IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Abiyselassie Gashaw

August 30, 2023

Outline

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

Executive Summary

Summary of methodologies

- Data Collection through API
- Data Collection with Webscraping
- Data Wrangling
- Exploratory Data Analysis with SQL
- Exploratory Data Analysis with Data Visualization
- Interactive Visual Analytics with Folium
- Interactive Dashboard with Plotly Dash
- Machine Learning Predictions

Summary of all results

- Exploratory Data Analysis Result
- Interactive Analytics in Screenshots
- Predictive Analytics Result

Introduction

Project background and context

- SpaceX has been the most successful in making space travel affordable for everyone over decades accomplishing tasks like sending spacecraft to the International Space Station, providing satellite Internet access & sending manned missions to Space.
- It advertises on its website, Falcon 9 rocket launches, which are known for incorporating several reusable technologies that have revolutionized space travel by significantly reducing the cost of launching payloads into space.
- It launches with a relatively inexpensive cost of 62 million dollars; but other rocket providers cost upward of 165 million dollars each.
- Much of its significant cost savings is due to SpaceX's Falcon 9 innovative technology to reuse 1st stage.
- Therefore, if the '1st stage will land' can be determined, the cost of each launch can also be determined.

Problems to find answers

- Predict if the 1st stage of Falcon 9 will land successfully.
- Predict if SpaceX will reuse the 1st stage if the cost of each launch can be determined.
- Determine what attributes (payload mass, flight No., Orbit, etc.) are correlated with successful landings.
- Identify factors that affect successful 1st stage landing of Falcon 9.
- Detect ML model that perform the best for predicting the successful landing.

Section 1

Methodology

Methodology

Executive Summary

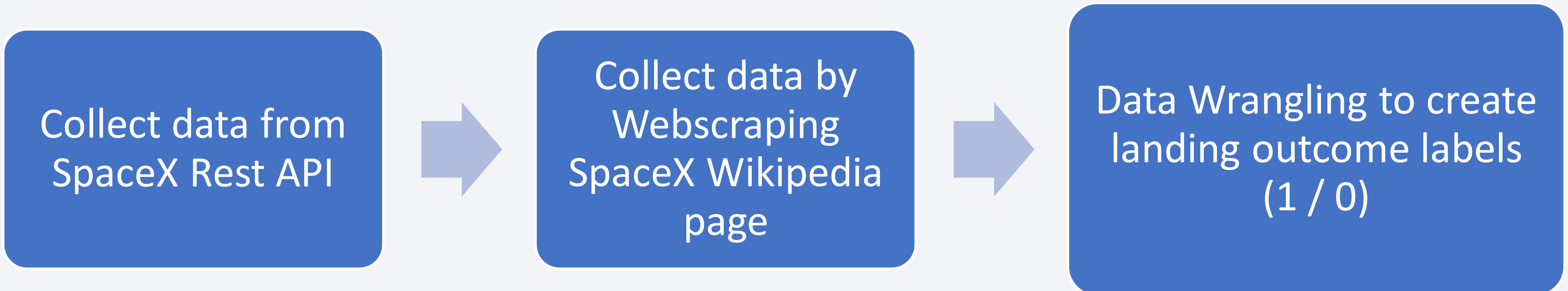
- Data collection methodology:
  - The data were collected from Public source using SpaceX API.
  - Falcon 9 launch records were webscraped from Wikipedia page using Beautiful Soup library.
- Perform Data Wrangling:
  - The data was wrangled to determine Success rate of Landing class & all successful landings were created.
- Perform Exploratory Data Analysis (EDA) using Data Visualization and SQL
  - Scatterplots, line plots & bar chart were plotted to identify factors for successful landing & yearly trend.
  - Total payload mass, Total No. of successful & failure mission outcomes, etc. were selected from SpaceXTBL
- Perform Interactive Visual Analytics using Folium and Plotly Dash
  - An Interactive Map with Folium was built to visualize a launch site proximity to geographical markers.
  - A Dashboard with Plotly Dash (Pie chart, Scatterplot, Payload range slider interaction) was built to visualize launch sites with most success & payload ranges.
- Perform Predictive Analysis using Classification Models
  - Successful landing outcomes were predicted by building ML models(LR, SVM, Decision Tree, KNN), tuning Hyperparameters & evaluating based on accuracy score and confusion matrix

Data Collection

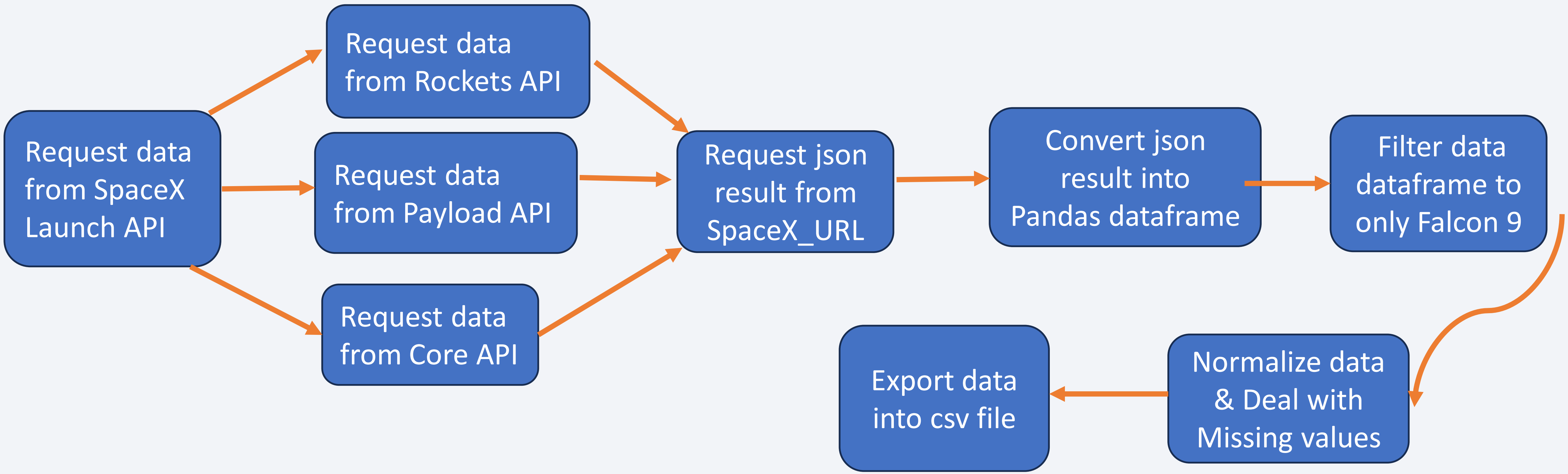
- The data were collected from Public source of SpaceX API &
  - Launch data was obtained using the get requests
- The data were collected by webscraping SpaceX Wikipedia page (updated on 9th June 2021)
  - Falcon 9 Launch data, another popular data source was obtained using Beautiful Soup library
- The data obtained from SpaceX API & Webscraping were wrangled
  - A landing outcome label was created from the Outcome column
  - Success rate of Landing class was determined

Landing outcome labels were trained with

- 1 means the booster successfully landed
- 0 means it was unsuccessful

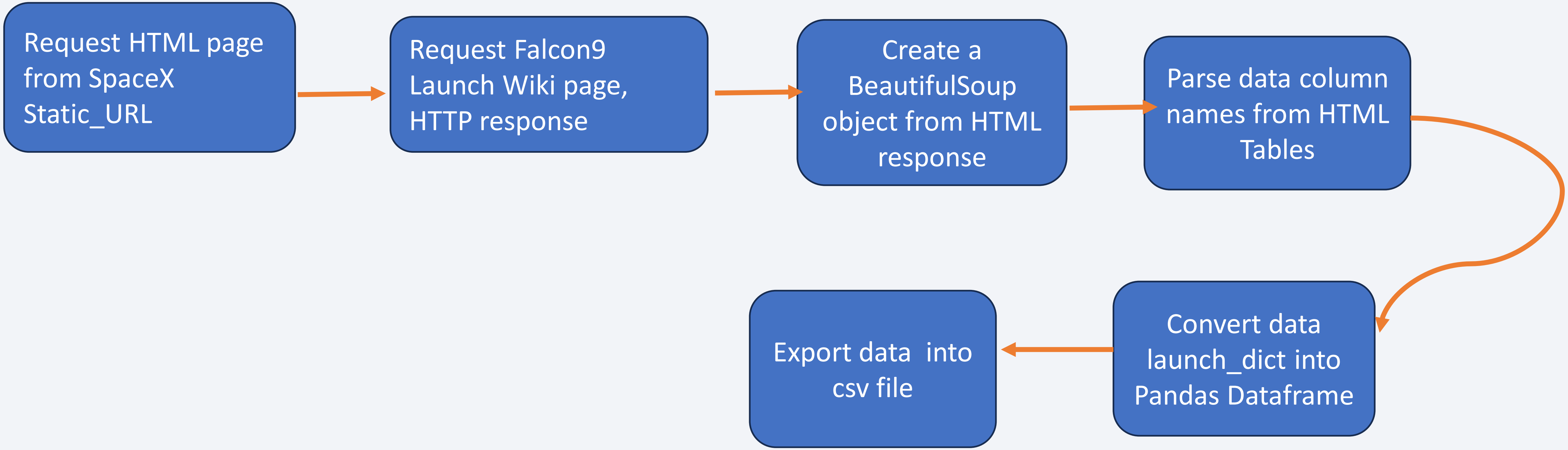


Data Collection – SpaceX API



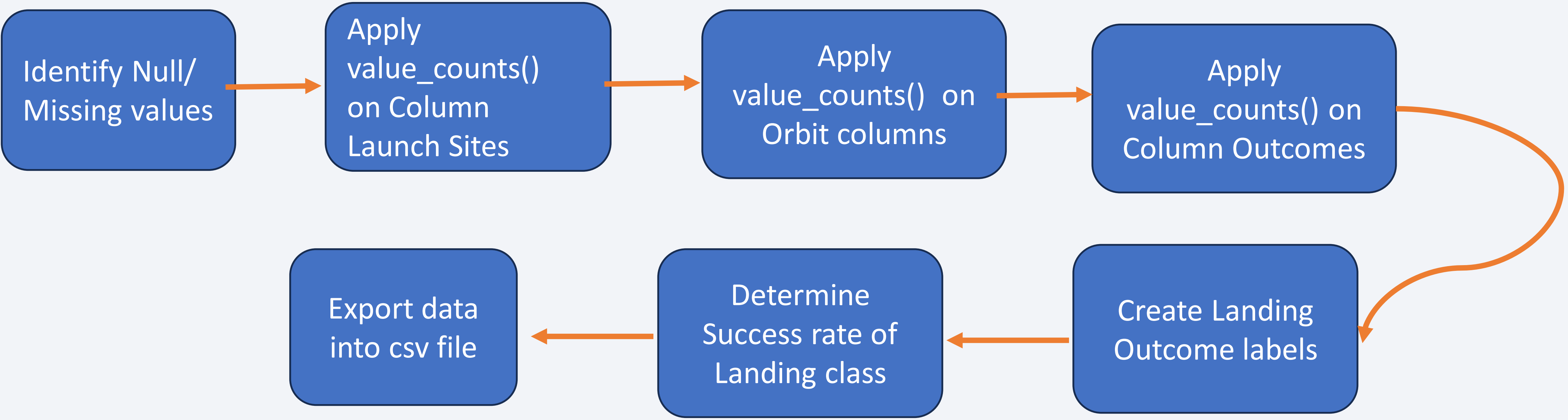
GitHub URL of the **completed SpaceX API calls notebook** (https://github.com/abiyselassie22/testpro/blob/master/jupyter-labs-spacex-data-collection-api%20(1).ipynb)

Data Collection - Scraping



GitHub URL of the completed **web scraping** **notebook**
(https://github.com/abiyselassie22/testpro/blob/master/jupyter-labs-webscraping.ipynb)

Data Wrangling



GitHub URL of your completed **data wrangling** related notebooks (https://github.com/abiyselassie22/testpro/blob/master/IBM-DS0321EN-SkillsNetwork_labs_module_1_L3_labs-jupyter-spacex-data_wrangling_jupyterlite.jupyterlite.ipynb)

EDA with Data Visualization

The following **attributes/features** can be used to determine if **Falcon 9's 1st stage** can be reused.

Scatterplots were plotted to show if there is any relationship between

- **Flight Numbers** & **Launch Site**;
- **Payload** & **Launch Site**;
- **Flight Number** & **Orbit type**;
- **Payload** & **Orbit type**

✓ **1st stage** is more likely to land **successfully** when **Flight No.** (continuous **launch attempts**) increases.

Bar chart

- was plotted to show if there are any relationship between **success rate** & each **orbit type**

Line chart

- was drawn to show the **increasing yearly trend** of **launch success rates**.

GitHub URL of your completed **EDA with data visualization notebook** (https://github.com/abiyselassie22/testpro/blob/master/IBM-DS0321EN-SkillsNetwork_labs_module_2_jupyter-labs-eda-dataviz.ipynb.jupyterlite.ipynb)

EDA with SQL

- **Names** of the **unique launch sites** were **selected** from the **space mission SPACEXTBL**
- **5 Records** of **launch sites** beginning with the **string 'CCA'** were **selected** from **SPACEXTBL**
- **Total payload mass** carried by **boosters** launched by **NASA** (**CRS**) were **selected**
- **Average payload mass** carried by **booster version F9 v1.1** was **selected**.
- **Date** when the **first successful landing outcome** in **ground pad** was achieved was **selected**.
- **Names** of **boosters** with **success** in **drone ship** & **payload** mass **>= 4000** &**<=6000** were **selected**
- **Total No.** of **successful** & **failure mission outcomes** were **selected**
- **Names** of the **booster_versions** which have carried the **maximum payload mass** were **selected**.
- **Count** of **landing Failure** / **Success outcomes** b/n **date 2010-06-04** & **2017-03-20** were **ranked**

GitHub URL of your completed **EDA** with **SQL notebook**

(https://github.com/abiyselassie22/testpro/blob/master/jupyter-labs-eda-sql-coursera_sqllite.ipynb)

Build an Interactive Map with Folium

- **map.add_child(circle)** & **map.add_child(marker) objects** were created & added to **folium map** to make the map look less congested & indicate points or areas of our **interest**.
- **MarkerCluster object** was added to simplify the **folium map** containing many **markers** having the same **coordinate** & group the multiple nearby **markers** into **clusters** based on **proximity**. This is used to **mark** all **launch sites, success**/**failed launches** for each site on an interactive map.
- **MousePosition object** was added on the **folium map** to get coordinate for a mouse over a point on the map & to easily find **coordinates** of any points of **interests** (e.g. **Launch site**).
- **Map.add_child(lines) objects** were used to add a **polyline** to the **Folium map** to visualize **connections** / draw **distances** between a **launch site** & their **close proximities** (selected **coastline, railway**, **highway**, **city**, etc.). This can be used to choose an **optimal launch site**.

GitHub URL of your **completed Interactive map with Folium map**
(https://nbviewer.org/github/abiyselassie22/testpro/blob/master/IBM-DS0321EN-SkillsNetwork_labs_module_3_lab_jupyter_launch_site_location.jupyterlite.ipynb)

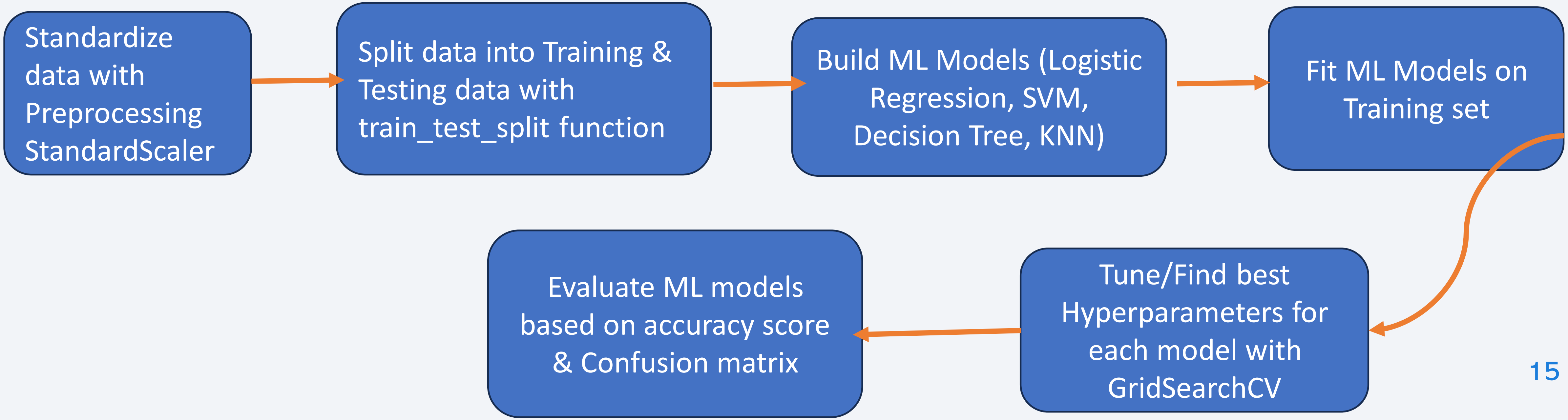
Build a Dashboard with Plotly Dash

- **SpaceX Launch Records Dashboard** with **dropdown list** was created to let us select **all sites** and different **launch sites**.
- **Pie charts** for **all sites** & **selected sites** were rendered & added to the **SpaceX Launch Records Dashboard** to visualize the **launch success counts** for **all sites** & **entered sites**.
- **Payload range slider interaction** was rendered & added to the **SpaceX Launch Records Dashboard** to find if selection of **variable payloads** is correlated with the **mission outcome**.
- **Scatterplot chart** were rendered & added to the **SpaceX Launch Records Dashboard** to visualize how **payload** may be correlated with **mission outcomes** for all **sites** & **selected site**(s).

GitHub URL of your **completed Plotly Dash lab**
(https://github.com/abiyselassie22/testpro/blob/master/spacex_dash_app.py)

Predictive Analysis (Classification)

GitHub URL of your completed **predictive analysis** lab,
https://github.com/abiyselassie22/testpro/blob/master/IBM-DS0321EN-SkillsNetwork_labs_module_4_SpaceX_Machine_Learning_Prediction_Part_5.jupyterlite.ipynb



Results

- **Exploratory Data Analysis results**
  - Visualization
  - SQL
- **Interactive Analytics Demo in screenshots**
  - Folium
  - Dashboard with Plotly Dash
- **Predictive Analysis results**
  - Machine Learning Prediction
  - Classification Accuracy
  - Confusion Matrix

Section 2

Insights drawn from EDA

Flight Number vs. Launch Site

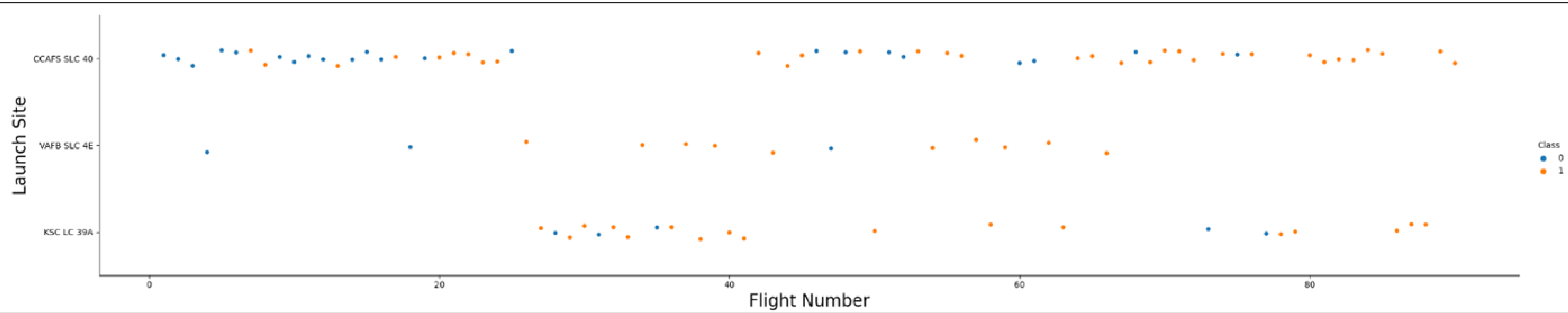


Figure 1. Scatterplot showing the relationship between **Flight Number** & **Launch Site**

- In CCAFS LC-40 launch site, the Success appears related to the Nos. of flights;

✓ as the **No.** of Success **labeled** class 1 increases with increase in **Flight Nos.**

- While in KSC LC-39A & VAFB SLC 4E launch sites, success have no relationship with Flight Nos.

✓ as the **No.** of their Successes show no difference when **Flight Nos.** increases.

Payload vs. Launch Site

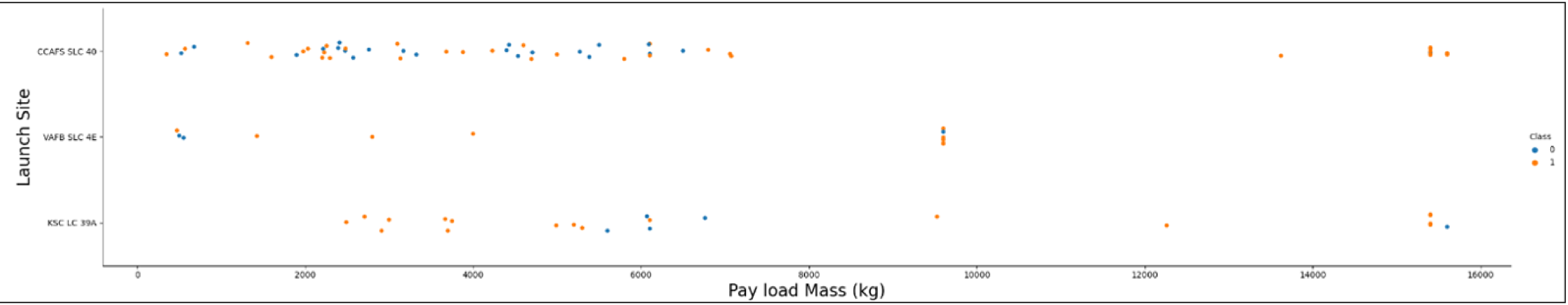


Figure 2. Scatterplot showing the relationship between **Payload** & **Launch Site**

The Payload vs. Launch Site scatter point chart shows that

- Lighter weighted payloads (1000-4000kg) has a success rate of 100% for the launch site VAFB-SLC.

✓ there are no rockets **launched** for heavy payload mass(>10000) for VAFB-SLC launch site.

- CCAFS LC-40 has a success rate of 60%,

✓ but if the **mass** is above > 10,000 kg, the success rate is 100% (Class 1).

Success Rate vs. Orbit Type

- The **plotted bar chart** shows that

✓ **orbits ES-L1**, **GEO**, **HEO**, & **SS0** have **highest success rates, 100%**

✓ **orbit VLEO**, **90%.**

✓ **Orbit, ISS, 60%**

having **lowest success rate**

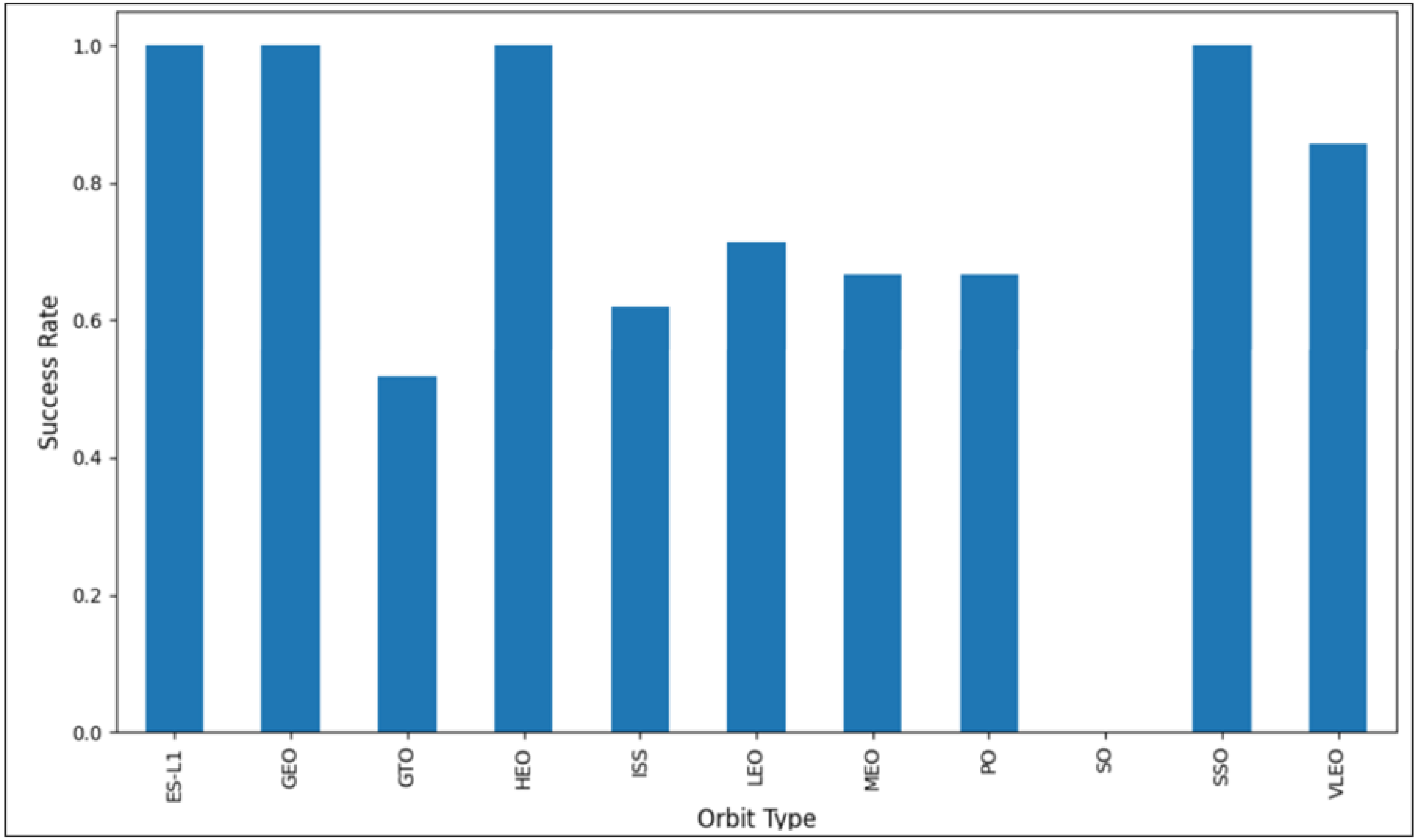


Figure 3. Bar chart showing the relationship between **Success Rate** & **Orbit Type.**

Flight Number vs. Orbit Type

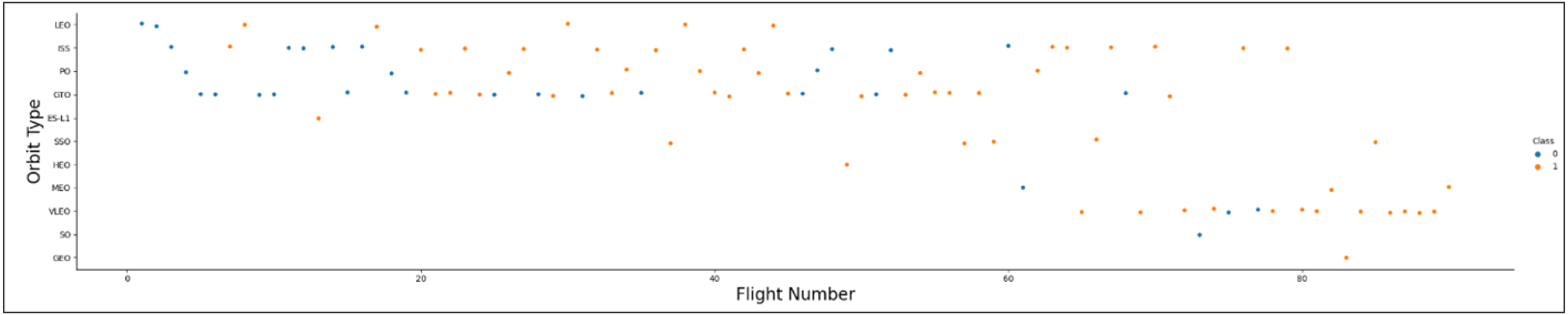


Figure 4. Scatterplot showing the relationship between **Orbit Type** & **Flight Number.**

- The Scatterplot shows that the Success appears related to the **Nos**. of flights in the LEO orbit;

✓ as the **No**. of Success **labeled** class 1 increases with increase in Flight Nos.

- while there seems to be no relationship between success rate & Flight Nos. in GTO orbit.

✓ as the **No**. of its Success shows no difference when Flight Nos. increases.

Payload vs. Orbit Type

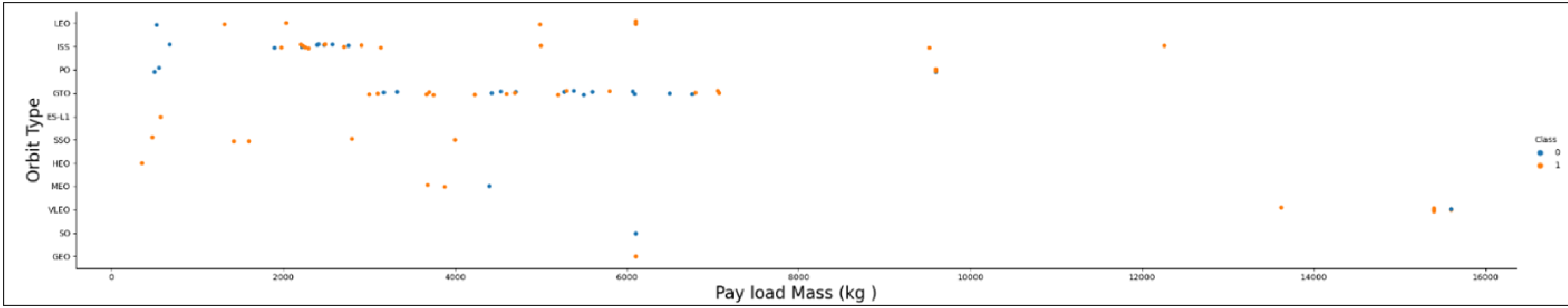


Figure 5. Scatterplot showing the relationship between **Payload** & **Orbit Type.**

- The **scatterplot** shows that with **heavy payloads**, the **successful landing** / **positive landing rate** are more for **Polar** (**PO**), **LEO** & **ISS**.
- However, **successful landing** /**positive landing rate** for **GTO** cannot be distinguished as both positive **landing rate** & **negative** landing rate (**unsuccessful mission**).

Launch Success Yearly Trend

The **line chart**

- shows that the **Launch success rate** since **2013** has kept **increasing** over time till **2020**.

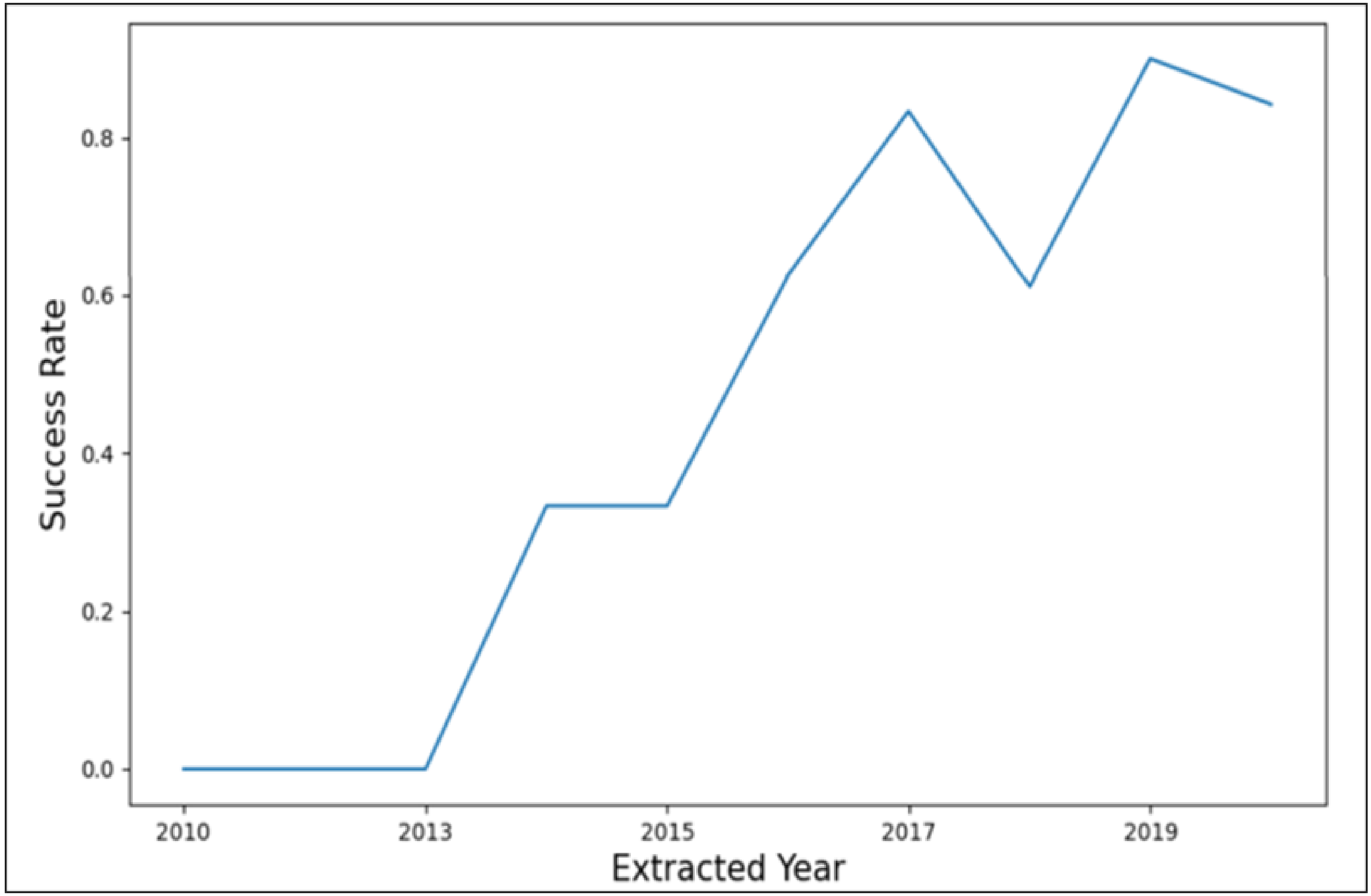


Figure 6. Line chart showing Launch Success Yearly Trend.

All Launch Site Names

Names of **unique launch sites** in space mission

- were distinctly selected from **SpaceX Table** as **CCAFS LC-40**, **VAFB SLC-4E**, **KSC LC-39A**, & **CCAFS SLC-40** in the query result on the **Right**.

```
[8]: %sql SELECT DISTINCT (Launch_Site) FROM SPACEXTBL;
 * sqlite:///my_data1.db
Done.
```

[8]:

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

Launch Site Names Begin with 'CCA'

```
[12]: %sql SELECT*FROM SPACEXTBL WHERE Launch_Site LIKE'%CCA%'LIMIT 5;
```

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

[12]:

Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer	Mission_Outcome	Landing_Outcome
2010-04-06	18:45:00	F9 v1.0 B0003	CCAFS LC-40	Dragon Spacecraft Qualification Unit	0	LEO	SpaceX	Success	Failure (parachute)
2010-08-12	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
2012-08-10	00:35:00	F9 v1.0 B0006	CCAFS LC-40	SpaceX CRS-1	500	LEO (ISS)	NASA (CRS)	Success	No attempt
2013-01-03	15:10:00	F9 v1.0 B0007	CCAFS LC-40	SpaceX CRS-2	677	LEO (ISS)	NASA (CRS)	Success	No attempt

5 Records where **launch sites** begin with the string **'CCA'**

- were selected from **SpaceX Table** as shown in the query result above.

Total Payload Mass

```
[12]: %sql SELECT SUM(PAYLOAD_MASS__KG_) AS Total_Payload_Mass, Customer FROM SPACEXTBL WHERE Customer = 'NASA (CRS)';
       * sqlite:///my_data1.db
      Done.
```

[12]:

Total_Payload_Mass	Customer
45596	NASA (CRS)

The **total payload mass** carried by **boosters launched** by **customer NASA (CRS)**

- was selected from **SpaceX Table** and calculated as **45596** kg in the query result above.

Average Payload Mass by F9 v1.1

```
[13]: %sql SELECT AVG(Payload_Mass__KG_) AS Average_Payload_Mass FROM SPACEXTBL WHERE Booster_Version LIKE'%F9 v1.1%';

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

[13]:

Average_Payload_Mass
2534.6666666666665

The **average payload mass** carried by **booster version F9 v1.1**

- was selected from **SpaceX Table** & displayed as **2534.67** in the query result above.

First Successful Ground Landing Date

```
[15]: %sql SELECT MIN(Date),Landing_Outcome FROM SPACEXTBL WHERE Landing_Outcome LIKE'%ground pad%';
 * sqlite:///my_data1.db
Done.
```

[15]:

MIN(Date)	Landing_Outcome
2015-12-22	Success (ground pad)

The **Date** when the **1st successful achievement** of **landing outcome** took place in **ground pad**

- was selected from **SpaceX Table** as **22 August 2015** in the query result above.

Successful Drone Ship Landing with Payload between 4000 and 6000

```
[15]: %sql SELECT Booster_Version,Landing_Outcome, PAYLOAD_MASS__KG_ FROM SPACEXTBL WHERE Landing_Outcome='Success (drone ship)'AND PAYLOAD_MASS__KG_>=4000 AND PAYLOAD_
```

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

[15]:

Booster_Version	Landing_Outcome	PAYLOAD_MASS__KG_
F9 FT B1022	Success (drone ship)	4696
F9 FT B1026	Success (drone ship)	4600
F9 FT B1021.2	Success (drone ship)	5300
F9 FT B1031.2	Success (drone ship)	5200

Names of the **boosters** which have **success** in **drone ship** &

have **payload mass greater than 4000** but **less than 6000 kg**

- were selected from **SpaceX Table** & listed down in the sql query result above.

Total Number of Successful and Failure Mission Outcomes

```
[16]: %sql SELECT COUNT(Mission_Outcome), Mission_Outcome FROM SPACEXTBL WHERE Mission_Outcome = 'Success' OR Mission_Outcome LIKE'%Failure%' Group By Mission_Outcome;
 * sqlite:///my_data1.db
Done.
```

[16]:

COUNT(Mission_Outcome)	Mission_Outcome
1	Failure (in flight)
98	Success

The **total no**. of **successful** & **failure mission outcomes**

- was selected from **SpaceX Table** &
- listed down as **1 Failure** & **98 Successful mission outcomes** in the sql query result above.

Boosters Carried Maximum Payload

The **names** of the **booster_versions** which have carried the **maximum payload mass**

- were selected from **SpaceX Table** & listed down in the sql query result below.

```
[17]: %sql SELECT Booster_Version, PAYLOAD_MASS__KG_ FROM SPACEXTBL WHERE PAYLOAD_MASS__KG_ = (select MAX(PAYLOAD_MASS__KG_) FROM SPACEXTBL);

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

[17]:

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

2015 Launch Records

```
[18]: %%sql
SELECT SUBSTR(Date, 1, 4) || '/' || SUBSTR(Date, 9, 2) || '/' || SUBSTR(Date, 6, 2) AS New_Date_Format,
Landing_Outcome, Booster_Version, Launch_Site FROM SPACEXTBL WHERE Landing_Outcome = 'Failure (drone ship)' AND SUBSTR(Date, 1, 4) = '2015';

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

[18]:

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

Launch records which will display the **month names**, **failure landing_outcomes** in **drone ship**, **booster versions**, & **launch_site** for the months in year **2015**

- were selected from **SpaceX Table** and listed down in the sql query result above.

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

```
[20]: %%sql
SELECT Landing_Outcome, COUNT(Landing_Outcome) AS COUNT, Rank() OVER (ORDER By COUNT(Landing_Outcome) DESC) FROM SPACEXTBL
WHERE Landing_Outcome = 'Failure (drone ship)' OR Landing_Outcome = 'Success (ground pad)' AND (Date between '2010-06-04' and '2017-03-20')
Group By Landing_Outcome
ORDER By COUNT DESC
```

* sqlite:///my_data1.db
Done.

[20]:

Landing_Outcome	COUNT	Rank() OVER (ORDER By COUNT(Landing_Outcome) DESC)
Success (ground pad)	5	1
Failure (drone ship)	5	1

Count of **landing outcomes** (such as **Failure** (**drone ship**) or **Success** (**ground pad**)) between the **date 2010-06-04** & **2017-03-20**

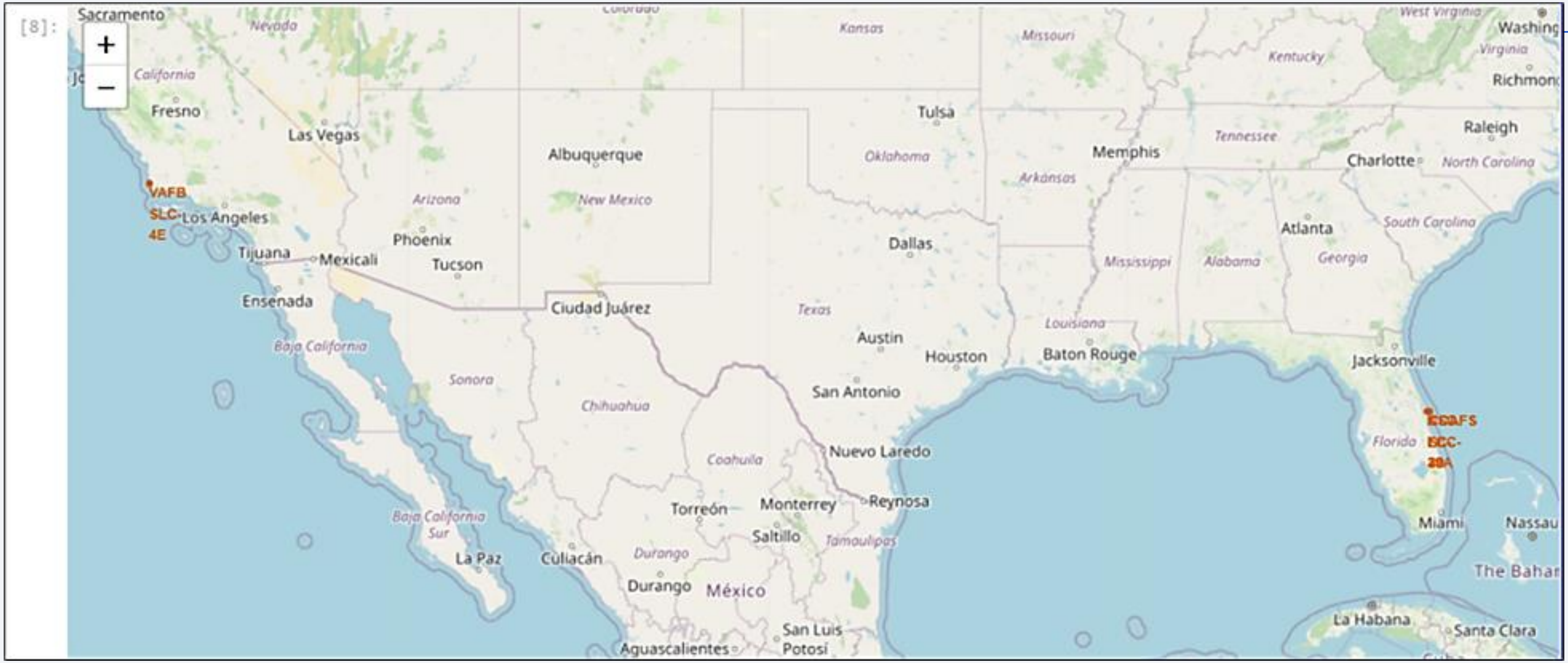
- was selected from **SpaceX Table** &
- **ranked** in **descending** order & **counted** as **5 Success** (**ground pad**)) or **5 Failure** (**drone ship**)

  as shown in the sql query result above.

Section 3

Launch Sites Proximities Analysis

Folium Map showing all Launch Sites Marked



- All **launch sites** were created & added on the **site map**

✓ using **folium.Circle** & **folium.Marker objects** as shown above.

- **Launch Sites marked** were **CCAFS LC-40**, **VAFB SLC-4E** , **KSC LC-39A** , & **CCAFS SLC-40**

Folium Map showing marked success/failed launches for Launch Sites

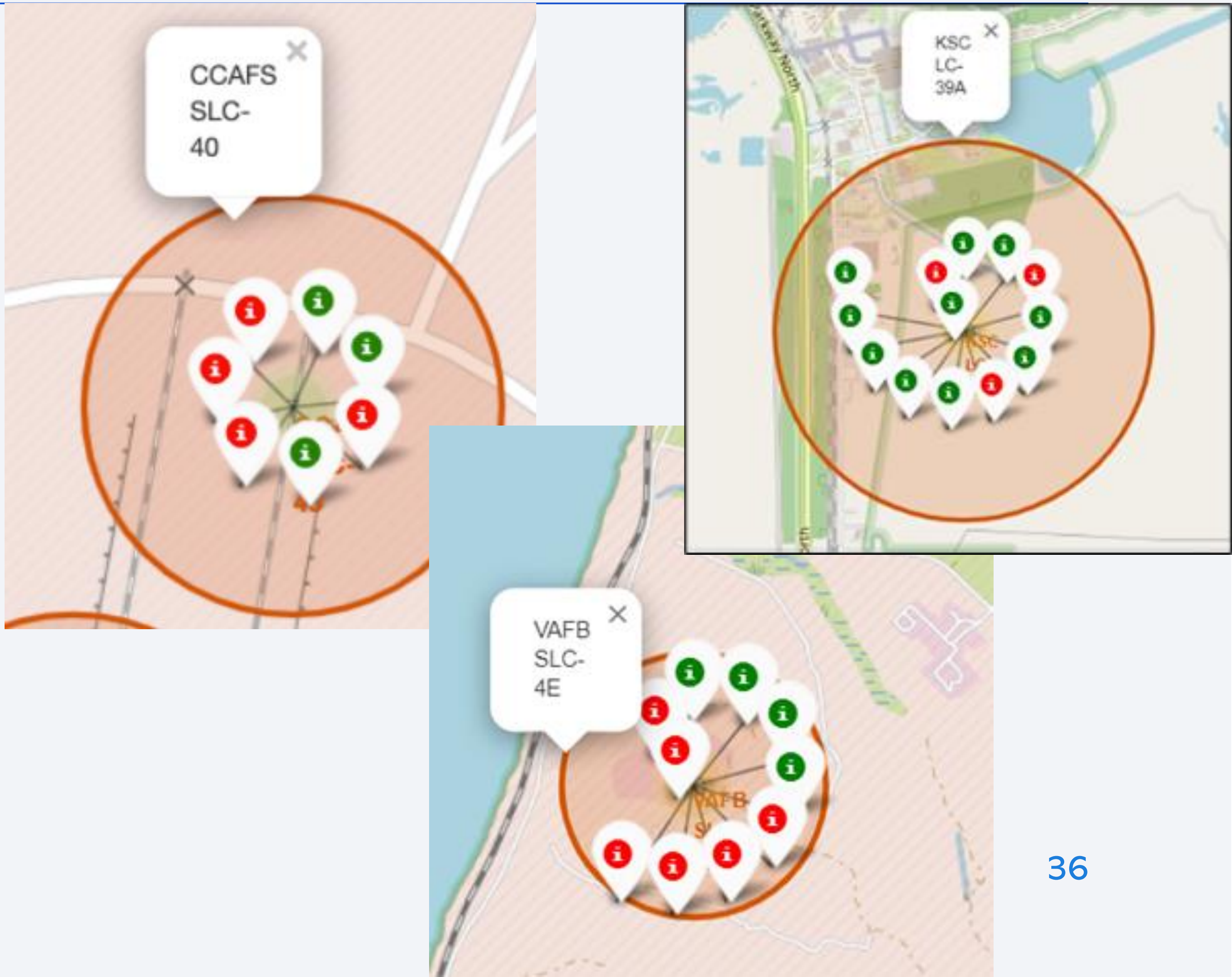
- The **success**/**failed launches** for each **launch site** were created and **marked** on the **site map** using **MarkerCluster()** **objects**
- **Zooming** in on the **launch sites**, **Launch sites** are **marked**; where

✓ each **green marker** represent a **launch** which was **successful** &

✓ each **red marker** stands for a **failed launch** as shown in the **screenshot** map on the **Right**.

- **Launch Site KSC LC-39A (more green markers)**

✓ shows relatively **higher success rate** than **CCAFS SLC-40** & **VAFB SLC-4E**



Folium Map showing distances b/n selected Launch Site to its proximities

The **distances** between a **launch site** **CCAFS SLC-40** to its **closest coastline**, **highway**, **railway** & **city**

- were calculated by finding their **coordinates** on the **map**

✓ first using **MousePosition objects**; &

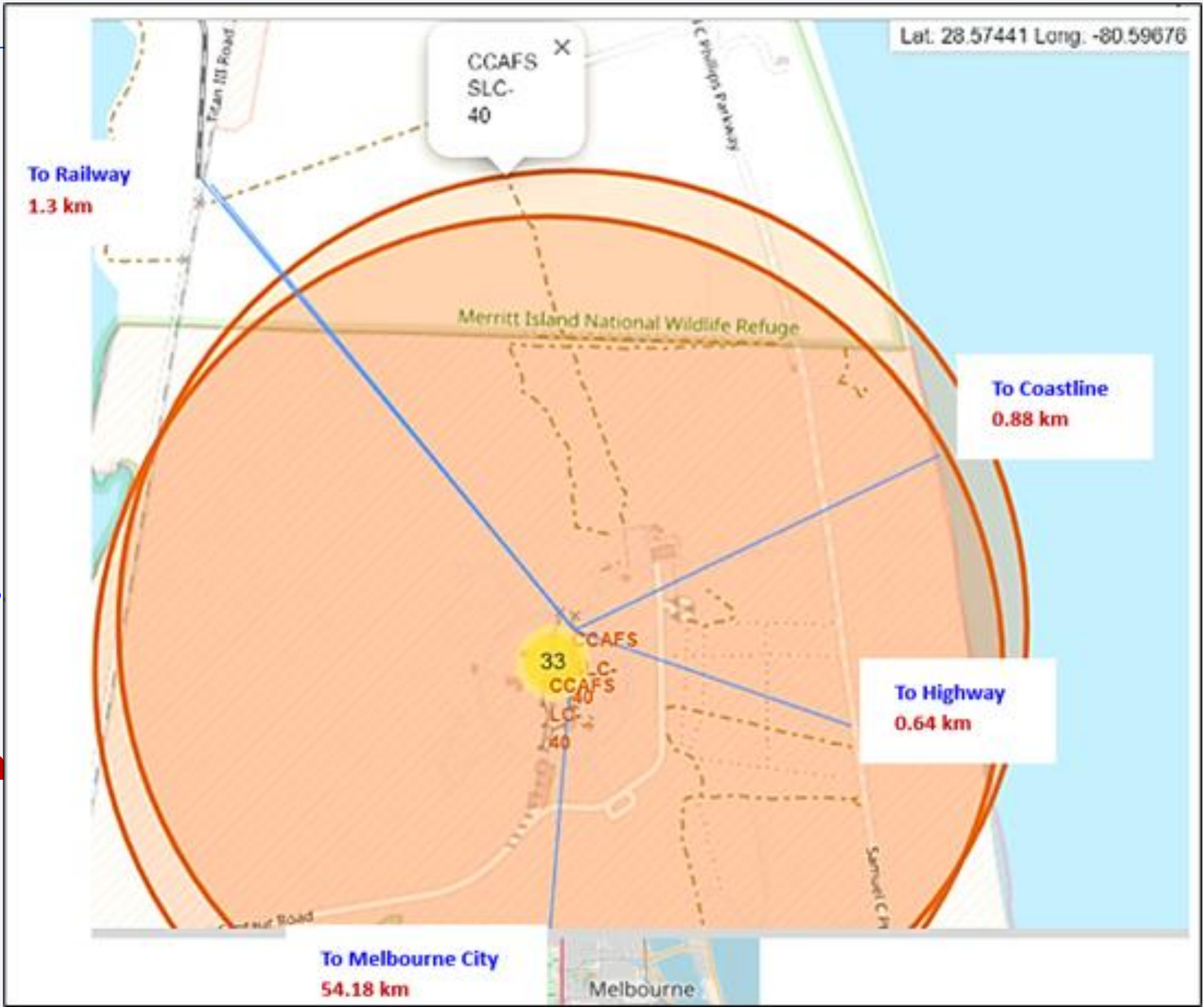
✓ then a **PolyLine** was drawn between the **CCAFS SLC-40** to its selected **proximities**

- Launch site is in close proximity to the

✓ **Highway** (**0.64 km**) & **Coastline** (**0.88 km**)

- But farther away from

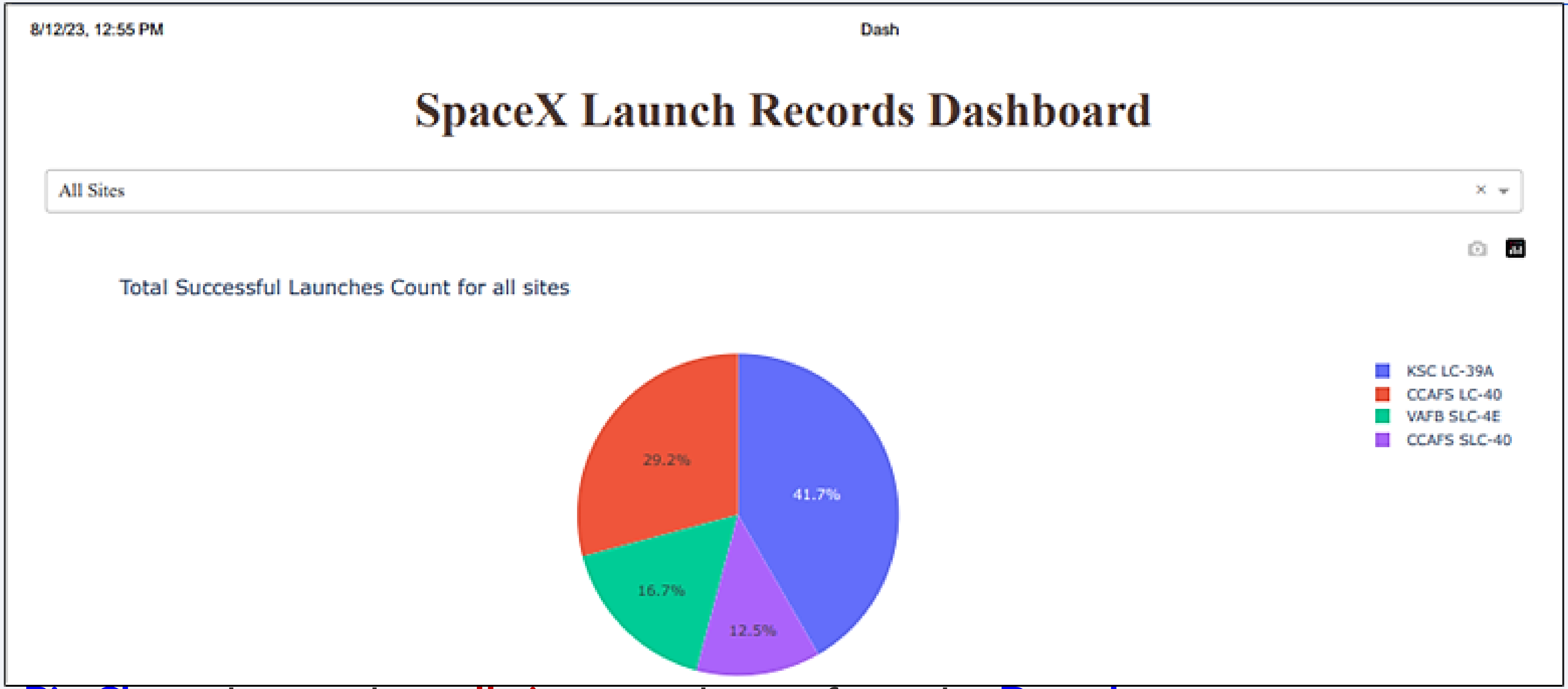
✓ **City** (**Melbourne**) (**54.18 km**).



Section 4

Build a Dashboard with Plotly Dash

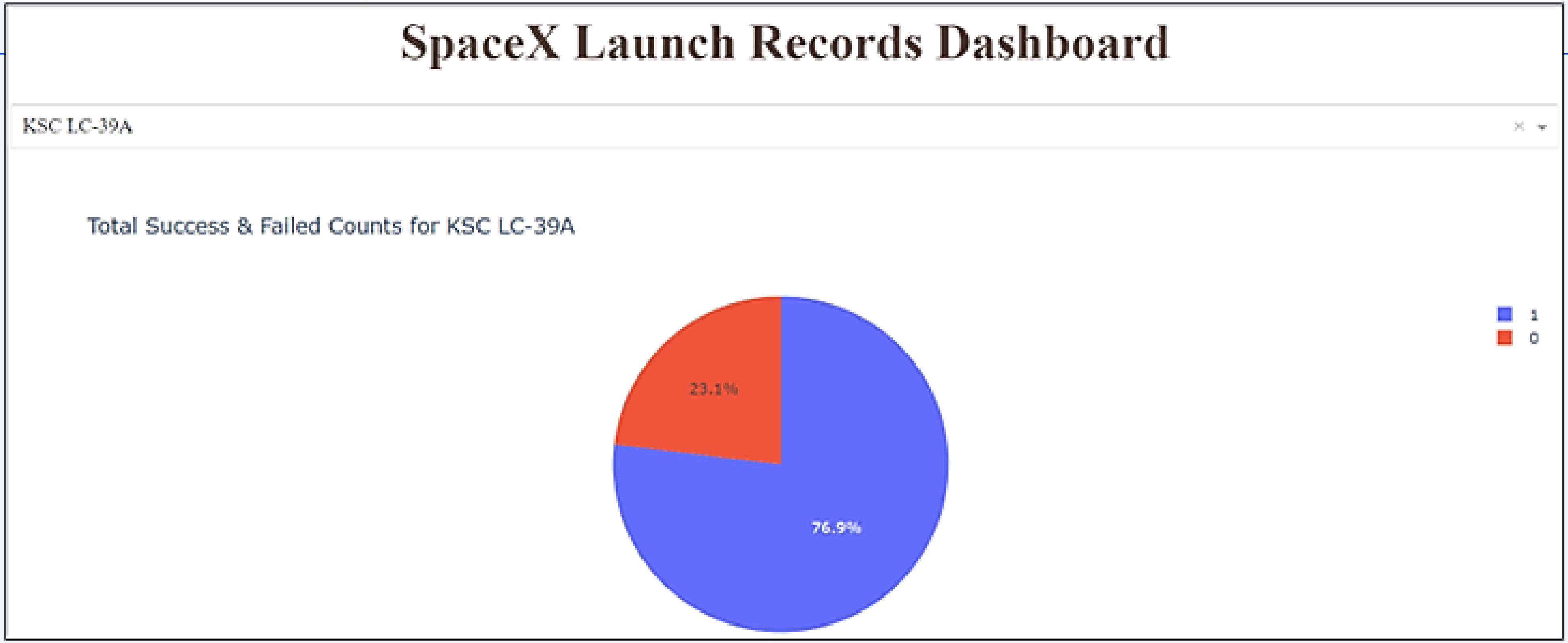
Dashboard showing Total Successful Launches Count for all Sites



The **Pie Chart** shows when **all sites** are chosen from the **Dropdown menu**.

- **KSC LC-39A** (**41.7%**) is the **launch site** with the **largest successful launches**.

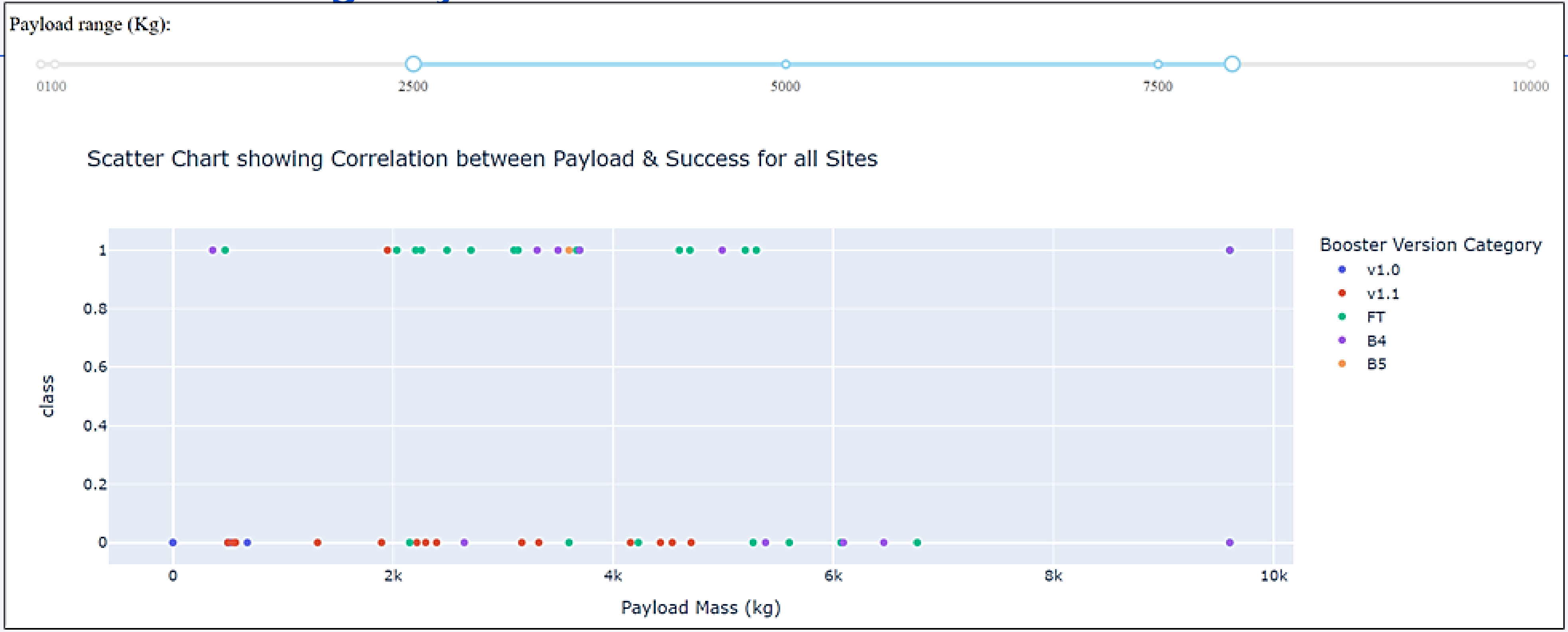
Dashboard showing Total Success & Failed Counts for KSC LC-39A



The Pie Chart shows when launch site KSC LC-39A is chosen from the Dropdown menu.

- 1 represents successful launches &
- 0 represents failed launches.
- 76.9% of the launches done at KSC LC-39A are successful launches.
- KSC LC-39A is displayed as the launch site with highest launch success ratio.

Dashboard showing Payload vs. Launch Outcome Scatter Plot for all sites



- The **Scatterplot** shows when **payload mass range** is selected b/n **2500** & **8000 kg** in the **range slider**.
- **Class 0** represents **failed launches** while **class 1** represents **successful launches**.
- The **payload range** of **20000-4000 kg** has the **highest launch success rate** whereas the **payload range** of **6000-8000 kg** has the **lowest launch success rate**.
- The **Booster version FT** tends to show the **highest launch success rate**

✓ with a greater number of **successful launches** reaching **class 1**.

Section 5

Predictive Analysis (Classification)

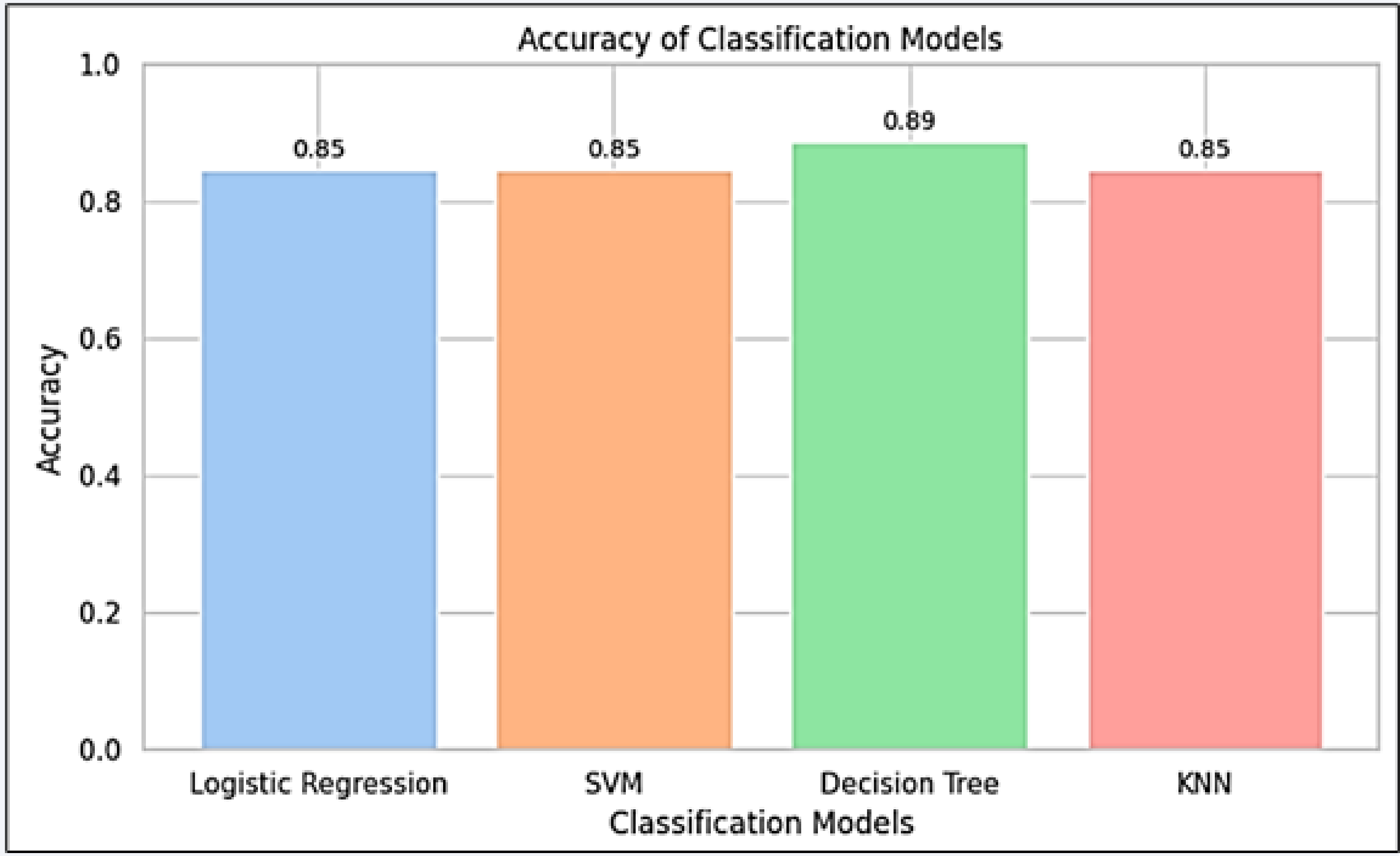
Classification Accuracy

- All the **Classification models** (**LR**, **SVM**, **Tree** and **KNN**)
- share the **same accuracy score** (**0.83**) on **test set**.

- **GridSearchCV.best_scores_**

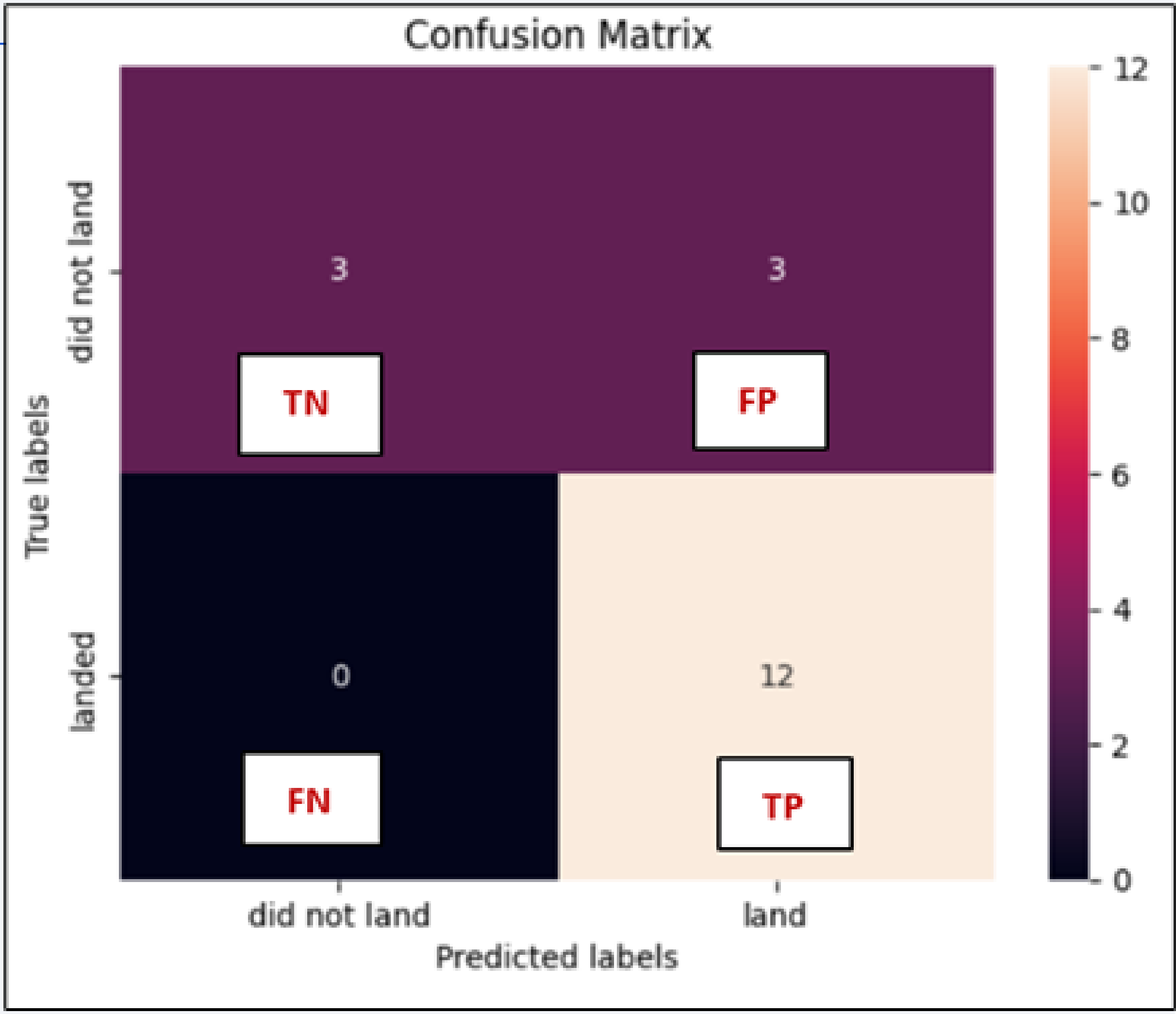
✓ used to compare **classification models** as shown in the **Bar chart**

- **Decision tree** with the **tree_cv.best_score_** of **0.89** has the **highest classification accuracy**.



Confusion Matrix

- The **confusion matrix plotted** after **Tree's accuracy** is tested on the **test data** as shown on the **Right**.
- It shows **corrected** & **wrong predictions** of **landings** in comparison with the **actual labels**.
- The **test set** has only **18 rows/outcomes**
- **True positive** (**TP**) is the **outcome (12)** when the **model correctly predicts** that a **landing** occurred when it actually did **land**.
- **False positive** (**FP**) is the **outcome (3)** when the **model predicts** a **landing** when it didn't actually **land**.
- **False Negative (FN) is the outcome (0) when model incorrectly predicted no landing when there was one.**
- **True Negative (TN) is the outcome (3) when model correctly predicted no landing.**



Confusion Matrix (Continued)

The **weighted Average of F1-Score** (0.81)

- balances **precision** & **recall** gained from **Confusion Matrix** as shown in query result below.

```
[35]: from sklearn.metrics import f1_score
      f1_score(Y_test, yhat, average='weighted')

[35]: 0.8148148148148149
```

```
[36]: from sklearn.metrics import classification_report, confusion_matrix
      import itertools
      cnf_matrix = confusion_matrix(Y_test, yhat)
      np.set_printoptions(precision=2)
      print (classification_report(Y_test, yhat))

              precision    recall  f1-score   support

           0       1.00      0.50      0.67         6
           1       0.80      1.00      0.89        12

    accuracy                           0.83        18
   macro avg       0.90      0.75      0.78        18
weighted avg       0.87      0.83      0.81        18
```

Conclusions

- All **models** (**LR**, **SVM**, **Decision Tree, KNN**) **perform best** w **same accuracy score** (**0.83**) on **test set**.
- **GridSearchCV_best scores_ (0.89) ranks** '**Decision Tree**' as the **best model, outperforming** the **other models**.

Confusion Matrix plays a good job in **predicting outcomes** with **Average Accuracy F1-score** (**0.81**)

- **correctly predicted 12** (**TP**) of **outcomes** as **landed** &
- **none** of them (**FN=0**) **wrongly predicted** as **not landed,**

✓ showing that **landing is not costly** & thus **SpaceX** can **reuse** the **Falcon 9 1st stage**.

The following **attributes/factors** are attributed to the **successful 1st stage landing** of **Falcon 9.**

- **ES-L1**, **GEO**, **HEO** & **SSO Orbits** display **highest success rates, 100%**
- **KSC LC-39A** (**41.7%**) has the **highest successful launches of all the launch sites.**
- **Lighter payloads** have **higher launch success rates** than **heavier payloads** for all **launch sites**.
- **Heavier payloads** perform **better** (**100%**) than **lighter payloads** for **CCAFS LC-40.**

Conclusions (continued)

- The correlation of attributes (**payload mass**, **flight No**., **Orbit**, etc.) with another also contributed to the **successful landings**.
- More successful **landing rate** was observed for Polar (PO), LEO & ISS orbits with heavy payloads.

The other **factor** that is responsible for successful **landing of Falcon 9** 1st **stage** is the fact that

- all launch sites are optimally **located** near the **Equator**, the **coastline**, **highways** & **railways**.
- Launch sites closer to the equator seems to gain a **rotational speed boost** due to the **Earth's rotation** which can be useful in terms of **fuel efficiency** & **payload capacity**.
- Their proximity to the coastline (**oceans**) allows them to have a safe **flight** over the **water** during the launch, minimizing risk to populated areas.
- Their accessibility to railways & highways facilitate easy **transportation** of **equipment** & **payloads**.
- Launch success rate has shown an increasing **yearly trend** since 2013 (year when reusable **technology was started**). This predicts **future successful landings** of Falcon 9 1st **stage** as more promising.

Appendix

[34]:
```
models = ['Logistic Regression', 'SVM', 'Decision Tree', 'KNN']
GridSearchCV.best_scores_ = [logreg_cv.best_score_, svm_cv.best_score_, tree_cv.best_score_, knn_cv.best_score_]
# Set the style using Seaborn
sns.set(style="whitegrid")

# Create a bar plot
plt.figure(figsize=(8, 5))
plt.bar(models, GridSearchCV.best_scores_, color=sns.color_palette("pastel"))

# Adding labels and title
plt.xlabel('Classification Models')
plt.ylabel('Accuracy')
plt.title('Accuracy of Classification Models')
plt.ylim(0, 1)  # Set the y-axis limits between 0 and 1

# Display the accuracy values on top of the bars
for i, v in enumerate(GridSearchCV.best_scores_):
    plt.text(i, v + 0.01, f'{v:.2f}', ha='center', va='bottom', fontsize=10)

# Show the plot
plt.tight_layout()
plt.show()
```

Thank you!