



IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Jose Javier Calderon Marin
07/04/2022



Outline

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

Executive Summary

- The purpose of this study is thus to build a predictive model of a rocket launch outcome.
- Falcon 9 rocket launches carried out in the last years are analyzed.
- Extensive rocket launches data sets are extracted from SpaceX API and from Wikipedia by means of Python programming language.
- Exploratory data analysis is performed using visualization and SQL, Folium and Plotly Dash. Following main conclusions are obtained:
 - Launch success rate since 2013 kept increasing till 2020. Successful mission rate awesomely high (98%), while successful landing rate is about 50%.
 - Most part of recent launches were made from CCAFS with a success rate of 0.75% (32 launches) and from KSC with a success rate of 80%.
 - Launch sites are located as close as possible to the Equator line and in very close proximity to the coast, railway and highway. On the contrary, they maintain a certain distance to the cities.

Executive Summary

- Most part of the launches are carried out with a payload mass which varies from 2000 to 7000 kg with a high success rate. Heavy payload missions (payload mass > 8000 kg) have a 100% success rate.
- Most recent missions use a VLEO orbit and present a high successful rate of 86%.
- Boosters F9 FT have successfully landed on drone ship.
- Finally, predictive analysis are carried out using classification models:
 - Models parameters are optimized.
 - Logistic regression, SVM and KNN provide accurate predictions on the landing success. They should be improved to avoid the false positives which would impact directly on the final cost of the mission.

Introduction

- The commercial space age is here. Our company SpaceY is entering the market and defining its strategy. One of the first steps is to determine the price of each launch.
- Regular rockets have its payload enclosed in the fairings; the second stage helps bring the payload to orbit, but most of the work is done by the first stage which is quite large and expensive.
- Actually, the rocket launch cost will depend mostly on the possibility of recovering that first stage, as demonstrated by SpaceX. It advertises **Falcon 9 rocket** launches on its website with a cost of 62 million dollars while other providers cost upwards of 165 million dollars each.
- First stage recovering strategy will be thus adopted. However, sometimes it will not land correctly and will crash. Therefore, if we can determine if the first stage will land, we can approximately determine the cost of a launch.
- The purpose of this study is thus to build a predictive model of a rocket launch outcome.

Section 1

Methodology

Methodology

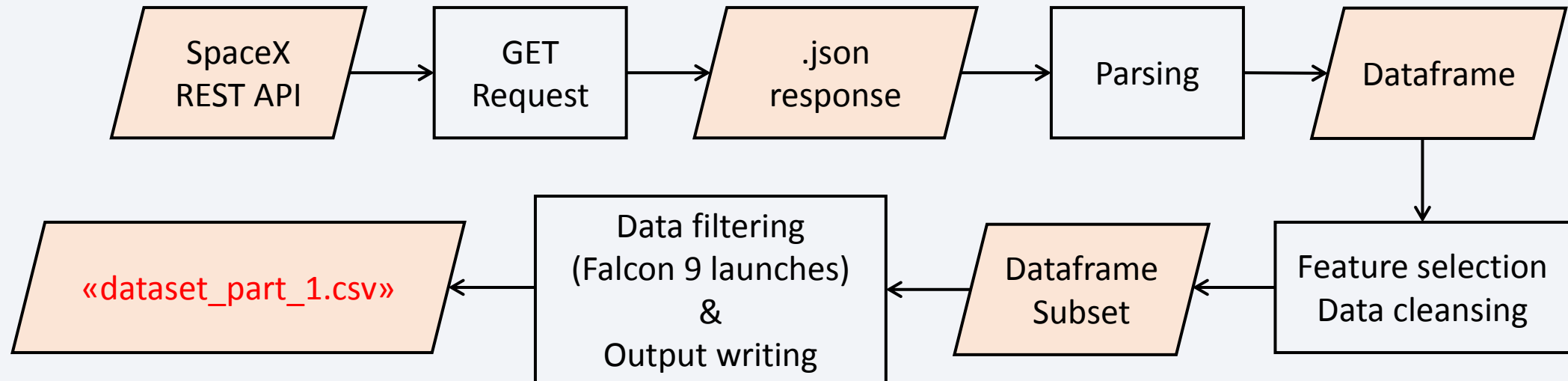
Executive Summary

- Data collection methodology:
 - Describe how data was collected
- Perform data wrangling
 - Describe how data was 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

- First step to create the predictive model consist in analyzing the Falcon 9 rocket launches carried out in the last years and identifying their relevant features.
- Extensive rocket launches data sets can be extracted from:
 - SpaceX REST API: “*<https://api.spacexdata.com/v4/launches/past>*”
 - Wikipedia website:
“*https://en.wikipedia.org/w/index.php?title=List_of_Falcon_9_and_Falcon_Heavy_launches&oldid=1027686922*”
- Python programming language has been used to perform the data collection.

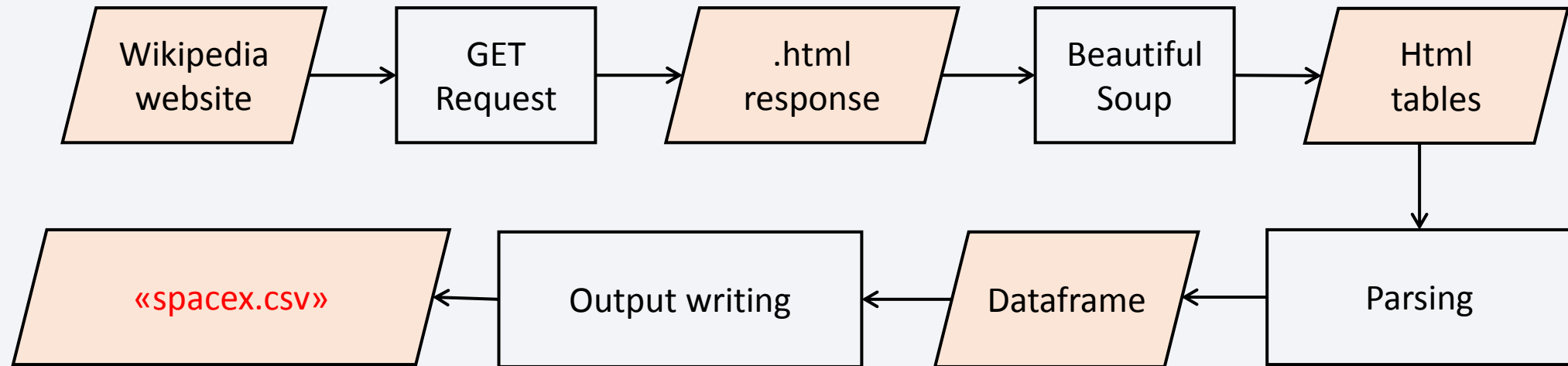
Data Collection – SpaceX API



- First rows of the dataset included in file “dataset_part_1.csv” are shown in appendix paragraph.
- Data collection details can be found in the following notebook:

https://github.com/JJCM2022/IBM-Data-Science-Professional-Certificate---Capstone-Project/blob/main/01_jupyter-labs-spacex-data-collection-api_DONE.ipynb

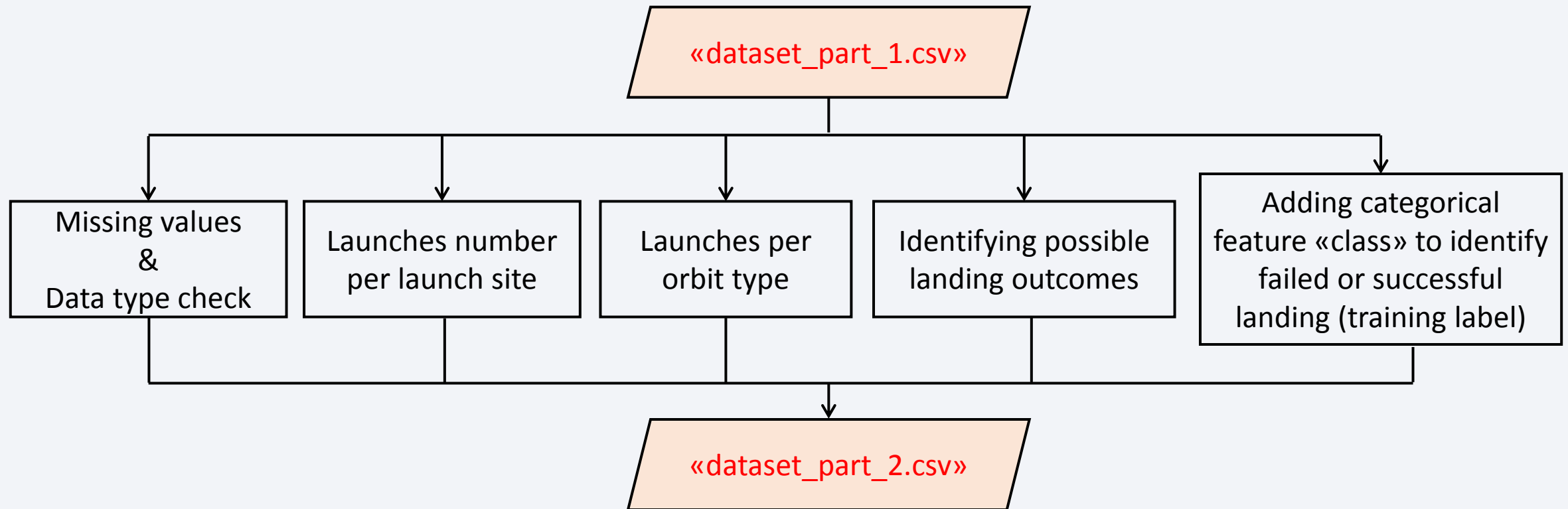
Data Collection - Scraping



- First rows of the dataset included in file “spacex.csv” are shown in appendix paragraph.
- Data collection details can be found in the following notebook:

https://github.com/JJCM2022/IBM-Data-Science-Professional-Certificate---Capstone-Project/blob/main/02_jupyter-labs-webscraping_DONE.ipynb

Data Wrangling



- Data wrangling details can be found in the following notebook:

https://github.com/JJCM2022/IBM-Data-Science-Professional-Certificate---Capstone-Project/blob/main/03_labs-jupyter-spacex-Data%20wrangling_DONE.ipynb

EDA with Data Visualization

- Dataset “**dataset_part_2.csv**” is analyzed by means of the following graphs:
 1. Scatter plots:
 - a. Flight number vs Payload mass (hue = class): most used payload and its influence in the mission outcome.
 - b. Flight number vs Launch site (hue = class): most used launch site and its influence in the mission outcome.
 - c. Flight number vs Orbit type (hue = class): most used orbit type and its relationship with the mission outcome.
 - d. Payload vs Orbit type (hue = class): relationship between Payload and Orbit type.
 2. Bar plot: Orbit type vs success rate: relationship between success rate and orbit type.
 3. Line chart: launch success yearly trend.

https://github.com/JJCM2022/IBM-Data-Science-Professional-Certificate---Capstone-Project/blob/main/05_jupyter-labs-eda-dataviz_DONE.ipynb

EDA with SQL

- Following SQL queries have been performed for EDA purposes of dataset “**spacex.csv**”:
 1. Names of the unique launch sites in the space mission.
 2. Count of landing outcomes between the date 2010-06-04 and 2017-03-20.
 3. Total number of successful and failure mission outcomes.
 4. Date when the first successful landing outcome in ground pad was achieved.
 5. Failed landing outcomes in drone ship in year 2015.
 6. Names of the boosters which have success in drone ship and have payload mass greater than 4000 but less than 6000.
 7. Total payload mass carried by boosters launched by NASA (CRS).
 8. Average payload mass carried by booster version F9 v1.1.
 9. Names of the booster versions which have carried the maximum payload mass.

https://github.com/JJCM2022/IBM-Data-Science-Professional-Certificate---Capstone-Project/blob/main/04_jupyter-labs-eda-sql-coursera_DONE.ipynb

Build an Interactive Map with Folium

- Following objects were added to a folium map:
 1. Circles and markers: they are used to locate the launch sites position on the map.
 2. Features group: it contains all the other features; this allows to add all the feature to the map at once.
 3. Marker cluster: launches markers are grouped by launch site.
 4. Mouse position: allows to get coordinate for a mouse over a point on the map; it is used to measure distances between relevant locations (launch sites and nearest point on the coast, railways, highways, ...).
 5. Polylines: helpful to indicate what distances are measured on the map.

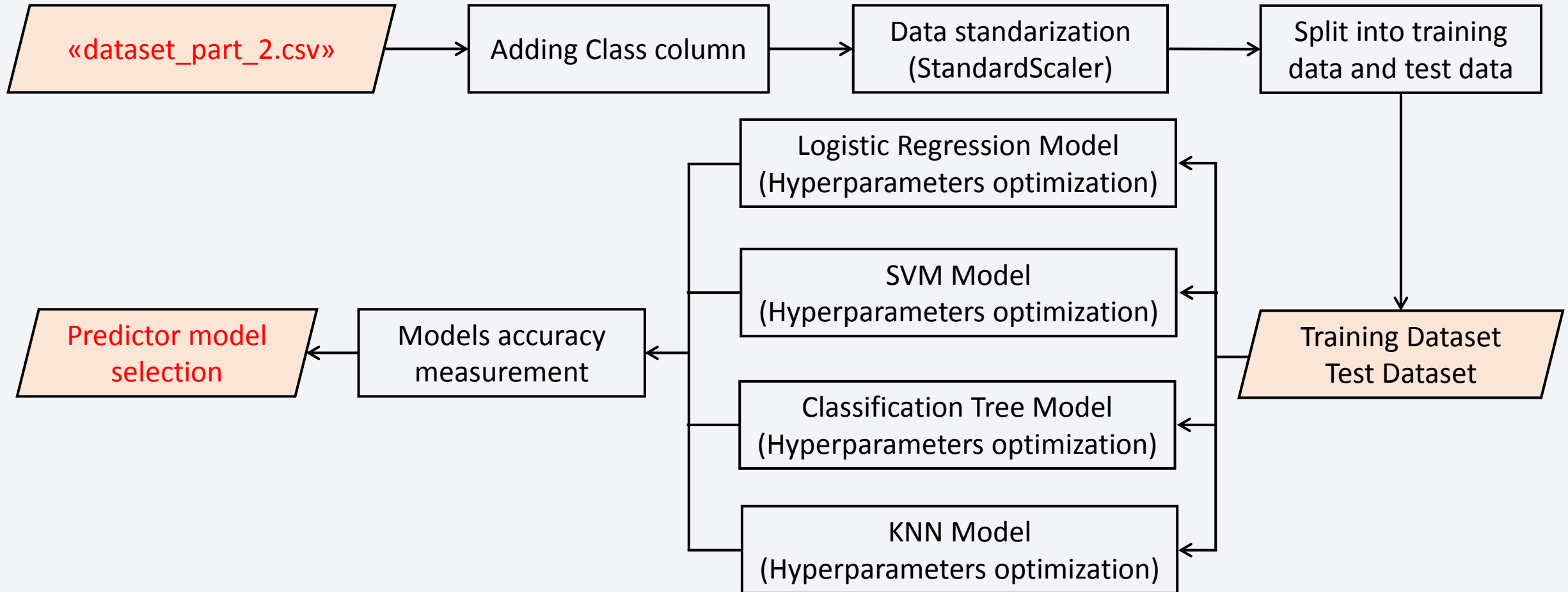
https://github.com/JJCM2022/IBM-Data-Science-Professional-Certificate---Capstone-Project/blob/main/06_lab_jupyter_launch_site_location_DONE_v2.ipynb

Build a Dashboard with Plotly Dash

- The following graphs have been built for EDA purposes:
 1. Pie chart to show the total successful launches by site.
 2. Pie chart to show the launch success rate for each site.
 3. Scatter plot to analyze the the payload influence on the launches success.

https://github.com/JJCM2022/IBM-Data-Science-Professional-Certificate---Capstone-Project/blob/main/07_spacex_dash_app_v3.py

Predictive Analysis (Classification)



Results

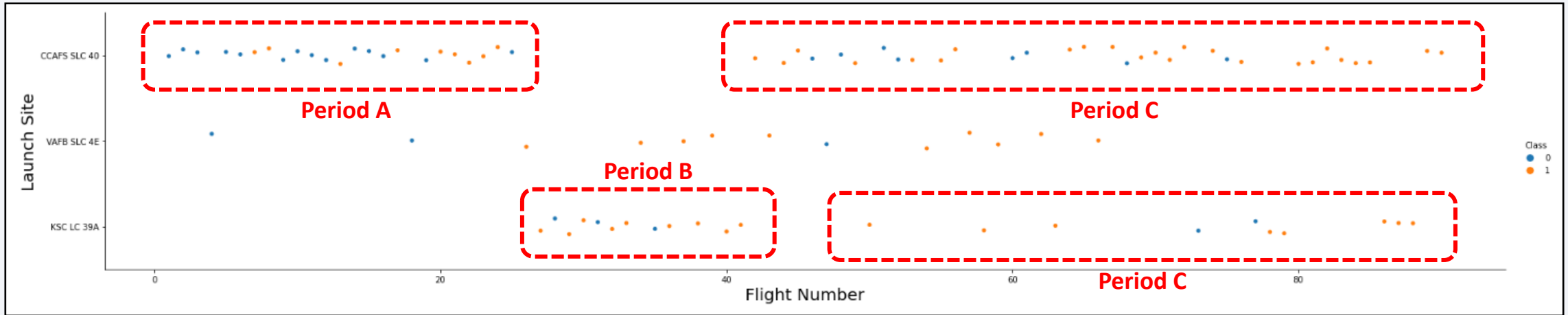
- Exploratory data analysis results
- Interactive analytics demo in screenshots
- Predictive analysis results

The background of the slide is an abstract composition. It features a dark blue base color. Overlaid on this are numerous diagonal streaks in shades of blue and red, creating a sense of motion or data flow. A faint, light blue grid pattern is also visible, particularly in the lower-left quadrant. The overall effect is high-tech and digital.

Section 2

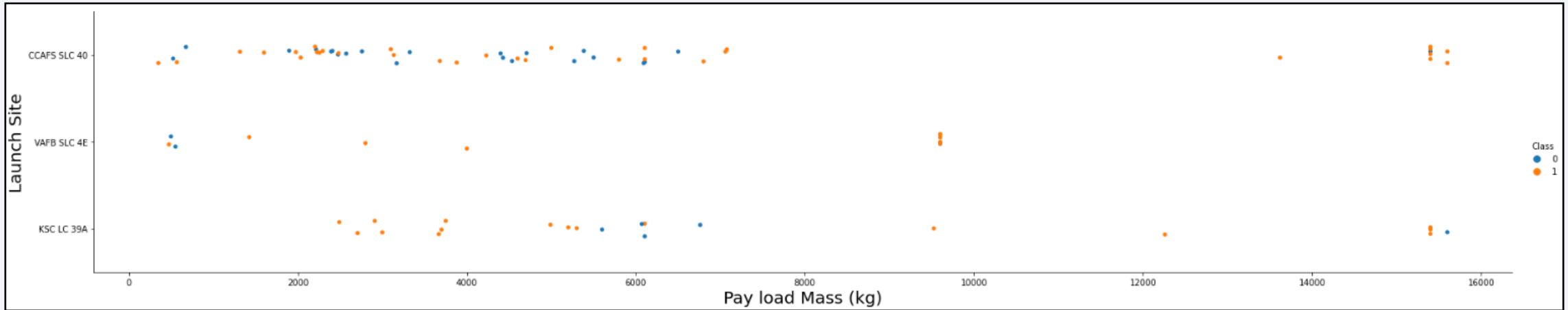
Insights drawn from EDA

Flight Number vs. Launch Site



- Launch success rate has increased for all the launch sites from the first to the last launch.
- Most part of the launches were performed from CCAFS and KSC launch sites.
- Launches made from VAFB present a great success rate of 77% (13 launches).
- Launches carried out from CCAFS and KSC can be analyzed in three main periods:
 - A. First 25 launches were carried out from CCAFS with a low success rate of 39% (23 launches).
 - B. Then launches 27 - 41 were performed from KSC with a success rate of 75% (12 launches).
 - C. From flight 42 on most of the launches were made from CCAFS again with a success rate of 0.75% (32 launches). In that period 10 launches were made from KSC with a success rate of 80%.

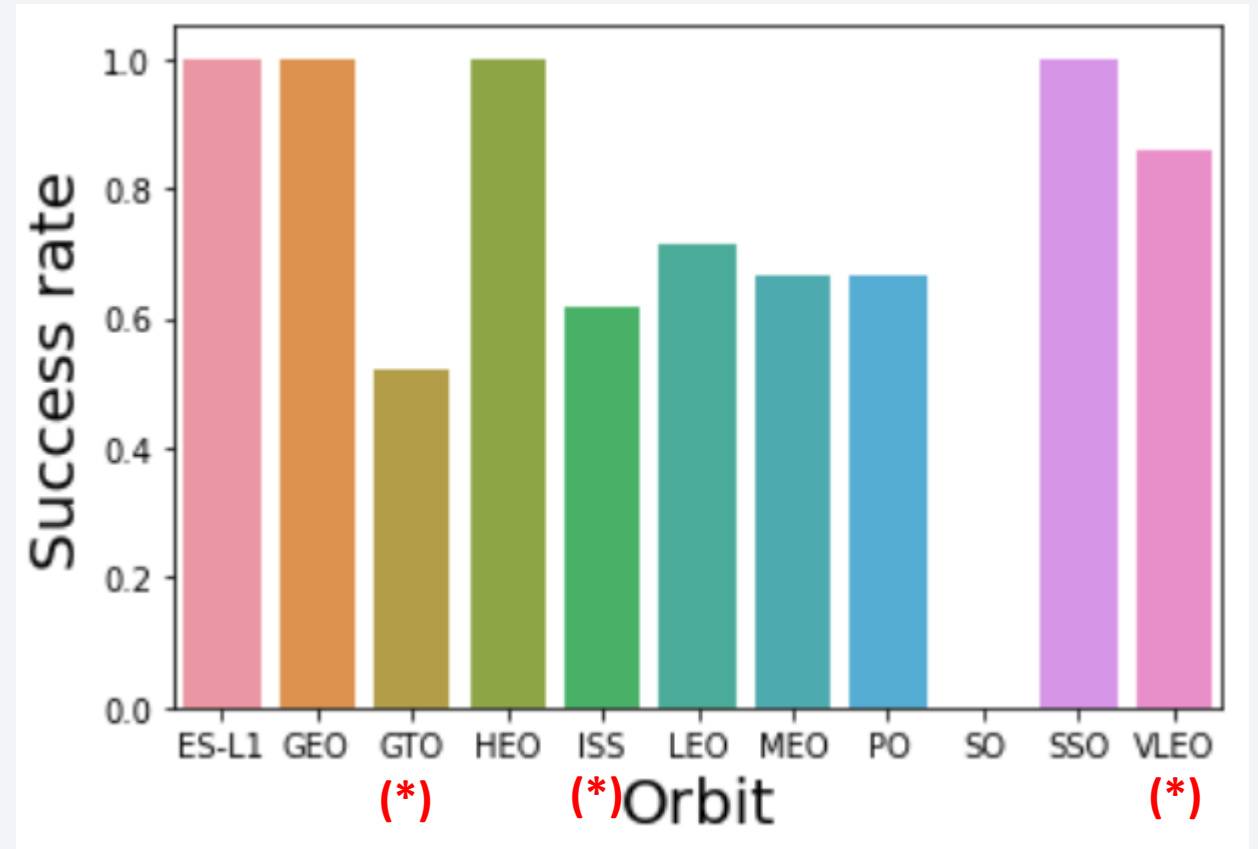
Payload vs. Launch Site



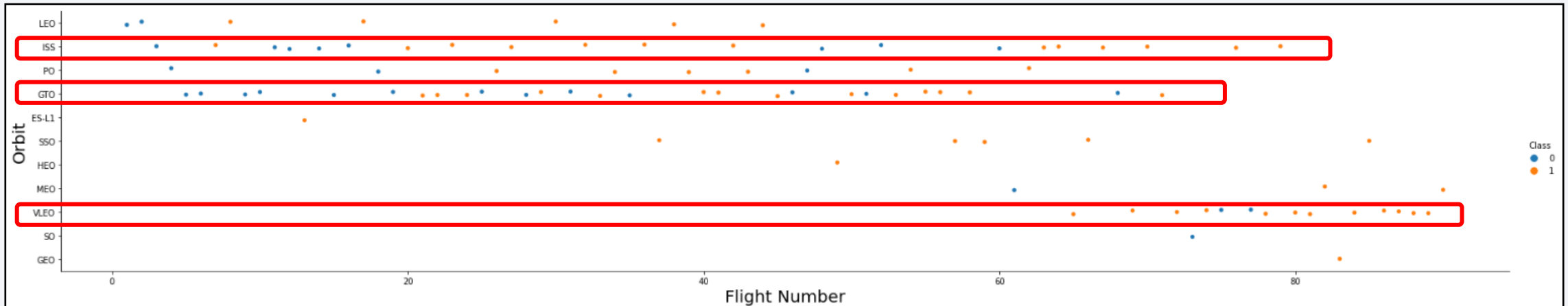
- For the VAFB-SLC launch site there are no rockets launched for heavy payload mass (> 10000 kg).
- Most part of the launches are carried out with a payload mass which varies approximately from 2000 to 7000 kg.
- High launch success rate for payload mass > 7000 kg.

Success Rate vs. Orbit Type

- No launch failure for orbits ES-L1, GEO, HEO and SSO.
- No successful launch found for orbit type SO.
- The rest of orbit types present a success rate of around 60%.
- Number of launches per orbit type would complete the information provided by this bar chart since the higher the number of flights is the more representative (*) is the calculated success rate (see next slide).

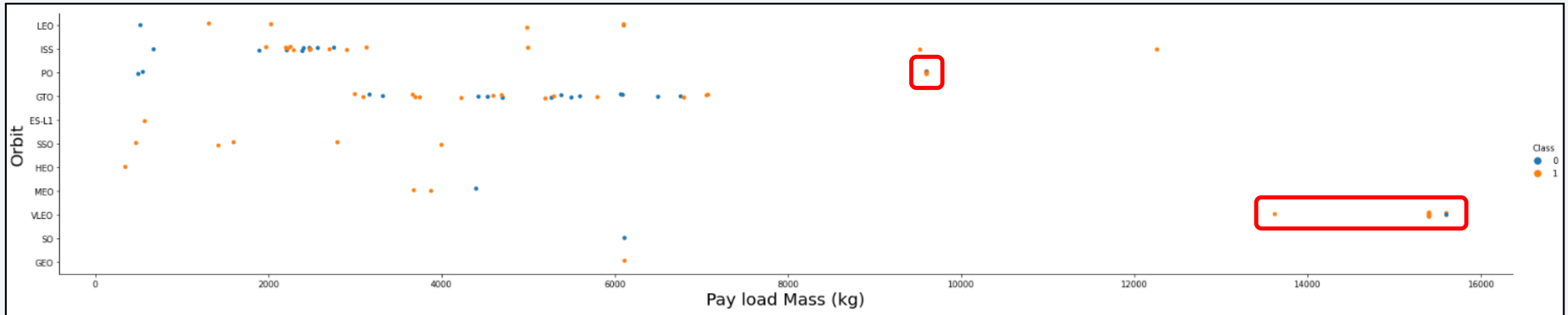


Flight Number vs. Orbit Type



- Higher number of launches correspond mainly to the orbit types ISS (21), GTO (27) and VLEO (14).
- In the LEO orbit the success appears related to the number of flights.
- There seems to be no relationship between orbit type and flight number for GTO and ISS orbits.
- VLEO orbit type has been chosen for the most recent launches (from flight 65 on) and present a high successful rate of 86%.

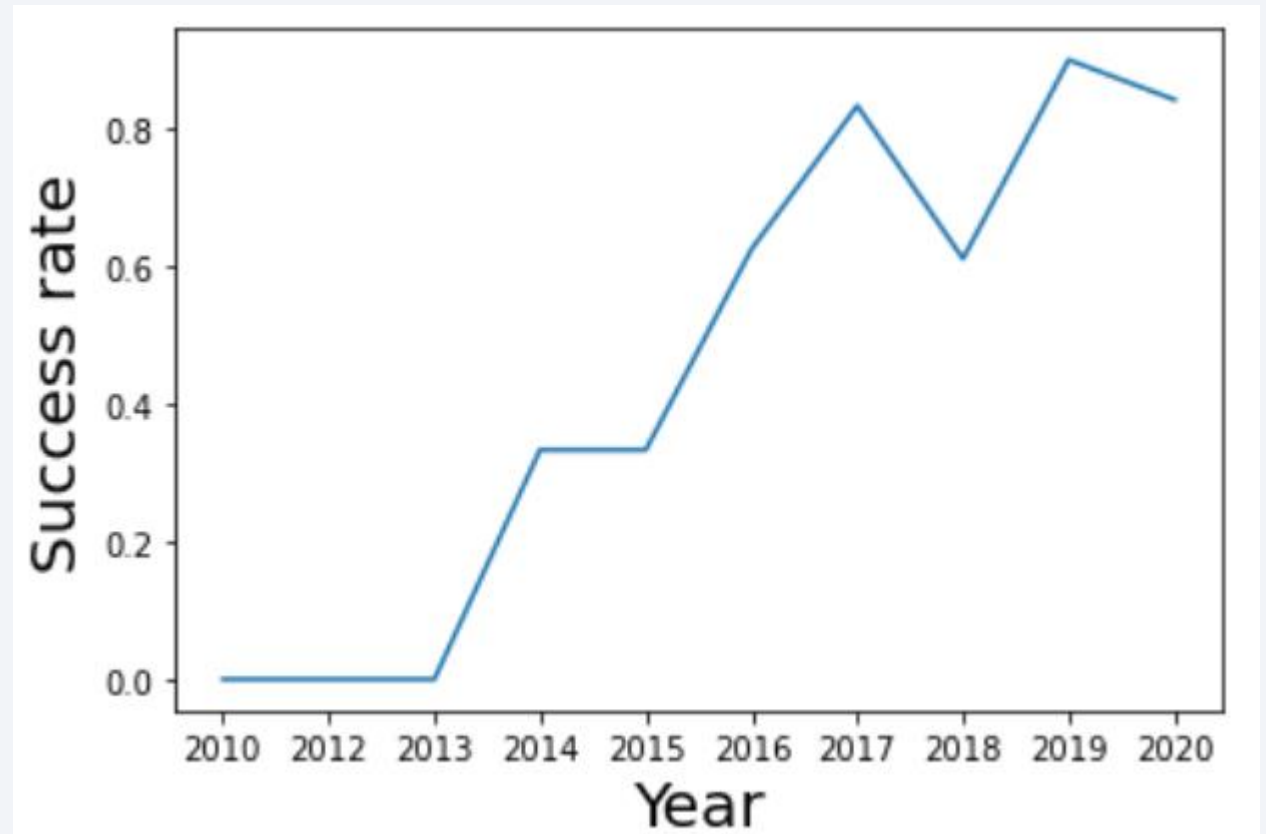
Payload vs. Orbit Type



- With heavy payloads the successful landing rate is higher for PO and VLEO orbits.
- 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

- The success rate since 2013 kept increasing till 2017.
- In 2018 the rate falls to around 60%, but it raised again to its maximum in 2019.



All Launch Site Names

- First step in the analysis of the dataset “**spacex.csv**” is to find the names of the unique launch sites in the space mission are:

```
In [26]: %sql SELECT DISTINCT(launch_site) FROM SPACEXTBL
```

```
* ibm_db_sa://srt20899:***@ea286ace-86c7-4d5b-8580-3fbfa46b1c66.bs2io90108kqb1od81cg.databases.appdomain.cloud:31505/bludb  
Done.
```

```
Out[26]:
```

launch_site

CCAFS LC-40

CCAFS SLC-40

KSC LC-39A

VAFB SLC-4E

Launch Site Names Begin with 'CCA'

- The following query gives 5 records where launch sites begin with 'CCA':

```
In [27]: %sql SELECT * FROM SPACEXTBL WHERE LAUNCH_SITE LIKE '%CCA%' LIMIT 5
```

```
* ibm_db_sa:///srt20899:***@ea286ace-86c7-4d5b-8580-3fbfa46b1c66.bs2io90l08kqb1od8l1cg.databases.appdomain.cloud:31505/blddb
Done.
```

Out[27]:

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

- The total payload carried by boosters from NASA is:

```
In [28]: %%sql SELECT SUM(PAYLOAD_MASS__KG_) FROM SPACEXTBL WHERE CUSTOMER = 'NASA (CRS)'  
%sql SELECT SUM(PAYLOAD_MASS__KG_) FROM SPACEXTBL WHERE CUSTOMER LIKE '%NASA (CRS)%'  
  
* ibm_db_sa://srt20899:***@ea286ace-86c7-4d5b-8580-3fbfa46b1c66.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:31505/bludb  
Done.  
  
Out[28]: 1  
48213
```

Average Payload Mass by F9 v1.1

- The average payload mass carried by booster version F9 v1.1 is:

```
In [29]: %sql SELECT AVG(PAYLOAD_MASS__KG_) FROM SPACEXTBL WHERE booster_version LIKE '%F9 v1.1%'
* ibm_db_sa://srt20899:***@ea286ace-86c7-4d5b-8580-3fbfa46b1c66.bs2io90l08kqb1od8l1cg.databases.appdomain.cloud:31505/bludb
Done.
Out[29]: 1
2534
```

- As can be seen, this booster has been used with very low payload masses.

First Successful Ground Landing Date

- The following query provides the date of the first successful landing outcome on ground pad:

```
In [30]: %sql SELECT DATE FROM SPACEXTBL WHERE LANDING__OUTCOME = 'Success (ground pad)' ORDER BY DATE LIMIT 1
* ibm_db_sa://srt20899:***@ea286ace-86c7-4d5b-8580-3fbfa46b1c66.bs2io90l08kqb1od8l1cg.databases.appdomain.cloud:31505/bludb
Done.
```

Out[30]:

DATE
2015-12-22

- It should be remarked that it took place 5 years after the first Falcon 9 launch (04/06/2010, according to dataset records shown in slide 47).

Successful Drone Ship Landing with Payload between 4000 and 6000

- The following figure shows a query to obtain the names of boosters which have successfully landed on drone ship and had a payload mass greater than 4000 but less than 6000:

```
In [31]: %sql SELECT booster_version, PAYLOAD_MASS__KG_ FROM SPACEXTBL WHERE LANDING__OUTCOME = 'Success (drone ship)' AND payload_mass__k
```

```
* ibm_db_sa://srt20899:***@ea286ace-86c7-4d5b-8580-3fbfa46b1c66.bs2io90108kqb1od81cg.databases.appdomain.cloud:31505/bludb  
Done.
```

```
Out[31]:
```

booster_version	payload_mass__kg_
F9 FT B1022	4696
F9 FT B1026	4600
F9 FT B1021.2	5300
F9 FT B1031.2	5200

Total Number of Successful and Failure Mission Outcomes

- It is interesting to know what is the total number of successful and failure mission outcomes:

```
In [11]: %sql SELECT MISSION_OUTCOME, COUNT(*) AS TOTAL FROM SPACEXTBL GROUP BY MISSION_OUTCOME
```

```
* ibm_db_sa://srt20899:***@ea286ace-86c7-4d5b-8580-3fbfa46b1c66.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:31505/bludb  
Done.
```

```
Out[11]:
```

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

- As can be seen, the successful mission rate is awesomely high (98%).

Boosters Carried Maximum Payload

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

```
In [33]: %sql SELECT DISTINCT(BOOSTER_VERSION) FROM SPACEXTBL WHERE PAYLOAD_MASS__KG_ = (SELECT MAX(PAYLOAD_MASS__KG_) FROM SPACEXTBL)
* ibm_db_sa://srt20899:***@ea286ace-86c7-4d5b-8580-3fbfa46b1c66.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:31505/bludb
Done.
```

```
Out[33]: 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
```

2015 Launch Records

- Here below a list of the failed "landing_outcomes" in drone ship is presented, together with their booster versions and launch site names during the year 2015:

```
In [34]: %sql SELECT date, landing__outcome, BOOSTER_VERSION, LAUNCH_SITE FROM SPACEXTBL WHERE YEAR(DATE) = 2015 AND LANDING__OUTCOME LIKE
```

* ibm_db_sa://srt20899:***@ea286ace-86c7-4d5b-8580-3fbfa46b1c66.bs2io90108kqb1od81cg.databases.appdomain.cloud:31505/bludb
Done.

```
Out[34]:
```

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

- First successful landing on a drone ship took of in 2016 with a new booster version:

```
In [14]: %sql SELECT date, landing__outcome, BOOSTER_VERSION, LAUNCH_SITE FROM SPACEXTBL WHERE LANDING__OUTCOME LIKE '%(drone ship)%' C
```

* ibm_db_sa://srt20899:***@ea286ace-86c7-4d5b-8580-3fbfa46b1c66.bs2io90108kqb1od81cg.databases.appdomain.cloud:31505/bludb
Done.

```
Out[14]:
```

DATE	landing__outcome	booster_version	launch_site
2015-01-10	Failure (drone ship)	F9 v1.1 B1012	CCAFS LC-40
2015-04-14	Failure (drone ship)	F9 v1.1 B1015	CCAFS LC-40
2015-06-28	Precluded (drone ship)	F9 v1.1 B1018	CCAFS LC-40
2016-01-17	Failure (drone ship)	F9 v1.1 B1017	VAFB SLC-4E
2016-03-04	Failure (drone ship)	F9 FT B1020	CCAFS LC-40
2016-04-08	Success (drone ship)	F9 FT B1021.1	CCAFS LC-40

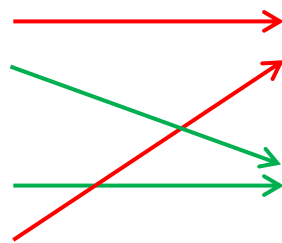
Rank Landing Outcomes Between 2010-06-04 and 2017-03-20

- Here below is presented a rank the count of landing outcomes in the period 2010-06-04 to 2017-03-20, in descending order:

```
In [24]: %sql SELECT LANDING__OUTCOME, COUNT(LANDING__OUTCOME) AS TOTAL FROM (SELECT * FROM SPACEXTBL WHERE DATE BETWEEN '2010-06-04' AND
* ibm_db_sa://srt20899:***@ea286ace-86c7-4d5b-8580-3fbfa46b1c66.bs2io90108kqb1od81cg.databases.appdomain.cloud:31505/bludb
Done.
```

Out[24]:

landing__outcome	total
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



Total failure landings = 7

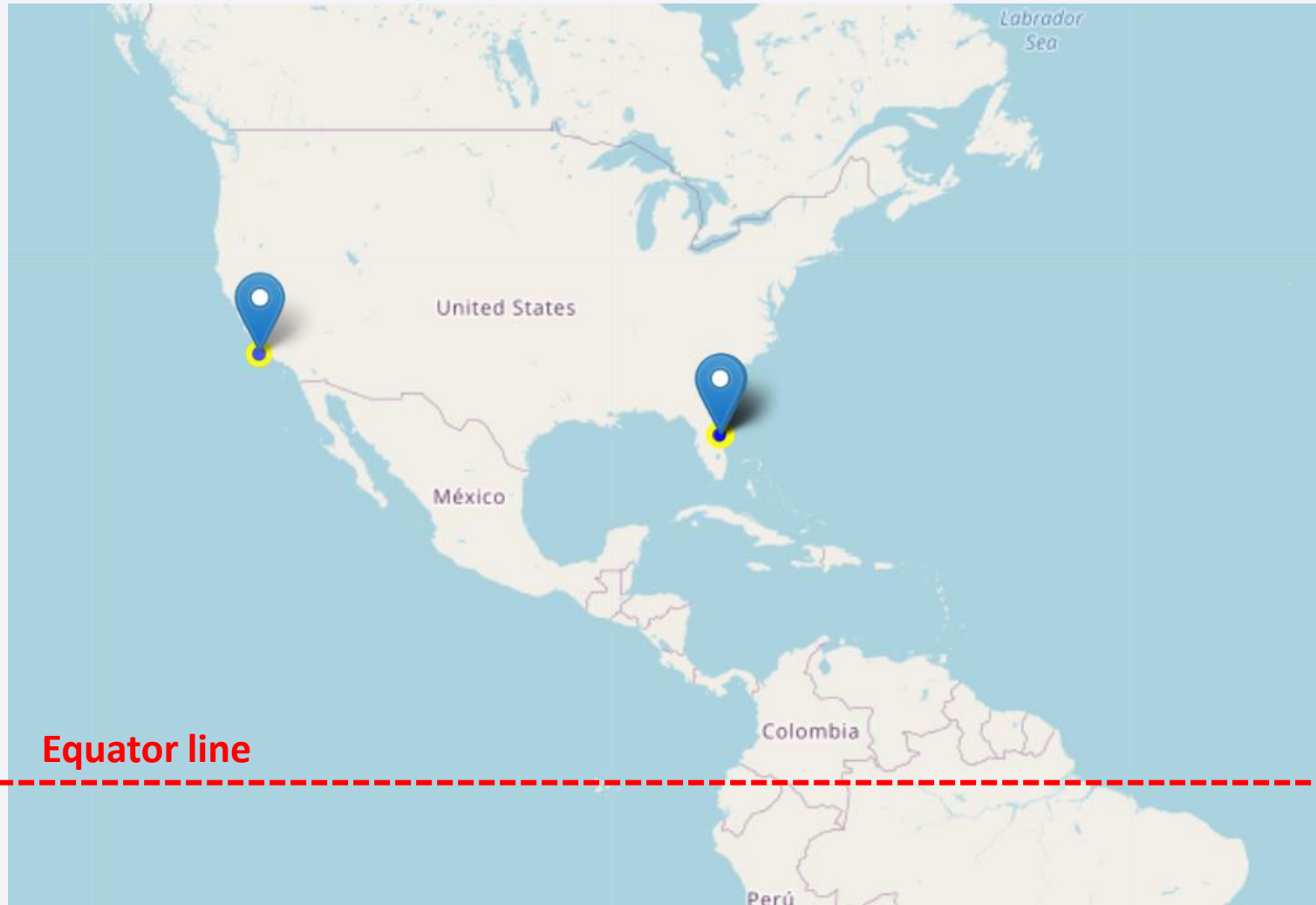
Total successful landings = 8

A satellite view of Earth from space, showing the curvature of the planet and city lights at night. The background is a deep blue gradient.

Section 3

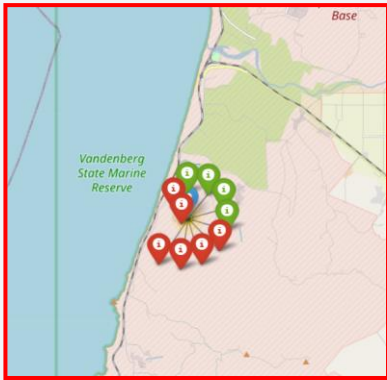
Launch Sites Proximities Analysis

Falcon 9 launch sites' location

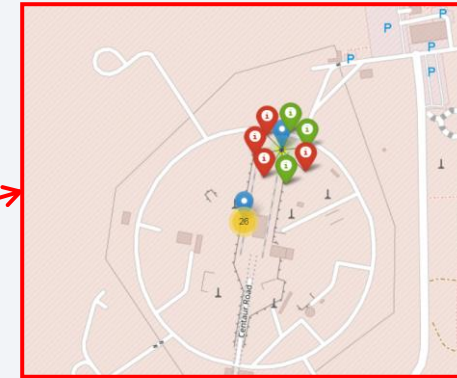
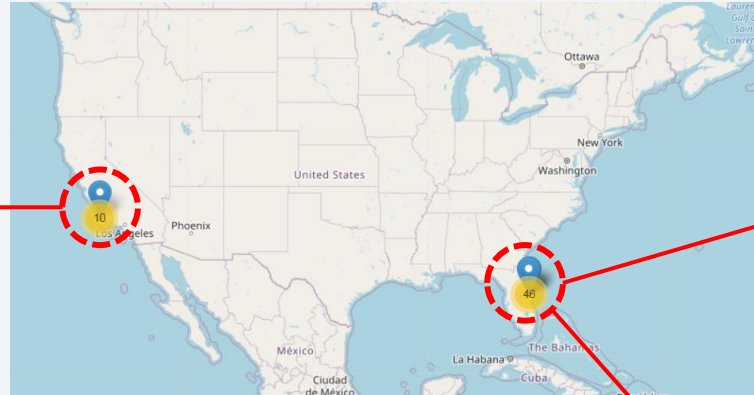


- Launch sites are located as close as possible to the Equator line.
- All launch sites are in very close proximity to the coast so that first stage can be thrown to the sea after the take-off.

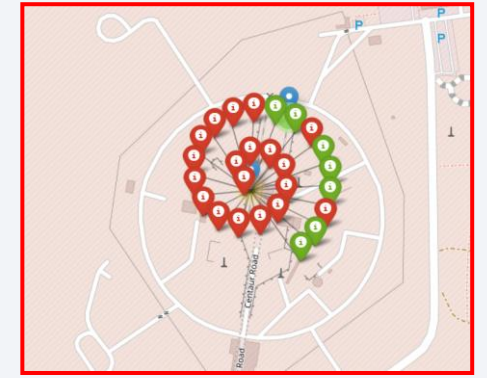
Launches success rate



VAFB SLC-4E

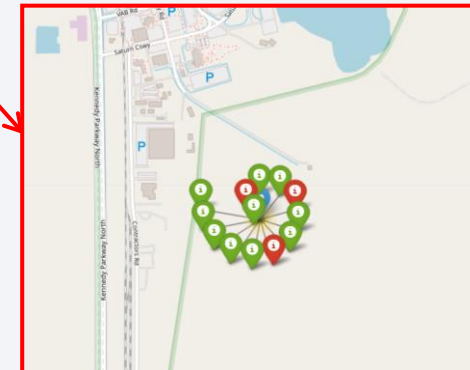


CCAFS SLC-40



CCAFS LC-40

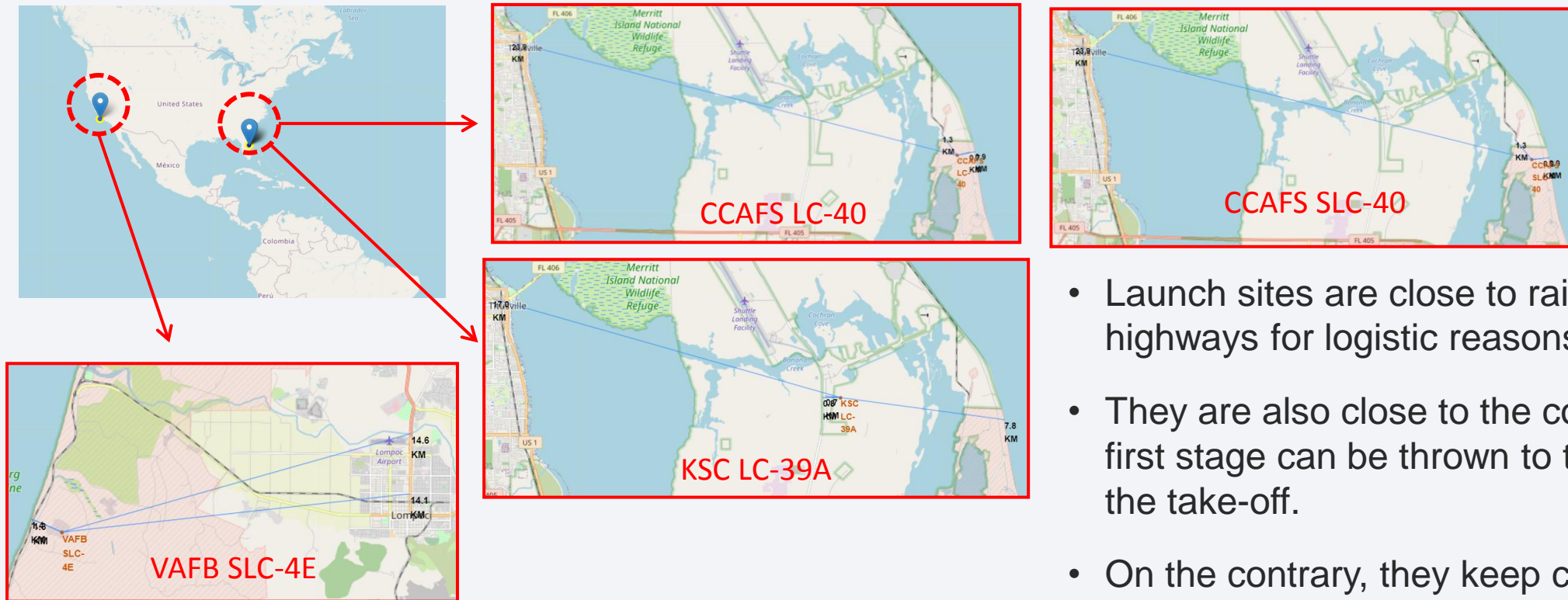
- The figures are color-labeled to show the launch outcomes on the map:
 - Green label indicates a successful launch.
 - Red label indicates a failure launch.
- Following success rates are derived from the figures:



KSC LC-39A

VAFB SLC-4E	CCAFS SLC-40	CCAFS LC-40	KSC LC-39A
40%	42.8%	26.9%	76.9%

Launch sites proximities



- Launch sites are close to railways and highways for logistic reasons.
- They are also close to the coastline so that first stage can be thrown to the sea after the take-off.
- On the contrary, they keep certain distance away from cities.

Launch Site	Distance to coast [km]	Distance to Railway [km]	Distance to Highway [km]	Distance to City [km]
CCAFS LC-40	0.93	1.34	0.66	23.92
CCAFS SLC-40	0.86	1.29	0.59	23.95
KSC LC-39A	7.77	0.72	0.84	17.05
VAFB SLC-4E	1.36	1.27	14.56	14.08



Section 4

Build a Dashboard with Plotly Dash

Launch success count

SpaceX Launch Records Dashboard

All Sites



Total Successful launches by site



- The graph shows that the major part of the successful launches have been carried out from the launch site KSC LC-39A.
- On the contrary, CCAFS SLC-40 presents the lowest number of successful launches.

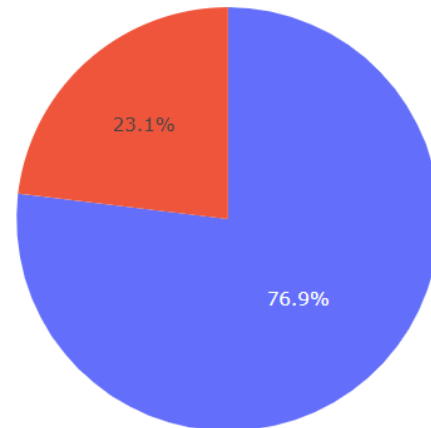
Launch success rate

SpaceX Launch Records Dashboard

KSC LC-39A

× ▼

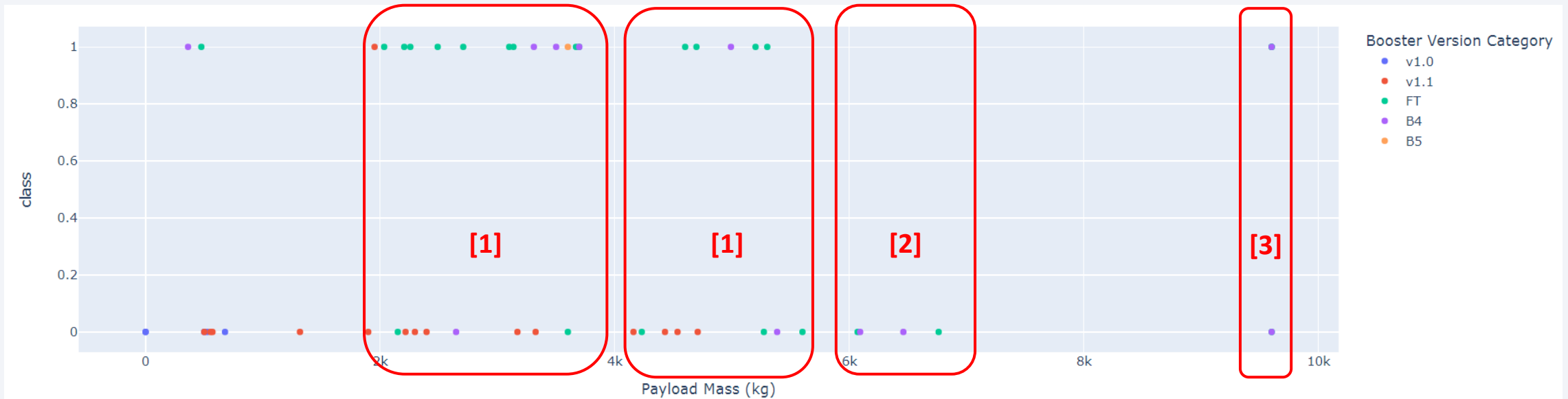
total successful launches for site KSC LC-39A



■ 1
■ 0

- The launch site with higher launch success rate (76.9%) is the KSC LC-39A (for more info about the rest of launch sites see Github link given in slide 15).
- Therefore KSC is not only the launch site with more successful launches but also with the highest success rate.

Correlation between payload and mission outcome



1. The highest launch success rates are observed for payload mass ranges [2000-4000] kg and [4500-5500] kg.
2. On the contrary, a success rate of 0% is obtained for payload masses in the range [6000-7000] kg.
3. Heavy payload missions (payload mass > 8000 kg) have a 100% success rate.

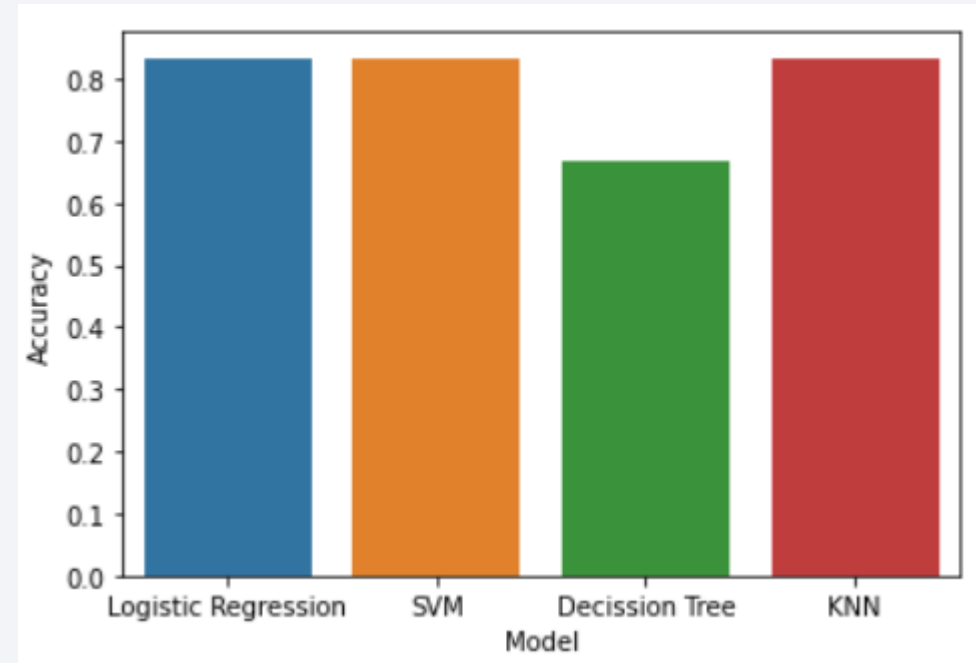


Section 5

Predictive Analysis (Classification)

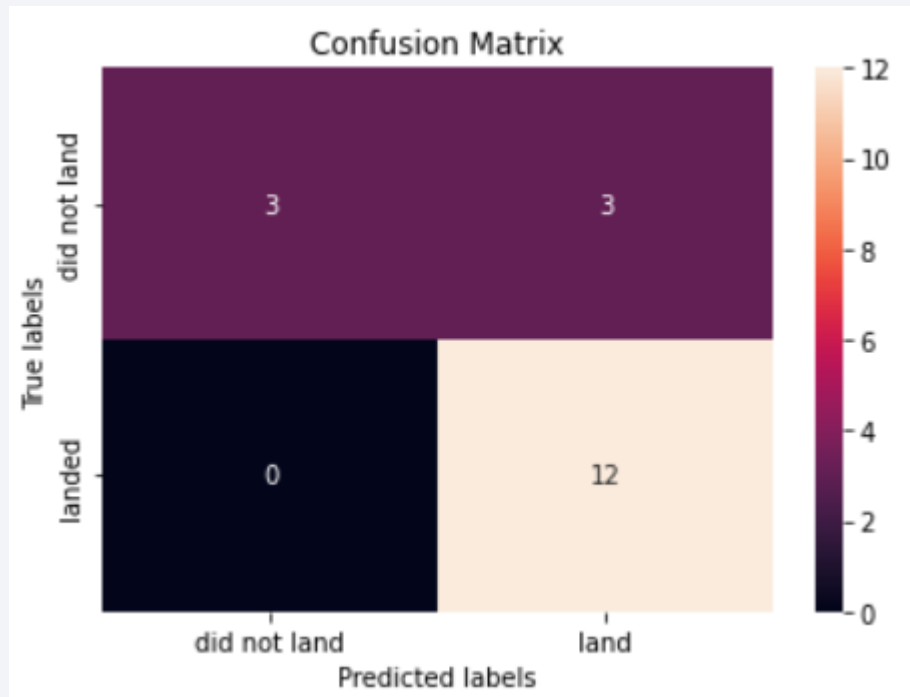
Classification Accuracy

- According to obtained results, logistic regression, SVM and KNN perform practically the same while decision tree classifier presents a lower accuracy.



Confusion Matrix

- The following figure shows the confusion matrix of the predictive models “logistic regression”, “SVM” and “KNN”:



- As can be seen, these models can distinguish between the different classes.
- The major problem is false positives (successful landing prediction when actually the mission failed). Models should be thus improved to avoid those false positives which would impact directly on the final cost of the mission.
- Confusion matrix of Decision Tree model can be found in appendix paragraph.

Conclusions

- Launch success rate since 2013 kept increasing till 2020. Successful mission rate awesomely high (98%), while successful landing rate is about 50%.
- Most part of recent launches were made from CCAFS with a success rate of 0.75% (32 launches) and from KSC with a success rate of 80%.
- Launch sites are located as close as possible to the Equator line and in very close proximity to the coast, railway and highway. On the contrary, they maintain a certain distance to the cities.
- Most part of the launches are carried out with a payload mass which varies from 2000 to 7000 kg with a high success rate. Heavy payload missions (payload mass > 8000 kg) have a 100% success rate.
- Most recent missions use a VLEO orbit and present a high successful rate of 86%.
- Boosters F9 FT have successfully landed on drone ship.
- Logistic regression, SVM and KNN provide accurate predictions on the landing success. They should be improved to avoid the false positives which would impact directly on the final cost of the mission.

Appendix

➤ “dataset_part_1.csv” file first rows:

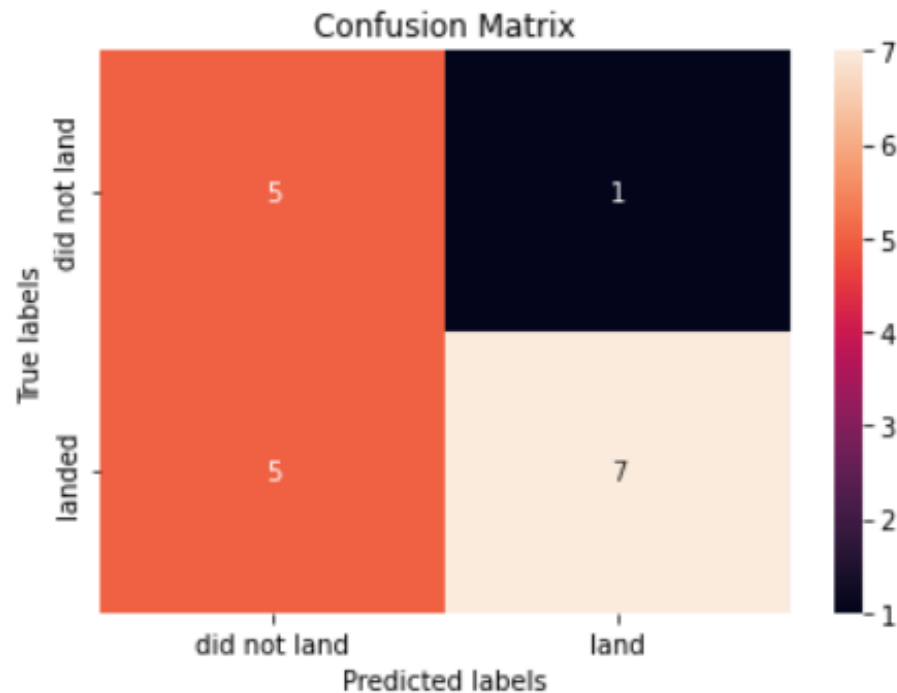
FlightNumber	Date	BoosterVersion	PayloadMass	Orbit	LaunchSite	Outcome	Flights	GridFins	Reused	Legs	LandingPad	Block	ReusedCount	Serial	Longitude	Latitude
1	04/06/2010	Falcon 9	6123.547647	LEO	CCSFS SLC 40	None None	1	FALSE	FALSE	FALSE		1	0	B0003	-80.577366	28.5618571
2	22/05/2012	Falcon 9	525	LEO	CCSFS SLC 40	None None	1	FALSE	FALSE	FALSE		1	0	B0005	-80.577366	28.5618571
3	01/03/2013	Falcon 9	677	ISS	CCSFS SLC 40	None None	1	FALSE	FALSE	FALSE		1	0	B0007	-80.577366	28.5618571
4	29/09/2013	Falcon 9	500	PO	VAFB SLC 4E	False Ocean	1	FALSE	FALSE	FALSE		1	0	B1003	-120.610829	34.632093
5	03/12/2013	Falcon 9	3170	GTO	CCSFS SLC 40	None None	1	FALSE	FALSE	FALSE		1	0	B1004	-80.577366	28.5618571

➤ “spacex.csv” file first rows:

Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS_KG_	Orbit	Customer	Mission_Outcome	Landing_Outcome
04/06/2010	18:45:00	F9 v1.0 B0003	CCAFS LC-40	Spacecraft Qualificati	0	LEO	SpaceX	Success	Failure (parachute)
08/12/2010	15:43:00	F9 v1.0 B0004	CCAFS LC-40	1, two CubeSats, barr	0	LEO (ISS)	NASA (COTS) NRO	Success	Failure (parachute)
22/05/2012	07:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525	LEO (ISS)	NASA (COTS)	Success	No attempt
08/10/2012	00:35:00	F9 v1.0 B0006	CCAFS LC-40	SpaceX CRS-1	500	LEO (ISS)	NASA (CRS)	Success	No attempt

Appendix

➤ Decision tree model confusion matrix:



- As can be seen, the decision tree model presents a low accuracy of 67%.
- Its main issue is false negatives (failed landing prediction when actually the landing takes place correctly). They could derive in a high mission price prediction which would not allow the company to enter competitively in the market.

Thank you!

