Rocket Launch Cost Prediction

Jhon

11/09/2024

IBM Developer

SKILLS NETWORK

OUTLINE

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

EXECUTIVE SUMMARY

- Forecasting the success rate of Falcon 9 first stage landings
- Identifying key factors that impact the chances of a successful landing
- Calculating the total expenses associated with a launch
- Offering strategic insights for companies aiming to challenge SpaceX
- Creating a predictive model to estimate the probability of successful landings for the first stage

INTRODUCTION

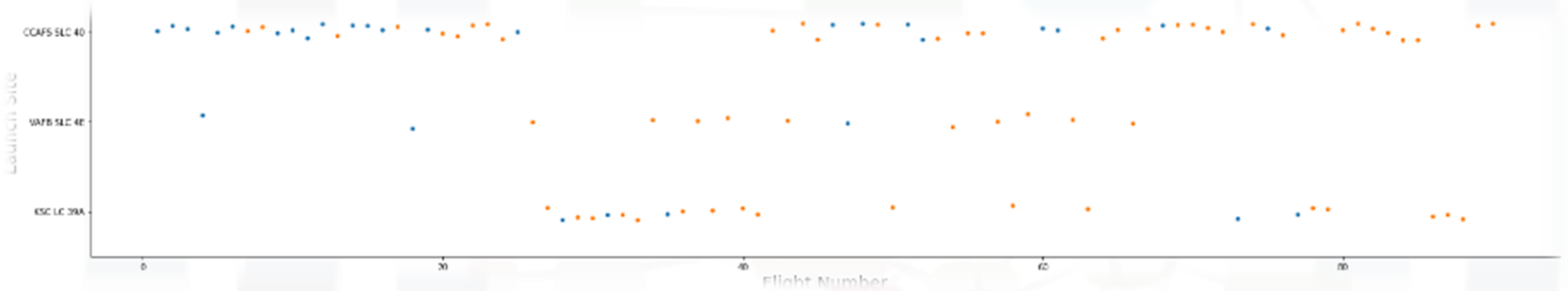
- Traditional rocket providers typically charge $165 million per launch.
- SpaceX's innovative approach of reusing the Falcon 9's first stage has significantly cut down launch costs.
- The Falcon 9 launch price is $62 million.
- Compared to traditional providers, SpaceX's reuse of the Falcon 9 first stage has substantially reduced the cost of launches.

METHODOLOGY

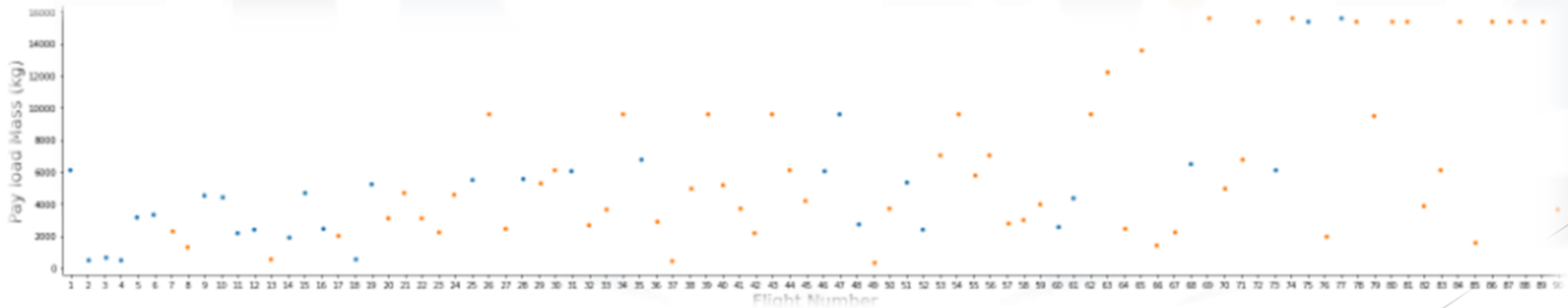
- Collecting Data via:
- SpaceX API
- Web Scraping
- Data Wrangling
- Analyzing and Visualizing Data using:
- SQL
- Pandas
- Machine Learning Prediction for Successful First Stage Landings

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

- Visualize the relationship between Flight Number and Launch Site

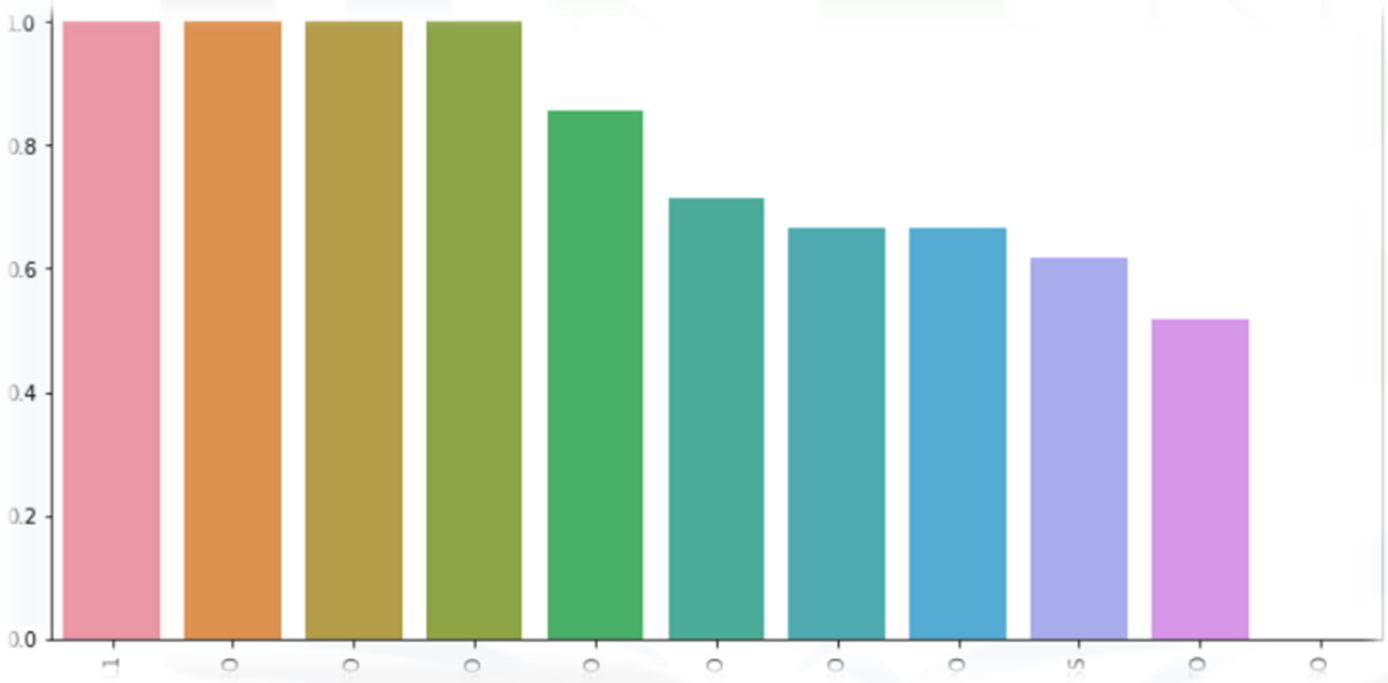


- Visualize the relationship between Payload and Launch Site



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

- Success rate of each orbit type



IBM Developer

SKILLS NETWORK

METHODOLOGY: Predictive Analysis

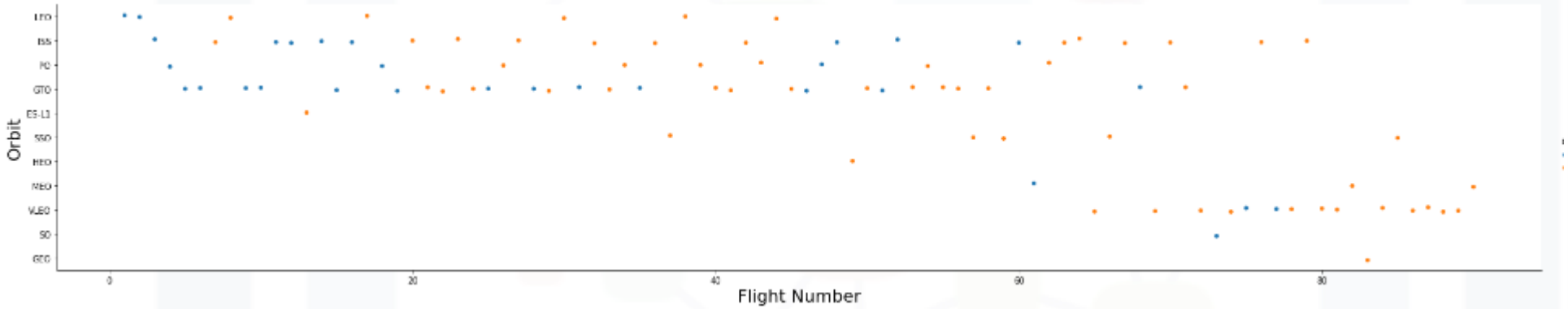
- A higher number of flights at a launch site is associated with a higher success rate at that site.
- The orbits ES-L1, GEO, HEO, SSO, and VLEO had the highest success rates.
- For the LEO orbit, success is linked to the number of flights, while for the GTO orbit, there is no significant correlation between flight numbers and success.

IBM Developer

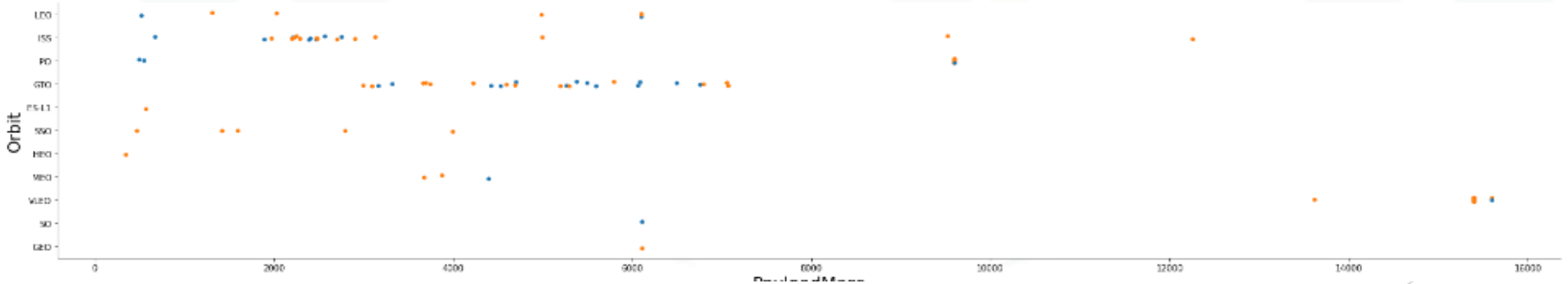
SKILLS NETWORK

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

	firstsuccessfull_landing_date
0	2015-12-22

- Booster_versions

	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)

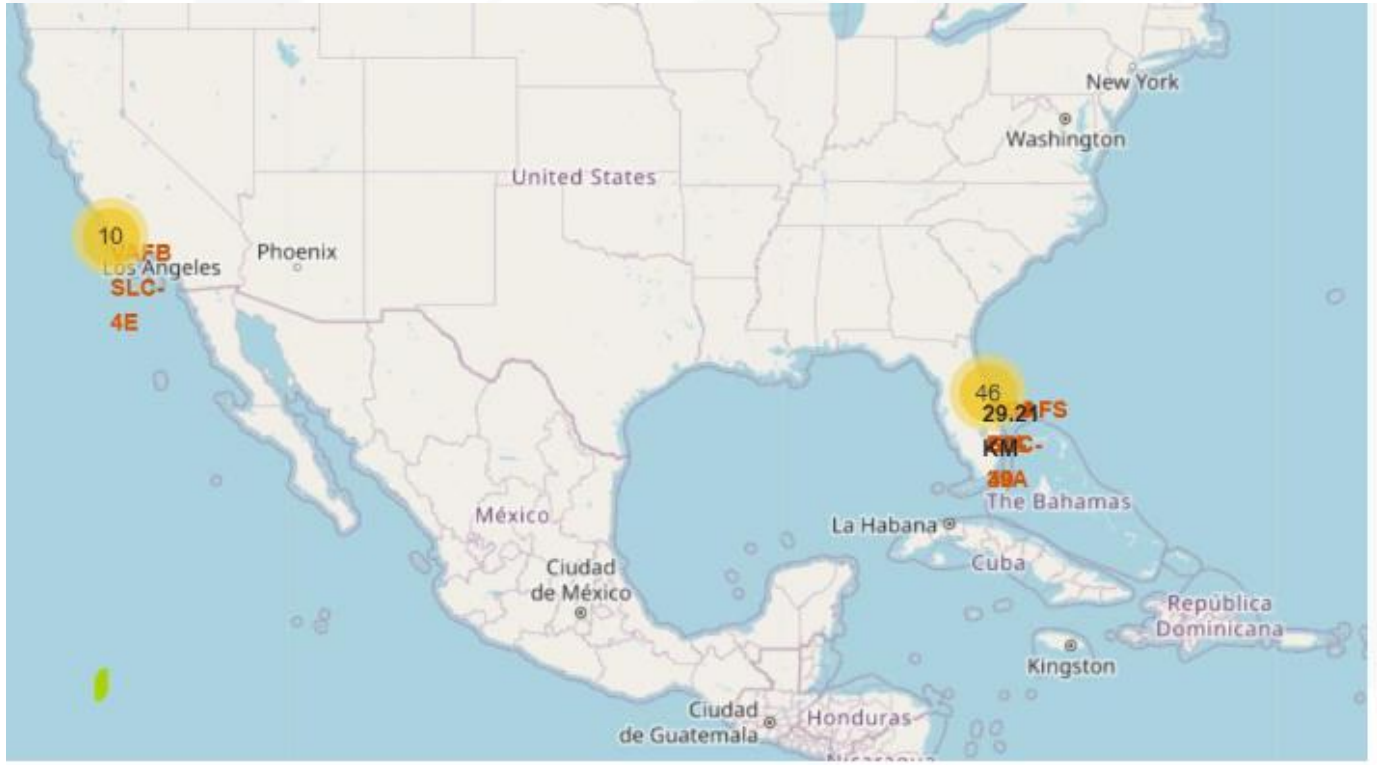
- all launch sites on a map

RESULTS: Interactive map with Folium (2/3)

- Success/Failed launches for each site on the map

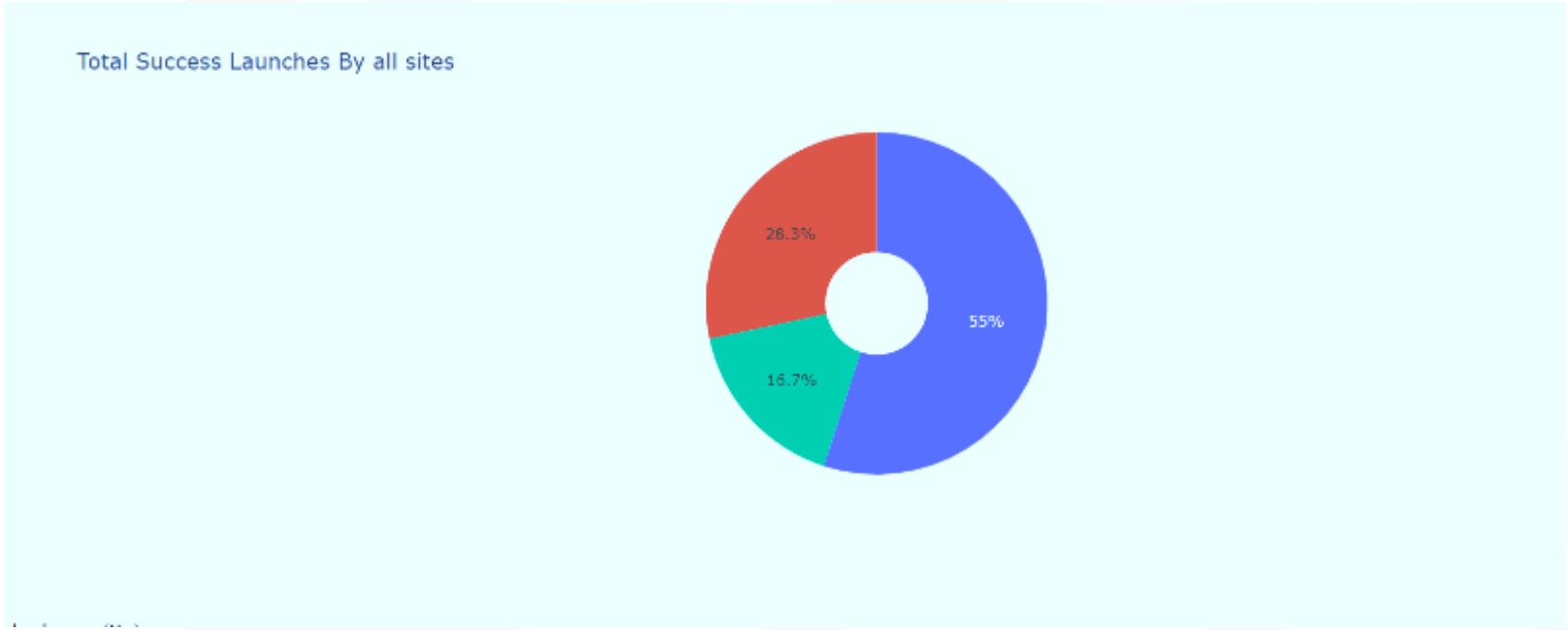
RESULTS: Interactive map with Folium (3/3)

- Distances between a launch site to its proximities



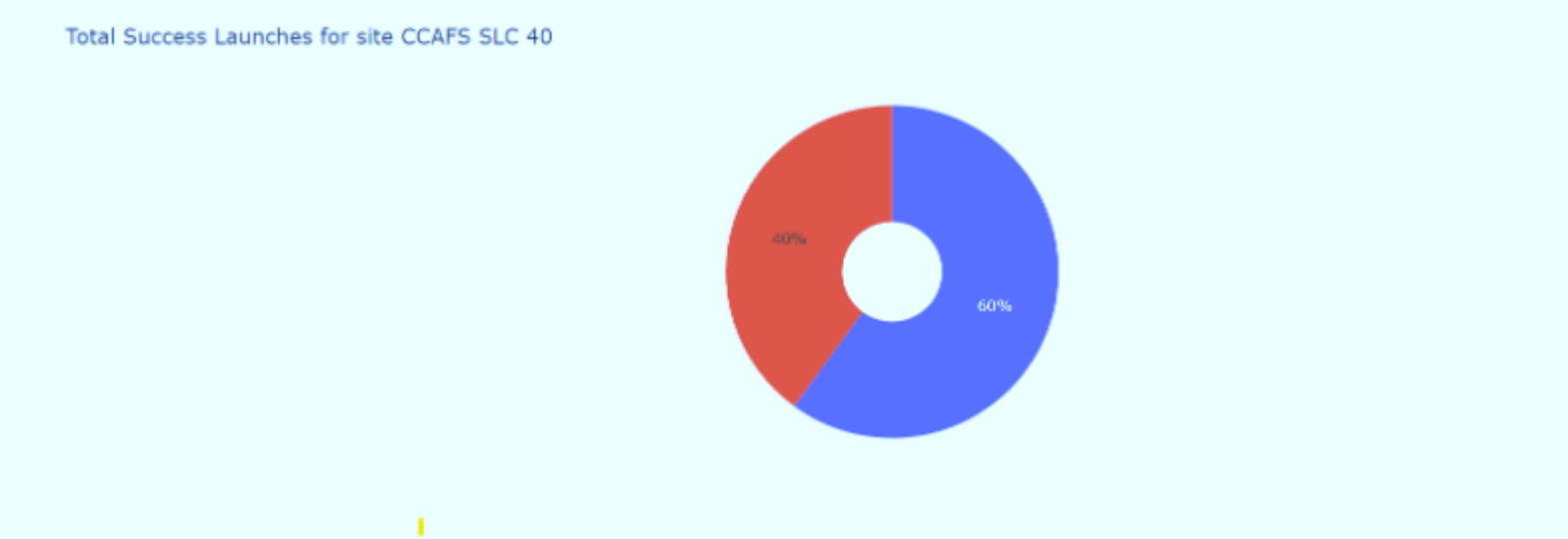
RESULTS: Plotly Dash dashboard (1/3)

- Plotly Dashboard: All sites launch record



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 (2/5) y (1/5)

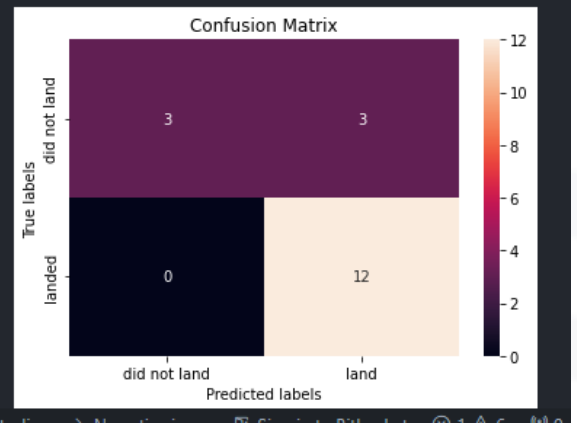
- Support Vector Machine: Accuracy on the test data using the meth

```
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)
```

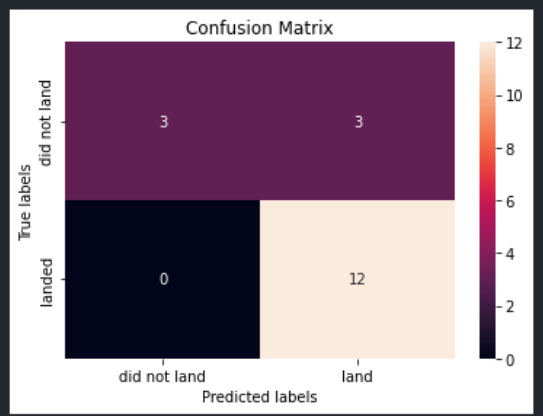


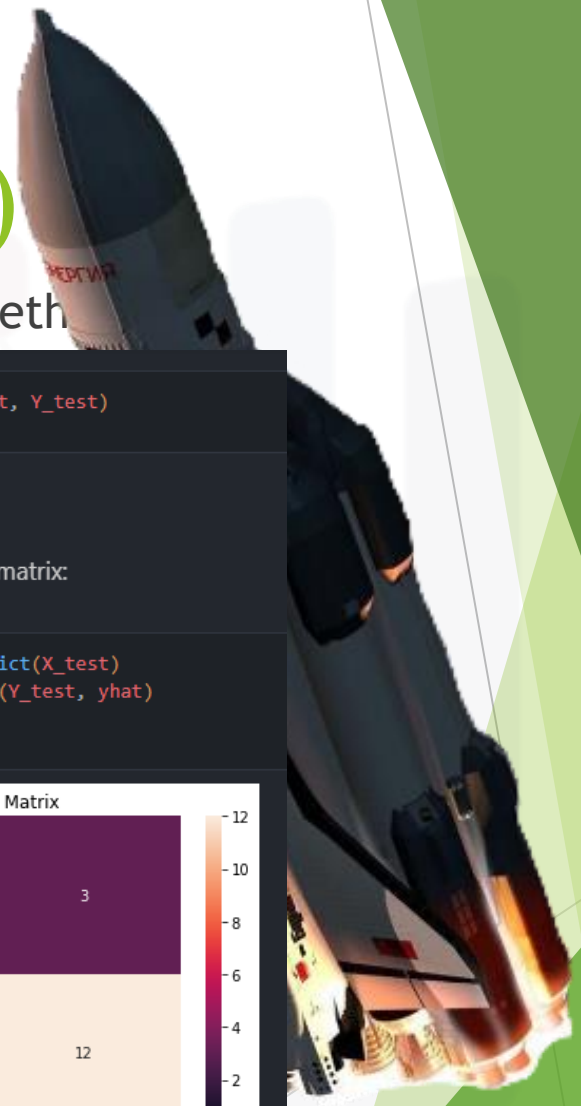

```
logreg_cv.score(X_test, Y_test)
```

```
0.8333333333333334
```

Lets look at the confusion matrix:

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





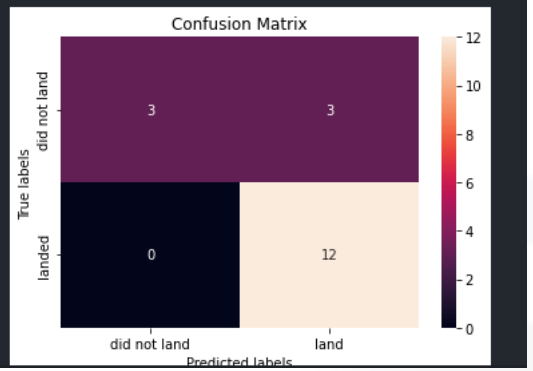
RESULTS: Predictive Analysis Classification Results (4/5) y (3/5)

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

```
[41]  knn_cv.score(X_test, Y_test)
...   0.8333333333333334
```

We can plot the confusion matrix

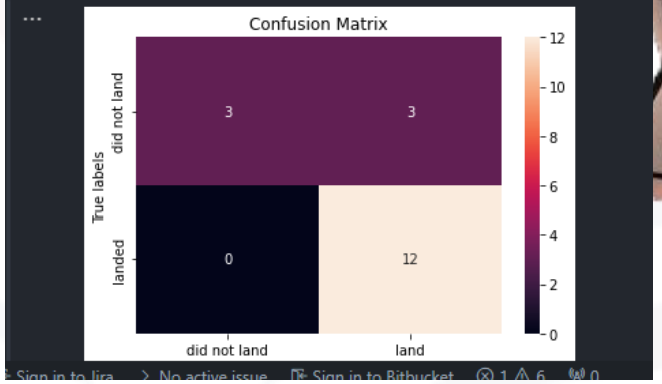
```
[42]  yhat = knn_cv.predict(X_test)
      plot_confusion_matrix(Y_test, yhat)
```




```
[34]  tree_cv.score(X_test, Y_test)
...   0.8888888888888888
```

We can plot the confusion matrix

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



IBM Developer

SKILLS NETWORK

RESULTS: Predictive Analysis Classification Results (5/5)

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

```
models = {
    "KNeighbors": knn_cv.best_score_,
    "DecisionTree": tree_cv.best_score_,
    "LogisticRegression": logreg_cv.best_score_,
    "SupportVector": svm_cv.best_score_,
}

bestalgorithm = max(models, key=models.get)
print(f"Best model is {bestalgorithm=}, with a score of, {models[bestalgorithm]}")

if bestalgorithm == "DecisionTree":
    print("Best params is :", tree_cv.best_params_)
if bestalgorithm == "KNeighbors":
    print("Best params is :", knn_cv.best_params_)
if bestalgorithm == "LogisticRegression":
    print("Best params is :", logreg_cv.best_params_)
if bestalgorithm == "SupportVector":
    print("Best params is :", svm_cv.best_params_)
```

Python

```
Best model is bestalgorithm='DecisionTree', with a score of, 0.8875
Best params is : {'criterion': 'entropy', 'max_depth': 12, 'max_features': 'sqrt', 'min_samples_leaf': 1, 'min_samples_split': 2, 'splitter': 'random'}
```

CONCLUSION

- A higher number of flights at a launch site correlates with a higher success rate.
- The success rate of launches began to rise from 2013 and continued improving until 2020.
- The orbits ES-L1, GEO, HEO, SSO, and VLEO exhibited the highest success rates.
- KSC LC-39A had the highest number of successful launches among all sites.
- The Decision Tree Classifier is the most effective machine learning algorithm for this task.