

SpaceX Falcon 9 first stage Landing Prediction:

First Stage Landing Success

© IBM Corporation. All rights reserved.



OUTLINE



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

EXECUTIVE SUMMARY



- Predicting the success of the Falcon 9 first stage landing
- Understanding the factors that contribute to successful landings
- Estimating the overall cost of a launch
- Providing valuable insights for companies seeking to bid against SpaceX
- Developing a predictive model that determines the likelihood of successful first stage landings.



INTRODUCTION



- Falcon 9 cost 62 million dollars whereas other cost 165 million
- Space X reuses Falcon9 first stage
- SpaceX's unique ability to reuse the first stage of the Falcon 9 rocket has significantly reduced launch costs compared to traditional providers.

METHODOLOGY

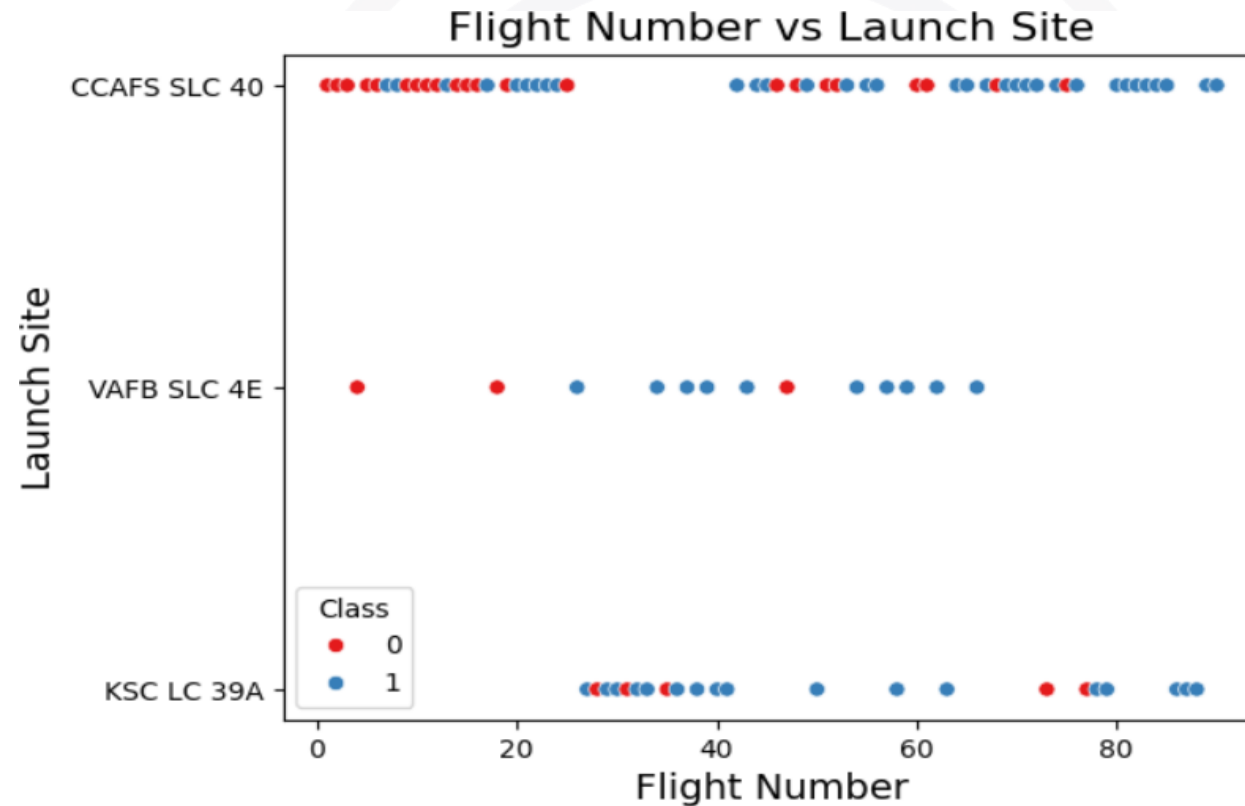


- Data Collection through:
 - SpaceX API
 - Web Scrapping
- Data Wrangling
- Data Analysis & Visualization using:
 - SQL
 - Pandas
 - Folium
 - Matplotlib
- ML Prediction of the Successful First stage Landing



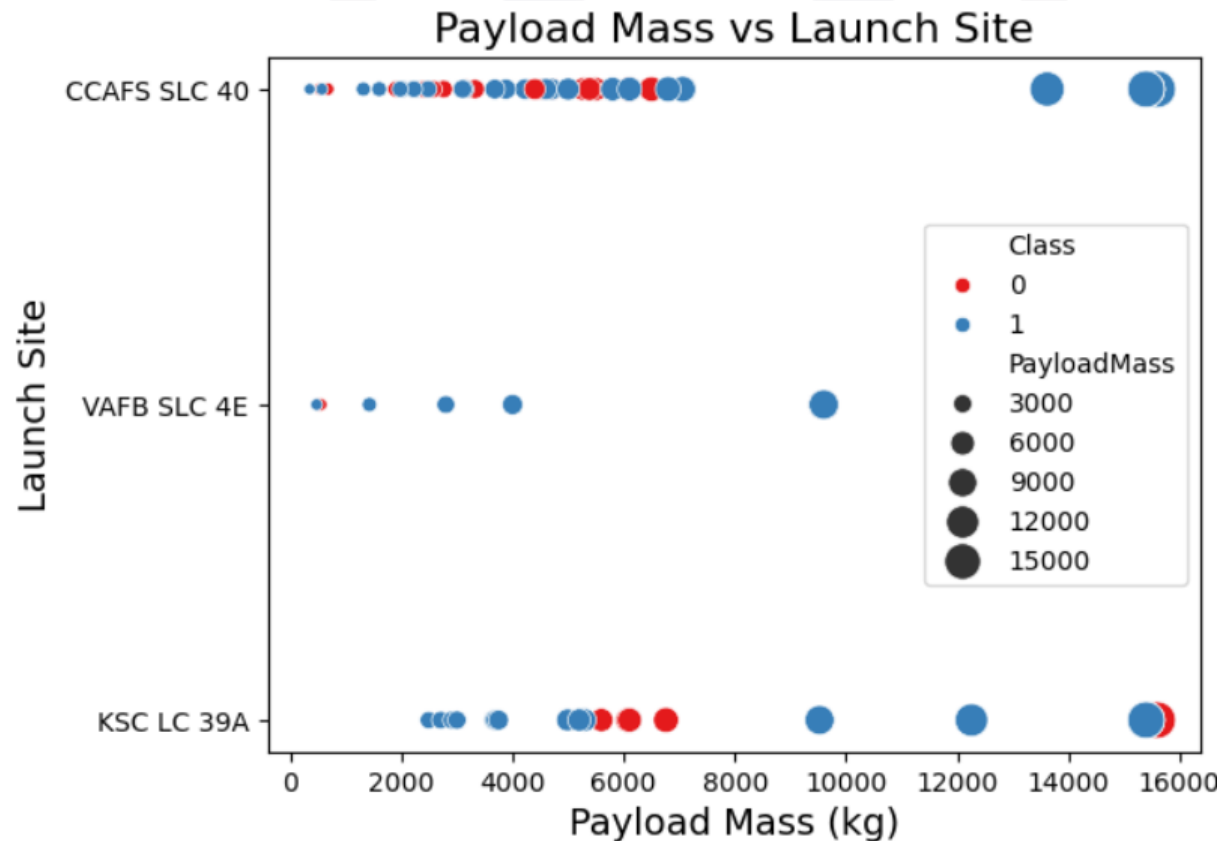
EDA & Interactive Visual Analytics Methodology (1/3)

- Visualize the relationship between Flight Number and Launch Site



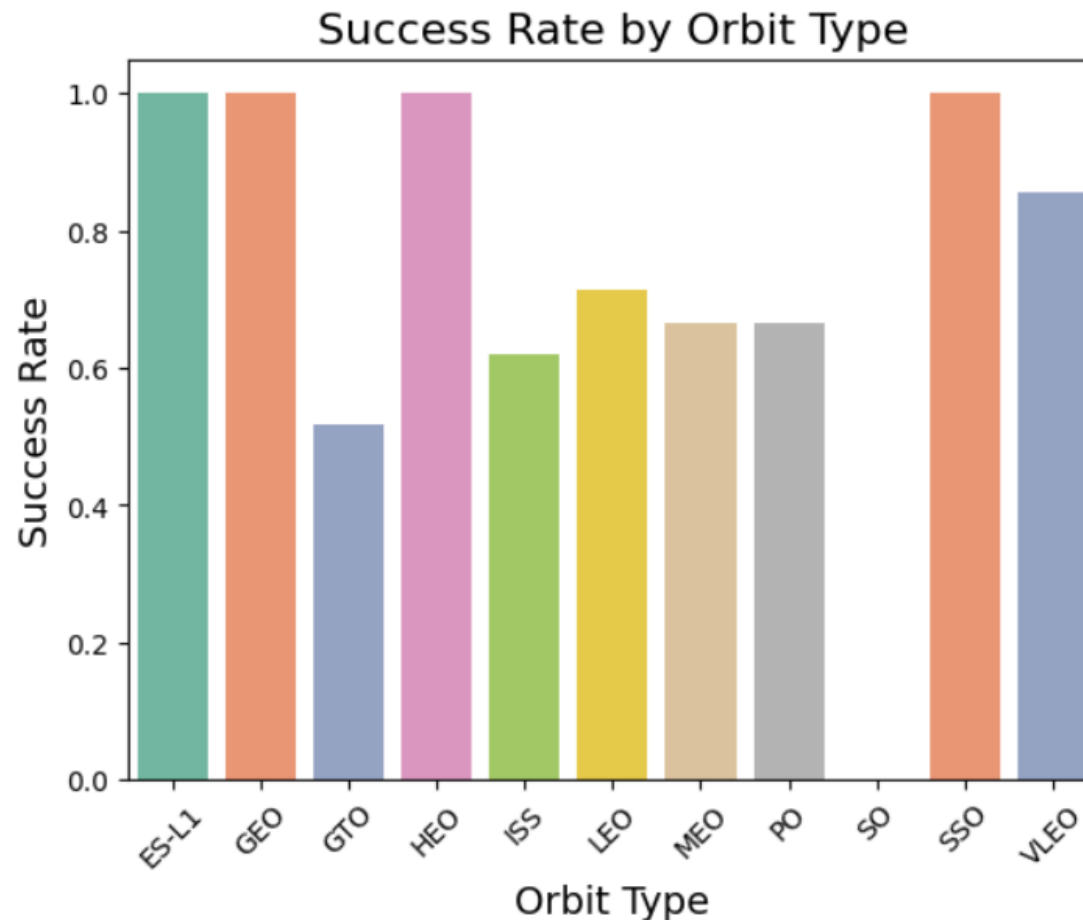
EDA & Interactive Visual Analytics Methodology (2/3)

- Visualize the relationship between Payload Mass and Launch Site



EDA & Interactive Visual Analytics Methodology (3/3)

- Visualize the relationship between success rate of each orbit type



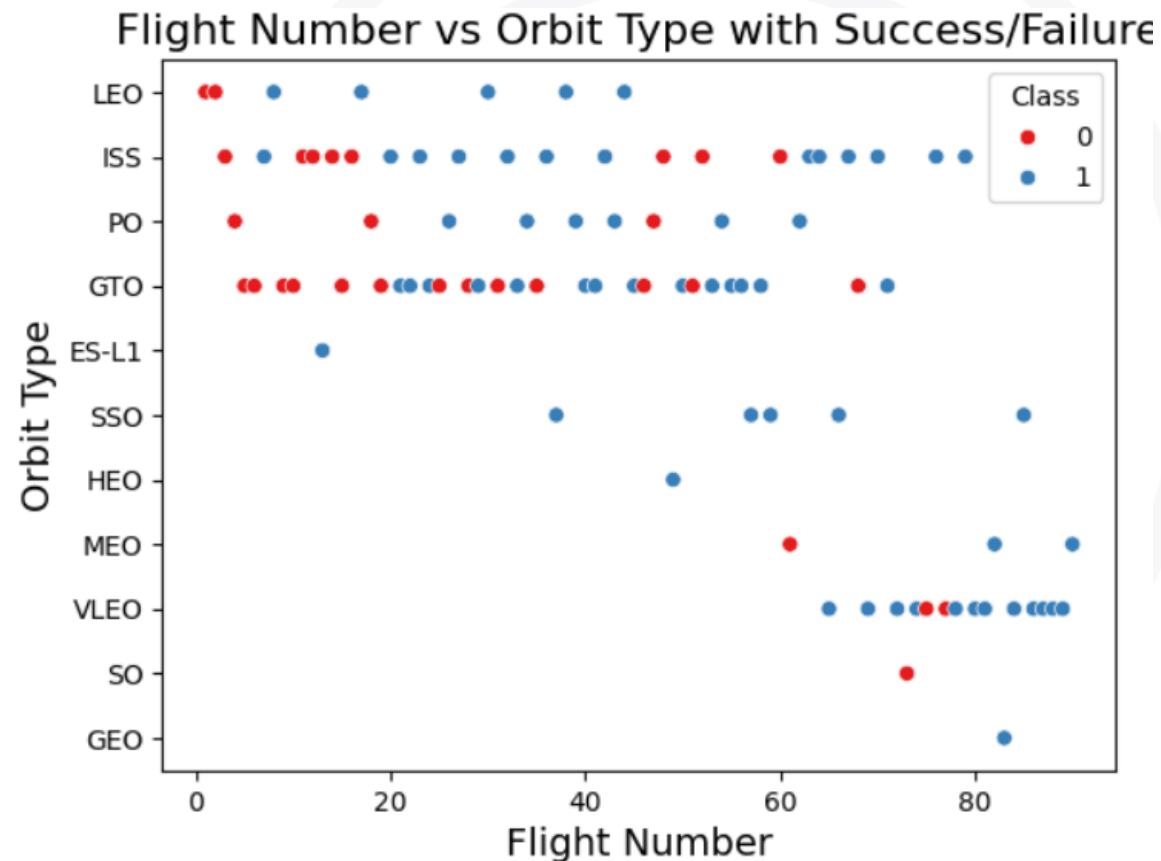
Predictive Analysis Methodology

- Larger the flight amount at a launch site, the greater the success rate at a launch site.
- ES-L1, GEO, HEO, SSO, VLEO had the most success rate.
- LEO orbit, success is related to the number of flights whereas in the GTO orbit, there is no relationship between flight number and the orbit.



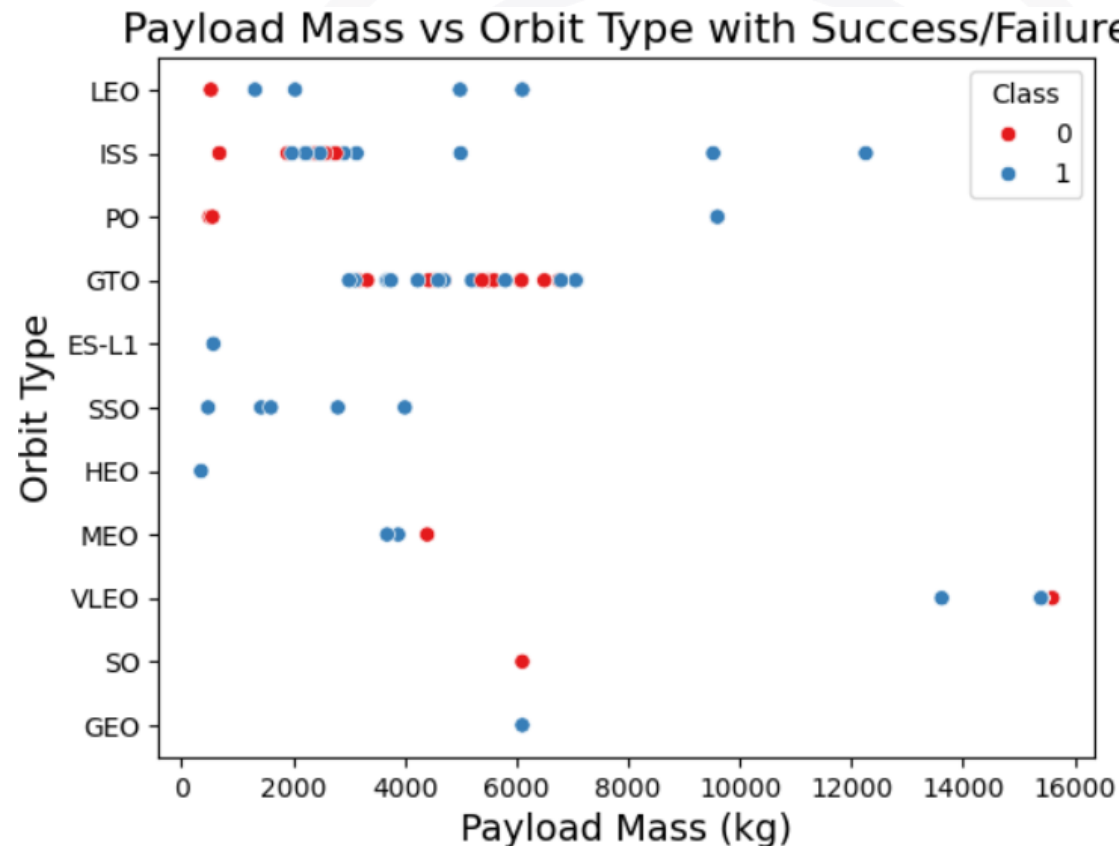
EDA & Interactive Visual Analytics Result (1/6)

- Visualize the relationship between FlightNumber and Orbit type



EDA & Interactive Visual Analytics Result (2/6)

- Visualize the relationship between Payload Mass and Orbit type



EDA & Interactive Visual Analytics

Result (3/6)

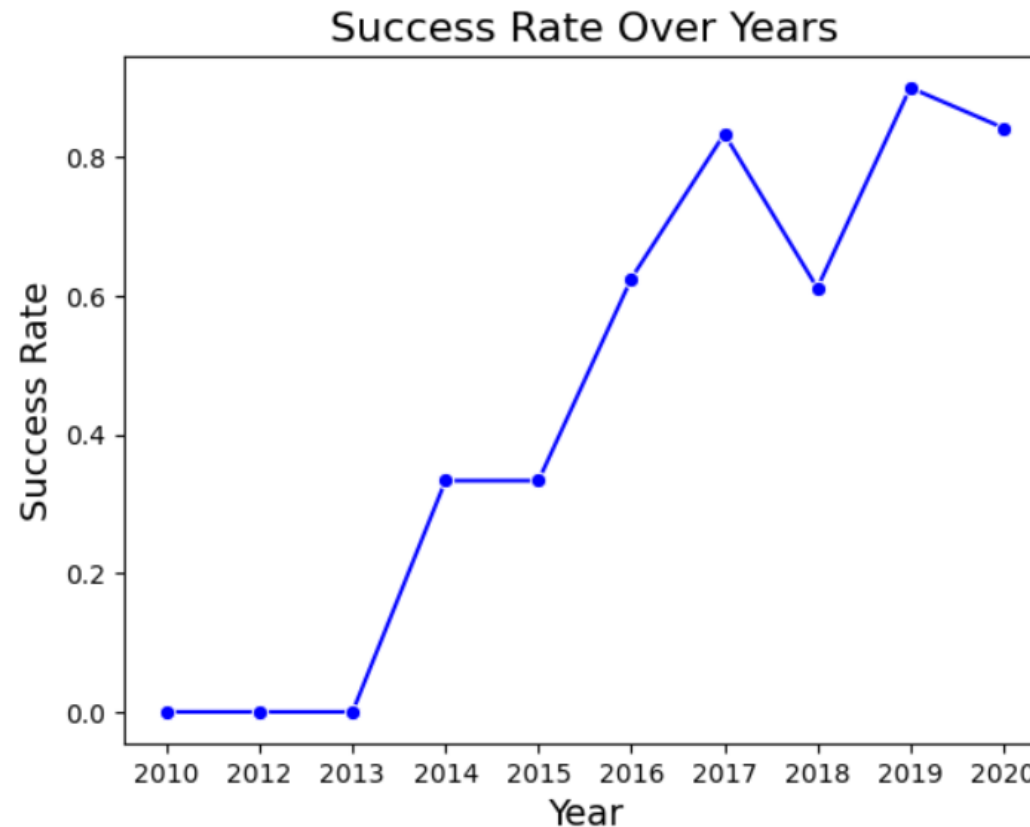
- Visualize the launch success yearly trend

	FlightNumber	Date	BoosterVersion	PayloadMass	Orbit	LaunchSite	Outcome	Flights	GridFins	Reused	Legs	LandingPad	Block
0	1	2010	Falcon 9	6104.959412	LEO	CCAFS SLC 40	None None	1	False	False	False	NaN	1.0
1	2	2012	Falcon 9	525.000000	LEO	CCAFS SLC 40	None None	1	False	False	False	NaN	1.0
2	3	2013	Falcon 9	677.000000	ISS	CCAFS SLC 40	None None	1	False	False	False	NaN	1.0
3	4	2013	Falcon 9	500.000000	PO	VAFB SLC 4E	False Ocean	1	False	False	False	NaN	1.0
4	5	2013	Falcon 9	3170.000000	GTO	CCAFS SLC 40	None None	1	False	False	False	NaN	1.0



EDA & Interactive Visual Analytics Result (4/6)

- Visualize the launch success yearly trend



EDA & Interactive Visual Analytics

Result (5/6)

- Create dummy variables to categorical columns

	FlightNumber	PayloadMass	Flights	GridFins	Reused	Legs	Block	ReusedCount	Orbit_ES-L1	Orbit_GEO	...	Serial_B1048	Serial_B10
0	1	6104.959412	1	False	False	False	1.0	0	False	False	...	False	Fa
1	2	525.000000	1	False	False	False	1.0	0	False	False	...	False	Fa
2	3	677.000000	1	False	False	False	1.0	0	False	False	...	False	Fa
3	4	500.000000	1	False	False	False	1.0	0	False	False	...	False	Fa
4	5	3170.000000	1	False	False	False	1.0	0	False	False	...	False	Fa

EDA & Interactive Visual Analytics

Result (6/6)

- Cast all numeric columns to float 64

	FlightNumber	PayloadMass	Flights	GridFins	Reused	Legs	Block	ReusedCount	Orbit_ES-L1	Orbit_GEO	...	Serial_B1048	Serial_B10
0	1.0	6104.959412	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	...	0.0	
1	2.0	525.000000	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	...	0.0	
2	3.0	677.000000	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	...	0.0	
3	4.0	500.000000	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	...	0.0	
4	5.0	3170.000000	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	...	0.0	

5 rows × 80 columns

EDA With SQL

Result (1-2/10)

- Display the names of the unique launch sites in the space mission
- Display 5 records where launch sites begin with the string 'CCA'

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

Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS_KG_	Orbit	Customer	Mission_Outcome	Landing_Outcome
2010-06-04	18:45:00	F9 v1.0 B0003	CCAFS LC-40	Dragon Spacecraft Qualification Unit	0	LEO	SpaceX	Success	Failure (parachute)
2010-12-08	15:43:00	F9 v1.0 B0004	CCAFS LC-40	Dragon demo flight C1, two CubeSats, barrel of Brouere cheese	0	LEO (ISS)	NASA (COTS) NRO	Success	Failure (parachute)
2012-05-22	7:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525	LEO (ISS)	NASA (COTS)	Success	No attempt
2012-10-08	0:35:00	F9 v1.0 B0006	CCAFS LC-40	SpaceX CRS-1	500	LEO (ISS)	NASA (CRS)	Success	No attempt
2013-03-01	15:10:00	F9 v1.0 B0007	CCAFS LC-40	SpaceX CRS-2	677	LEO (ISS)	NASA (CRS)	Success	No attempt



EDA With SQL

Result (3-4/10)

- Display the total payload mass carried by boosters launched by NASA (CRS)
- Display average payload mass carried by booster version F9 v1.1

Total_Payload_Mass_kg

Average_Payload_Mass_kg

2928.4

EDA With SQL

Result (5-6/10)

- List the date when the first successful landing outcome in ground pad was achieved.
- List the names of the boosters which have success in drone ship and have payload mass greater than 4000 but less than 6000

First_Successful_Ground_Landing_Date

2015-12-22

Booster_Version

F9 FT B1022

F9 FT B1026

F9 FT B1021.2

F9 FT B1031.2



EDA With SQL

Result (7-8/10)

- List the total number of successful and failure mission outcomes
- List all the booster_versions that have carried the maximum payload mass. Use a subquery.

Mission_Outcome	Total_Count
Failure (in flight)	1
Success	98
Success	1
Success (payload status unclear)	1

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



EDA With SQL

Result (9-10/10)

- List the records which will display the month names, failure landing_outcomes in drone ship, booster versions, launch_site for the months in year 2015.
- Rank the count of landing outcomes (such as Failure (drone ship) or Success (ground pad)) between the date 2010-06-04 and 2017-03-20, in descending order.

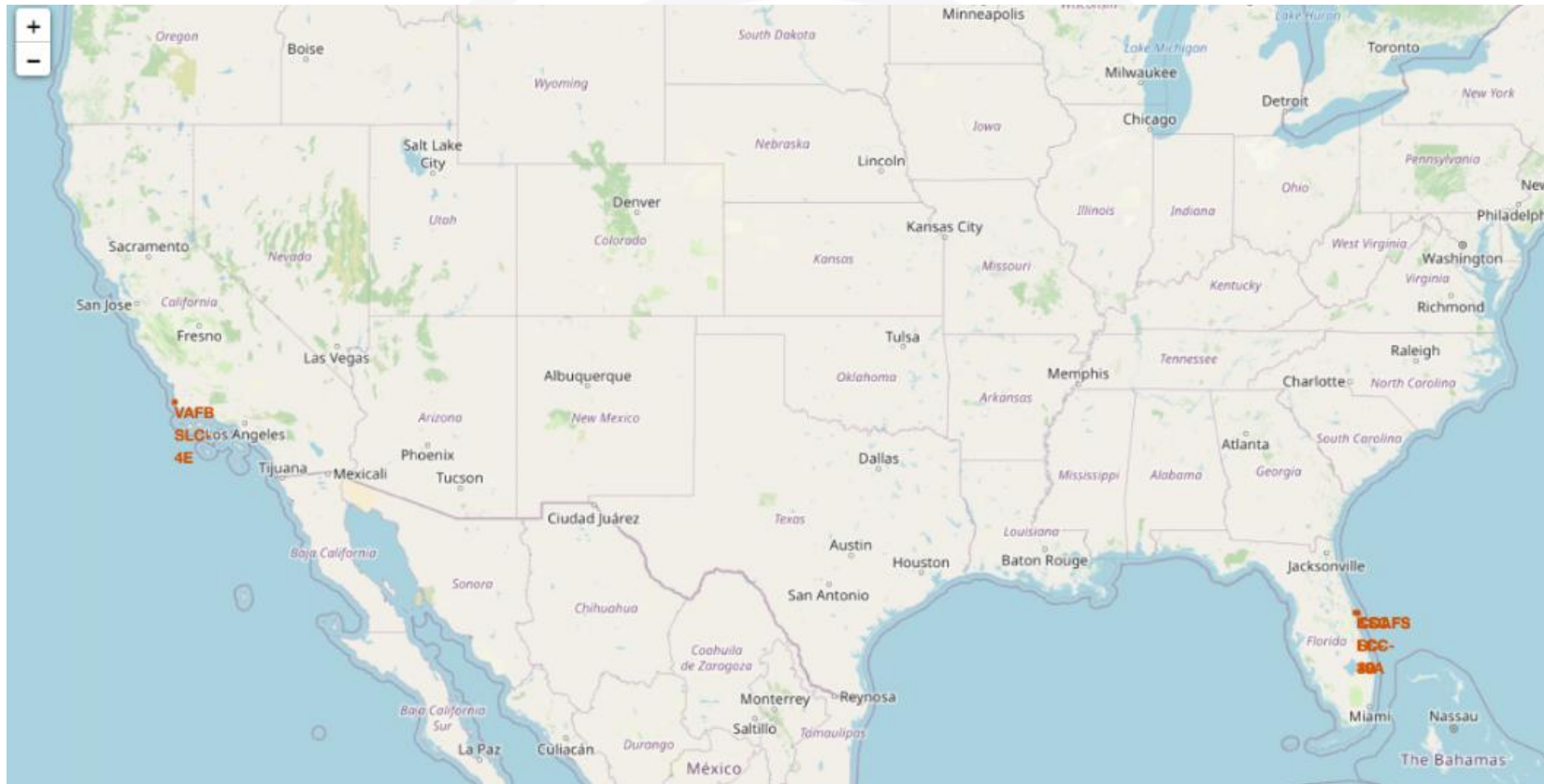
Month	Landing_Outcome	Booster_Version	Launch_Site
01	Failure (drone ship)	F9 v1.1 B1012	CCAFS LC-40
04	Failure (drone ship)	F9 v1.1 B1015	CCAFS LC-40

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



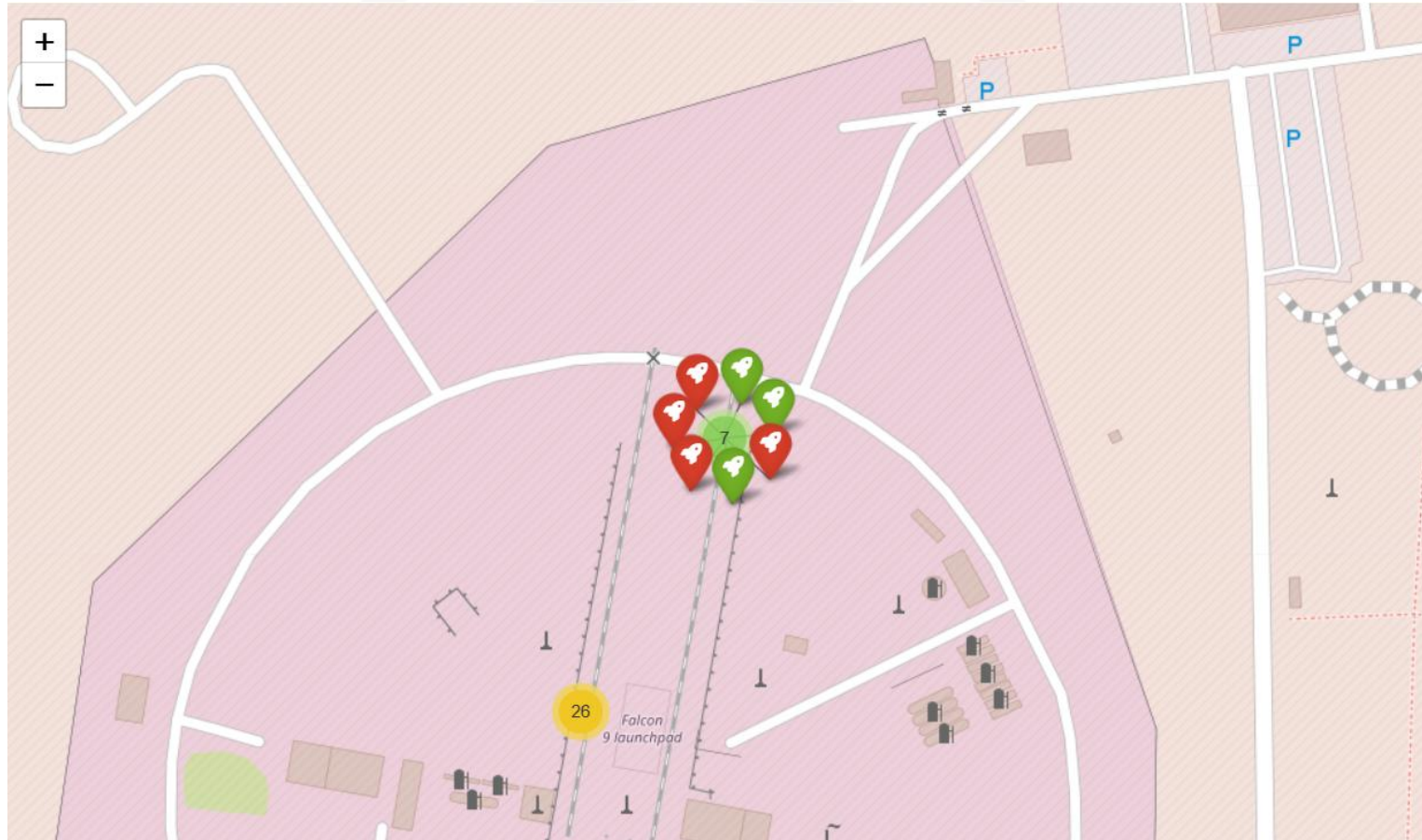
Interactive Map with Folium Results (1/3)

- All launch sites on a map



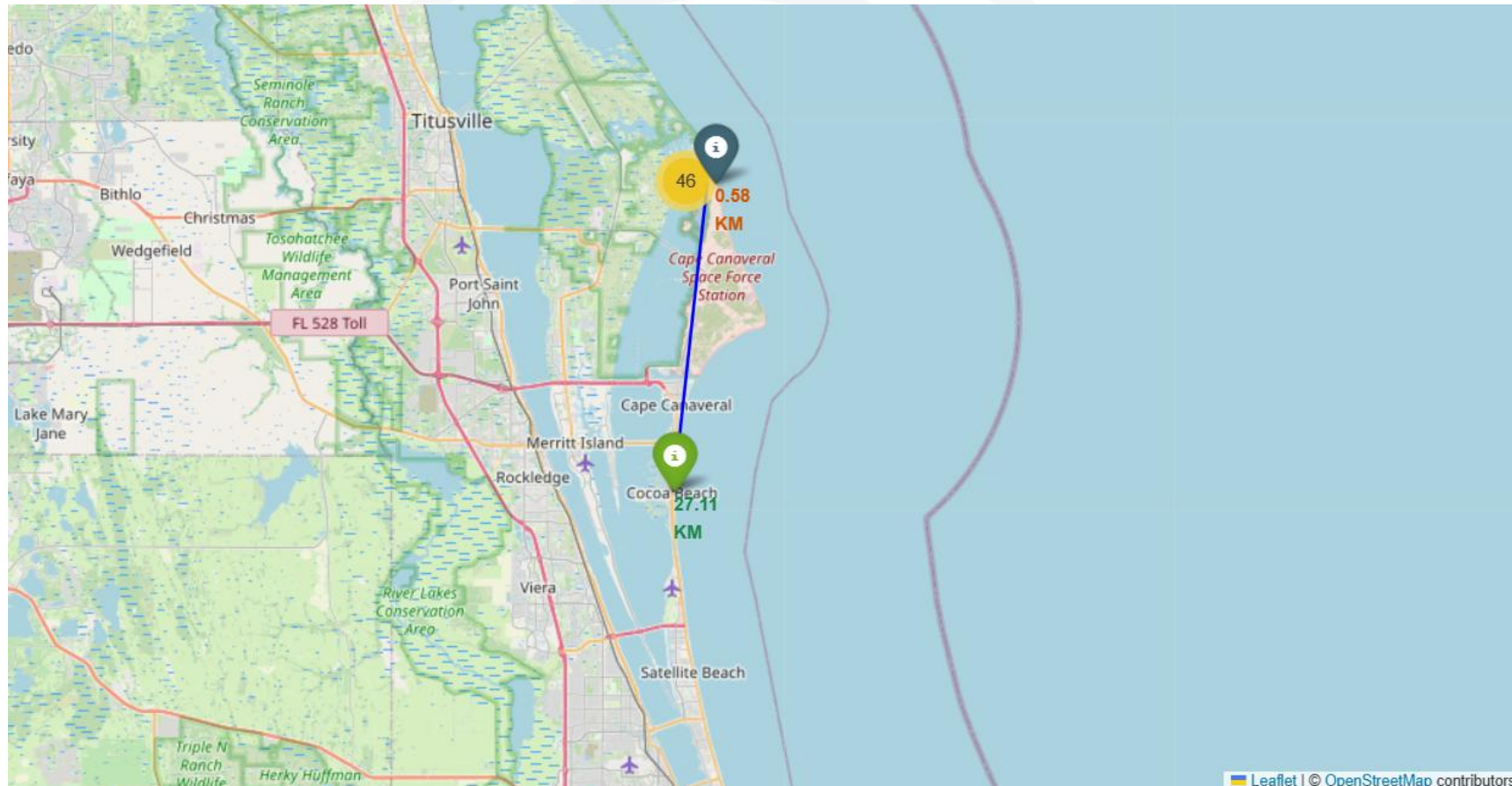
Interactive Map with Folium Results (2/3)

- Success/Failed launches for each site on the map



Interactive Map with Folium Results (3/3)

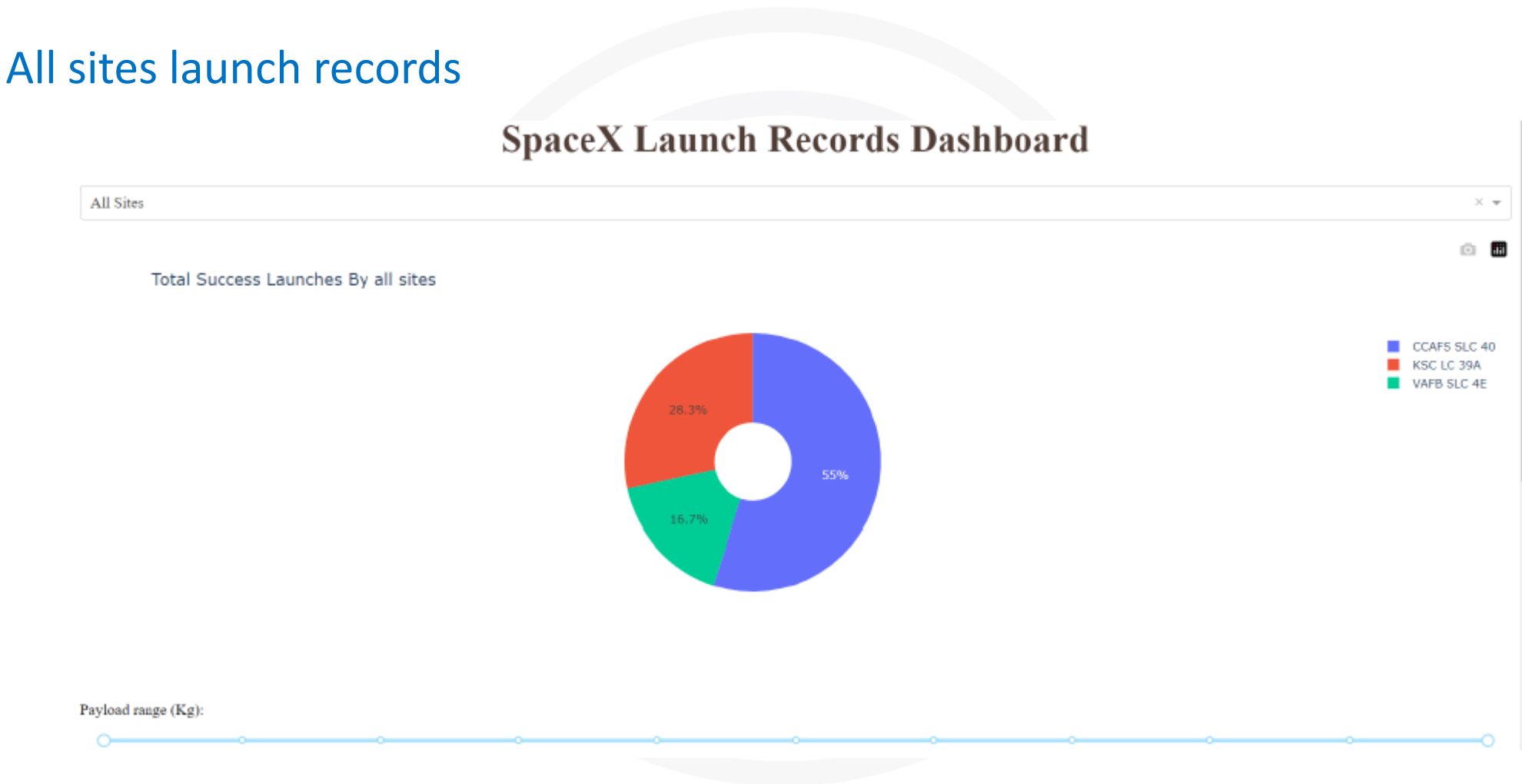
- Distance to a closest city, railway, highway, etc



Plotly Dash dashboard

Results (1/3)

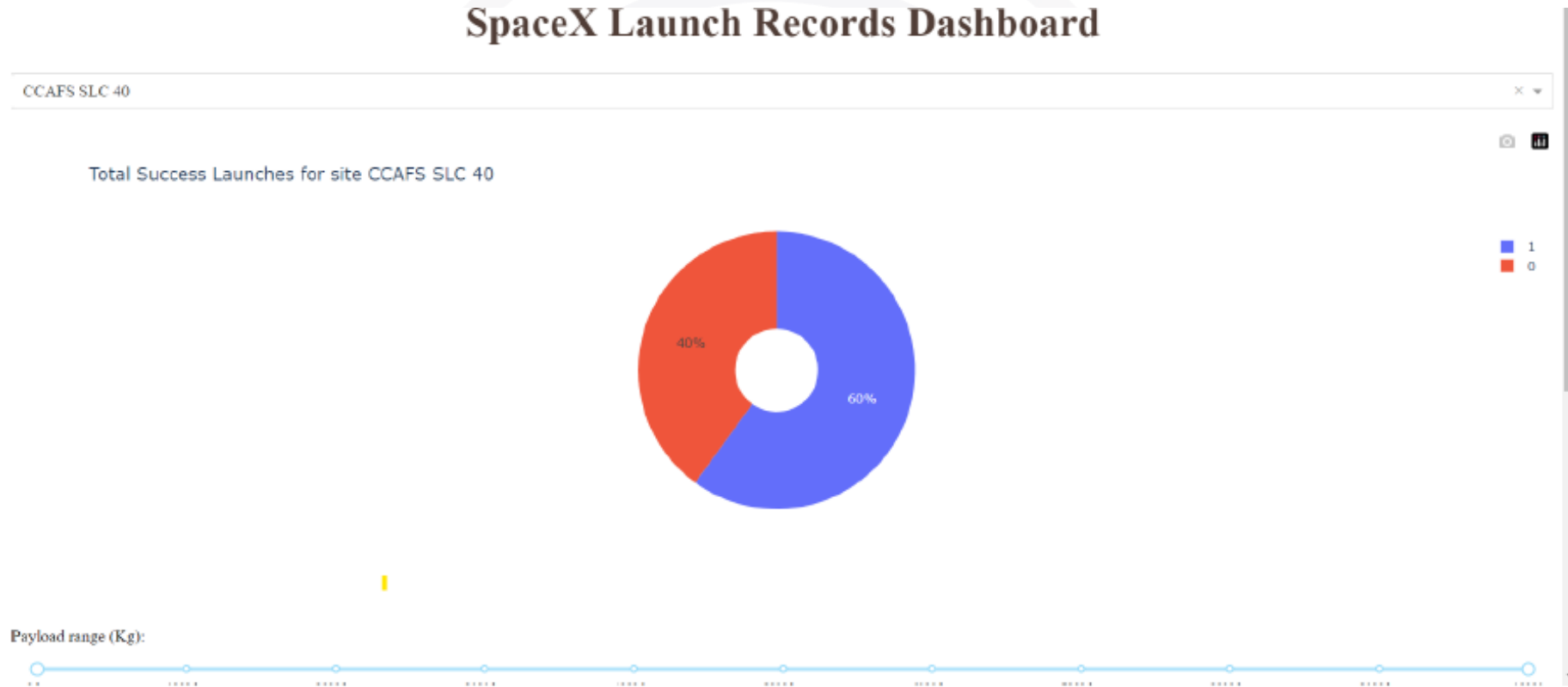
- All sites launch records



Plotly Dash dashboard

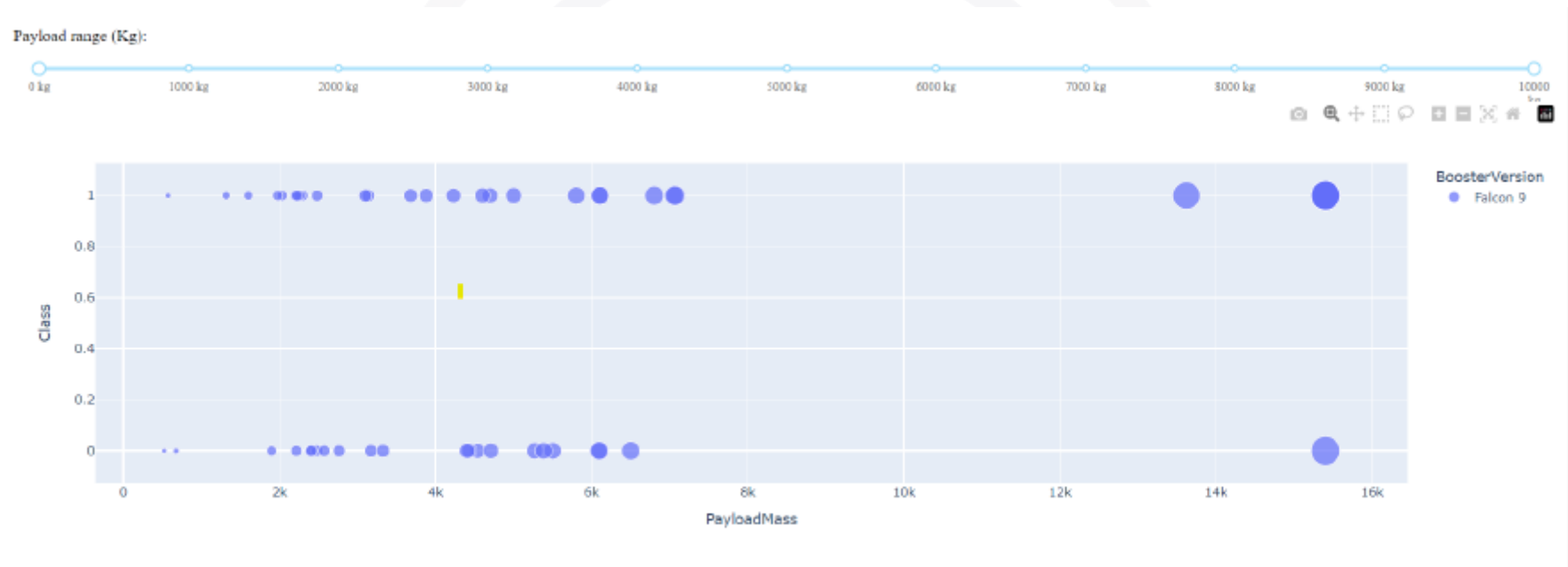
Results (2/3)

- Site CCAFS SLC 40 launch records



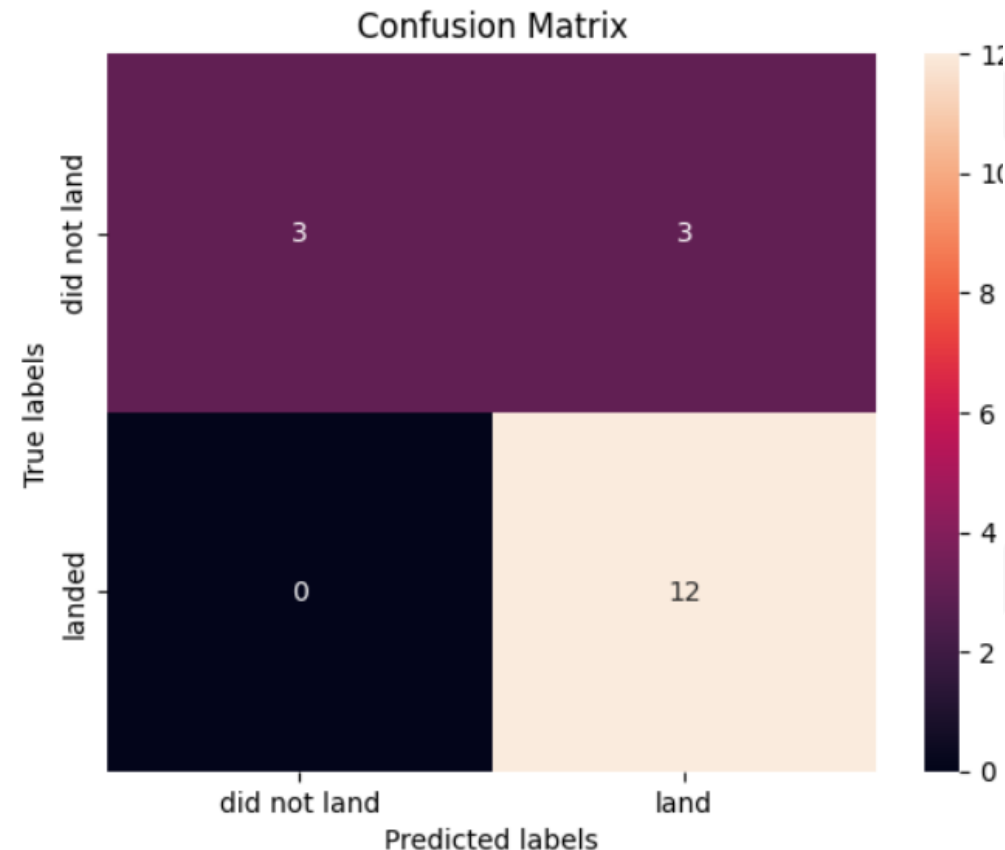
Plotly Dash dashboard Results (3/3)

- Booster version and Payload



Predictive Analysis (classification) Results (1/6)

- Accuracy on the test data using the method score



Predictive Analysis (classification)

Results (2/6)

- Best parameters from the dictionary parameters

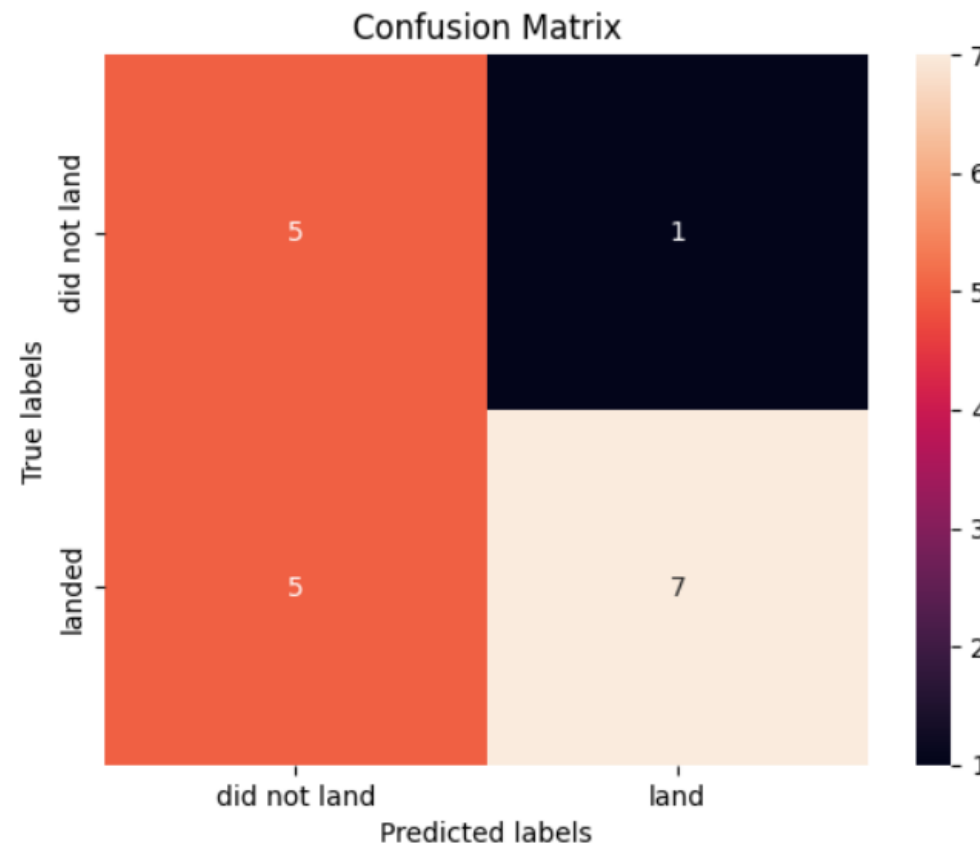
```
tuned hpyerparameters :(best parameters) {'C': 1.0, 'gamma': 0.03162277660168379, 'kernel': 'sigmoid'}  
accuracy : 0.8482142857142856
```



Predictive Analysis (classification)

Results (3/6)

- Accuracy of tree_cv on the test data using the method score



Predictive Analysis (classification)

Results (4/6)

- k nearest neighbors object then create a GridSearchCV object knn_cv with cv = 10. Fit the object to find the best parameters from the dictionary parameters.

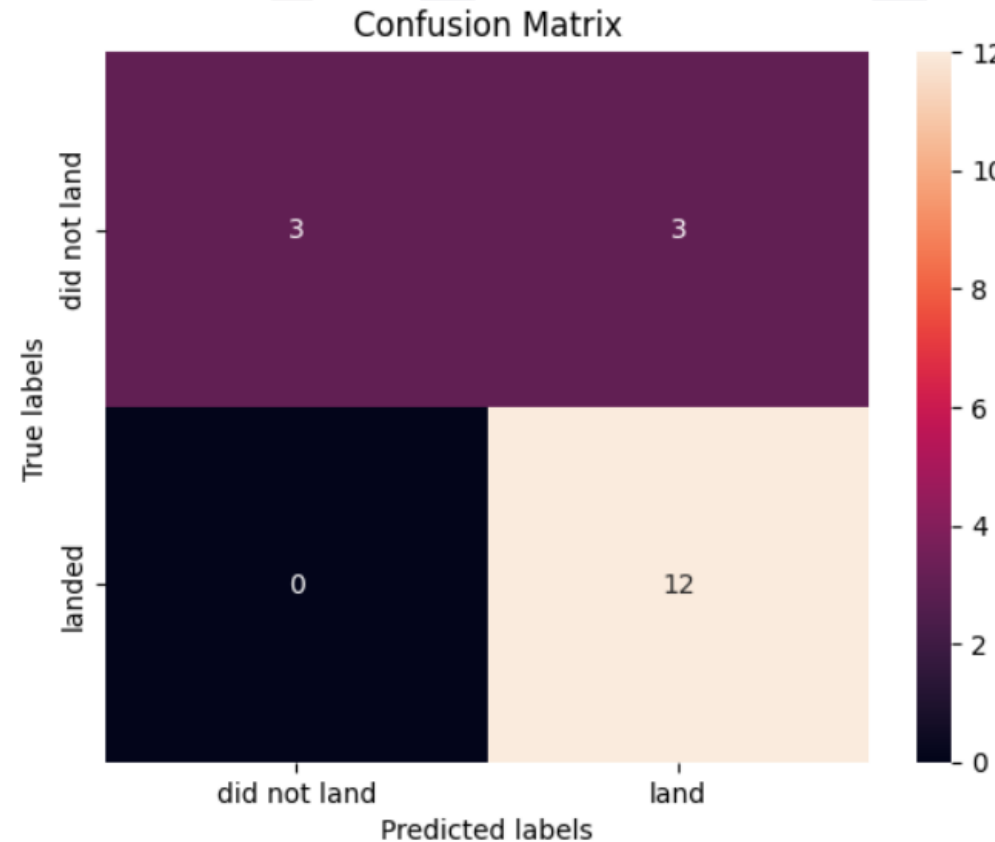
```
tuned hpyerparameters :(best parameters) {'algorithm': 'auto', 'n_neighbors': 10, 'p': 1}  
accuracy : 0.8482142857142858
```



Predictive Analysis (classification)

Results (5/6)

- Accuracy of knn_cv on the test data using the method score



Predictive Analysis (classification) Results (6/6)

- Best Performs

Logistic Regression Accuracy: 0.8333333333333334

SVM Accuracy: 0.8333333333333334

Decision Tree Accuracy: 0.6666666666666666

KNN Accuracy: 0.8333333333333334

Best performing model: Logistic Regression



Conclusion

- The larger the flight amount at a launch site, the greater the success rate at a launch site.
- Launch success rate started to increase in 2013 till 2020.
- Orbits ES-L1, GEO, HEO, SSO, VLEO had the most success rate.
- KSC LC-39A had the most successful launches of any sites.
- The Decision tree classifier is the best machine learning algorithm for this task.

