IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Debo Prosad Dutta
01.05.2022

Outline

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

Executive Summary

- **Summary of project –**

  Now a days we are in commercial space age era. In the market so many companies are there who are trying to make space object affordable. SpaceX is the most successful among them. In this project we are going to find out success factor of launch of space object in the cheapest affordable price.

- **Summary of methodologies –**

  To get the insights we have used SpaceX API and Web Scrapping Technique for Data Collection. We have used Data Wrangling methodology to prepare data for analysis. We have done exploratory data analysis using SQL and matplotlib. After that we have used folium methodology and dash to create interactive Dashboard. We have used Logistics Regression, SVC, Decision Tree Classifier, KNN for machine learning prediction.

- **Summary of all results –**

  From the analysis we have found that affordable launching of space object of SpaceX is mainly dependent on few parameters. Those are launching site location, payload and booster version category.

Introduction

- Project background and context –

  World is in commercial space age. Now every private company wants to make launching affordable to remain competitive in the space business. SpaceX has the lead in the business. SpaceX is able to make launching cost cheapest. In this project we are going to analyze factors responsible for SpaceX successful launch.

- Problems you want to find answers –

  Here we are going analyze how the launch of space object could be successful based on some factors within minimal cost.

Section 1

Methodology

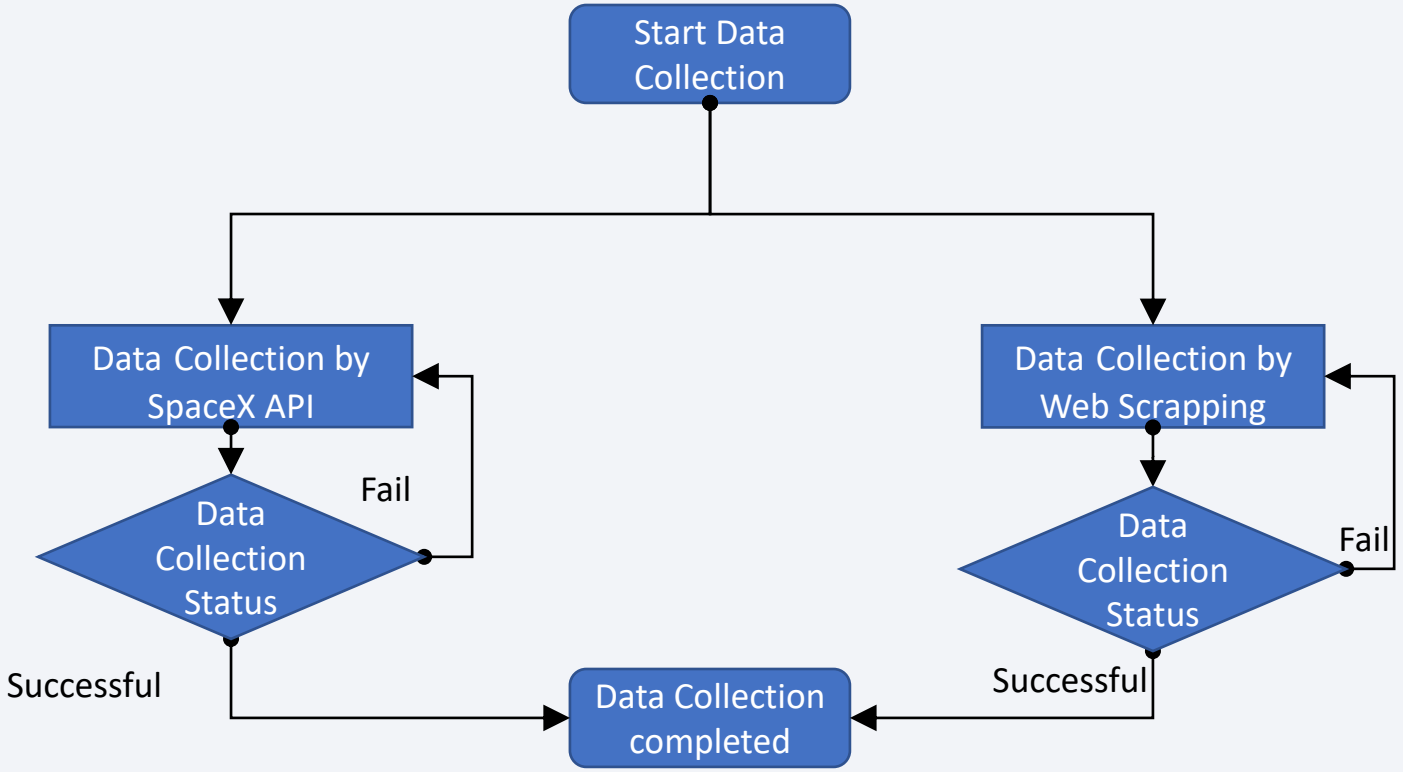
Methodology

Executive Summary

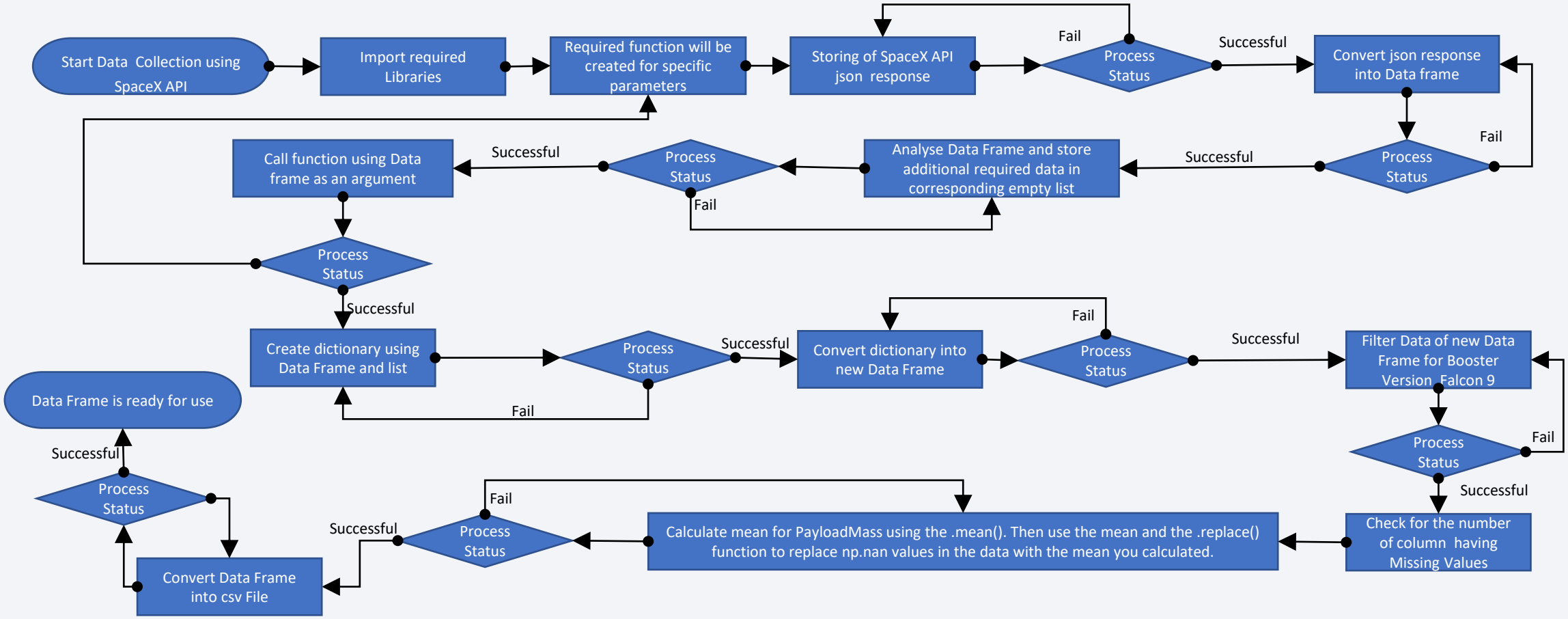
- Data collection methodology:
  - We have used SpaceX API and Web Scrapping to collect Data.
- Perform data wrangling
  - In the collected data there are several there are several different cases where the booster did not land successfully. Here we will convert failed landing with 0 and successful landing in 1.
- Perform exploratory data analysis (EDA) using visualization and SQL
- Perform interactive visual analytics using Folium and Plotly Dash
- Perform predictive analysis using classification models
  - We will perform EDA , determine Training Labels, create column for class, standardize data, split data into training and test data. After that we will apply SVM, KNN, Classification trees, Logistic Regression and will find which method is best.

Data Collection

- We have collected data using SpacX API and web scrapping described below by flowchart.



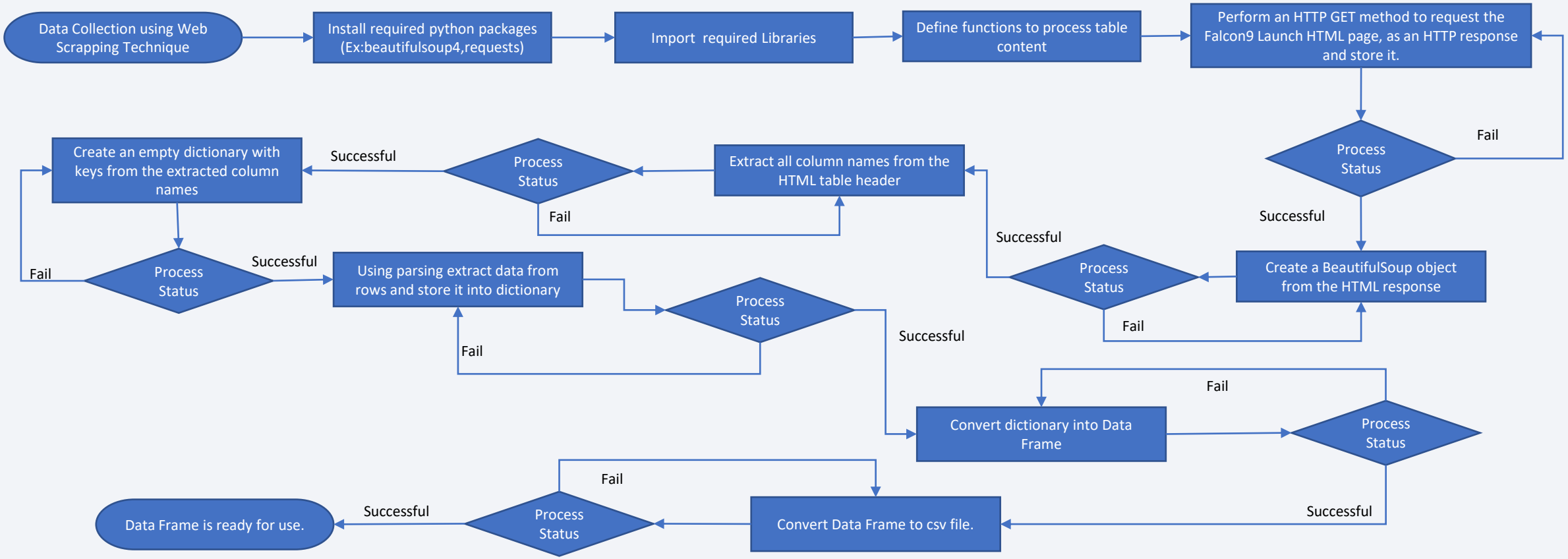
Data Collection – SpaceX API



For more details of Data Collection using SpaceX API , GitHub details are given below.

GitHub Link - https://github.com/daboprasaddutta0/Applied-Data-Science-Capstone/blob/master/Week%201/F_A_Data_Collecting.ipynb

Data Collection – Web Scraping



For more details of Data Collection using Web Scrapping Technique, GitHub details are given below.

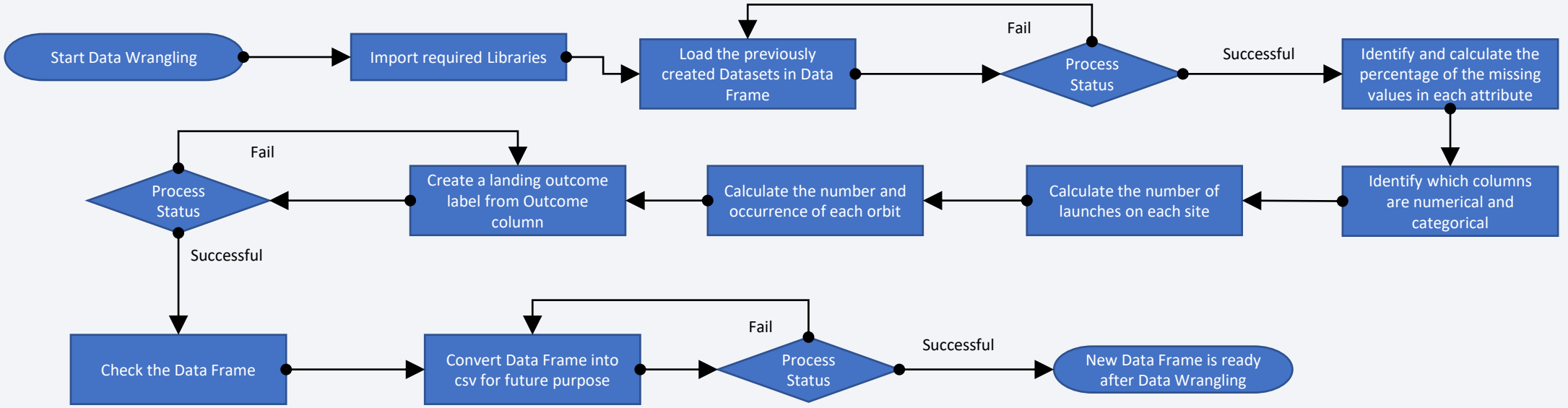
GitHub Link - https://github.com/daboprasaddutta0/Applied-Data-Science-Capstone/blob/master/Week%201/F_A_Web_Scrapping.ipynb

Data Wrangling

In data wrangling part of the project , 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.

In the data set, there are several different cases where the booster did not land successfully. Sometimes a landing was attempted but failed due to an accident; for example, True Ocean means the mission outcome was successfully landed to a specific region of the ocean while False Ocean means the mission outcome was unsuccessfully landed to a specific region of the ocean. True RTLS means the mission outcome was successfully landed to a ground pad False RTLS means the mission outcome was unsuccessfully landed to a ground pad. True ASDS means the mission outcome was successfully landed on a drone ship False ASDS means the mission outcome was unsuccessfully landed on a drone ship.

In this lab we will mainly convert those outcomes into Training Labels with 1 means the booster successfully landed 0 means it was unsuccessful.



For more details of Data Wrangling, GitHub details are given below.

GitHub Link - https://github.com/daboprasaddutta0/Applied-Data-Science-Capstone/blob/master/Week%201/F_A_Data_Wrangling.ipynb

EDA with Data Visualization

- The FlightNumber vs. PayloadMass and scatterplot overlay the outcome of the launch. It seems the more massive the payload, the less likely the first stage will return. We found that 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%.
- The FlightNumber vs. LaunchSite scatterplot. It seems the launch sites which have higher success rate have higher number of flight number.
- The PayloadMass vs. LaunchSite scatterplot. The VAFB-SLC launchsite there are no rockets launched for heavypayload mass(greater than 10000).
- Bar Chart to Visualize the relationship between success rate of each orbit type. ES-L1, GEO, HEO, SSO orbits have highest success rate among other orbits.
- The FlightNumber vs. Orbit type scatterplot. In the 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.
- The PayloadMass vs. Orbit type scatterplot. 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.
- Visualization of the launch success yearly trend. The success rate since 2013 kept increasing till 2020.

For more details of EDA with Data Visualization, GitHub details are given below.

GitHub Link - https://github.com/daboprasaddutta0/Applied-Data-Science-Capstone/blob/master/Week%202/F_A_Exploratory_Data_Analysis_Pandas_Matplotlib.ipynb

EDA with SQL

- SELECT DISTINCT LAUNCH_SITE FROM SPACEXTBL. (Display the names of the unique launch sites in the space mission).
- SELECT * FROM SPACEXTBL WHERE LAUNCH_SITE LIKE 'CCA%' LIMIT 5. (Display 5 records where launch sites begin with the string 'CCA').
- SELECT SUM(PAYLOAD_MASS__KG_) FROM SPACEXTBL WHERE CUSTOMER = 'NASA (CRS)'. (Display the total payload mass carried by boosters launched by NASA (CRS)).
- SELECT AVG(PAYLOAD_MASS__KG_) FROM SPACEXTBL WHERE BOOSTER_VERSION = 'F9 v1.1'. (Display average payload mass carried by booster version F9 v1.1)
- SELECT MIN(DATE) FROM SPACEXTBL WHERE LANDING__OUTCOME ='Success (ground pad)'. (List the date when the first successful landing outcome in ground pad was achieved.)
- SELECT BOOSTER_VERSION FROM SPACEXTBL WHERE PAYLOAD_MASS__KG_ > 4000 AND PAYLOAD_MASS__KG_<6000 AND LANDING__OUTCOME = 'Success (drone ship)'. (List the names of the boosters which have success in drone ship and have payload mass greater than 4000 but less than 6000)
- SELECT COUNT(*) AS "TOTAL NUMBER OF SUCCESSFUL AND FAILURE" FROM SPACEXTBL. (List the total number of successful and failure mission outcomes)
- SELECT DISTINCT BOOSTER_VERSION FROM SPACEXTBL WHERE PAYLOAD_MASS__KG_ = (SELECT MAX(PAYLOAD_MASS__KG_) FROM SPACEXTBL). (List the names of the booster_versions which have carried the maximum payload mass. Use a subquery)
- SELECT BOOSTER_VERSION, LAUNCH_SITE FROM SPACEXTBL WHERE DATE LIKE '2015%' AND LANDING__OUTCOME = 'Failure (drone ship)'. (List the names of the booster_versions which have carried the maximum payload mass. Use a subquery)
- SELECT BOOSTER_VERSION, LAUNCH_SITE FROM SPACEXTBL WHERE DATE LIKE '2015%' AND LANDING__OUTCOME = 'Failure (drone ship)'. (List the failed landing_outcomes in drone ship, their booster versions, and launch site names for in year 2015)
- SELECT LANDING__OUTCOME, COUNT(*) FROM SPACEXTBL WHERE DATE BETWEEN '2010-06-04' AND '2017-03-20' GROUP BY LANDING__OUTCOME ORDER BY COUNT(*) DESC . (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)

For more details of EDA with SQL, GitHub details are given below.

GitHub Link - https://github.com/daboprasaddutta0/Applied-Data-Science-Capstone/blob/master/Week%202/F_A_Exploratory_Data_Analysis_SQL.ipynb

Build an Interactive Map with Folium

- Here some map objects such as map, circles, markers, PolyLines, MarkerCluster, MousePosition, Popup are added to a folium map
- Map objects are created to pinning launch sites on the map.
- Circle objects are created to add a highlighted circle area with a text label on a specific coordinate.
- Marker objects are created to mark location on map.
- PolyLines objects created to draw line between two co-ordinates.
- MarkerCluster objects are created because it can be a good way to simplify a map containing many markers having the same coordinate.
- MousePosition objects are created to get coordinate for a mouse over a point on the map. As such, while you are exploring the map, you can easily find the coordinates of any points of interests (such as railway).
- Popup objects are created to show the name on the map.

For more details of Map with Folium, GitHub details are given below.

GitHub Link - https://github.com/daboprasaddutta0/Applied-Data-Science-Capstone/blob/master/Week%203/F_A_Data_Visualization_Folium.ipynb

Build a Dashboard with Plotly Dash

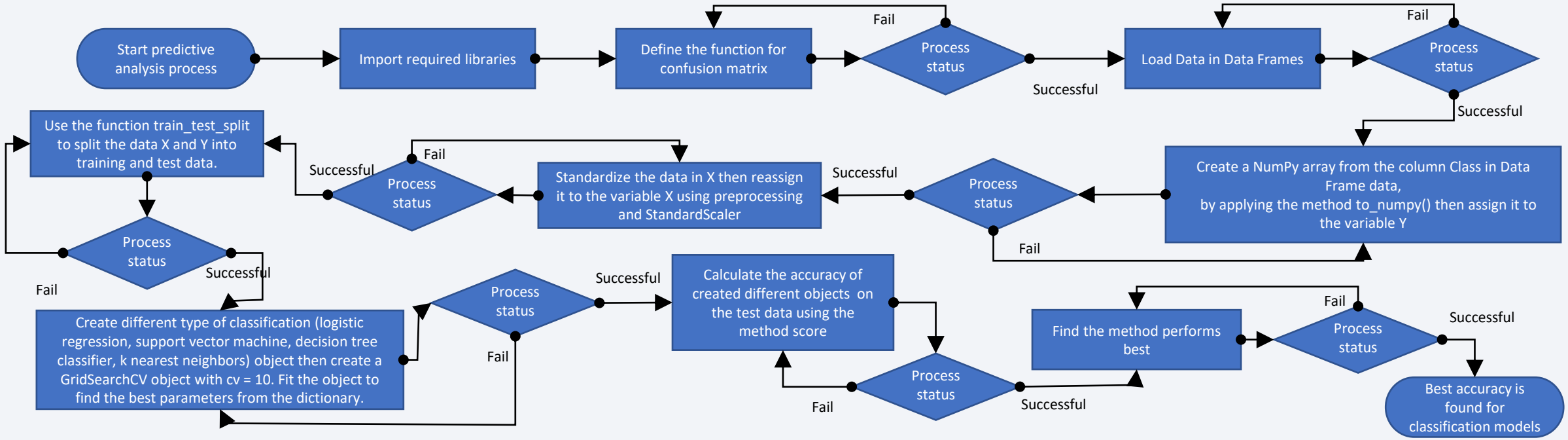
- Here we have used dropdown, pie chart, range-slider, scatter chart, two call back function for our interactive dashboard.
- In the dropdown we can select all launch sites or a specific launch sites. If we select all launch site or specific launch site then for that we can observe success rate of all launch sites in pie chart and scatter chart for Payload Mass(Kg) vs class.
- Pie chart is created for check success rate for all launch sites and success and failed rate of selected launching sites.
- Scatter chart is created to observe Payload Mass(Kg) vs class relation for all launching sites or specific launching sites.
- Range Slider is used to determine range of Payload Mass.
- Inside call back function pie and scatter chart mechanism is defined.

For more details of Dashboard with Plotly Dash, GitHub details are given below.

GitHub Link - https://github.com/daboprasaddutta0/Applied-Data-Science-Capstone/blob/master/Week%203/F_A_Hands_On_Interactive_Dashboard_with_Plotly_Dash.txt

Predictive Analysis (Classification)

- After importing required libraries and defining confusion matrix function data is loaded into Data Frame. NumPy array is created from the column Class in data, by applying the method to_numpy() then assign it to the variable Y. Standardize the data in X then reassign it to the variable X. Function train_test_split to split the data X and Y into training and test data. Classification (logistic regression, support vector machine, decision tree classifier, k nearest neighbors) object is created then GridSearchCV object with cv = 10 is created. Fit the object to find the best parameters from the dictionary. Calculate the accuracy of created different objects on the test data using the method score. At last the methods performs best are identified.



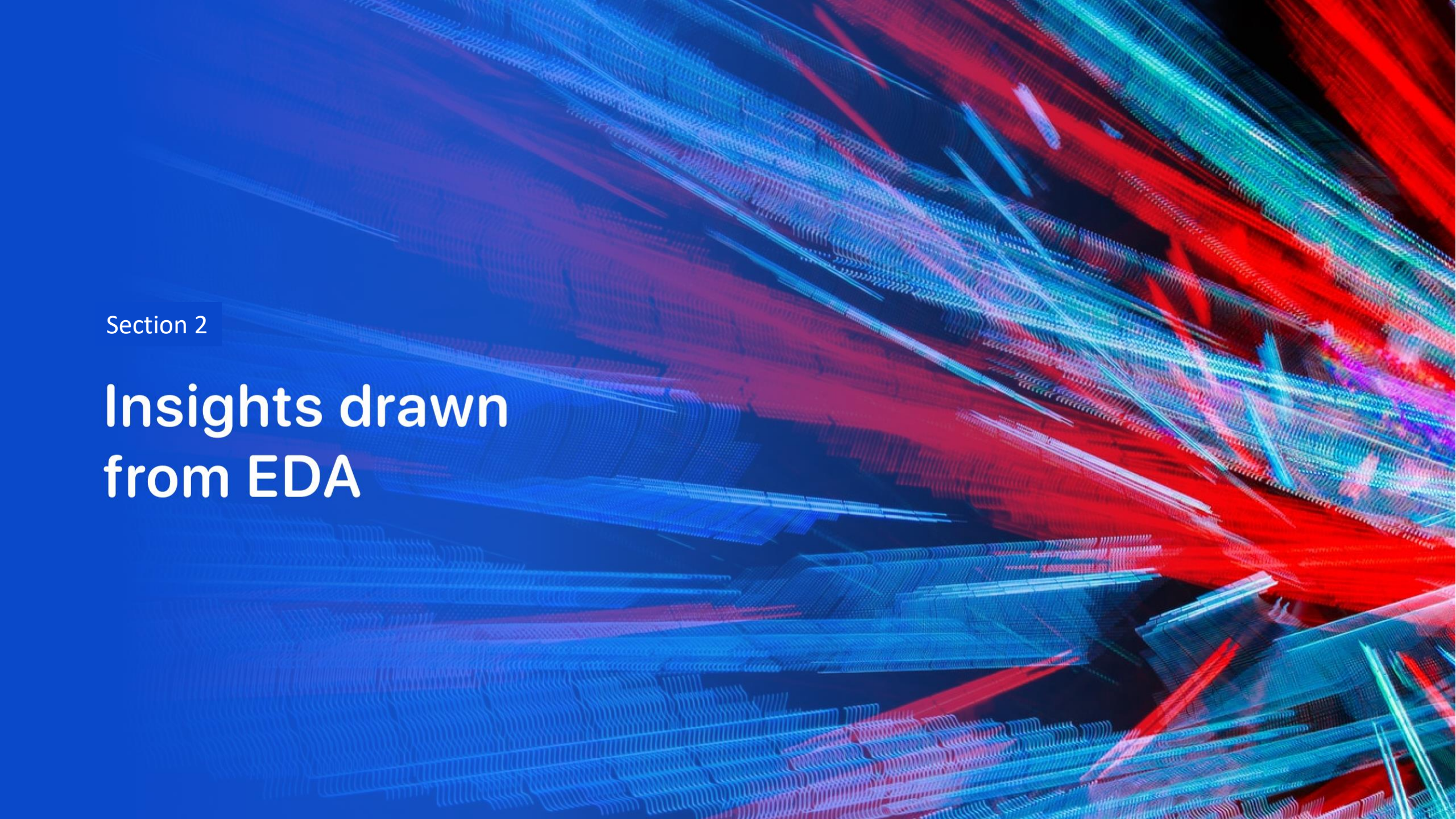
For more details of Predictive Analysis (Classification), GitHub details are given below.

GitHub Link - https://github.com/daboprasaddutta0/Applied-Data-Science-Capstone/blob/master/Week%204/F_A_Machine_Learning_Prediction.ipynb

Results

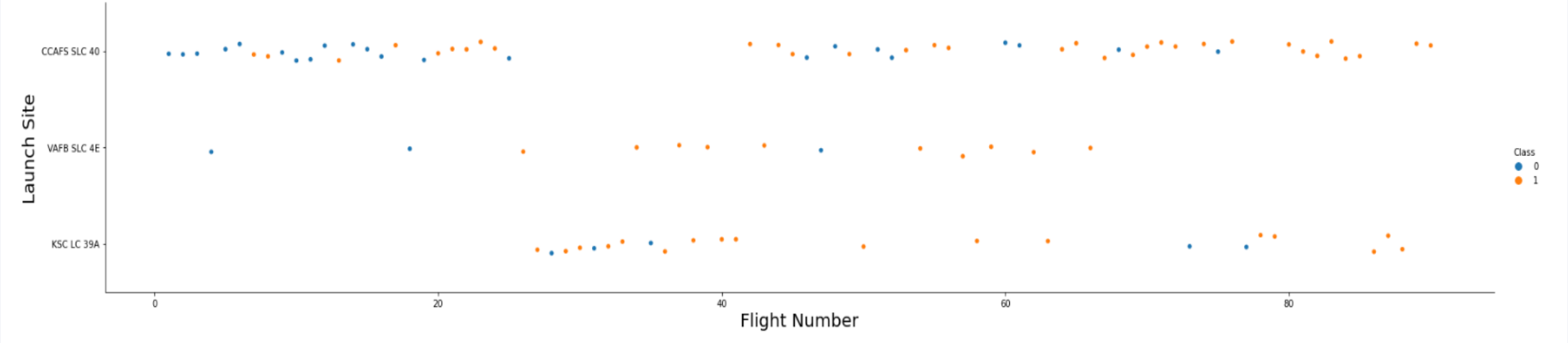
- Exploratory data analysis is done using SQL and Matplotlib. From FlightNumber vs. PayloadMass scatter plot we see that different launch sites have different success rates. CCAFS LC-40, has a success rate of 60 %, while KSC LC-39A and VAFB SLC 4E has a success rate of 77%. From FlightNumber vs LaunchSite scatter plot we found that highly successful launch sites have higher flight number. From Payload vs Launch Site scatter plot we observe Payload Vs. Launch Site scatter point chart you will find for the VAFB-SLC launch site there are no rockets launched for heavypayload mass(greater than 10000). From Orbit vs Success rate bar chart we can see that launching in orbit ES-L1, GEO, HEO, SSO have the highest rate of success. From FlightNumber vs Orbit scatter plot we see that in the 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. From Payload vs Orbit scatter plot we see that 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. From year vs success rate line chart we can observe that the sucess rate since 2013 kept increasing till 2020.

- In Predictive analysis maximum accuracy is 83.3%


Section 2
Insights drawn from EDA

Flight Number vs. Launch Site

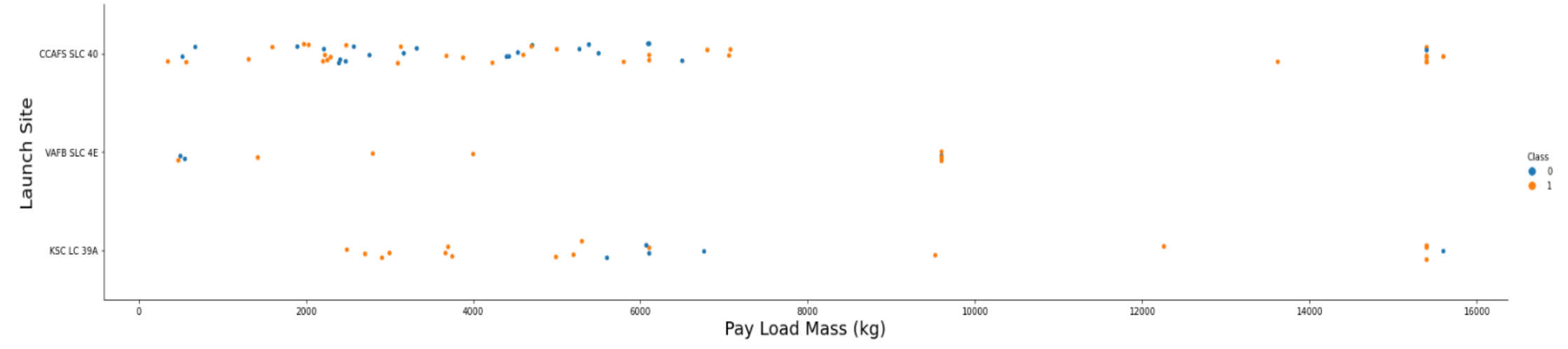
- Scatter plot of Flight Number vs. Launch Site



- We can see that launch site CCAFS SLC 40 has maximum flight number.

Payload vs. Launch Site

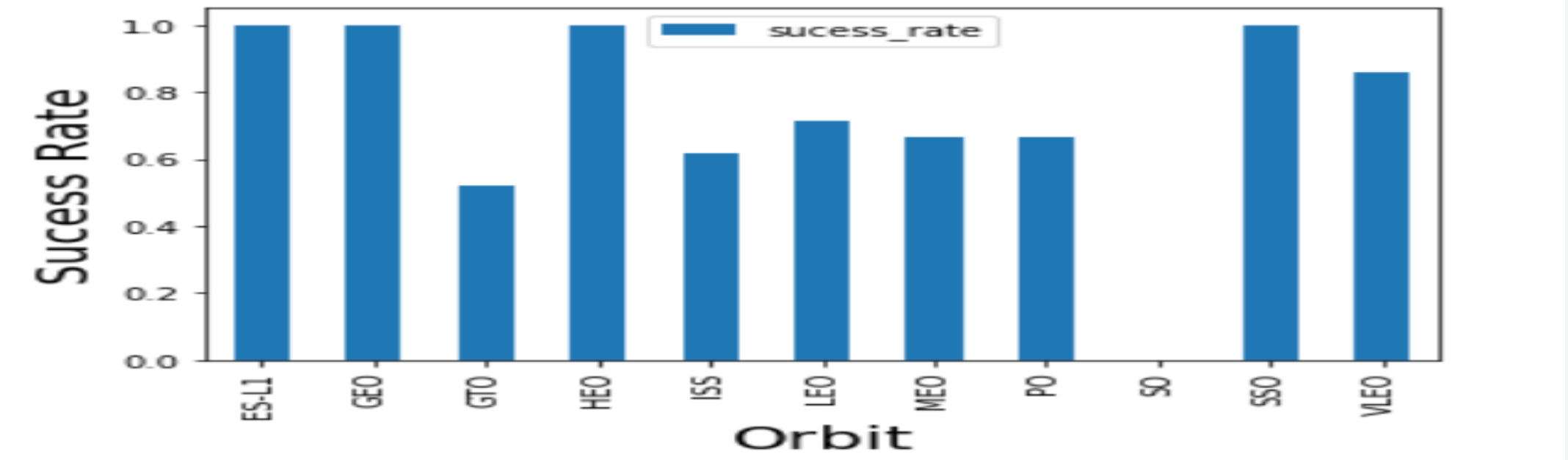
- Scatter plot of Payload vs. Launch Site



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

Success Rate vs. Orbit Type

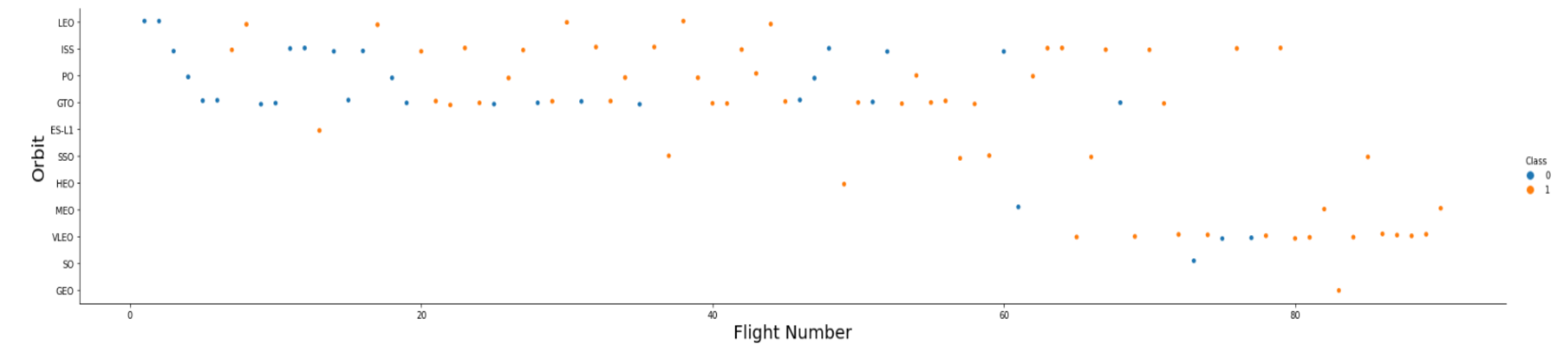
- Bar chart for the success rate of each orbit type



- Show Analyze the plotted bar chart try to find which orbits have high success rate.

Flight Number vs. Orbit Type

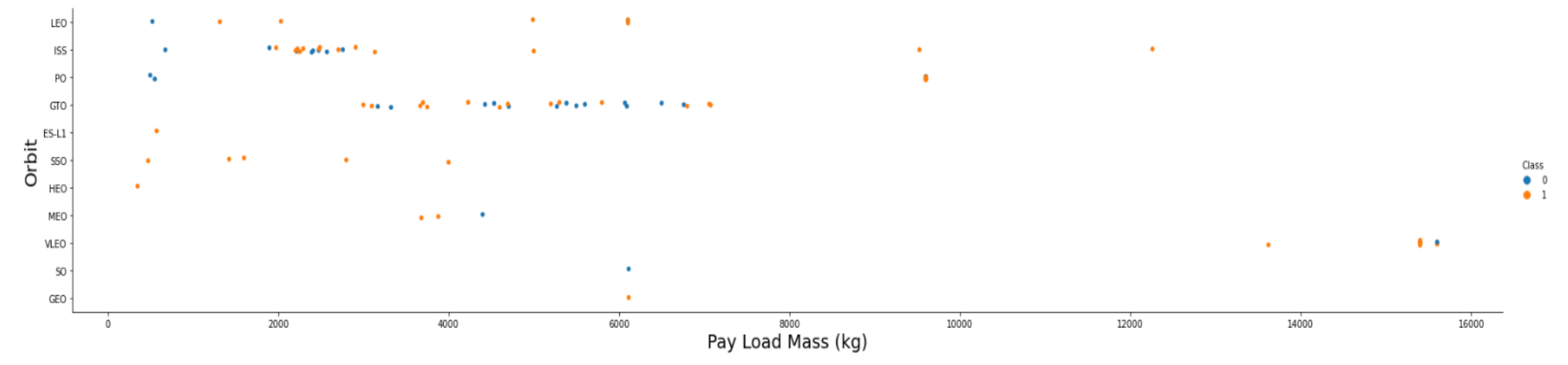
- Scater point of Flight number vs. Orbit type



- We can see that in the 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.

Payload vs. Orbit Type

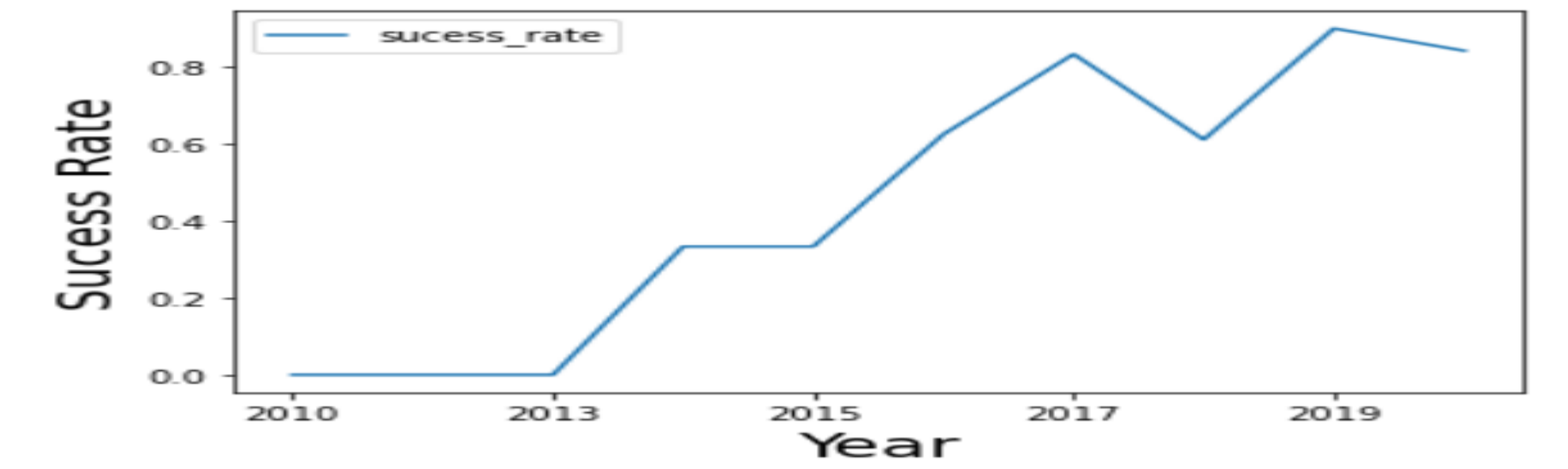
- Scater point of payload vs. orbit type



- 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) are both there here.

Launch Success Yearly Trend

- Line chart of yearly average success rate



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

All Launch Site Names

- Find the names of the unique launch sites

```
In [5]: %sql SELECT DISTINCT LAUNCH_SITE FROM SPACEXTBL
```

```
 * ibm_db_sa://zlm91362:***@824dfd4d-99de-440d-9991-629c01b3832d.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:30119/bludb
Done.
```

Out[5]:

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

- By this Select statement we are able to find unique launch site name from SPACEXTBL table.

Launch Site Names Begin with 'CCA'

- Find 5 records where launch sites begin with `CCA`.

In [33]:
```
%sql SELECT *  FROM SPACEXTBL WHERE LAUNCH_SITE LIKE 'CCA%' LIMIT 5
```

```
 * ibm_db_sa://zlm91362:***@824dfd4d-99de-440d-9991-629c01b3832d.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:30119/bludb
Done.
```

Out[33]:

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

- In this select statement we use like operator to identifies records where launch site starts with ‘CCA’ in SPACEXTBL table.

Total Payload Mass

- Calculate the total payload carried by boosters from NASA

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

In [7]:
```
%sql SELECT SUM(PAYLOAD_MASS__KG_)  FROM SPACEXTBL WHERE CUSTOMER = 'NASA (CRS)'
```

```
 * ibm_db_sa://zlm91362:***@824dfd4d-99de-440d-9991-629c01b3832d.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:30119/bludb
Done.
```

Out[7]:

1
45596

- In this select statement we have used simple where clause and aggregate function SUM() to find relevant records from table SPACEXTBL table.

Average Payload Mass by F9 v1.1

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

```
In [8]: %sql SELECT AVG(PAYLOAD_MASS__KG_)  FROM SPACEXTBL WHERE BOOSTER_VERSION = 'F9 v1.1'

 * ibm_db_sa://zlm91362:***@824dfd4d-99de-440d-9991-629c01b3832d.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:30119/bludb
Done.

Out[8]:   1

          2928
```

- In this select statement we have used simple where clause and aggregate function AVG() to find relevant records from table SPACEXTBL table.

First Successful Ground Landing Date

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

```
In [12]: %sql SELECT MIN(DATE)  FROM SPACEXTBL WHERE LANDING__OUTCOME ='Success (ground pad)'

 * ibm_db_sa://zlm91362:***@824dfd4d-99de-440d-9991-629c01b3832d.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:30119/bludb
Done.

Out[12]:        1

2015-12-22
```

- In this select statement we have used simple where clause and aggregate function AVG() to find relevant records from table SPACEXTBL table.

Successful Drone Ship Landing with Payload between 4000 and 6000

- List the names of boosters which have successfully landed on drone ship and had payload mass greater than 4000 but less than 6000

```
In [13]: %sql SELECT BOOSTER_VERSION  FROM SPACEXTBL WHERE PAYLOAD_MASS__KG_ > 4000 AND PAYLOAD_MASS__KG_<6000 AND LANDING__OUTCOME = 'Success (drone ship)'
 * ibm_db_sa://zlm91362:***@824dfd4d-99de-440d-9991-629c01b3832d.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:30119/bludb
Done.
```

Out[13]:

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

- In this select statement we have used simple where clause ; less than (<), greater than (>), and operator to find relevant records from table SPACEXTBL table.

Total Number of Successful and Failure Mission Outcomes

- Calculate the total number of successful and failure mission outcomes

```
In [17]: %sql SELECT  COUNT(*) AS "TOTAL NUMBER OF SUCCESSFUL AND FAILURE" FROM SPACEXTBL
 * ibm_db_sa://zlm91362:***@824dfd4d-99de-440d-9991-629c01b3832d.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:30119/bludb
Done.
Out[17]: TOTAL NUMBER OF SUCCESSFUL AND FAILURE
                                        101
```

- In this select statement we have used count() and aliasing technique to find relevant records from table SPACEXTBL table.

Boosters Carried Maximum Payload

- List the names of the booster which have carried the maximum payload mass

```
In [19]: %sql SELECT DISTINCT BOOSTER_VERSION FROM SPACEXTBL WHERE PAYLOAD_MASS__KG_ = (SELECT  MAX(PAYLOAD_MASS__KG_)  FROM SPACEXTBL)
 * ibm_db_sa://zlm91362:***@824dfd4d-99de-440d-9991-629c01b3832d.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:30119/bludb
Done.
```

Out[19]:

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

- Here we have used subquery and inside subquery aggregate function MAX() is used to find maximum payload which we have used in main query to find out the booster version carried out maximum payload from SPACEXTBL table.

2015 Launch Records

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

```
In [20]: %sql SELECT BOOSTER_VERSION, LAUNCH_SITE FROM SPACEXTBL WHERE DATE LIKE '2015%' AND LANDING__OUTCOME = 'Failure (drone ship)'
 * ibm_db_sa://zlm91362:***@824dfd4d-99de-440d-9991-629c01b3832d.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:30119/bludb
Done.
```

Out[20]:

booster_version	launch_site
F9 v1.1 B1012	CCAFS LC-40
F9 v1.1 B1015	CCAFS LC-40

- In this select statement we use like operator to identifies records where year is ‘2015’ and landing outcome is ‘Failure (drone ship)’ in SPACEXTBL table.

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

```
In [32]: %sql SELECT LANDING__OUTCOME, COUNT(*) FROM SPACEXTBL WHERE DATE BETWEEN '2010-06-04' AND '2017-03-20' GROUP BY LANDING__OUTCOME
         ORDER BY COUNT(*) DESC
```

```
 * ibm_db_sa://zlm91362:***@824dfd4d-99de-440d-9991-629c01b3832d.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:30119/bludb
Done.
```

Out[32]:

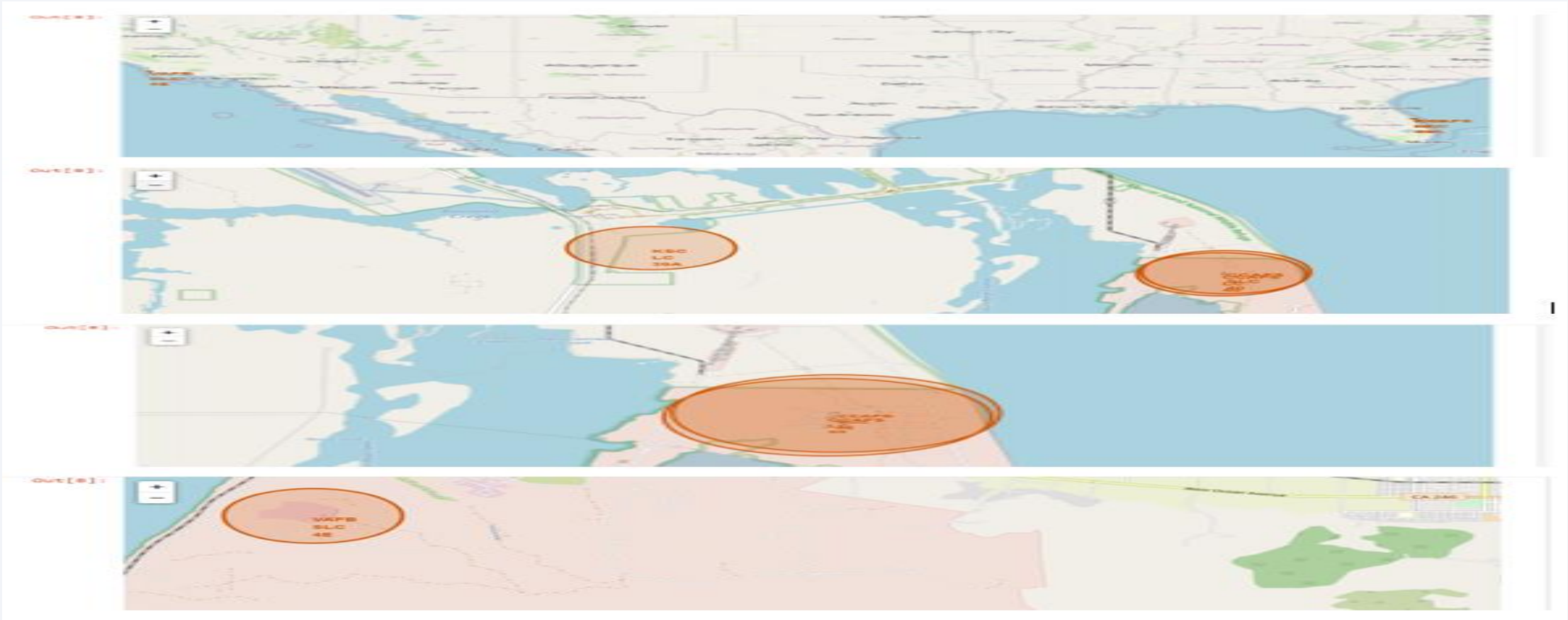
landing__outcome	2
No attempt	10
Failure (drone ship)	5
Success (drone ship)	5
Controlled (ocean)	3
Success (ground pad)	3
Failure (parachute)	2
Uncontrolled (ocean)	2
Precluded (drone ship)	1

- In this select statement we have used count() function; between, and operator in where clause; group by and order by desc to identifies relevant records from SPACEXTBL table.

Section 3

Launch Sites Proximities Analysis

All Launch Sites in Folium Map



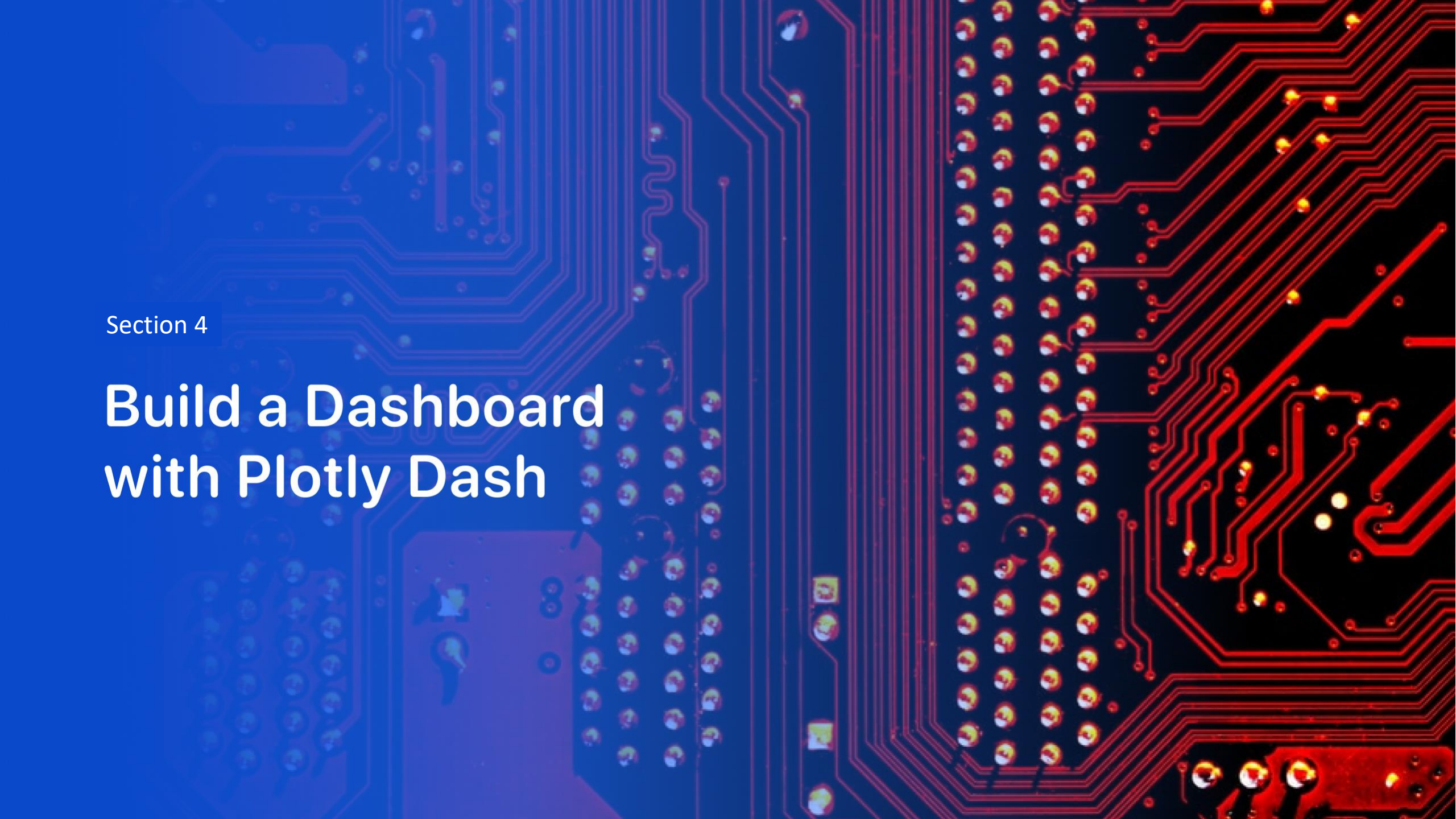
- All successful launching sites are nearer to Ecuador.

Folium Map With color-labeled launch outcomes on the map

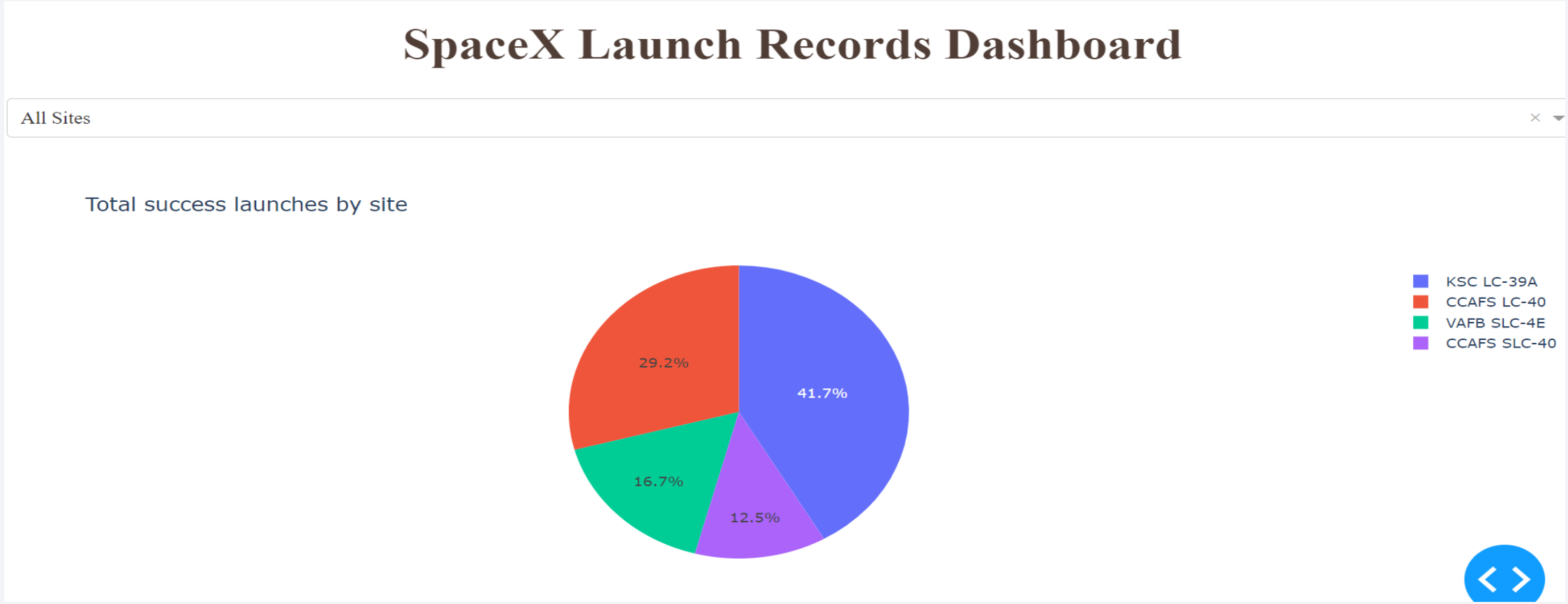
- We can observe that site KSC LC 39A have highest number of green markers.

Folium Map for launch site ‘CCAFS SLC 40’ to its proximities such as railway, highway, coastline, with distance calculated and displayed

- In the 1st figure nearest coastline distance from launch site CCAFS SLC 40 is 0.86 Km. In the 2nd figure nearest highway distance from launch site CCAFS SLC 40 is 0.58 Km. In the 3rd and 4th figure nearest railway distance from launch site CCAFS SLC 40 is 1.28 Km. In the 5th and 6th figure nearest city distance from launch site CCAFS SLC 40 is 51.43 Km

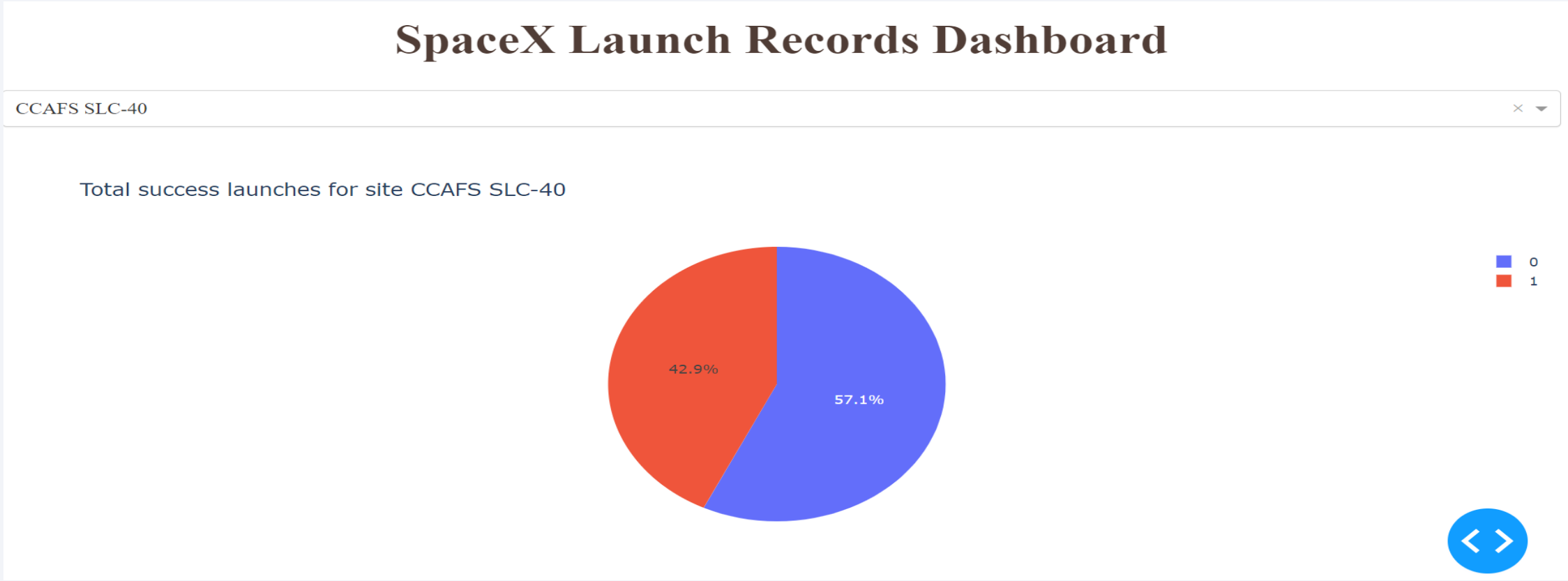

Section 4
Build a Dashboard with Plotly Dash

Launch success count for all sites, in a pie chart



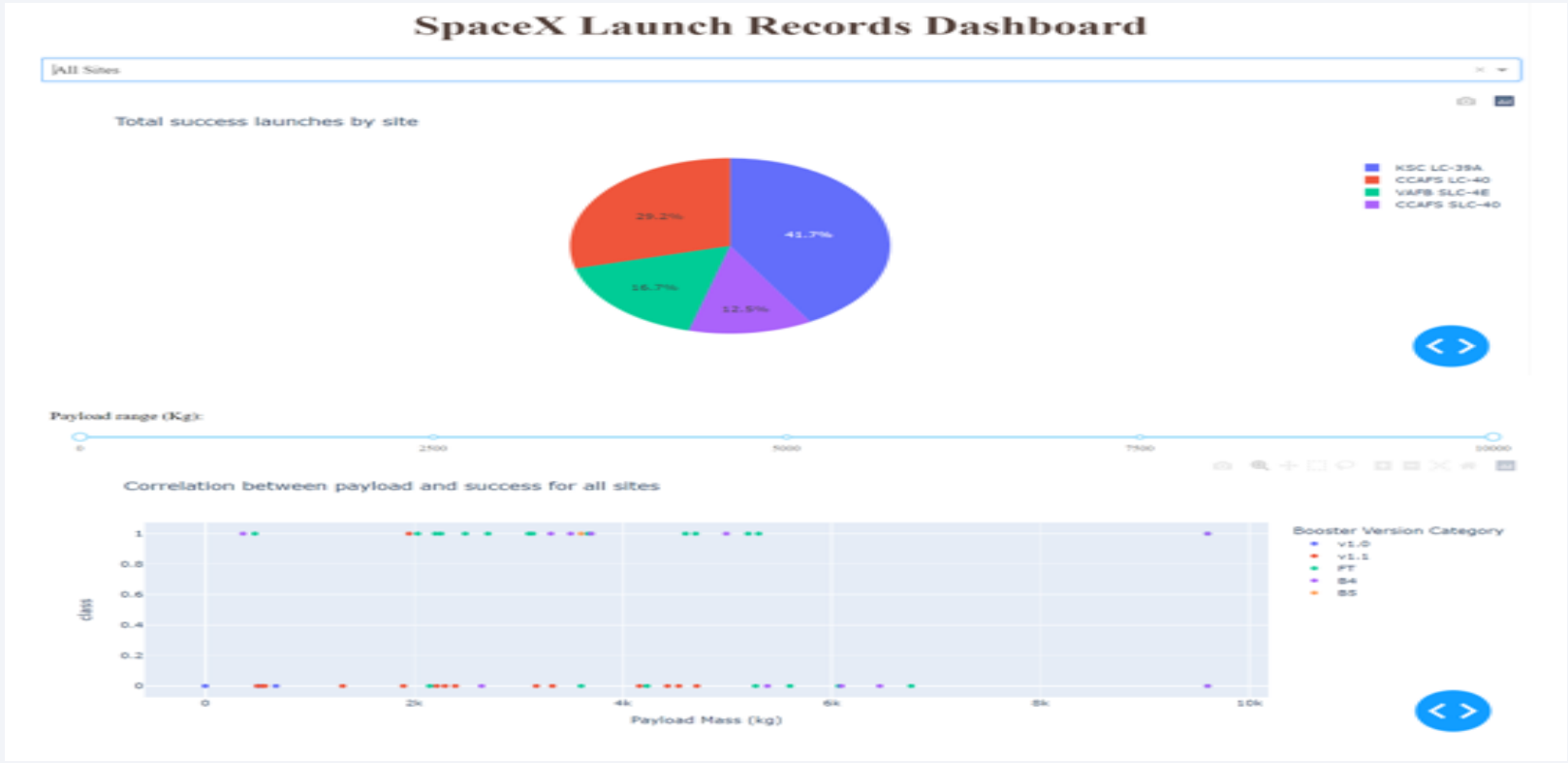
- In the above pie chart success rate for all sites is shown

The pie chart for the launch site with highest launch success ratio



- Launch site CCAFS SLC-40 has the highest launch success ratio.

Payload vs. Launch Outcome scatter plot for all sites, with different payload selected in the range slider

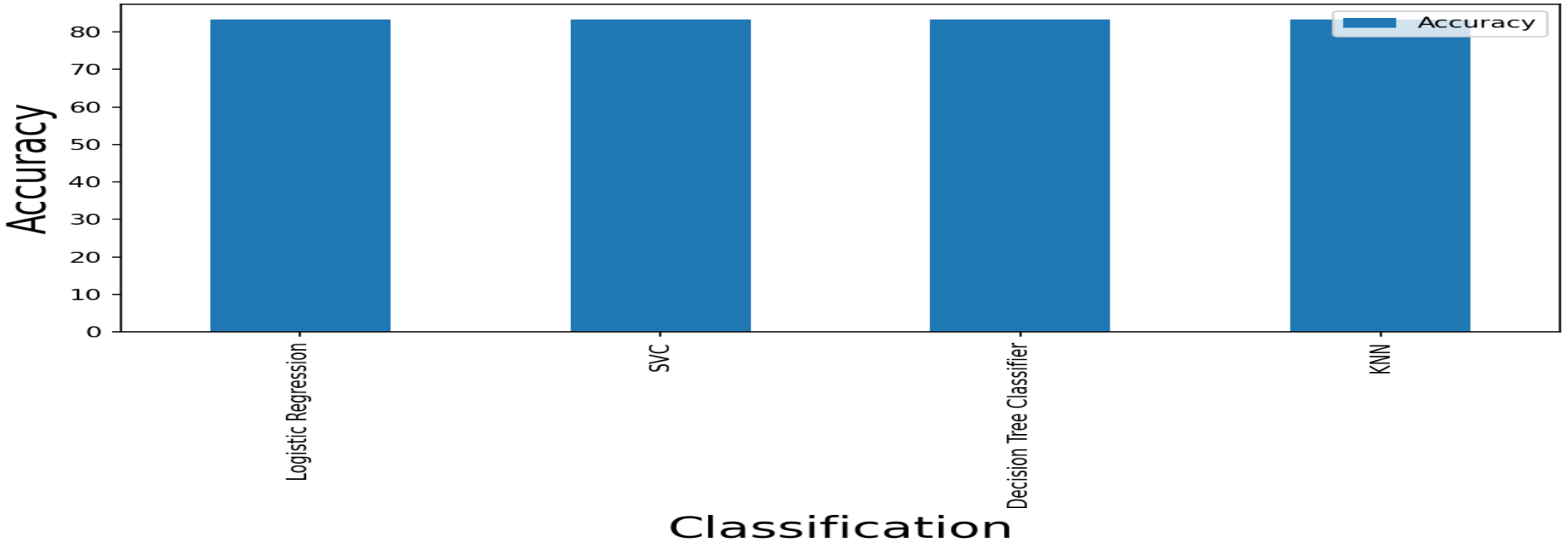


- In payload range 2k to 4k and booster version ‘FT’ have the largest success rate. e

Section 5

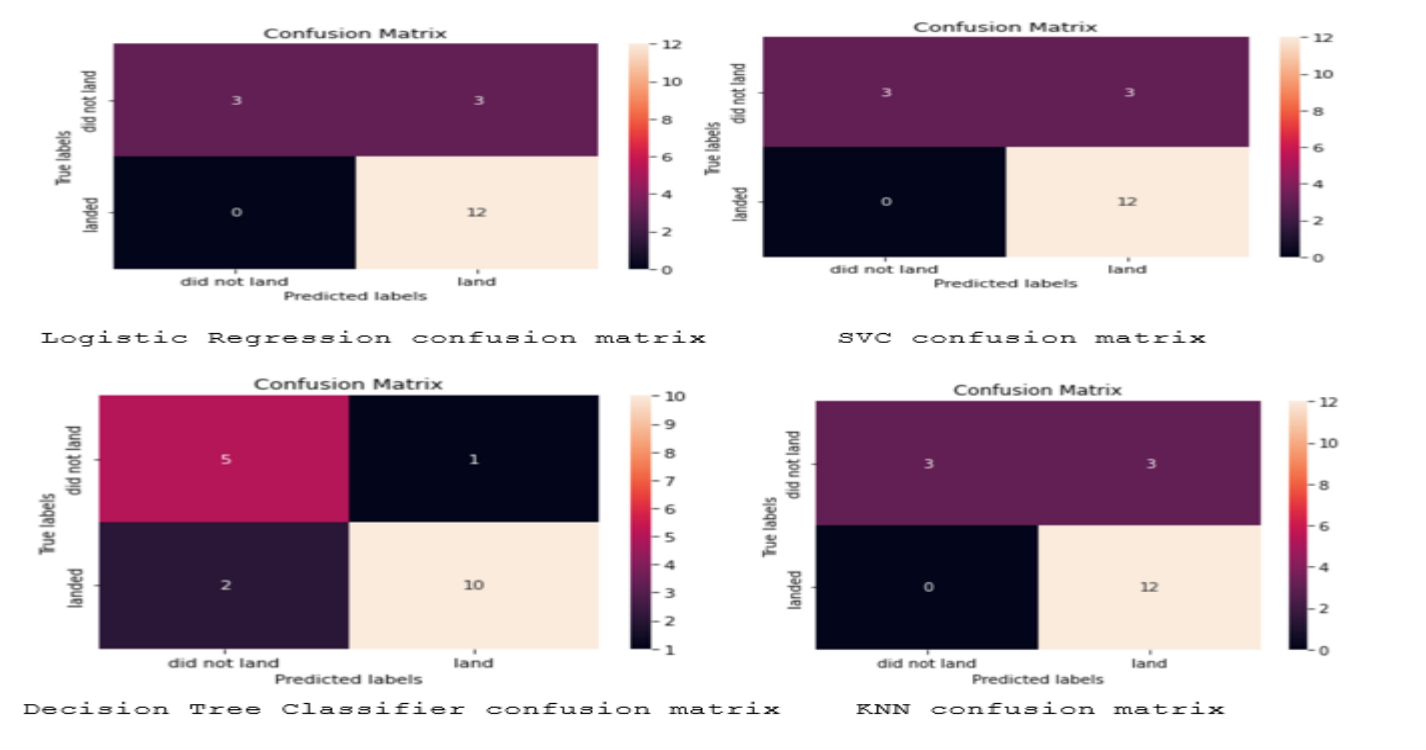
Predictive Analysis (Classification)

Classification Accuracy



- All models have same accuracy for test data that is 83.3 %.

Confusion Matrix



- Since accuracy of all model with test data are same hence we have shown confusion matrix for all.

Conclusions

- Falcon 9 Falcon 9 rocket launches on its website with a cost of 62 million dollars; other providers cost upward of 165 million dollars each, much of the savings is because Space X can reuse the first stage.
- Launch sites are in close proximity to Ecuador to minimize fuel consumption by using Earth's ~ 30km/sec eastward spin to help spaceships get into orbit.
- Launch sites are in close proximity to coastline so they can fly over the ocean during launch, for at least two safety reasons - (1) crew has option to abort launch and attempt water landing, (2) minimize people and property at risk from falling debris.
- Launch sites are in close proximity to highways, which allows for easily transport required people and property.
- Launch sites are in close proximity to railways, which allows transport for heavy cargo. Launch sites are not in close proximity to cities, which minimizes danger to population dense areas.
- Launch sites are not nearer to any city for example Melbourne city from launch site CCAFS SLC 40 is 51.43 Km away.

Appendix

- SPACEXTBL table is created in IBM DB 2.
- Links for created Data Sets.
- Data Link1 - https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DS0321EN-SkillsNetwork/datasets/dataset_part_1.csv
- Data Link 2 - https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DS0321EN-SkillsNetwork/datasets/dataset_part_2.csv
- Data Link 3 - https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DS0321EN-SkillsNetwork/datasets/dataset_part_3.csv
- Link for SQL Query - https://github.com/daboprasaddutta0/Applied-Data-Science-Capstone/blob/master/Week%202/F_A_Exploratory_Data_Analysis_SQL.ipynb

Thank you!