IBM Developer SKILLS NETWORK

Winning Space Race with Data Science

MSc Alexandre Gil Moreno
19 October 2021

Outline

EXECUTIVE SUMMARY

INTRODUCTION

METHODOLOGY

RESULTS

CONCLUSION

APPENDIX

Executive Summary

Project Mission

- Predict if Falcon 9 first stage will land successfully and hence its launch cost.

Data Collection

- Extract data from SpaceX API (Booster Version, Launch Site, Outcome, etc).
- Format and Clean data (E.g. Substituting Payload mass null values by its mean value).

Exploratory Data Analysis

- Create training labels by Outcome , where successful landing (class = 1) and failed landing (class = 0).
- Check success rate by launch site, payload mass, orbit type, booster version, launch attempts, etc.

Highlights

- Perform predictive analysis using different classification models, where Tree classification performs the best.
- Payload mass, launch site and orbit type have a huge influence on success rate.
- Overall 66% success rate with a yearly increasing trend.

Introduction

SPACEX PERFORMS CONTROLLED LANDINGS. IF LANDING IS SUCCESSFUL, FIRST STAGE OF THEIR ROCKETS CAN BE REUSED AND HENCE THEY SAVE AROUND 100 MILLION DOLLARS COMPARED TO OTHER PROVIDERS.

OUR GOAL IS TO UNDERSTAND UNDER WHICH CONDITIONS FALCON 9 WILL LAND SUCCESSFULLY AND HENCE DETERMINE ITS LAUNCH COST.

Section 1

Methodology

Methodology

Data collection methodology:

Request rocket launch data from SpaceX API and create a dataframe including the rocket type, its payload, launchpad and cores.

Web scrap Falcon 9 launch records stored in a HTML table from Wiki.

Perform data wrangling

Parse dataframe and filter it to only include Falcon 9 launches information.

Deal with missing/null values from the Payload mass column by substituting them by column mean value.

Perform exploratory data analysis (EDA) using visualization and SQL

Perform interactive visual analytics using Folium and Plotly Dash

Perform predictive analysis using classification models

Standardize data and split it into training (80%) and testing (20%) data.

Using *sklearn* create and train using the training data different classification models such as logistic regression, SVM, decision tree and K-neighbors.

Perform cross-validation using *GridSearchCV* and tune the model with the best parameters.

Estimate the accuracy of the model and check its confusion matrix.

SpaceX API url: https://api.spacexdata.com/v4/launches/past
Wikipedia url: https://en.wikipedia.org/wiki/List_of_Falcon_9_and_Falcon_Heavy_launches

Data Collection

SpaceX API

- Request launch data
- Extract relevant data for the project
- Store data into a new dataframe

Web Scraping

- Request Falcon 9 launch data from Wiki HTML page
- Extract tables
- Create dataframe with relevant data from the table

Data Collection – SpaceX API

SpaceX REST API

- Request data and decode to JSON file:

```
import requests
response = requests.get(url).json()
```

Convert to Dataframe

```
import pandas as pd
data = pd.json_normalize(response)
```

Get relevant information

```
data = data[['rocket','payloads','launchpad','cores','flight_number','date_utc']]
```

Extract info about launches

- Use helper functions to extract info using IDs
- Store info into a dictionary launch_dict

Create new dataframe

```
df = pd.DataFrame(launch_dict)
```

Github URL: https://github.com/xelivy/CapstoneProject/blob/80f052676ff3629f3bcfeecc45d1b808e37c59e6/Data%20Collection%20API%20Lab.ipynb

Data Collection – Scraping

Launch records on Wiki

• Request the Falcon9 launch Wiki page from its URL:

```
import requests
response = requests.get(url)
html = response.text
```

Create BeautifulSoup object

```
from bs4 import BeautifulSoup
soup = BeautifulSoup(html, 'html5lib')
```

Extract column name from HTML table

```
html_tables = soup.find_all('table')
columns = html_tables[2].find_all('th')
column_names.append(extract_column_from_header(col_name )) for col_name in columns
```

Extract info about launches

- Create a dictionary launch_dict with column names as keys
- Parse table content and fill up the previous dictionary with relevant data about Falcon9 launches

Create new dataframe

```
df = pd.DataFrame(launch_dict)
```

Github URL: https://github.com/xelivy/CapstoneProject/blob/7035382f8ef50dbc92fc4dd8fd6ec1136c46b2b4/Data%20Collection%20with%20Web%20Scraping.ipynb

Data Wrangling

SpaceX API dataframe

Filter dataframe:

•Only include Falcon9 launches info

```
data_falcon9 = df[df['BoosterVersion'] != 'Falcon 1']
```

Check missing values

```
data_falcon9.isnull().sum()
```

Replace missing values by column mean value

```
mean=data_falcon9['PayloadMass'].mean()
data_falcon9['PayloadMass']=data_falcon9['PayloadMass'].replace(np.nan,mean)
```

Calculate the number and occurrence of:

- Launch Sites
  `df['LaunchSite'].value_counts()`
- Orbit types
  `df['Orbit'].value_counts()`
- Mission Outcome
  `landing_outcomes=df['Outcome'].value_counts()`

Create label/column called 'Class' for successful and failed mission

Failed mission:

```
If outcome in bad_outcomes:
    landing_class.append(int(0))
df['Class'] = landing_class
```

Successful mission:

```
else:
    landing_class.append(int(1))
df['Class'] = landing_class
```

Github URL: https://github.com/xelivy/CapstoneProject/blob/093c5e6515b2a8c42d7049eb8805789726b2aaab/Data%20Wrangling.ipynb

EDA with Data Visualization

- Chart summary:
  I. Flight Number vs Launch Site

     See how the launch attempts would affect the launch outcome with respect to the launch site.

  II. Payload mass vs Launch Site

     Check if there is any relationship between launch sites and their payload mass.

  III. Success rate vs Orbit type

     Visually check the relationship between the success rate of each orbit type.

  I. Flight Number vs Orbit type

     Check if there is any relationship between launch attempts and orbit type and how this would affect the launch outcome.

  II. Payload mass vs Orbit type

     Reveal the relationship between payload mass and orbit type and how this affects the launch outcome.

  III. Year vs Success rate

     Visualize the launch success rate yearly trend.

Github URL: https://github.com/xelivy/CapstoneProject/blob/28dd89f39420978c106d234466b257192e39277c/EDA%20with%20Visualization.ipynb

EDA with SQL

- SQL queries:
    I. List all the unique launch sites in the SpaceX data set.
    II. Display the first 5 records where launch site begins with 'KSC'.
    III. Calculate the total payload mass carried by boosters launched by NASA (CRS).
    IV. Calculate the average payload mass carried by booster version F9 v1.1 .
    V. List the date where the first successful landing outcome in drone ship was achieved.
    VI. List the booster names which landed successfully in ground pad and have payload mass between 4000 and 6000.
    VII. Count and list the total number of successful and failure mission outcomes.
    VIII. List the names of the booster versions which carried the maximum payload mass in ascending order.
    IX. Display the month names, successful landing outcomes in ground pad, booster versions, launch sites records in 2017,
    X. Order by the number of successful landing outcomes between the date 2010-06-04 and 2017-03-20 in descending order.

Github URL: https://github.com/xelivy/CapstoneProject/blob/c017987ffaf3c796336da75452f625b14364a430/EDA%20with%20SQL%20.ipynb

Build an Interactive Map with Folium

- Folium map items:
    1. We located on the map and added a circle in each launch site together with a marker containing the name of the launch site.
    2. In each launch site we added a marker cluster gathering all landing outcomes (green = success and red = failure) per each site in order to easily identify which place has the higher success rate.
    3. For the launch site with the best success rate, we draw a line towards the closest coastline, railway, highway and city with markers indicating the distance in kms in order to determine what could be an optimal location parameters by looking at its proximities.

Github URL: https://github.com/xelivy/CapstoneProject/blob/3d2bc81b1c75e8d9620d199cea6d6582ec6cbd15/Interactive%20Visual%20Analytics%20with%20Folium.ipynb

Build a Dashboard with Plotly Dash

- Dashboard items:
  - ➢ Dropdown bar to select a single launch site or all launch sites together.
    - ➢ After selection, it displays a **pie chart** showing the total success rate by site (with 0 as a failure and 1 as success) or the total success rate for all sites.
    - ➢ Check the success rate per site and check the success rate between sites.
  - ➢ Range slider for payload mass to select the range we want to show in the following scatter plot.
    - ➢ An **interactive scatter plot** representing the landing outcome with respect to the payload mass where the color per each point indicates the booster version used for each launch mission.
    - ➢ Check the correlation between the payload mass and the success per all sites and per each booster version.

Github URL: https://github.com/xelivy/CapstoneProject/blob/c774a906792a61ed42c5f575e1df82dd4e319041/spacex_dash_app.py

Predictive Analysis (Classification)

Load data

- Define Y as 'Class' column — `Y = pd.DataFrame(data['Class'].to_numpy())`
- Standardized and transformed data(X) into a suitable form — `X = preprocessing.StandardScaler().fit(X).transform(X)`

Split data into random train and test subsets

```
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size = 0.2, random_state = 2)
```

Build a classifier algorithm with cross-validation (E.g. logistic regression, SVM, tree and K-neighbors)

E.g.

```
lr = LogisticRegression()
logreg_cv = GridSearchCV(lr, cv = 10, param_grid = parameters)
```

Fit model using train subset and find the best parameters

```
logreg_cv.fit(X_train, Y_train)
logreg_cv.best_params_
```

Run predictions using test subset

```
yhat = logreg_cv.predict(X_test)
```

Evaluate the model

- Accuracy: `logreg_cv.score(X_test, Y_test)`
- Confusion Matrix `confusion_matrix(Y_test, yhat)`

Github URL: https://github.com/xelivy/CapstoneProject/blob/f7ddccd2a0469b079d7fdab4b16340e6b160fa4d/Machine%20Learning%20Prediction%20lab.ipynb

Results

Exploratory data analysis results

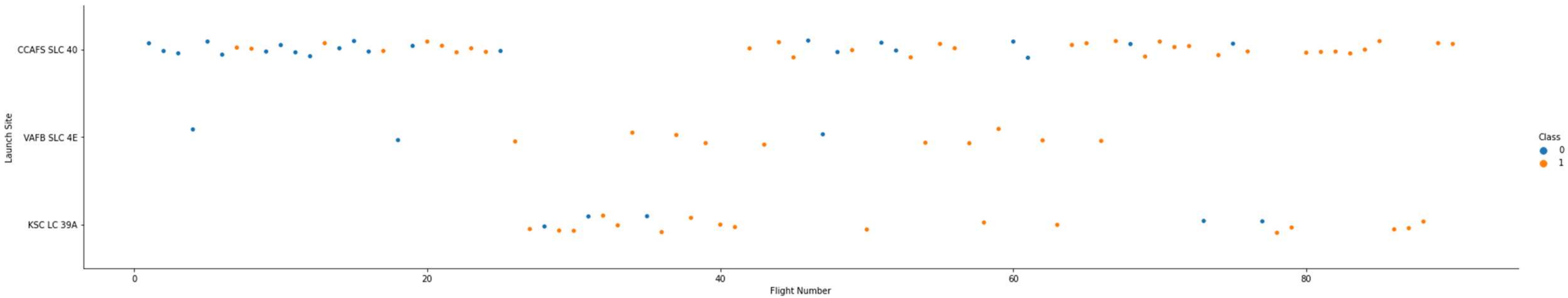
Interactive analytics demo in screenshots

Predictive analysis results

Section 2

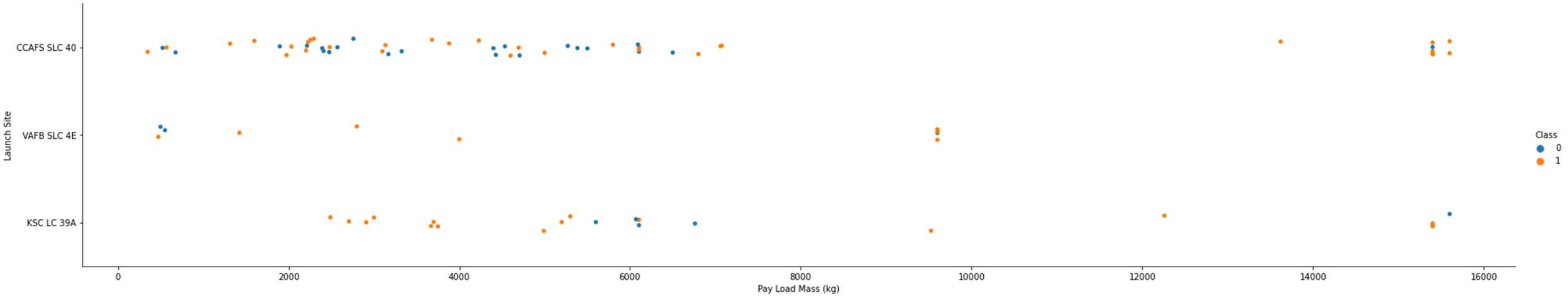
Insights drawn from EDA

Flight Number vs. Launch Site



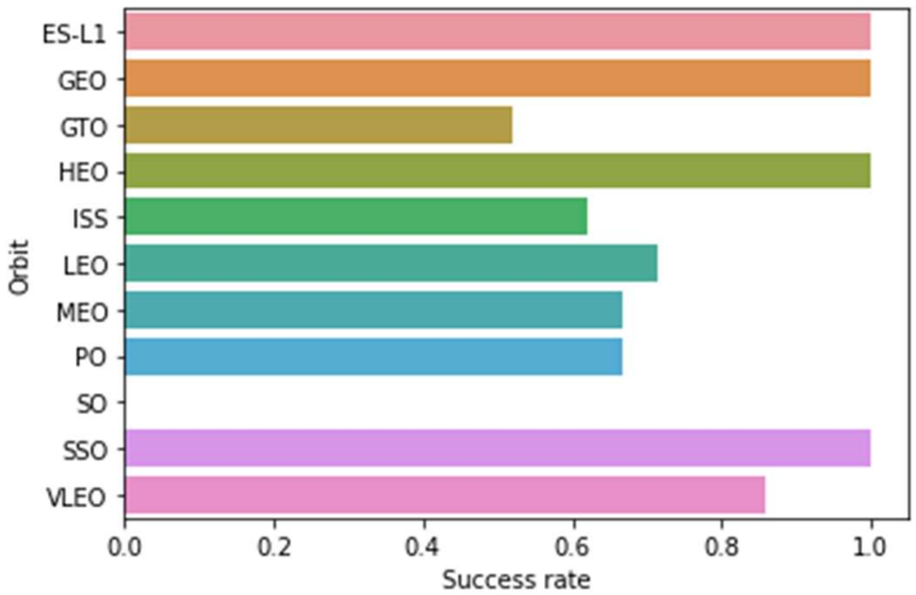
- Two trends can be observed:
  - We observe an increasing landing success for all launch sites, the more launches are performed.
  - Different launch sites have different success rate. CCAFS SLC-40, has a 60% success rate, while KSC LC 39A and VAFB SLC 4E have a 77% success rate.

Payload vs. Launch Site



- Observations:
  - Overall, we observe that more launches were performed by lower payload masses. And also, that higher payload masses have more success rate.
  - Specifically, CCAFS SLC 40 launches have a very low success rate for lighter payloads with respect to the other two launches sites, while KSC LC 39A has a region between 5500 to 7000 kg where it is less likely to succeed in.
  - Finally, VAFB SLC 4E has an overall a higher successful landing outcome rate. However, there have not perform any launch with a payload mass above 1000 kg.

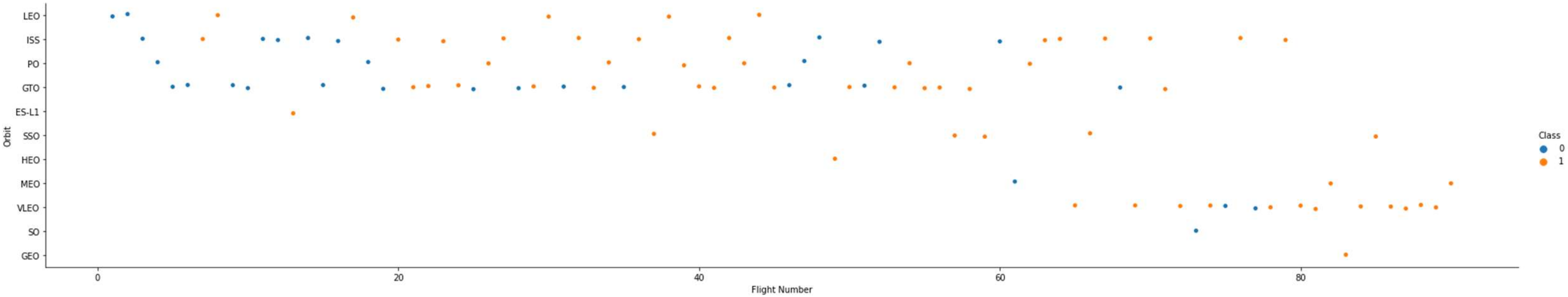
Payload vs. Launch Site



➢ Observations:

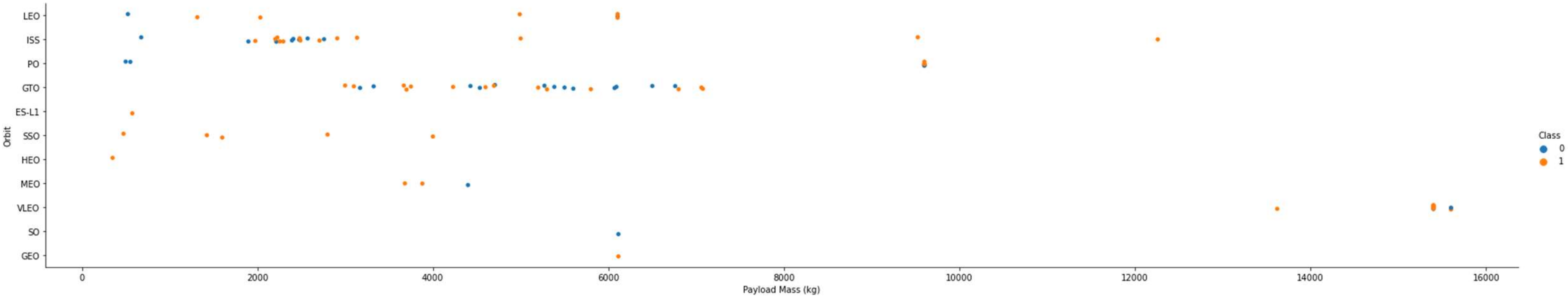
- ➢ Mission in ES-L1, GEO, HEO and SSO orbits have a 100% chance to land successfully, whereas SO orbit has a 0% success rate.
- ➢ Other orbits like GTO has a 50% success rate while the rest has a success rate above 60% (e.g., VLEO orbits has a success rate around 85%)

Flight Number vs. Orbit Type



- **Observations:**
  - LEO orbit launches seem to have an increasing success with increasing the flight number, whereas GTO orbit missions seem to not have such relation since failures can be seen either at higher flight numbers or lower flight numbers.
  - ES-L1, SO, HEO and SO orbits have only one single launch attempt. Therefore, from such missions we cannot extract any relevant insights. Similarly, MEO orbit missions have only three attempts, which is also rather small to gain further insights.
  - Finally, SSO and VLEO orbit missions have an overall high success rate with a considerable high number of attempts.

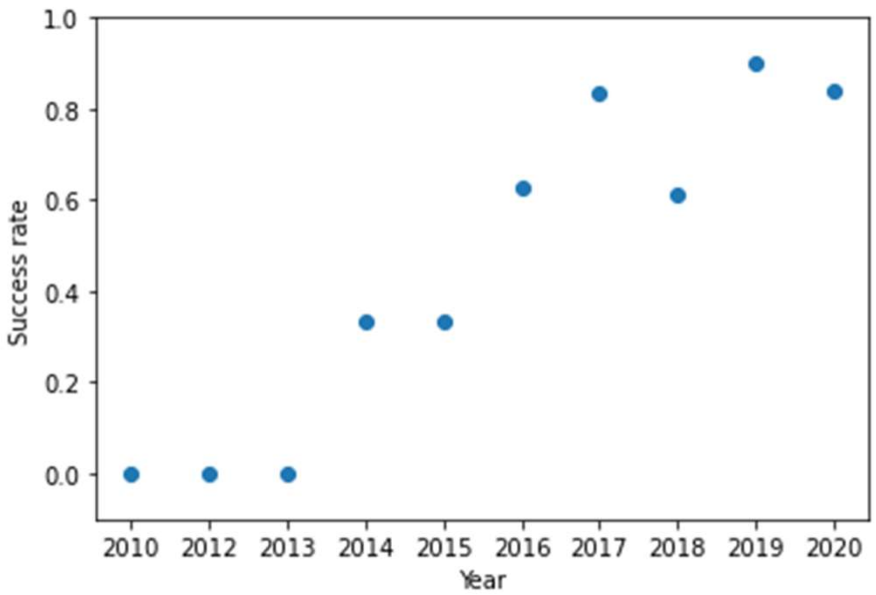
Payload vs. Orbit Type



➢ Observations:

- ➢ LEO and ISS orbit launches success seems to be positively related for heavy payload, whereas GTO orbit success seems to be negative correlated.
- ➢ SSO orbit launches have been performed for light payloads with an impressive 100% success rate, while VLEO orbit launches have been performed at extremely heavy payloads with also an impressive high success rate.

Launch Success Yearly Trend



➢ Note:

- ➢ There has been a continuous increase on the success rate since 2013 until 2020.
- ➢ However, the success curve seems to flattened since 2017 with a success rate above 0.8.

All Launch Site Names

- SQL query:

```
%sql SELECT DISTINCT(LAUNCH_SITE) FROM SPACEXDATASET
```

- Outcome:

Out[5]:

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

Comments:

- Four different launch sites are displayed by selecting all the unique launch sites from the SpaceX data set given.

Launch Site Names Begin with ‘KSC’

- SQL query:

```
%sql SELECT * FROM SPACEXDATASET WHERE LAUNCH_SITE LIKE 'KSC%' LIMIT 5
```

- Outcome:

Out[10]:

DATE	time__utc_	booster_version	launch_site	payload	payload_mass__kg_	orbit	customer	mission_outcome	landing__outcome
2017-02-19	14:39:00	F9 FT B1031.1	KSC LC-39A	SpaceX CRS-10	2490	LEO (ISS)	NASA (CRS)	Success	Success (ground pad)
2017-03-16	06:00:00	F9 FT B1030	KSC LC-39A	EchoStar 23	5600	GTO	EchoStar	Success	No attempt
2017-03-30	22:27:00	F9 FT B1021.2	KSC LC-39A	SES-10	5300	GTO	SES	Success	Success (drone ship)
2017-05-01	11:15:00	F9 FT B1032.1	KSC LC-39A	NROL-76	5300	LEO	NRO	Success	Success (ground pad)
2017-05-15	23:21:00	F9 FT B1034	KSC LC-39A	Inmarsat-5 F4	6070	GTO	Inmarsat	Success	No attempt

Comments:

- We displayed only the first 5 rows of the data set that contain ‘KSC’ at the beginning of the launch_site column.

Total Payload Mass

- SQL query:

```
%sql SELECT SUM(PAYLOAD_MASS__KG_) AS TOTAL_PAYLOAD_MASS FROM SPACEXDATASET WHERE CUSTOMER = 'NASA (CRS)'
```

- Outcome:

Out[15]:

total_payload_mass
45596

Comments:

- We sum up over the payload_mass__kg_ column the values that have the NASA (CRS) as a booster provider/customer.
- The total payload is 45596 kg.

Average Payload Mass by F9 v1.1

- SQL query:

```
%sql SELECT AVG(PAYLOAD_MASS__KG_) AS AVG_PAYLOAD_MASS FROM SPACEXDATASET WHERE BOOSTER_VERSION = 'F9 v1.1'
```

- Outcome:

Out[19]:

avg_payload_mass
2928

Comments:

- We average over the payload_mass__kg_ column of the launches that had a booster F9 v1.1.
- The average payload carried is 2928 kg.

First Successful Drone Ship Landing Date

- SQL query:

```
%sql SELECT MIN(DATE) AS DATE_SUCCESSFUL_LANDING FROM SPACEXDATASET WHERE LANDING__OUTCOME = 'Success (drone ship)'
```

- Outcome:

Out[25]:

date_successful_landing
2016-04-08

Comments:

- We select the first date (minimum) that has as a landing outcome in the landing__outcome column 'Success (drone ship)'.
- The first successful drone ship landing was on April 8th 2016.

Successful Ground Landing with Payload between 4000 and 6000

- SQL query:

```
%sql SELECT BOOSTER_VERSION,LANDING__OUTCOME,PAYLOAD_MASS__KG_ FROM SPACEXDATASET WHERE (LANDING__OUTCOME = ‘Success (ground pad)’ AND
PAYLOAD_MASS__KG_ BETWEEN 4000 AND 6000)
```

- Outcome:

Out[32]:

booster_version	landing__outcome	payload_mass__kg_
F9 FT B1032.1	Success (ground pad)	5300
F9 B4 B1040.1	Success (ground pad)	4990
F9 B4 B1043.1	Success (ground pad)	5000

Comments:

- We displayed all values from the booster_version, landing__outcome and payload_mass__kg_ columns from which the dataset fulfils the two conditions.
    1. Landing outcome = Success (ground pad)
    2. 4000 < Payload mass < 6000
- Only 3 launches fulfil these conditions.

Total Number of Successful and Failure Mission Outcomes

- SQL query:

%sql SELECT COUNT(LANDING__OUTCOME) AS SUCCESSFUL_MISSION, (SELECT COUNT(LANDING__OUTCOME) AS FAIL_MISSION FROM SPACEXDATASET WHERE LANDING__OUTCOME LIKE 'Failure %') FROM SPACEXDATASET WHERE LANDING__OUTCOME LIKE 'Success %'

- Outcome:

Out[40]:

successful_mission	fail_mission
23	7

Comments:

- We display a table that counts all successful and failed launches by using a subquery for the second column.
- Results show that there are more successful missions rather than failures.

Boosters Carried Maximum Payload

- SQL query:

```
%sql SELECT DISTINCT(BOOSTER_VERSION), PAYLOAD_MASS__KG_ FROM SPACEXDATASET WHERE (SELECT MAX(PAYLOAD_MASS__KG_) FROM SPACEXDATASET) ORDER BY PAYLOAD_MASS__KG_ DESC
```

- Outcome:

Out[52]:

booster_version	payload_mass__kg_
F9 B5 B1048.4	15600
F9 B5 B1048.5	15600
F9 B5 B1049.4	15600
F9 B5 B1049.5	15600
F9 B5 B1049.7	15600
F9 B5 B1051.3	15600
F9 B5 B1051.4	15600
F9 B5 B1051.6	15600
F9 B5 B1056.4	15600
F9 B5 B1058.3	15600
F9 B5 B1060.2	15600
F9 B5 B1060.3	15600

Comments:

- We display a table that shows the maximum payloads carried for all unique boosters versions ordered in ascending order.

Note:

- The list is longer but this list only shows the booster versions that have the heaviest payload mass.

2017 Launch Records

- SQL query:

%sql SELECT MONTHNAME(DATE) AS Month, LANDING__OUTCOME AS OUTCOME, BOOSTER_VERSION, LAUNCH_SITE FROM SPACEXDATASET WHERE LANDING__OUTCOME = ‘Success (ground pad)’ AND YEAR(DATE) = ‘2017’

- Outcome:

Out[60]:

MONTH	outcome	booster_version	launch_site
February	Success (ground pad)	F9 FT B1031.1	KSC LC-39A
May	Success (ground pad)	F9 FT B1032.1	KSC LC-39A
June	Success (ground pad)	F9 FT B1035.1	KSC LC-39A
August	Success (ground pad)	F9 B4 B1039.1	KSC LC-39A
September	Success (ground pad)	F9 B4 B1040.1	KSC LC-39A
December	Success (ground pad)	F9 FT B1035.2	CCAFS SLC-40

Comments:

- We list all launches that successfully landed on ground that occurred in 2017 displaying the month of the launch, the mission outcome, the booster version and the launch site.

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

- SQL query:

%sql SELECT LANDING__OUTCOME AS LANDING, COUNT(LANDING__OUTCOME) AS TIMES FROM SPACEXDATASET WHERE LANDING__OUTCOME LIKE 'Success%' AND DATE BETWEEN '2010-06-04' AND '2017-03-20' GROUP BY LANDING__OUTCOME ORDER BY COUNT(LANDING_OUTCOME)

- Outcome:

Out[86]:

landing	times
Success (ground pad)	3
Success (drone ship)	5

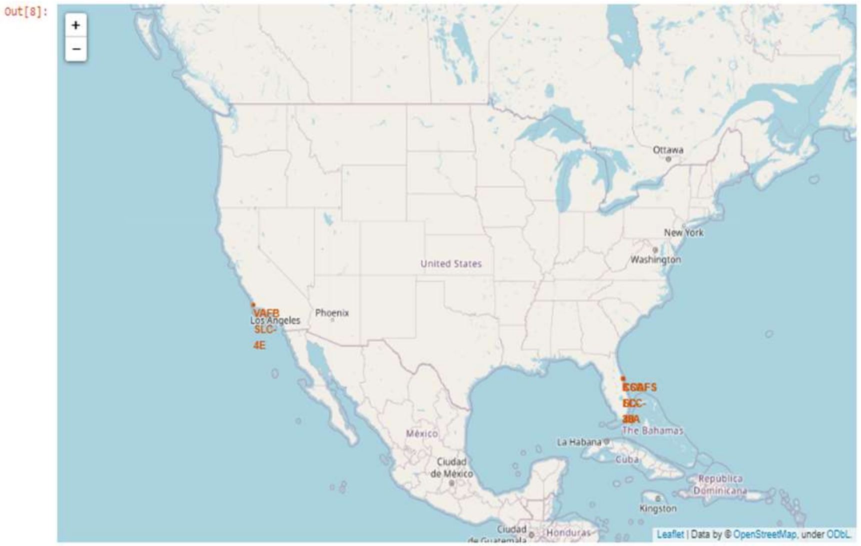
Comments:

- We display a table that shows the number of successful landing per each type of landing between the dates 2010-06-04 and 2017-03-20 in descending order.
- Results show that there were more missions that landed successfully on a drone ship rather than on the ground.

Section 4

Launch Sites Proximities Analysis

Locating All Launch Sites



- Folium map containing all launch sites with an orange circle.

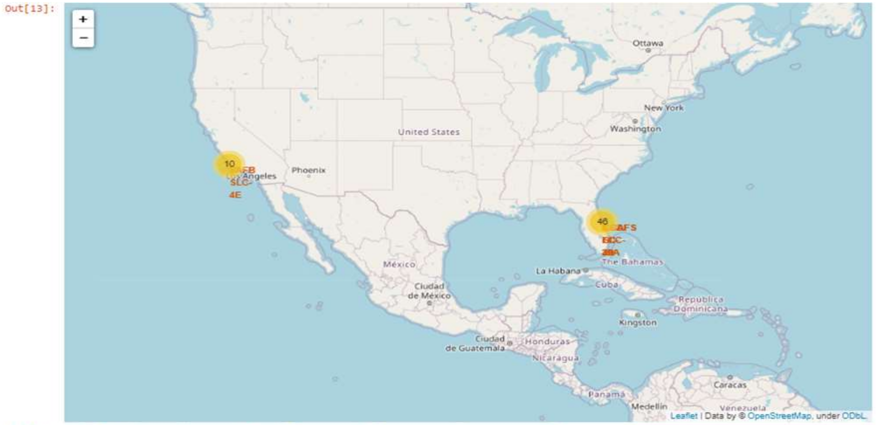
- Highlight:

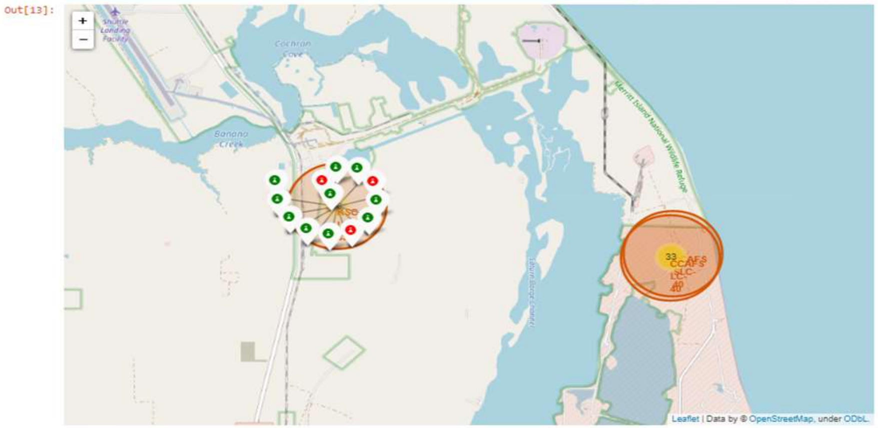
    1. All launch sites are located at two locations in USA that are close to the equatorial line.

    2. VAFB SLC 4E is located at the southwest, whereas KSC LC 39A, CCAFS SLC 40 and CCAFS LC 40 are located to the southeast.

    3. All launch sites are located really close to the coastline.

Success and Failed Launches

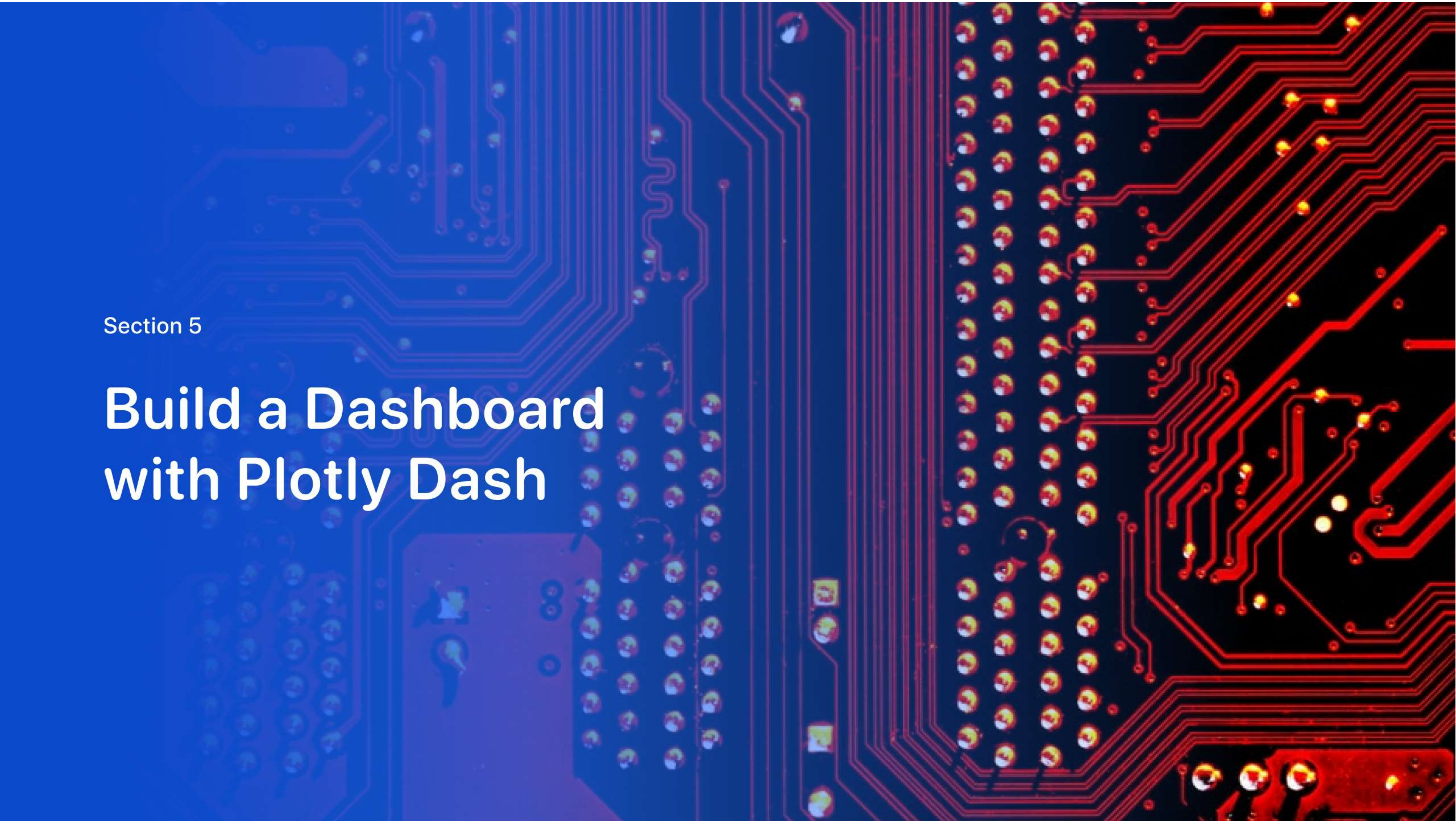




- Per each launch site we add a cluster marker indicating the successful (green) and failed (red) missions.
- Highlight:
    1. The launch site with a greater success rate is located on the east-coast and it is KSC LC 39A.
    2. At the same coastline, CCAFS SLC 40 and CCAFS LC 40 have a rather lower success rate compared to KSC LC 39A.
    3. On the other side, launches in VAFB SLC 4E have a 40% success rate.

Launch Sites and their Proximities

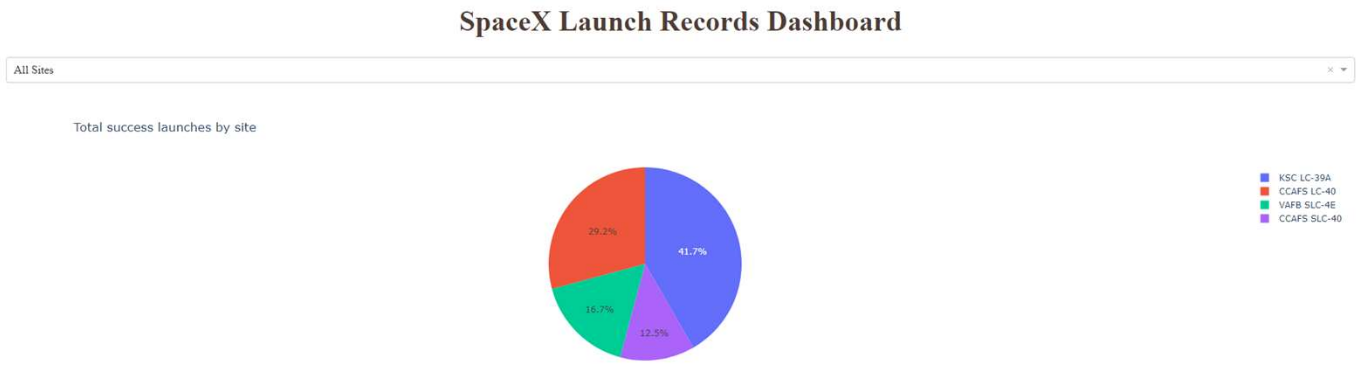
- ➢ We explore the distances of CCAFS-SLC 40 to its proximities using a Folium map.
- ➢ Specifically,
  1. It is closely located to the highway, coastline and train railway, where the highway it is the closest at 0.59 km.
  2. Launch sites in general are located far away from cities as shown in this case where the closest city (Melbourne) is at 51.42 km.



Section 5

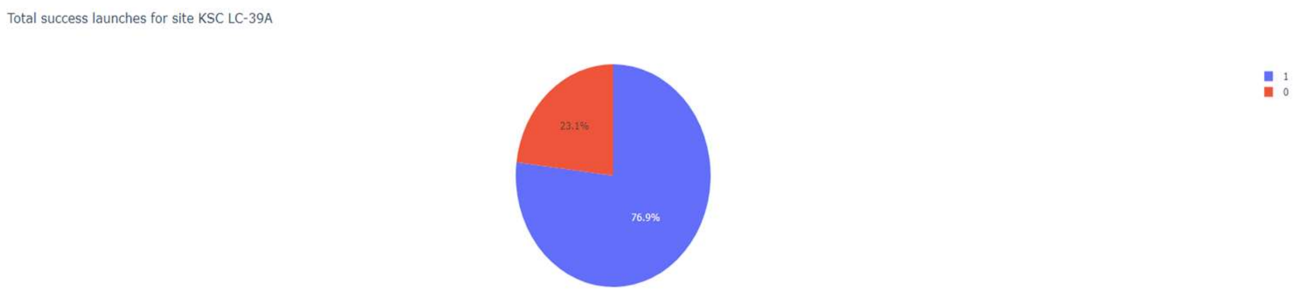
Build a Dashboard with Plotly Dash

Total Success Launch by Site



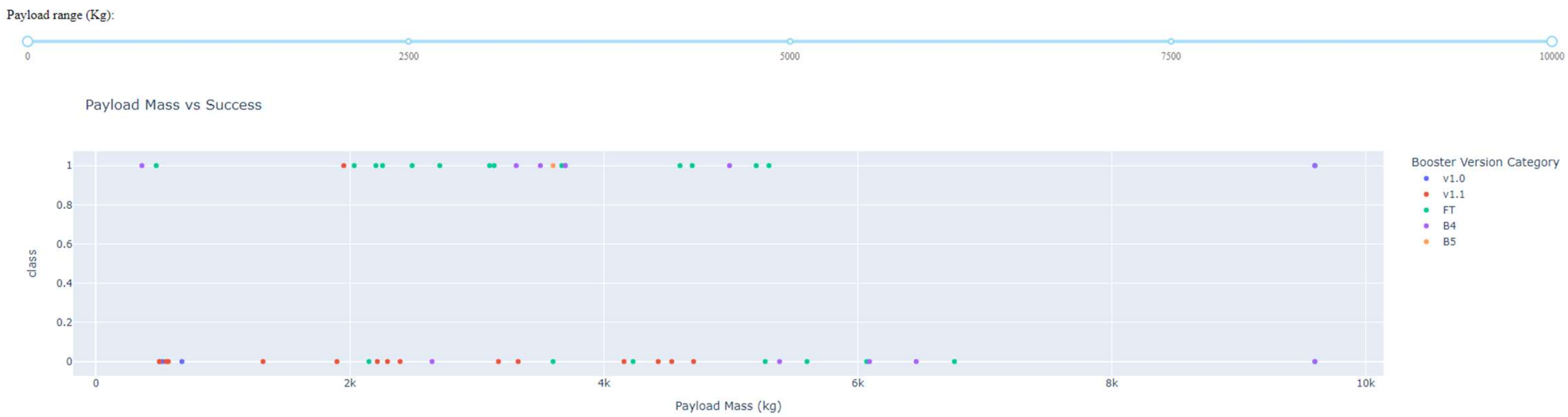
- **Observations:**
  - The launch site with the highest number of successful launches is KSC LC-39A with a 41.7% of the total launches which landed successfully, whereas launches at CCAFS SLC-40 has the lowest with 12.5%.

KSC LC-39A Success ratio



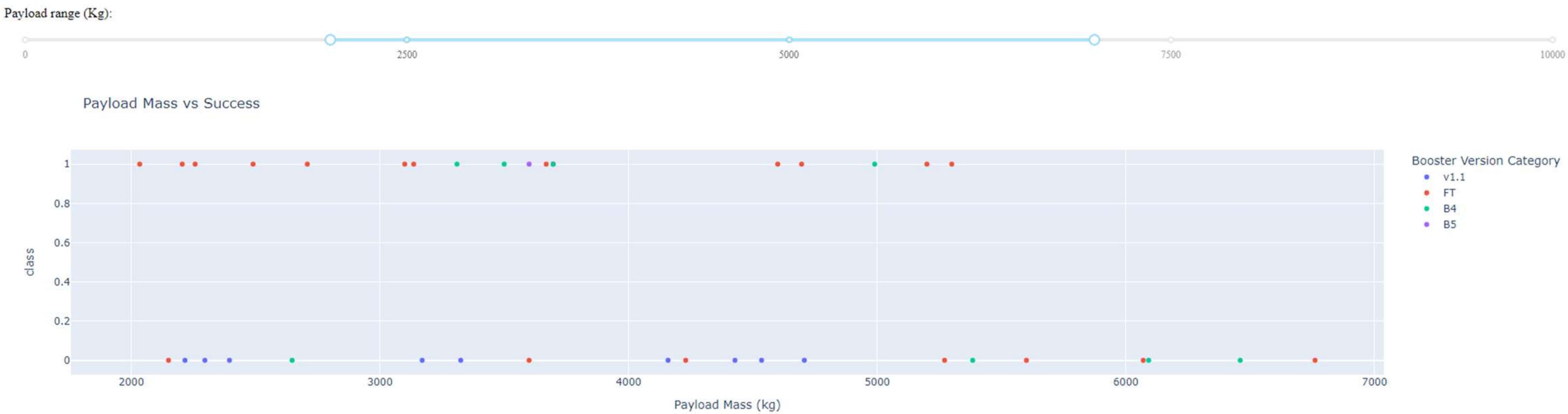
- **Observations:**
  - We notice that launches occurring in KSC LC-39A has overall the largest success rate with a value of 76.9% with a total of 10 launches performed.

Payload vs. Launch Outcome - Overall



- **Observations:**
  - At lighter payloads (below 2000 kg) and at heavier payloads (above 6000 kg) the success rate is not favorable and most landing are not successful.
  - For payloads between 2000 kg and 6000 kg we can observe that successful landing are more likely to occur.

Payload vs. Launch Outcome – Between 2000 kg and 6000 kg

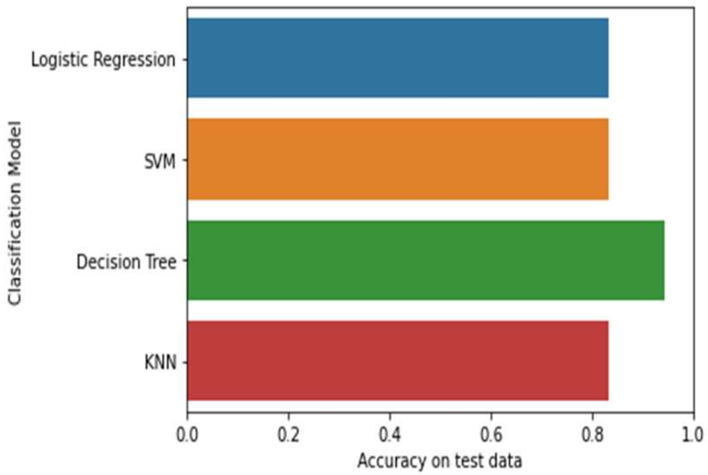


- **Observations:**
  - We observe that the F9 boosters that have the highest success rate are FT boosters, whereas booster's version v1.1 have the lowest success rate.
  - We could also observe that for F9 B4 boosters are more likely to land successfully when payload mass carried is below 4000 kg and less likely when it is above 5000 kg.

Section 6

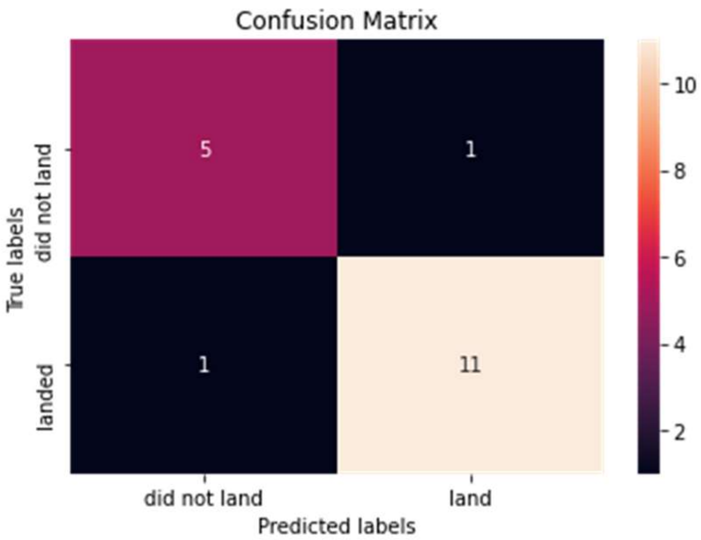
Predictive Analysis (Classification)

Classification Accuracy



- We trained 4 different classification models with a cross-validation check:
  - Logistic regression
  - SVM
  - Decision Tree
  - K-nearest neighbors (KNN)
- Note:
  1. Among the models trained the Decision Tree classification model has the highest accuracy with a value of 0.94.

Confusion Matrix



- This confusion matrix belong to a tree classification model with an accuracy 94.4%.
- Comments:
  1. We observe that the model predict with rather high accuracy and distinguish between landing classes.
  2. False positives and false negatives do not play a big role.

Conclusions

<u>EDA with visualization:</u>

- We observe an improving mission outcome over the years.
- Heavier payloads have, in general, a higher success rate than lighter payloads.
- Launches to SSO and VLEO orbits have relatively high success rate with lighter and higher payload masses, respectively.

<u>Interactive Folium maps:</u>

- KSC LC 39A launch site has the best mission outcome rates.
- Launch sites are, in general, close to the coastline, highways and railways but far away from cities.

<u>Interactive Dashboards:</u>

- Payloads between 2000-6000 kg have higher success rates, where rockets with a F9 FT boosters are predominating within this range.

<u>Classification Analysis:</u>

- The classification model used that classifies the best is a decision tree classification model which performs with a low count on false positives and false negatives.

Appendix

- Github URLs:
  I. Data Collection – SpaceX API
     https://github.com/xelivy/CapstoneProject/blob/80f052676ff3629f3bcfeecc45d1b808e37c59e6/Data%20Collection%20API%20Lab.ipynb
  II. Data Collection – Webscrapping
     https://github.com/xelivy/CapstoneProject/blob/096d0ab992e5b8c1a1e409c01b4d28bf221c6762/Data%20Collection%20with%20Web%20Scraping.ipynb
  III. Data Wrangling
     https://github.com/xelivy/CapstoneProject/blob/096d0ab992e5b8c1a1e409c01b4d28bf221c6762/Data%20Wrangling.ipynb
  IV. EDA with Data Visualization
     https://github.com/xelivy/CapstoneProject/blob/096d0ab992e5b8c1a1e409c01b4d28bf221c6762/EDA%20with%20Visualization.ipynb
  V. EDA with SQL
     https://github.com/xelivy/CapstoneProject/blob/096d0ab992e5b8c1a1e409c01b4d28bf221c6762/EDA%20with%20SQL%20.ipynb
  VI. Build an Interactive Map with Folium
     https://github.com/xelivy/CapstoneProject/blob/3d2bc81b1c75e8d9620d199cea6d6582ec6cbd15/Interactive%20Visual%20Analytics%20with%20Folium.ipynb
  VII. Build a Dashboard with Plotly Dash
     https://github.com/xelivy/CapstoneProject/blob/096d0ab992e5b8c1a1e409c01b4d28bf221c6762/spacex_dash_app.py
  VIII. Predicitve Analysis (Classification)
     https://github.com/xelivy/CapstoneProject/blob/f7ddccd2a0469b079d7fdab4b16340e6b160fa4d/Machine%20Learning%20Prediction%20lab.ipynb

Thank you!