IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Frederico Assis
April 22th, 2024

Outline

- Executive Summary
- Introduction
- Methodology
- Results
- Conclusion
- Appendix

Executive Summary

Research attempts to identify the key factors for a successful rocket landing. To achieve that, the following analysis were made:

- Collecting the data
- Web scraping Falcon 9 and Falcon Heavy Launches Records from Wikipedia
- Overview of the DataSet
- Exploring and Preparing Data
- Perform exploratory Data Analysis and determine Training Labels
- Launch Sites Locations Analysis with Folium
- Machine Learning Prediction

Summary of all results

- Exploratory Data Analysis: Understanding in absolute numbers the quantity of success and failure across different time windows among other variables.
- Launch Sites Locations Analysis with Folium: The closer the launch site to the equator, the easier it is to launch to equatorial orbit, and the more help you get from Earth's rotation for a prograde orbit. Rockets launched from sites near the equator get an additional natural boost - due to the rotational speed of earth - that helps save the cost of putting in extra fuel and boosters.
- Predictions: All models performed similarly; however, one of them had a slight edge with slightly higher accuracy, tree decisions.

Introduction

- SpaceX, founded by Elon Musk in 2002, is a pioneering aerospace manufacturer and space transportation company. It aims to revolutionize space technology, making access to space more affordable and sustainable. Known for its Falcon rockets and Dragon spacecraft, SpaceX has achieved remarkable milestones, including the first privately-funded spacecraft to reach orbit and dock with the International Space Station (ISS). With ambitious plans for Mars colonization and satellite internet constellation, SpaceX continues to push the boundaries of space exploration and commercial spaceflight.
- We will use the data provided by SpaceX to try to understand what would be the key factors for a successful land landing.

Section 1

Methodology

Methodology

Executive Summary

- Data collection methodology:
  - The data was obtained directly through SpaceX's provision as well as web scraping from other sources.
- Perform data wrangling
  - The data was cleaned, filtered, missing values were replaced, and organized to be used for the analyses to be conducted.
- Perform exploratory data analysis (EDA) using visualization and SQL
- Perform interactive visual analytics using Folium and Plotly Dash
- Perform predictive analysis using classification models
  - Several models were created and compared to each other, aiming to select the best model for the current situation.

Data Collection – SpaceX API and Web Scraping

- We utilized the SpaceX API to collect data for the analyses. Additionally, a web scraping task was conducted on Wikipedia to gather other company-related data.

```python
# Takes the dataset and uses the rocket column to call the API and append the data to the list
def getBoosterVersion(data):
    for x in data['rocket']:
      if x:
        response = requests.get("https://api.spacexdata.com/v4/rockets/"+str(x)).json()
        BoosterVersion.append(response['name'])
```

https://github.com/fredpessoa/IBM-Capstone/blob/main/1%20-%20Data%20%20Collection%20API.ipynb

https://api.spacexdata.com/v4/rockets/

[hide] Flight No.	Date and time (UTC)	Version, Booster[b]	Launch site	Payload[c]	Payload mass	Orbit	Customer	Launch outcome	Booster landing
78	7 January 2020, 02:19:21[492]	F9 B5 ♺ B1049.4	CCAFS, SLC-40	Starlink 2 v1.0 (60 satellites)	15,600 kg (34,400 lb)[5]	LEO	SpaceX	Success	Success (drone ship)
	Third large batch and second operational flight of Starlink constellation. One of the 60 satellites included a test coating to make the satellite less reflective, and thus less likely to interfere with ground-based astronomical observations.[493]								
	19 January 2020, 15:30[494]	F9 B5 ♺ B1046.4	KSC, LC-39A	Crew Dragon in-flight abort test[495] (Dragon C205.1)	12,050 kg (26,570 lb)	Sub-orbital[496]	NASA (CTS)[497]	Success	No attempt

https://en.wikipedia.org/wiki/List_of_Falcon_9_and_Falcon_Heavy_launches

https://github.com/fredpessoa/IBM-Capstone/blob/main/2%20-%20Webscraping.ipynb

Data Wrangling

After completing Exploratory Data Analysis and determining Training Labels:

- Exploratory Data Analysis
- Valuable insights were gained from data exploration, identifying key patterns and factors influencing launch outcomes.
- Determine Training Labels
- Training labels were established based on launch success criteria, crucial for model training and accurate predictions on new data.

https://github.com/fredpessoa/IBM-Capstone/blob/main/3%20-%20Data%20Wrangling.ipynb

EDA with SQL

- The SpaceX dataset has been loaded and stored in a database table, with subquery routines already implemented to gain insights into the obtained data.

Out[9]:

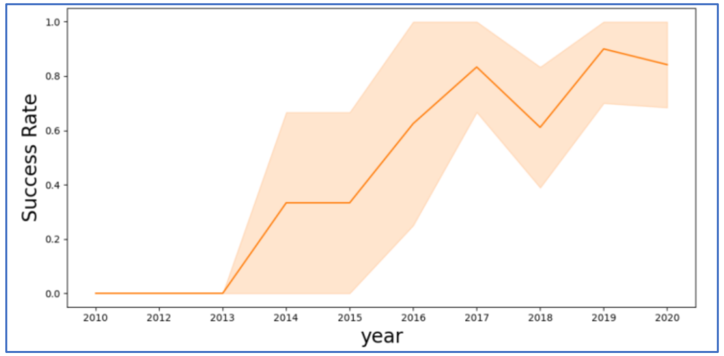
Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer	Mission_Outcome	Landing _Outcome
04-06-2010	18:45:00	F9 v1.0 B0003	CCAFS LC-40	Dragon Spacecraft Qualification Unit	0	LEO	SpaceX	Success	Failure (parachute)
08-12-2010	15:43:00	F9 v1.0 B0004	CCAFS LC-40	Dragon demo flight C1, two CubeSats, barrel of Brouere cheese	0	LEO (ISS)	NASA (COTS) NRO	Success	Failure (parachute)
22-05-2012	07:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525	LEO (ISS)	NASA (COTS)	Success	No attempt
08-10-2012	00:35:00	F9 v1.0 B0006	CCAFS LC-40	SpaceX CRS-1	500	LEO (ISS)	NASA (CRS)	Success	No attempt
01-03-2013	15:10:00	F9 v1.0 B0007	CCAFS LC-40	SpaceX CRS-2	677	LEO (ISS)	NASA (CRS)	Success	No attempt

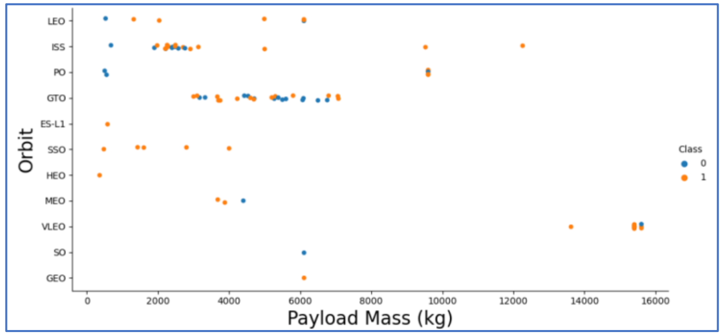
https://github.com/fredpessoa/IBM-Capstone/blob/main/4%20-%20EDA%20SQL.ipynb

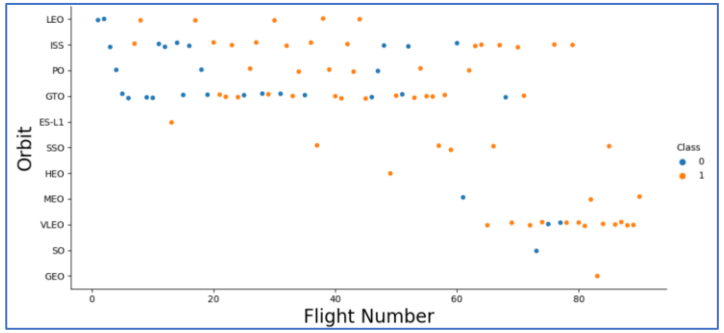
EDA with Data Visualization

Exploratory Data Analysis has been conducted, alongside Data Feature Engineering using Pandas and Matplotlib, preparing the dataset for further analysis and modeling.

https://github.com/fredpessoa/IBM-Capstone/blob/main/5%20-%20EDA%20Viz.ipynb







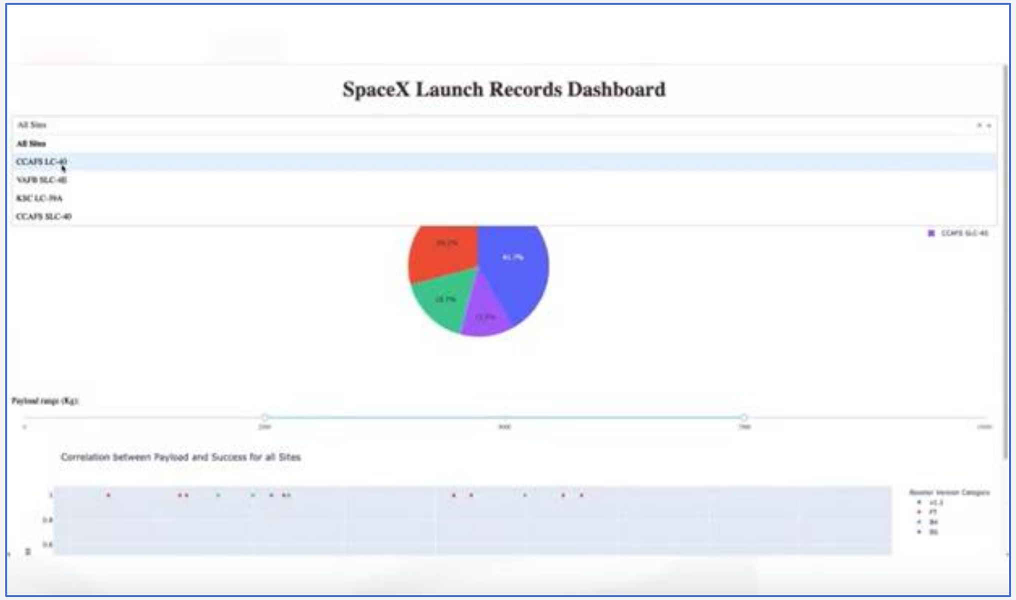
Build an Interactive Map with Folium

- We successfully completed the following tasks: marking all launch sites on a map, indicating success/failed launches for each site, and calculating distances between a launch site and its proximities. This allowed us to uncover valuable geographical patterns surrounding the launch sites.

https://github.com/fredpessoa/IBM-Capstone/blob/main/6%20-%20Analysis%20with%20Folium.ipynb

Build a Dashboard with Plotly Dash

- Plots/graphs and interactions you have added to a dashboard



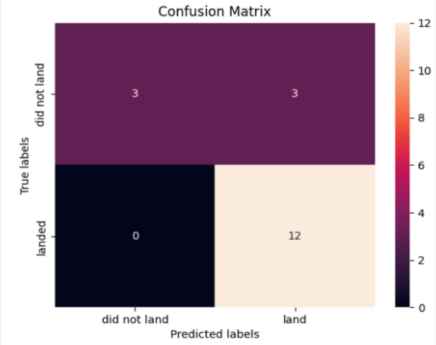
https://github.com/fredpessoa/IBM-Capstone/blob/main/7%20-%20Analytics%20Plotly.py

Predictive Analysis (Classification)

- Having completed the tasks, we've performed exploratory data analysis, determined training labels by creating a column for the class, standardized the data, and split it into training and test datasets. Furthermore, we found the best hyperparameters for SVM, Classification Trees, and Logistic Regression models. Finally, we identified the best-performing method using the test data.

Out[7]:

	FlightNumber	PayloadMass	Flights	Block	ReusedCount	Orbit_ES-L1	Orbit_GEO	Orbit_GTO	Orbit_HEO	Orbit_ISS	...	Serial_B1058
0	1.0	6104.959412	1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0
1	2.0	525.000000	1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0
2	3.0	677.000000	1.0	1.0	0.0	0.0	0.0	0.0	0.0	1.0	...	0.0
3	4.0	500.000000	1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0
4	5.0	3170.000000	1.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	...	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...
85	86.0	15400.000000	2.0	5.0	2.0	0.0	0.0	0.0	0.0	0.0	...	0.0
86	87.0	15400.000000	3.0	5.0	2.0	0.0	0.0	0.0	0.0	0.0	...	1.0
87	88.0	15400.000000	6.0	5.0	5.0	0.0	0.0	0.0	0.0	0.0	...	0.0
88	89.0	15400.000000	3.0	5.0	2.0	0.0	0.0	0.0	0.0	0.0	...	0.0
89	90.0	3681.000000	1.0	5.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0



https://github.com/fredpessoa/IBM-Capstone/blob/main/8%20-%20Predictive%20Analytics.ipynb

Results

Data Exploration:

Over time, there has been a noticeable improvement in launch success rates.
KSC LC-39A stands out as the landing site with the highest success rate.
Orbits ES-L1, GEO, HEO, and SSO have all achieved a perfect 100% success rate.

Visual Analysis:

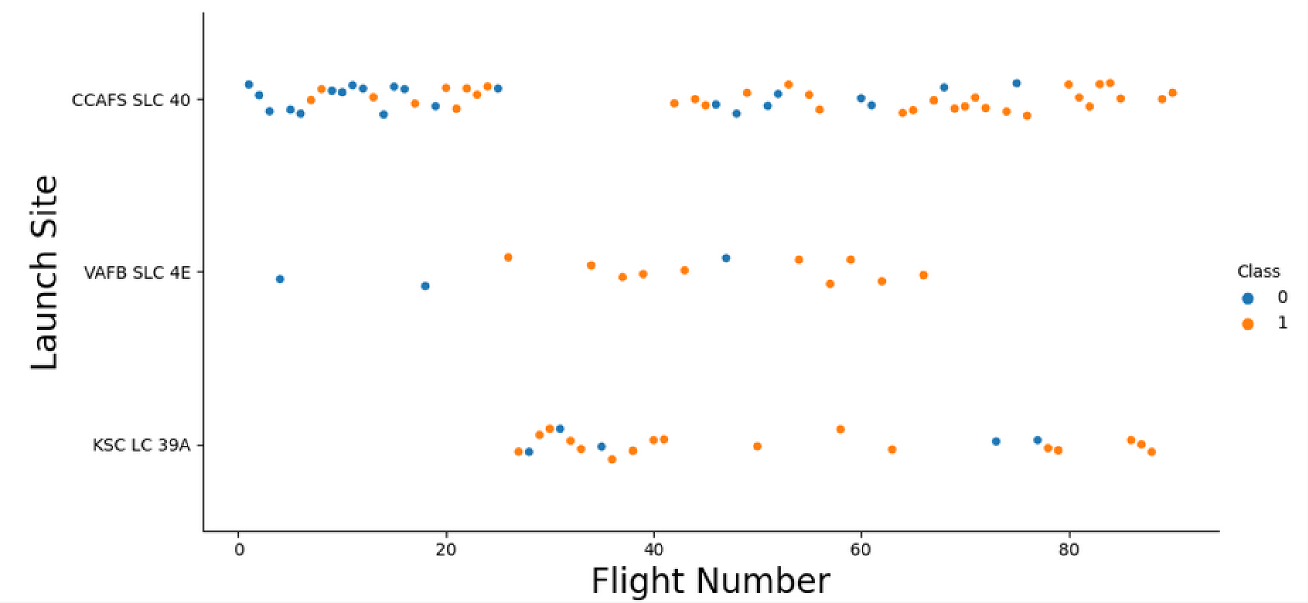
Most launch sites are located near the equator and are adjacent to coastlines.
Launch sites are strategically situated at a safe distance from populated areas (cities, highways, railways) to mitigate potential damage from failed launches, while remaining accessible for logistical support.

Predictive Modeling:

Based on our analysis, the Decision Tree model emerges as the most effective predictive model for the dataset.

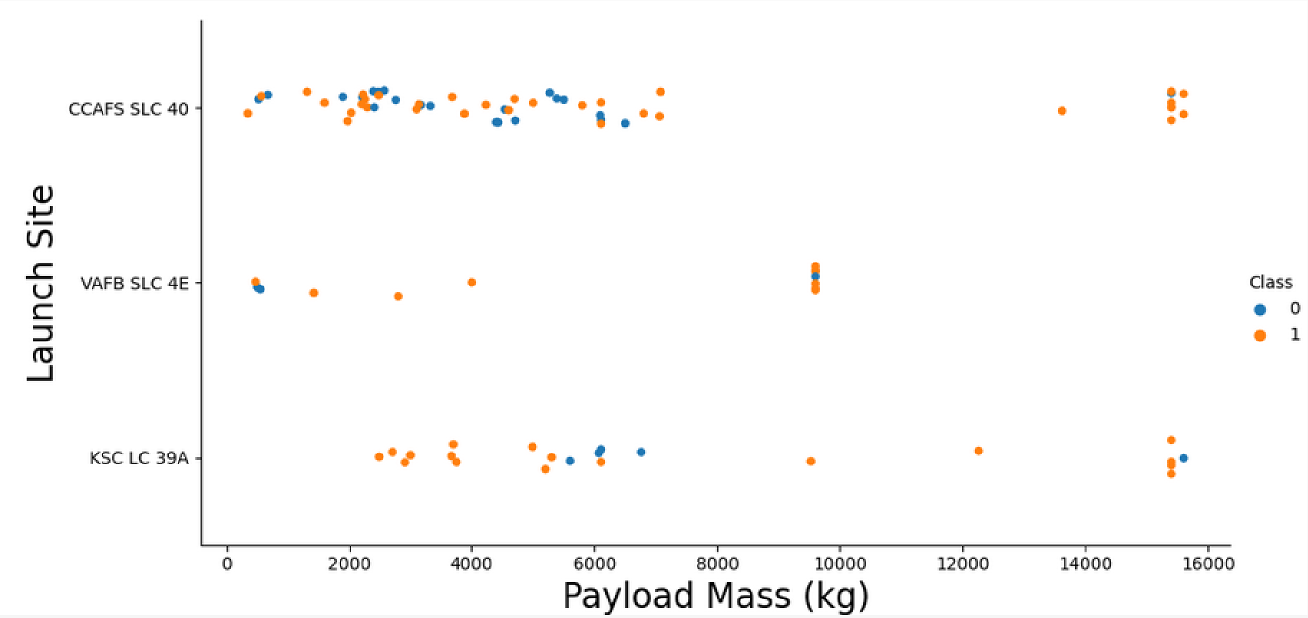
Section 2

Insights drawn from EDA



Flight Number vs. Launch Site

- Blue = Fail and Orange = Success
- VAFB SLC 4E and KSC LC 39A have higher success rates than CCAFS SLC 40

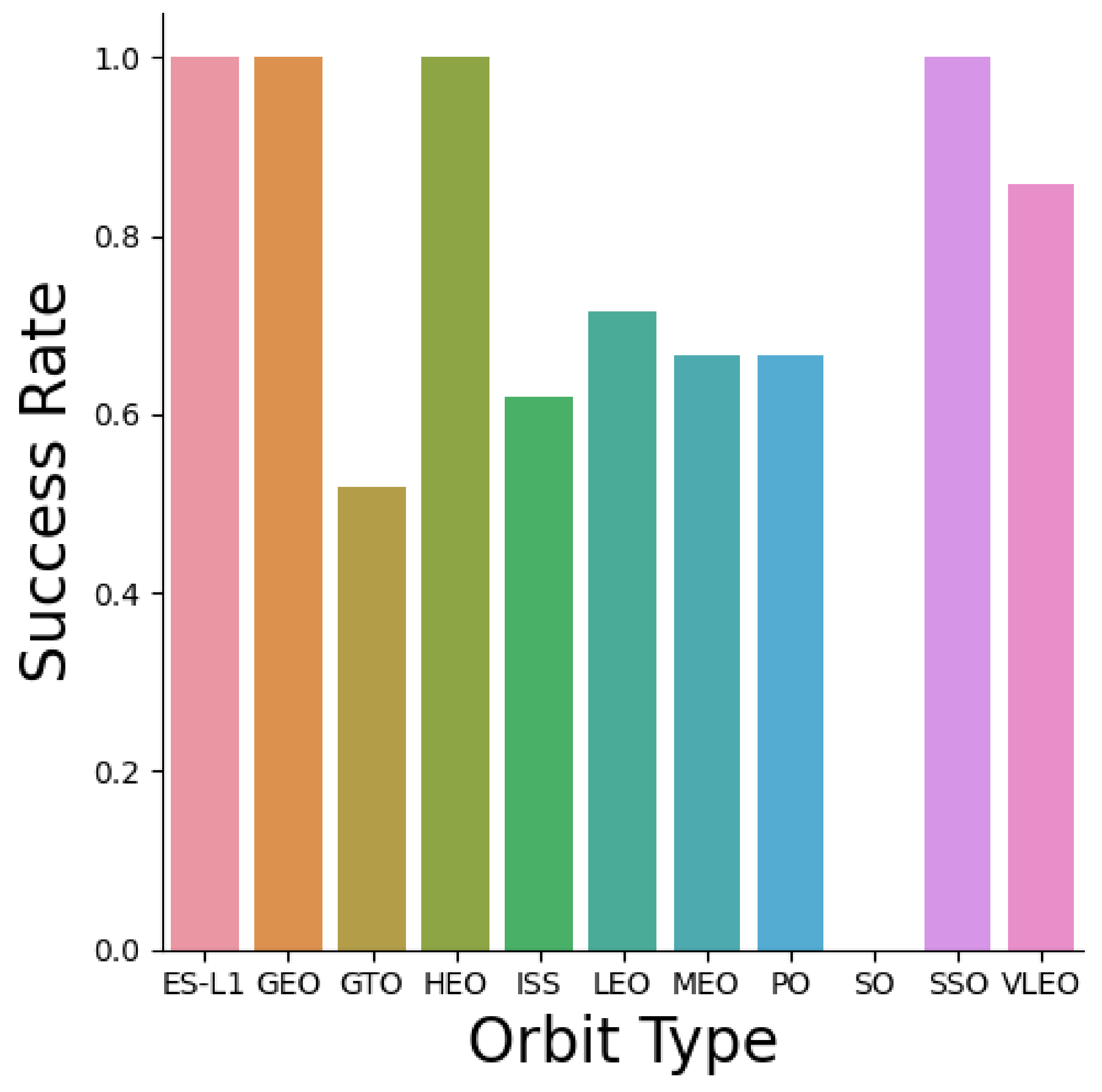


Payload vs. Launch Site

- Blue = Fail and Orange = Success

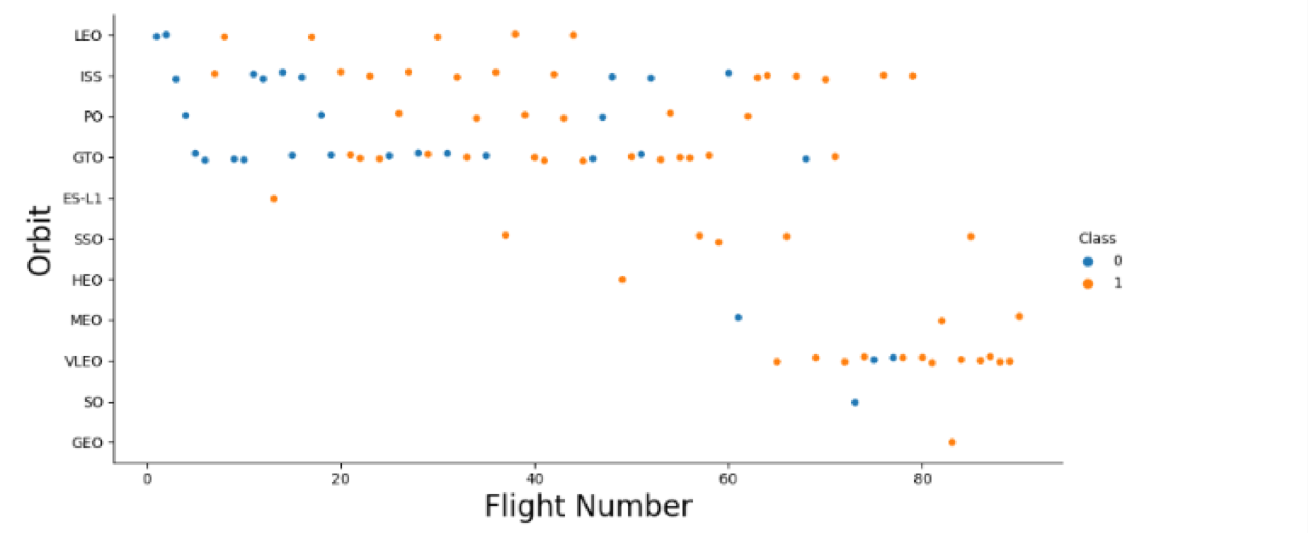
- Most launces with a payload greater than 8,000 kg were successful

- KSC LC 39A has a 100% success rate for launches less than 4,000 kg



Success Rate vs. Orbit Type

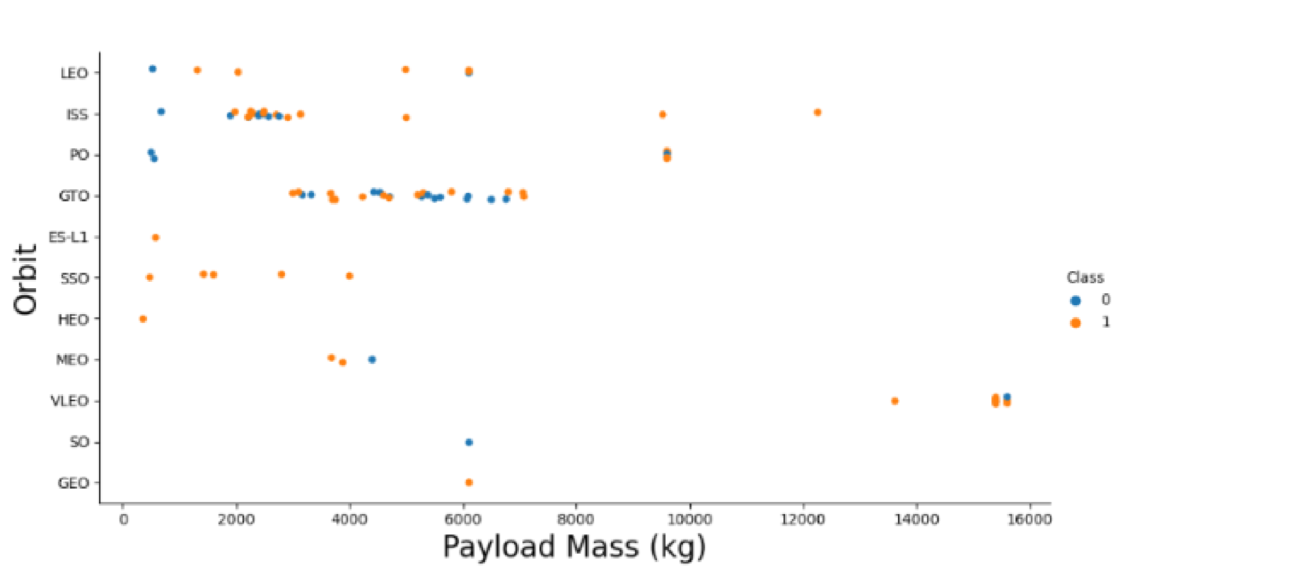
- 100% Success Rate: ES-L1, GEO, HEO and SSO
- 50%-80% Success Rate: GTO, ISS, LEO, MEO, PO
- 0% Success Rate: SO



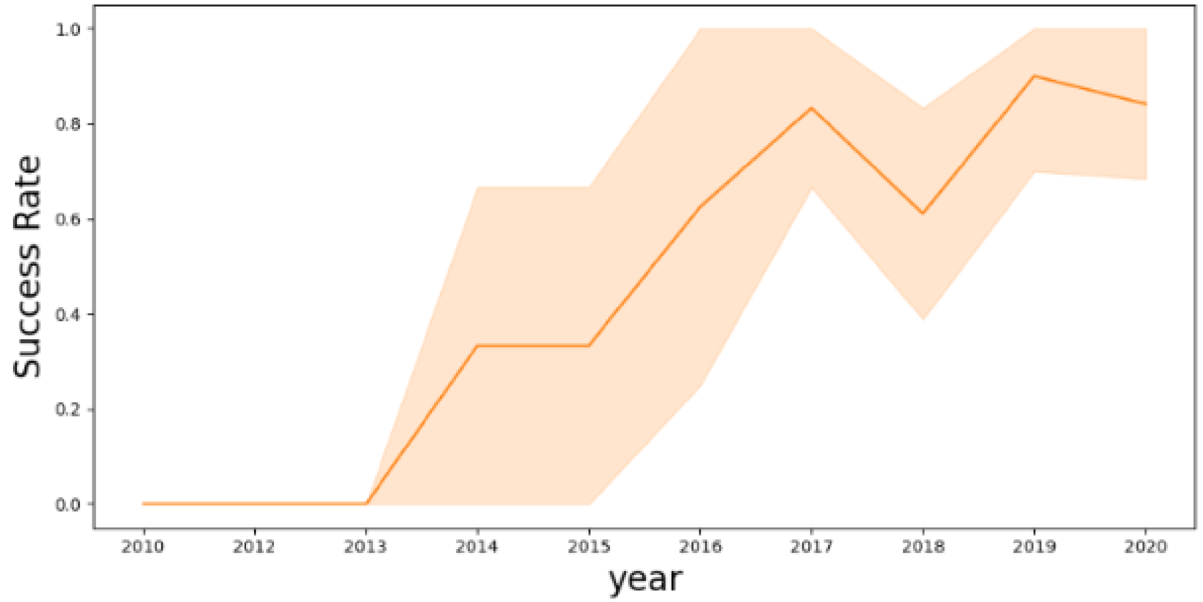
Flight Number vs. Orbit Type

- Blue = Fail and Orange = Success
- The success rate increases with the number of flights for each orbit

Payload vs. Orbit Type



- Blue = Fail and Orange = Success
- Without any clear indication of trend



Launch Success Yearly Trend

- An upward trend over the years, indicating that experience seems to lead to an increasing success rate as time goes by

Task 1

Display the names of the unique launch sites in the space mission

[17]:
```
%sql ibm_db_sa://yyy33800:dwNKg8J3L0IBd6CP@1bbf73c5-d84a-4bb0-85b9-ab1a4348f4a4.c3n41cmd0nqnrk39u98g.databases.appdomain.cloud:32286/BLUDB?security=SSL
%sql SELECT DISTINCT LAUNCH_SITE FROM SPACEXTBL;
```

```
(ibm_db_dbi.OperationalError) ibm_db_dbi::OperationalError: [IBM][CLI Driver] SQL30082N  Security processing failed with reason "24" ("USERNAME AND/OR PASSWORD INVALID").  SQLSTATE=08001 SQLCODE=-3008
2
(Background on this error at: http://sqlalche.me/e/e3q8)
Connection info needed in SQLAlchemy format, example:
               postgresql://username:password@hostname/dbname
               or an existing connection: dict_keys(['sqlite:///my_data1.db'])
 * sqlite:///my_data1.db
Done.
```

[17]:

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

All Launch Site Names

Four launch sites:

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

Task 2

Display 5 records where launch sites begin with the string 'CCA'

[18]:
```
%sql SELECT * \
    FROM SPACEXTBL \
    WHERE LAUNCH_SITE LIKE'CCA%' LIMIT 5;
```

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

[18]:

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

Launch Site Names Begin with 'CCA'

- Displaying 5 of the first records above

Task 3

Display the total payload mass carried by boosters launched by NASA (CRS)

```
[19]: %sql SELECT SUM(PAYLOAD_MASS__KG_) \
          FROM SPACEXTBL \
          WHERE CUSTOMER = 'NASA (CRS)';

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

[19]:

SUM(PAYLOAD_MASS__KG_)
45596

Total Payload Mass

- 45,596 Kg carried by boosters launched by NASA (CRS)

Task 4

Display average payload mass carried by booster version F9 v1.1

```
[20]: %sql SELECT AVG(PAYLOAD_MASS__KG_) \
          FROM SPACEXTBL \
          WHERE BOOSTER_VERSION = 'F9 v1.1';

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

[20]:

AVG(PAYLOAD_MASS__KG_)
2928.4

Average Payload Mass by F9 v1.1

- Average payload mass – 292,4 Kg
- Average payload mass carried by booster version F9 V1.1

Task 5

List the date when the first succesful landing outcome in ground pad was acheived.

Hint:Use min function

[26]:
```
%sql SELECT * FROM SPACEXTBL LIMIT 20;
```

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

[26]:

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
2013-09-29	16:00:00	F9 v1.1 B1003	VAFB SLC-4E	CASSIOPE	500	Polar LEO	MDA	Success	Uncontrolled (ocean)
2013-12-03	22:41:00	F9 v1.1	CCAFS LC-40	SES-8	3170	GTO	SES	Success	No attempt
2014-01-06	22:06:00	F9 v1.1	CCAFS LC-40	Thaicom 6	3325	GTO	Thaicom	Success	No attempt
2014-04-18	19:25:00	F9 v1.1	CCAFS LC-40	SpaceX CRS-3	2296	LEO (ISS)	NASA (CRS)	Success	Controlled (ocean)
2014-07-14	15:15:00	F9 v1.1	CCAFS LC-40	OG2 Mission 1 6 Orbcomm-OG2 satellites	1316	LEO	Orbcomm	Success	Controlled (ocean)
2014-08-05	8:00:00	F9 v1.1	CCAFS LC-40	AsiaSat 8	4535	GTO	AsiaSat	Success	No attempt
2014-09-07	5:00:00	F9 v1.1 B1011	CCAFS LC-40	AsiaSat 6	4428	GTO	AsiaSat	Success	No attempt
2014-09-21	5:52:00	F9 v1.1 B1010	CCAFS LC-40	SpaceX CRS-4	2216	LEO (ISS)	NASA (CRS)	Success	Uncontrolled (ocean)
2015-01-10	9:47:00	F9 v1.1 B1012	CCAFS LC-40	SpaceX CRS-5	2395	LEO (ISS)	NASA (CRS)	Success	Failure (drone ship)
2015-02-11	23:03:00	F9 v1.1 B1013	CCAFS LC-40	DSCOVR	570	HEO	U.S. Air Force NASA NOAA	Success	Controlled (ocean)
2015-03-02	3:50:00	F9 v1.1 B1014	CCAFS LC-40	ABS-3A Eutelsat 115 West B	4159	GTO	ABS Eutelsat	Success	No attempt
2015-04-14	20:10:00	F9 v1.1 B1015	CCAFS LC-40	SpaceX CRS-6	1898	LEO (ISS)	NASA (CRS)	Success	Failure (drone ship)
2015-04-27	23:03:00	F9 v1.1 B1016	CCAFS LC-40	Turkmen 52 / MonacoSAT	4707	GTO	Turkmenistan National Space Agency	Success	No attempt
2015-06-28	14:21:00	F9 v1.1 B1018	CCAFS LC-40	SpaceX CRS-7	1952	LEO (ISS)	NASA (CRS)	Failure (in flight)	Precluded (drone ship)
2015-12-22	1:29:00	F9 FT B1019	CCAFS LC-40	OG2 Mission 2 11 Orbcomm-OG2 satellites	2034	LEO	Orbcomm	Success	Success (ground pad)

First Successful Ground Landing Date

- 2015-12-22
- Success (ground pad)

Task 6

List the names of the boosters which have success in drone ship and have payload mass greater than 4000 but less than 6000

[33]:
```
%sql SELECT PAYLOAD \
FROM SPACEXTBL \
WHERE Landing_Outcome = 'Success (drone ship)' \
AND PAYLOAD_MASS__KG_ BETWEEN 4000 AND 6000;
```

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

[33]:

Payload
JCSAT-14
JCSAT-16
SES-10
SES-11 / EchoStar 105

Successful Drone Ship Landing with Payload between 4000 and 6000

- JCSAT-14
- JCSAT-16
- SES-10
- SES-11/EchoStar 105

Task 7

List the total number of successful and failure mission outcomes

[34]:
```
%sql SELECT MISSION_OUTCOME, COUNT(*) as total_number \
FROM SPACEXTBL \
GROUP BY MISSION_OUTCOME;
```

* sqlite:///my_data1.db
Done.

[34]:

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

Total Number of Successful and Failure Mission Outcomes

Task 8

List the names of the booster_versions which have carried the maximum payload mass. Use a subquery

```
[35]: %sql SELECT BOOSTER_VERSION \
FROM SPACEXTBL \
WHERE PAYLOAD_MASS__KG_ = (SELECT MAX(PAYLOAD_MASS__KG_) FROM SPACEXTBL);
```

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

[35]:

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

Boosters Carried Maximum Payload

- Names of the booster versions which have carried the maximum payload mass

Task 9

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.

Note: SQLLite does not support monthnames. So you need to use substr(Date, 6,2) as month to get the months and substr(Date,0,5)='2015' for year.

[40]:
```
%sql SELECT substr(Date,6,2) as month, DATE, BOOSTER_VERSION, LAUNCH_SITE, [Landing_Outcome] \
FROM SPACEXTBL \
where [Landing_Outcome] = 'Failure (drone ship)' and substr(Date,0,5)='2015';
```

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

[40]:

month	Date	Booster_Version	Launch_Site	Landing_Outcome
01	2015-01-10	F9 v1.1 B1012	CCAFS LC-40	Failure (drone ship)
04	2015-04-14	F9 v1.1 B1015	CCAFS LC-40	Failure (drone ship)

2015 Launch Records

- Failed landing outcomes in drone ship, their booster versions, and launch site names in year 2015

Task 10

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.

[42]:
```
%sql SELECT [Landing_Outcome], count(*) as count_outcomes \
FROM SPACEXTBL \
WHERE DATE between '2010-06-04' and '2017-03-20' group by [Landing_Outcome] order by count_outcomes DESC;
```

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

[42]:

Landing_Outcome	count_outcomes
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

Rank Landing Outcomes Between 2010-06-04 and 2017-03-20

- 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

Section 3

Launch Sites Proximities Analysis

Folium Map

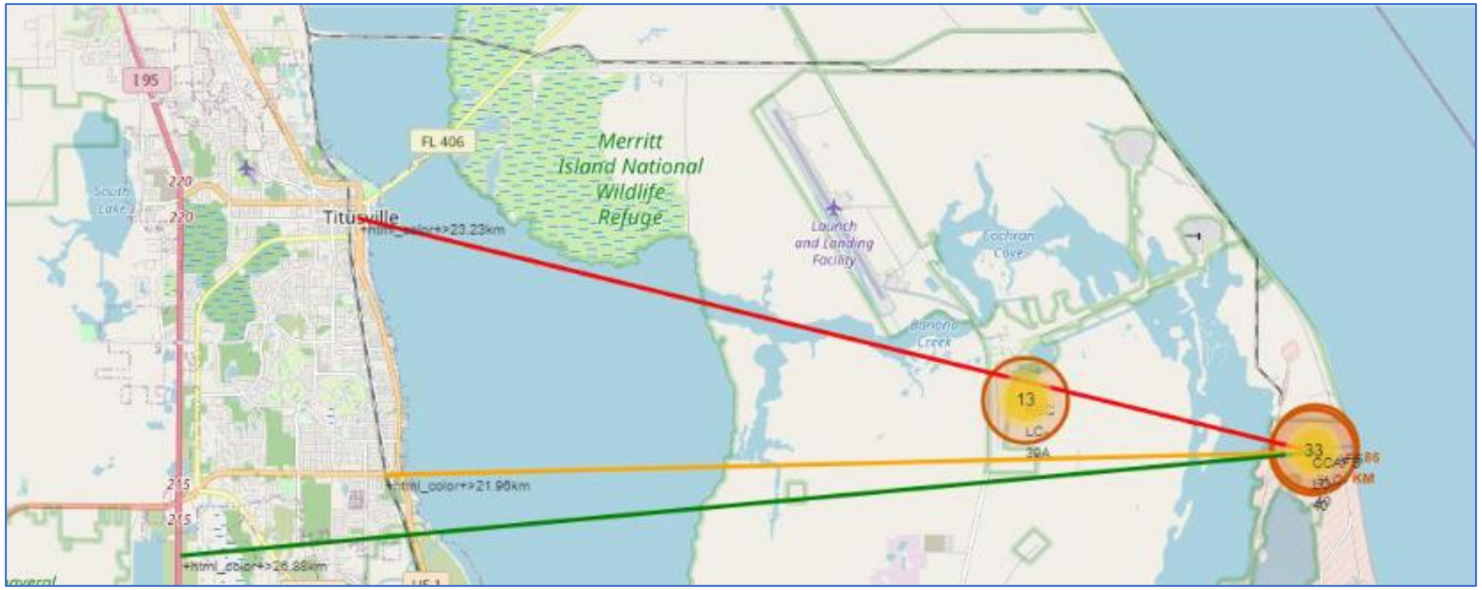
- 10 West coast
- 46 East coast
- All near to Equator Line

Folium Map

- Green – Successful launches
- Red – Unsuccessful launches

Folium Map

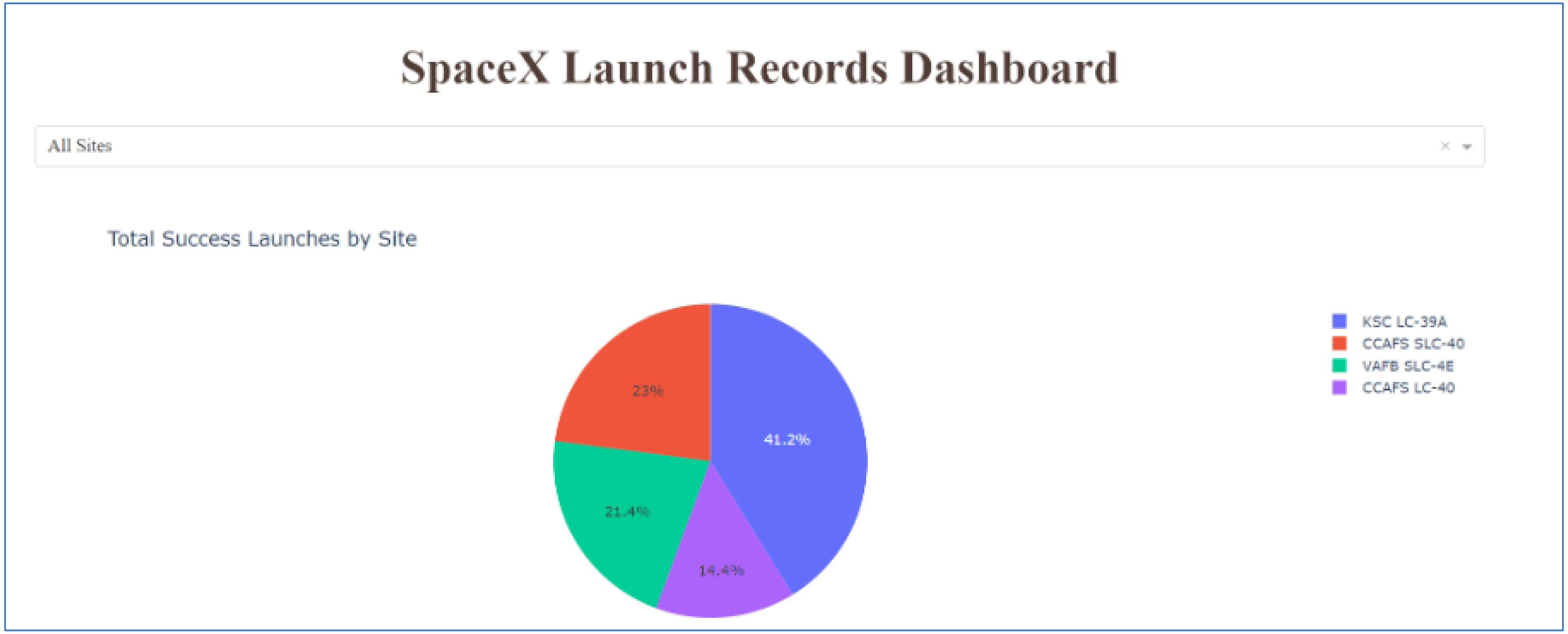
- Away from busy zones



PÚBLICA

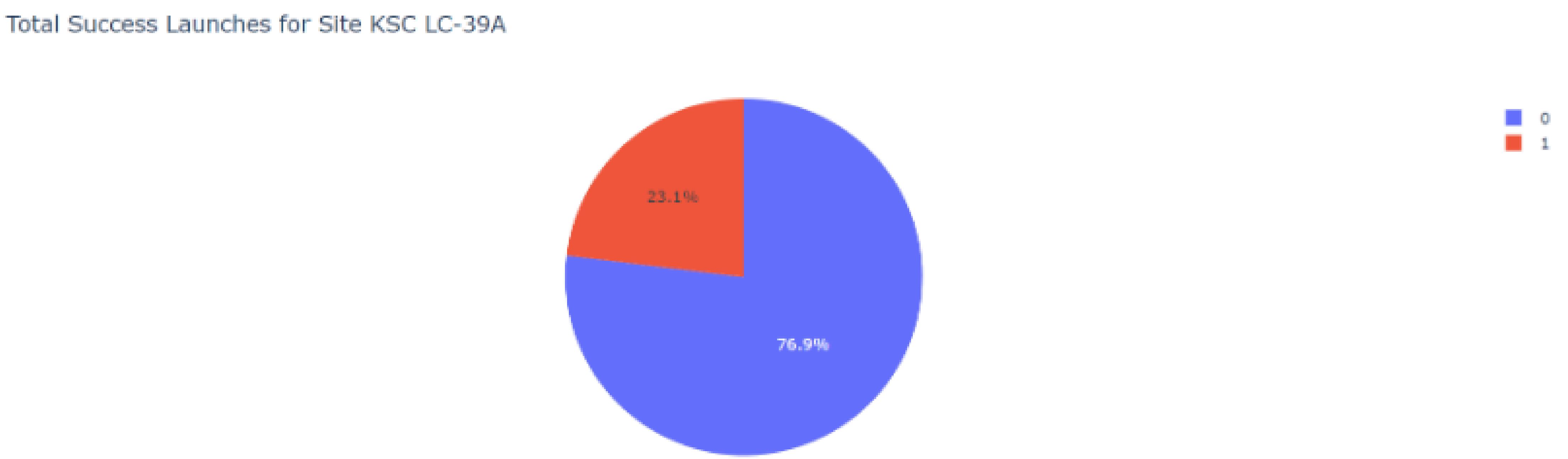
Section 4

Build a Dashboard with Plotly Dash



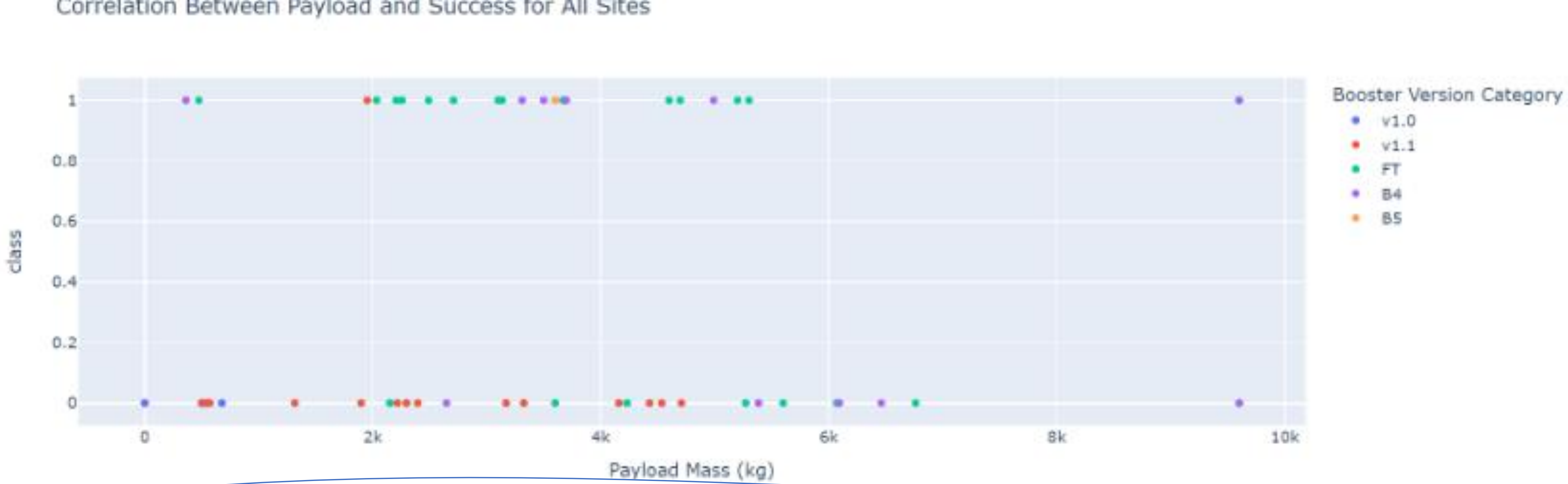
Dashboard - Total

- KSC LC-39A has the biggest success rate among all the launch sites



Dashboard – KSC LC-29A

- Highest success rate among other launch sites



Dashboard Payload and Success

- 1 = Success
- 0 = Fail
- Without a very clear trend defined

Section 5

Predictive Analysis (Classification)

TASK 8

Create a decision tree classifier object then create a `GridSearchCV` object `tree_cv` with cv = 10. Fit the object to find the best parameters from the dictionary `parameters`.

```
parameters = {'criterion': ['gini', 'entropy'],
     'splitter': ['best', 'random'],
     'max_depth': [2*n for n in range(1,10)],
     'max_features': ['auto', 'sqrt'],
     'min_samples_leaf': [1, 2, 4],
     'min_samples_split': [2, 5, 10]}

tree = DecisionTreeClassifier()
tree_cv = GridSearchCV(estimator=tree, cv=10, param_grid=parameters).fit(X_train, Y_train)
```

```
print("tuned hpyerparameters :(best parameters) ",tree_cv.best_params_)
print("accuracy :",tree_cv.best_score_)
```

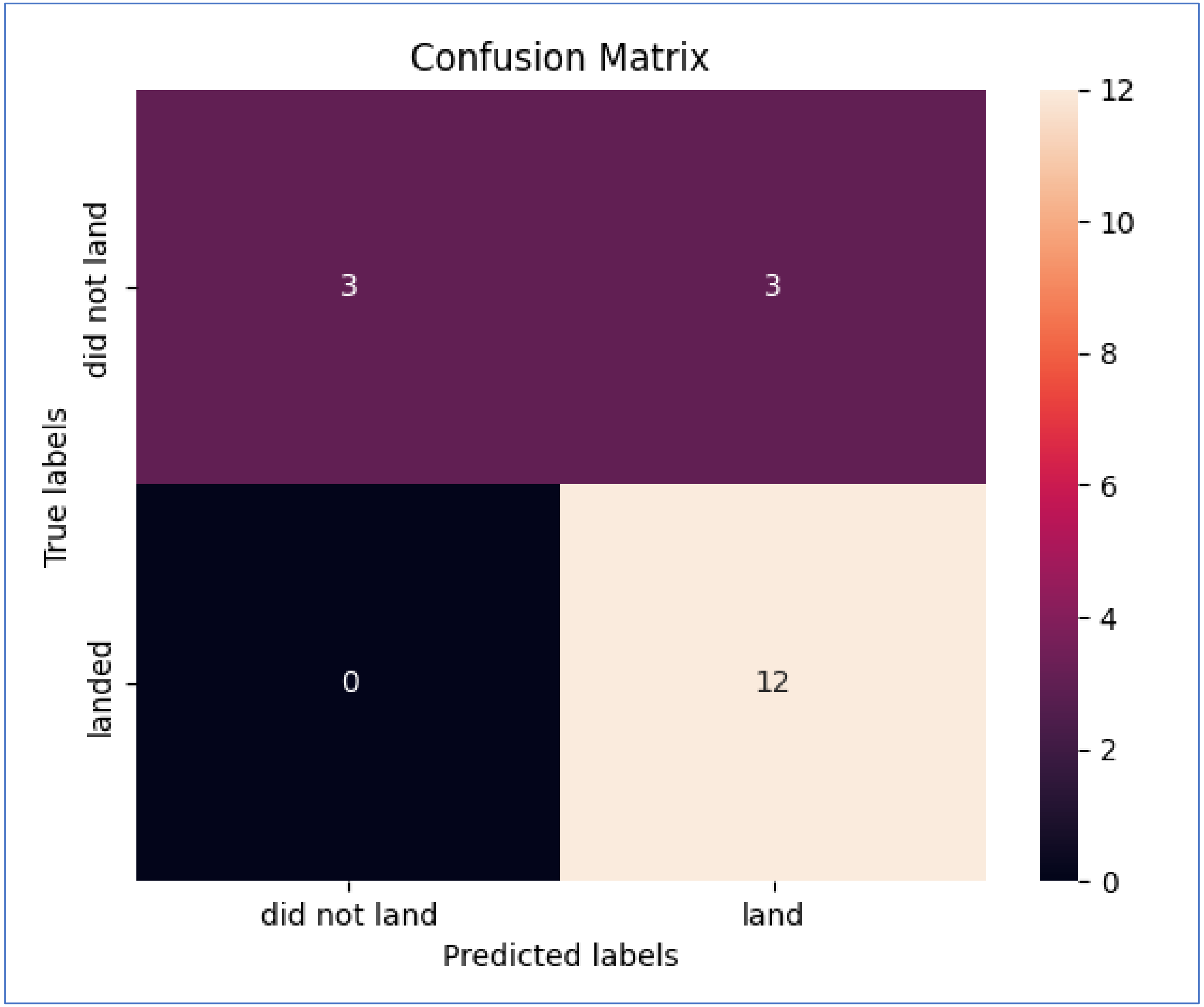
```
tuned hpyerparameters :(best parameters)  {'criterion': 'gini', 'max_depth': 18, 'max_features': 'sqrt', 'min_samples_leaf': 2, 'min_samples_split': 2, 'splitter': 'best'}
accuracy : 0.9017857142857144
```

	LogReg	SVM	Tree	KNN
Jaccard_Score	0.800000	0.800000	0.800000	0.800000
F1_Score	0.888889	0.888889	0.888889	0.888889
Accuracy	0.833333	0.833333	0.833333	0.833333

Classification Accuracy

- Visualization of the built model accuracy for all built classification models
- The Decision Tree model performed the best accuracy rate ~ 90%

Confusion Matrix



Conclusions

- The Decision Tree model performed the best accuracy rate ~ 90%
- Location – Near to Equator Line and close to coast
- The success rate has been increasing over the years, which indicates a clear evolution, largely due to the lessons learned over the years
- KSC LC-29A has the most impressive success rate
- Would be interesting to have a learning curve vs the success rate, I believe this would show the clear evolution in the success rate over the years with descriptive engineering improvements

Appendix

- https://github.com/fredpessoa/IBM-Capstone

Thank you!