PREDICT SPACEX LANDING SUCCESS RATES AND ANALYZE RELATED DATA

DATA COLLECTION, WRANGLING, EDA, PREDICTIVE MODELING, AND INTERACTIVE VISUALIZATIONS.

SpaceX

Launch

Insights

INSIGHTS INTO LAUNCH SUCCESS FACTORS, PREDICTIVE MODEL ACCURACY, AND INTERACTIVE TOOLS FOR ANALYSIS.

EXECUTIVE SUMMARY

Project Objective:
The project aims to analyze SpaceX launch data to identify factors influencing successful landings, build predictive models to forecast outcomes, and create interactive tools for data exploration and insight generation.

Key Achievements:

Conducted comprehensive data collection through APIs, web scraping, and data wrangling techniques.
Developed interactive dashboards and visualizations to uncover trends in launch success rates.
Built and optimized machine learning models, achieving accurate predictions of landing success.
Outcomes:

Identified payload mass and launch site as key factors for success.
Created a Python Dash application for real-time analysis and visualization.
Delivered actionable insights for improving SpaceX's mission outcomes.

INTRODUCTION

Background:
SpaceX revolutionized the space industry by developing reusable rockets, significantly reducing the cost of space travel. This project focuses on understanding the factors contributing to SpaceX's success in landing rocket boosters.

Problem Statement:
With hundreds of launches conducted, SpaceX aims to enhance mission efficiency and success rates. Identifying trends in past launches and predicting outcomes for future missions is critical.

Goals:

Explore SpaceX data to uncover patterns and insights using advanced visualizations.
Build predictive models to forecast landing success based on key factors like payload mass and launch site.
Develop an interactive dashboard for stakeholders to analyze data dynamically.
Scope:
The project encompasses data collection, cleaning, exploratory data analysis (EDA), machine learning, and dashboard development to create a comprehensive solution for decision-making.

DATA COLLECTION AND WRANGLING

Data Sources: APIs, web scraping (Beautiful Soup), and provided datasets.
Wrangling Steps:
Cleaning inconsistencies.
Formatting payload and success metrics.
Transforming and merging data for analysis.

6	1
7	1

In [25]:
```
df.head(5)
complete_failures = df['Outcome'].str.contains('None', na=False).sum()
complete_failures
```
Out[25]:
```
np.int64(21)
```

We can use the following line of code to determine the success rate:

In [17]:
```
df["Class"].mean()
```
Out[17]:
```
np.float64(0.6666666666666666)
```

We can now export it to a CSV for the next section,but to make the answers consistent, in the next lab we will provide data in a pre-selected date range

```
versus-a-copy
  """
```

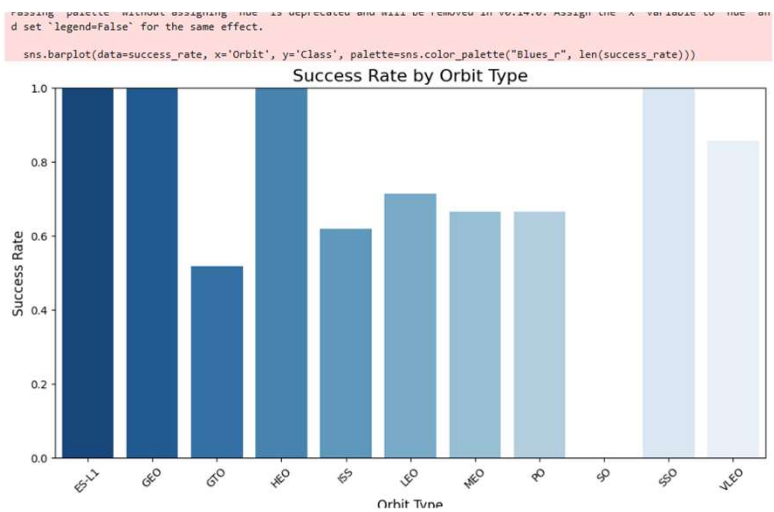
t[40]:

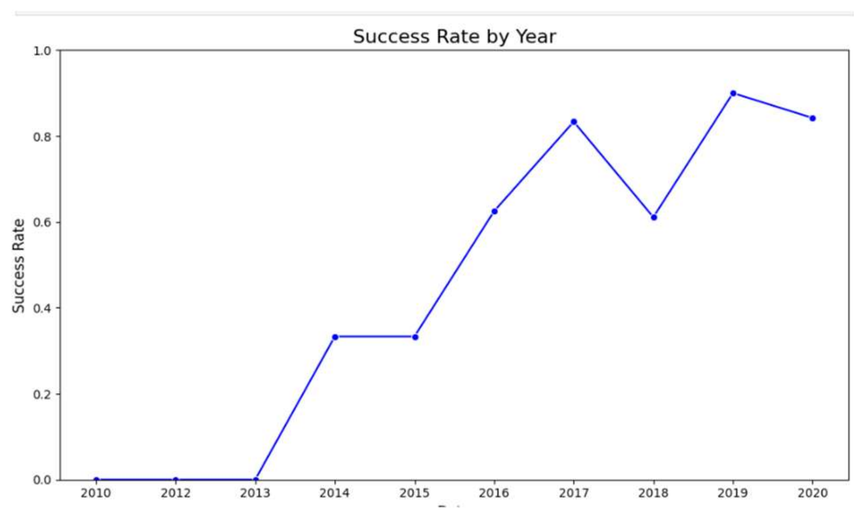
	FlightNumber	Date	BoosterVersion	PayloadMass	Orbit	LaunchSite	Outcome	Flights	GridFins	Reused	Legs	La
4	1	2010-06-04	Falcon 9	6123.547647	LEO	CCSFS SLC 40	None None	1	False	False	False	
5	2	2012-05-22	Falcon 9	525.000000	LEO	CCSFS SLC 40	None None	1	False	False	False	
6	3	2013-03-01	Falcon 9	677.000000	ISS	CCSFS SLC 40	None None	1	False	False	False	
7	4	2013-09-29	Falcon 9	500.000000	PO	VAFB SLC 4E	False Ocean	1	False	False	False	
8	5	2013-12-03	Falcon 9	3170.000000	GTO	CCSFS SLC 40	None None	1	False	False	False	
...	...	...	...	...	...	...	...	...	...	...	...	
89	86	2020-09-03	Falcon 9	15600.000000	VLEO	KSC LC 39A	True ASDS	2	True	True	True	5e9e3032383ecb6t
90	87	2020-10-06	Falcon 9	15600.000000	VLEO	KSC LC 39A	True ASDS	3	True	True	True	5e9e3032383ecb6t
91	88	2020-10-18	Falcon 9	15600.000000	VLEO	KSC LC 39A	True ASDS	6	True	True	True	5e9e3032383ecb6t
92	89	2020-10-24	Falcon 9	15600.000000	VLEO	CCSFS SLC 40	True ASDS	3	True	True	True	5e9e3033383ecbb!
93	90	2020-11-05	Falcon 9	3681.000000	MEO	CCSFS SLC 40	True ASDS	1	True	False	True	5e9e3032383ecb6t

EXPLORATORY DATA ANALYSIS (EDA)

Perform exploratory Data Analysis and Feature Engineering using Pandas and Matplotlib

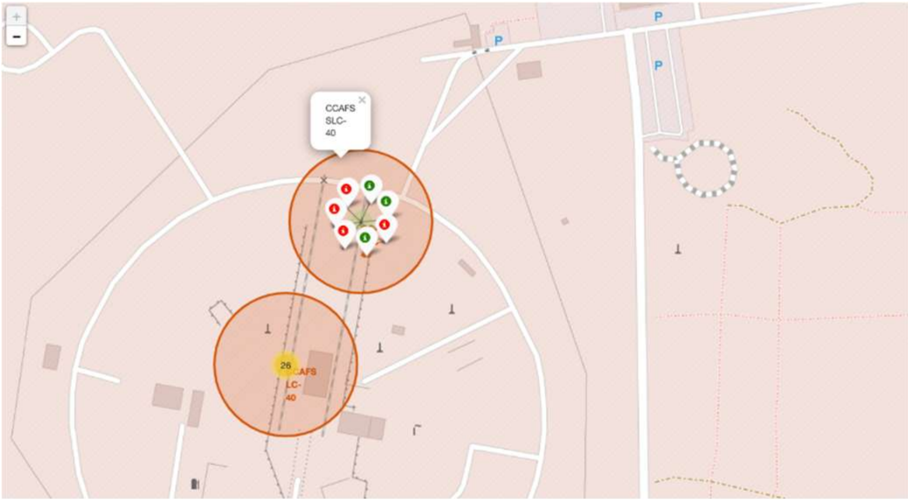
Exploratory Data Analysis
Preparing Data Feature Engineering

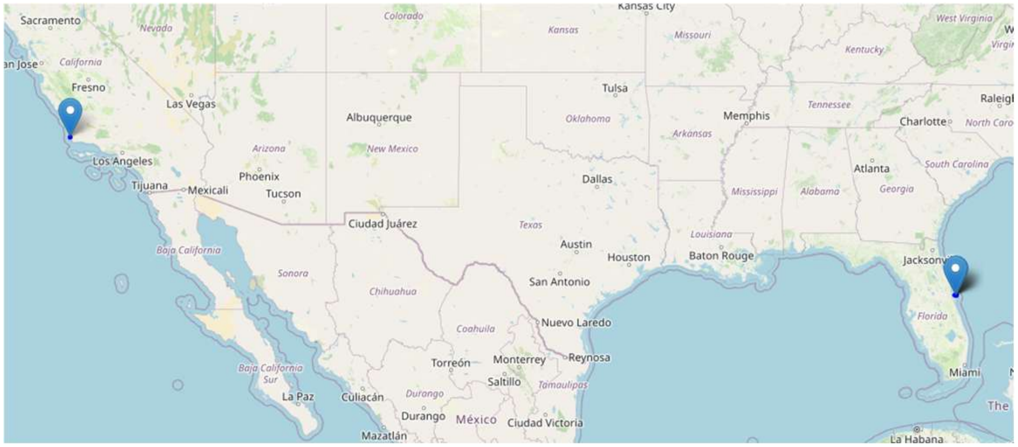




USING MAPS IN ANALYSIS

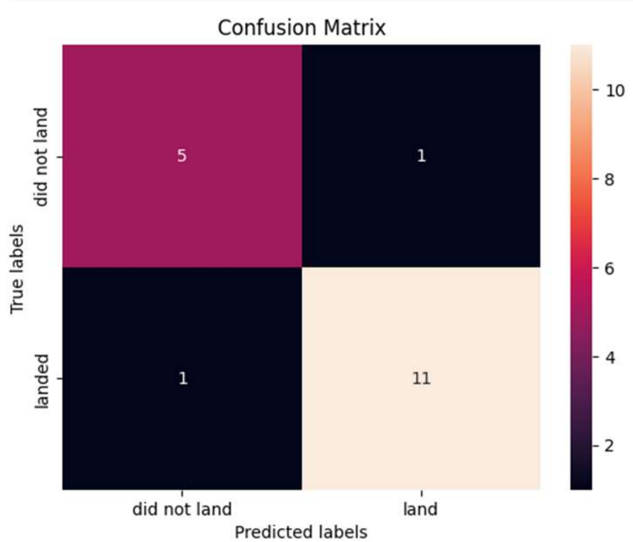
The launch success rate may depend on many factors such as payload mass, orbit type, and so on. It may also depend on the location and proximities of a launch site, i.e., the initial position of rocket trajectories. Finding an optimal location for building a launch site certainly involves many factors and hopefully we could discover some of the factors by analyzing the existing launch site locations.

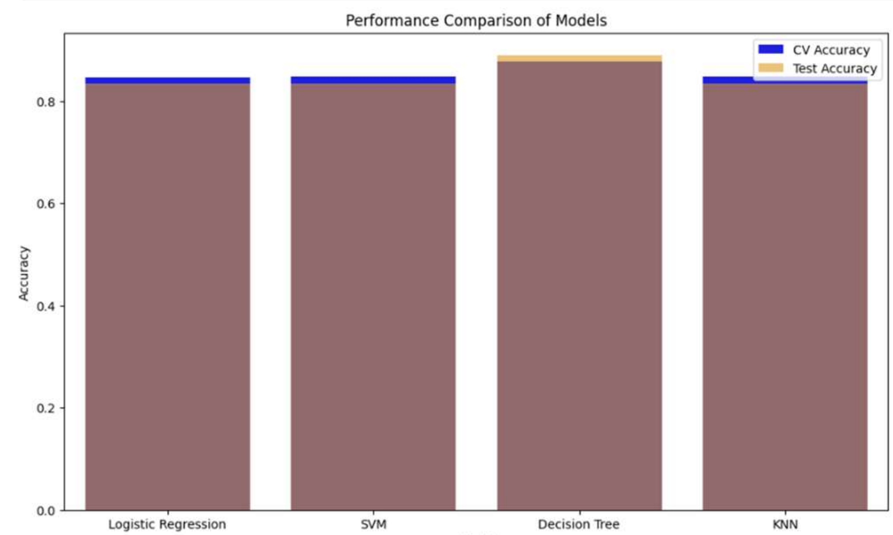




MODEL TRAINING FOR PREDICTION

Space X advertises Falcon 9 rocket launches on its website with a cost of 62 million dollars; other providers cost upward of 165 million dollars each, much of the savings is because Space X can reuse the first stage. Therefore if we can determine if the first stage will land, we can determine the cost of a launch. This information can be used if an alternate company wants to bid against space X for a rocket launch.





CONCLUSION

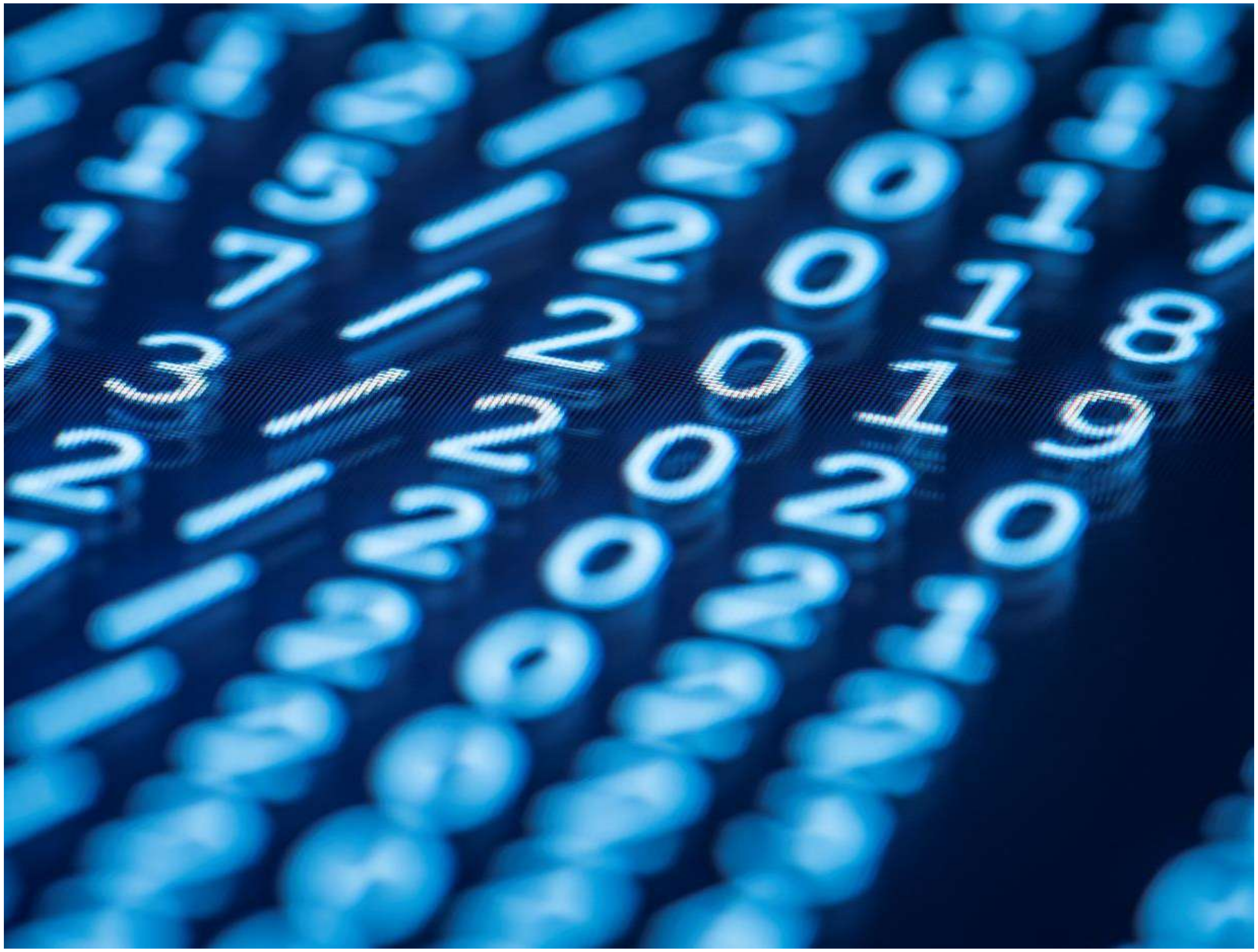
Key Findings:

Launch success rates are significantly influenced by payload mass and launch site.
KSC LC-39A demonstrated the highest success rate among all launch sites.
Optimal payload range for success lies within the mid-range of the payload capacity.
Predictive Modeling:

Machine learning models, especially Decision Trees and Logistic Regression, provided reliable predictions of landing success.
GridSearchCV optimization improved model accuracy, offering practical utility for future mission planning.
Impact:

The insights derived empower SpaceX to enhance mission strategies, focus on high-performing sites, and optimize payload design for improved outcomes.
The developed tools (interactive maps and dashboards) enable real-time analysis and decision-making.
Future Scope:

Incorporating additional features like weather data and mission type for improved predictions.
Expanding the dashboard's capabilities to include live mission tracking and anomaly detection.



THANK YOU