IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Mariusz Ochla

Sep 18th, 2022

Outline

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

Executive Summary

Prediction of SpaceX Successful Landing with Machine Learning and Predictive Analytics
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.
Determination of the first stage landing success allows to predict the total cost of the launch.

- Summary of methodologies
  - **Data Collection**
    The data has been collected from: SpaceX Launch data using SpaceX API and Historical launch records from Wikipedia page
  - **Exploratory Data Analysis and Interactive Data Analysis**
    Data Analysis allowed to identify launch parameters that are key for the mission success.
  - **Predictive Analysis using Machine Learning** - to train the model so that the prediction could be done
- Summary of all results
  - Trained Machine Learning model **predicts Launch Success with 88% accuracy.**

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.

If we can determine whether 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 investigated with the Data Analysis and Machine Learning

Is it possible to predict if the Falcon 9 first stage lands successfully?

What are the key success parameters of the mission?

What Machine Learning classification model works the best for the given problem?

What is the accuracy of the prediction?

Section 1

Methodology

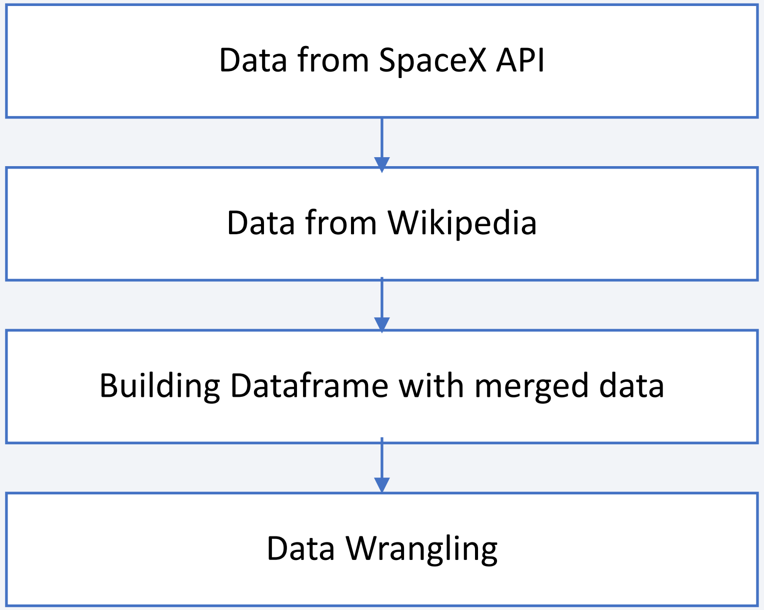
Methodology

Executive Summary

- **Data collection methodology**
  - SpaceX launch data has been collected using SpaceX API
  - Historical launch records has been collected with web-scrapping from a Wikipedia page - “List of Falcon 9 and Falcon Heavy launches.”
  - The data of 90 launches has been collected, merged and analysed.
- **Data wrangling**
  - Missing values of the payload has been replaced by the mean value.
  - Different launch sites, different orbits and the mission success has been merged in the dataset.
- **Exploratory data analysis** (EDA) using visualisation and SQL, Interactive visual analysis using Folium and Plotly Dash have been applied to find patterns in the data.
- **Predictive analysis using classification models**
  - Logistic regression, Support Vector Machines, Decision Tree and K-NN have been used to train and optimise the model
  - GridSearchCV has been used to optimise the model parameters

Data Collection

- Launch data has been collected from SpaceX site API
    - spacex_url="https://api.spacexdata.com/v4/launches/past"
- Historical launch records has been collected with web-scrapping from a Wikipedia page -
    - "List of Falcon 9 and Falcon Heavy launches."



https://github.com/mariuszochla/Applied_DS_Capstone

Data Collection – SpaceX API

- Data collection with SpaceX REST calls

Requesting rocket launch data from SpaceX API

```
spacex_url="https://api.spacexdata.com/v4/launches/past"
response = requests.get(spacex_url)
data = pd.json_normalize( response.json(), max_level=1 )
```

Requesting again for specific data about **rocket, payload, launchpad, and cores**

```
requests.get("https://api.spacexdata.com/v4/rockets/"+str(x)).json()
getBoosterVersion(data)
getPayloadData(data)
getLaunchSite(data)
getCoreData(data)
```

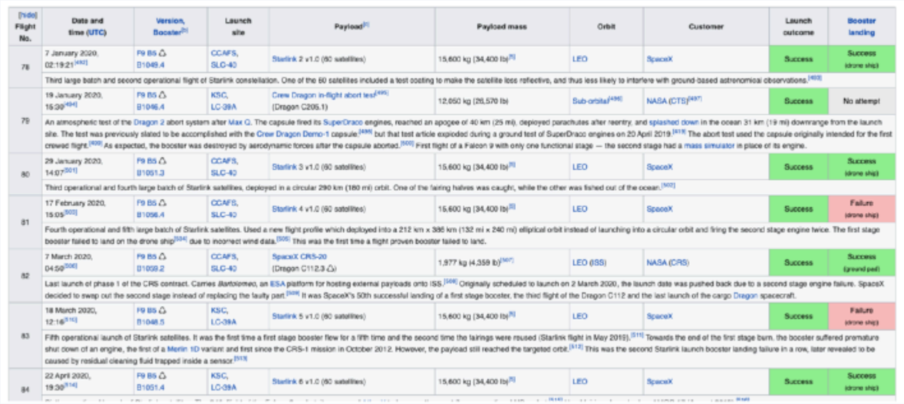
Pandas DataFrame

https://github.com/mariuszochla/Applied_DS_Capstone

Data Collection - Scraping

- Wikipedia page

  List of Falcon 9 and Falcon Heavy launches.



Requesting page content from Wikipedia

`https://en.wikipedia.org/w/index.php?title=List_of_Falcon_9_and_Falcon_Heavy_launches&`

↓

Parsing with BeautifulSoup

```
BeautifulSoup(response, 'html.parser')
```

↓

Building a Pandas Dataframe

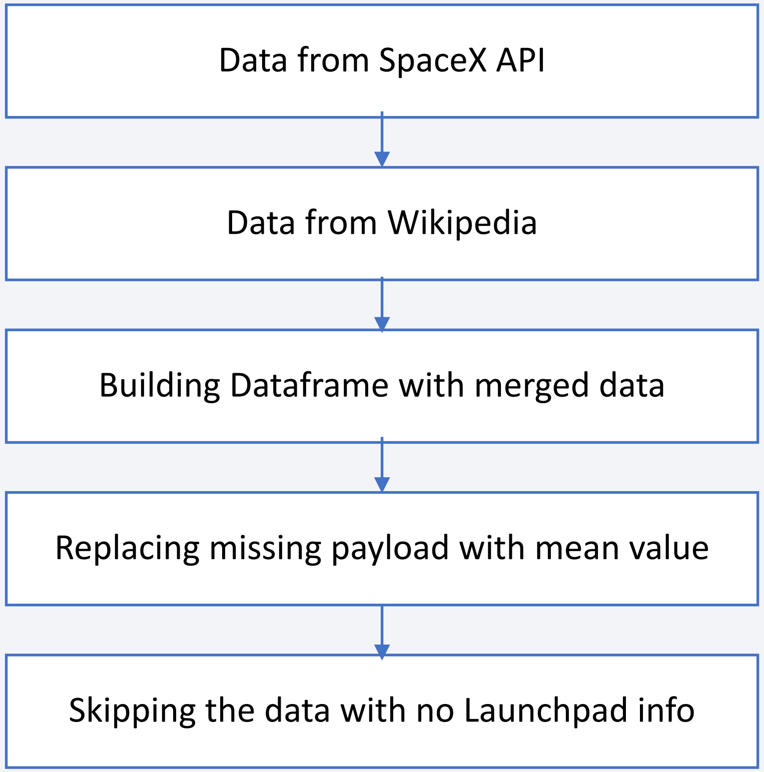
Rearranging data into columns

```
'Flight No.', 'Date and time ( )', 'Launch site', 'Payload', 'Payload mass', 'Orbit', 'Customer', 'Launch outcome
```

https://github.com/mariuszochla/Applied_DS_Capstone

Data Wrangling

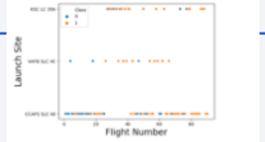
- Data collected with SpaceX API and obtained from Wikipedia page has been merged
- Missing values of the payload mass has been replaced by the mean value.
- Missing values of Launchpad has been omitted as the Launch site is crucial for the mission success



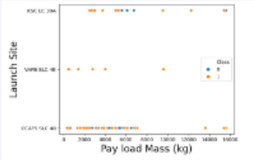
https://github.com/mariuszochla/Applied_DS_Capstone

EDA with Data Visualization

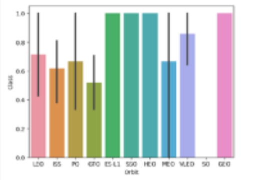
- Scatter plot to check the dependency of success rate on Flight Number and Launch Site



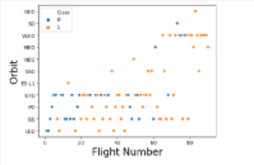
- Scatter plot to check the dependency of success rate on Payload and Launch Site



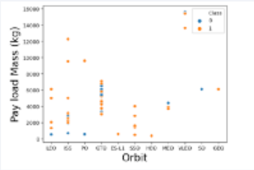
- Bar plot to check the mean success rate for each orbit type



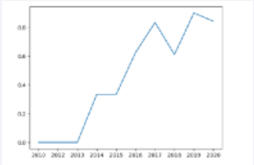
- Scatter plot to check the dependency of success rate on the Orbit Type



- Scatter plot to check the dependency of success rate on the Payload Mass



- Line plot to observe the improvement (growing success rate over time)



https://github.com/mariuszochla/Applied_DS_Capstone

EDA with SQL

Summary the SQL queries

```
%sql select distinct LAUNCH_SITE from SPACEXTBL
```

```
%sql select * from SPACEXTBL where substring(LAUNCH_SITE,1,3)=='CCA' limit 5
```

```
%sql select * from SPACEXTBL where substring(LAUNCH_SITE,1,3)=='CCA' limit 5
```

```
%sql select avg(PAYLOAD_MASS__KG_) from SPACEXTBL where substring(BOOSTER_VERSION, 1,7)="F9 v1.1"
```

```
%sql select min(DATE) from SPACEXTBL where MISSION_OUTCOME="Success"
```

```
%sql select BOOSTER_VERSION from SPACEXTBL where PAYLOAD_MASS__KG_ >4000 and PAYLOAD_MASS__KG_<6000 and [Landing _OUTCOME] like '%drone%' and MISSION_OUTCOME like '%Success%'
```

```
%sql select count(*) as FAILURE from SPACEXTBL where [Landing _OUTCOME] not like '%Success%'
```

```
%sql select BOOSTER_VERSION from SPACEXTBL where PAYLOAD_MASS__KG_ = (select max(PAYLOAD_MASS__KG_) from SPACEXTBL)
```

```
%sql select substr(DATE,4,2) as MONTH, substr(DATE,7,4) as YEAR, BOOSTER_VERSION, LAUNCH_SITE, [LANDING _OUTCOME] from SPACEXTBL where [LANDING _OUTCOME] not like '%Success%' and YEAR='2015'
```

```
%sql select [LANDING _OUTCOME] as LO, DATE, BOOSTER_VERSION, LAUNCH_SITE from SPACEXTBL where DATE>"04-06-2010" and DATE<"20-03-2017" order by LO, DATE DESC
```

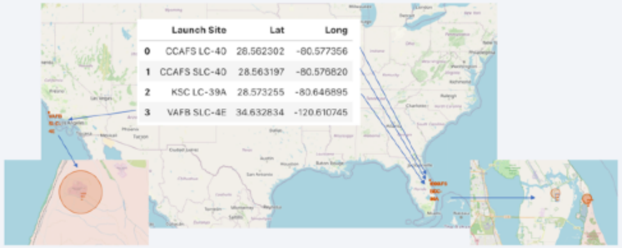
```
%sql select count(*) as SUCCESS from SPACEXTBL where [Landing _OUTCOME] like '%Success%'
```

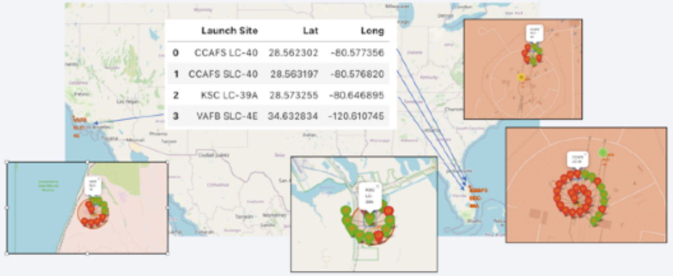
```
%sql select [LANDING _OUTCOME] as Outcome_Type, count(*) as Attempts from SPACEXTBL where DATE>"04-06-2010" and DATE<"20-03-2017" group by Outcome_Type order by Outcome_Type, DATE DESC
```

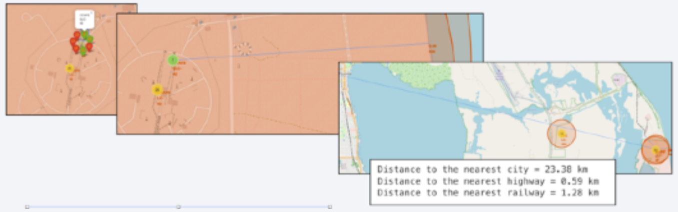
https://github.com/mariuszochla/Applied_DS_Capstone

Build an Interactive Map with Folium

- During the Interactive Analysis with Folium the following has been performed:
  - **Launch Sites locations marked on the map** - in order to analyse whether the geographical location is important for the success of a mission
  - **Success Rate calculated for each Launch Site** indicates a dependency of the success rate and the location of the site
  - **Proximities to the nearest coastline, city, highway and railway** has been calculated to determine their impact to the success of the missions

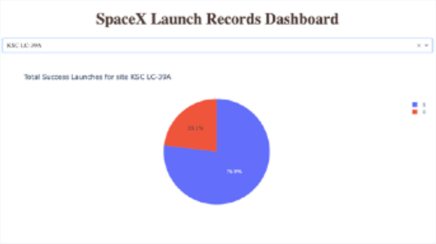


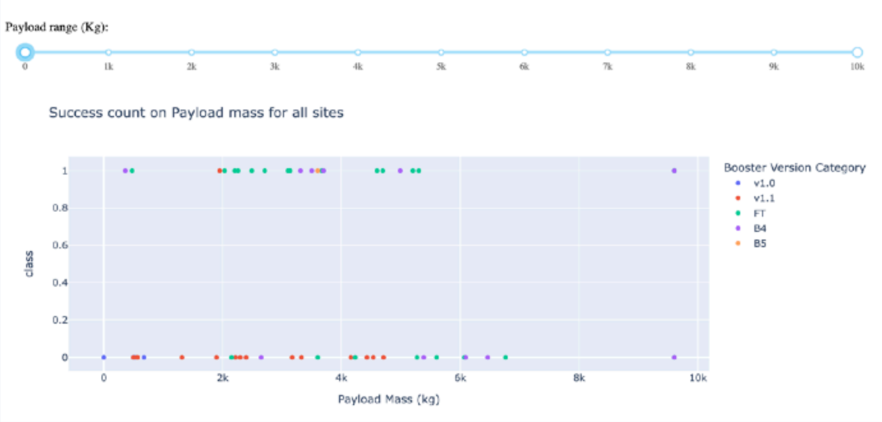




Build a Dashboard with Plotly Dash

- During the Interactive Analysis with Plotly Dashboard the following has been performed:
    - **SpaceX Launch Records Dashboard** has been built. It allowed to identify which Launch Site has the best Success Rate.
    - **Interactive Pie Charts** for Success Rate analysis
    - **Interactive Scatter Plots** for analysis of how the success of the mission depends on Payload mass and Launch Site

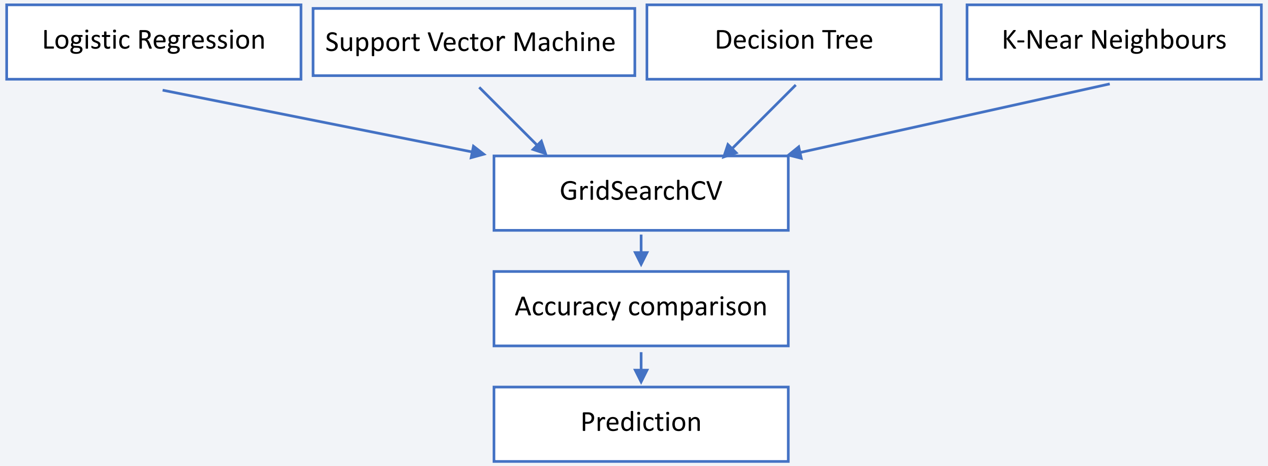




https://github.com/mariuszochla/Applied_DS_Capstone

Predictive Analysis (Classification)

- The following classifiers have been evaluated:
  - Logistic Regression
  - Support Vector Machine
  - Decision Tree
  - K-Near Neighbours

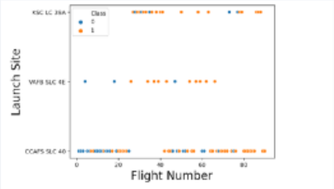


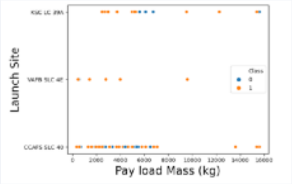
- The optimal hyper parameters for each classifier has been found using GridSearchCV with cross validation generator of 10

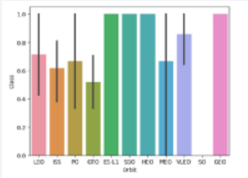
https://github.com/mariuszochla/Applied_DS_Capstone

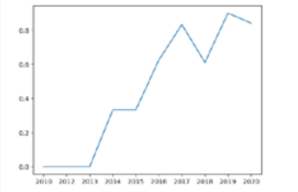
Results

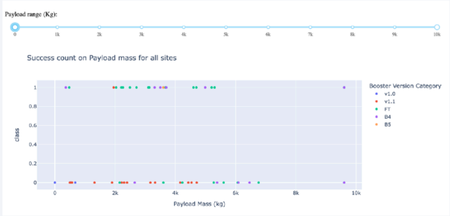
- Exploratory data analysis results
  - Significant improvement in the mission results can be observed since 2013
  - Launch site, orbit type and the payload are important factors of success
  - Booster version play an important role especially for high range of payload values
- Interactive analytics demo in screenshots
  - The launch site KSC LC-39A was used in 41.7% of successful mission.
  - The payload range 3-4k kg has the biggest success rate and the boosters FT and B4 performed the best
- Predictive analysis results
  - The best performing classifier is Decision Tree
  - The trained model performs with the accuracy of 0.88







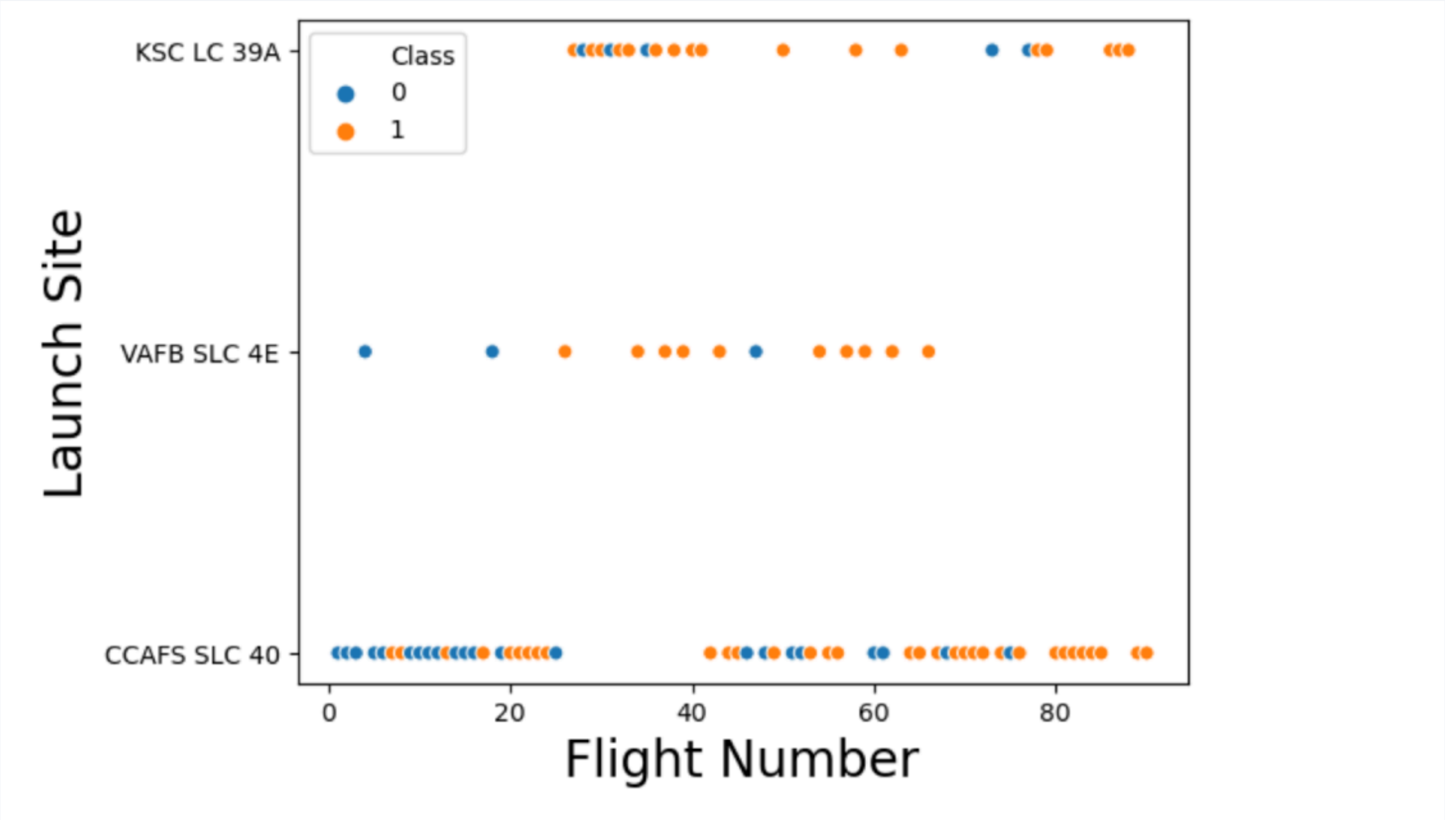




Section 2

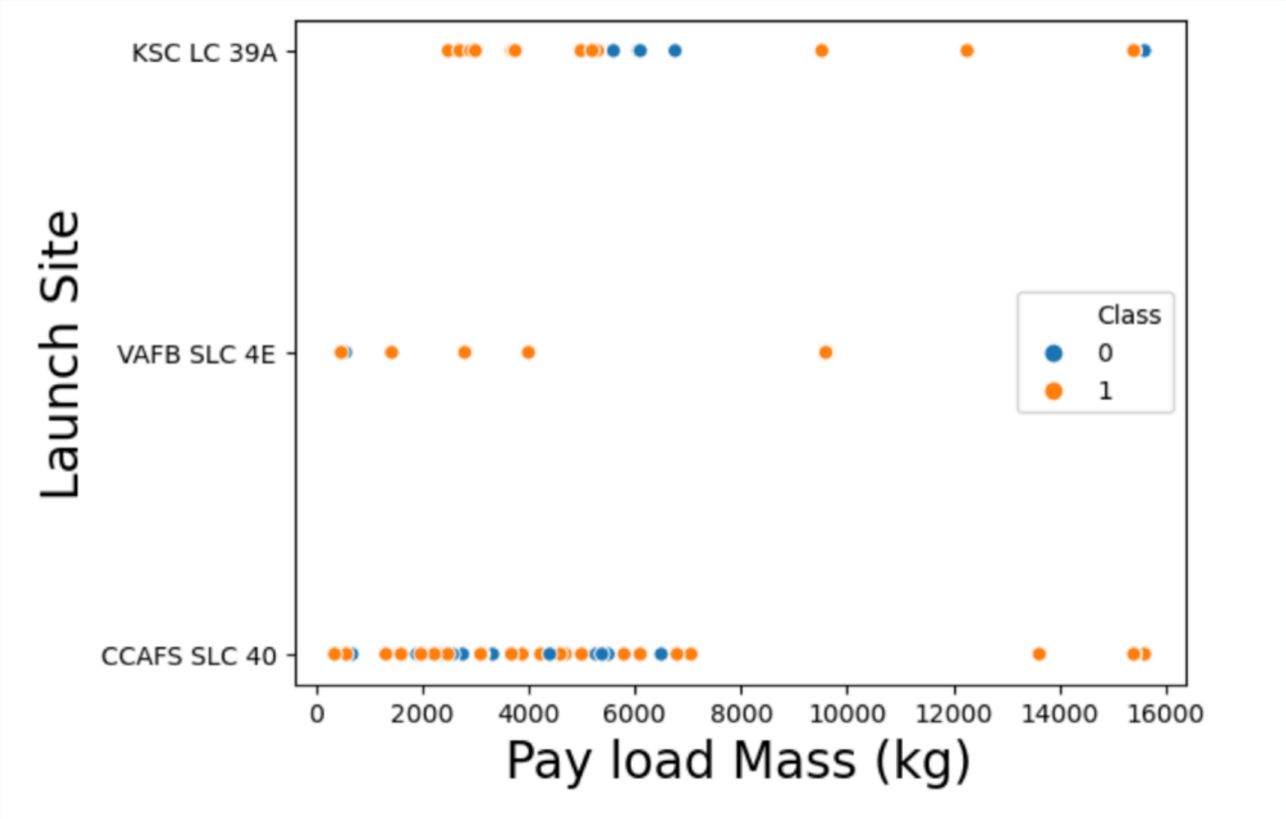
Insights drawn from EDA

Flight Number vs. Launch Site



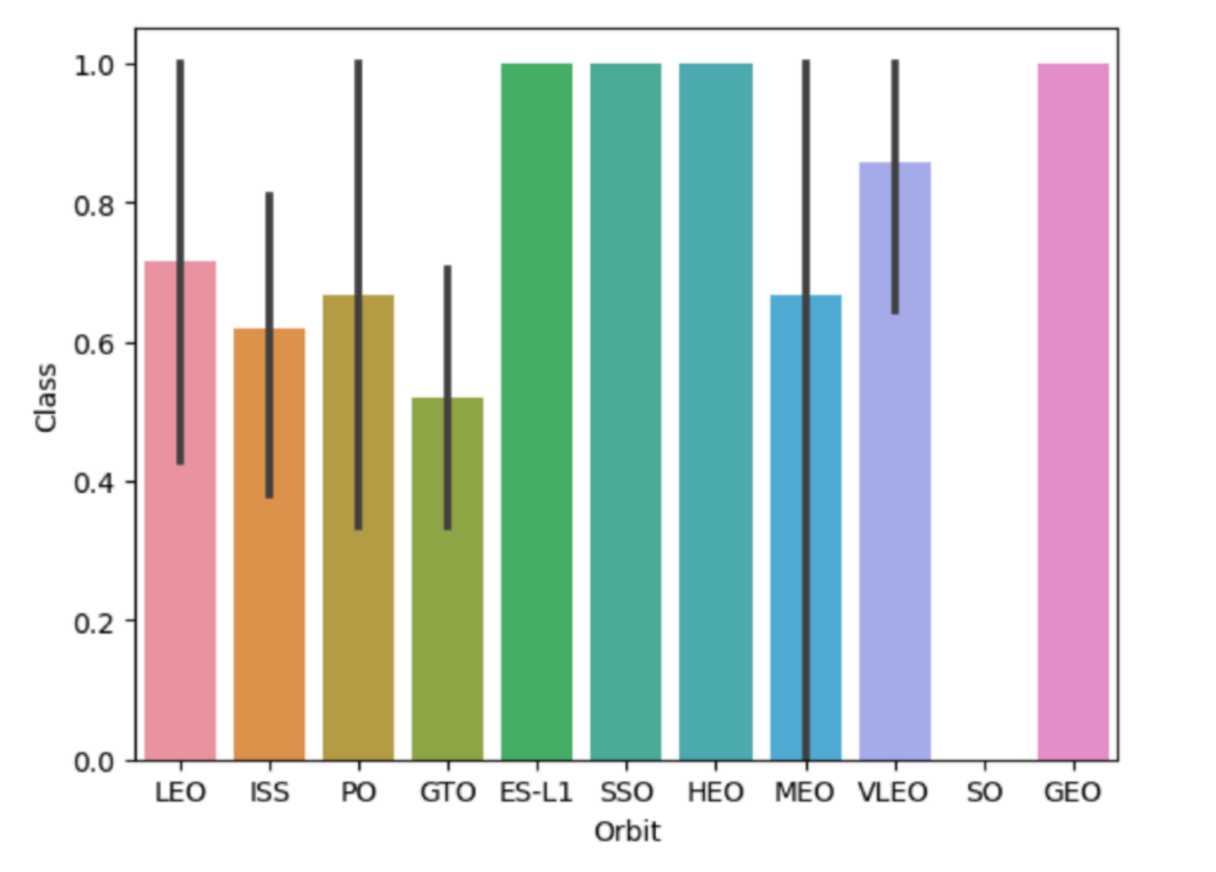
- Success Rate depends on the Launch Site and increases with the Flight Number

Payload vs. Launch Site



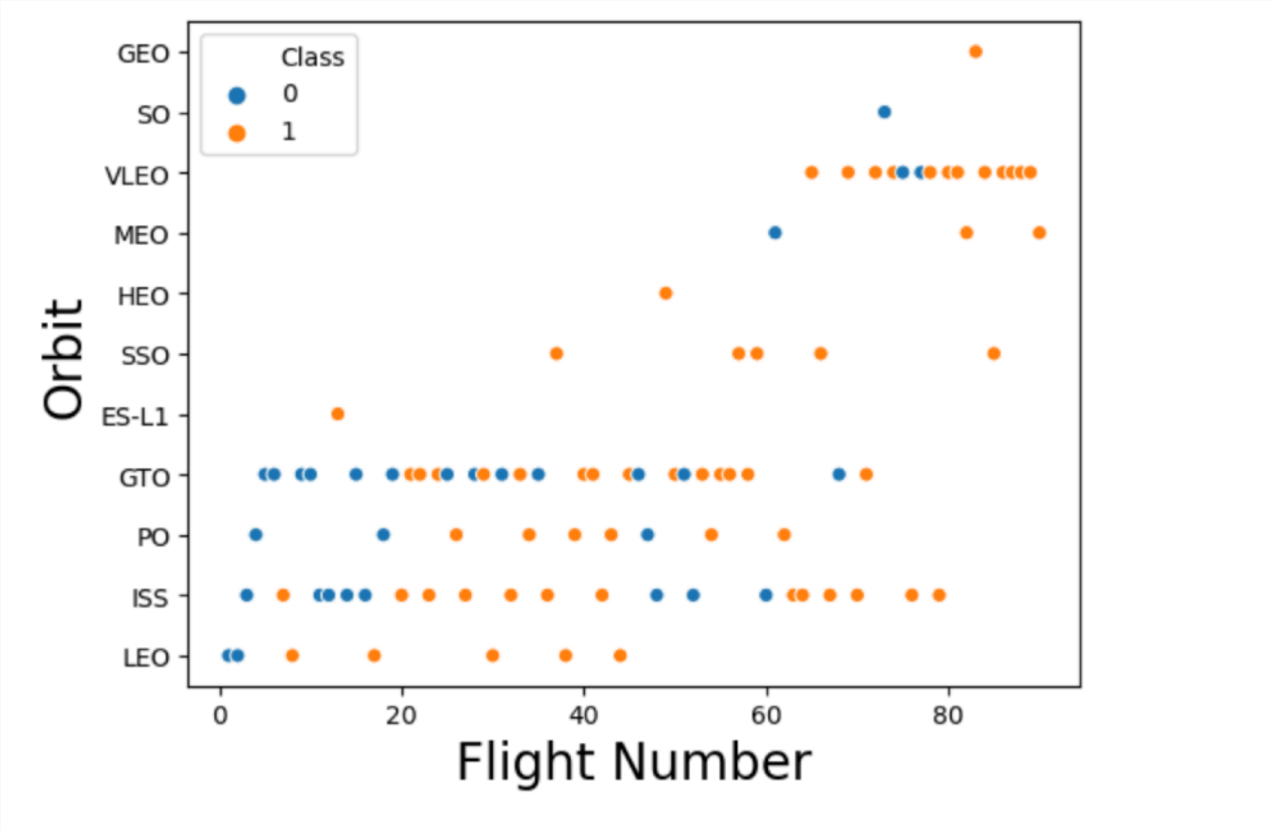
- Payload affects the Success of the mission.
- VAFB SLC 4E is successful in medium ranges of payloads

Success Rate vs. Orbit Type



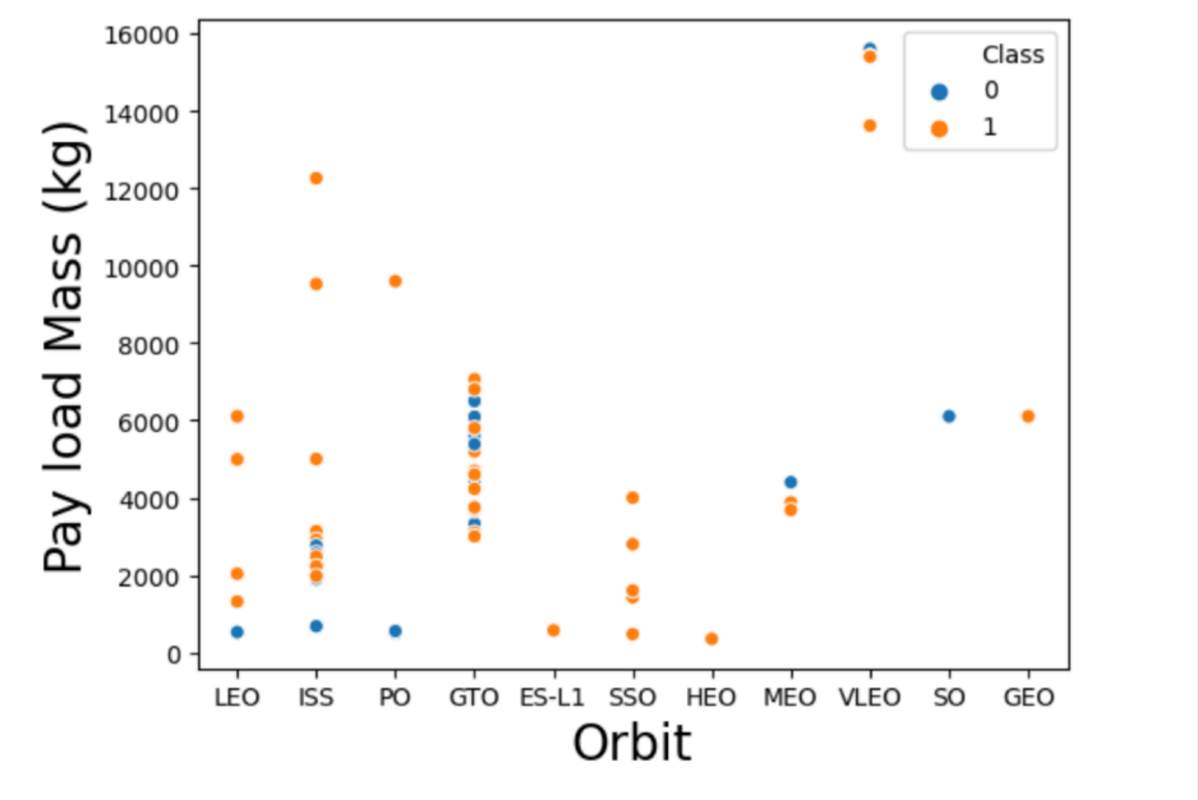
- The average success rate per Orbit Type
- Missions to GEO, ES-L1, SSO, HEO had the highest success rate

Flight Number vs. Orbit Type



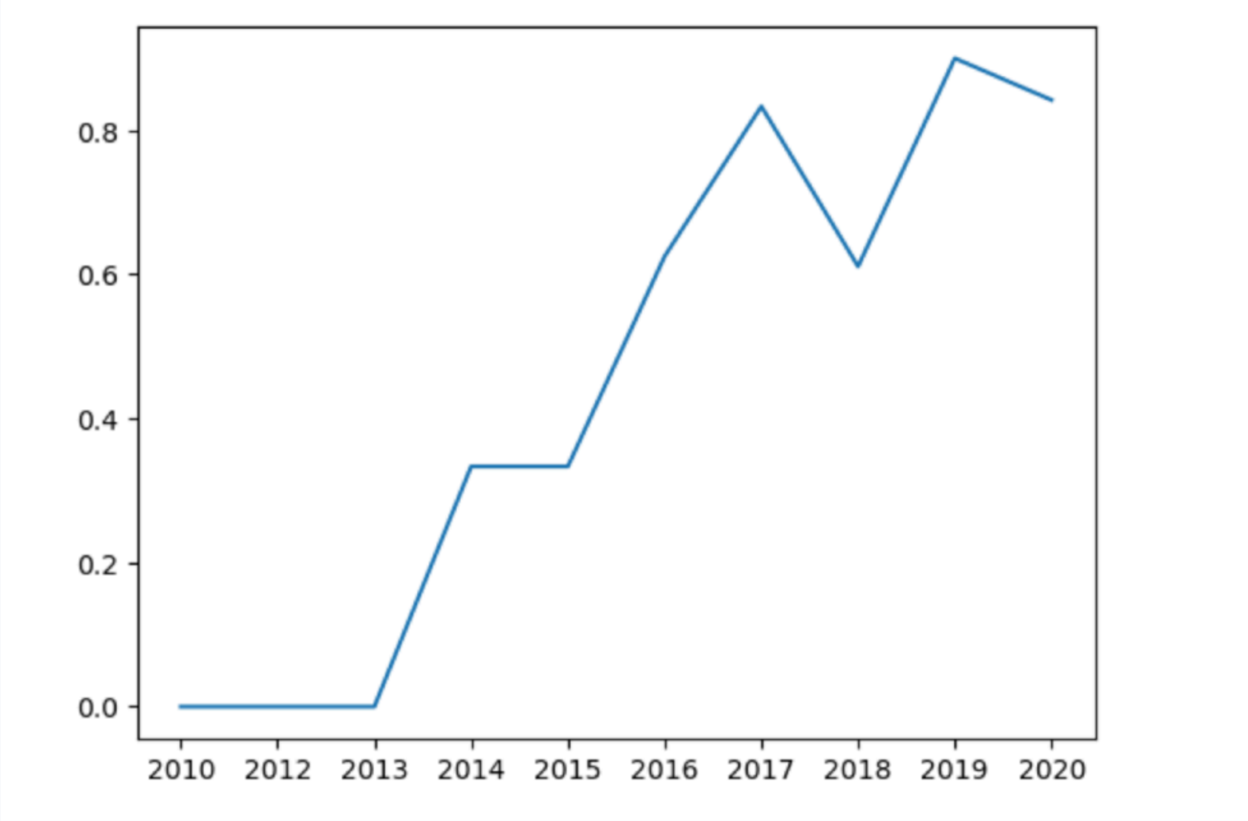
- Missions with higher number have more successful outcomes - for different orbit types

Payload vs. Orbit Type



- Payload affects the outcome of the mission
- For the payload below 6000 the missions to SSO were very successful

Launch Success Yearly Trend



- Success Rate grows with time. Late missions were on average more successful.
- We can see a significant increase since 2013.

All Launch Site Names

```
%sql select distinct LAUNCH_SITE from SPACEXTBL
```

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

- Four different Launch Sites have been used - three in Florida, one in California

Launch Site Names Begin with 'CCA'

```
%sql select * from SPACEXTBL where substring(LAUNCH_SITE,1,3)=='CCA' limit 5
```

Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer
04-06-2010	18:45:00	F9 v1.0 B0003	CCAFS LC-40	Dragon Spacecraft Qualification Unit	0	LEO	SpaceX
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
22-05-2012	07:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525	LEO (ISS)	NASA (COTS)
08-10-2012	00:35:00	F9 v1.0 B0006	CCAFS LC-40	SpaceX CRS-1	500	LEO (ISS)	NASA (CRS)
01-03-2013	15:10:00	F9 v1.0 B0007	CCAFS LC-40	SpaceX CRS-2	677	LEO (ISS)	NASA (CRS)

- CCA Launch Sites have been used to carry payload contracted mainly by NASA to carry it to Low Earth Orbit

Total Payload Mass

```
%sql select sum(payload_mass__kg_) from SPACEXTBL where CUSTOMER="NASA (CRS)"
```

sum(payload_mass__kg_)
45596

- Total payload carried by boosters from NASA was **45 596 kg**

Average Payload Mass by F9 v1.1

```
%sql select avg(PAYLOAD_MASS__KG_) from SPACEXTBL where substring(BOOSTER_VERSION, 1,7)="F9 v1.1"
```

avg(PAYLOAD_MASS__KG_)
2534.6666666666665

- On average the payload mass carried by booster version F9 v1.1 was 2535 kg

First Successful Ground Landing Date

```
%sql select min(DATE) from SPACEXTBL where MISSION_OUTCOME="Success"
```

min(DATE)
01-03-2013

- The first successful landing outcome on ground pad

Successful Drone Ship Landing with Payload between 4000 and 6000

```
%sql select BOOSTER_VERSION from SPACEXTBL where PAYLOAD_MASS__KG_ >4000 and PAYLOAD_MASS__KG_<6000 and [Landing _OUTCOME] like '%drone%' and MISSION_OUTCOME like '%Success%'
```

Booster_Version
F9 FT B1020
F9 FT B1022
F9 FT B1026
F9 FT B1021.2
F9 FT B1031.2

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

Total Number of Successful and Failure Mission Outcomes

```
%sql select count(*) as FAILURE from SPACEXTBL where [Landing _OUTCOME] not like '%Success%'
```

```
%sql select count(*) as SUCCESS from SPACEXTBL where [Landing _OUTCOME] like '%Success%'
```

FAILURE
40

SUCCESS
61

- Total number of successful and failure mission outcomes
- Although the success rate grows with time, the overall success rate is 60.4%

Boosters Carried Maximum Payload

```
%sql select BOOSTER_VERSION from SPACEXTBL where PAYLOAD_MASS__KG_ = (select max(PAYLOAD_MASS__KG_) from SPACEXTBL)
```

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

- Booster versions which have carried the maximum payload mass
- For the maximum load Falcon 9 B5 has been used

2015 Launch Records

```sql
%sql select substr(DATE,4,2) as MONTH, substr(DATE,7,4) as YEAR, BOOSTER_VERSION, LAUNCH_SITE, [LANDING _OUTCOME] from SPACEXTBL where [LANDING _OUTCOME] not like '%Success%' and YEAR='2015'
```

MONTH	YEAR	Booster_Version	Launch_Site	Landing _Outcome
01	2015	F9 v1.1 B1012	CCAFS LC-40	Failure (drone ship)
02	2015	F9 v1.1 B1013	CCAFS LC-40	Controlled (ocean)
03	2015	F9 v1.1 B1014	CCAFS LC-40	No attempt
04	2015	F9 v1.1 B1015	CCAFS LC-40	Failure (drone ship)
04	2015	F9 v1.1 B1016	CCAFS LC-40	No attempt
06	2015	F9 v1.1 B1018	CCAFS LC-40	Precluded (drone ship)

- There were several failed landings (drone ship) in 2015

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

```
%sql select [LANDING _OUTCOME] as LO, DATE, BOOSTER_VERSION, LAUNCH_SITE from SPACEXTBL where DATE>"04-06-2010" and DATE<"20-03-2017" order by LO, DATE DESC
```

LO	Date	Booster_Version	Launch_Site
Controlled (ocean)	18-04-2014	F9 v1.1	CCAFS LC-40
Controlled (ocean)	14-07-2014	F9 v1.1	CCAFS LC-40
Controlled (ocean)	11-02-2015	F9 v1.1 B1013	CCAFS LC-40
Failure	18-03-2020	F9 B5 B1048.5	KSC LC-39A
Failure	17-02-2020	F9 B5 B1056.4	CCAFS SLC-40
Failure	05-12-2018	F9 B5B1050	CCAFS SLC-40
Failure (drone ship)	17-01-2016	F9 v1.1 B1017	VAFB SLC-4E
Failure (drone ship)	15-06-2016	F9 FT B1024	CCAFS LC-40
Failure (drone ship)	14-04-2015	F9 v1.1 B1015	CCAFS LC-40
Failure (drone ship)	10-01-2015	F9 v1.1 B1012	CCAFS LC-40
Failure (parachute)	08-12-2010	F9 v1.0 B0004	CCAFS LC-40
No attempt	19-01-2020	F9 B5 B1046.4	KSC LC-39A
No attempt	16-03-2017	F9 FT B1030	KSC LC-39A
No attempt	15-05-2017	F9 FT B1034	KSC LC-39A
No attempt	08-10-2012	F9 v1.0 B0006	CCAFS LC-40
No attempt	07-09-2014	F9 v1.1 B1011	CCAFS LC-40
No attempt	06-03-2018	F9 B4 B1044	CCAFS SLC-40
No attempt	06-01-2014	F9 v1.1	CCAFS LC-40
No attempt	05-08-2014	F9 v1.1	CCAFS LC-40
No attempt	05-07-2017	F9 FT B1037	KSC LC-39A
No attempt	04-06-2018	F9 B4 B1040.2	CCAFS SLC-40
No attempt	06-08-2019	F9 B5 B1047.3	CCAFS SLC-40
Success	18-10-2020	F9 B5 B1051.6	KSC LC-39A
Success	18-08-2020	F9 B5 B1049.6	CCAFS SLC-40
Success	17-12-2019	F9 B5 B1056.3	CCAFS SLC-40
Success	16-11-2020	F9 B5B1061.1	KSC LC-39A
Success	15-11-2019	F9 B5 B1047.2	KSC LC-39A
Success	13-06-2020	F9 B5 B1059.3	CCAFS SLC-40
Success	12-06-2019	F9 B5 B1051.2	VAFB SLC-4E
Success	11-11-2019	F9 B5 B1048.4	CCAFS SLC-40
Success	11-01-2019	F9 B5 B1049.2	VAFB SLC-4E
Success	10-09-2018	F9 B5B1049.1	CCAFS SLC-40
Success	08-10-2018	F9 B5 B1048.2	VAFB SLC-4E
Success	07-08-2020	F9 B5 B1051.5	KSC LC-39A
Success	07-08-2018	F9 B5 B1046.2	CCAFS SLC-40
Success	07-03-2020	F9 B5 B1059.2	CCAFS SLC-40
Success	07-01-2020	F9 B5 B1049.4	CCAFS SLC-40
Success	06-12-2020	F9 B5 B1058.4	KSC LC-39A
Success	06-10-2020	F9 B5 B1058.3	KSC LC-39A
Success	05-12-2019	F9 B5B1059.1	CCAFS SLC-40
Success	05-11-2020	F9 B5B1062.1	CCAFS SLC-40
Success	04-06-2020	F9 B5 B1049.5	CCAFS SLC-40
Success (drone ship)	18-04-2018	F9 B4 B1045.1	CCAFS SLC-40
Success (drone ship)	14-08-2016	F9 FT B1026	CCAFS LC-40
Success (drone ship)	14-01-2017	F9 FT B1029.1	VAFB SLC-4E
Success (drone ship)	11-10-2017	F9 FT B1031.2	KSC LC-39A
Success (drone ship)	11-05-2018	F9 B5 B1046.1	KSC LC-39A
Success (drone ship)	09-10-2017	F9 B4 B1041.1	VAFB SLC-4E
Success (drone ship)	08-04-2016	F9 FT B1021.1	CCAFS LC-40
Success (drone ship)	06-05-2016	F9 FT B1022	CCAFS LC-40
Success (ground pad)	19-02-2017	F9 FT B1031.1	KSC LC-39A
Success (ground pad)	18-07-2016	F9 FT B1025.1	CCAFS LC-40
Success (ground pad)	15-12-2017	F9 FT B1035.2	CCAFS SLC-40
Success (ground pad)	14-08-2017	F9 B4 B1039.1	KSC LC-39A
Success (ground pad)	08-01-2018	F9 B4 B1043.1	CCAFS SLC-40
Success (ground pad)	07-09-2017	F9 B4 B1040.1	KSC LC-39A

Outcome_Type	Attempts
Controlled (ocean)	3
Failure	3
Failure (drone ship)	4
Failure (parachute)	1
No attempt	10
No attempt	1
Success	20
Success (drone ship)	8
Success (ground pad)	6

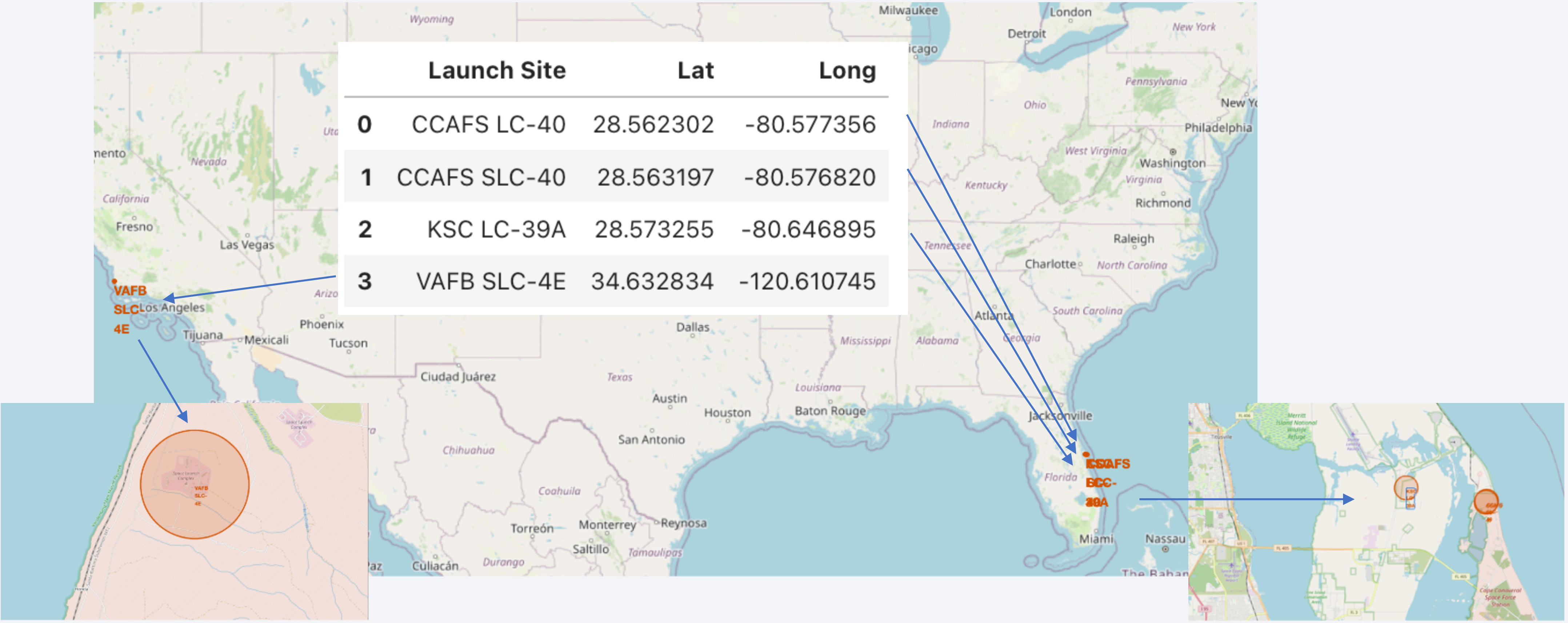
```
%sql select [LANDING _OUTCOME] as Outcome_Type, count(*) as Attempts from SPACEXTBL where DATE>"04-06-2010" and DATE<"20-03-2017" group by Outcome_Type order by Outcome_Type, DATE DESC
```

- There were many successful landings between the date 2010-06-04 and 2017-03-20

Section 3

Launch Sites Proximities Analysis

Launch Sites Location

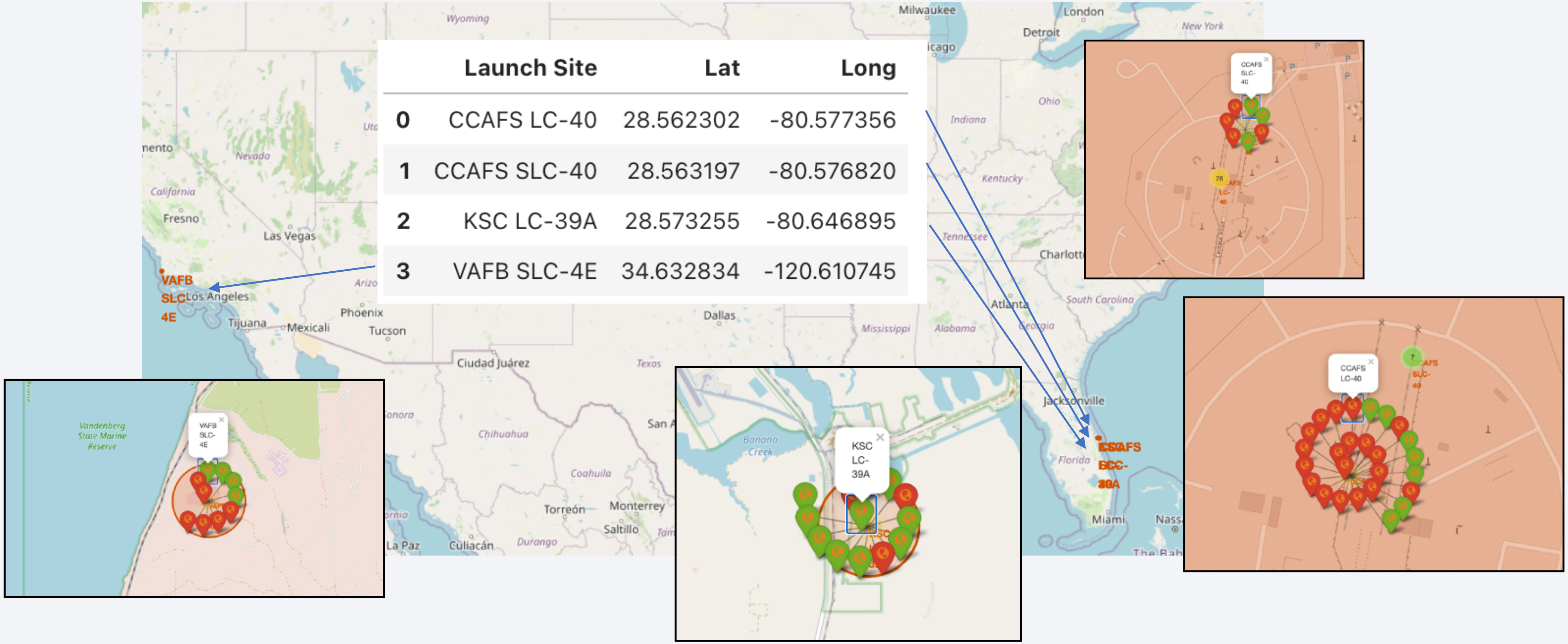


	Launch Site	Lat	Long
0	CCAFS LC-40	28.562302	-80.577356
1	CCAFS SLC-40	28.563197	-80.576820
2	KSC LC-39A	28.573255	-80.646895
3	VAFB SLC-4E	34.632834	-120.610745

- Main four launch sites are located in Florida and California. Their success rates depend on various parameters.
- Although we do not analyse the weather conditions at the time of the launches, but we can look closer at the locations of the sites.

Success Rate of the Launch Sites

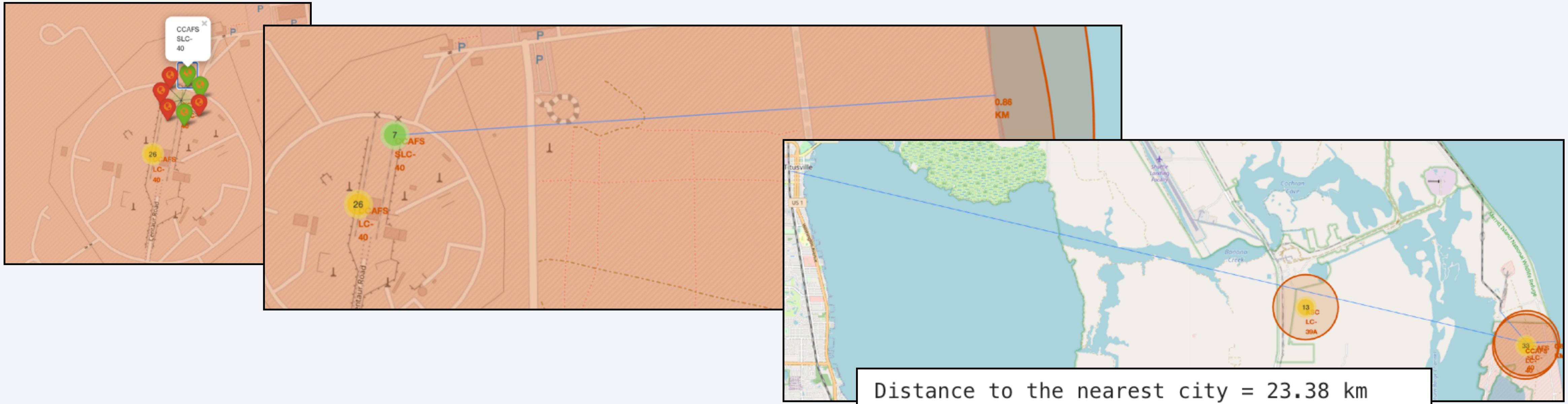
	Launch Site	Lat	Long
0	CCAFS LC-40	28.562302	-80.577356
1	CCAFS SLC-40	28.563197	-80.576820
2	KSC LC-39A	28.573255	-80.646895
3	VAFB SLC-4E	34.632834	-120.610745



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

Proximities to coastline, railway, highway, and city

- One of the most important elements that affects the launch success rate is the launch site proximity to the coastline the CCAFS SLC-40 is 0.88 km




```
Distance to the nearest city = 23.38 km
Distance to the nearest highway = 0.59 km
Distance to the nearest railway = 1.28 km
```

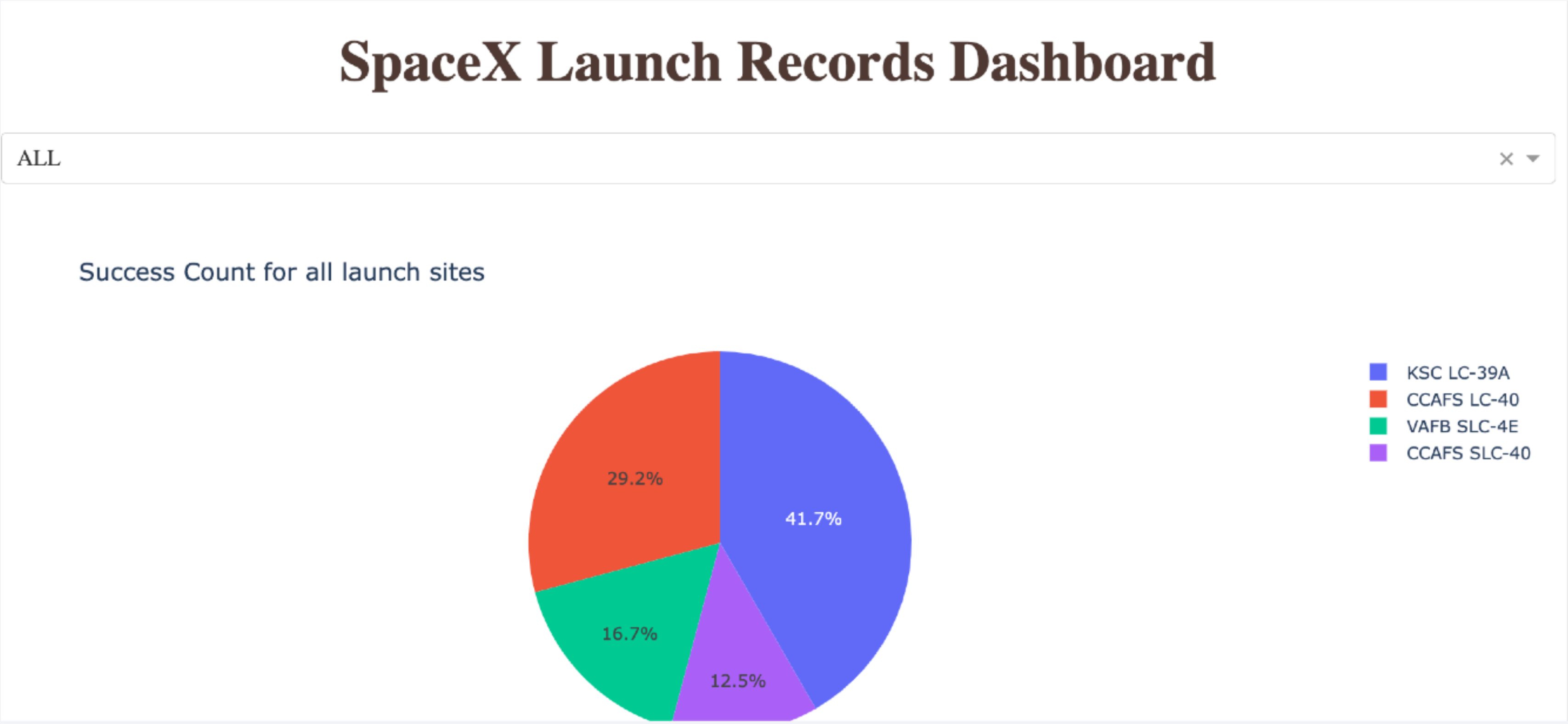
- Distances to railway, highway and the closest city are also important and for CCAFS SLC-40 are:

Section 4

Build a Dashboard with Plotly Dash

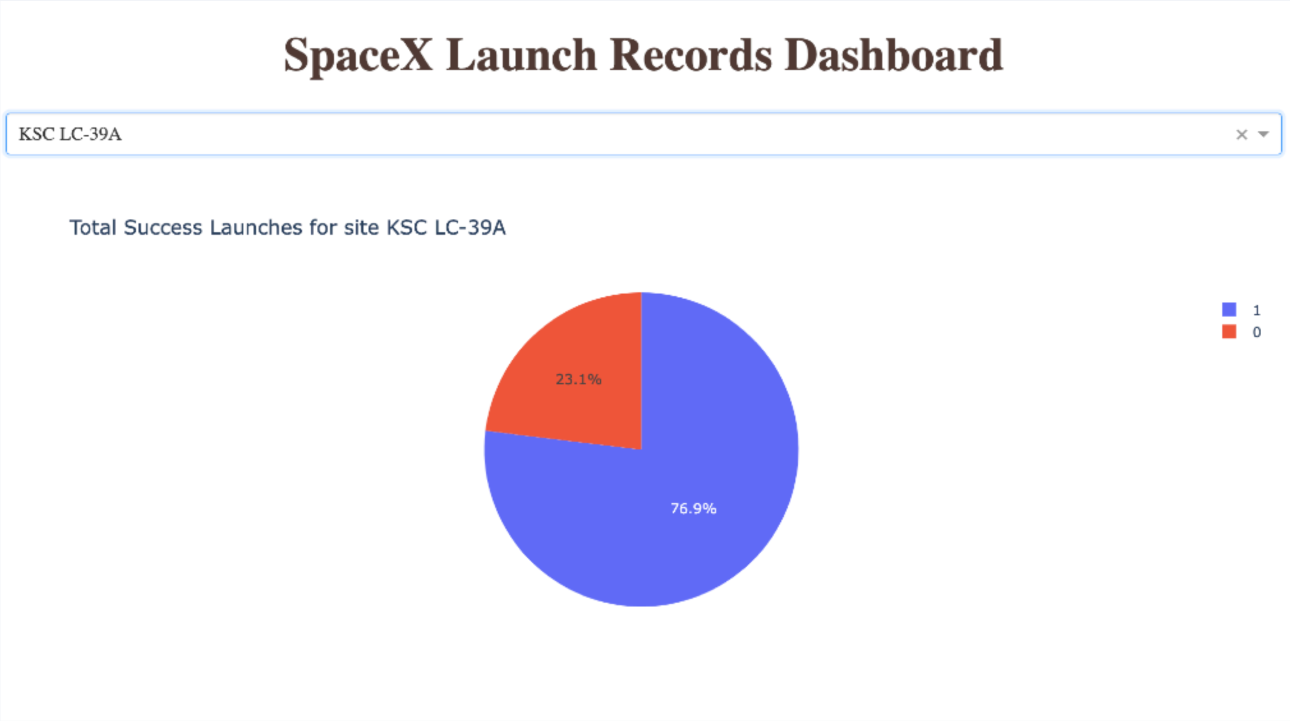
SpaceX Launch Records Dashboard - Success Count

- Interactive data analysis shows that the success of the launch highly depends on the launch site



The launch site KSC LC-39A was used in 41.7% of successful mission

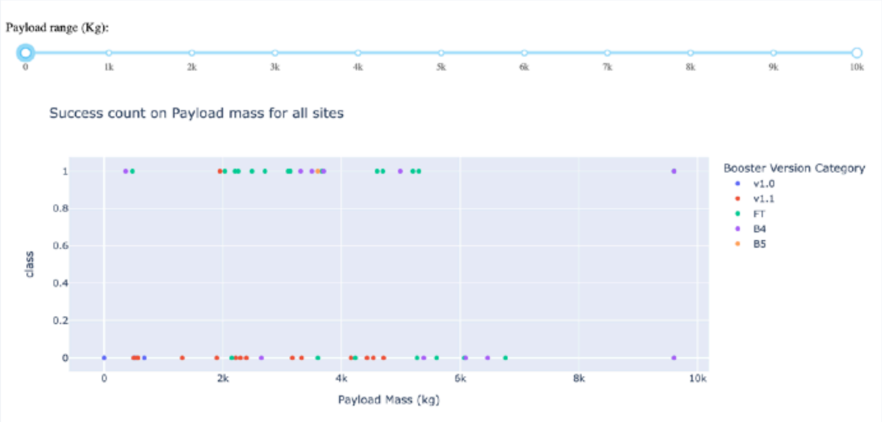
Total Success Launches for site KSC LC-39A



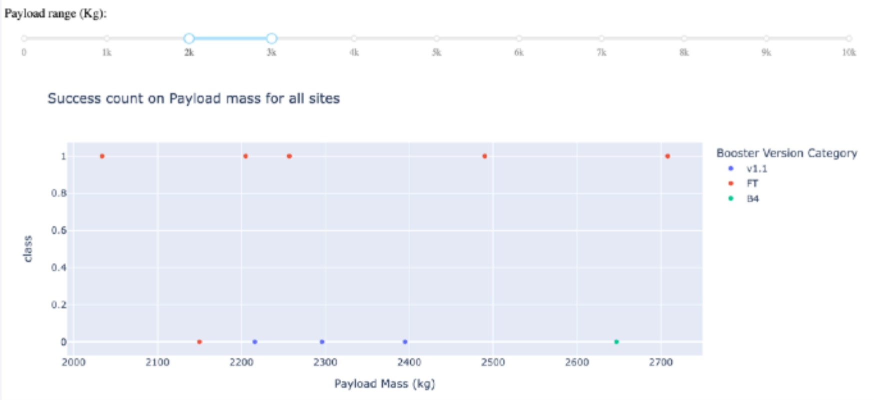
- All launches from KSC LC-39A gave a successful outcome in 78,1% of cases

Payload vs. Launch Success

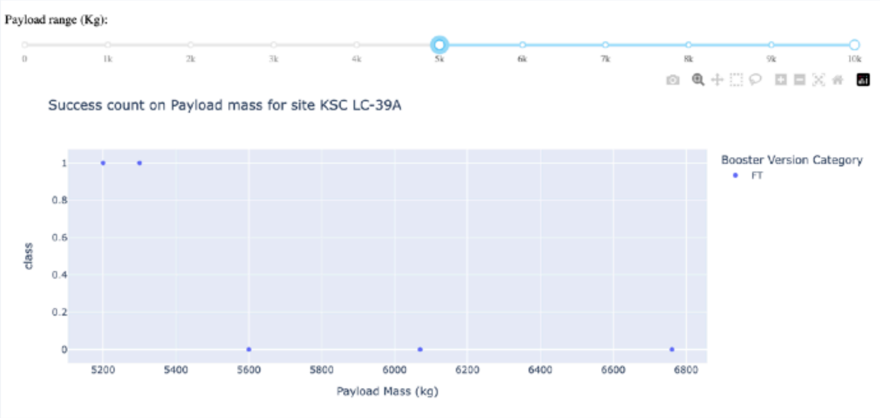
a.



b.



c.



d.

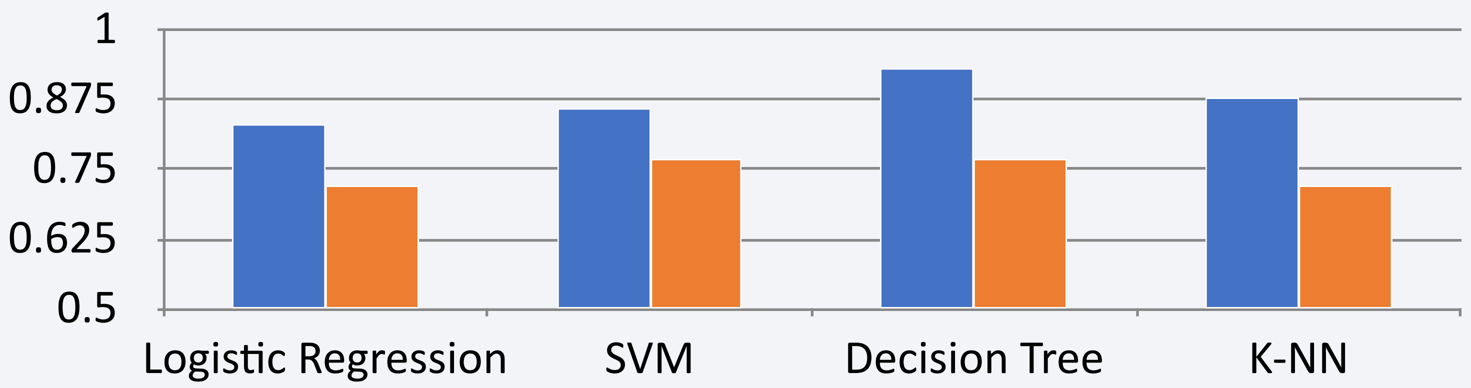
a. Launch Success depends on payload mass and the type of booster, **b.** Booster FT is the most successful in the payload mass range of 2-3k kg, **c.** The most successful launch site (KSC LC-39A) also fails when the payload mass is greater than 5.4k kg, **d.** The range 3-4k kg has the biggest success rate and the boosters FT and B4 perform the best.

Section 5

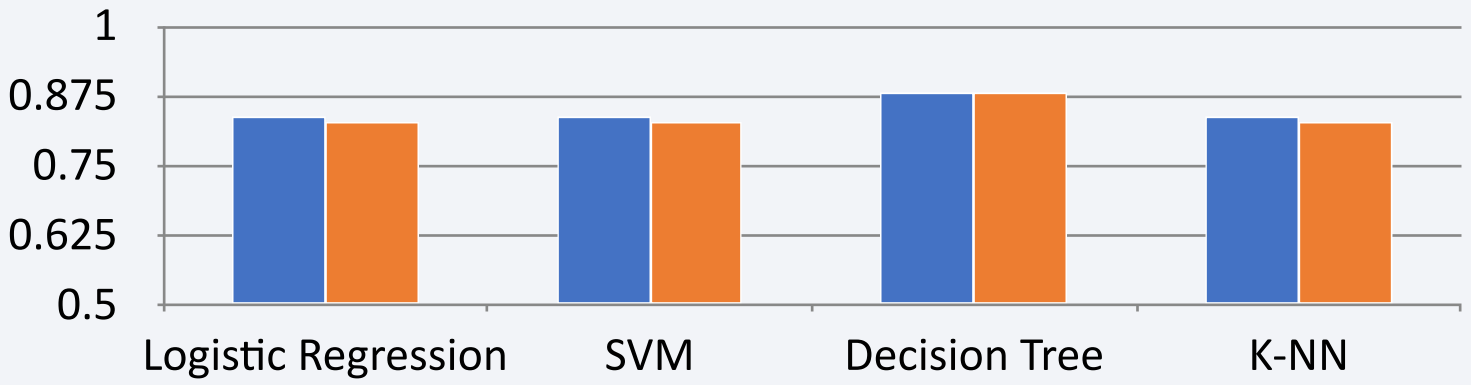
Predictive Analysis (Classification)

Classification Accuracy

- Comparison of the Classifiers - Train / Test - 0.8/0.2 - shuffle 4



- Comparison of the Classifiers - Train / Test - 0.8/0.2 - shuffle 2



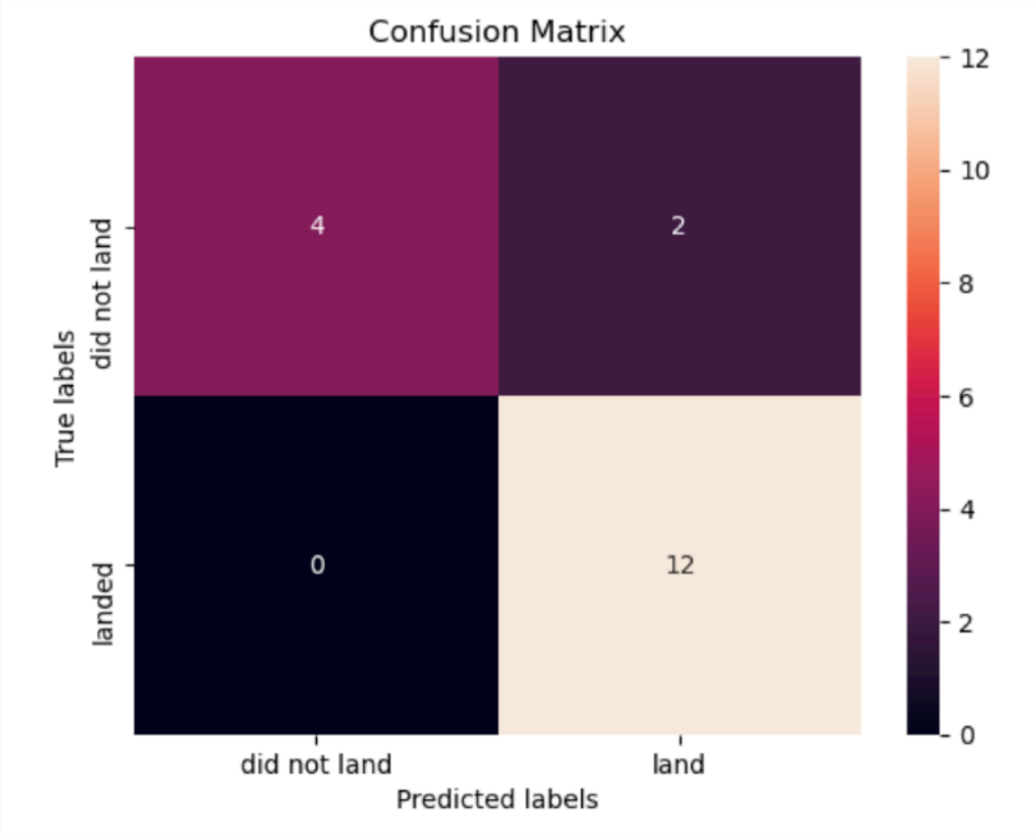
Decision Tree - the best classifier for the problem.

Performs best regardless the way the initial data is shuffled.

Trained model results with accuracy of 0.88 for both train and test sets.

Confusion Matrix

- The best performing model
- Decision Tree - accuracy 0.88
- Hyperparameters:
  - criterion='gini',
  - max_depth=10,
  - max_features='sqrt',
  - min_samples_leaf=1,
  - min_samples_split=10,
  - splitter='best'



Zero false negative and only two false positives using Decision Tree model

Conclusions

- It is possible to train a Machine Learning model and predict whether a SpaceX launch will end with successful landing.
- Decision Tree is the best classifier for the given problem.
- In 88% of cases the model predicts the correct outcome.
- Confusion matrix shows that the model has a slight tendency towards false positive.

Appendix

Links for Data collection

https://api.spacexdata.com/v4/launches/past

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

https://en.wikipedia.org/w/index.php?title=List_of_Falcon_9_and_Falcon_Heavy_launches&

Link to GitHub repository with completed Jupiter notebooks

https://github.com/mariuszochla/Applied_DS_Capstone

Appendix

Summary the SQL queries

```
%sql select distinct LAUNCH_SITE from SPACEXTBL
```

```
%sql select * from SPACEXTBL where substring(LAUNCH_SITE,1,3)=='CCA' limit 5
```

```
%sql select * from SPACEXTBL where substring(LAUNCH_SITE,1,3)=='CCA' limit 5
```

```
%sql select avg(PAYLOAD_MASS__KG_) from SPACEXTBL where substring(BOOSTER_VERSION, 1,7)="F9 v1.1"
```

```
%sql select min(DATE) from SPACEXTBL where MISSION_OUTCOME="Success"
```

```
%sql select BOOSTER_VERSION from SPACEXTBL where PAYLOAD_MASS__KG_ >4000 and PAYLOAD_MASS__KG_<6000 and [Landing _OUTCOME] like '%drone%' and MISSION_OUTCOME like '%Success%'
```

```
%sql select count(*) as FAILURE from SPACEXTBL where [Landing _OUTCOME] not like '%Success%'
```

```
%sql select BOOSTER_VERSION from SPACEXTBL where PAYLOAD_MASS__KG_ = (select max(PAYLOAD_MASS__KG_) from SPACEXTBL)
```

```
%sql select substr(DATE,4,2) as MONTH, substr(DATE,7,4) as YEAR, BOOSTER_VERSION, LAUNCH_SITE, [LANDING _OUTCOME] from SPACEXTBL where [LANDING _OUTCOME] not like '%Success%' and YEAR='2015'
```

```
%sql select [LANDING _OUTCOME] as LO, DATE, BOOSTER_VERSION, LAUNCH_SITE from SPACEXTBL where DATE>"04-06-2010" and DATE<"20-03-2017" order by LO, DATE DESC
```

```
%sql select count(*) as SUCCESS from SPACEXTBL where [Landing _OUTCOME] like '%Success%'
```

```
%sql select [LANDING _OUTCOME] as Outcome_Type, count(*) as Attempts from SPACEXTBL where DATE>"04-06-2010" and DATE<"20-03-2017" group by Outcome_Type order by Outcome_Type, DATE DESC
```

Appendix

Python code for Interactive Dashboard

Layout

```
app.layout = html.Div(children=[html.H1('SpaceX Launch Records Dashboard',
                                        style={'textAlign': 'center', 'color': '#503D36',
                                               'font-size': 40}),

                                dcc.Dropdown(id='site-dropdown',
                                options=[
                                  {'label': x, 'value': x} for x in sites
                                ],
                                value="ALL",
                                placeholder="Select a Launch Site here",
                                searchable=True
                                ),

                                html.Br(),

                                html.Div(dcc.Graph(id='success-pie-chart')),
                                html.Br(),

                                html.P("Payload range (Kg):"),
                                # TASK 3: Add a slider to select payload range
                                dcc.RangeSlider(id='payload-slider',
                                         min=0,
                                         max=10000,
                                         step=1000,
                                         value=[min_payload, max_payload]
                                         ),


                                html.Div(dcc.Graph(id='success-payload-scatter-chart')),
                                ])
```

Callback

```
@app.callback(Output(component_id='success-payload-scatter-
chart',component_property='figure'),
              [Input(component_id='site-
dropdown',component_property='value'),
              Input(component_id='payload-
slider',component_property='value')])
```

```
@app.callback( Output(component_id='su
ccess-pie-chart',
component_property='figure'),
             Input(component_id='site-
dropdown', component_property='value'))
```

Thank you!