SpaceX Launch Insights

Diego B.N.

April 4th, 2024

OUTLINE

- Executive Summary
- Introduction
- Methodology
- Results
  - Visualization of most relevant insights.
  - Dashboard Insights
- Discussion
  - Findings & Implications of the data.
- Conclusion
- Appendix
  - A few extra slides for data figures!

EXECUTIVE SUMMARY

Opting for launches at KSC LC-39A presents clients with the highest potential for success while also offering cost-saving benefits. Clients are advised to choose this launch site to maximize the likelihood of successful launches and optimize their budget allocation.

Clients utilizing SpaceX reusable rockets stand to benefit from increased cost savings and improved launch success rates over time. By partnering with us, clients can expect to launch payloads at significantly reduced costs compared to market prices, enhancing their overall efficiency and competitiveness.

Leveraging machine learning-based insights, our clients gain access to strategic recommendations on optimal launch locations and timing, tailored to their specific needs. This advanced analytical support enhances decision-making processes, enabling clients to capitalize on discounted rates and maximize the success of their payload launches.

INTRODUCTION

Welcome to my data capstone project, where I revolutionize the world of payload launches through cutting-edge machine learning and data-driven insights. As a forward-thinking project, I specialize in providing our clientele with strategic recommendations on when to opt for Falcon 9 rockets and SpaceX launch services, ensuring not only cost-efficiency but also a guaranteed success rate for payload deployments. Our commitment to leveraging advanced technologies and predictive analytics enables us to offer unparalleled advantages to clients seeking optimal launch solutions. In this introductory guide, we delve into the compelling findings and implications of our research, demonstrating how partnering with us translates into enhanced launch success, substantial cost savings, and a competitive edge in the aerospace industry.

METHODOLOGY

- In conducting our comprehensive analysis to provide clients with actionable insights on optimal payload launch strategies, we employed a multifaceted approach leveraging a diverse array of tools and techniques. Our methodology encompassed data acquisition, preprocessing, analysis, and machine learning model development, all aimed at delivering robust and reliable recommendations.
  - Data Acquisition:
    - We accessed relevant data sources using a combination of the Pandas API and web scraping techniques. By tapping into publicly available repositories and proprietary databases, we compiled a rich dataset comprising historical launch records, performance metrics, and contextual information vital for our analyses.
  - Data Preprocessing:
    - Prior to analysis, we subjected the raw data to meticulous wrangling processes to ensure its integrity and suitability for modeling. This involved handling missing values, standardizing formats, and performing feature engineering to extract meaningful insights from the dataset.
  - Analysis Tools:
    - Our analytical arsenal comprised a suite of powerful Python libraries, including Pandas, Dash, Plotly, Matplotlib, Seaborn, and Scikit-learn. These tools facilitated exploratory data analysis, visualization, and machine learning model development, enabling us to extract actionable intelligence from the dataset.
  - Machine Learning Model Development:
    - To predict launch success and optimize decision-making, we deployed advanced machine learning algorithms such as K-nearest Neighbors, logistic regression, and Support Vector Machines (SVM). Through rigorous training, validation, and fine-tuning, we iteratively refined our models to achieve high levels of accuracy and predictive performance.
  - Exploratory Data Analysis (EDA) with SQL:
    - In addition to Python-based analysis, we conducted exploratory data analysis using SQL, leveraging its querying capabilities to unearth hidden patterns, trends, and correlations within the dataset. This complementary approach provided deeper insights into the underlying dynamics of launch success and informed our modeling efforts.
- In summary, our methodology represents a holistic and data-driven approach to delivering actionable insights to clients, combining cutting-edge technologies with rigorous analytical processes to optimize payload launch strategies and drive business success.

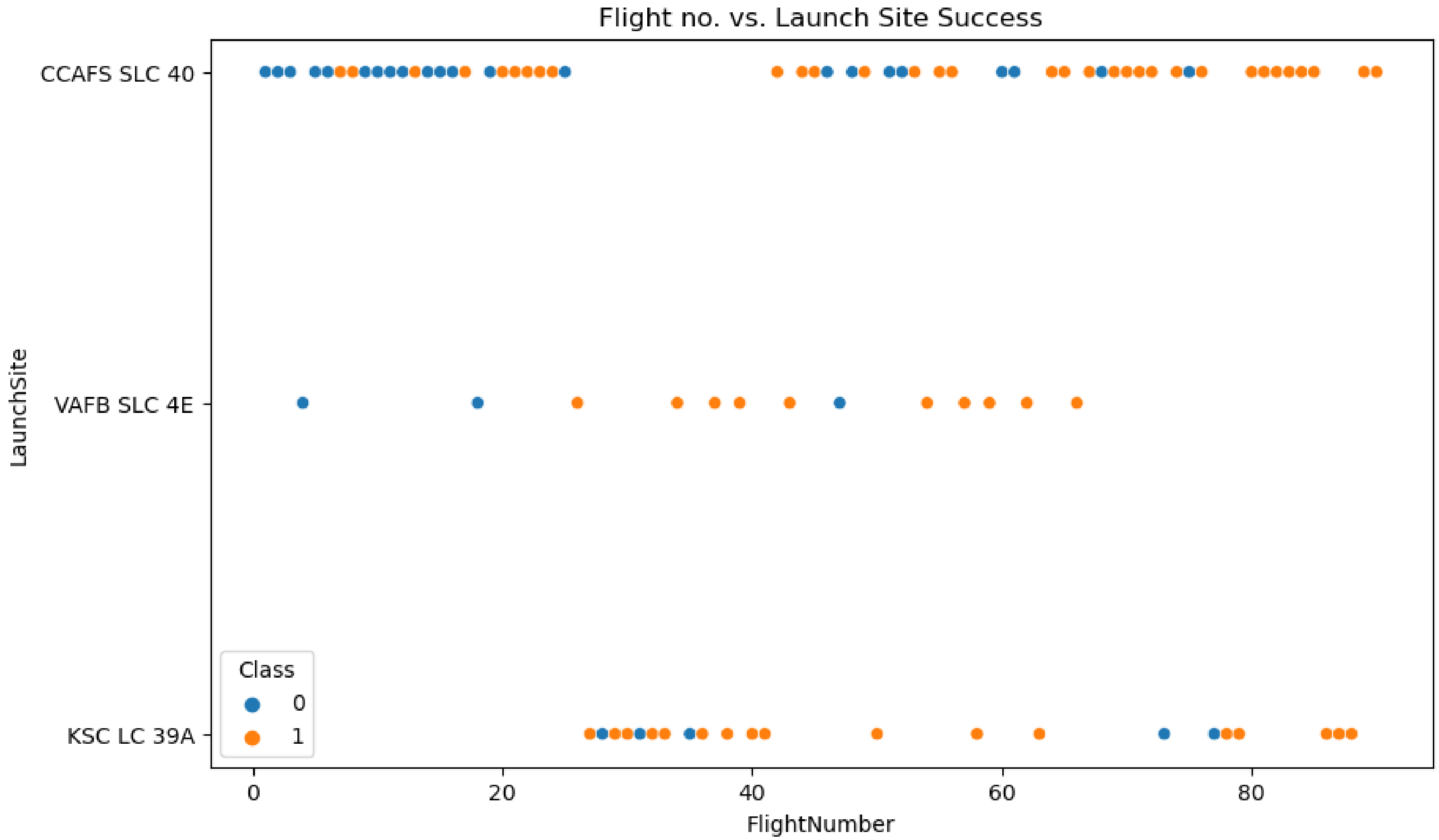
RESULTS

Our analysis confirms KSC LC 39A as the most successful standalone launch site, showing a consistent upward trend in launch success rates. Utilizing machine learning algorithms, including K-nearest Neighbors, logistic regression, and Support Vector Machines, we achieved an impressive prediction accuracy of approximately 83% for launch success.

Clients opting for launches at KSC LC 39A can expect substantial cost savings, often at half the market price, by leveraging SpaceX reusable rockets and our predictive analytics-driven insights. Our machine learning-based recommendations provide tailored insights on optimal launch locations and timing, enhancing success rates and cost-efficiency for clients.

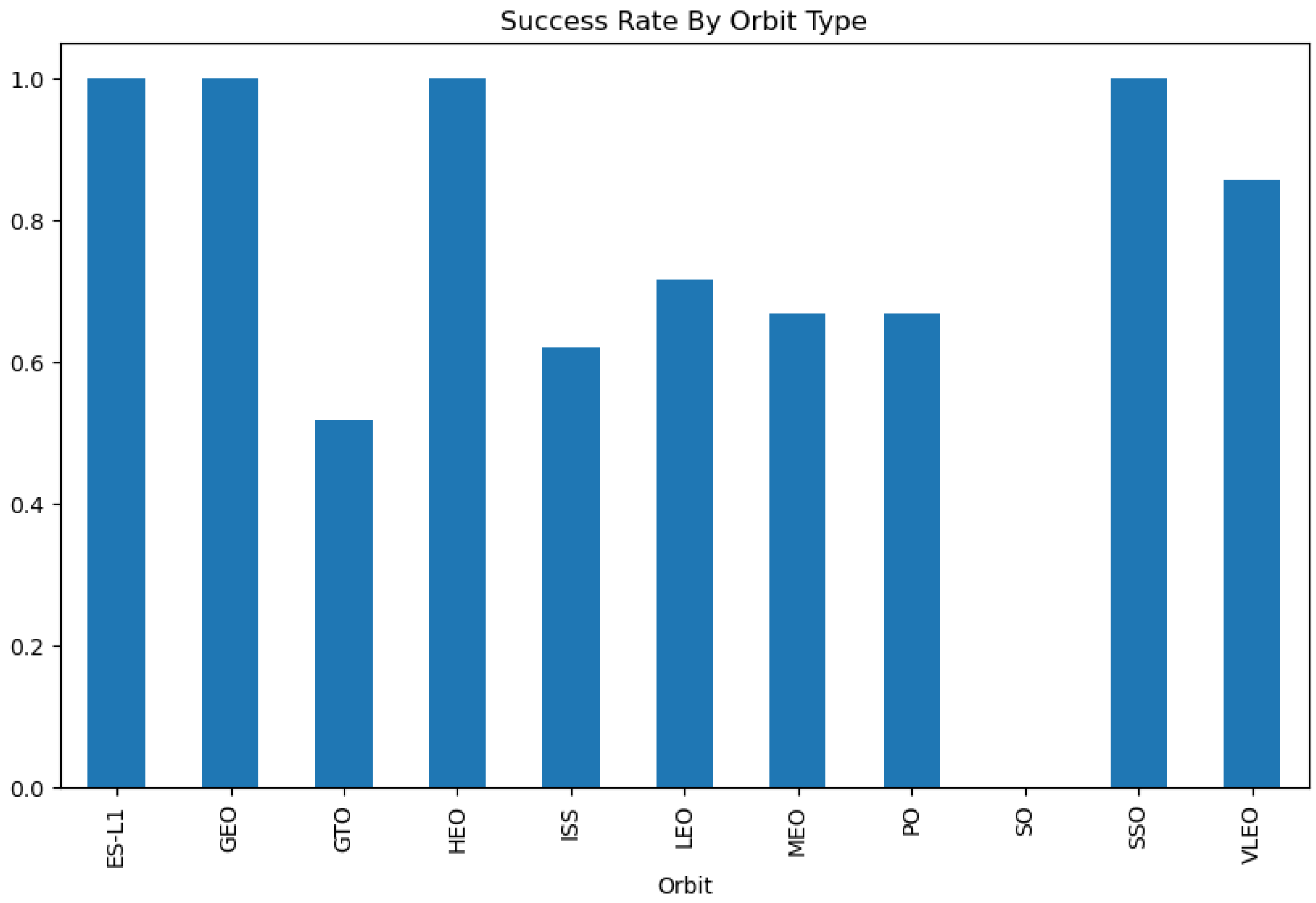
In summary, our results highlight the strategic advantages of adopting data-driven approaches to payload launch planning, empowering clients with actionable insights to optimize their launch strategies and minimize risks.

Flight no. Vs. Launch Sites

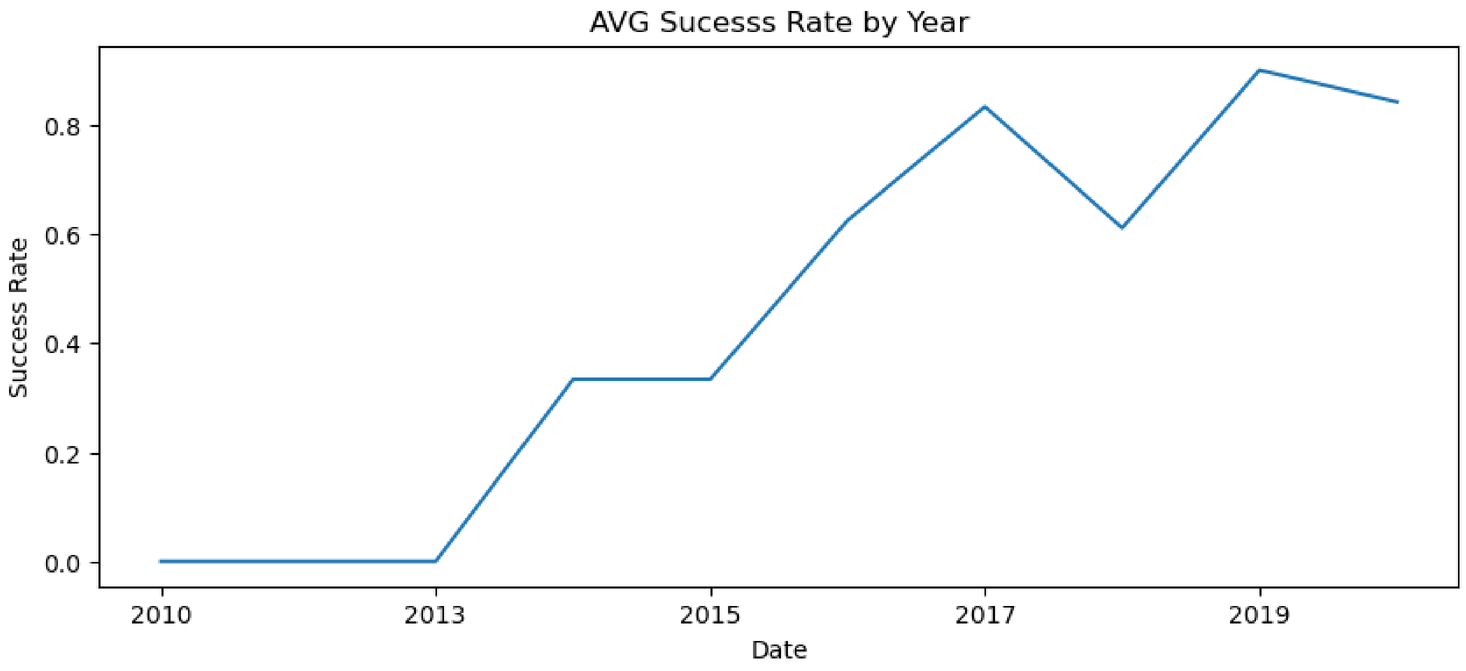


As the flight numbers increase, the success rate of the launches is also increasing.

Success Rate by Orbits



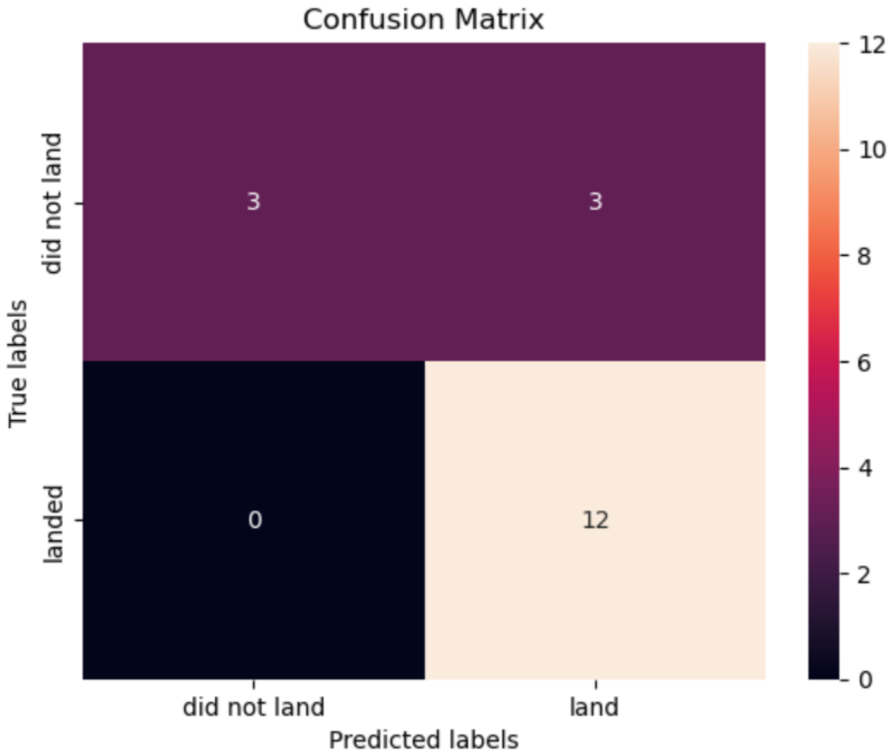
Launch Success Year Over Year



Year over year the success rate of launches independent of site, payload, or booster version is increasing and trending towards more success.

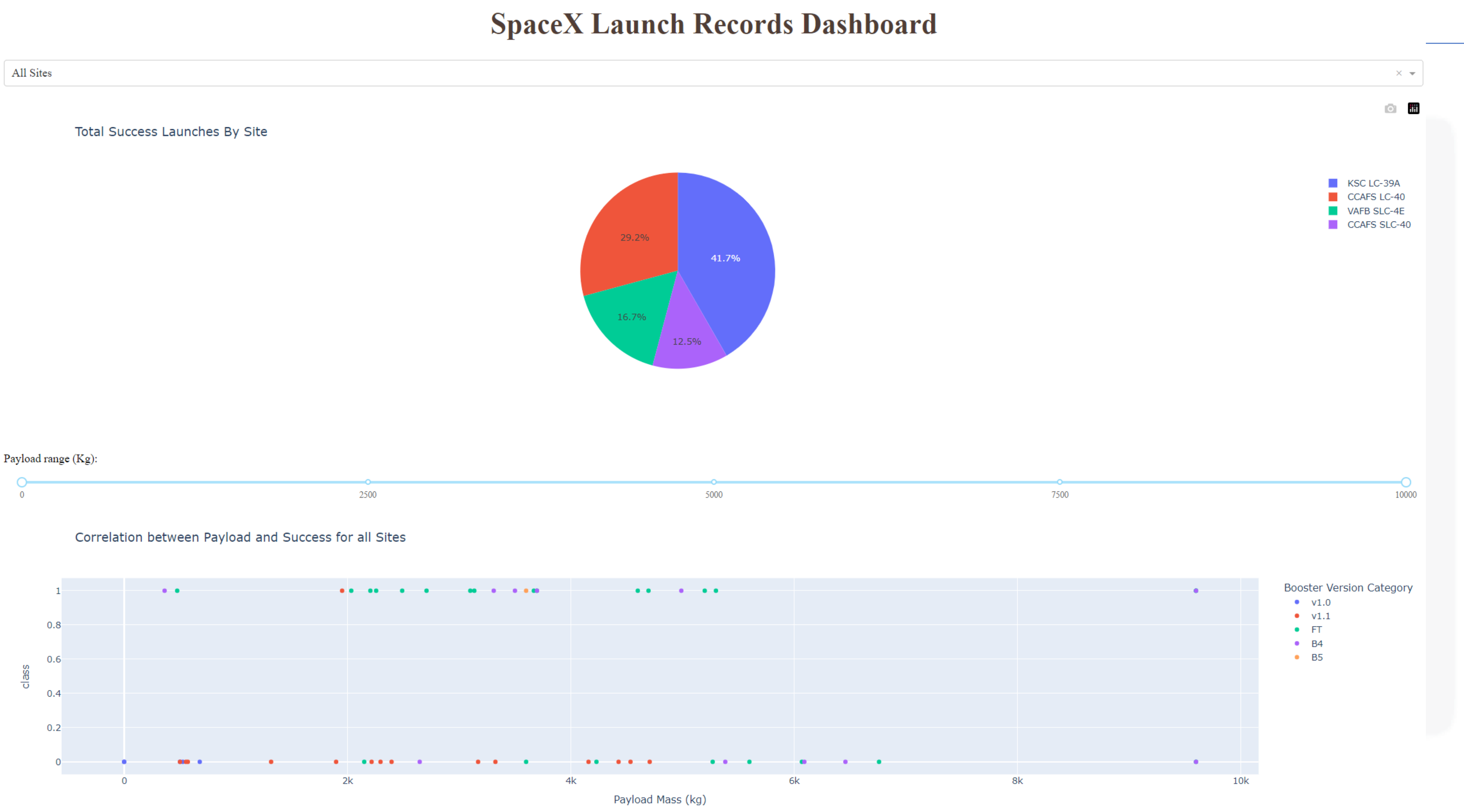
Accuray of ML models

```
Accuracy on test data: 0.8333333333333334
```



The confusion matrix demonstrates that our predictions is often correct when predicting the overall launch outcome. The Accuracy of the model is ~83%

DASHBOARD Success of all Sites

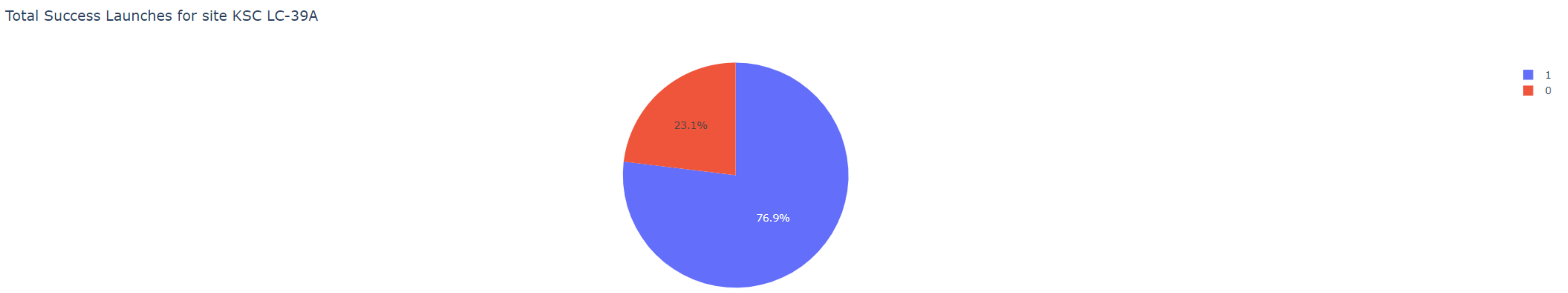


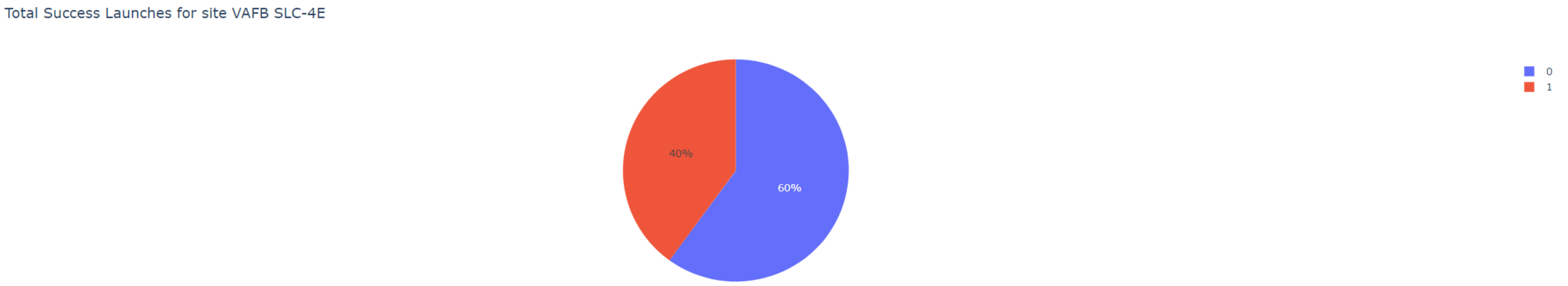
DASHBOARD Success By Site Part 1

Total Success Launches for site CCAFS LC-40

Total Success Launches for site CCAFS SLC-40

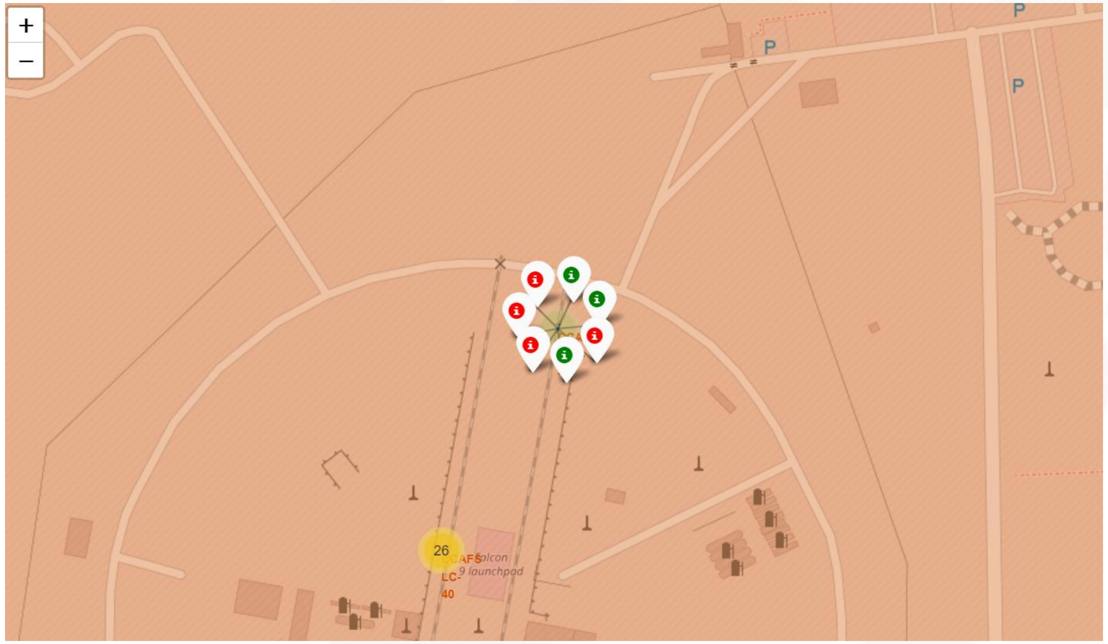
DASHBOARD Success By Site Part 2



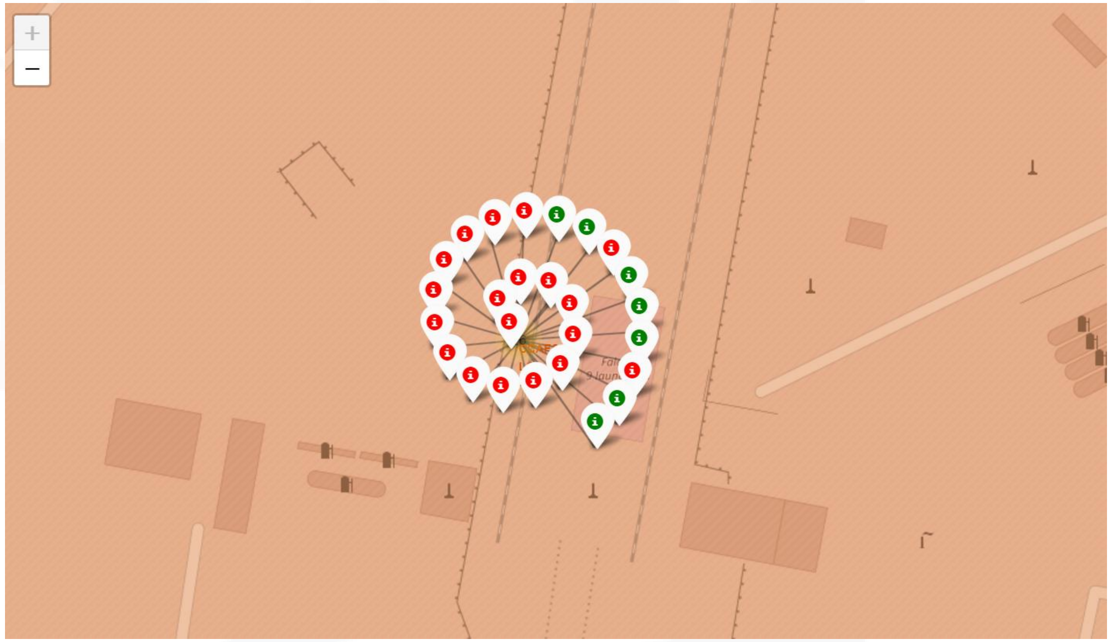


Success Maps by Site Part 1

CCAFS LC-40



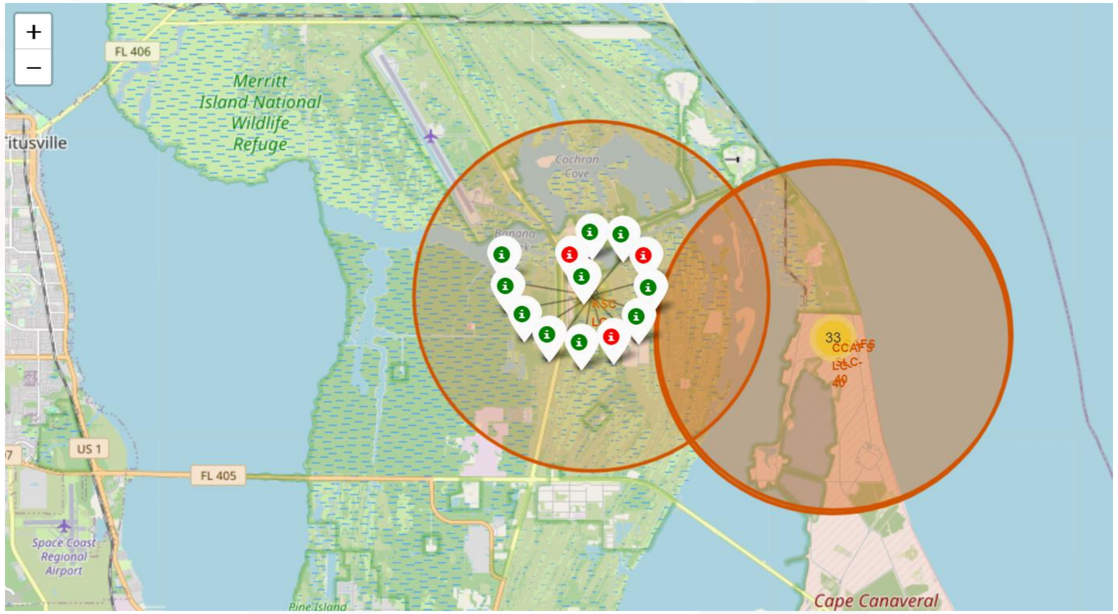
CCFAS SLC-40



Success Maps by Site Part 2

VAFB SLC-4E

KSC LC-39A



OVERALL FINDINGS & IMPLICATIONS

Findings:

- KSC LC 39A
    - Most successful, standalone launch site.
- Overall, throughout the years launch success is ever increasing.
- Prediction Rate
    - We can predict with ~83% accuracy the success of the launch.

Implications

- It would be best for clients to opt-in for launches to take place in KSC LC-39A, in order to have the highest potential of success and saving money.
- As clients continue to launch payloads using SpaceX reusable rockets there odds to save money and successfully launch payloads is likely to increase. Clients working with us will most of the time be able to launch payloads at half the market price.
- As of today's date our clients will be able to receive machine learning based insights into where and when will be best to launch their respective payload at the desired discounted rate.

CONCLUSION

- This study highlights the transformative potential of data-driven decision-making in payload launch planning. Through meticulous analysis and application of advanced machine learning techniques, valuable insights have been provided to optimize launch strategies. The findings confirm the effectiveness of KSC LC 39A as a premier launch site and emphasize the substantial cost savings achievable through strategic use of SpaceX reusable rockets.
- Moreover, the high accuracy in predicting launch success empowers individual clients to make informed decisions, mitigating risks and maximizing success rates. Continued refinement and expansion of analytical capabilities promise further advancements in optimizing payload launch operations.
- In summary, this work demonstrates the tangible benefits of leveraging cutting-edge technology and rigorous data analysis for individual stakeholders in the aerospace industry. By embracing data-driven methodologies, a future is envisioned where space accessibility and affordability are within reach for all.

APPENDIX 1- SQL EDA

]:
```
%%sql
SELECT

    COUNT(*) as cnt,
    "Landing_Outcome"

FROM
    spacextable

WHERE
    "Date" BETWEEN '2010-06-04' AND '2017-03-20'

GROUP BY
    "Landing_Outcome"

ORDER BY
    cnt DESC
```

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

:[82]:

cnt	Landing_Outcome
10	No attempt
5	Success (drone ship)
5	Failure (drone ship)
3	Success (ground pad)
3	Controlled (ocean)
2	Uncontrolled (ocean)
2	Failure (parachute)
1	Precluded (drone ship)

List the total number of successful and failure mission outcomes

n [50]:
```
%%sql
select
    SUM(Case when "Mission_Outcome" LIKE "S%" then 1 else 0 end) as tot_success,
    SUM(Case when "Mission_Outcome" LIKE "F%" then 1 else 0 end) as tot_failure
from
    spacextable
```

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

Out[50]:

tot_success	tot_failure
100	1

I did not find these queries useful enough to be effective as a part of my main presentation.

APPENDIX 2- Github Link IBM Capstone

https://github.com/Diego-A-B/IBM-CAPSTONE