SPACEX Launches

Sidra

4 August 2023

OUTLINE

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

EXECUTIVE SUMMARY

- This data science project aims to leverage historical launch data and machine learning techniques to predict the likelihood of SpaceX reusing the first stage of their rockets. The insights gained from the analysis could have far-reaching implications for SpaceX's operational strategies and the broader aerospace sector's approach to reusability.
- The aim of this data exploratory project is to develop a machine learning model using publicly available data to predict the likelihood of SpaceX reusing the first stage of their rockets after a launch. By analyzing historical launch data and mission-specific factors, the project seeks to provide insights into the reusability strategy of SpaceX.
- The project will gather a dataset of SpaceX launch information, encompassing details about launch dates, launch sites, payload mass, and outcomes (successful reuse or not). Publicly accessible sources, such as SpaceX's official records and launch databases, will serve as the primary data sources. The collected data will undergo thorough preprocessing to handle missing values.
- After preprocessing, relevant features will be selected to train a machine learning model. Various machine learning algorithms will be evaluated, such as logistic regression, and decision trees to determine the most effective predictive model.
- The dataset will be divided into training and testing sets to simulate real-world performance. Metrics such as accuracy score and confusion matrix will be employed to quantify the model's predictive capability.

INTRODUCTION

- In the rapidly evolving landscape of space exploration, SpaceX has emerged as a pioneering force, revolutionizing the way we access and utilize space. One of the company's groundbreaking innovations has been the concept of reusing rocket components, specifically the first stage of their Falcon 9 rockets.
- This data exploratory project uses data science to uncover the factors that influence SpaceX's first stage reusability. By using machine learning and analyzing publicly available data, this project seeks to predict the likelihood of SpaceX reusing the first stage after each launch. The insights garnered from this predictive model could provide a deeper understanding of SpaceX's operational decisions and reveal patterns that contribute to successful reusability.
- The project revolves around the idea that historical launch data, when combined with payload specifications, and mission-specific outcomes, can offer valuable insights into SpaceX's reusability strategy. The analysis and predictions derived from this model could potentially guide decisions, streamline operational processes, while maintaining efficiency for the new SpaceY's space missions.

METHODOLOGY

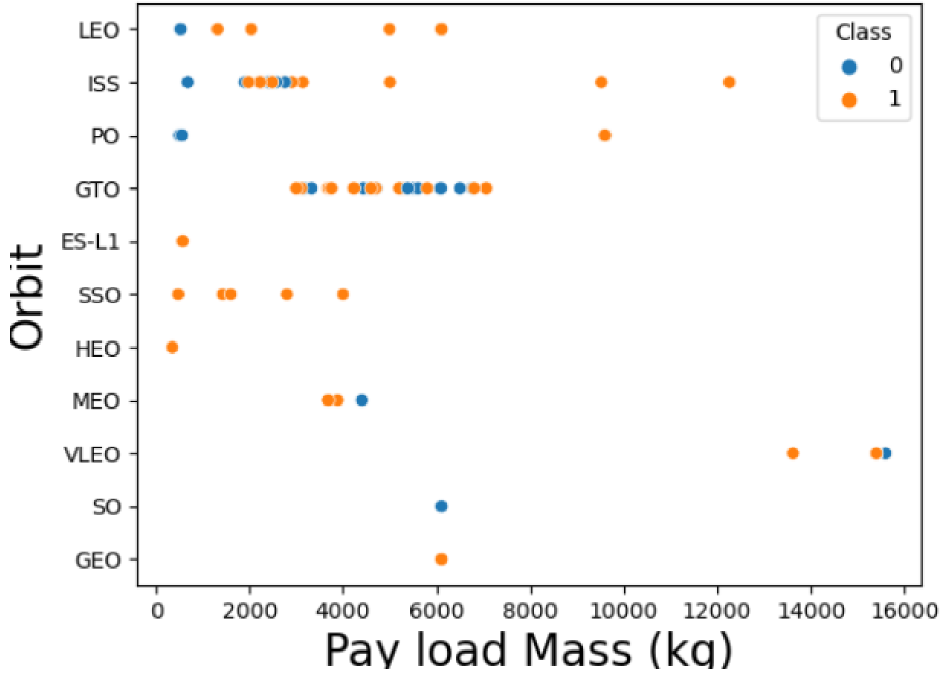
- The primary source of data for this analysis will be SpaceX's official records, including launch dates, mission outcomes, and details about the first stage reusability status. These records will serve as the foundation for building the predictive model.
- The project will begin by systematically collecting data from SpaceX's official records and launch databases. Information regarding launch dates, mission objectives, payload types, and reusability outcomes will be compiled.
- Exploratory Data Analysis (EDA) will involve visualizing and summarizing the data to gain initial insights into potential correlations between variables and reusability outcomes. This step will guide feature selection and model development.
- Following the EDA a model will be developed based on selected features such as launch data, payload mass and outcomes. Based on these features an appropriate model will be tested and trained, before the model's hyperparameters can be fine-tuned and the model can be evaluated.

RESULTS

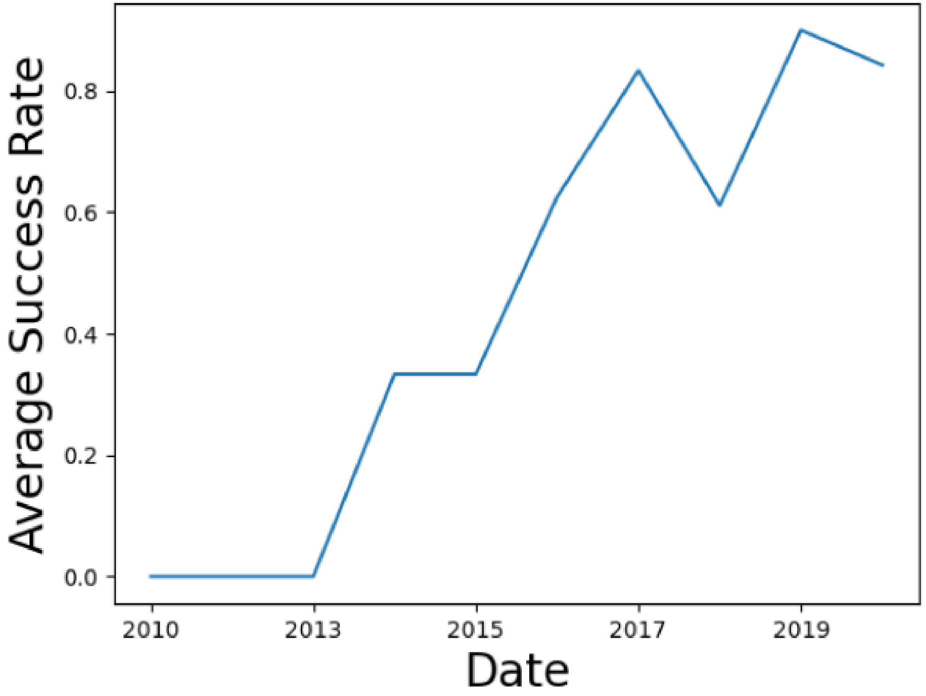
- SpaceX's official records, launch databases, and mission reports, were systematically collected to create a comprehensive dataset.
- Information encompassing launch details, payload specifications, orbit, booster version and mission outcomes was organized into an appropriate dataframe.
- Exploratory Data Analysis (EDA) was employed to visualize and uncover patterns between variables and reusability outcomes.
- Key features influencing reusability were selected based on EDA insights, and machine learning models (e.g., logistic regression, decision trees) were developed using these features.
- Models were evaluated using metrics like accuracy score, and confusion matrix to gauge predictive performance.

EDA with visualization result

Relationship between Payload and Orbit type



Average success rate over the years



EDA with visualization result FINDINGS & IMPLICATIONS

Findings

- It can be observed that the success rate since 2013 kept increasing till 2020
- With heavy payloads the successful landing or positive landing rate are more for Polar,LEO and ISS.
- However for GTO we cannot distinguish this well as both positive landing rate and negative landing(unsuccessful mission) are both there here.

Implications

- The increasing success rate of first stage reusability, could simply mean that SpaceX has been refining their launch processes, and that continuous efforts have led to enhanced reusability success, contributing to cost savings.
- The varying success rates based on payload type and orbits (Polar, LEO, ISS) suggest that SpaceX might be tailoring their reusability strategies to different mission profiles. This may mean that payload mass and mission profiles play a crucial role in determining the reusability success.
- The inability to clearly distinguish positive landing rates from negative landing rates for missions to Geostationary Transfer Orbit (GTO) raises questions about the unique challenges associated with these types of missions, and may require further investigation.

EDA with SQL results

Boosters with success in drone ship and a payload mass > 4000 but < 6000

Booster_Version
F9 FT B1021.1
F9 FT B1022
F9 FT B1023.1
F9 FT B1026
F9 FT B1021.2
F9 FT B1029.2
F9 FT B1038.1
F9 FT B1031.2
F9 B4 B1042.1
F9 B5 B1046.1

Total payload mass carried by boosters launched by NASA (CRS)

```
%sql select SUM("PAYLOAD_MASS__KG_") from SPACEXTABLE where "Customer"= "NASA (CRS)"
```

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

SUM("PAYLOAD_MASS__KG_")
45596

First successful landing outcome in ground pad was achieved

```
%sql select MIN(Date) from SPACEXTABLE where "Landing_Outcome"= "Success (ground pad)"
```

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

MIN(Date)
2015-12-22

EDA with SQL results

Launch sites in the space mission

```
%sql select distinct "Launch_Site" from SPACEXTABLE
```

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

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

Booster versions which have carried the maximum payload mass

Booster_VersionF9 B5
B1048.4F9 B5 B1049.4F9
B5 B1051.3F9 B5
B1056.4F9 B5 B1048.5F9
B5 B1051.4F9 B5
B1049.5F9 B5 B1060.2F9
B5 B1058.3F9 B5
B1051.6F9 B5 B1060.3F9
B5 B1049.7

EDA with SQL - FINDINGS & IMPLICATIONS

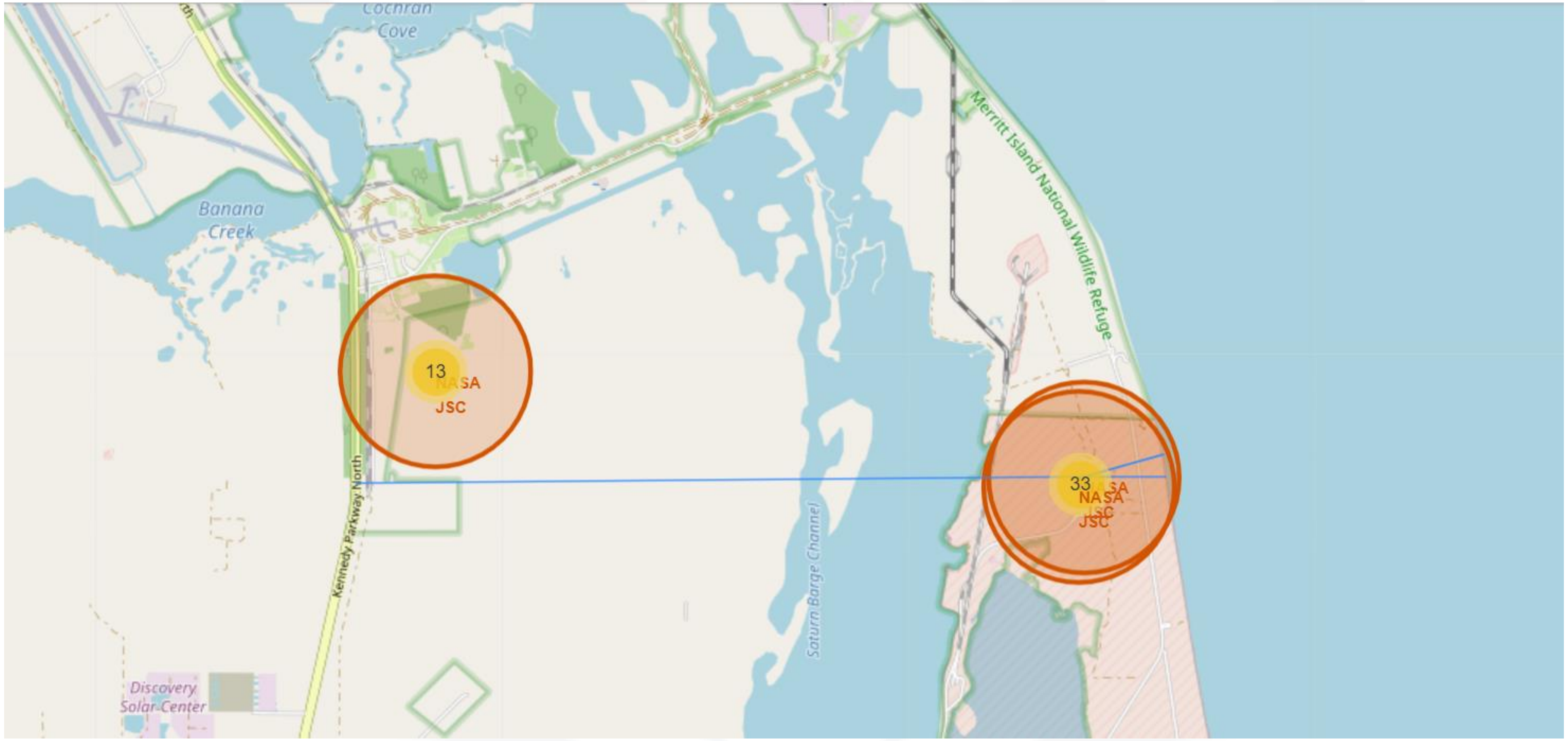
Findings

- The 4 launch sites for these missions include CCAFS LC-40, VAFB SLC-4E, KSC LC-39A and CCAFS SLC-40.
- Most 'F9 FT' booster versions had a success in drone ship and a payload mass > 4000 but < 6000
- Total payload mass carried by boosters launched by NASA (CRS) was 45596 kg
- First successful landing outcome in ground pad was achieved on 22nd of December in 2015.

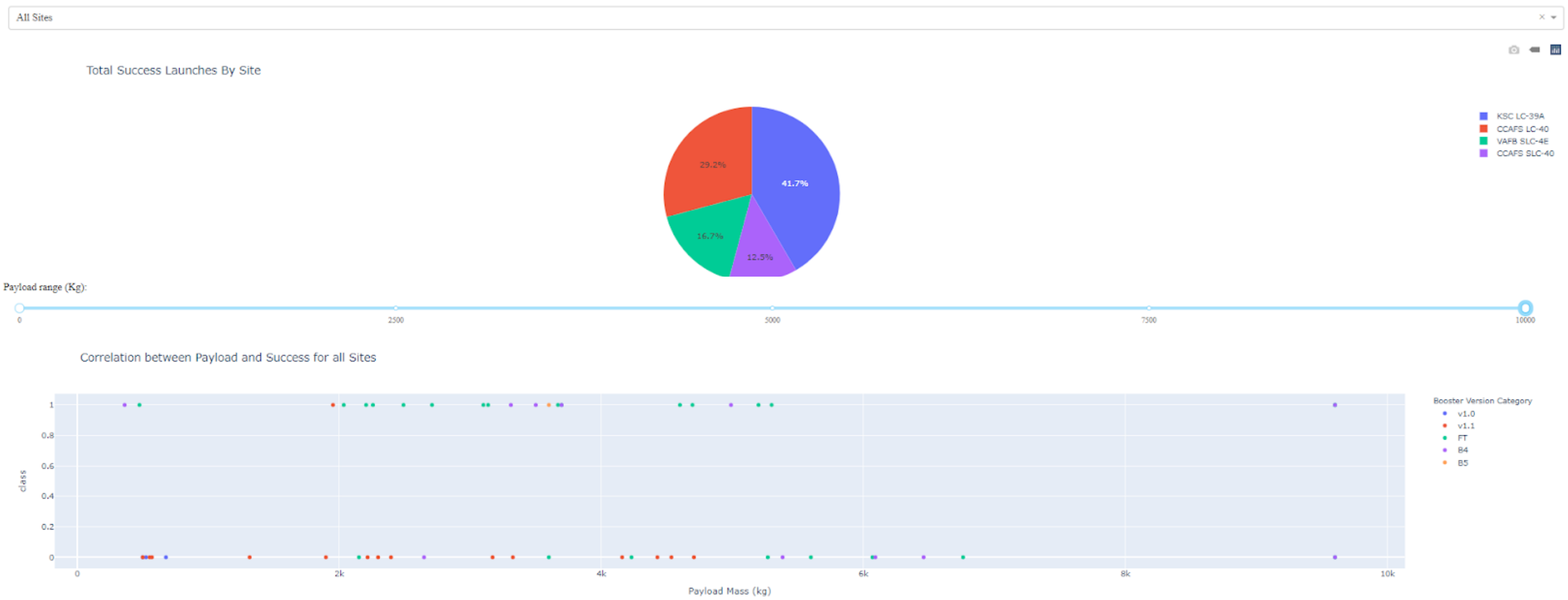
Implications

- The presence of four distinct launch sites indicates that SpaceX might be strategically selecting launch sites based on mission requirements which could be driven by considerations such as proximity to desired orbits, allowing SpaceX to maximize mission success and reusability outcomes.
- The correlation between 'F9 FT' booster versions, success in drone ship landings, and a specific payload mass range (4000-6000 kg) suggests that SpaceX may have designed these booster versions to excel in particular scenarios.
- The observation of payload mass carried by boosters launched by NASA (CRS) may imply that NASA's collaboration with SpaceX is significant in terms of payload utilization and that the partnership between them is fostering technological advancements and contributing by efficiently delivering substantial payload masses to space.

Interactive map with Folium results slides



Plotly Dash dashboard results



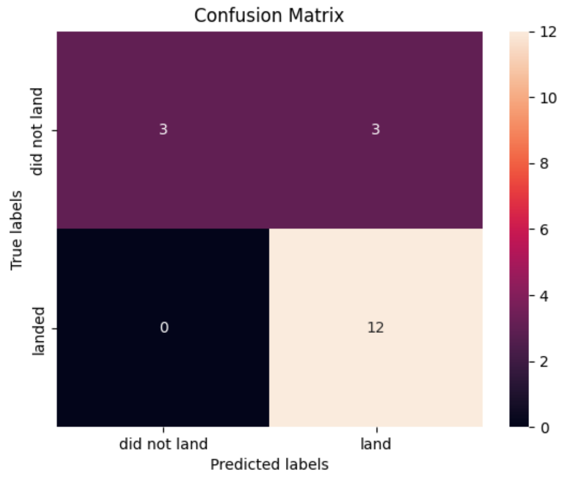
Predictive analysis (classification) results

Logistic Regression

```
[36]: logreg_cv.score(X_test,Y_test)
[36]: 0.8333333333333334
```

Lets look at the confusion matrix:

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



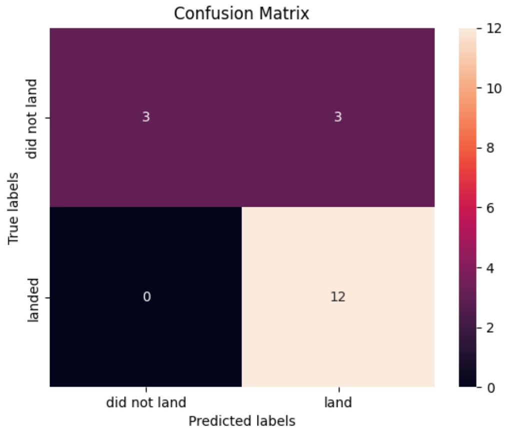
SVM

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

```
[60]: svm_cv.score(X_test,Y_test)
[60]: 0.8333333333333334
```

We can plot the confusion matrix

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



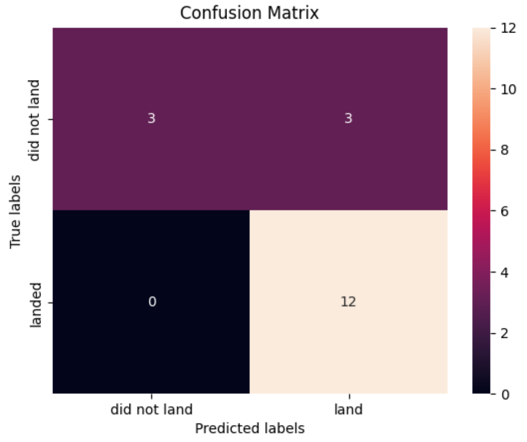
Decision Trees

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

```
[52]: tree_cv.score(X_test,Y_test)
[52]: 0.9444444444444444
```

We can plot the confusion matrix

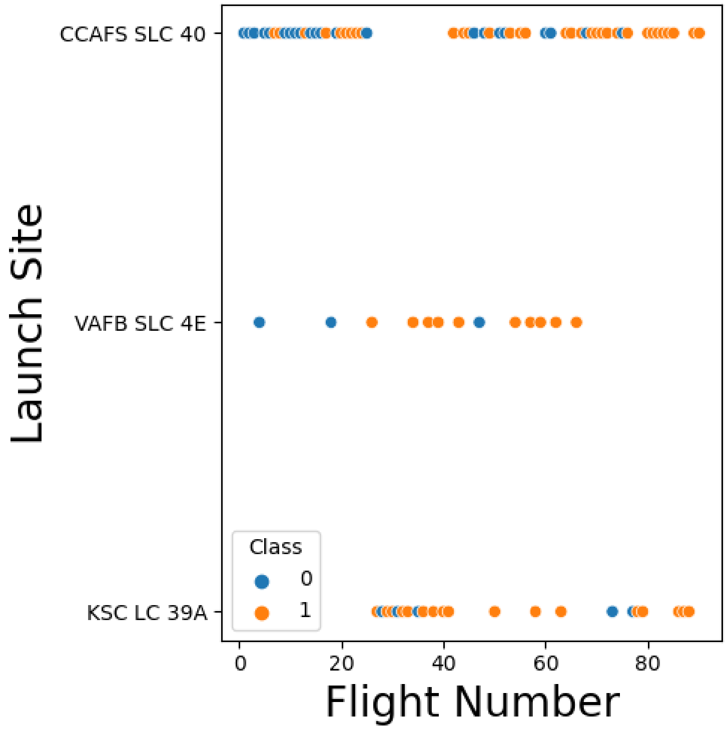
```
[53]: yhat = svm_cv.predict(X_test)
      plot_confusion_matrix(Y_test,yhat)
```

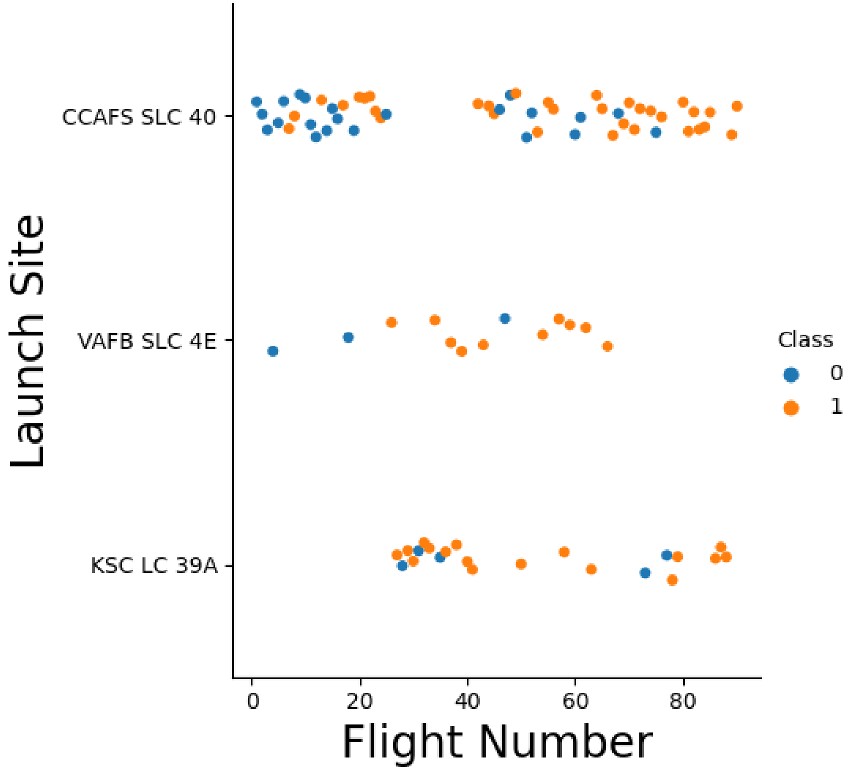


CONCLUSION

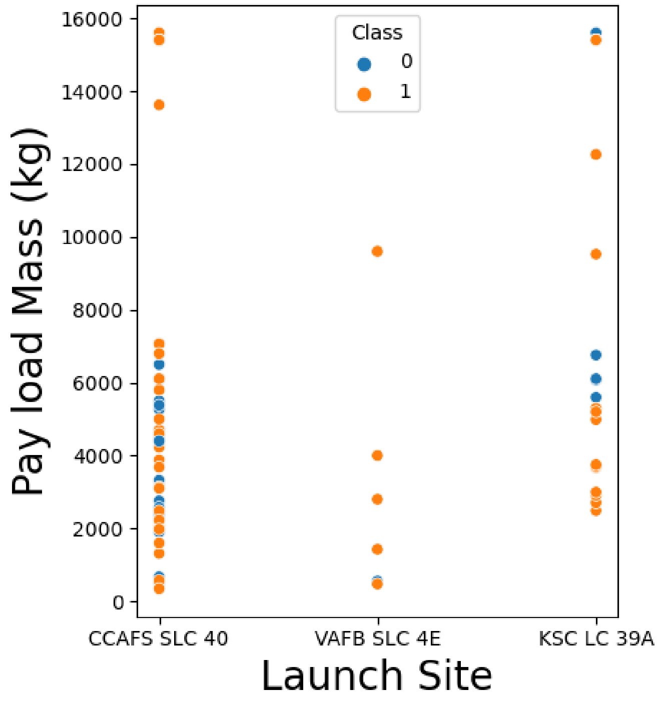
- The observed upward trend in first stage reusability success rates from 2013 to 2020 shows SpaceX's continuous commitment. This evolution signifies the ability to learn from each launch, refine processes, and apply lessons to subsequent missions, resulting in improved reusability outcomes
- The varying success rates across different payload types and mission destinations indicate a strategic approach to reusability based on specific mission requirements. The higher success rates for heavy payloads to Polar, LEO, and ISS orbits suggest tailored strategies for these missions.
- The challenges associated with Geostationary Transfer Orbit (GTO) missions, where distinguishing between successful and unsuccessful outcomes is less clear, highlight the intricacies of achieving reusability under specific conditions.
- Incorporating these insights into its business strategy and operations can empower the new company SPACEY to navigate challenges, optimize reusability, and establish itself as a competitive player in the industry. By learning from SpaceX's experiences, the new company can accelerate its growth and contribute to the space exploration.

APPENDIX

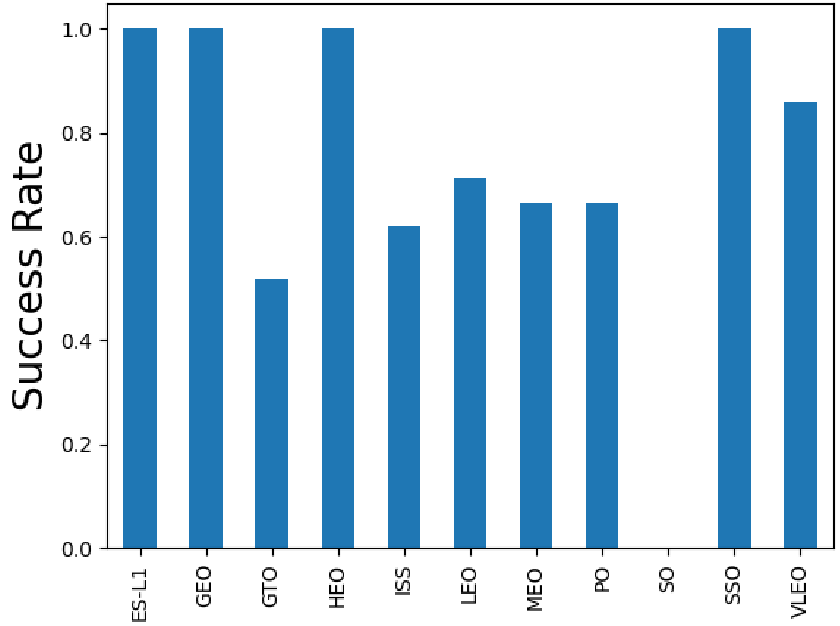

Launch Site
CCAFS SLC 40
VAFB SLC 4E
KSC LC 39A
Class
0
1
0
20
40
60
80
Flight Number


Launch Site
CCAFS SLC 40
VAFB SLC 4E
KSC LC 39A
Class
0
1
0
20
40
60
80
Flight Number

APPENDIX



Class
0
1
Pay load Mass (kg)
16000
14000
12000
10000
8000
6000
4000
2000
0
CCAFS SLC 40
VAFB SLC 4E
KSC LC 39A
Launch Site



Success Rate
1.0
0.8
0.6
0.4
0.2
0.0
ES-L1
GEO
GTO
HEO
ISS
LEO
MEO
PO
SO
SSO
VLEO