Rocket Launch Cost Prediction: First Stage Landing Success

VidyaHarkal

13 Apr 2025

OUTLINE

- Executive Summary
- Introduction
- Methodology
- Results
  - Visualization – Charts
  - Dashboard
- Discussion
  - Findings & Implications
- Conclusion
- Appendix

EXECUTIVE SUMMARY

- Predicting the success of the Falcon 9 first stage landing
- Understanding the factors that contribute to successful landings
- Estimating the overall cost of a launch
- Providing valuable insights for companies seeking to bid against SpaceX
- Developing a predictive model that determines the likelihood of successful first stage landings.

INTRODUCTION

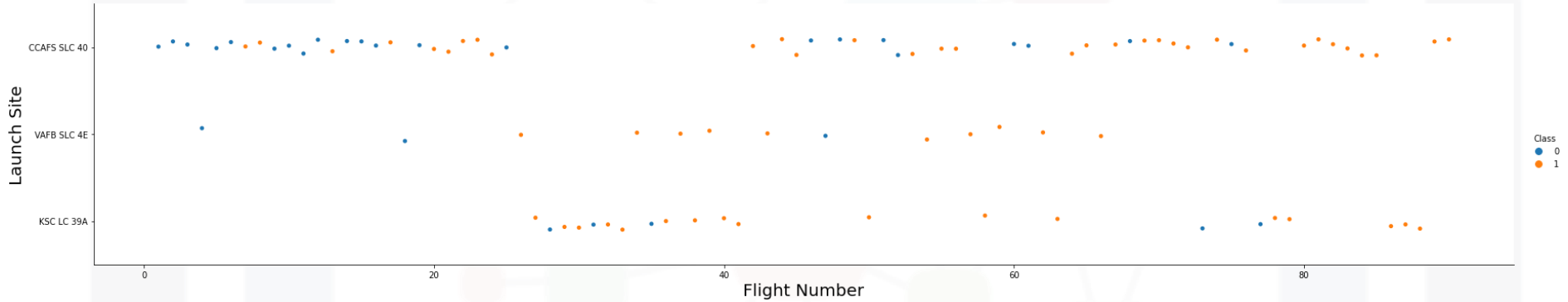
- Falcon 9 cost 62 million dollars whereas other cost 165 million
- Space X reuses Falcon9 first stage
- SpaceX's unique ability to reuse the first stage of the Falcon 9 rocket has significantly reduced launch costs compared to traditional providers.

METHODOLOGY

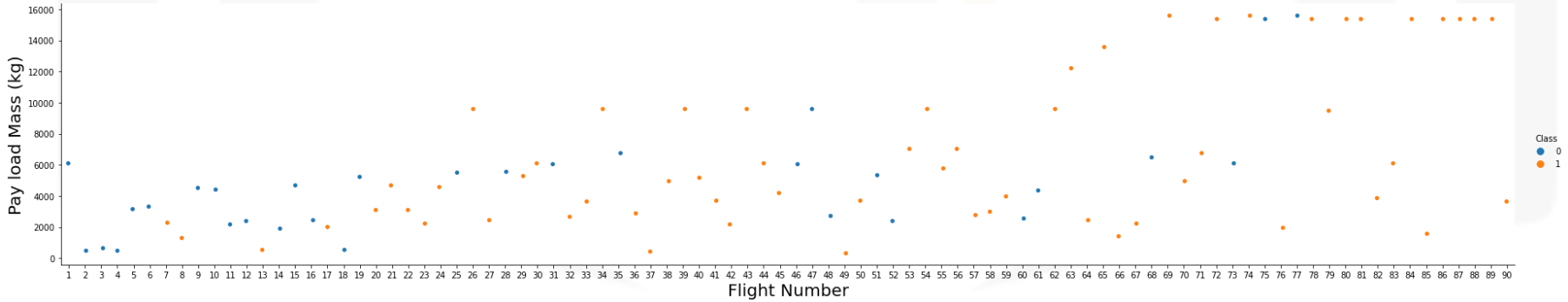
- Data Collection through:
  - SpaceX API
  - Web Scrapping
- Data Wrangling
- Data Analysis & Visualization using:
  - SQL
  - Pandas
  - Folium
  - Matplotlib & Seaborn
- ML Prediction of the Successful First stage Landing

METHODOLOGY: EDA & Interactive Visual Analytics (1/3)

- Visualize the relationship between Flight Number and Launch Site

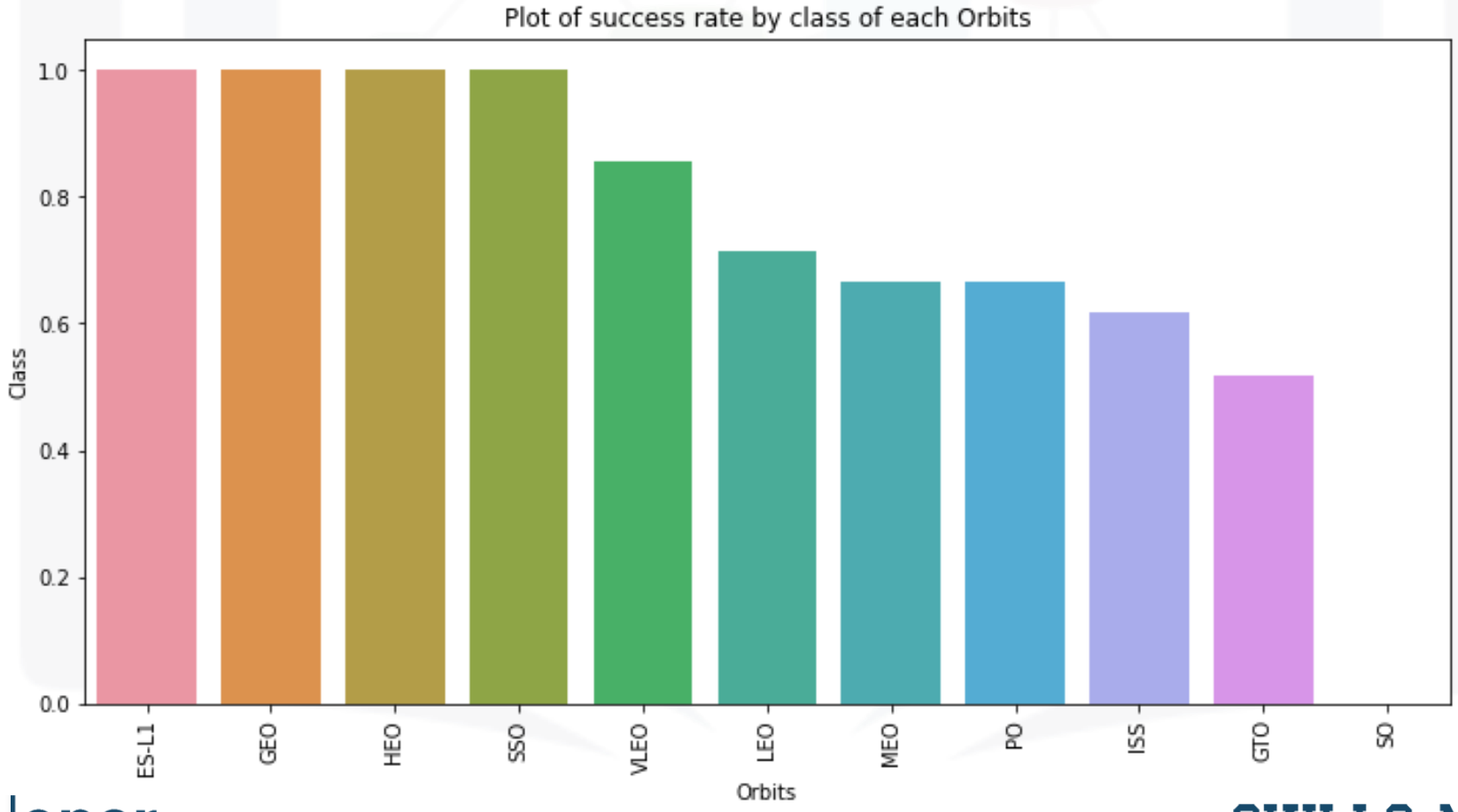


- Visualize the relationship between Payload and Launch Site



METHODOLOGY: EDA & Interactive Visual Analytics (2/3)

- Success rate of each orbit type

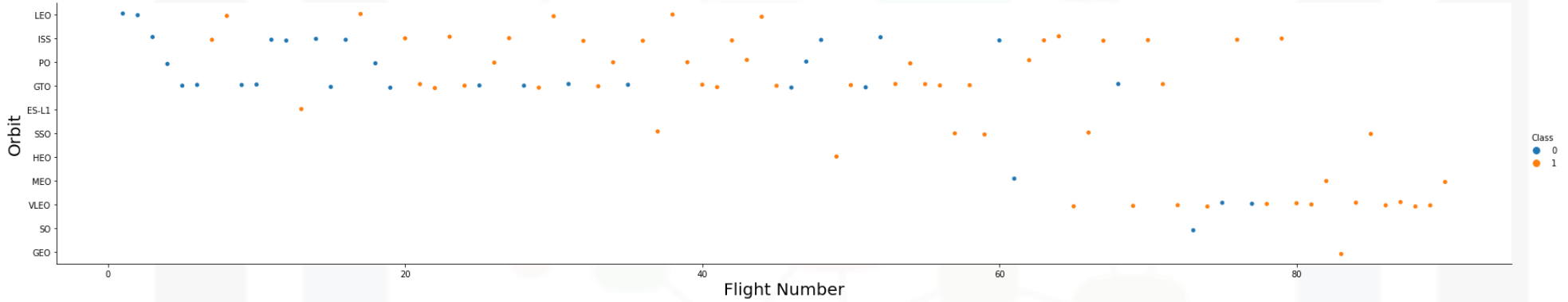


METHODOLOGY: Predictive Analysis

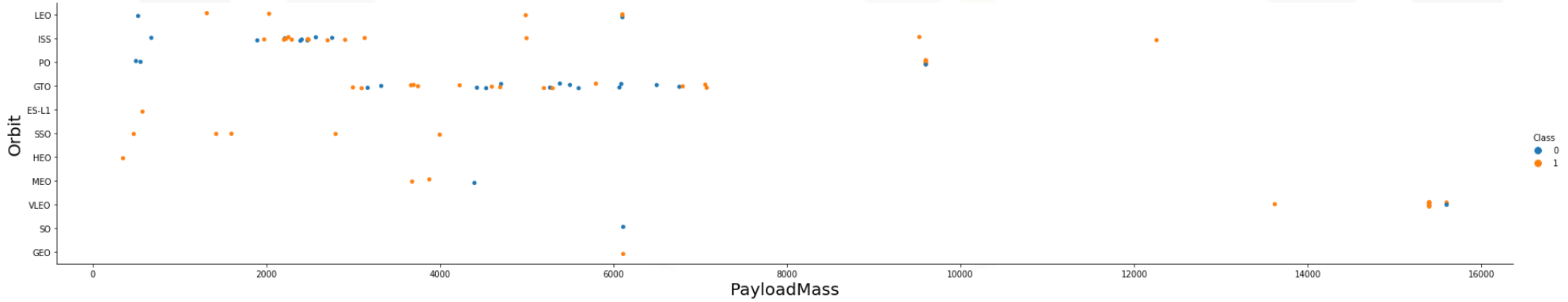
- Larger the flight amount at a launch site, the greater the success rate at a launch site.
- ES-L1, GEO, HEO, SSO, VLEO had the most success rate.
- LEO orbit, success is related to the number of flights whereas in the GTO orbit, there is no relationship between flight number and the orbit.

RESULTS: EDA & Interactive Visual Analytics (3/3)

- Visualize the relationship between FlightNumber and Orbit type



- Visualize the relationship between Payload Mass and Orbit type



RESULTS: EDA With SQL (1/4)

- Unique launch sites in the space mission

	launchsite
0	KSC LC-39A
1	CCAFS LC-40
2	CCAFS SLC-40
3	VAFB SLC-4E

- Total & Average payload mass carried by boosters launched by NASA (CRS)

	total_payloadmass
0	45596

	avg_payloadmass
0	2928.4

RESULTS: EDA With SQL (2/4)

- First successful landing date

	firstsuccessfull_landing_date
0	2015-12-22

- Booster Version

	boosterversion
0	F9 FT B1022
1	F9 FT B1026
2	F9 FT B1021.2
3	F9 FT B1031.2

RESULTS: EDA With SQL (3/4)

- Successful Outcomes

	successoutcome
0	100

- Booster_versions which have carried the maximum payload mass

	boosterversion	payloadmasskg
0	F9 B5 B1048.4	15600
1	F9 B5 B1048.5	15600
2	F9 B5 B1049.4	15600
3	F9 B5 B1049.5	15600
4	F9 B5 B1049.7	15600
5	F9 B5 B1051.3	15600
6	F9 B5 B1051.4	15600
7	F9 B5 B1051.6	15600
8	F9 B5 B1056.4	15600
9	F9 B5 B1058.3	15600
10	F9 B5 B1060.2	15600
11	F9 B5 B1060.3	15600

RESULTS: EDA With SQL (4/4)

- Failed landing outcomes in drone ship

	boosterversion	launchsite	landingoutcome
0	F9 v1.1 B1012	CCAFS LC-40	Failure (drone ship)
1	F9 v1.1 B1015	CCAFS LC-40	Failure (drone ship)

- Count of landing outcomes

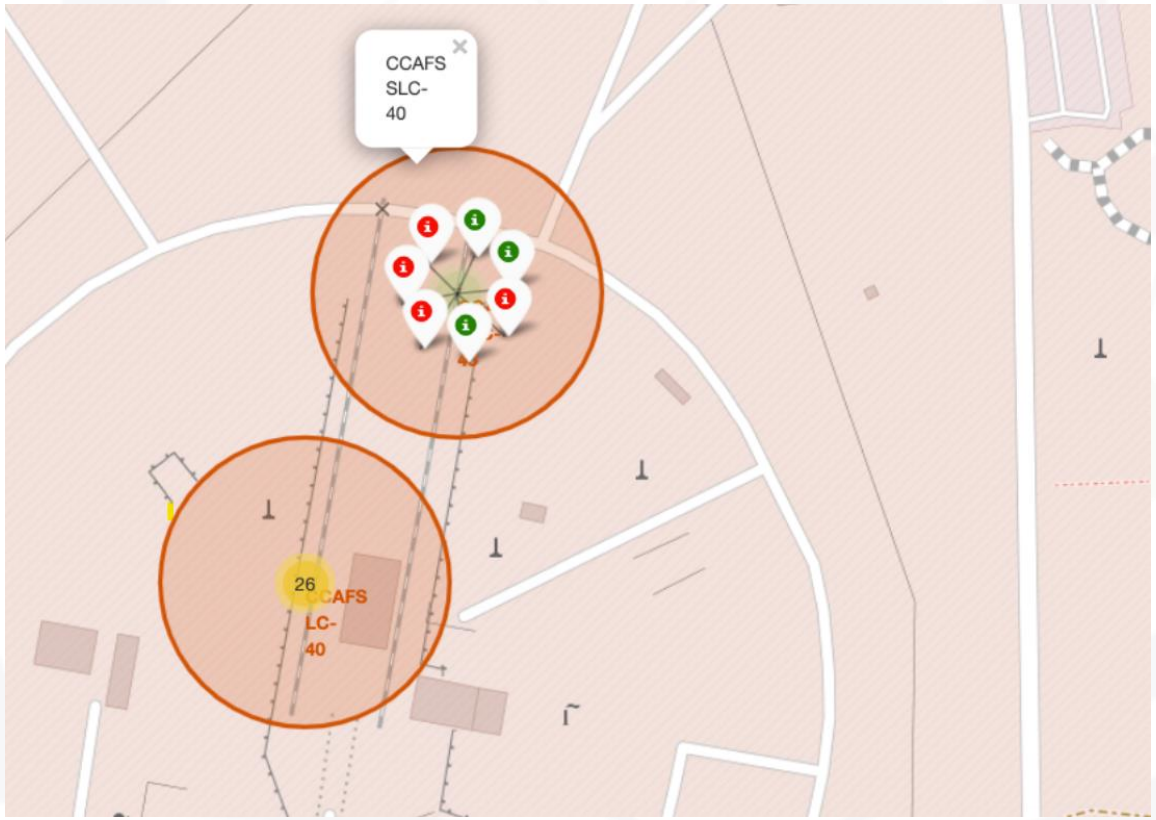
	landingoutcome	count
0	No attempt	10
1	Success (drone ship)	5
2	Failure (drone ship)	5
3	Success (ground pad)	5
4	Controlled (ocean)	3
5	Uncontrolled (ocean)	2
6	Precluded (drone ship)	1
7	Failure (parachute)	1

RESULTS: Interactive map with Folium (1/3)

- All launch sites on a map

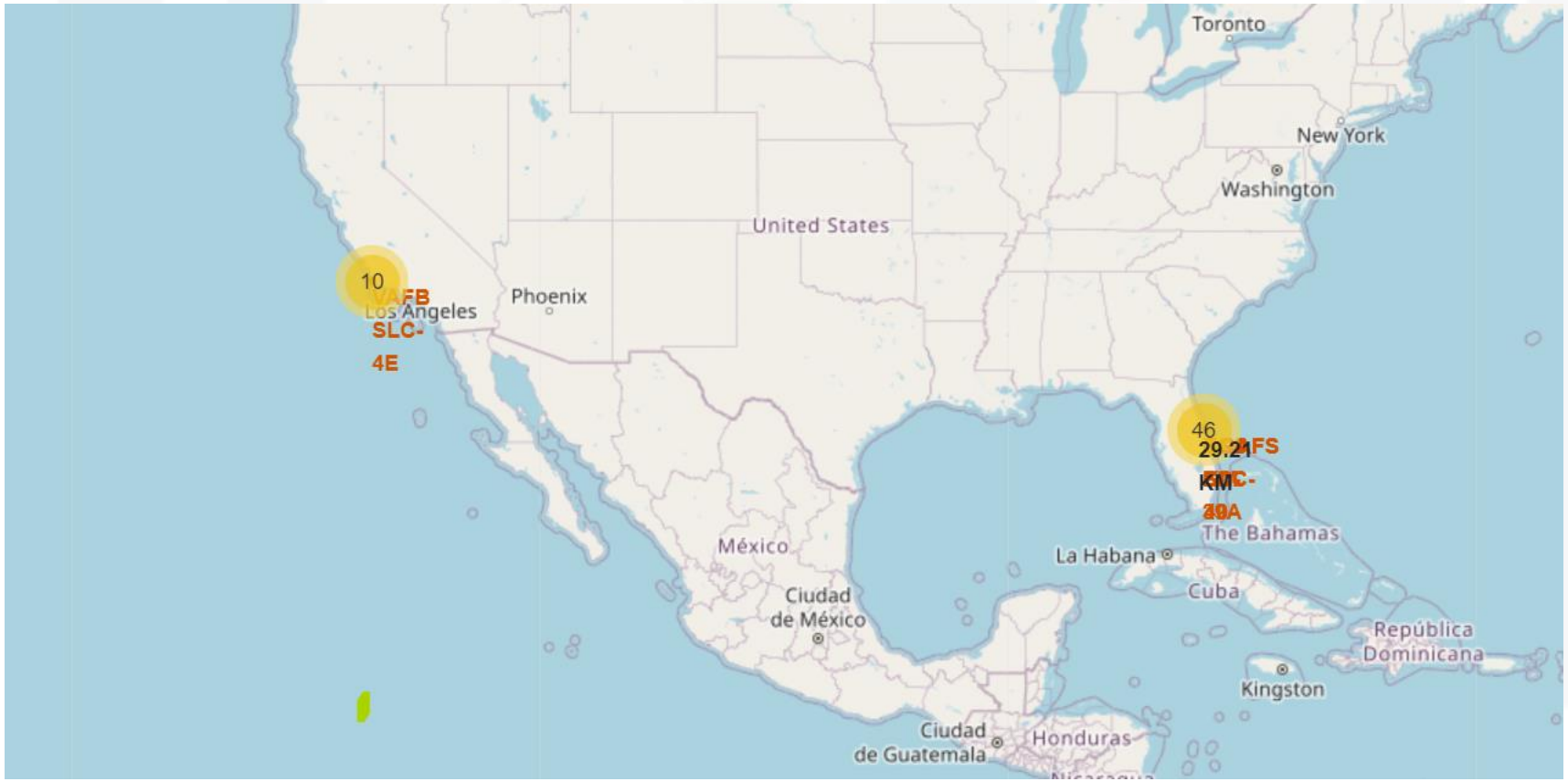
RESULTS: Interactive map with Folium (2/3)

- Success/Failed launches for each site on the map



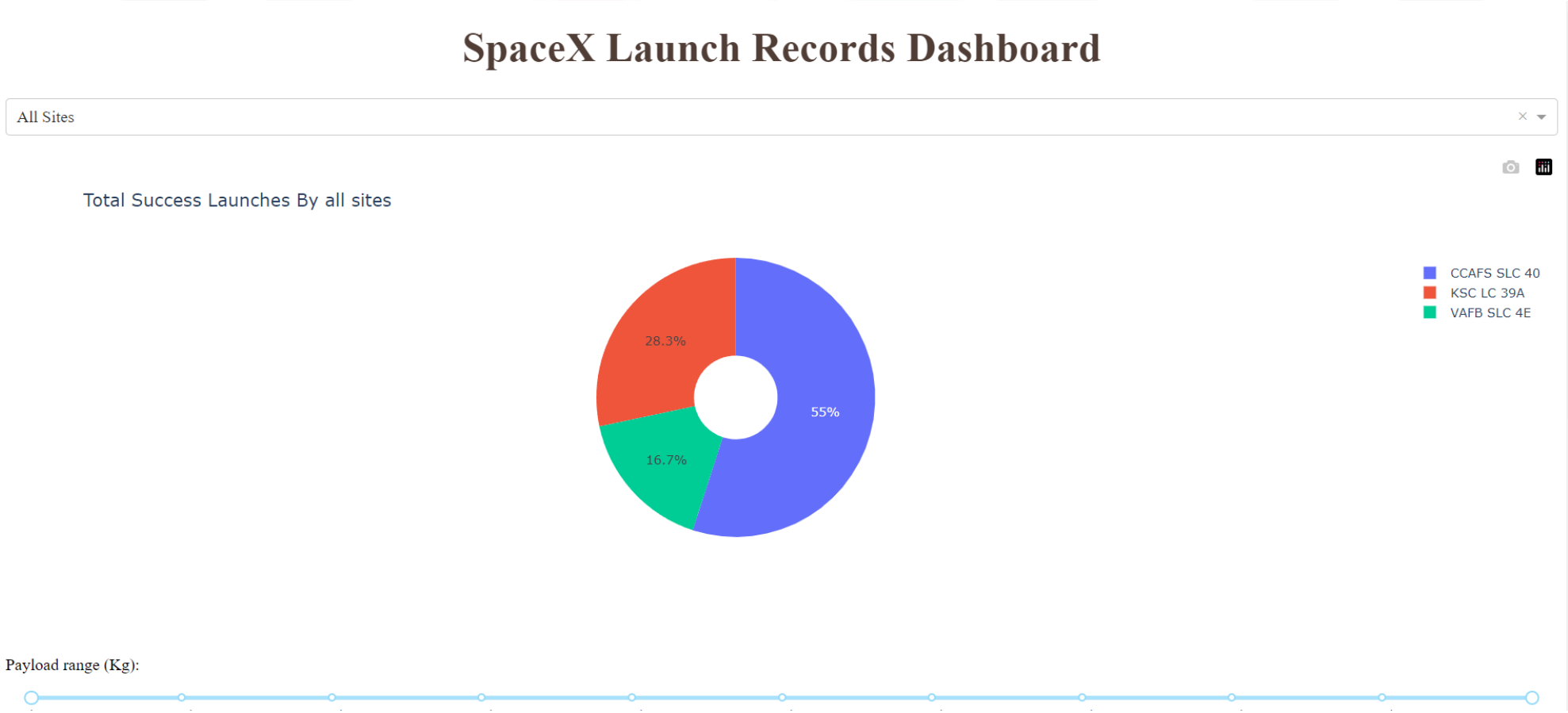
RESULTS: Interactive map with Folium (3/3)

- Distances between a launch site to its proximities



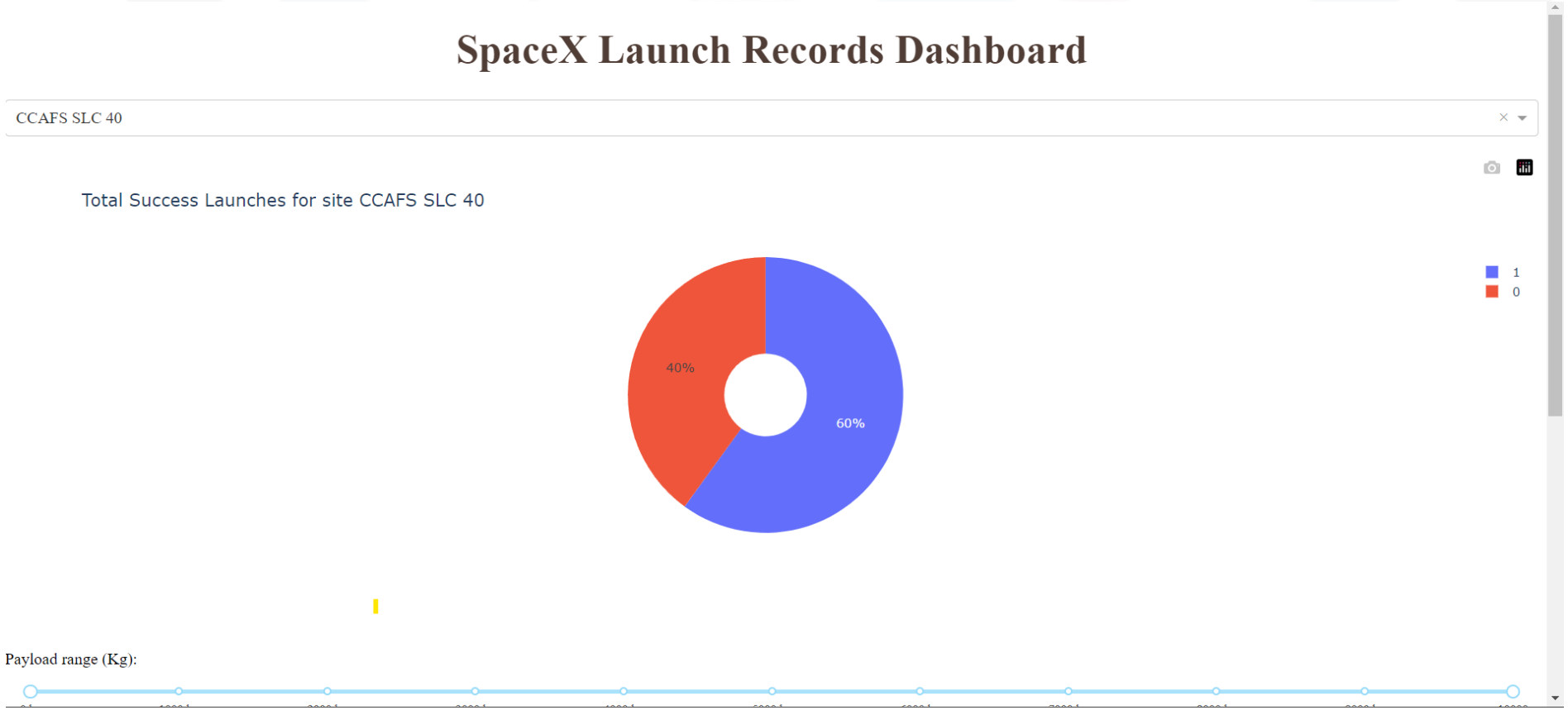
RESULTS: Plotly Dash dashboard (1/3)

- Plotly Dashboard: All sites launch records



RESULTS: Plotly Dash dashboard (2/3)

- Plotly Dashboard: Site CCAFS SLC 40 launch records



RESULTS: Plotly Dash dashboard (3/3)

- Plotly Dashboard: Booster version and Payload

RESULTS: Predictive Analysis Classification Results (1/5)

- Grid Search: Accuracy on the test data using the method score

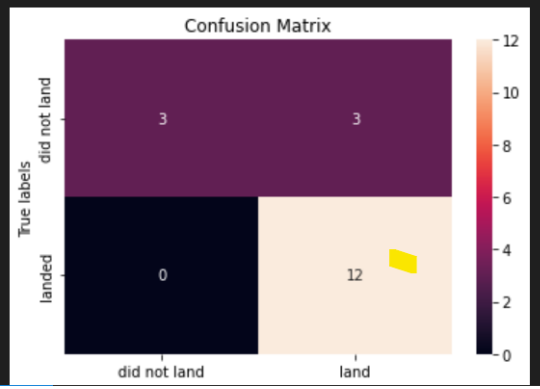
Calculate the accuracy on the test data using the method `score`:

```
[20]
logreg_cv.score(X_test, Y_test)
```

```
0.8333333333333334
```

Lets look at the confusion matrix:

```
[21]
yhat = logreg_cv.predict(X_test)
plot_confusion_matrix(Y_test, yhat)
```



RESULTS: Predictive Analysis Classification Results (2/5)

- Support Vector Machine: Accuracy on the test data using the method score

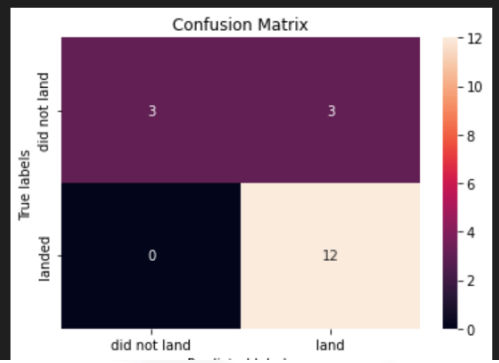
Calculate the accuracy on the test data using the method `score`:

```
svm_cv.score(X_test, Y_test)
```

```
0.8333333333333334
```

We can plot the confusion matrix

```
yhat = svm_cv.predict(X_test)
plot_confusion_matrix(Y_test, yhat)
```



RESULTS: Predictive Analysis Classification Results (3/5)

- Tree Classifier: Accuracy on the test data using the method score

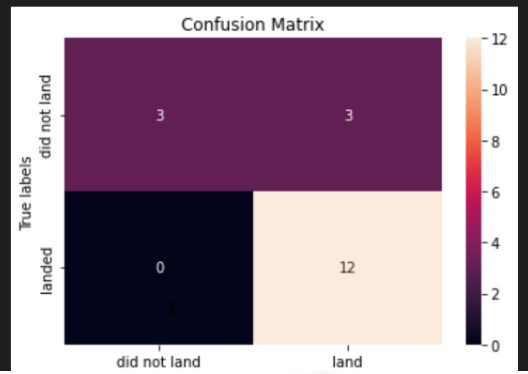
Calculate the accuracy of tree_cv on the test data using the method `score`:

```
tree_cv.score(X_test, Y_test)
```

```
0.8888888888888888
```

We can plot the confusion matrix

```
yhat = svm_cv.predict(X_test)
plot_confusion_matrix(Y_test, yhat)
```



RESULTS: Predictive Analysis Classification Results (4/5)

- K nearest neighbour: Accuracy on the test data using the method score

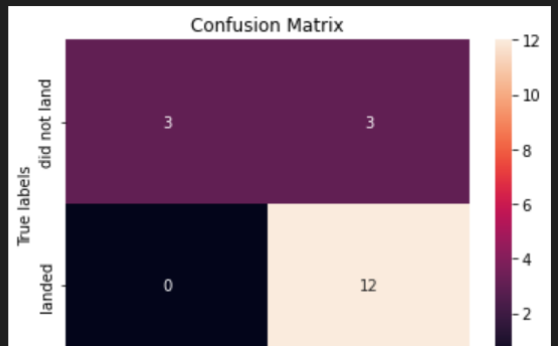
Calculate the accuracy of knn_cv on the test data using the method `score`:

```
[41]
knn_cv.score(X_test, Y_test)
```

```
0.8333333333333334
```

We can plot the confusion matrix

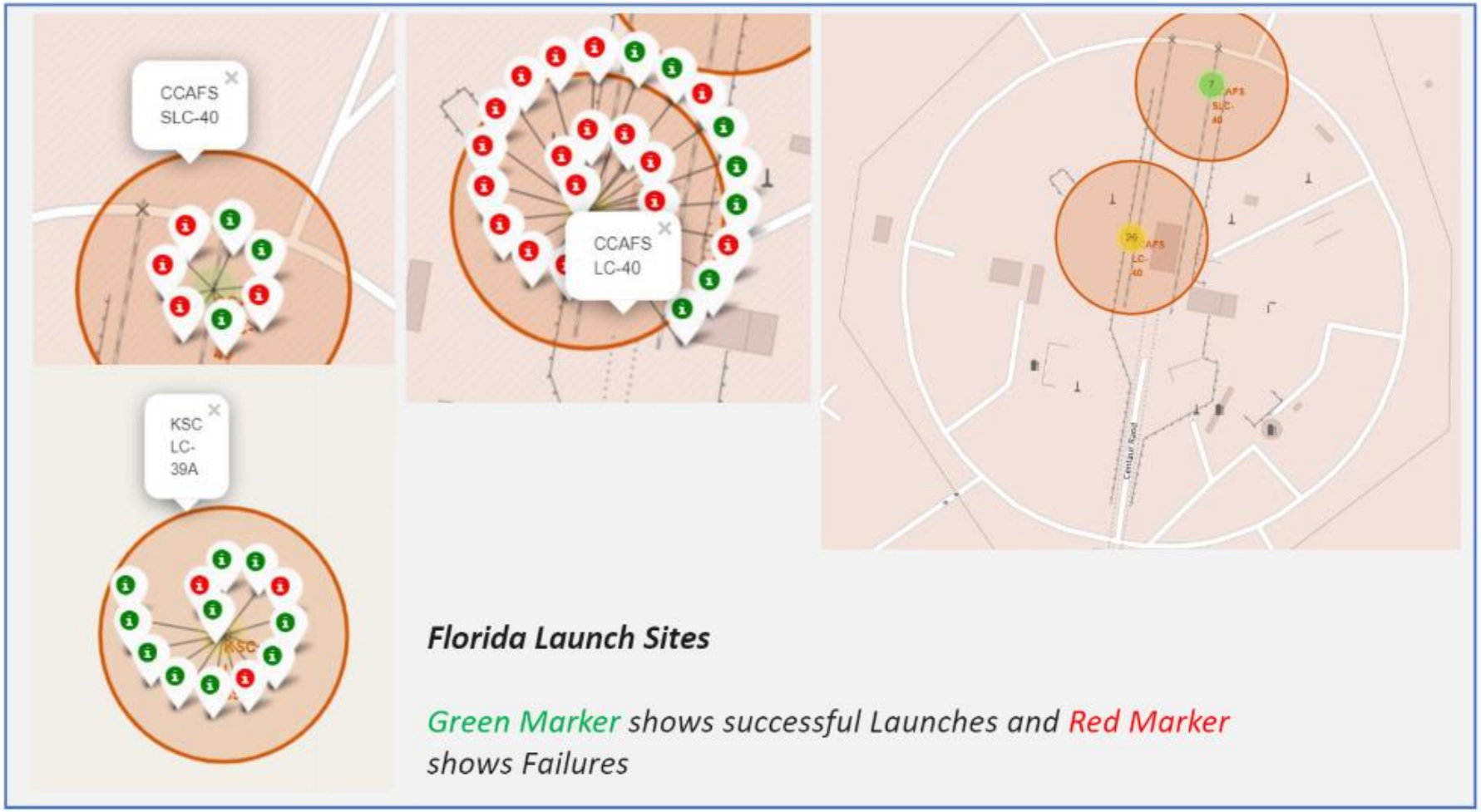
```
[42]
yhat = knn_cv.predict(X_test)
plot_confusion_matrix(Y_test, yhat)
```



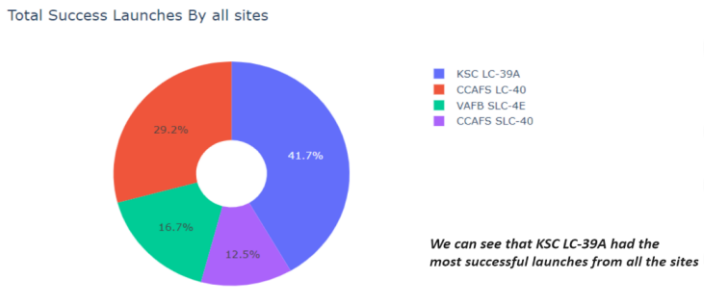
RESULTS: Predictive Analysis Classification Results (5/5)

- Best performing Model: Decision Tree (Score=88%)

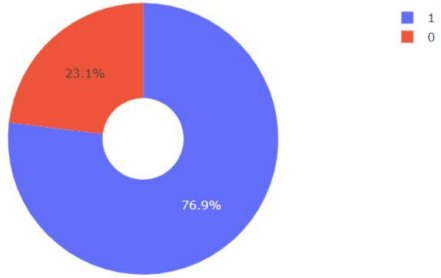
Launch sites



CONCLUSION



We can see that KSC LC-39A had the most successful launches from all the sites



KSC LC-39A achieved a 76.9% success rate while getting a 23.1% failure rate

- The larger the flight amount at a launch site, the greater the success rate at a launch site.
- Launch success rate started to increase in 2013 till 2020.
- Orbits ES-L1, GEO, HEO, SSO, VLEO had the most success rate.
- KSC LC-39A had the most successful launches of any sites.
- The Decision tree classifier is the best machine learning algorithm for this task.