SpaceX Launch Analysis

Presented by
YouJia Chen

Agenda

- Executive Summary
- Introduction
- Exploratory Data Analysis
- Statistical Analysis
- Results
- Conclusion

Executive Summary

SpaceX revolutionizes the space industry by offering Falcon 9 rocket launches at a cost of $62 million, significantly lower than competitors' prices.

This remarkable cost reduction is achieved through the innovative practice of reusing the first stage.

In this project, our goal is to predict the landing success of the first stage, enabling accurate cost estimation for rocket launches. By analyzing historical data and applying machine learning algorithms, we aim to develop a predictive model that can be used for competitive bidding against SpaceX.

Additionally, we consider the geographical significance of launch sites, which are strategically located near coastlines and the equator to optimize trajectories and facilitate access to various orbital inclinations.

This project holds the potential to provide valuable insights for the industry, empowering decision-makers with precise cost estimates and enhancing bidding strategies.

Introduction

- SpaceX offers Falcon 9 rocket launches at an unbeatable price of 62 million dollars.
- Competitors charge over 165 million dollars for similar services.
- The cost reduction is made possible by SpaceX's groundbreaking approach of reusing the first stage of their rockets.
- This innovation has transformed the economics of space travel.

Data Collection And Data Wrangling Methodology

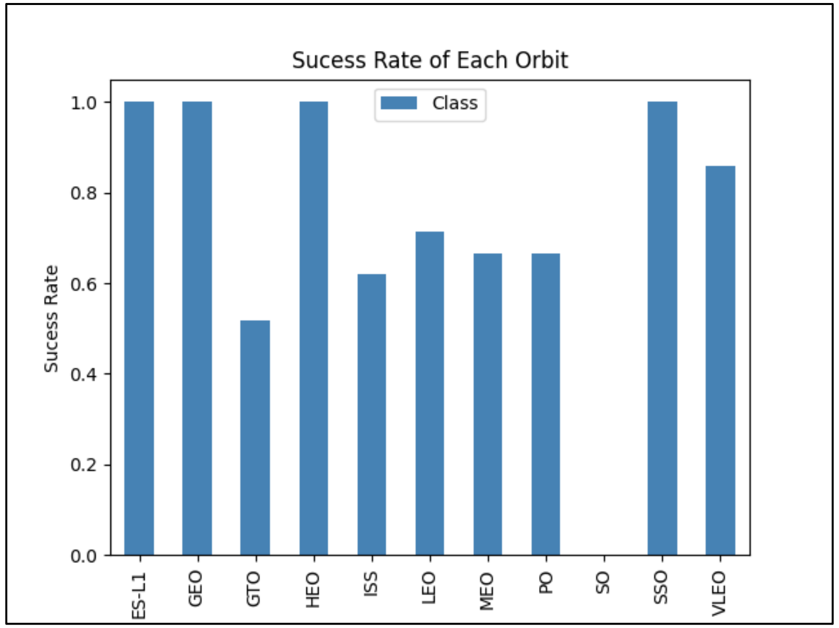
- API
- Web Scrabbing from Wikipedia

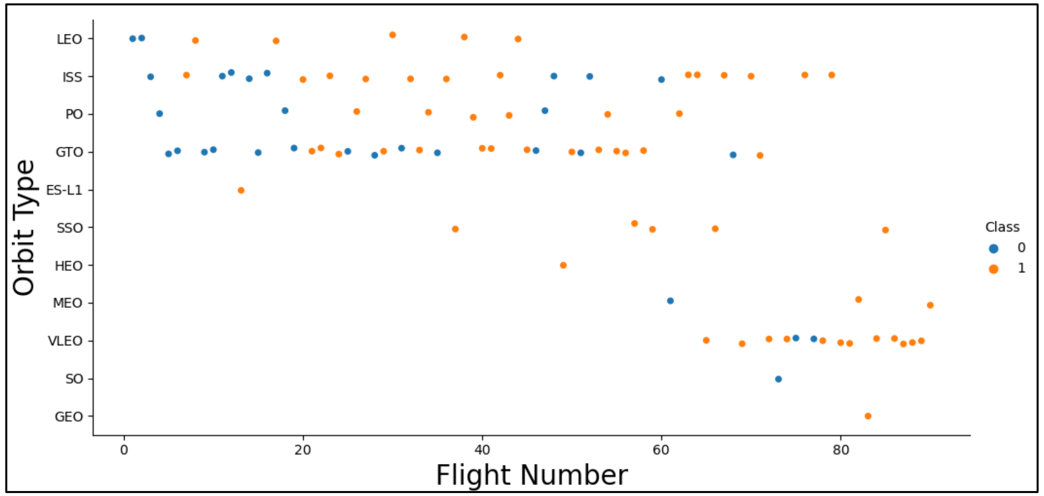
- Create new column ‘Class’, 1 for success, 0 for fail

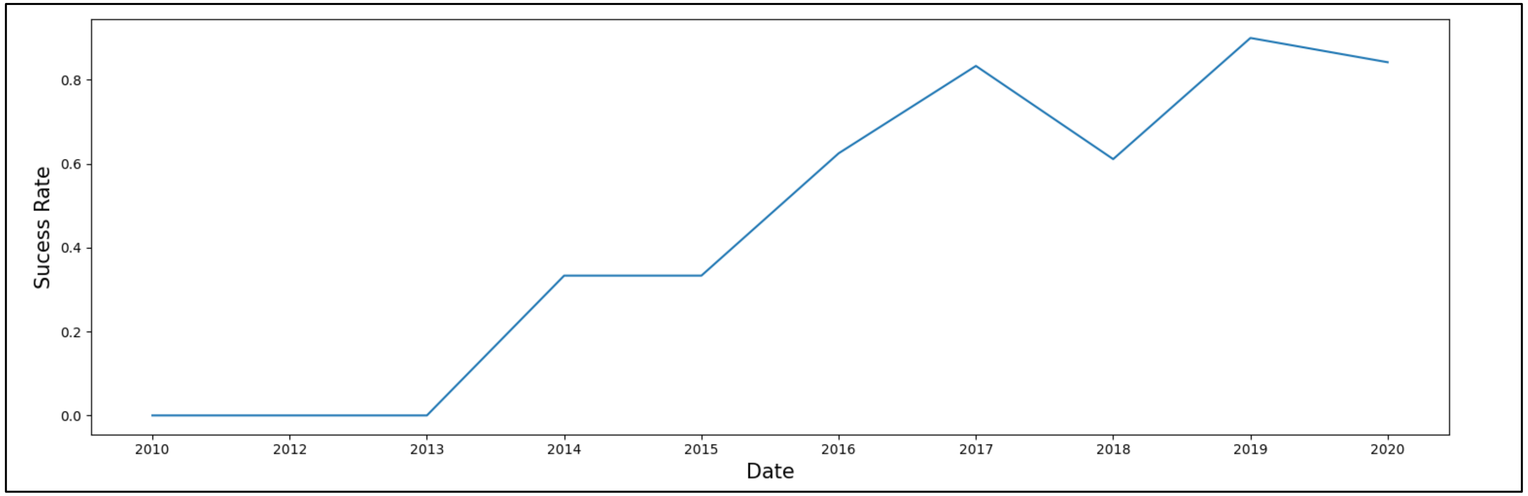
```
df.dtypes

FlightNumber        int64
Date               object
BoosterVersion     object
PayloadMass       float64
Orbit              object
LaunchSite         object
Outcome            object
Flights             int64
GridFins             bool
Reused               bool
Legs                 bool
LandingPad         object
Block             float64
ReusedCount         int64
Serial             object
Longitude         float64
Latitude          float64
dtype: object
```

EDA with visualization results

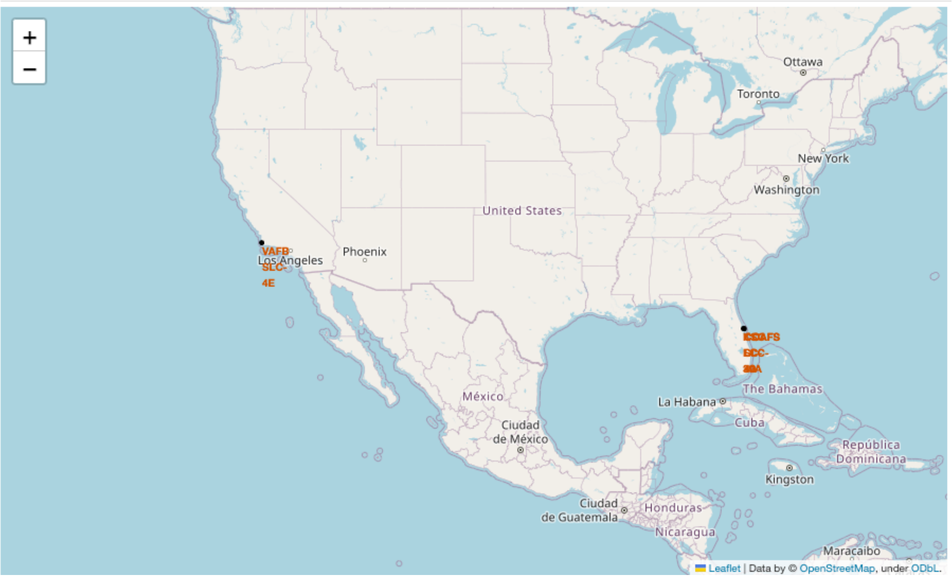


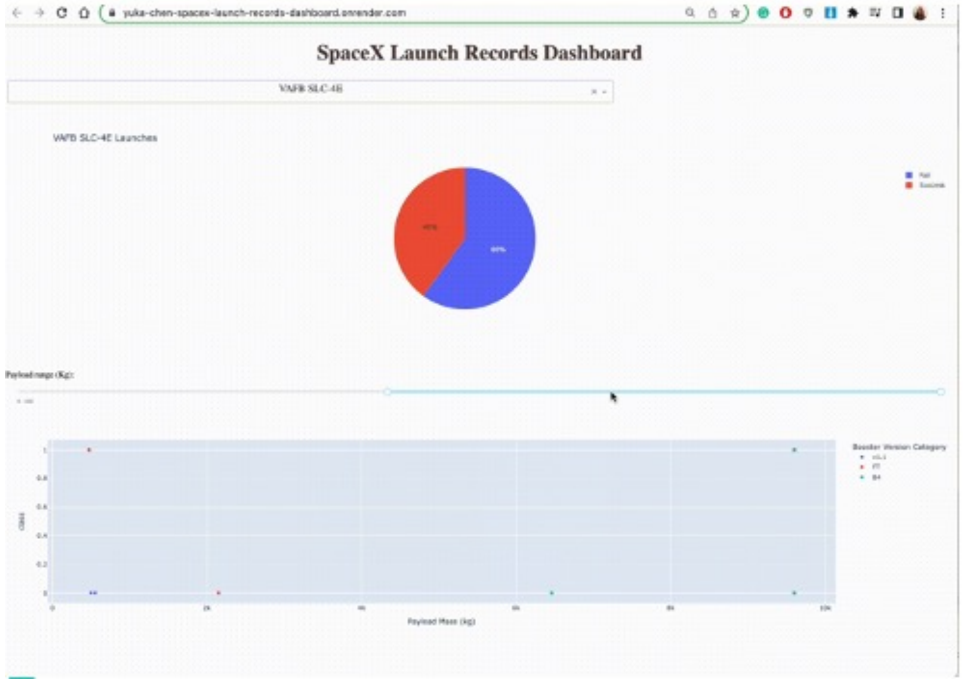




EDA and interactive visual analytics methodology

- Folium for interactive map analysis
- Plotly for interactive visual analysis (pie and scatter plot)





Predictive Analysis Methodology

- Create a column for the class
- Standardize the data
- Split into training data and test data
- Find best hyperparameter for **SVM**, **Classification trees** and **Logistic Regression**

EDA with SQL(ite)

```
%load_ext sql
```

```
import csv, sqlite3

con = sqlite3.connect("my_data1.db")
cur = con.cursor()
```

```
!pip install -q pandas==1.1.5
```

```
%sql sqlite:///my_data1.db
```

```
'Connected: @my_data1.db'
```

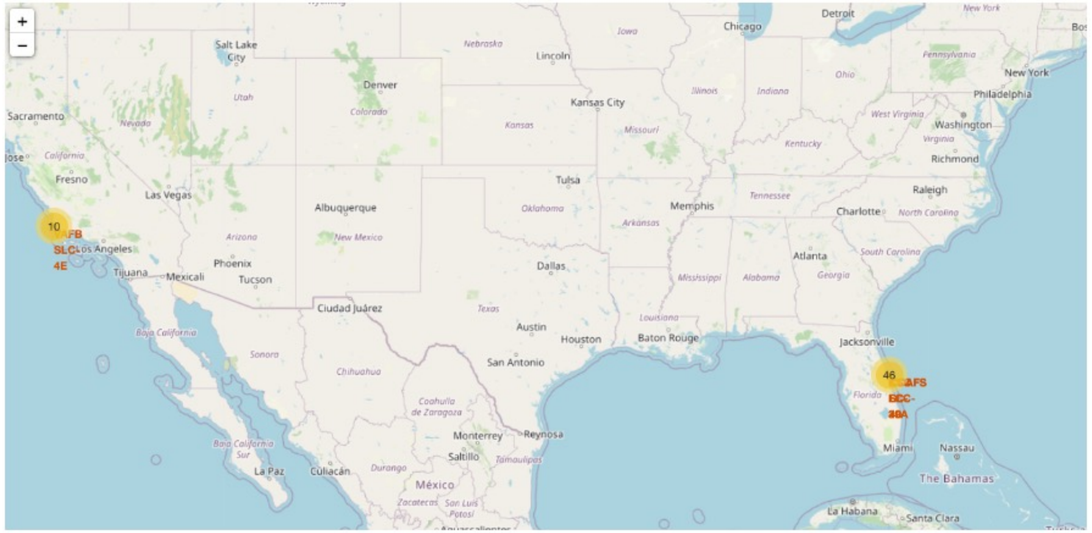
Mission_Outcome	total_count
None	898
Failure (in flight)	1
Success	98
Success	1
Success (payload status unclear)	1

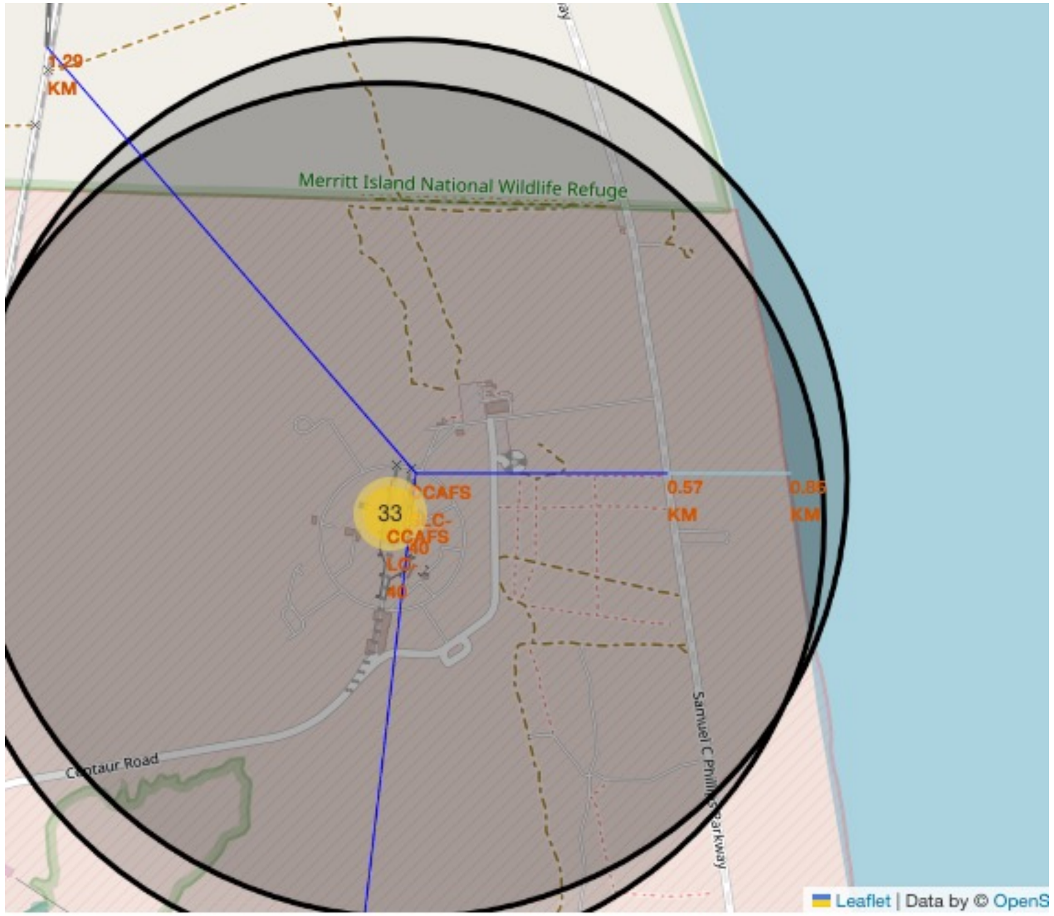
Launch_Site
CCAFS LC-40
VAFB SLC-4E
KSC LC-39A
CCAFS SLC-40
None

Booster_Version
F9 FT B1022
F9 FT B1026
F9 FT B1021.2
F9 FT B1031.2

Landing_Outcome	Count
Success	30
No attempt	15
Success (drone ship)	10
Success (ground pad)	6
Failure (drone ship)	3
Failure	3
Controlled (ocean)	3
Uncontrolled (ocean)	2
Failure (parachute)	2
Precluded (drone ship)	1
No attempt	1

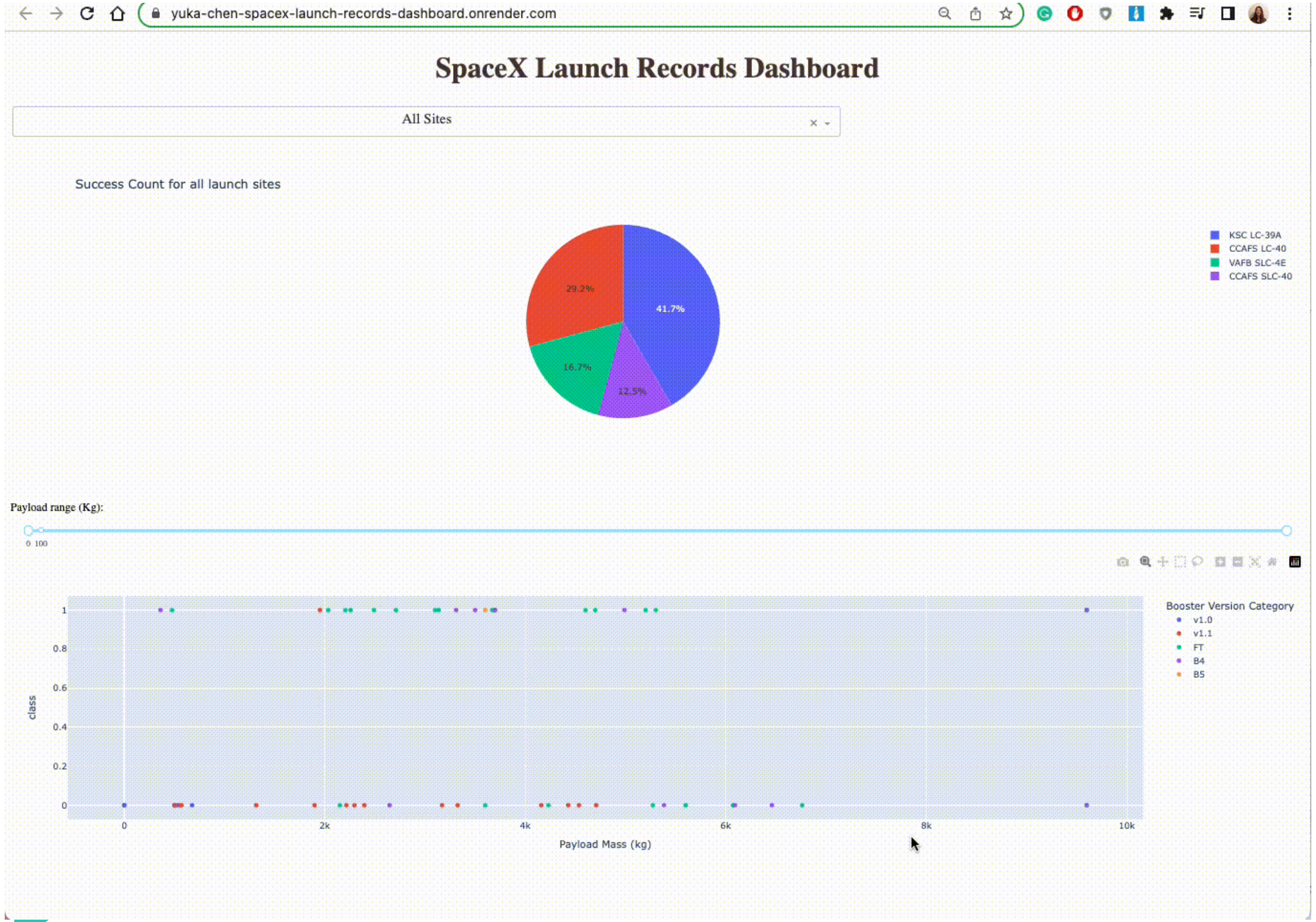
Interactive map with Folium





location	distances
highway	0.57
railway	1.29
coastline	0.85
cities	51.2

Plotly Dash dashboard



https://yuka-chen-spacex-launch-records-dashboard.onrender.com/

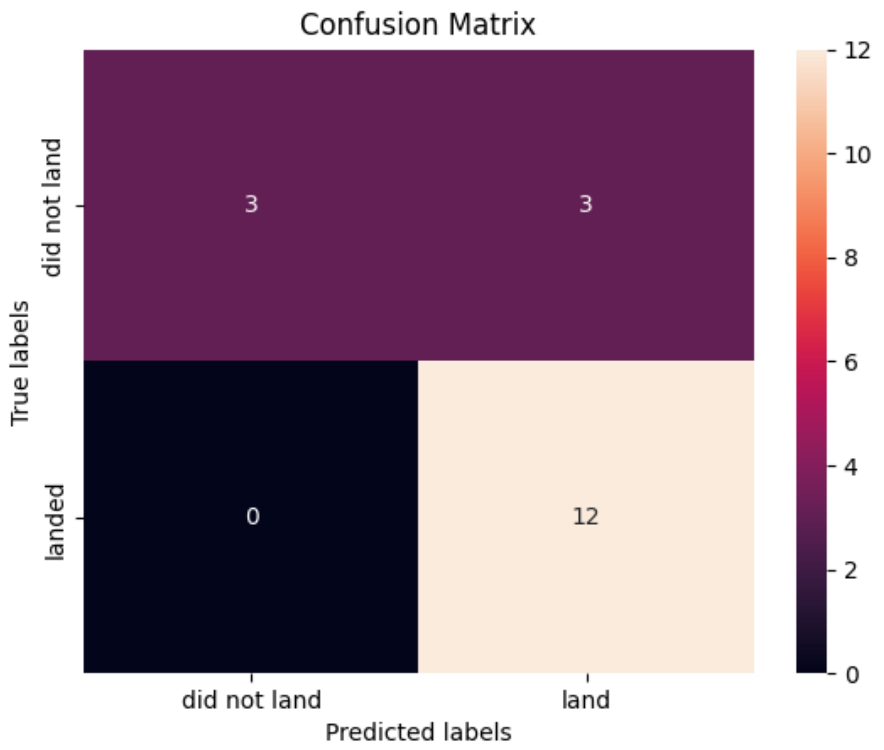
Predictive Analysis (Classification)

- Three methods all have 83.33% accuracy

```
logreg_accuracy = logreg_cv.score(X_test, Y_test)
svm_accuracy = svm_cv.score(X_test, Y_test)
tree_accuracy = tree_cv.score(X_test, Y_test)

print("Logistic Regression Accuracy:", logreg_accuracy)
print("Support Vector Machine Accuracy:", svm_accuracy)
print("Decision Tree Accuracy:", tree_accuracy)
```

```
Logistic Regression Accuracy: 0.8333333333333334
Support Vector Machine Accuracy: 0.8333333333333334
Decision Tree Accuracy: 0.8333333333333334
```



Conclusion

Launch Sites Success Rate

- CCAFS LC-40: 26.9% (Florida)
- CCAFS SLC-40 : 42.9% (Florida)
- KSC LC-39A : 23.1% (Florida)
- VAFB SLC-4E : 40% (California)

- Proximity to coastlines:
  - Allows for safer abort scenarios and emergency contingencies during launch.
  - Simplifies transportation logistics for rocket components.
  - Facilitates integration and assembly processes.
- Proximity to the equator:
  - Provides an inherent velocity boost from the Earth's rotation.
  - Reduces the amount of fuel required for achieving orbit or reaching destinations.
  - Increases payload capacity or extends the operational lifespan of satellites.
  - Enables access to a wider range of orbital inclinations.
- Strategic advantages:
  - Maximizes operational efficiency and safety.
  - Enhances payload capabilities.
  - Offers flexibility and versatility for space missions.

Appendix