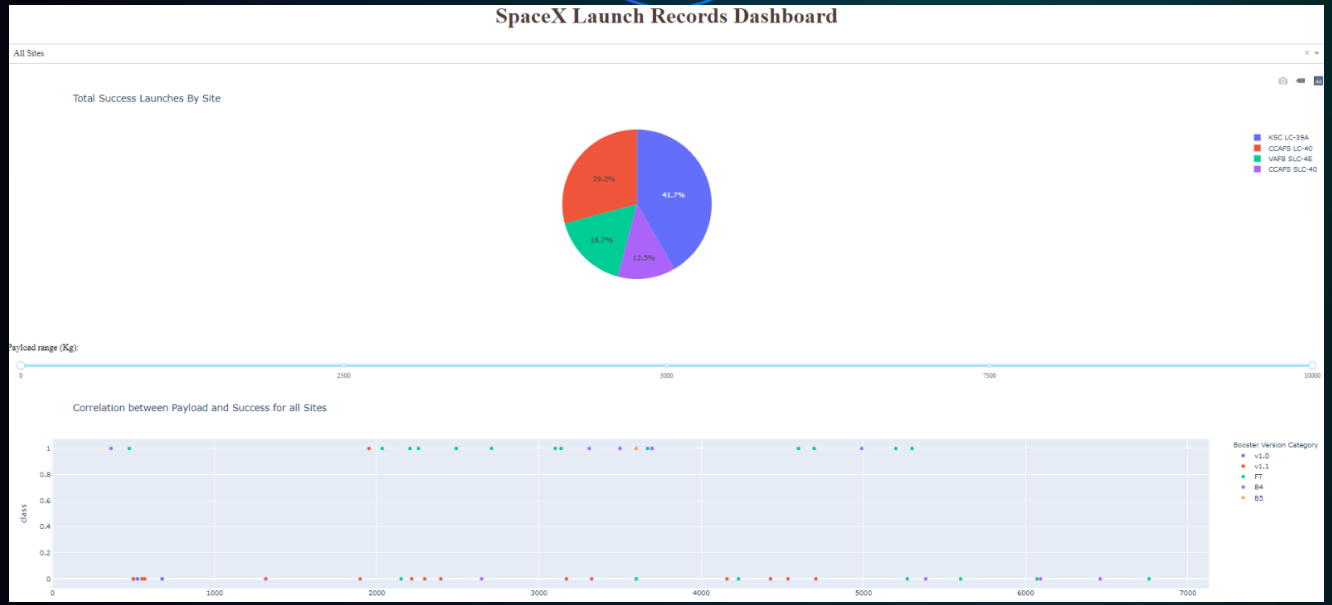


https://github.com/Shubham2376G/trainingIBM/blob/main/jupyter-labs-eda-dataviz.ipynb.jupyterlite.ipynb



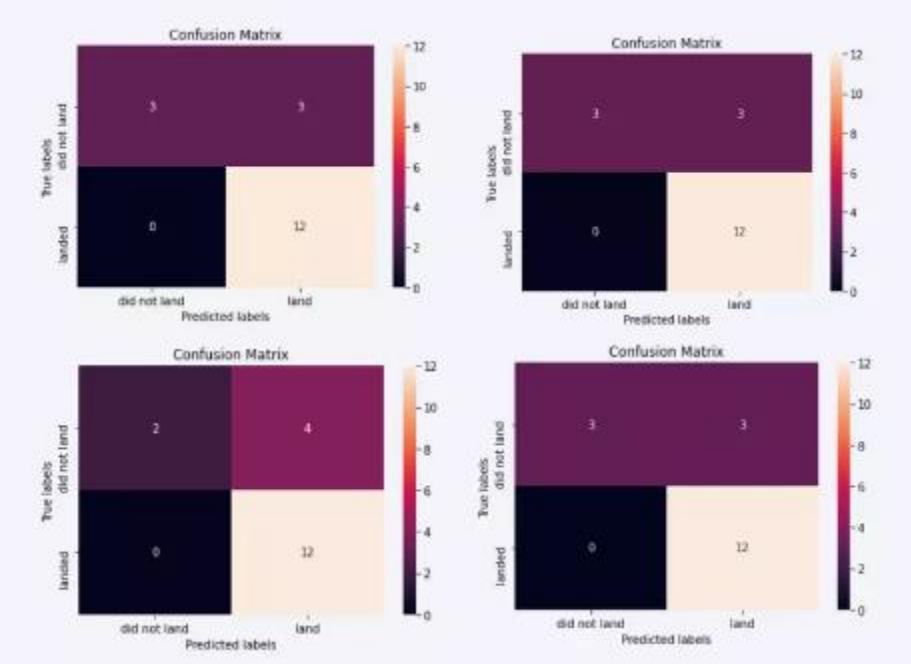
Folium Map

https://github.com/Shubham2376G/trainingIBM/blob/main/lab_jupyter_launch_site_location.jupyterlite_folium.ipynb



Plotly Dashboard

Methodology-Predictive Classification



https://github.com/Shubham2376G/trainingIBM/blob/main/SpaceX_Machine_Learning_Prediction_Part_5_jupyterlite.ipynb

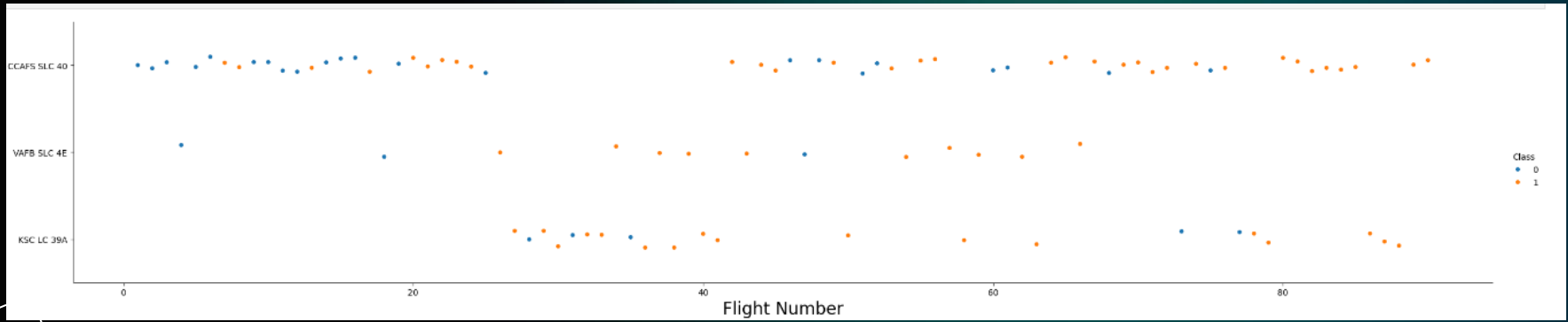
03

Results

Results:

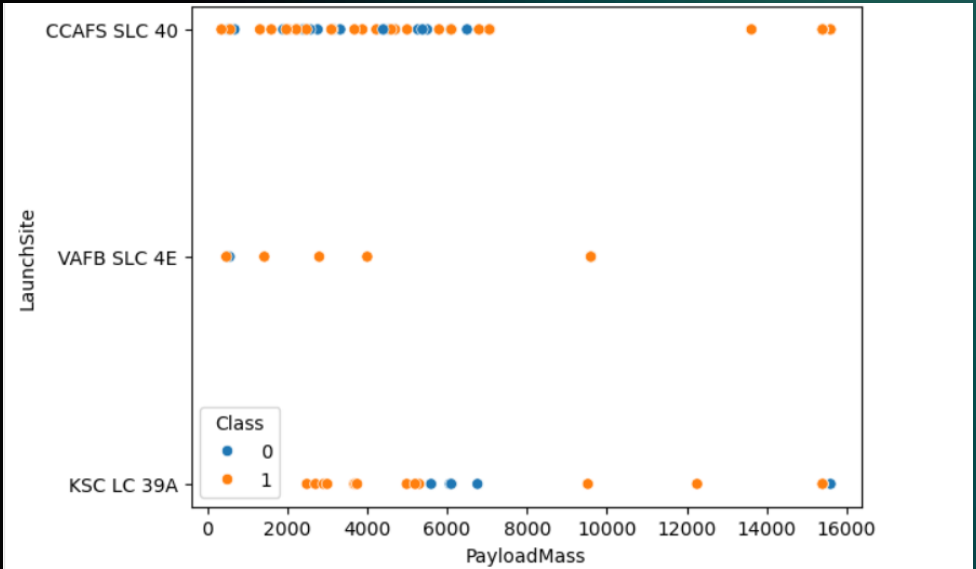
- Low Weighted payloads have more success chances than higher weighted payloads
- Success rate for spaceX launches is directly proportional to time (in year).
- KSC LC 39A launch site had the most success rate
- Orbit GEO, SSO,HEO,ES L1 has the most success rate
- SVM model has the Highest prediction accuracy as compared to other predictive models

Results:



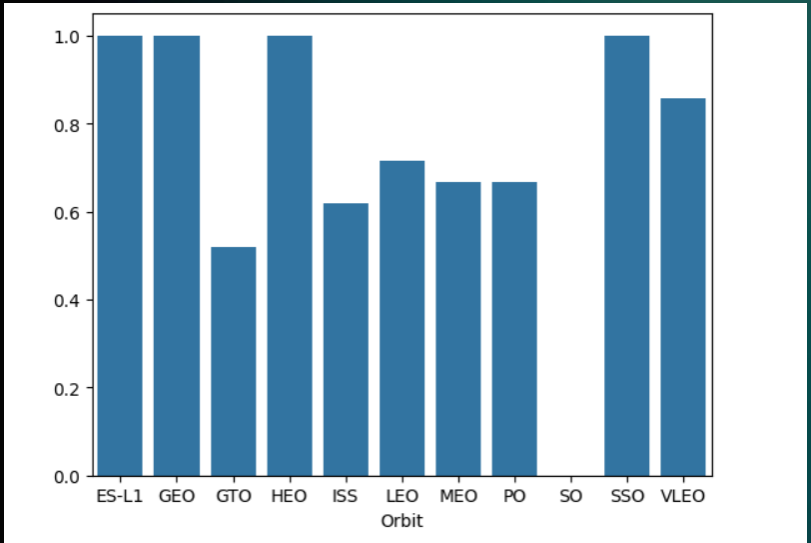
- There is higher number of launches from site CCAFS SLC 40 as compared to other sites

Results:



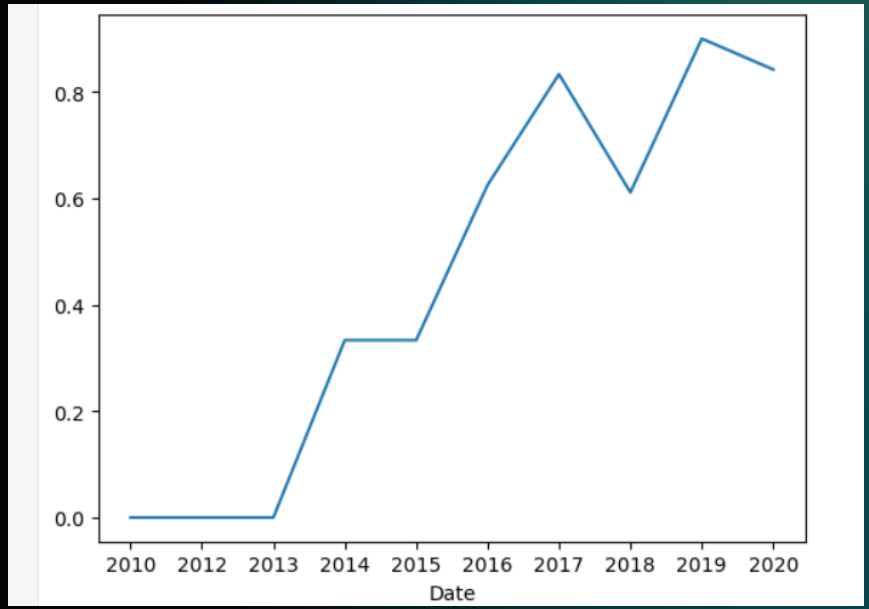
- Lower payloads shows high success rate at all the sites

Results:



- Orbit GEO, SSO,HEO,ES L1 has the most success rate

Results:



- Success rate for spaceX launches is directly proportional to time (in year)

Results:

- Display the names of the unique launch sites in the space mission

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

Results:

- Display 5 records where launch sites begin with the string '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	07:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525	LEO (ISS)	NASA (COTS)	Success	No attempt
2012-10-08	00:35:00	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

Results:

- Display the total payload mass carried by boosters launched by NASA (CRS)
  - 45596
- Display average payload mass carried by booster version F9 v1.1
  - 2928.40
- List the date when the first succesful landing outcome in ground pad was acheived.
  - 2015-12-22
- List the total number of successful and failure mission outcomes
  - 100

Results:

- List the names of the boosters which have success in drone ship and have payload mass greater than 4000 but less than 6000

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

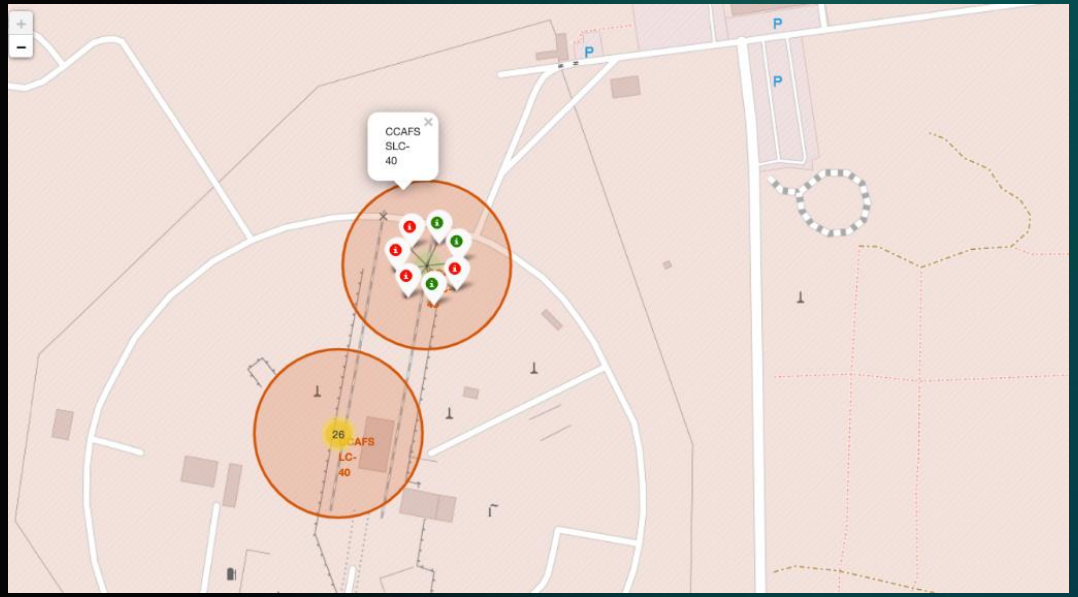
Results:

- List the names of the booster_versions which have carried the maximum payload mass

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

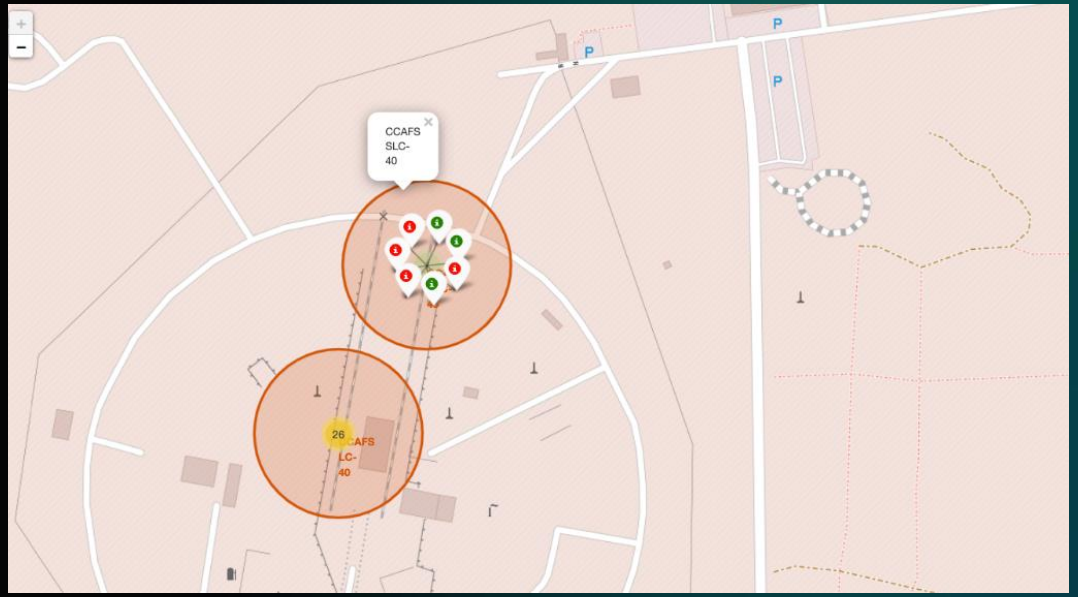
Results:

- Folium color marked all launches



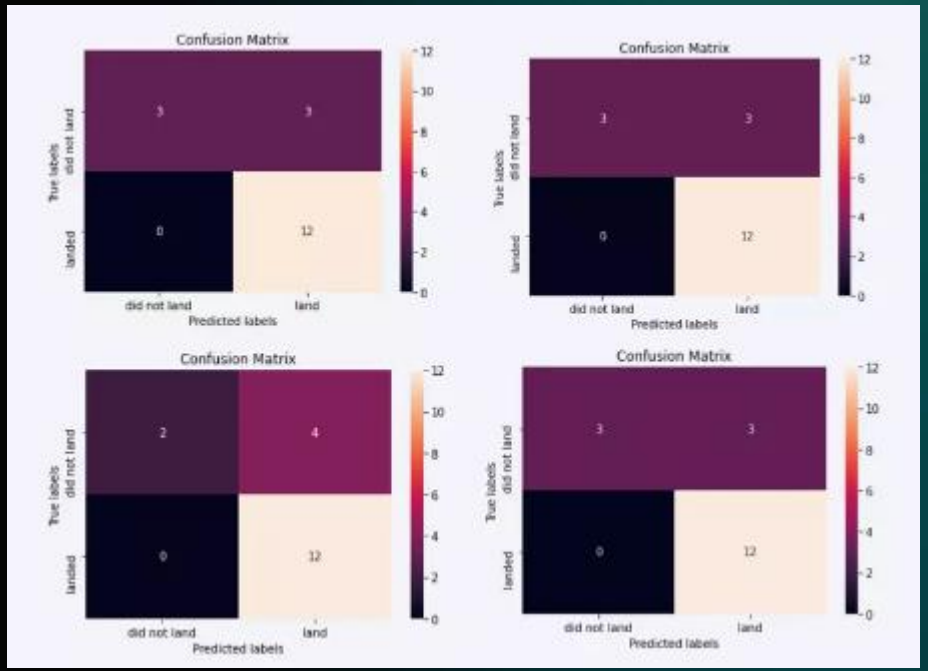
Results:

- Folium color marked all launches



Results:

- Confusion Matrix : All models have almost same accuracy



Conclusion

- SVM model has the Highest prediction accuracy as compared to other predictive models
- Low Weighted payloads have more success chances than higher weighted payloads
- Success rate for spaceX launches is directly proportional to time (in year).
- KSC LC 39A launch site had the most success rate
- Orbit GEO, SSO,HEO,ES L1 has the most success rate