IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Bogdanis Ioannis - Ioulianos

15th May 2022

Outline

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

Executive Summary

Purpose of Report

Can SpaceY bid against SpaceX for a rocket launch?

The purposes of this report are as follows:

- predict if the Falcon 9 first stage will land successfully,
- determine the cost of a launch.

Methods Used

i. Collect and prepare the data from an API
ii. Perform web scraping to collect Falcon 9 historical launch records
iii. Perform some Exploratory Data Analysis (EDA) to find some patterns in the data
iv. Determine what would be the label for training supervised models.
v. Finally build a Dash application for users to perform interactive visual analytics on SpaceX launch data in real-time.

Findings

Falcon9 FT
Rocket

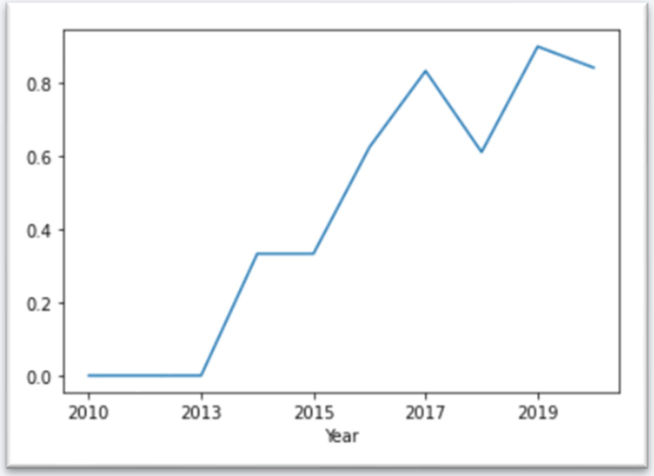
- **90 total flights**
- **60 successful**
- **Increasing Success Rate**

KFS LC-39A
76.9%
Success

- **Best Launch Site KFS**
- **Located at East Coast**
- **Offering best overall conditions**

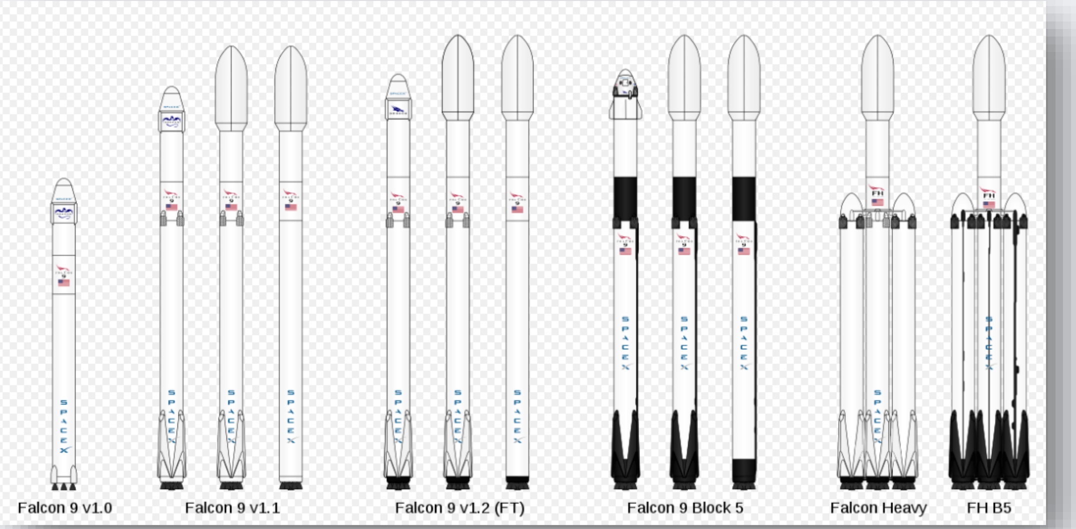
91.4%
Prediction Model
Accuracy

- **Machine Learning**
- **Hyperparameters**
- **Decision Tree Classifier**



Introduction

- SpaceX's cost of launch is always lower than the competition: 62 million dollars compered to 165 million dollars each
- Over 10 years of data SpaceX's accomplished:
  - Sending spacecraft to the International Space Station.
  - Starlink, a satellite internet constellation providing satellite Internet access.
  - Sending manned missions to Space
- Compared to the competition SpaceX's has achieved greatest rates of rocket reusability.



Introduction

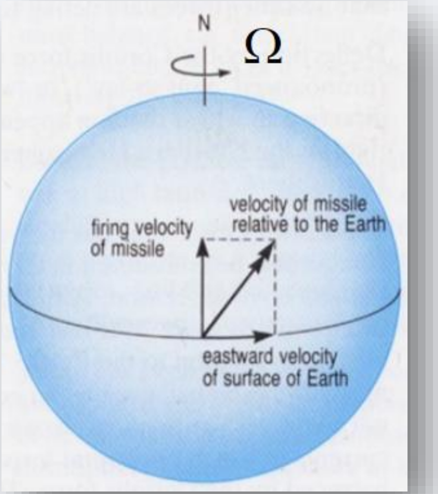
Some Physics

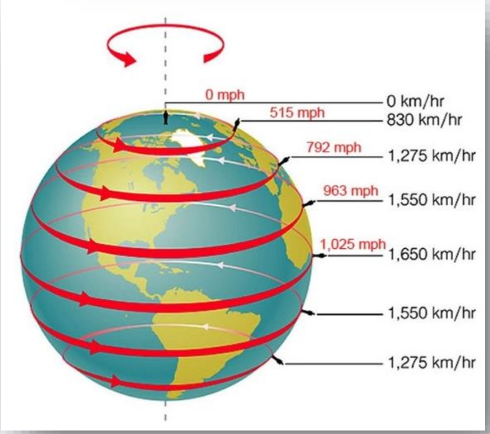
What affects a rocket launch?

- **Coriolis effect:** makes things (like planes or currents of air) traveling long distances around the Earth appear to move at a curve as opposed to a straight line

- **Spin of Earth:** Earth's rotation or Earth's spin is the rotation of planet Earth around its own axis, as well as changes in the orientation of the rotation axis in space. Earth rotates eastward, in prograde motion. As viewed from the north pole star Polaris, Earth turns counterclockwise.

- **Locations latitude:** Latitude is a geographic coordinate that specifies the north–south position of a point on the Earth's surface. Earth's velocity is relative to earth's radius, The equator has the bigger radius, so the velocity is the highest. The closer to the equator the greater the surface velocity.





Upper image: https://slideplayer.com/slide/5075634/ (slide 3)
Lower image: https://sciengsustainability.blogspot.com/2013/04/faster-than-sound-our-velocity-due-to.html
Theory from Wiki: https://en.wikipedia.org/wiki/Latitude , https://en.wikipedia.org/wiki/Coriolis_force , https://en.wikipedia.org/wiki/Earth%27s_rotation

Section 1

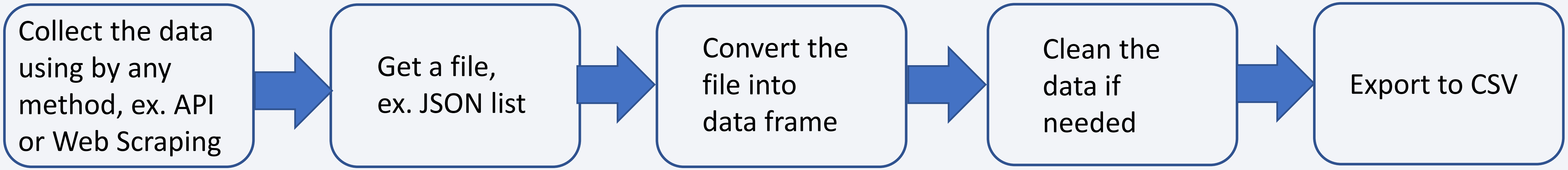
Methodology

Methodology

Data Collection	• SpaceX REST API • Web Scraping from Wikipedia
Data Wrangling	• Transform and Map the Data • Apply OneHotEncoder
Exploratory Data Analysis	• Using SQL • Using Visualization
Interactive Visual Analytics & Dashboard	• Using Plotly Dash for building a Dashboard • Using Folium for Interactive Map
Predictive Analysis	• Train the Model • Perform Grid Search • Build a Machine Learning Pipeline

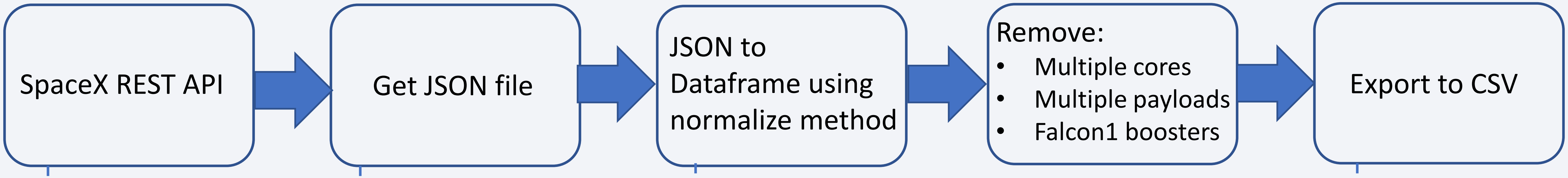
Data Collection

- Data was collected
  - Via SpaceX REST API
  - Via Web Scraping a Wiki page.



Data Collection – SpaceX API

USE: SpaceX REST API or URL: https://api.spacexdata.com/v4/




```
data = pd.json_normalize(response.json())
```

```
data_falcon9.to_csv('dataset_part_1.csv', index=False)
```

```
response = requests.get(spacex_url)
```

```
spacex_url="https://api.spacexdata.com/v4/launches/past"
```

	FlightNumber	Date	BoosterVersion	PayloadMass	Orbit	LaunchSite	Outcome	Flights	GridFins	Reused	Legs	LandingPad	Block	Reus
4	1	2010-06-04	Falcon 9	6123.547647	LEO	CCSFS SLC 40	None None	1	False	False	False	None	1.0	
5	2	2012-05-22	Falcon 9	525.000000	LEO	CCSFS SLC 40	None None	1	False	False	False	None	1.0	
6	3	2013-03-01	Falcon 9	677.000000	ISS	CCSFS SLC 40	None None	1	False	False	False	None	1.0	
7	4	2013-09-29	Falcon 9	500.000000	PO	VAFB SLC 4E	False Ocean	1	False	False	False	None	1.0	
8	5	2013-12-03	Falcon 9	3170.000000	GTO	CCSFS SLC 40	None None	1	False	False	False	None	1.0	

https://github.com/ibogdanis/testrepo/blob/6637bcbd9e2faf51d754cdd2d825378c9d8efe90/Final_Assignment.ipynb

Data Collection - Scraping

USE: Web Scraping Wiki pages with BeautifulSoup

Get a response from HTML page → Create BeautifulSoup object → Extract tables → Create a Dataframe by parsing the HTML tables → Export to CSV

```
static_url = "https://en.wikipedia.org/w/index.php?title=List_of_Falcon_9_and_Falcon_Heavy_launches&oldid=1027686922"
response = requests.get(static_url)
html = response.content
```

```
soup = BeautifulSoup(html, 'html5lib')
```

```
html_tables = soup.find_all('table')
```

```
launch_dict= dict.fromkeys(column_names)
df=pd.DataFrame(launch_dict)
```

```
df.to_csv('spacex_web_scraped.csv', index=False)
```

	Flight No.	Launch site	Payload	Payload mass	Orbit	Customer	Launch outcome	Version Booster	Booster landing	Date	Time
0	1	CCAFS	Dragon Spacecraft Qualification Unit	0	LEO	SpaceX	Success\n	F9 v1.0B0003.1	Failure	4 June 2010	18:45
1	2	CCAFS	Dragon	0	LEO	NASA	Success	F9 v1.0B0004.1	Failure	8 December 2010	15:43
2	3	CCAFS	Dragon	525 kg	LEO	NASA	Success	F9 v1.0B0005.1	No attempt\n	22 May 2012	07:44
3	4	CCAFS	SpaceX CRS-1	4,700 kg	LEO	NASA	Success\n	F9 v1.0B0006.1	No attempt	8 October 2012	00:35
4	5	CCAFS	SpaceX CRS-2	4,877 kg	LEO	NASA	Success\n	F9 v1.0B0007.1	No attempt\n	1 March 2013	15:10

https://github.com/ibogdanis/testrepo/blob/6637bcbd9e2faf51d754cdd2d825378c9d8efe90/Final_Assignment_v2.ipynb

Data Wrangling

Data wrangling is the process of cleaning and unifying messy and complex data sets for easy access and analysis.

We will perform:

- some Exploratory Data Analysis (EDA) to find some patterns in the data and
- determine what would be the label for training supervised models.

Load SpaceX dataset → Calculate: Launches of each site, Occurrences of each orbit, Occurrences of mission outcome per orbit. → Create a set of bad landing outcomes → Create a landing outcome label → Export to CSV

```python
df.to_csv("dataset_part_2.csv", index=False)
```

```python
bad_outcomes=set(landing_outcomes.keys()[[1,3,5,6,7]])
```

```python
landing_class = []

for i in df['Outcome']:
    if i in set(bad_outcomes):
        landing_class.append(0)
    else:
        landing_class.append(1)
```

```python
df.LaunchSite.value_counts()
df.Orbit.value_counts()
landing_outcomes = df.Outcome.value_counts()
```



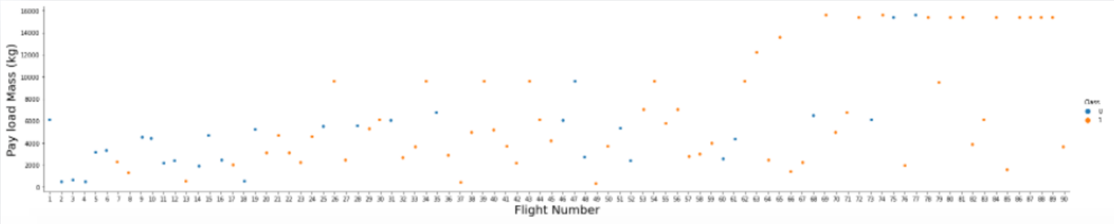
```python
df=pd.read_csv("https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DS0321EN-SkillsNetwork/datasets/dataset_part_1.csv")
```

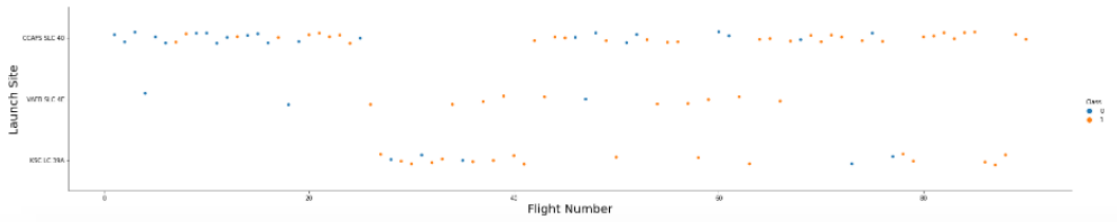
https://github.com/ibogdanis/testrepo/blob/6637bcbd9e2faf51d754cdd2d825378c9d8efe90/Final_Assignment_v3.ipynb

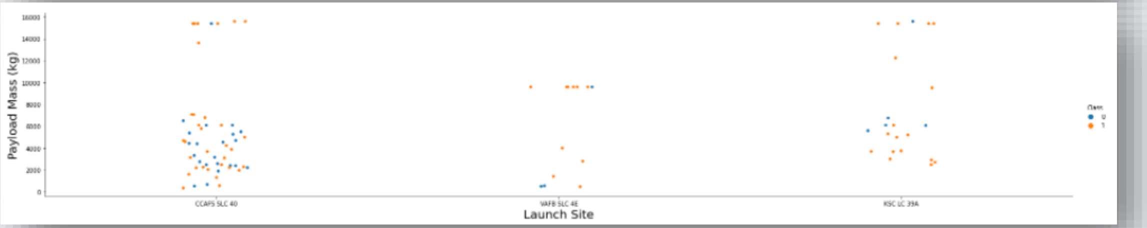
EDA with Data Visualization

Catplot: in order to determine the relationship between:

- Flight Number
- Launch Site
- Payload Mass

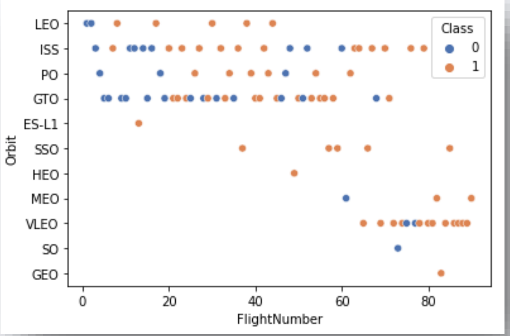


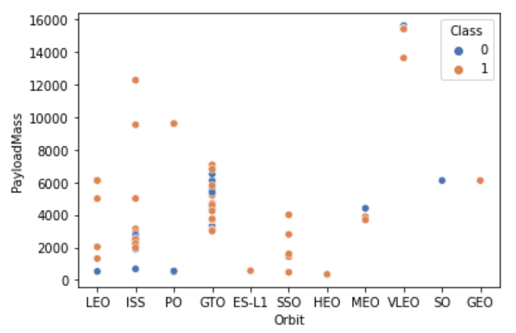




Scatterplot: in order to determine the relationship between:

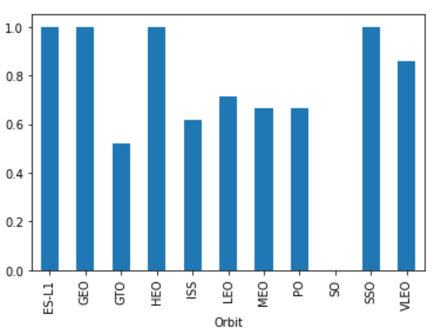
- Flight Number
- Orbit
- Payload Mass

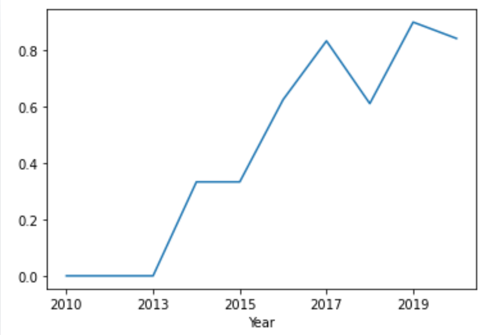




Bar Chart for the success rate of each orbit

Line Chart for the yearly launch success





https://github.com/ibogdanis/testrepo/blob/6637bcbd9e2faf51d754cdd2d825378c9d8efe90/Complete%20the%20EDA%20with%20Visualization%20lab.ipynb

EDA with SQL

Summary of SQL queries that were used:

- Display the names of the unique launch sites in the space mission,
- Display 5 records where launch sites begin with the string 'CCA',
- Compare the payload mass with boosters launched by NASA (CRS),
- Display average payload mass carried by booster version F9 v1.1,
- Listing the date where the successful landing outcome in drone ship was achieved,
- Listing the names of the boosters which have success in ground pad and have payload mass greater than 4000 but less than 6000,
- List the total number of successful and failure mission outcomes,
- Listing the names of the booster versions which have carried the maximum payload mass,
- Listing the failed landing outcomes in drone ship, their booster version and the launch site names for the year 2015,
- Ranking the count of landing outcomes between the date 2010-06-04 and 2017-03-20, in descending order.

https://github.com/ibogdanis/testrepo/blob/6637bcbd9e2faf51d754cdd2d825378c9d8efe90/Final_Assignment_Complete%20the%20EDA%20with%20SQL%20lab.ipynb

Build an Interactive Map with Folium

We will use Folium to build an interactive map and perform interactive visual analytics.

- Firstly we will be focused on analyzing launch site geo and proximities with Folium.
  - We will mark the launch site locations and their close proximities on an interactive map.
  - Then, we can explore the map with those markers and try to discover any patterns from them.
- Finally, we should be able to explain **how to choose an optimal launch site**.

Map object/function	Reason
folium.Marker()	To create launch sites locations with coordinates
folium.Circle()	To add a highlighted circle area
folium.Polyline()	To draw a line between two points
MousePosition()	To get coordinate for a mouse over a point on the map
MarkerCluster()	To simplify a map containing many markers having the same coordinate
folium.Map()	To define the map
DivIcon()	To create an icon as a text label

https://github.com/ibogdanis/testrepo/blob/6637bcbd9e2faf51d754cdd2d825378c9d8efe90/Interactive%20Visual%20Analytics%20with%20Folium%20lab.ipynb

Build a Dashboard with Plotly Dash

We will use Plotly Dash to build a dashboard to perform interactive visual analytics.

- We will build a dashboard application with the Python Plotly Dash package containing input components such as
  - a dropdown list,
  - a range slider to interact with a pie chart,
  - a pie chart and a scatter point chart.
- After the dashboard is built, we can use it to find more insights from the SpaceX dataset more easily than with static graphs.

Object/Function	Reason
dcc.Dropdown()	Create a dropdown menu for launch sites
dcc.RangeSlider()	Create a range slider for payload mass
px.pie()	Create a pie chart for success percentage
px.scatter()	Create a scatter chart for correlation

https://github.com/ibogdanis/testrepo/blob/6637bcbd9e2faf51d754cdd2d825378c9d8efe90/dash_interactivity.py

Predictive Analysis (Classification)

Load the data

```
data = pd.read_csv("https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DS0321EN-SkillsNetwork/datasets/dataset_part_2.csv")
X = pd.read_csv('https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DS0321EN-SkillsNetwork/datasets/dataset_part_3.csv')
```

Prepare the data

```
Y = data["Class"].to_numpy()
transform = preprocessing.StandardScaler()
X = transform.fit_transform(X)
```

Set training and test data

```
X_train, X_test, Y_train, Y_test = train_test_split( X, Y, test_size=0.2, random_state=2)
```

Try different classification models

```
lr=LogisticRegression()
svm = SVC()
tree = DecisionTreeClassifier()
KNN = KNeighborsClassifier()
```

```
logreg_cv = GridSearchCV(lr, parameters, cv=10).fit(X_train, Y_train)
svm_cv = GridSearchCV(svm, parameters, cv=10).fit(X_train, Y_train)
tree_cv = GridSearchCV(tree, parameters, cv=10).fit(X_train, Y_train)
knn_cv = GridSearchCV(KNN, parameters, cv=10).fit(X_train, Y_train)
```

Most accurate model

```
Accuracy for Logistics Regression (LR): 0.8464285714285713
Accuracy for Support Vector Machine (SVM): 0.8482142857142856
Accuracy for Decision Tree (DT): 0.9142857142857144
Accuracy for K nearsdt neighbors (KNN): 0.8482142857142858
```

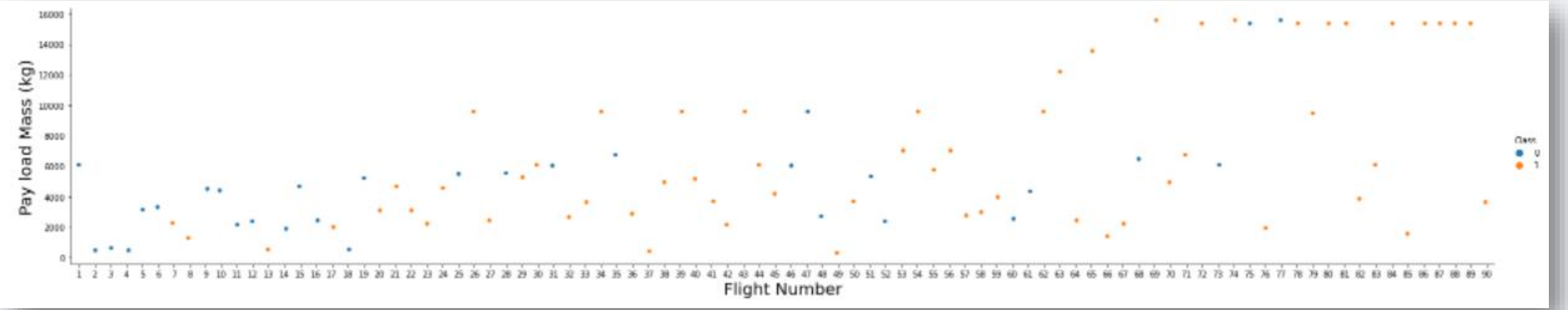

https://github.com/ibogdanis/testrepo/blob/6637bcbd9e2faf51d754cdd2d825378c9d8efe90/Machine%20Learning%20Prediction%20lab.ipynb

Section 2

Insights drawn from EDA

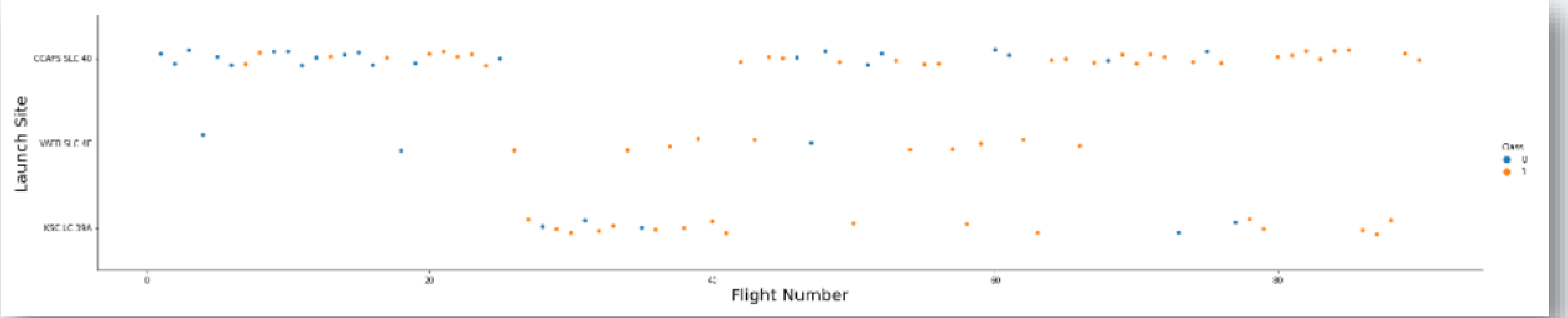
Flight Number vs. Payload Mass

- The higher the flight number, the higher the success landing rate
- The higher the flight number, the greater the size of the payload attempting to launch



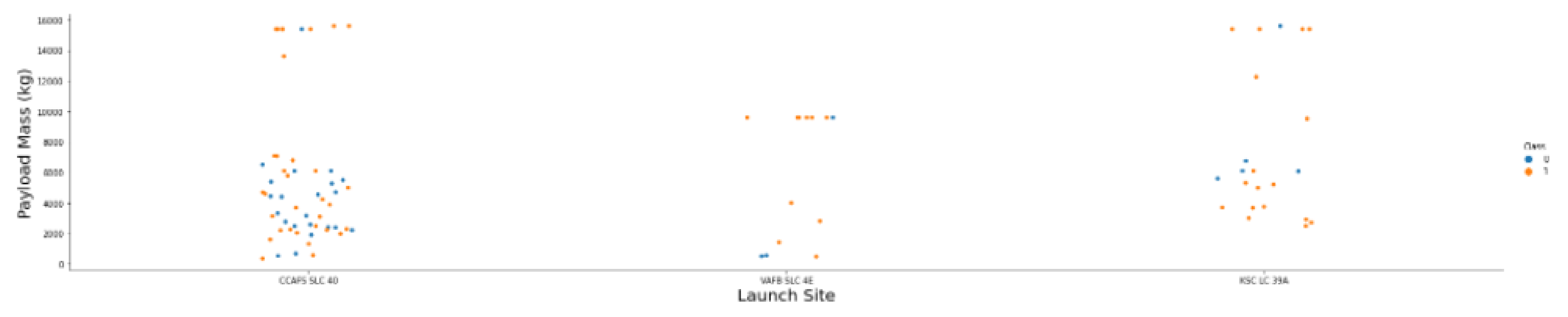
Flight Number vs. Launch Site

- The most launch attempts have been made from CCAFS,
- The CCAFS has also the highest number of successful landings
- But the most successful sites are KSC and VAFB with a 77% success rate but with fewer attempts



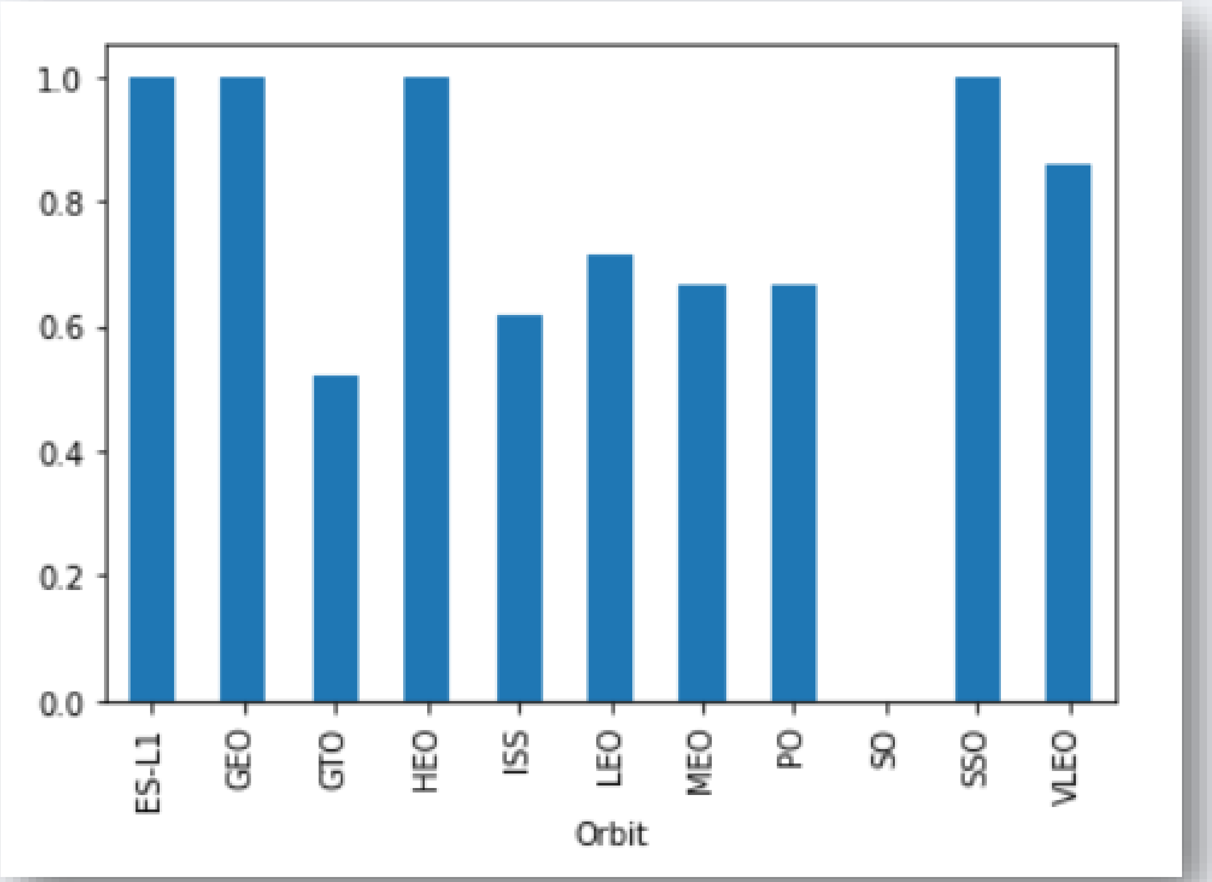
Payload vs. Launch Site

- The maximum payload attempts were performed from CCAFS and KSC, over 15000kg,
- From VAFB the payload does not exceed the 10000 kg,
- The VAFB launch site has also the fewer failed landings.



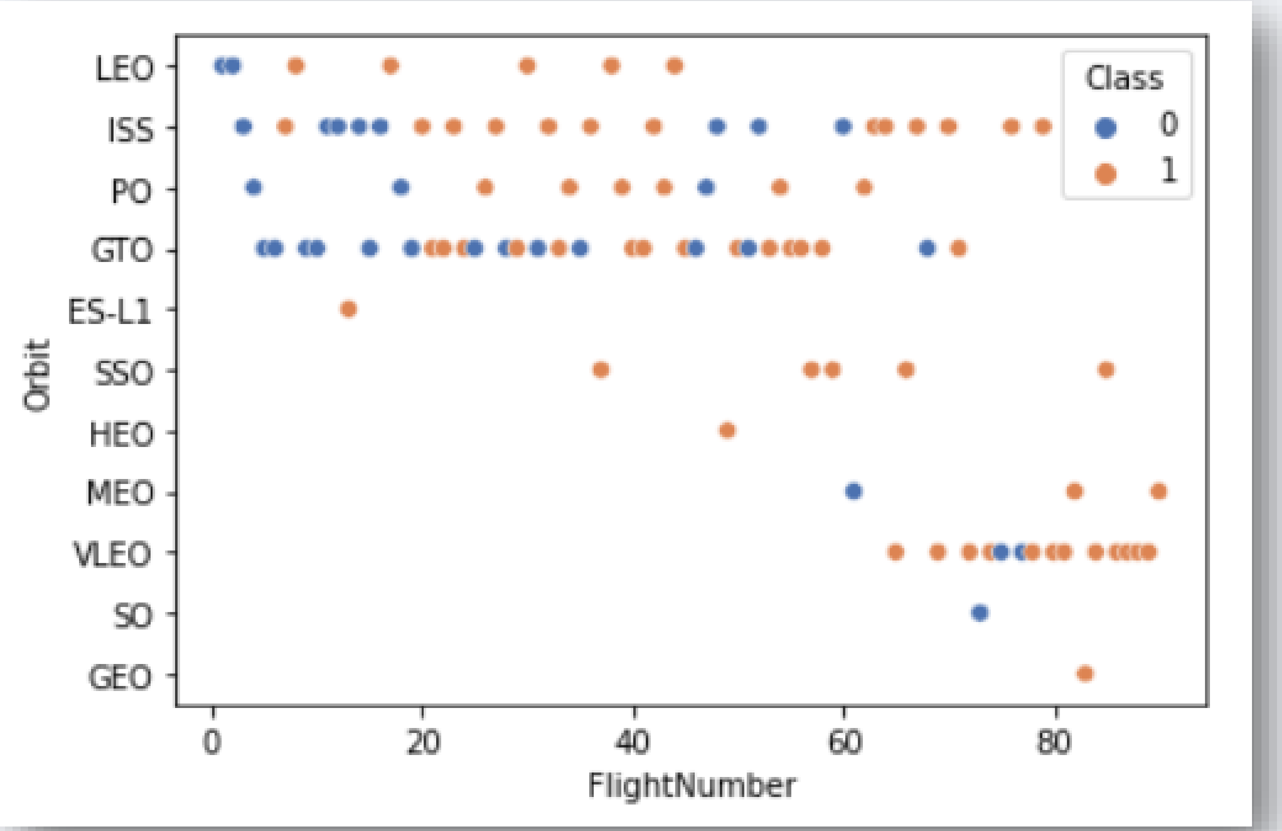
Success Rate vs. Orbit Type

- The orbits with the higher success rate: SSO, ES-L1, HEO, GEO



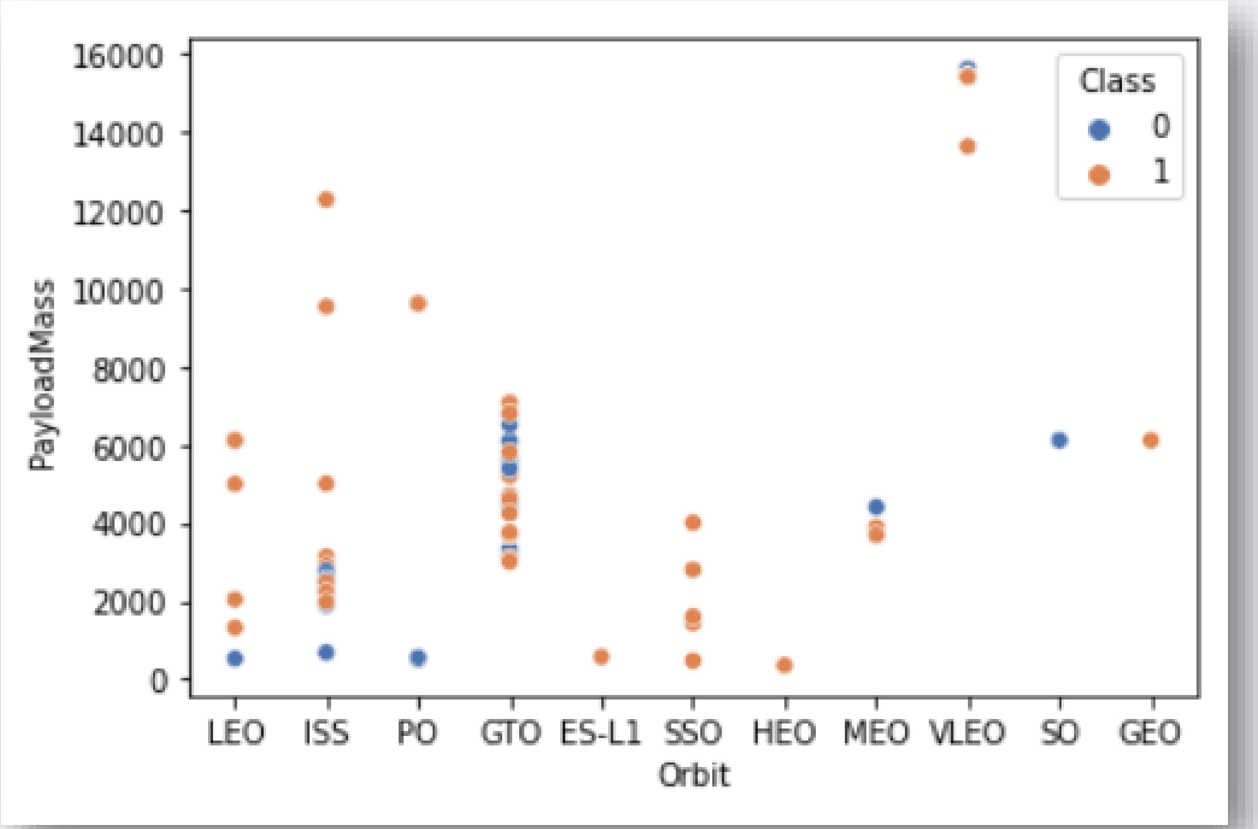
Flight Number vs. Orbit Type

- The orbits with higher attempts are: ISS, GTO, VLEO
- The higher the flight number, the more different orbits are attempted



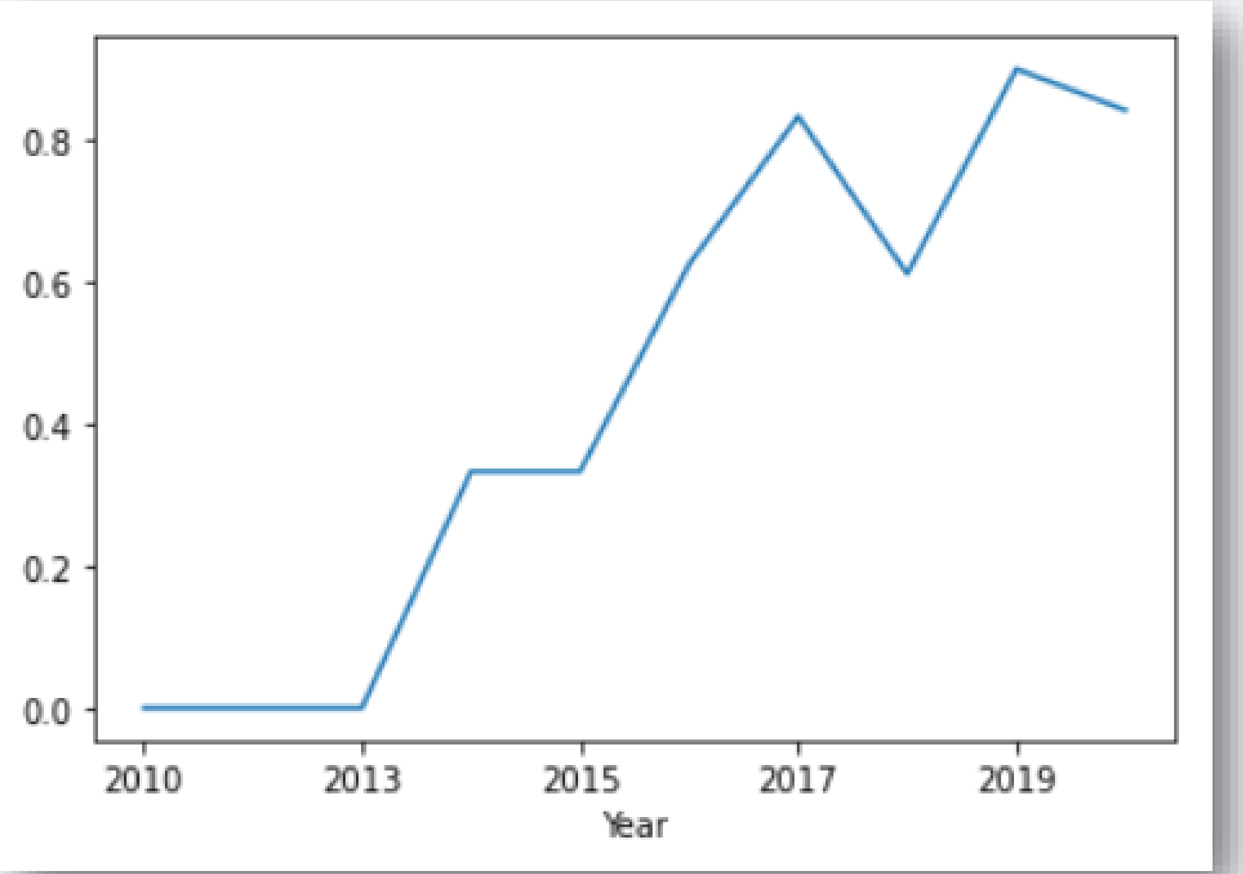
Payload vs. Orbit Type

- The orbits with the highest Payload Mass are : ISS, PO, VLEO



Launch Success Yearly Trend

- We can observe that the success rate since 2013 kept increasing till 2020



All Launch Site Names

- SQL Query : *Display the names of the unique launch sites in the space mission*

```
%%sql
SELECT distinct launch_site FROM SPACEXTBL;
```

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

Launch Site Names Begin with 'CCA'

- SQL Query : *Display 5 records where launch sites begin with the string 'CCA'*

```
%%sql

SELECT * FROM SPACEXTBL WHERE launch_site LIKE 'CCA%' LIMIT 5;
```

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	07:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525	LEO (ISS)	NASA (COTS)	Success	No attempt
2012-10-08	00:35:00	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

Total Payload Mass

- SQL Query : *Display the total payload mass carried by boosters launched by NASA (CRS)*

```
%%sql

SELECT SUM (payload_mass__kg_) FROM SPACEXTBL WHERE customer LIKE ('NASA (CRS)')
```

1
45596

Average Payload Mass by F9 v1.1

- SQL Query : *Display average payload mass carried by booster version F9 v1.1*

```
%%sql

SELECT AVG(payload_mass__kg_) FROM SPACEXTBL WHERE booster_version LIKE ('F9 v1.1%')
```

1
2534

First Successful Ground Landing Date

- SQL Query : *List the date when the first successful landing outcome in ground pad was acheived*

```
%%sql

SELECT MIN(Date) FROM SPACEXTBL WHERE landing__outcome LIKE ('Success (ground pad)')
```

1
2015-12-22

Successful Drone Ship Landing with Payload between 4000 and 6000

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

```
%%sql

SELECT distinct booster_version
FROM SPACEXTBL
WHERE landing__outcome LIKE ('Success (drone ship)') AND payload_mass__kg_ > 4000 AND payload_mass__kg_ < 6000;
```

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

Total Number of Successful and Failure Mission Outcomes

- SQL Query : *List the total number of successful and failure mission outcomes*

```
%%sql

SELECT mission_outcome, COUNT(mission_outcome) FROM SPACEXTBL GROUP BY (mission_outcome)
```

mission_outcome	2
Failure (in flight)	1
Success	99
Success (payload status unclear)	1

Boosters Carried Maximum Payload

- SQL Query : *List the names of the booster_versions which have carried the maximum payload mass*

```
%%sql

SELECT booster_version, payload_mass__kg_ FROM SPACEXTBL WHERE payload_mass__kg_ = (SELECT MAX(payload_mass__kg_) FROM SPACEXTB
L)
```

booster_version	payload_mass__kg_
F9 B5 B1048.4	15600
F9 B5 B1049.4	15600
F9 B5 B1051.3	15600
F9 B5 B1056.4	15600
F9 B5 B1048.5	15600
F9 B5 B1051.4	15600
F9 B5 B1049.5	15600
F9 B5 B1060.2	15600
F9 B5 B1058.3	15600
F9 B5 B1051.6	15600
F9 B5 B1060.3	15600
F9 B5 B1049.7	15600

2015 Launch Records

- SQL Query : *List the failed landing_outcomes in drone ship, their booster versions, and launch site names for in year 2015*

```
%%sql
SELECT landing__outcome, booster_version, launch_site, date FROM SPACEXTBL WHERE landing__outcome LIKE ('Failure (drone ship)')
AND year(Date) = 2015;
```

landing__outcome	booster_version	launch_site	DATE
Failure (drone ship)	F9 v1.1 B1012	CCAFS LC-40	2015-01-10
Failure (drone ship)	F9 v1.1 B1015	CCAFS LC-40	2015-04-14

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

- SQL Query : *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*

```
%%sql

SELECT landing__outcome, COUNT(landing__outcome)
FROM SPACEXTBL WHERE date BETWEEN '2010-06-04' AND '2017-03-20'
GROUP BY (landing__outcome) ORDER BY COUNT(landing__outcome) DESC
```

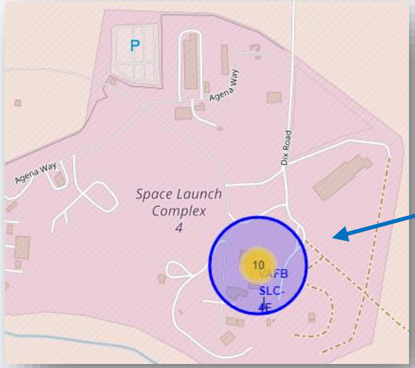
landing__outcome	2
No attempt	10
Failure (drone ship)	5
Success (drone ship)	5
Controlled (ocean)	3
Success (ground pad)	3
Failure (parachute)	2
Uncontrolled (ocean)	2
Precluded (drone ship)	1

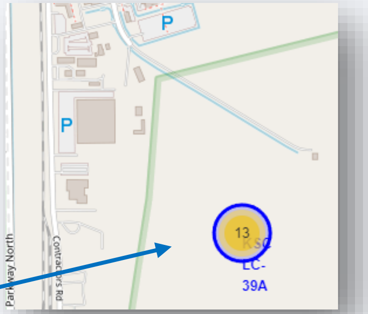
Section 3

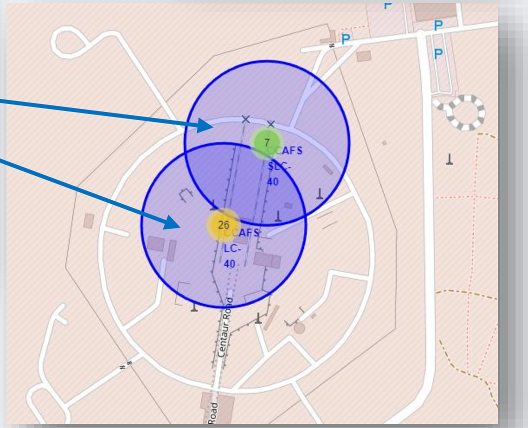
Launch Sites Proximities Analysis

All Launch Sites with Folium Map

- We can observe that the most launch sites are located at the east coast
- The east sites are 6 degrees closer to equator, which means higher velocity and lower pressure





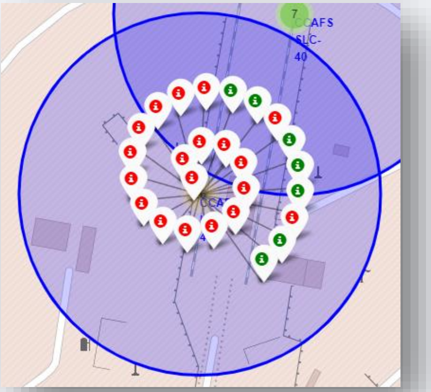


	Launch Site	Lat	Long
0	CCAFS LC-40	28.562302	-80.577356
1	CCAFS SLC-40	28.563197	-80.576820
2	KSC LC-39A	28.573255	-80.646895
3	VAFB SLC-4E	34.632834	-120.610745

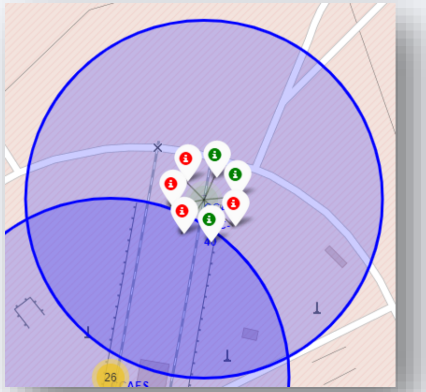
All Launch Site Records Color Labeled

- The most successful launch site is the KSC LC-39A
- Green label : Successful launches
- Red label : Failed launches

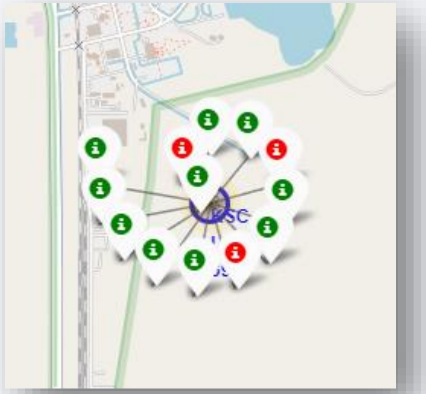
CCAFS LC-40



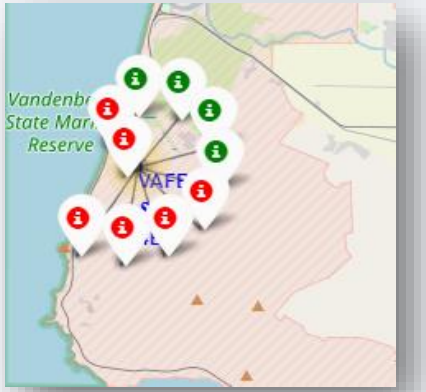
CCAFS SLC-40



KSC LC-39A



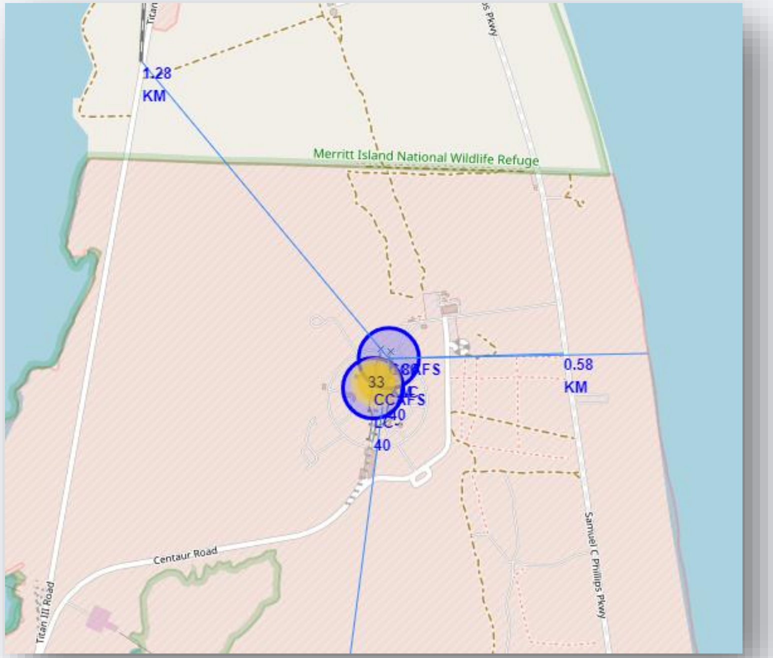
VAFB SLC-4E

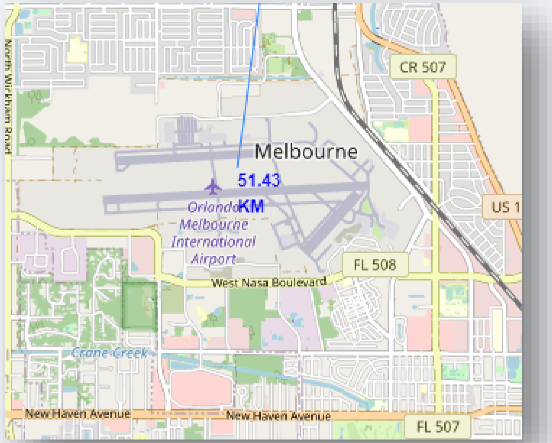


All significant points distances from launch sites

- We observe that:
  - Launch Sites are close to railways and highways, less than 10 km, due to easiest access allowing for better logistics overall.
  - Launch Sites are close to coastlines because the launch attempts are made towards the sea due to safety reasons,
  - And finally the launch sites are far from cities, more than 30 km, for safety reasons also.

```
distance highway to CCAFS SLC 40 = 0.5834695366934144  km
distance highway to CCAFS LC 40 = 0.6461880763256931  km
distance highway to KSC LC 39A = 0.8427976971406239  km
distance highway to VAFB SLC 4E = 5.761773850878405  km
distance railway to CCAFS SLC 40 = 1.2845344718142522  km
distance railway to CCAFS LC 40 = 1.3314378832303038  km
distance railway to KSC LC 39A = 0.840130287647207  km
distance railway to VAFB SLC 4E = 1.2627243006445117  km
distance city to CCAFS SLC 40 = 51.43416999517233  km
distance city to CCAFS LC 40 = 51.32864612692812  km
distance city to KSC LC 39A = 52.11418862316215  km
distance city to VAFB SLC 4E = 38.5014605150241  km
```



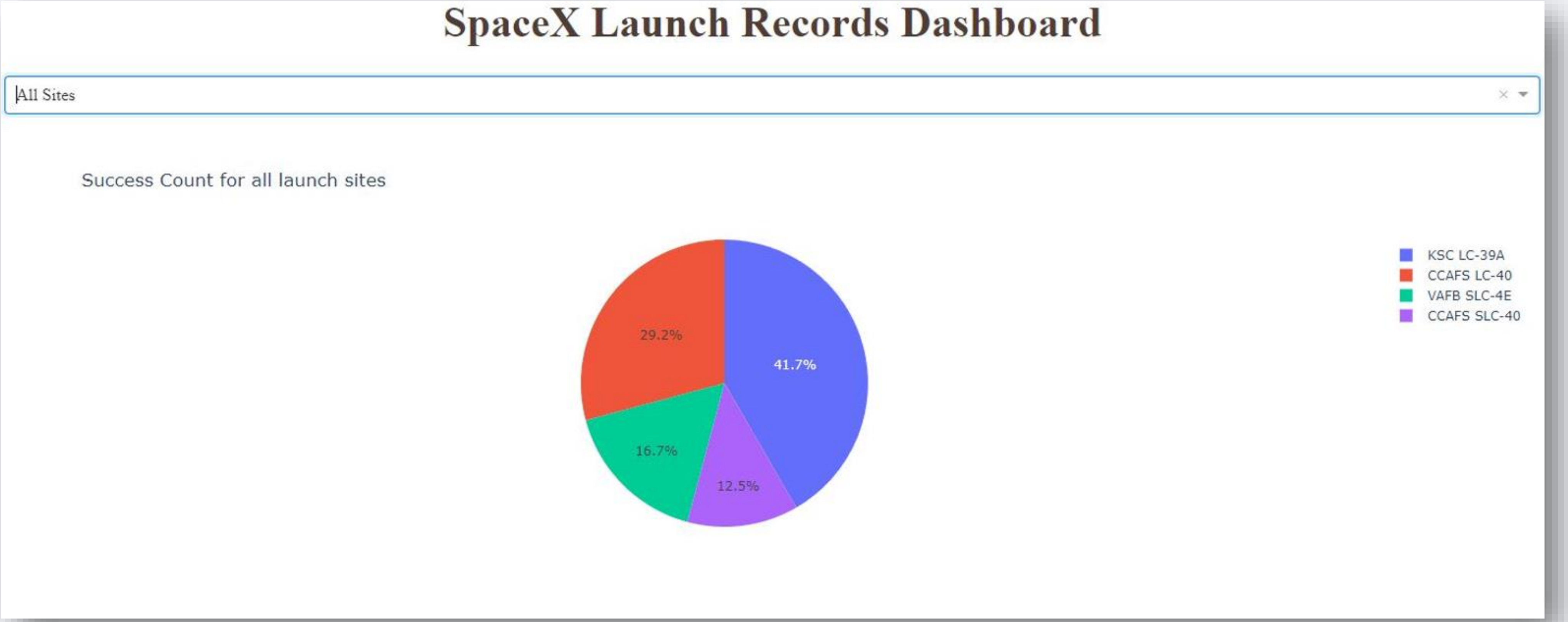


Section 4

Build a Dashboard with Plotly Dash

Launch Success count for all sites

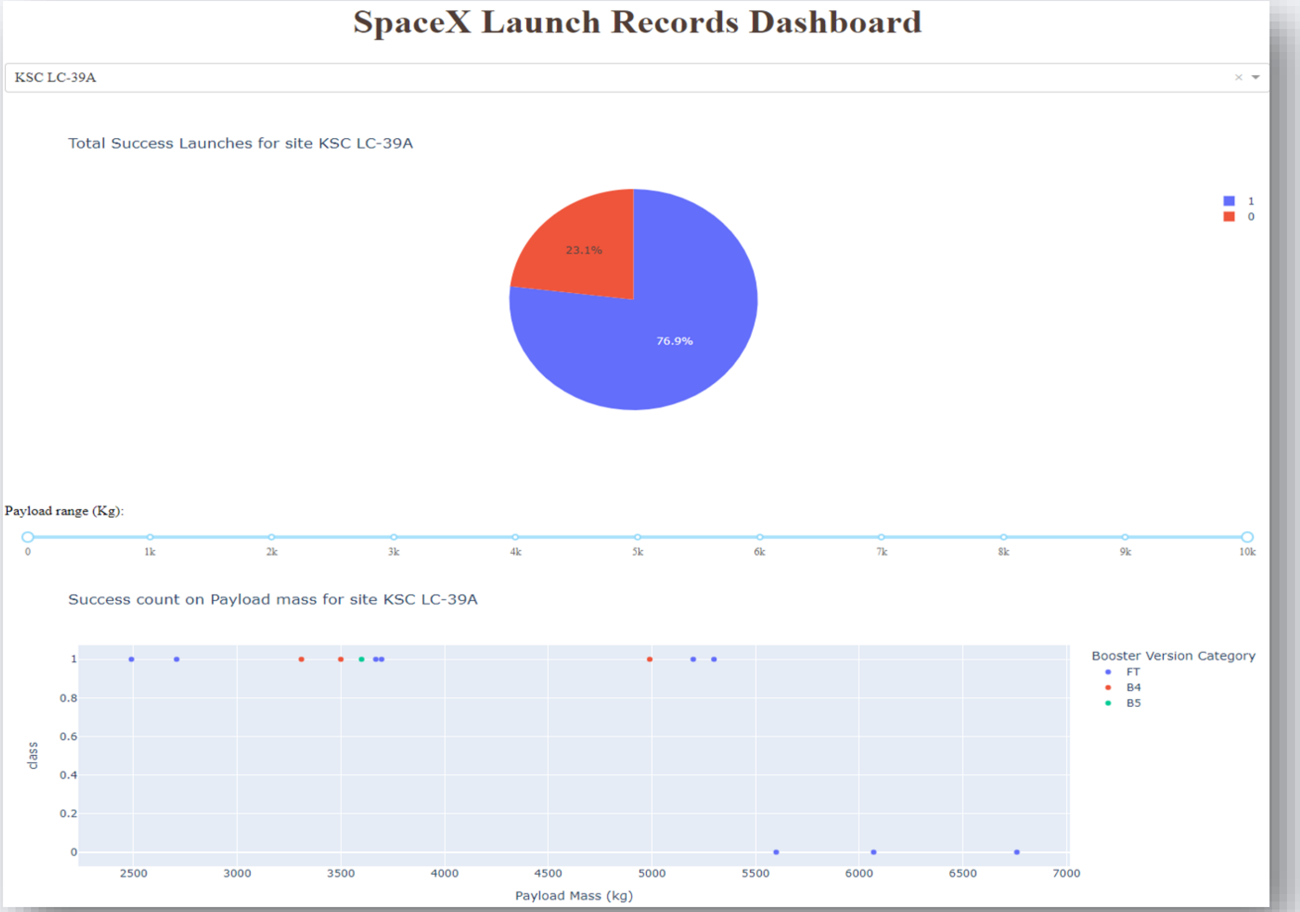
- Dashboard showing the success rate for all launch sites, KSC has the highest → **41.7%**



https://github.com/ibogdanis/testrepo/blob/4996600a63fc40e5eb0c40da37cdc82f9d3bdb5a/image%20files/01%20-%20SpaceX%20Dash%20All%20Sites.JPG

Launch site with highest launch success ratio

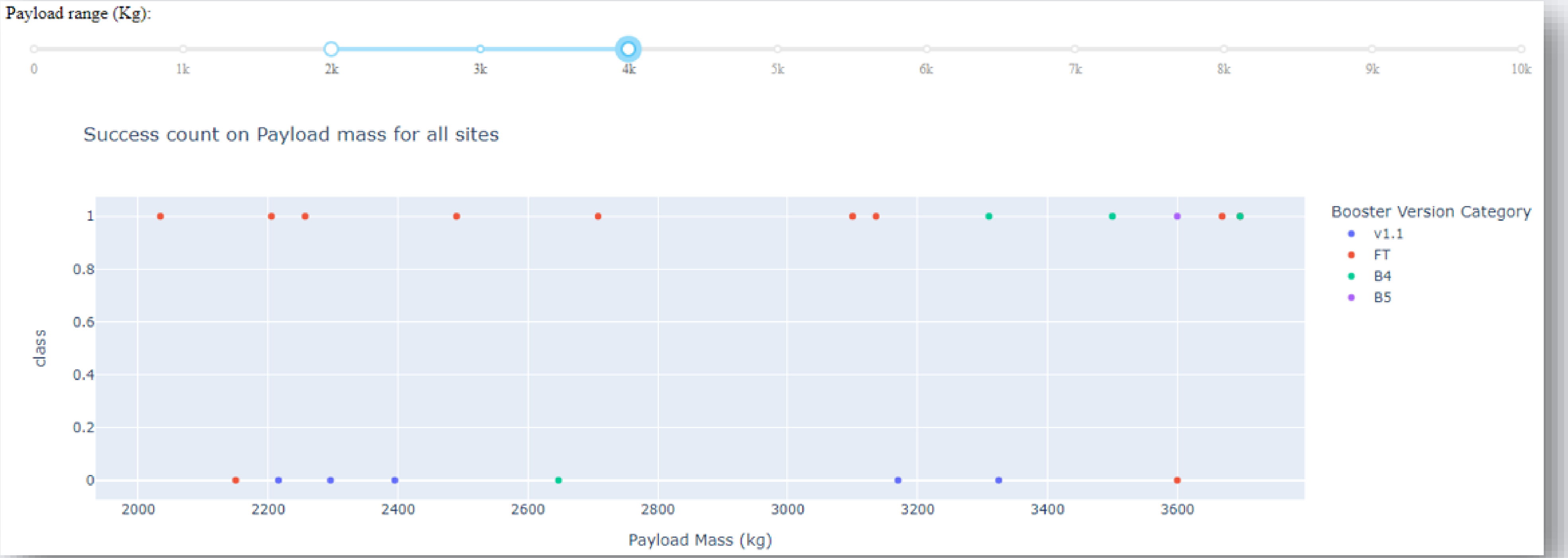
- The KSC launch site has also the most successful attempts with payload mass between 2500 and 5500 kg.



https://github.com/ibogdanis/testrepo/blob/4996600a63fc40e5eb0c40da37cdc82f9d3bdb5a/image%20files/04%20-%20KSC%20LC-39A.JPG

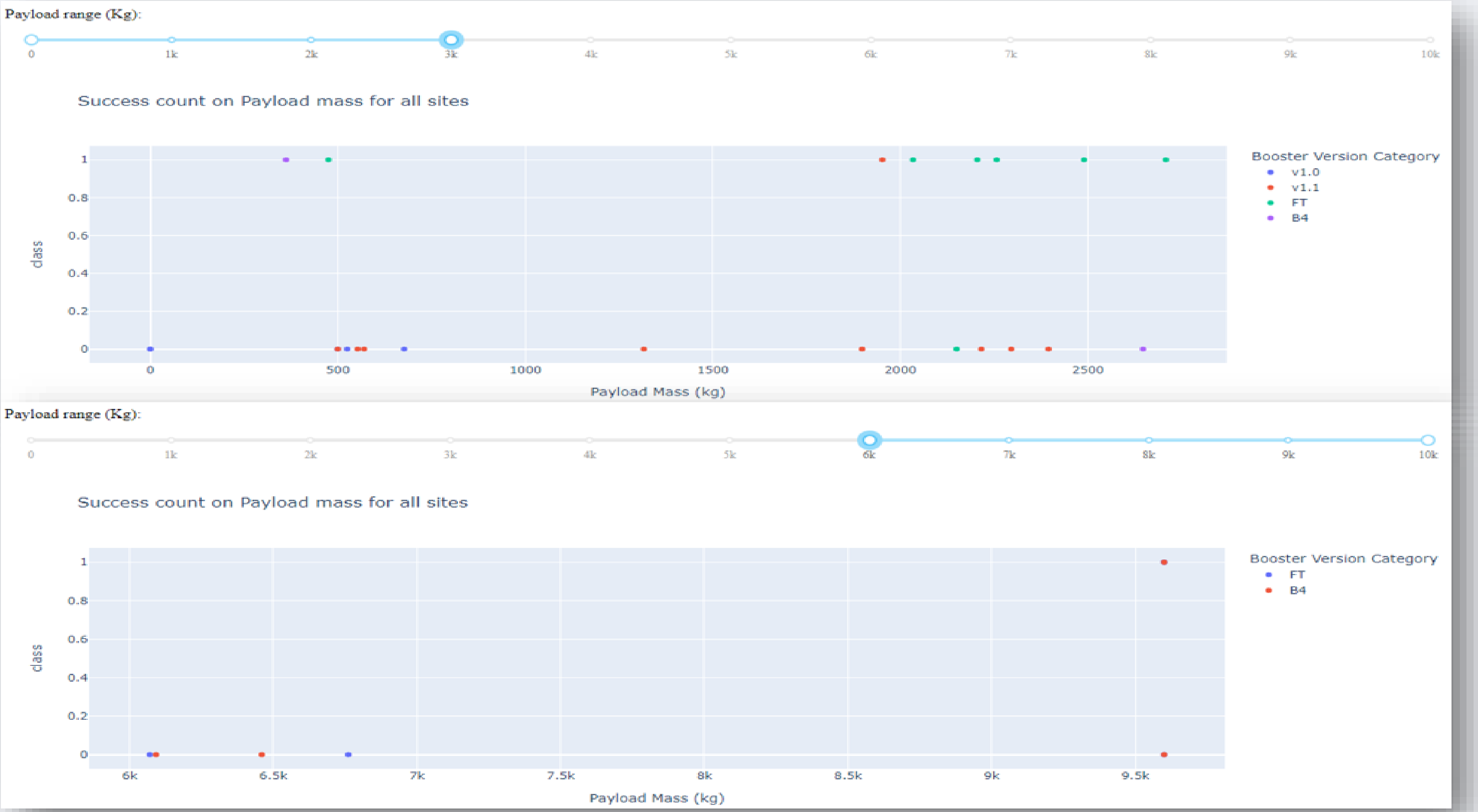
Payload vs Launch Outcome scatter plot

- The payload range which has the highest launch success rate is between 2000 and 4000 kg



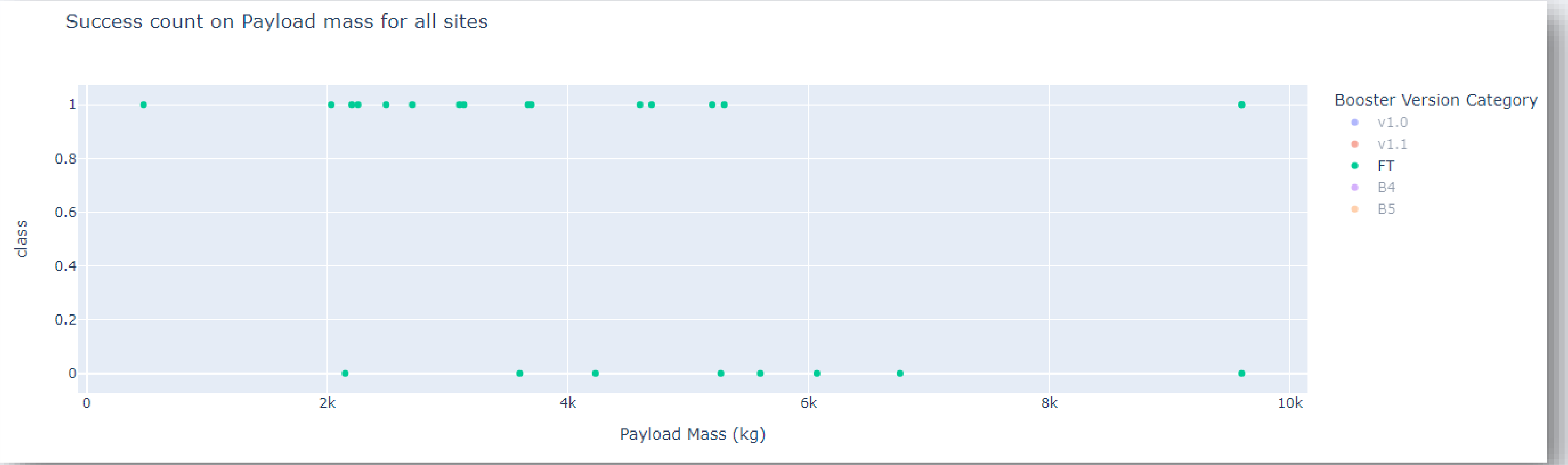
Payload vs Launch Outcome scatter plot

- The payload ranges which have the lowest launch success rate are:
  - between 0 and 3000 kg and
  - between 6000 and 10000 kg



Booster with highest success

- **FT,** Falcon 9 booster with the highest success rate.



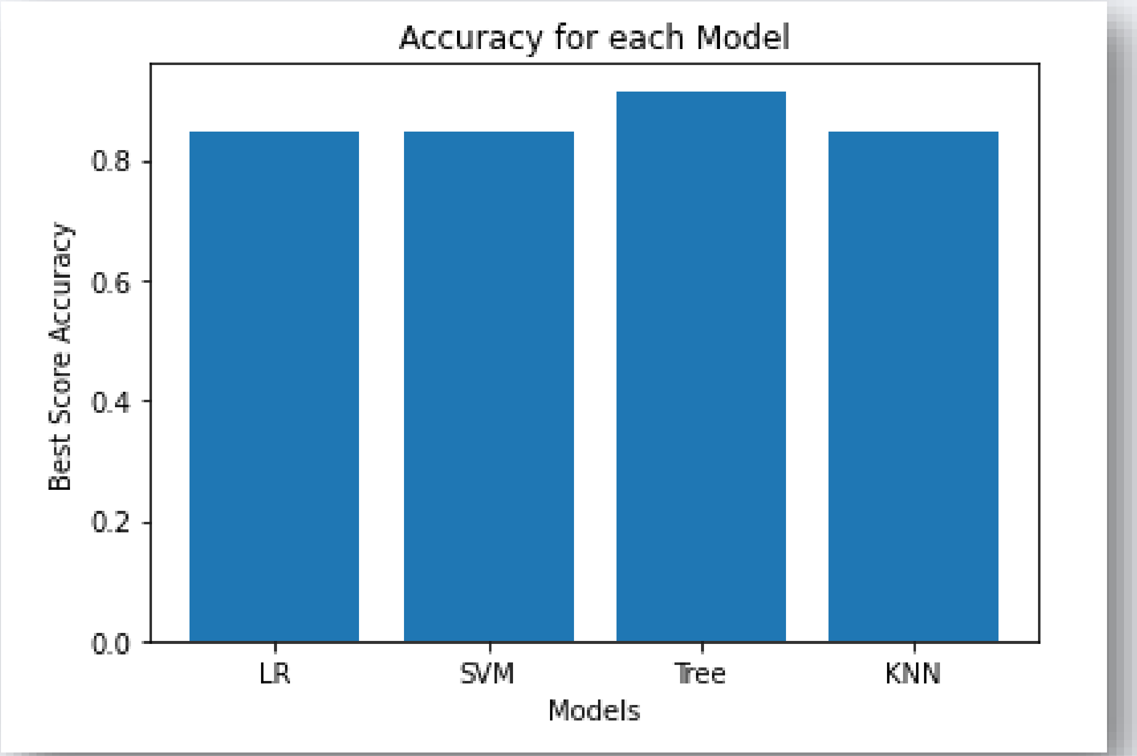
https://github.com/ibogdanis/testrepo/blob/4996600a63fc40e5eb0c40da37cdc82f9d3bdb5a/image%20files/FT.JPG

Section 5

Predictive Analysis (Classification)

Classification Accuracy

- The model with highest accuracy is Decision Tree with 91.4%

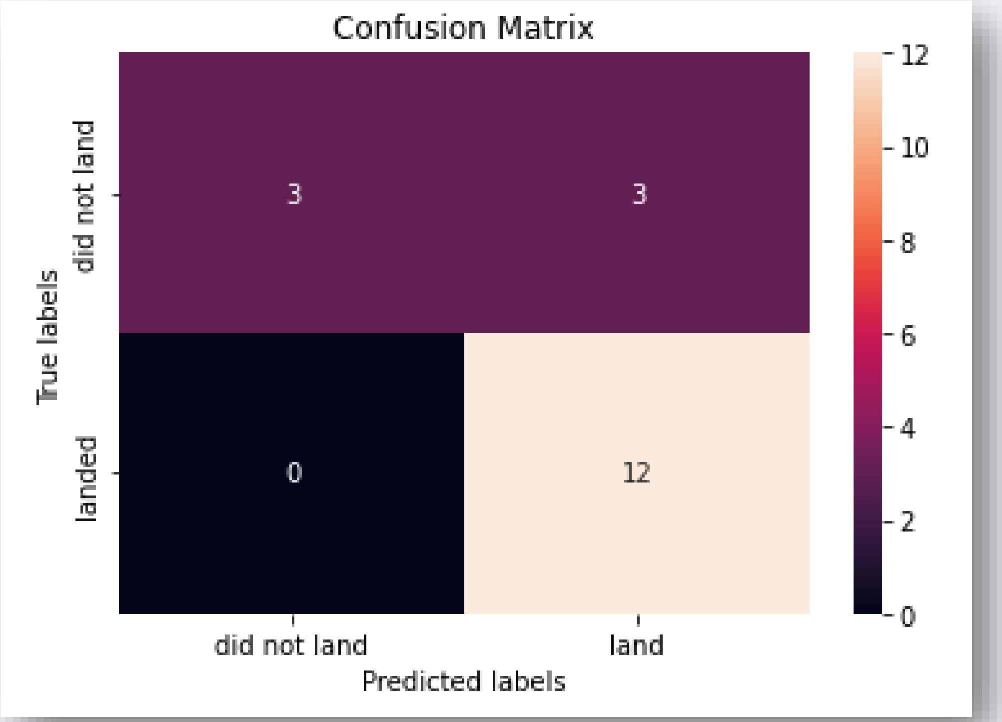



```
Accuracy for Logistics Regression (LR): 0.8464285714285713
Accuracy for Support Vector Machine (SVM): 0.8482142857142856
Accuracy for Decision Tree (DT): 0.9142857142857144
Accuracy for K nearsdt neighbors (KNN): 0.8482142857142858
```

Confusion Matrix

- All models tested have the same confusion matrix,
- From total 18 samples from our test set, 12 landed successfully and 6 failed to land.
- Our prediction model predicted:
  - All the 12 successful landings
  - The 3 failed landings, and
  - Missed the other 3 fails.
- Accuracy: 83.3%,
  - (TP + TN) / Total Samples = (12 + 3) / 18 = 0.83333

	Predicted: **NO**	Predicted: **YES**	
Actual: **NO**	True Negative **TN = 3**	False Positive **FP = 3**	**6**
Actual: **YES**	False Negative **FN = 0**	True Positive **TP = 12**	**12**
	3	**15**	**= 18**



Conclusions

1. Most successful site: **Kennedy Space Center Launch Complex 39A (KSC LC-39A)**
2. Most successful booster: **Falcon 9 Full Thrust (FT)**
   i. Landed on **drone ship**, with a medium payload between **4000kg** and **6000kg,**
   ii. With the highest success launch rate
3. Most successful landing pad: **Drone Ship Landing Pad**
4. Most accurate prediction model: **Decision Tree Classifier**

KSC LC-39A

Falcon9 FT

Drone Ship Landing Pad

Photos are taken from Wikipedia

Conclusions – In Details

- We should keep the cost low by consuming less fuel and ensuring a safe landing,
  1. By choosing the east sites, where the sea is located to the east and the rotation of the earth is also eastward, we gain an extra boost in velocity and maintain the safety of the inhabitants.
  2. Also, the east sites are 6 degrees closer to the equator where the earth's velocity is higher, so we have lower fuel consumption.
  3. The most launch attempts and the most successful landings were took place at the eastern launch sites.
  4. The most successful landings took place on drone ships.
  5. The most successful launch site is the Kennedy Space Center Launch Complex 39A (KSC LC 39A) - **41.7%** between all sites,
  6. KSC LC-39A, also, has the most successful launches - **76.9%** from total launches that made from this site,
- The most successful boosters and payloads,
  1. The heaviest payloads were launched from the eastern launch sites.
  2. The most newer the booster version the higher the landing success, due to success rate is related to the year.
  3. The payload range with highest success is between **2000kg** and **4000kg**,
  4. The booster with the highest success launch rate is the **F9 FT,**
  5. The most successful boosters landed on drone ship, with a medium payload between **4000kg** and **6000kg**, are:
     - **F9 FT** B1021.2
     - **F9 FT** B1031.2
     - **F9 FT** B1022
     - **F9 FT** B1026
- And a prediction model of **91.4%** accuracy, Decision Tree Classifier.

Appendix

Lab1: Collecting Data

https://github.com/ibogdanis/testrepo/blob/6637bcbd9e2faf51d754cdd2d825378c9d8efe90/Final_Assignment.ipynb

Lab2 : Web Scraping

https://github.com/ibogdanis/testrepo/blob/6637bcbd9e2faf51d754cdd2d825378c9d8efe90/Final_Assignment_v2.ipynb

Lab3 : Data Wrangling

https://github.com/ibogdanis/testrepo/blob/6637bcbd9e2faf51d754cdd2d825378c9d8efe90/Final_Assignment_v3.ipynb

Lab4 : EDA with SQL

https://github.com/ibogdanis/testrepo/blob/6637bcbd9e2faf51d754cdd2d825378c9d8efe90/Final_Assignment_Complete%20the%20EDA%20with%20SQL%20lab.ipynb

Lab5 : EDA with Visualization

https://github.com/ibogdanis/testrepo/blob/6637bcbd9e2faf51d754cdd2d825378c9d8efe90/Complete%20the%20EDA%20with%20Visualization%20lab.ipynb

Lab6 : Location Analysis with Folium

https://github.com/ibogdanis/testrepo/blob/6637bcbd9e2faf51d754cdd2d825378c9d8efe90/Interactive%20Visual%20Analytics%20with%20Folium%20lab.ipynb

Lab7 : Machine Learning Prediction

https://github.com/ibogdanis/testrepo/blob/6637bcbd9e2faf51d754cdd2d825378c9d8efe90/Machine%20Learning%20Prediction%20lab.ipynb

Dashboard code

https://github.com/ibogdanis/testrepo/blob/6637bcbd9e2faf51d754cdd2d825378c9d8efe90/dash_interactivity.py

Thank you!