Capstone – Coursera Data Science

BY: Mouaz

Outline

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

Executive Summary

- Summary of methodologies
  - SpaceX Data Collection using SpaceX API
  - SpaceX Data Collection with Web Scraping
  - SpaceX Data Wrangling
  - SpaceX Exploratory Data Analysis using SQL
  - Space-X EDA DataViz Using Python Pandas and Matplotlib
  - Space-X Launch Sites Analysis with Folium-Interactive Visual Analytics and Ploty Dash
  - SpaceX Machine Learning Landing Prediction
- Summary of all results

  EDA results

  Interactive Visual Analytics and Dashboards

  Predictive Analysis(Classification)

Introduction

Project background and context

SpaceX 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 SpaceX 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 SpaceX for a rocket launch.

Problems you want to find answers

In this capstone, we will predict if the Falcon 9 first stage will land successfully using data from Falcon 9 rocket launches advertised on its website.

Methodology

Executive Summary

- Data collection methodology:
  - Describes how data sets were collected
- Perform data wrangling
  - Describes how data were processed
- Perform exploratory data analysis (EDA) using visualization and SQL
- Perform interactive visual analytics using Folium and Plotly Dash
- Perform predictive analysis using classification models
- How to build, tune, evaluate classification models

Data Collection

- Description of how SpaceX Falcon9 data was collected.
  - Data was first collected using SpaceX API (a RESTful API) by making a get request to the SpaceX API. This was done by first defining a series helper functions that would help in the use of the API to extract information using identification numbers in the launch data and then requesting rocket launch data from the SpaceX API url.
  - Finally to make the requested JSON results more consistent, the SpaceX launch data was requested and parsed using the GET request and then decoded the response content as a Json result which was then converted into a Pandas data frame. •
  - Also performed web scraping to collect Falcon 9 historical launch records from a Wikipedia page titled List of Falcon 9 and Falcon Heavy launches of the launch records are stored in a HTML. Using BeautifulSoup and request Libraries, I extract the Falcon 9 launch HTML table records from the Wikipedia page, Parsed the table and converted it into a Pandas data frame

Data Collection - SpaceX API

- Data collected using SpaceX API (a RESTful API) by making a get request to the SpaceX API then requested and parsed the SpaceX launch data using the GET request and decoded the response content as a Json result which was then converted into a Pandas data frame

- Here is the GitHub URL of the completed SpaceX API calls notebook (https://github.com/mouazsa/SpaceX-Falcon-9-Landing_prediction)

Task 1: Request and parse the SpaceX launch data using the GET request

To make the requested JSON results more consistent, we will use the following static response object for this project:

static ison url='https://cf-courses-data.53.us.clomj-Qbiect-storaee.apDJoinain.claud/IBH-D59321EN-SkillsHetwork/data5ets/API call spacer: api.ison'

We should see that the request was successful! with the 200 status response code

response.status_code

203

Now we decode the response content as a Json using . json{) and turn it into a Pandas dataframe using . json jortializef)

```
# Use json_n™?tize iieetM to convert t/ie json result into n dataframe j|j'Jj
respjson = response.json()
data = pd.jsonjomalizefrespjson)
```

Data Collection - WebScraping

- Performed web scraping to collect Falcon 9 historical launch records from a Wikipedia using BeautifulSoup and request, to extract the Falcon 9 launch records from HTML table of the Wikipedia page, then created a data frame by parsing the launch HTML.

- [Here](https://github.com/mouazsa/SpaceX-Falcon-9-Landing_prediction/blob/main/2.%20Space-X%20Web%20scraping%20Falcon%209%20and%20Falcon%20Heavy%20Launches%20Records%20from%20Wikipedia.ipynb)is the GitHub URL of the completed web scraping notebook.

- https://github.com/mouazsa/SpaceX-Falcon-9-Landing_prediction/blob/main/2.%20Space-X%20Web%20scraping%20Falcon%209%20and%20Falcon%20Heavy%20Launches%20Records%20from%20Wikipedia.ipynb

In [4]:

```
static_url = "https://en.wikipedia.org/w/index.php?title=List_of_Falcon_9_and_Falcon_Heavy_launches&oldid=Iti27686922'
```

Next, request the HTML page from the above URL and get a response object

TASK 1: Request the Falcon9 Launch Wiki page from its URL

First, let's perform an HTTP GET method to request the Faicon9 Lajrch HTML page, as an HTTP response.

```
9 use requests,get() method with the provided static_urL 9 assign
the response to a object response = requests.get(static_url)
```

Create a BeautifulSoup object from the HTML response

```
9 Use BeautifulSoup0 to create a BeautifulSoup object from a response text content soup =
BeautifulSoup(response.content, 'ht»l.parser')
```

Prirt the page title to verify if the BeautifulSoup object was created properly

```
9 Use soup.titLe attribute
soup.title
```

: List of falcon 9 and falcon Heavy launches - Wikipedia

TASK 2: Extract all column/variable names from the HTML table header

Next, we want to collect all relevant column names from the HTML table header

Let's try to f rd al tab es on the wiki page first. If you need to refresh your memory about BeautifulSoup , please check the external re this lab

```
9 Use the find_all function in the BeautifulSoup object, with element type table' w Assign the result to a list called 'html_tables'
i i
```

Data Wrangling

- After obtaining and creating a Pandas DF from the collected data, data was filtered using the *BoosterVersion* column to only keep the Falcon 9 launches, then dealt with the missing data values in the *LandingPad* and *PayloadMass* columns. For the *PayloadMass*, missing data values were replaced using mean value of column.
- Also performed some Exploratory Data Analysis (EDA) to find some patterns in the data and determine what would be the label for training supervised models
- Here is the GitHub URL of the completed data wrangling related notebooks.

TASK 4: Create a landing outcome label from Outcome column

Using the Outcome , create a list where the element is zero if the corresponding row in Outcome is variable landing_class :

```
#   Landing_cLass = 0 if bad_outcome
#   Landing_cLass = 1 otherwise
df['Class'] = df['Outcome'].apply(lambda x: 0 if x in bad_outcomes
else 1) df['Class'].value_counts()
```

```
1   60
0   30
Name: Class, dtype: int64
```

This variable will represent the classification variable that represents the outcome of each launch., If tl first stage landed Successfully

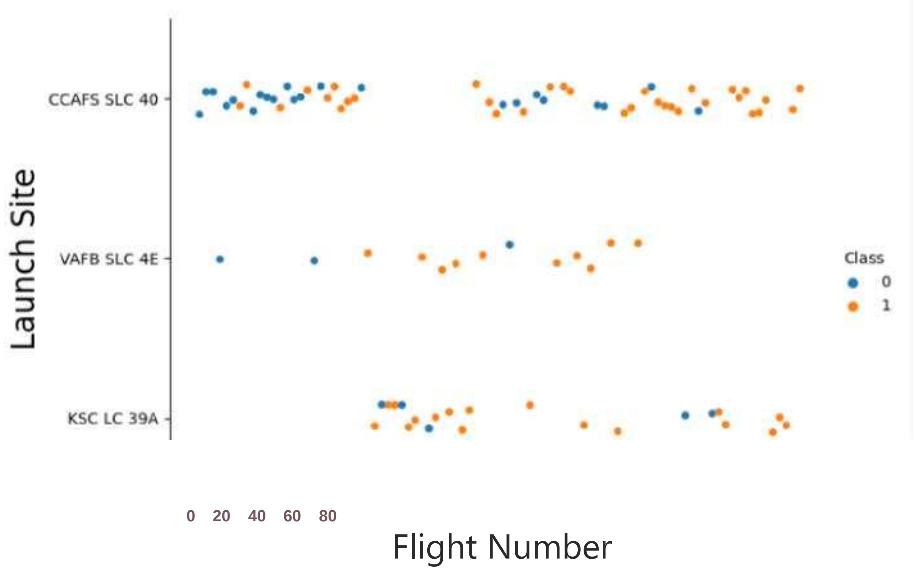
```
landing_class=df['Clas
s'] df[['Class']].head(8)
```

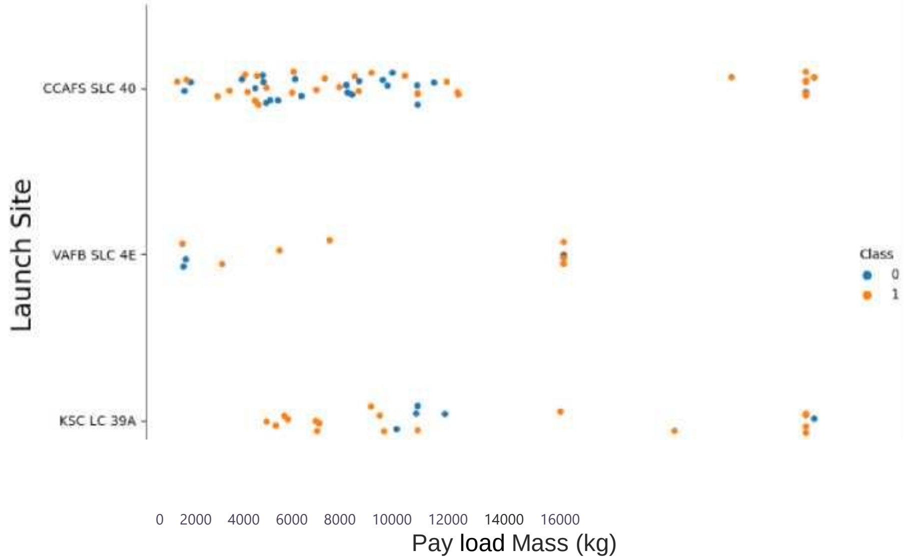
EDA with Data Visualization

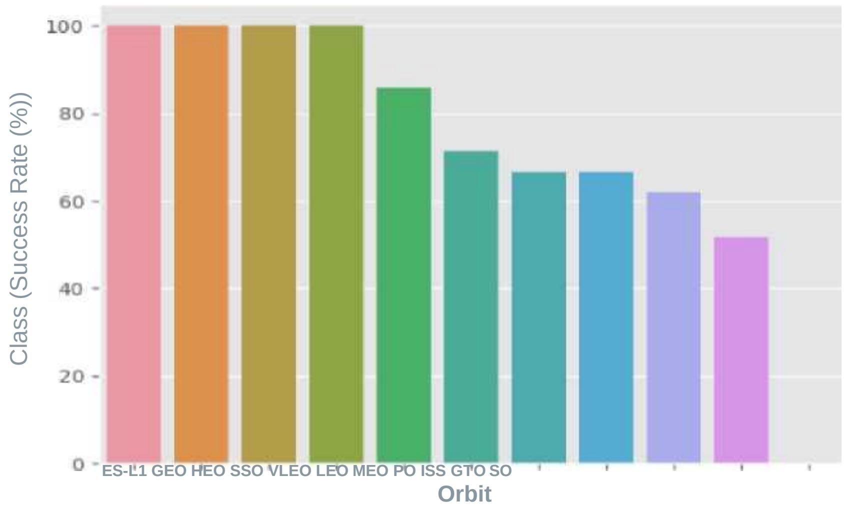
- Performed data Analysis and Feature Engineering using Pandas and Matplotlib.i.e.
- Exploratory Data Analysis
- Preparing Data Feature Engineering
- Used scatter plots to Visualize the relationship between Flight Number and Launch Site, Payload and Launch Site, FlightNumber and Orbit type, Payload and Orbit type.
- Used Bar chart to Visualize the relationship between success rate of each orbit type
- Line plot to Visualize the launch success yearly trend.
- Here is the GitHub URL of your completed EDA with data visualization notebook,

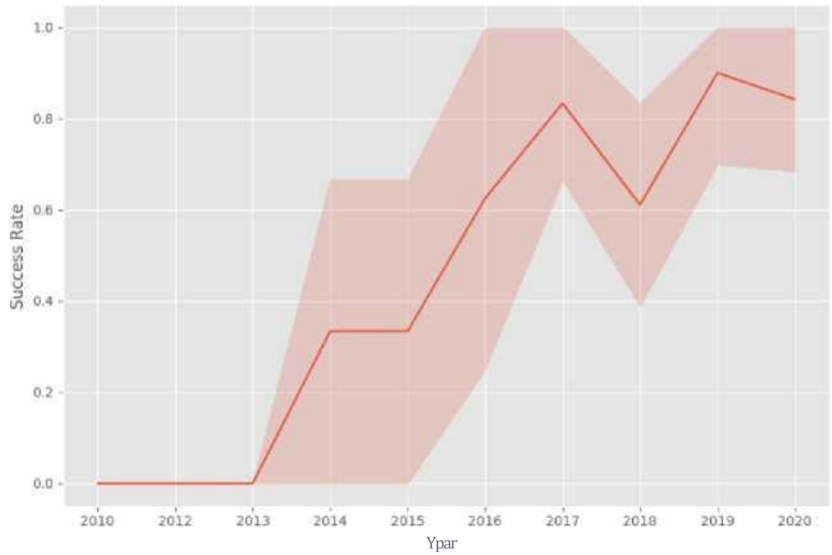
(*https://github. com/cgatama/SpaceX-Falcon-9- 1st-stage-Success-Landing-Prediction/blob/main/5. %>20Space-X%20EDA%20Data Viz°%20Using %20Pandas%20and%20Matplotlib %20- %20SpaceX. ipynb)*

EDA with Data Visualization (Plots Cont....)









EDA with SQL

- The following SQL queries were performed for EDA

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

    %sq- SELECT DISTINCT LAIUHCH_SXTE as "Launch_Sites11 FROM SPACEXTBL;

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

    SsqL SELECT * FROM SPACEXTBL' WHERE Latin ch_Site LIKE FCCAS' LIMIT 5;

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

    Ssql SELECT SUN(PAYLQADHASS_K6 } as "Total Payload Mass (Kgs)", Customer FROM 1SPACEXTBL1 L4HERE Customer = 'NASA (CRS)';

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

    Ssql SELECT AVG PAYLOADJWSS_tG_ as "Payload Mass Kgs", Customer, Boos turners ion FROM 'SPACEXTBL' WHERE Booster_Versior LIKE ' F9 vl.lY;

EDA with SQL (Cont....)

- List the date when the first successful landing outcome in ground pad was achieved
  fcql SELECT MIN DATE) FROM 'SPACEXTBL' VJHERE "Landing _Qutcome" = "Success (ground pad)";
- List the names of the boosters which have success in drone ship and have payload mass greater than 4000 but less than 6000 *(%sq/SELECTDISTINCTBooster_Version, Payload FROM SPACEXTBL WHERE "Landing _Outcome" = "Success (drone ship)" AND PAYLOAD_MASS__KG_ > 4000AND PAYLOAD_MASS_KG_ < 6000)*
- List the total number of successful and failure mission outcomes
  rsq: SELECT "Mission.OutcoiK" , CQUNTf'Ttission.Qutcarrie") as Total FROM SPACEXTBL GROUP BY "HissiorjOiitcane"

Here is the GitHub URL of your completed EDA with SQL notebook.

(https://github.com/mouazsa/SpaceX-Falcon-9-Landing_prediction/blob/main/4.%20Space-X%20EDA%20Using%20SQL.ipynb)

Build an Interactive Map with Folium

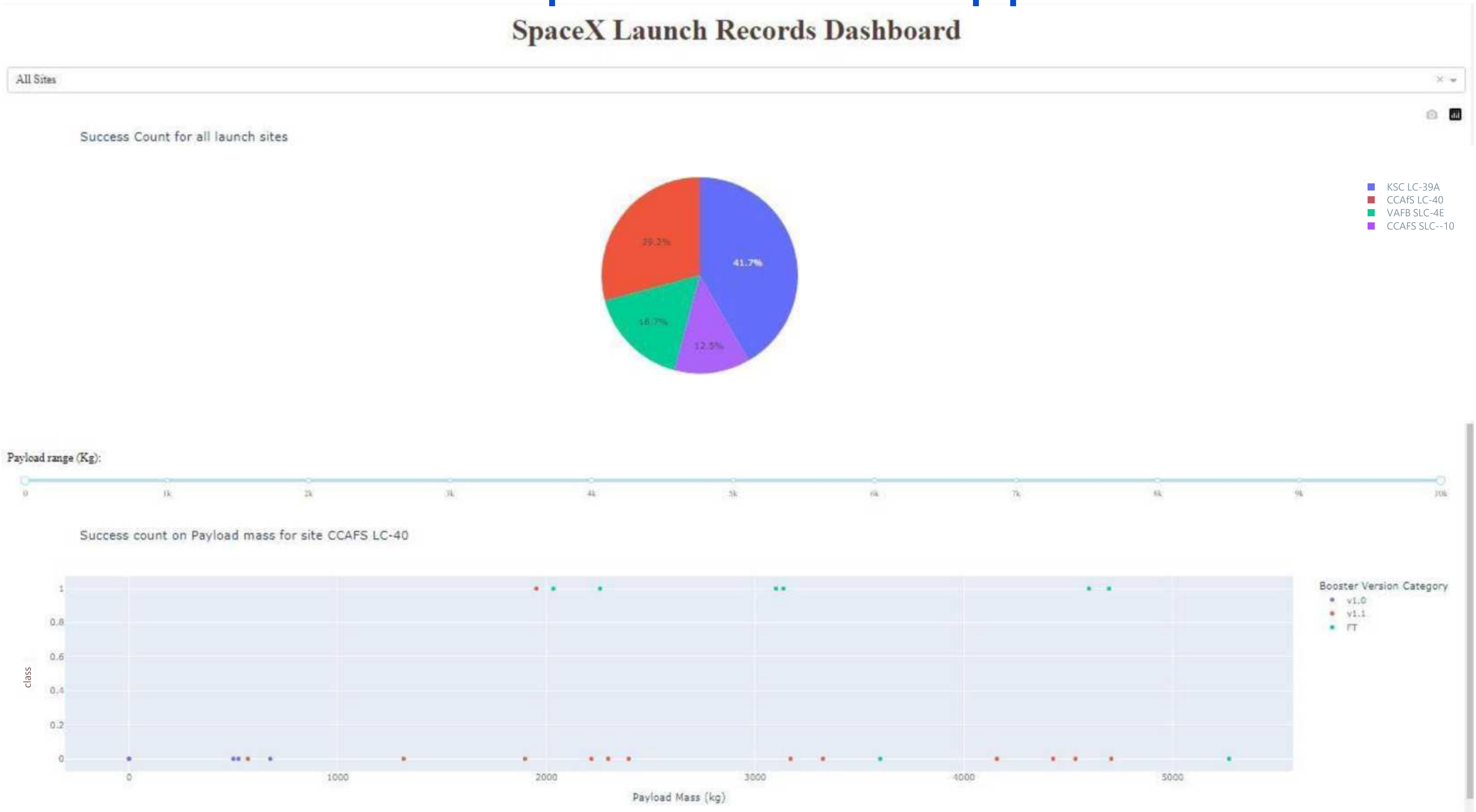
- Created folium map to marked all the launch sites, and created map objects such as markers, circles, lines to mark the success or failure of launches for each launch site.
- Created a launch set outcomes (failure=0 or success=1).
- Here is the GitHub URL of the completed interactive map with Folium map, as an external reference and peer-review purpose (*https://github.com/mouazsa/SpaceX-Falcon-9-Landing_prediction/blob/main/6.Space-X%20Launch%20Sites%20Locations%20Analysis%20with%20Folium-Interactive%20Visual%20Analytics.ipynb)*

Build a Dashboard with Plotly Dash

- Built an interactive dashboard application with Plotly dash by:
- Adding a Launch Site Drop-down Input Component
- Adding a callback function to render success-pie-chart based on selected site dropdown
- Adding a Range Slider to Select Payload
- Addeng a callback function to render the success-payload-scatter-chart scatter plot
- Here is the GitHub URL of your completed Plotly Dash lab

(https://github.com/mouazsa/SpaceX-Falcon-9-Landing_prediction/blob/main/7.%20Build%20an%20Interactive%20Dashboard%20with%20Ploty%20Dash%20-%20spacex_dash_app.py)

SpaceX Dash App



Predictive Analysis (Classification)

- Summary of how I built, evaluated, improved, and found the best performing classification model
- After loading the data as a Pandas Dataframe, I set out to perform exploratory Data Analysis and determine Training Labels by;
- creating a NumPy array from the column Class in data, by applying the method to_numpy() then assigned it to the variable Y as the outcome variable.
- Then standardized the feature dataset (x) by transforming it using preprocessing.StandardScaler() function from Sklearn. •

  - After which the data was split into training and testing sets using the function train_test_split from sklearn.model_selection with the test_size parameter set to 0.2 and random state to 2.

Predictive Analysis (Classification)

- In order to find the best ML model/ method that would performs best using the test data between SVM, Classification Trees, k nearest neighbors and Logistic Regression;
  - First created an object for each of the algorithms then created a GridSearchCV object and assigned them a set of parameters for each model.
  - For each of the models under evaluation, the GridsearchCV object was created with cv=10, then fit the training data into the GridSearch object for each to Find best Hyperparameter. •
  - After fitting the training set, we output GridSearchCV object for each of the models, then displayed the best parameters using the data attribute best_params_ and the accuracy on the validation data using the data attribute best_score_.
  - Finally using the method score to calculate the accuracy on the test data for each model and plotted a confussion matrix for each using the test and predicted outcomes.

Predictive Analysis (Classification)

- The table below shows the test data accuracy score for each of the methods comparing them to show which performed best using the test data
  between SVM, Classification Trees, k nearest neighbors and Logistic Regression; out[6S]: o

Method	Da:a AccjraQ,'
Logistic_Reg	C.S33333
SVM	;.S33333
Decision Tree	^.533333
KNN	;.S33333

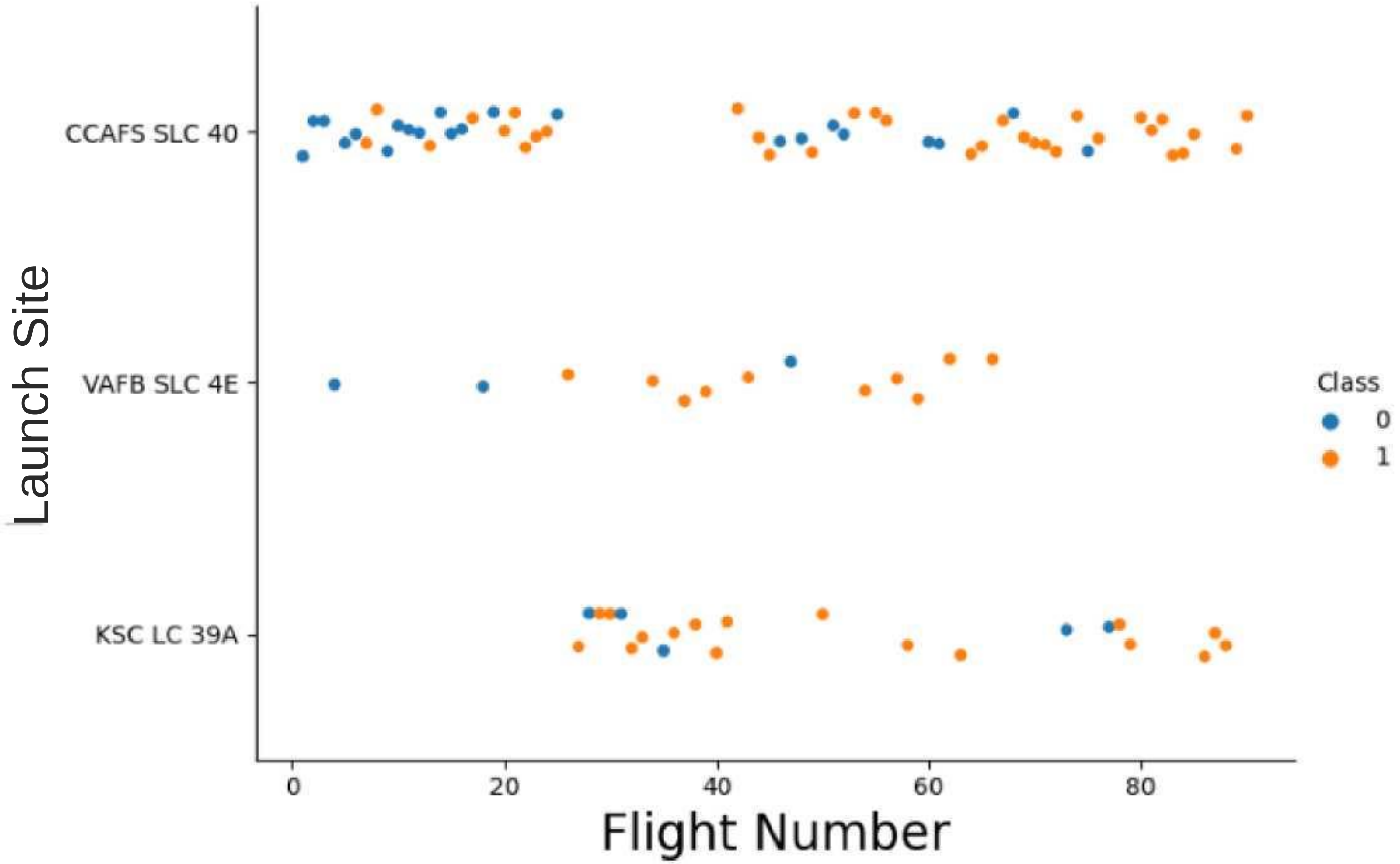
- GitHub URL of the completed predictive analysis lab
  (https://github.com/mouazsa/SpaceX-Falcon-9-
  Landing_prediction/blob/main/8.%20SpaceX%20Machine%20Learning%20Prediction.ipynb)

Results

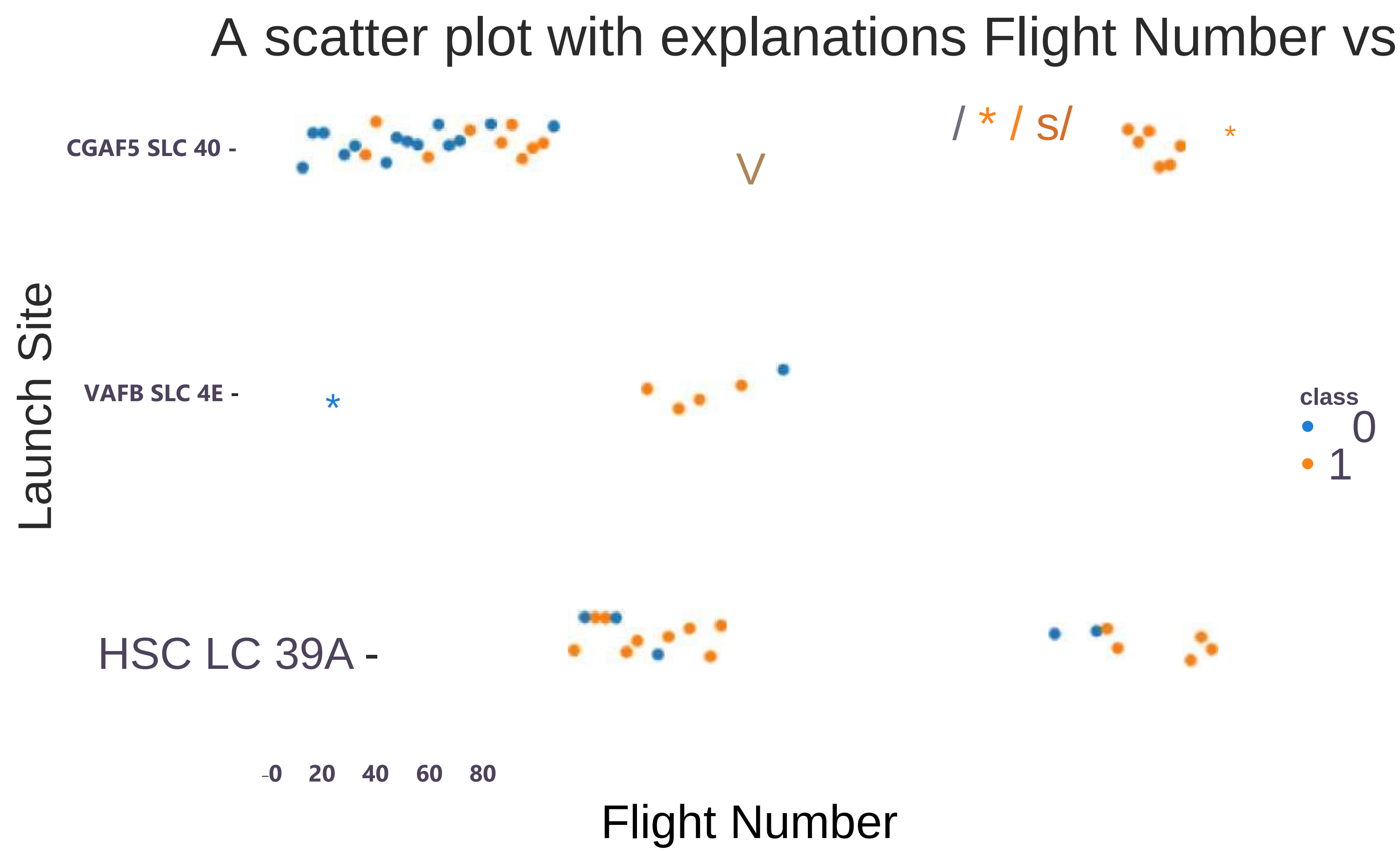
- Exploratory data analysis results
- Interactive analytics demo in screenshots
- Predictive analysis results

Flight Number vs. Launch Site

A scatter plot of Flight Number vs. Launch Site



Flight Number vs. Launch Site with explanations

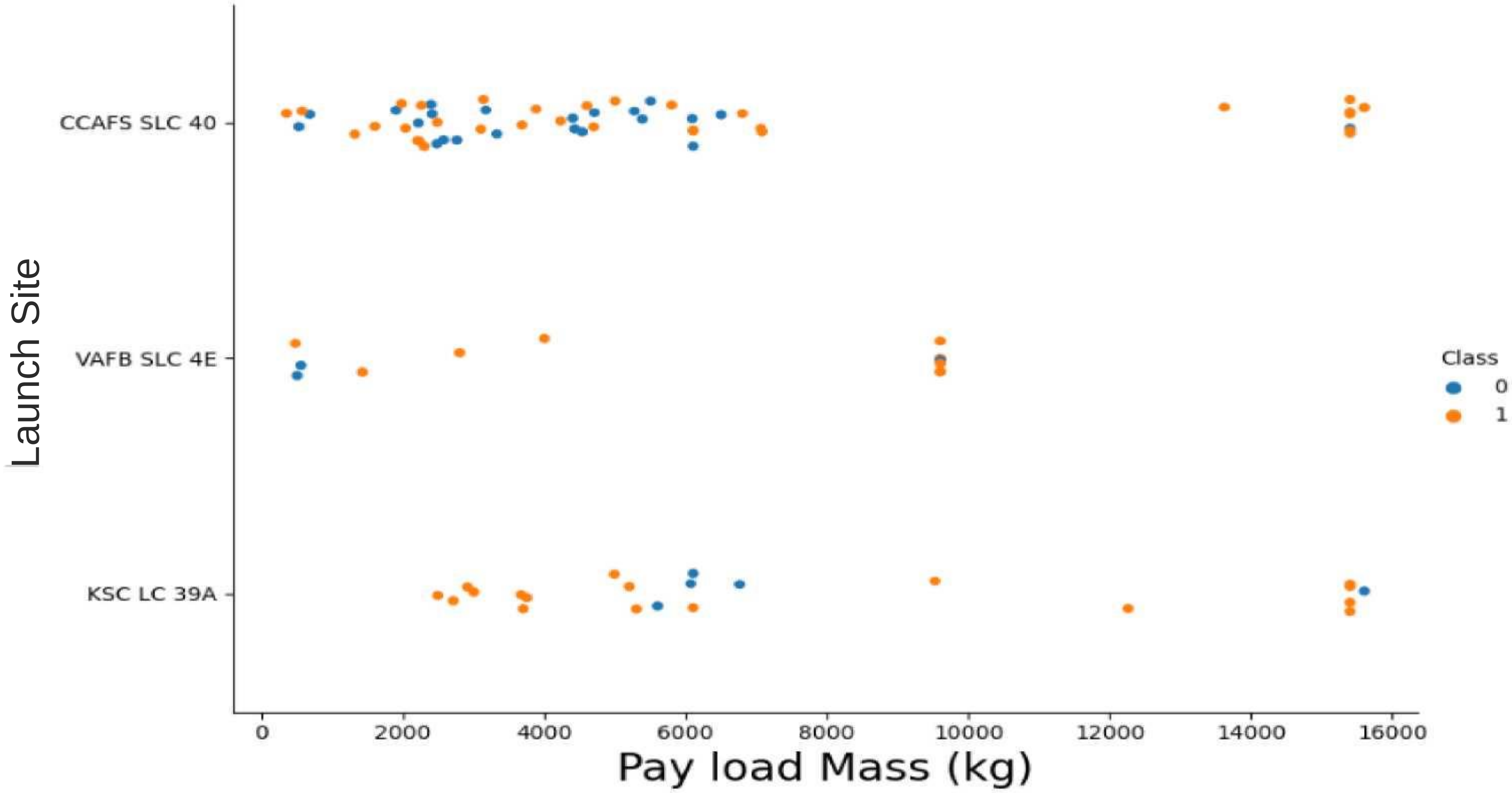


Now try to explain the patterns you found in the Flight Number vs. Launch Site scatter point plots.

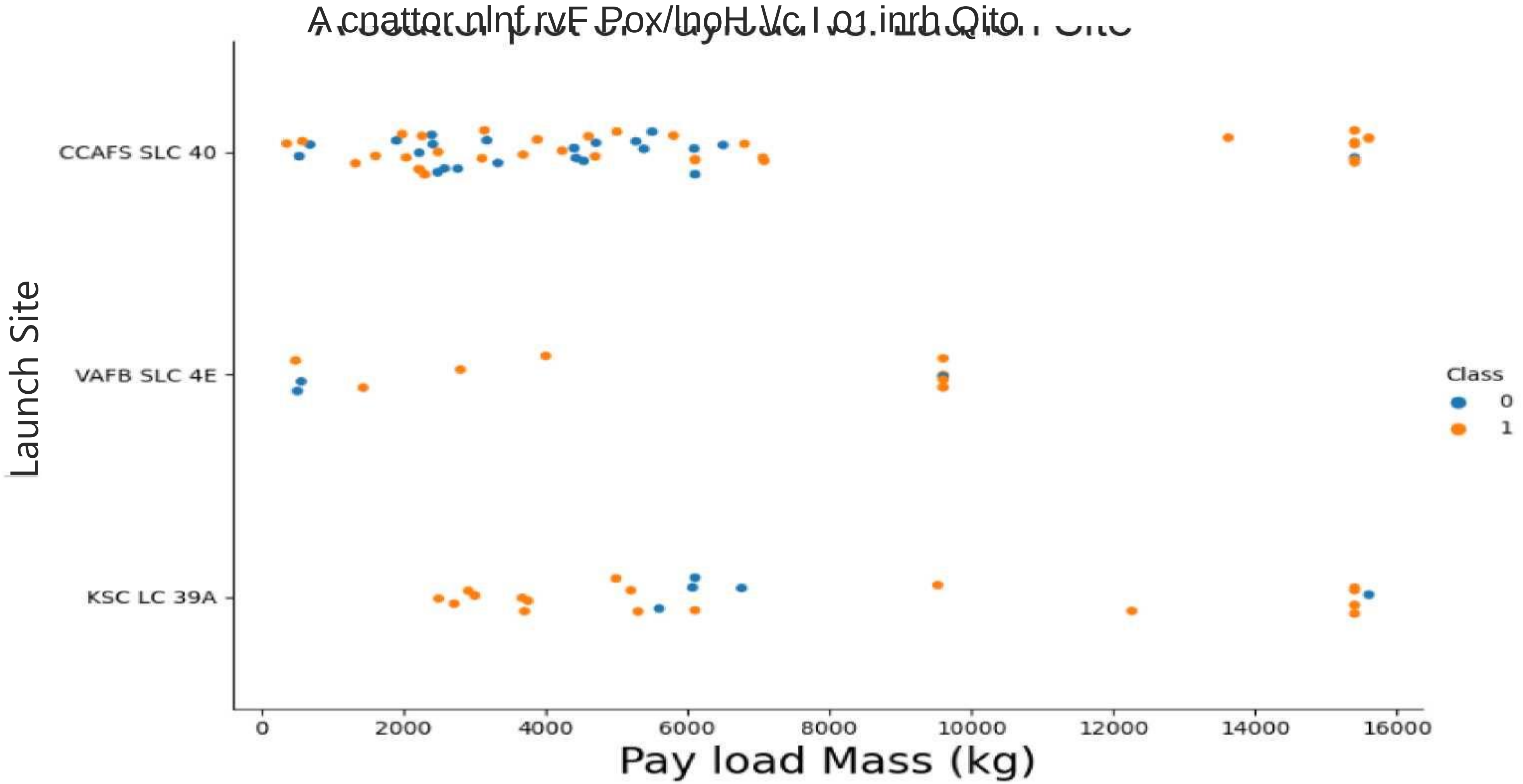
We can deduce that, as the flight number increases in each of the 3 launeg sites, so does me success -ate. For nstance. the success rate for the VAFB SLC 4E launch site is 100% afte- the Flight number 50. Both KSC LC 39A and CCAFS SLC 40 have a 100% success rates after 80th f igh:,

Payload vs. Launch Site

A scatter plot of Payload vs. Launch Site



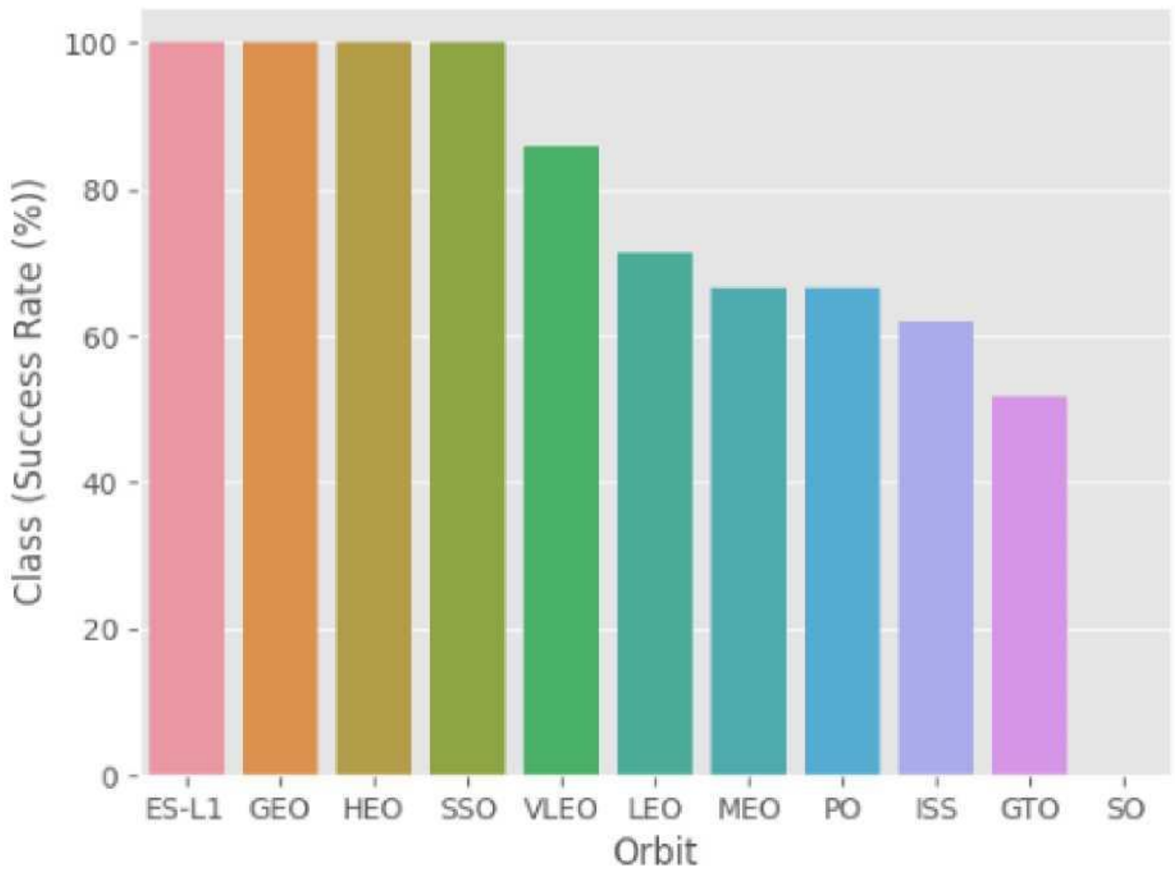
Payload vs. Launch Site with explanations



Now if you observe Payload Vs. Launch Site scatter point chart you will find for the VAFB-SLC launchsite there are no rockets launched for heavypayload mass(greater than 10000).

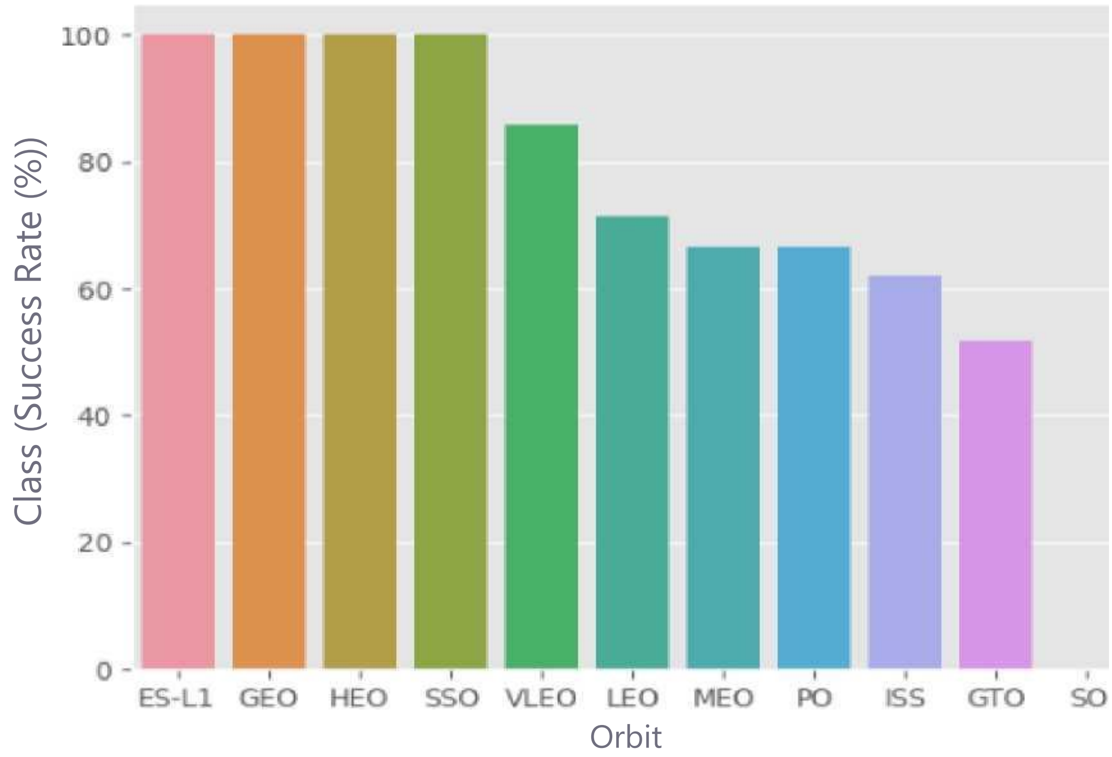
Success Rate vs. Orbit Type

- Show a bar chart for the success rate of each orbit type



Success Rate vs. Orbit Type with explanations

- Show the screenshot of the bar chart with explanations

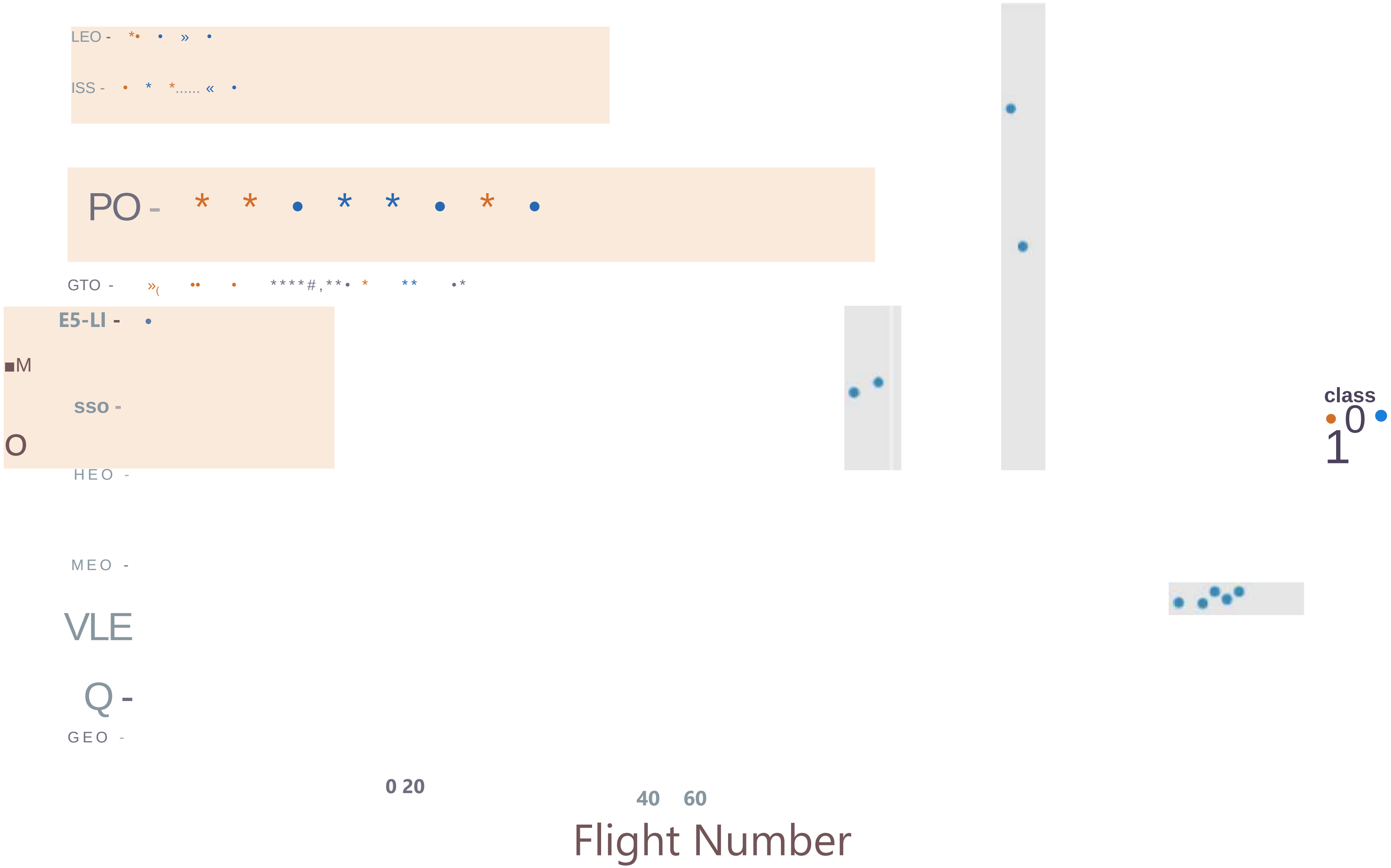


Analyze the ploted bar chart try to find which orbits have high sucess rate.

Orbits ES-L'I. GEO. HEO & SSO nave the highest success rates at 1C0%.. with SO orlbit having the lowest success rate at -50%. Orbit SO has 0% success rate,

Flight Number vs. Orbit Type

- A scatter point of Flight number vs. Orbit type



Flight Number vs. Orbit Typewith explanations

You snould 58e that in the _=0 orbit the Success appears related to the number of flights; on the other hand, there seems to be no relationship between flight number when lin GTO orbit.

Payload vs. Orbit Type

With heavy payloads the successful landing or positive landing rate are more for Polar,LEO and ISS.

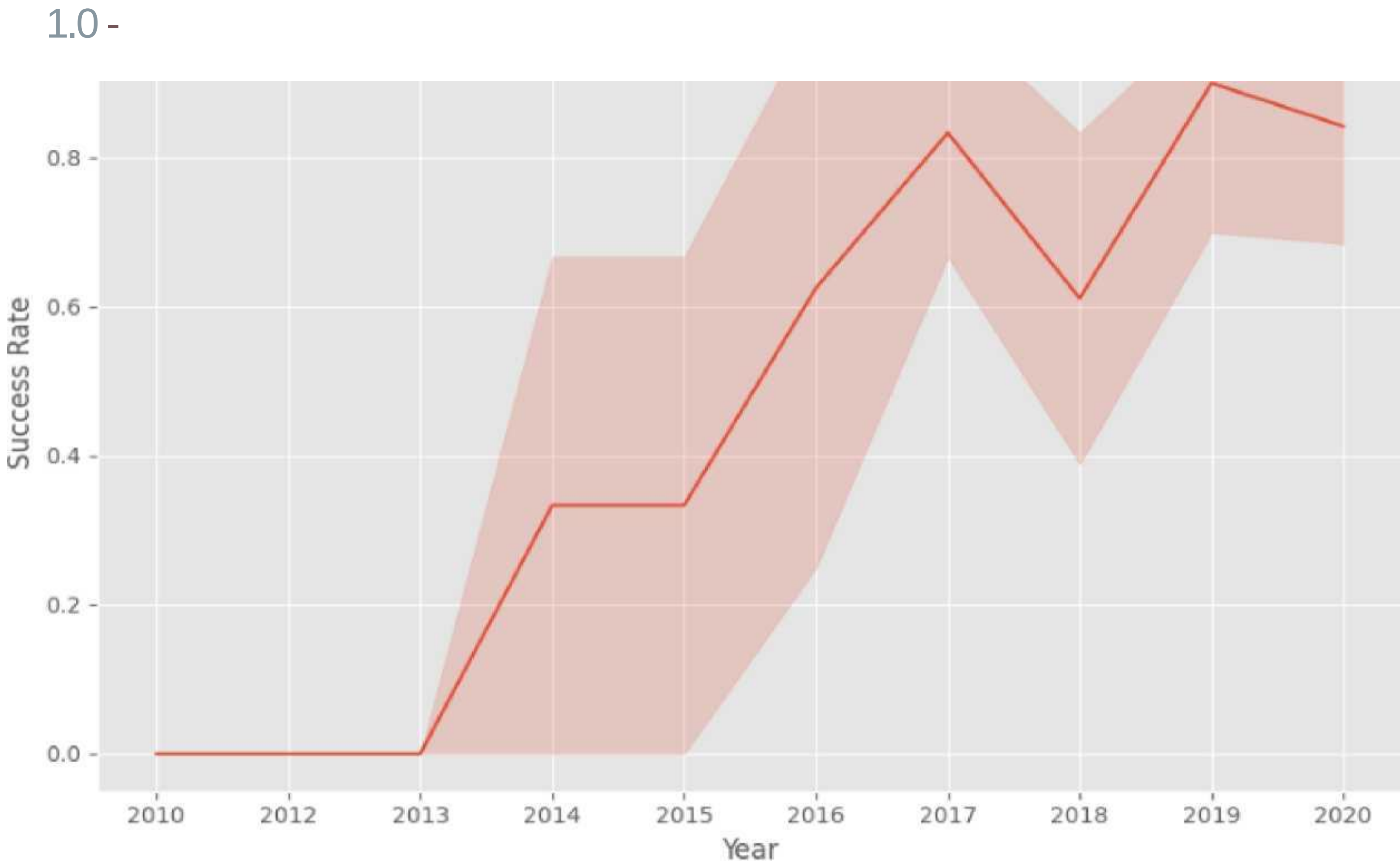
However for GTO we cannot distinguish this well as both positive landing rate and negative

landing(unsuccessful mission) both have near equal chances.

Launch Success Yearly Trend

- Since 2013, the success rate kept going up till 2020

A line chart of yearly average success rate



All Launch Site Names

- Find the names of the unique launch sites

- Used ‘SELECT DISTINCT’ statement to return only the unique launch sites from the ‘LAUNCH_SITE’ column of the SPACEXTBL table

Task 1

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

: fisql SELECT DISTINCT LflUNCH_SITE as " Laurieh_Siies" FROM SPACEXTBL;

* **sql ite: ///mydatal, d b** Dene.

Laurdi_Sites

CCAF5 L£40 VAFE

KSC LC-E9A CCAFSSLC4D

Launch Site Names Begin with 'CCA'

- Find 5 records where launch sites begin with 'CCA'

Task 2

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

In [72] :

```
Ssql SELECT * FROM ' SPACEXTBL' WHERE Launch, Si^e LIKE 'COWS' LIMIT 5;
```

K sclite: ///my_detal„db

Done.

Out[72]:

Date	Time <UTQ	Booster, Version	Launch_Site	Payload	PAYLOAD_MASS_KG_	Orbit	Customer	Missit>n_Outcome	Larding _Outcome
04-06-2010	18:45:00	F9 v' ,0 B0003	CCAFS LC-40	Dragcr Spacecraft Qua 1 '"cat on Jni:	0	.EO	SoaceX	Success	Failure (parachute)
08-12-2010	154-3:00	F9 vIJO BOOM	CCAFS LC-40	Dragon demc f ght Cl, two C jbe5ats, barre of Erouere cheese	0	.EO 1S5)	NASA [COTS) NRQ	Success	Fa' ure [parachute)
22-05-2012	0744:00	F9v .0 30005	CCAFS LC-40	Dragon demo flight C2	523	.EO OSS)	NASA [COTS)	Success	No atsempt
QS-02012	00:35:00	F9 vIJO B0006	CCAFS LC-40	SpaceX CRS-'	5'00	.EO OSS)	NASA (CRS)	Success	No atsemp:
01-03-2C1E-	15:10:00	F9 ;◌ʲ,0 300Q7	CCAFS LC-40	SpaceX CRS-2	677	.EO OSS)	NASA (CRS)	Success	No atsemp:

Used ‘LIKE’ command with ‘%’ wildcard in ‘WHERE’ clause to select and dispay a table of all records where **launch sites begin with the string 'CCA'**

Total Payload Mass

- Calculate and Display the total payload carried by boosters from NASA

Task 3

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

ftsql SELECT SUn(PAYLOADJttSS_lffiJ as "Total Payloa: MassOtgs)", Customer FROM 'SPACEXTEiL' WHERE Customer = 'NASA (CRS)[1];

* 5clite:///my_clatal.db Done.

Ou- [: Tota l Pay l oad M ass (Kgs) Cu stonier

455% NASA (CRS)

Used the ‘SUM()' function to return and dispaly the total sum of ‘ PAYLOAD_MASS_KG' column for Customer NASA(CRS'

Average Payload Mass by F9 v1.1

- Calculate the average payload mass carried by booster version F9 v1.1

Task 4

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

```
ftsql SELECT AVG(PAYLMD_MASS_KG_) as "Payload Mass Kgs"., Customer, BoosterVersicn FROM 1SPACEXTBL' WHERE Boos terJfers ion LIKE 'F=9 vl.lST
```

K sqlite:///iny_datal,db Dene,

Payload Mass Kgs	Customer	Booster_Version
1514,666E666666565	MHA	F9 v1.J B10C3

Used the ‘AVG()’ function to return and dispaly the average payload mass carried by booster version F9 v1.1

First Successful Ground Landing Date

- Find the dates of the first successful landing outcome on ground pad

Task 5

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

Hin£:Use min function

ttsql SELECT MIN (DATE) ZW1 1SPACEXTBL1 L4HERE "Landing .Outcome1' = "Success (grounc pad)";

K iclite:///my_datal.db Dene.

MIN (DATE)

•*y* -05-2017

Used the ‘MIN()' function to return and dispaly the first (oldest) date when first successful landing outcome on ground pad *Success (groundpad)happened.*

Successful Drone Ship Landing with Payload between 4000 and 6000

List of Boosters which have successfully landed on drone ship and had payload mass greater than 4000 but less than 6000

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

```
# %sql SELECT * FROM 'SPACEXTBL'
```

```
%sql SELECT DISTINCT Booster_Version, Payload FROM SPACEXTBL WHERE "Landing _Outcome" = "Success (drone ship)" AND PAYLOAD_MASS__KG_ > 4000 AND PAYLOA
```

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

Booster_Version	Payload
F9 FT B1022	JCSAT-14
F9 FT B1026	JCSAT-16
F9 FT B1021.2	SES-10
F9 FT B1031.2	SES-11 / EchoStar 105

Used 'Select Distinct' statement to return and list the 'unique' names of boosters with operators >4000 and <6000 to only list booster with payloads btween 4000-6000 with landing outcome of 'Success (drone ship)'.

Total Number of Successful and Failure Mission Outcomes

- Calculate the total number of successful and failure mission outcomes

Task 7

List the total number of successful and failure mission outcomes

fcql SELECT "Mission_Qutcome", COUNT ("Mis sicn_Outcome") as Total FROM SPACEXTEL GROUP BV 'TtLssionJ^utcome"

K sc Lite:///my_datal.db Dene .

MissionjOutcome Total

ʳa' ure (in f 'ght)

Success 98 Success

Success (payload status unc ear)

Used the ‘COUNT()' together with the ‘GROUP BY' statement to return total number of missions outcomes

Boosters Carried Maximum Payload

- List of the boosters which have carried the maximum payload mass

Sisql SELECT "Boo5ter_Vers Lon" , Payload, "PAYLQAD_RASS KG_" FROM SPACEXTBL '.JHERE " PAYLOAD_MASS KG_" = (SELECT HAX(" PAYLQAD HASS KG_") FROM SPACEXTBL)

* sqlite: Done. datal.db

Booster Version

Payload PAYLOAD, M A55_KJS_

F9 BE B104S4 F9 B5 5tarl nk 1 vl.O, SpaceX CRS-19 ar nk 2 v1.C Crew IB ECO 156CO 15600 15 ECO 156CO 15600 15600 15600 15600 15600 15600 15600

B10494 S' F9 BE Dragon r-f ; ht abort test Star nk 3 vl.O. Star ink 4 vl .0

B105' .3 F9 BE B 5tarlnk 4 vl.O, 5paceX CRS-20 Sta-nk B vl.O. Star ink

1056.4 F9 BE 6 vl .0 5star nkG v'.O, Cnew Dragon Demo-2 Sta - 'nk 7

B104S.E F9 BE vl ,0. Star ink Svl .0 Starlink 11 vl .0. Starlink ' 2 vl .0

B10514 F9 BE B Star'rk 12 vl.O, Star'rk '3 vl.0 StarTnk 13 vl .0. Starlink

1049.5 F9 BE B 14 Vl .0 Star nk 14v1.0, GPE II-04 5tarlink 1 5 vl.O,

1060.2 F9 BE 5paceX CRS-2J

B105S.3 F9 BE

BIOE',6 F9 BE B

1060.3 F9 BE B

1049.7

Using a Subquerry to return and pass the Max payload and used it list all the boosters that have carried the Max payload of 15600kgs

2015 Launch Records

- List of failed landing outcomes in drone ship, with their booster versions, and launch site names in 2015

Task 9

List the records which will disDlavthe month names, failure landing_outcomes in drone ship .booster versions, launch_siteforthe months in year 2015.

ftsql SELECT subsii^Date, 7,4) ^ £U9str(Datej 4j 2),"0ot>ster_Versior", "Launci^Site";, Payload, "PAYLOADJ-IASS KG_l", "Mis sion_Dut come1'., "Landing jOutcome"

* sqllte:///niy_da1:al.db Dene.

substr(Date,7,4)	substrfDate, 4,2)	Booster_ Version	LaiinchSite	Payload	PAYLOAD_MAS5_KG_	Mission_OiJtcome	Landing _Outcome
2015	01	F9v1.1 31012	CCAFS LC-40	5paceXCRS-5	2355	5uccess	Mure (drone ship)
2015	04	F9 vl.1 31015	CCAFS LC-40	5paceXCRS-5	1B98	5uccess	Fai ure (drone ship)

Used the ‘subsrt()' in the select statement to get the month and year from the date column where substr(Date,7,4)='2015' for year and Landing_outcome was ‘Failure (drone ship’) and return the records nmatching the filter.

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

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

Task 10

Rank the count of successful landmgoutcomes between the date 04-06-2010 and 20-03-2017 in descending order.

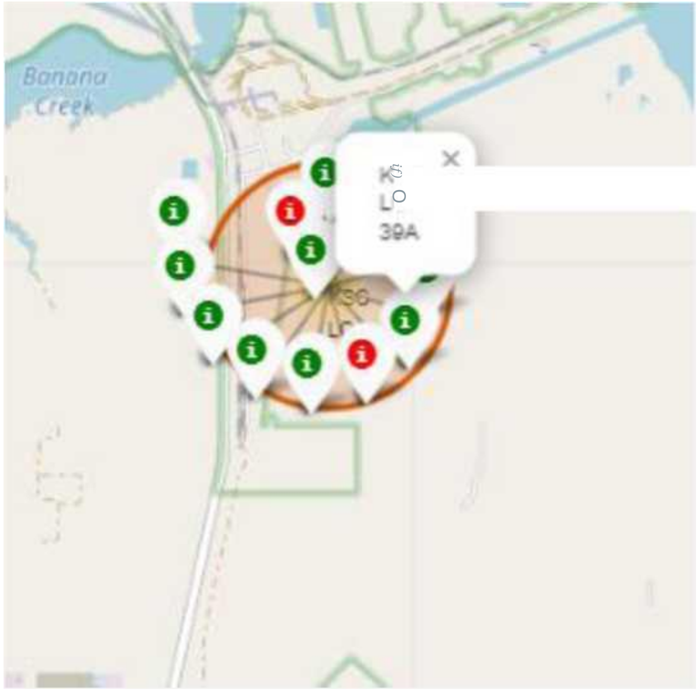
```
%sql SELECT * FROM 5PACEXTEL WHERE "Landing .Outcome1' LIKE 'Success*1 AFC [Date BETWEEN "04-05-2019' AND '20-33-2017') ORDER BY Date DESCj
```

K s:3:te:///iny_datal.db
Dene.

□ate	Time {UTQ	Booster Version	Launch Site	Payload	PAYLOAD_MASS_KG_	Orbit	Customer	Mission Outcome	Landing Outcome
19-02¬2017	14:39:00	FS 31031,'	K5C .C-3SA	SpaceX CR5-12	2493	.EO	NASA (CRS)	Success	Success [groLrd pad;
18-1G-2020	12:25:57	FSB5 31051,6	K5C .C-3SA	Star -k 13 vl 0, Star ink 14 V1.l	15600	.50	SpaceX	Success	Success
18-08¬2020	14:31:00	F9 B5 31043,5	OCAFS SLC-43	Starlmk 10 v1 0. SkyS2t-19, -20, -21, SAOCOM 13	15440	.30	SpaceX.3 anet Labs, F aretlQ	Success	Success
18-07¬2016	04:45:00	FS B1025,'	CCAFS LC-43	SpaceX CR5-9	2257	.30 OSS)	NASA CRS)	Success	Success [groLrd pad)
18-04¬2018	22:51:00	FSB4 31045,'	OCAFS SLC-43	Transiting Exoplaret Survey Sa:e lite "ESS;	562	H30	NASA [LSP)	Success	Success [drone ship)

Launch outcomes for each site on the map With Color Markers

Florida Sites



In the Eastern coast (Florida) Launch site KSC LC-39A has relatively high success rates compared to CCAFS SLC-40 & CCAFS LC-40.

Launch outcomes for each site on the map With Color Markers

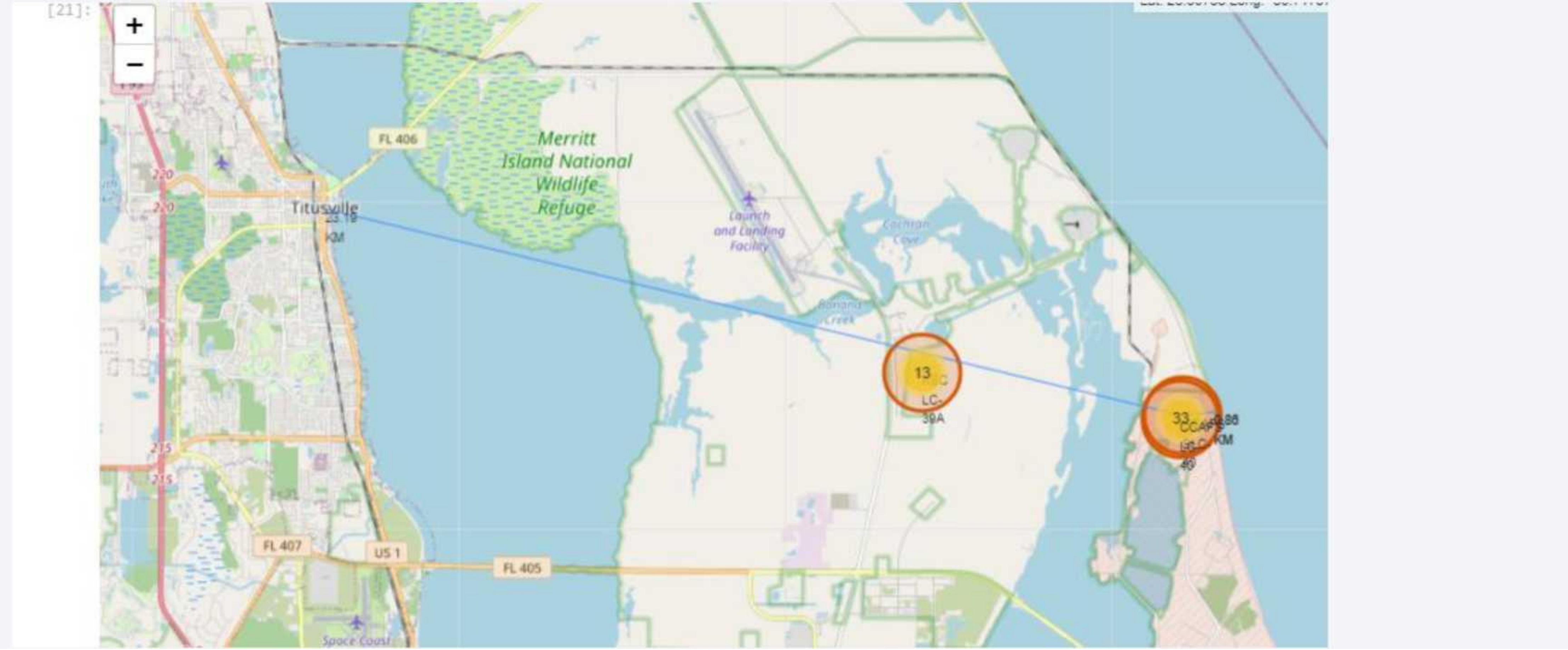
West Coast/ Carlifonia

- In the West Coast (Californai) Launch site VAFB SLC-4E has relatively lower success rates 4/10 compared to KSC LC-39A launch site in the Eastern Coast of Florida.

Distances between a launch site to its proximities

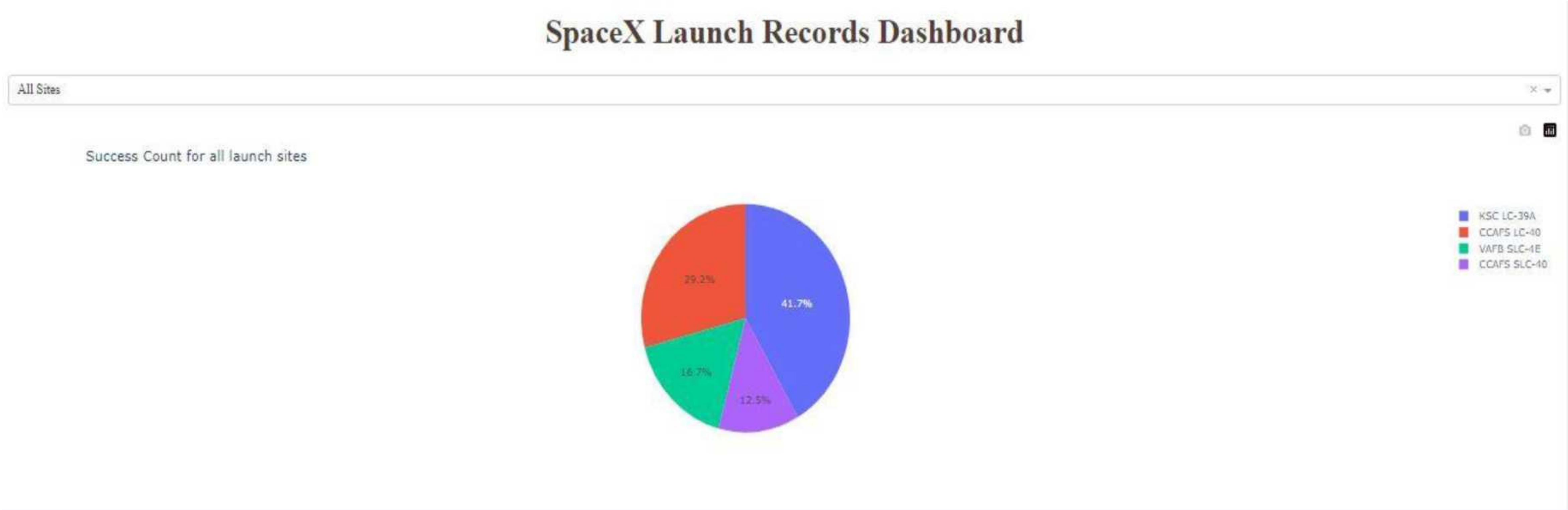
- Launch site CCAFS SLC-40 proximity to coastline is 0.86km

Distances between a launch site to its proximities



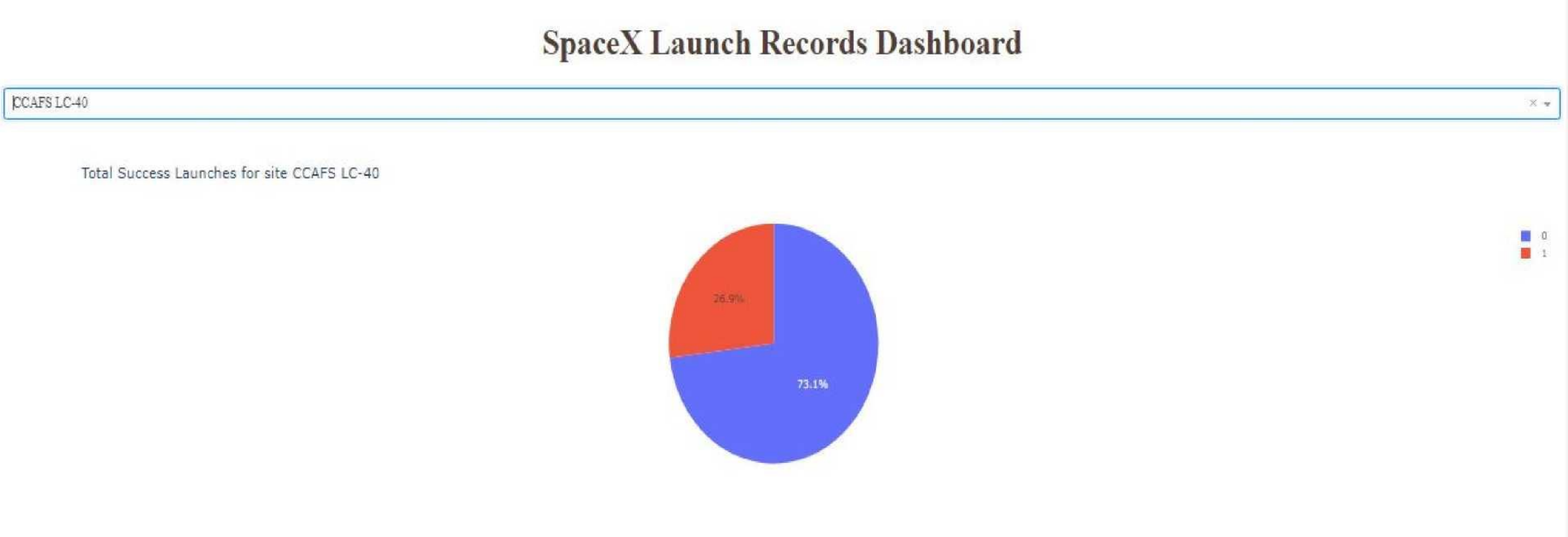
- Launch site CCAFS SLC-40 closest to highway (Washington Avenue) is 23.19km

Pie-Chart for launch success count for all sites •



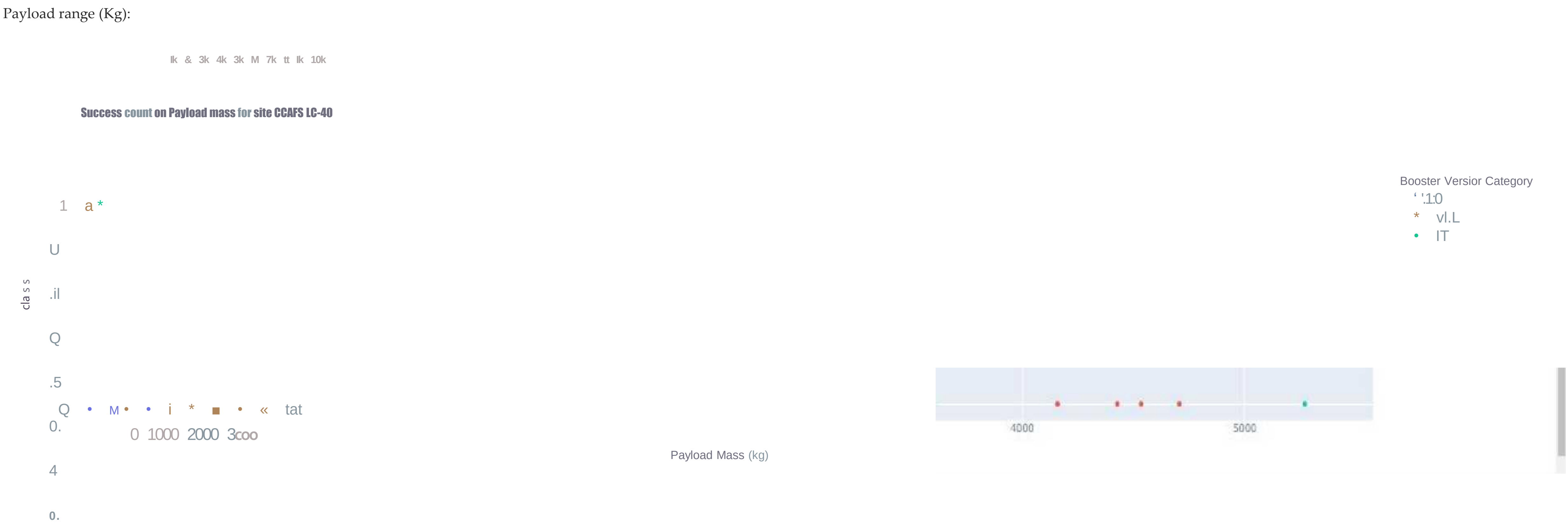
- Launch site KSC LC-39A has the highest launch success rate at 42% followed by CCAFS LC-40 at 29%, VAFB SLC-4E at 17% and lastly launch site CCAFS SLC-40 with a success rate of 13%

Pie chart for the launch site with 2nd highest launch success ratio



Launch site CCAFS LC-40 had the 2nd highest success ratio of 73% success against 27% failed launches

Payload vs. Launch Outcome scatter plot for all sites



For Launch site CCAFS LC-40 the booster version FT has the largest success rate from a payload mass of >2000kg

Classification Models Accuracy

Out[68]: 0

Meth od "est Data Accu raty Logi&tic_Reg C.S33333

SVM C.S33333

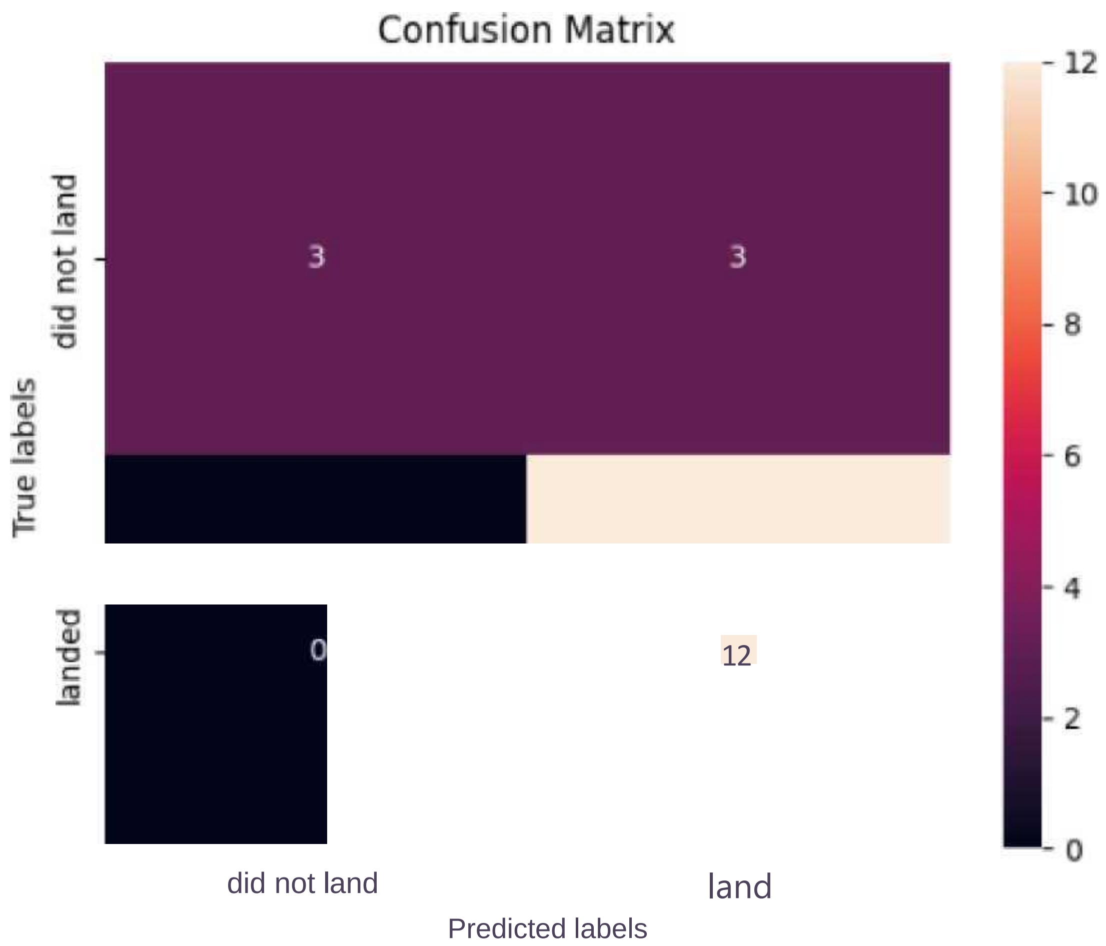
Deriaon Tree :.S33333

KNN J.B33333

*All the **methods** perform equally* on *the test data: ie, They alt have the same accuracy of €.833333 on the test Data*

Confusion Matrix

- All the 4 classification model had the same confusion matrixes and were able equally distinguish between the different classes. The major problem is false positives for all the models.



Conclusions

- Different launch sites have different success rates. CCAFS LC-40, has a success rate of 60 %, while KSC LC-39A and VAFB SLC 4E has a success rate of 77%.
- We can deduce that, as the flight number increases in each of the 3 launcg sites, so does the success rate. For instance, the success rate for the VAFB SLC 4E launch site is 100% after the Flight number 50. Both KSC LC 39A and CCAFS SLC 40 have a 100% success rates after 80th flight
- If you observe Payload Vs. Launch Site scatter point chart you will find for the VAFB-SLC launchsite there are no rockets launched for heavypayload mass(greater than 10000).
- Orbits ES-L1, GEO, HEO & SSO have the highest success rates at 100%, with SO orbit having the lowest success rate at ~50%. Orbit SO has 0% success rate.
- LEO orbit the Success appears related to the number of flights; on the other hand, there seems to be no relationship between flight number when in GTO orbit

Conclusions Cont....

- With heavy payloads the successful landing or positive landing rate are more for Polar,LEO and ISS. However for GTO we cannot distinguish this well as both positive landing rate and negative landing(unsuccessful mission) are both there here
- Anf finally the sucess rate since 2013 kept increasing till 2020.