IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Emre Nuray
June 26, 2022

Outline

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

Executive Summary

- **Summary of methodologies**
  - Data collection with API and web scrapping
  - Data wrangling
  - Exploratory data analysis with SQL
  - Exploratory data analysis with visualization
  - Building interactive map using Folium
  - Building a dashboard with Plotly Dash
  - Applying machine learning techniques
- **Summary of all results**
  - Exploratory data analysis
  - Analysis with interactive map
  - Analysis of machine learning methods

Introduction

- **Project background and context**

The commercial space age is here, companies are making space travel affordable for everyone. The most successful among these companies is undoubtedly SpaceX. Of course, the biggest reason behind this success is the cost for customer.

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.

In this project, we will predict whether the first stage of Falcon 9 will land successfully. If we determine this correctly, we can also determine the cost of the launch.

- **Problems you want to find answers**
  - Relation between each rocket data
  - Finding best successful landing rate

Section 1

Methodology

Methodology

Executive Summary

- Data collection methodology:
    - SpaceX API
    - Web scraping from Wikipedia page of Falcon 9
- Perform data wrangling
    - Labelling outcomes respect to the boosters successful/fail landed data
- Perform exploratory data analysis (EDA) using visualization and SQL
- Perform interactive visual analytics using Folium and Plotly Dash
- Perform predictive analysis using classification models
    - SVM, Classification Tress, Logistic Regression

Data Collection

Data collected in two steps; SpaceX API and web scraping from Wikipedia pages of Falcon 9 and Falcon Heavy Launches Records from Wikipedia

- You need to present your data collection process use key phrases and flowcharts

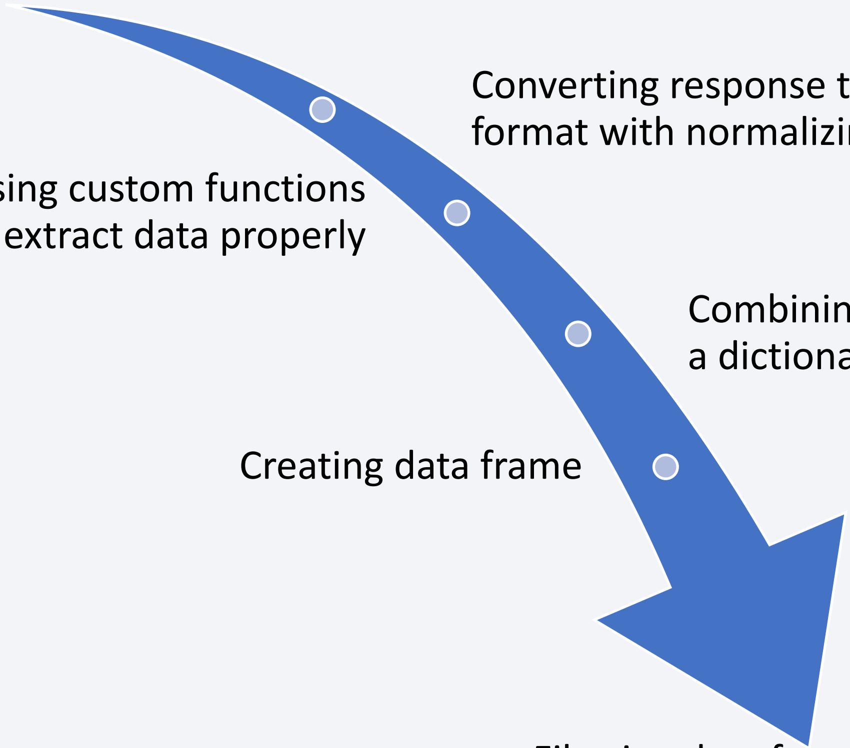
SpaceX API

1. Request data from API
2. API returns data in .JSON format
3. Normalize data into CSV format using Pandas library

Web Scraping

1. Get HTML response from Wikipedia page
2. Extract data using Beautiful Soup library
3. Normalize data into CSV format using Pandas library

Data Collection – SpaceX API

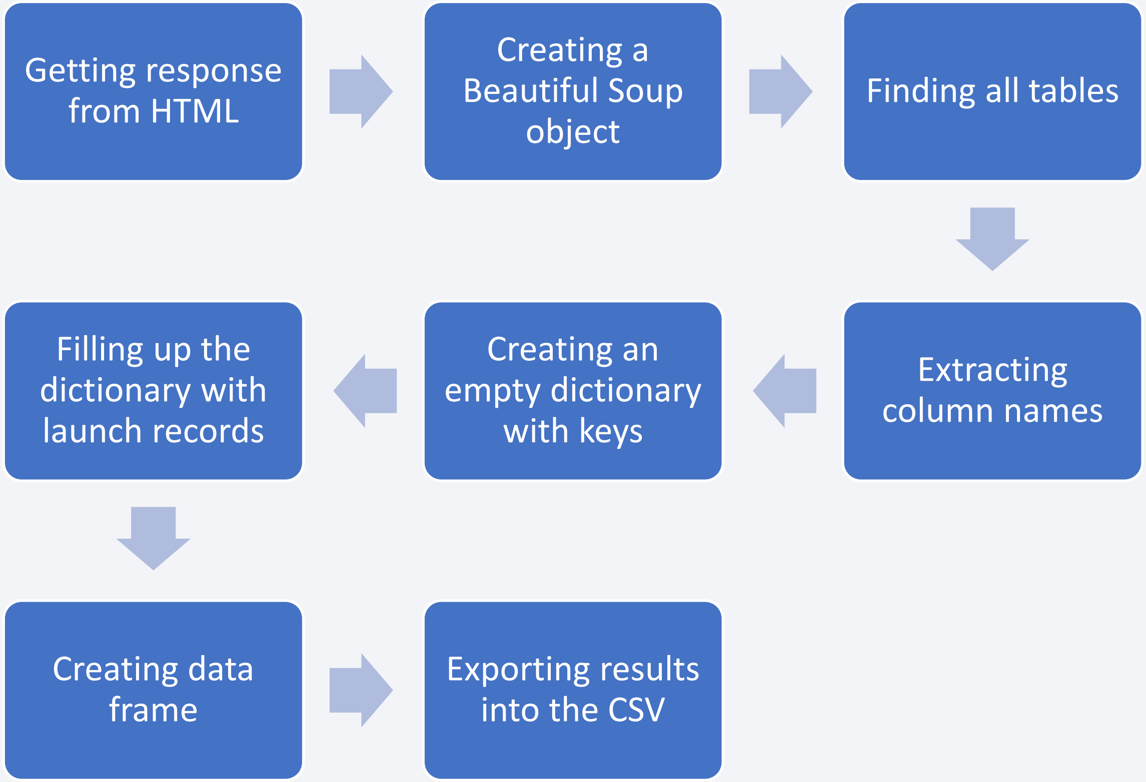


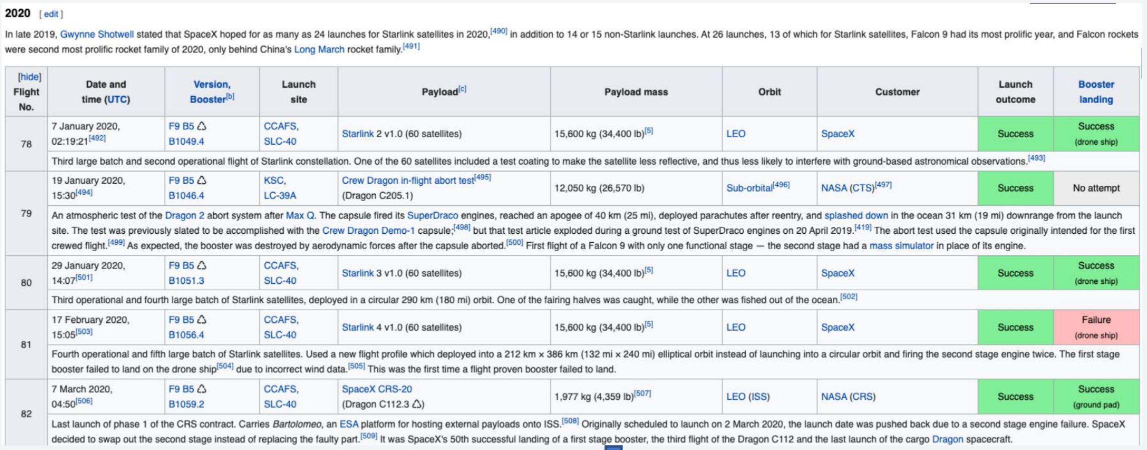
kbhqo/anasisii_launch_campaign_thread","launch":"https://www.reddit.com/r/spacex/comments/hu
a":"https://www.reddit.com/r/spacex/comments/hun4pv/rspacex_anasisii_media_thread_photograph
mments/hvgjk9/anasisii_recovery_thread/"},"flickr":{"small":[],"original":["https://live.sta
live.staticflickr.com/65535/50137510881_4618ba6c84_o.jpg","https://live.staticflickr.com/655
r.com/65535/50136967658_9347d7c575_o.jpg"]},"presskit":null,"webcast":"https://youtu.be/Tshv
eflightnow.com/2020/07/20/spacex-delivers-south-koreas-first-military-satellite-into-on-targ
7-11T17:58:00.000Z","static_fire_date_unix":1594490280,"net":false,"window":0,"rocket":"5e9c
s":"SpaceX will launch ANASIS-II, a South Korean geostationary military communication satell
ea\'s first dedicated military communications satellite. Falcon 9 will deliver the satellite
to land downrange on an ASDS.","crew":[],"ships":["5ea6ed2e080df4000697c908","5ea6ed2e080df4
oads":["5eb0e4d2b6c3bb0006eeb25b"],"launchpad":"5e9e4501f509094ba4566f84","flight_number":98
date_unix":1595280600,"date_local":"2020-07-20T17:30:00-04:00","date_precision":"hour","upco
ight":2,"gridfins":true,"legs":true,"reused":true,"landing_attempt":true,"landing_success":t
cc"}],"auto_update":true,"tbd":false,"launch_library_id":null,"id":"5eb87d50ffd86e000604b394
vered":true,"ships":["5ea6ed2e080df4000697c908","5ea6ed2e080df4000697c907"]},"links":{"patch
s://imgur.com/573IfGk.png"},"reddit":{"campaign":"https://www.reddit.com/r/spacex/comments/h
www.reddit.com/r/spacex/comments/i4ozw3/rspacex_starlink9_launch_discussion_updates/","media
starlink9_media_thread_photographer/","recovery":"https://www.reddit.com/r/spacex/comments/i
l":[],"original":["https://live.staticflickr.com/65535/50198901143_0bb53a499e_o.jpg","https:

	FlightNumber	Date	BoosterVersion	PayloadMass	Orbit	LaunchSite	Outcome	Flights	GridFins	Reused	Legs	LandingPad	Block	ReusedCount	Serial
4	1	2010-06-04	Falcon 9	NaN	LEO	CCSFS SLC 40	None None	1	False	False	False	None	1.0	0	B0003
5	2	2012-05-22	Falcon 9	525.0	LEO	CCSFS SLC 40	None None	1	False	False	False	None	1.0	0	B0005
6	3	2013-03-01	Falcon 9	677.0	ISS	CCSFS SLC 40	None None	1	False	False	False	None	1.0	0	B0007
7	4	2013-09-29	Falcon 9	500.0	PO	VAFB SLC 4E	False Ocean	1	False	False	False	None	1.0	0	B1003
8	5	2013-12-03	Falcon 9	3170.0	GTO	CCSFS SLC 40	None None	1	False	False	False	None	1.0	0	B1004
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
89	86	2020-09-03	Falcon 9	15600.0	VLEO	KSC LC 39A	True ASDS	2	True	True	True	5e9e3032383ecb6bb234e7ca	5.0	7	B1060
90	87	2020-10-06	Falcon 9	15600.0	VLEO	KSC LC 39A	True ASDS	3	True	True	True	5e9e3032383ecb6bb234e7ca	5.0	7	B1058
91	88	2020-10-18	Falcon 9	15600.0	VLEO	KSC LC 39A	True ASDS	6	True	True	True	5e9e3032383ecb6bb234e7ca	5.0	9	B1051
92	89	2020-10-24	Falcon 9	15600.0	VLEO	CCSFS SLC 40	True ASDS	3	True	True	True	5e9e3033383ecbb9e534e7cc	5.0	7	B1060
93	90	2020-11-05	Falcon 9	3681.0	MEO	CCSFS SLC 40	True ASDS	1	True	False	True	5e9e3032383ecb6bb234e7ca	5.0	2	B1062

[Click to go GitHub]

Data Collection - Scraping

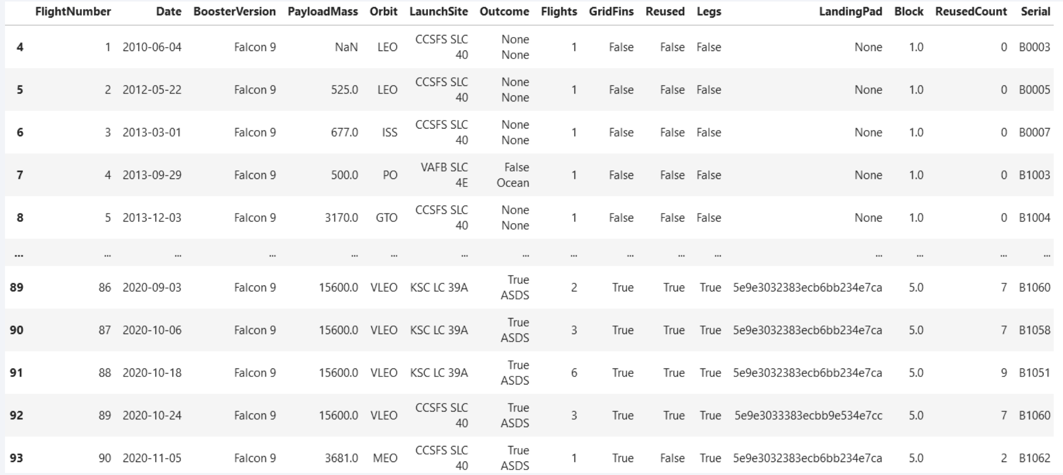




2020 [edit]

In late 2019, Gwynne Shotwell stated that SpaceX hoped for as many as 24 launches for Starlink satellites in 2020,[490] in addition to 14 or 15 non-Starlink launches. At 26 launches, 13 of which for Starlink satellites, Falcon 9 had its most prolific year, and Falcon rockets were second most prolific rocket family of 2020, only behind China's Long March rocket family.[491]

[hide] Flight No.	Date and time (UTC)	Version, Booster[b]	Launch site	Payload[c]	Payload mass	Orbit	Customer	Launch outcome	Booster landing
78	7 January 2020, 02:19:21[492]	F9 B5 ♺ B1049.4	CCAFS, SLC-40	Starlink 2 v1.0 (60 satellites)	15,600 kg (34,400 lb)[5]	LEO	SpaceX	Success	Success (drone ship)
	Third large batch and second operational flight of Starlink constellation. One of the 60 satellites included a test coating to make the satellite less reflective, and thus less likely to interfere with ground-based astronomical observations.[493]								
79	19 January 2020, 15:30[494]	F9 B5 ♺ B1046.4	KSC, LC-39A	Crew Dragon in-flight abort test[495] (Dragon C205.1)	12,050 kg (26,570 lb)	Sub-orbital[496]	NASA (CTS)[497]	Success	No attempt
	An atmospheric test of the Dragon 2 abort system after Max Q. The capsule fired its SuperDraco engines, reached an apogee of 40 km (25 mi), deployed parachutes after reentry, and splashed down in the ocean 31 km (19 mi) downrange from the launch site. The test was previously slated to be accomplished with the Crew Dragon Demo-1 capsule;[498] but that test article exploded during a ground test of SuperDraco engines on 20 April 2019.[419] The abort test used the capsule originally intended for the first crewed flight.[499] As expected, the booster was destroyed by aerodynamic forces after the capsule aborted.[500] First flight of a Falcon 9 with only one functional stage — the second stage had a mass simulator in place of its engine.								
80	29 January 2020, 14:07[501]	F9 B5 ♺ B1051.3	CCAFS, SLC-40	Starlink 3 v1.0 (60 satellites)	15,600 kg (34,400 lb)[5]	LEO	SpaceX	Success	Success (drone ship)
	Third operational and fourth large batch of Starlink satellites, deployed in a circular 290 km (180 mi) orbit. One of the fairing halves was caught, while the other was fished out of the ocean.[502]								
81	17 February 2020, 15:05[503]	F9 B5 ♺ B1056.4	CCAFS, SLC-40	Starlink 4 v1.0 (60 satellites)	15,600 kg (34,400 lb)[5]	LEO	SpaceX	Success	Failure (drone ship)
	Fourth operational and fifth large batch of Starlink satellites. Used a new flight profile which deployed into a 212 km × 386 km (132 mi × 240 mi) elliptical orbit instead of launching into a circular orbit and firing the second stage engine twice. The first stage booster failed to land on the drone ship[504] due to incorrect wind data.[505] This was the first time a flight proven booster failed to land.								
82	7 March 2020, 04:50[506]	F9 B5 ♺ B1059.2	CCAFS, SLC-40	SpaceX CRS-20 (Dragon C112.3 ♺)	1,977 kg (4,359 lb)[507]	LEO (ISS)	NASA (CRS)	Success	Success (ground pad)
	Last launch of phase 1 of the CRS contract. Carries *Bartolomeo*, an ESA platform for hosting external payloads onto ISS.[508] Originally scheduled to launch on 2 March 2020, the launch date was pushed back due to a second stage engine failure. SpaceX decided to swap out the second stage instead of replacing the faulty part.[509] It was SpaceX's 50th successful landing of a first stage booster, the third flight of the Dragon C112 and the last launch of the cargo Dragon spacecraft.								



	FlightNumber	Date	BoosterVersion	PayloadMass	Orbit	LaunchSite	Outcome	Flights	GridFins	Reused	Legs	LandingPad	Block	ReusedCount	Serial
4	1	2010-06-04	Falcon 9	NaN	LEO	CCSFS SLC 40	None None	1	False	False	False	None	1.0	0	B0003
5	2	2012-05-22	Falcon 9	525.0	LEO	CCSFS SLC 40	None None	1	False	False	False	None	1.0	0	B0005
6	3	2013-03-01	Falcon 9	677.0	ISS	CCSFS SLC 40	None None	1	False	False	False	None	1.0	0	B0007
7	4	2013-09-29	Falcon 9	500.0	PO	VAFB SLC 4E	False Ocean	1	False	False	False	None	1.0	0	B1003
8	5	2013-12-03	Falcon 9	3170.0	GTO	CCSFS SLC 40	None None	1	False	False	False	None	1.0	0	B1004
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
89	86	2020-09-03	Falcon 9	15600.0	VLEO	KSC LC 39A	True ASDS	2	True	True	True	5e9e3032383ecb6bb234e7ca	5.0	7	B1060
90	87	2020-10-06	Falcon 9	15600.0	VLEO	KSC LC 39A	True ASDS	3	True	True	True	5e9e3032383ecb6bb234e7ca	5.0	7	B1058
91	88	2020-10-18	Falcon 9	15600.0	VLEO	KSC LC 39A	True ASDS	6	True	True	True	5e9e3032383ecb6bb234e7ca	5.0	9	B1051
92	89	2020-10-24	Falcon 9	15600.0	VLEO	CCSFS SLC 40	True ASDS	3	True	True	True	5e9e3033383ecbb9e534e7cc	5.0	7	B1060
93	90	2020-11-05	Falcon 9	3681.0	MEO	CCSFS SLC 40	True ASDS	1	True	False	True	5e9e3032383ecb6bb234e7ca	5.0	2	B1062

[Click to go GitHub]

Data Wrangling

- **Defining cases**
  - `True Ocean`: the mission result has successfully landed in a specific area of the ocean
  - `False Ocean`: the mission result has not successfully landed in a specific area of the ocean
  - `True RTLS`: the mission result successfully landed on the ground pad
  - `False RTLS`: the mission result has not successfully landed on the ground pad
  - `True ASDS`: the mission result has successfully landed on the drone ship
  - `False ASDS`: the mission result has not landed on the drone ship
- **Labelling**
  - **0** → Failure
  - **1** → Successful

[Click to go GitHub]

Data Wrangling

- Calculating the number of launches at each site
- Calculating the number and occurrence of each orbit
- Calculating the number and occurrence of mission outcome per orbit type
- Creating a landing outcome label
- Calculating the success rate for every landing
- Exporting dataset to a CSV

LandingPad	Block	ReusedCount	Serial	Longitude	Latitude	Class
NaN	1.0	0	B0003	-80.577366	28.561857	0
NaN	1.0	0	B0005	-80.577366	28.561857	0
NaN	1.0	0	B0007	-80.577366	28.561857	0
NaN	1.0	0	B1003	-120.610829	34.632093	0
NaN	1.0	0	B1004	-80.577366	28.561857	0

Click to go GitHub

EDA with Data Visualization

Data Visualization

- Scatter Chart
  - Flight Number vs Launch Site
  - Payload vs Launch Site
  - Flight Number vs Orbit Type
  - Payload vs Orbit Type
- Bar Chart
  - Orbit Type vs Success Rate
- Line Chart
  - Year vs Success Rate

Purpose of charts is to show the relationship between the two axes.

Click to go GitHub

EDA with SQL

- Executing SQL queries to answer following questions

Unique launch sites in the space mission

Launch sites begin with the string 'CCA’

Total payload mass carried by boosters CRS

Average payload mass carried by booster version v1.1

Date when the first successful landing outcome in ground pad was achieved

Names of the boosters which have success in drone ship and have payload mass greater than 4000 but less than 6000

Total number of successful and failure mission outcomes

Names of the booster versions which have carried the maximum payload mass

Failed landing outcomes in drone ship, their booster versions, and launch site names for in year 2015

Ranking the count of landing outcomes between the date 2010-06-04 and 2017-03-20

[Click to go GitHub]

Build an Interactive Map with Folium

- **Objects created and added to a map**
  - Marker 1 → shows all launch sites on the map
  - Marker 2 → shows the success/failed launches for each site

- **Founded geographical patterns about launch sites**
  - The launch sites are close to the railways
  - The launch sites are close to the highways
  - The launch sites are close to the coastline
  - The launch sites are keep certain distance away from cities

Click to go GitHub

Build a Dashboard with Plotly Dash

Dashboard

Pie chart

Scatter chart

Shows total success launches by sites

It can be selected to indicate a successful landing distribution across all launch sites or to indicate the success rate of individual launch sites.

Shows the relationship between Outcomes and Payload mass by different boosters

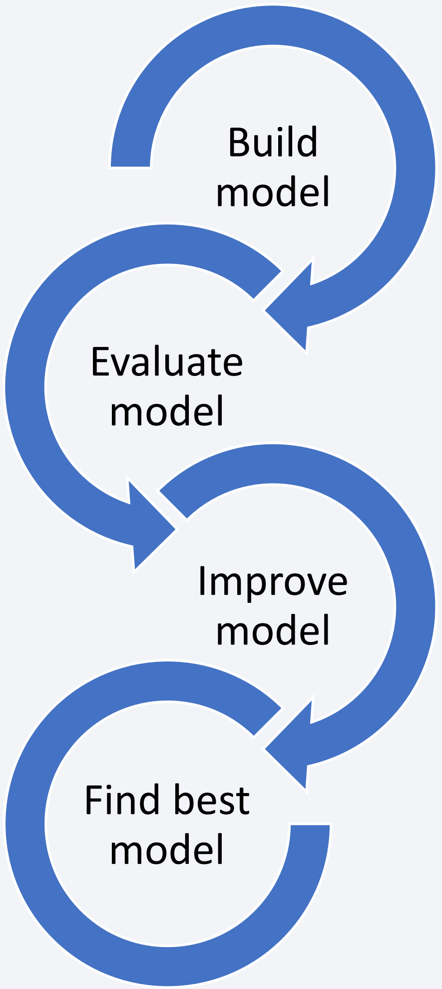
Helps determine how success depends on the launch point, payload mass, and booster version categories.

Click to go GitHub

Predictive Analysis (Classification)

- **Performing EDA and determining training labels**
    - Creating column for the class
    - Standardizing the data
    - Splitting into training and test dataset

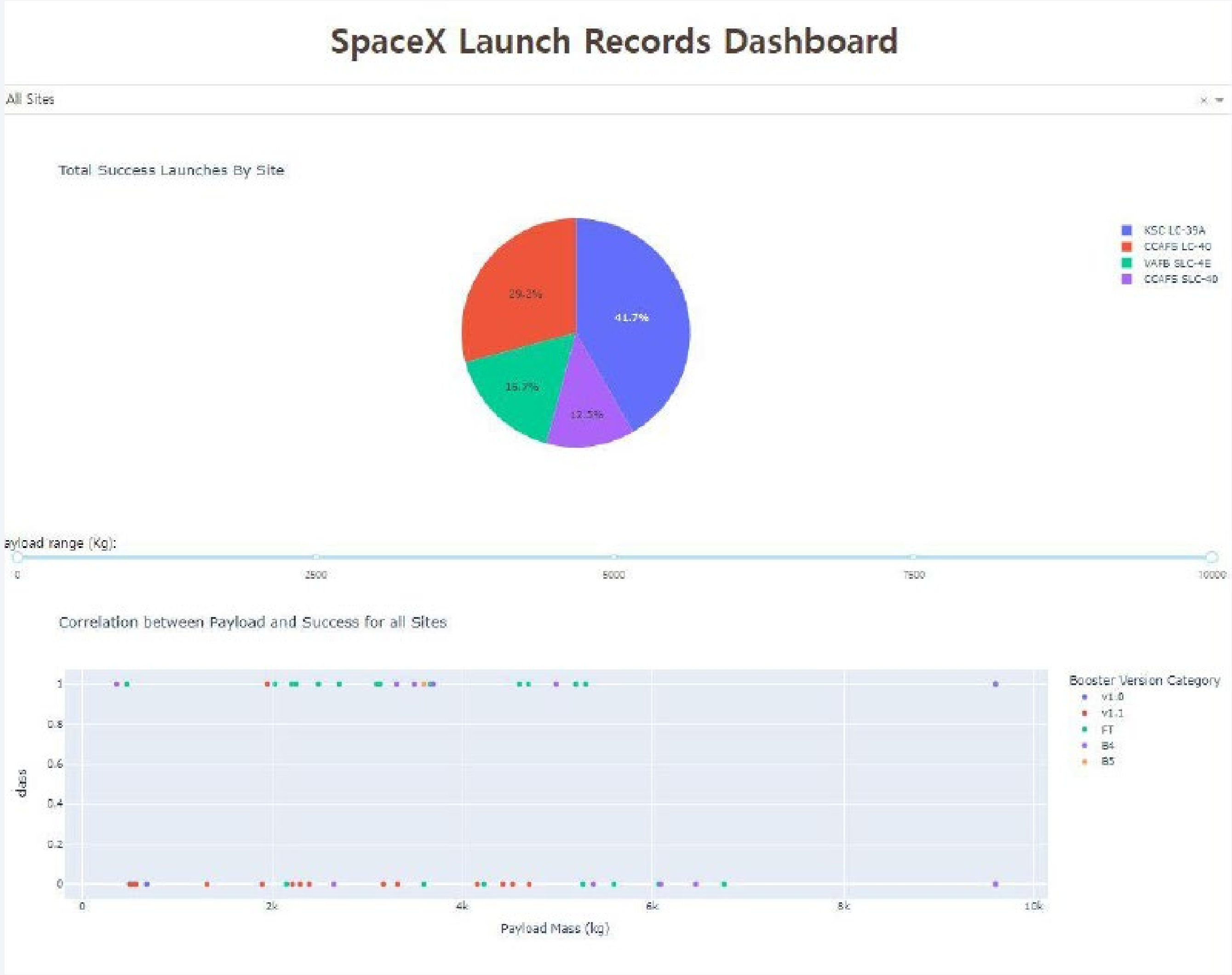
- **Applying machine learning techniques**
    - SVM
    - Classification Trees
    - Logistic Regression



Click to go GitHub

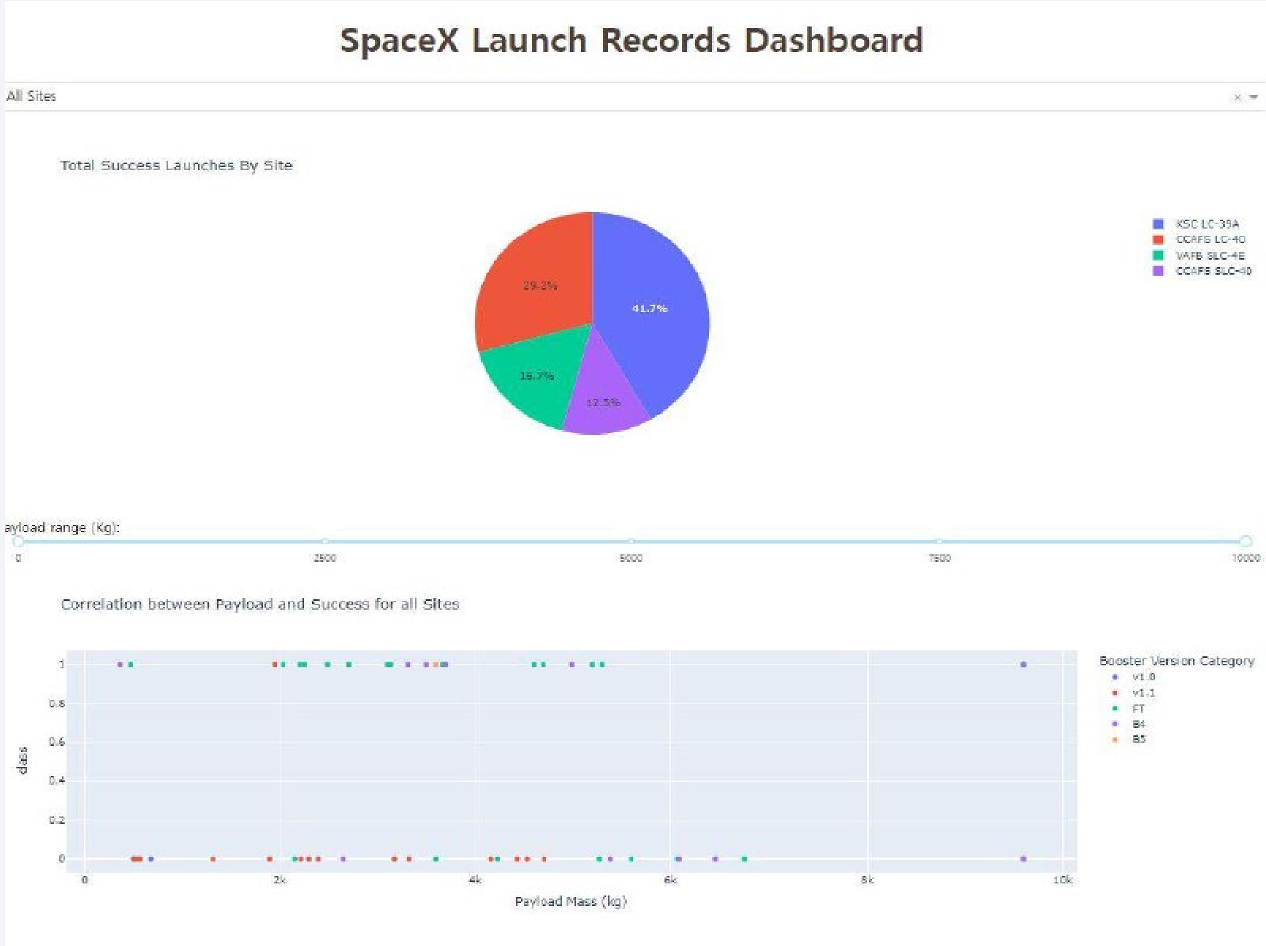
Results

Preview of Dashboard with Plotly Dash

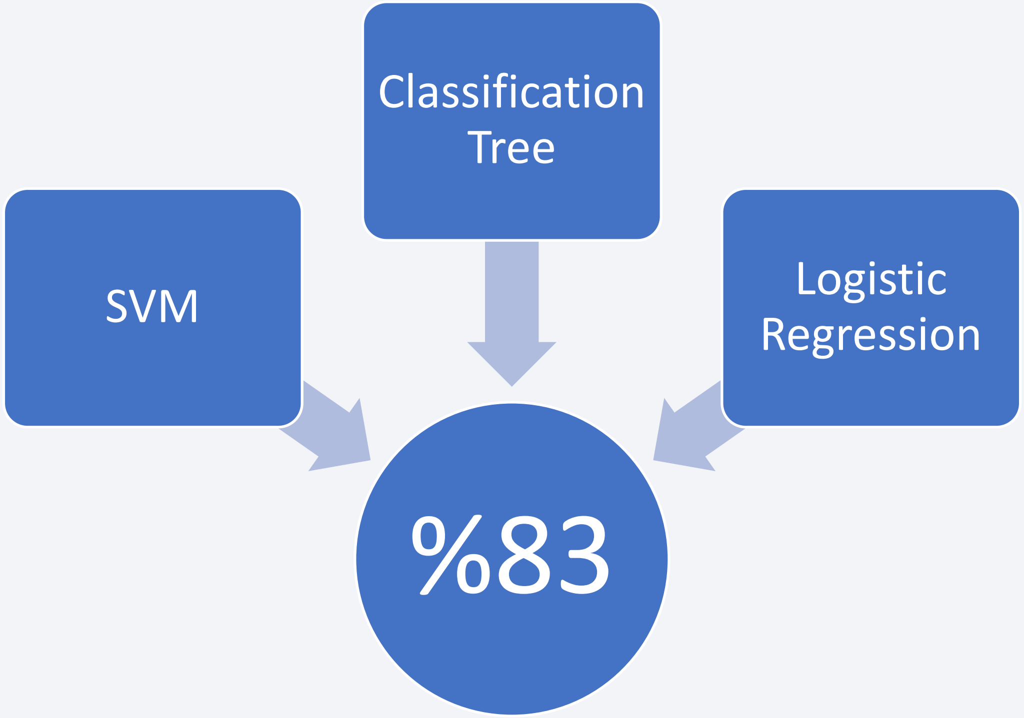


Results

Preview of Dashboard with Plotly Dash



Accuracy of Models



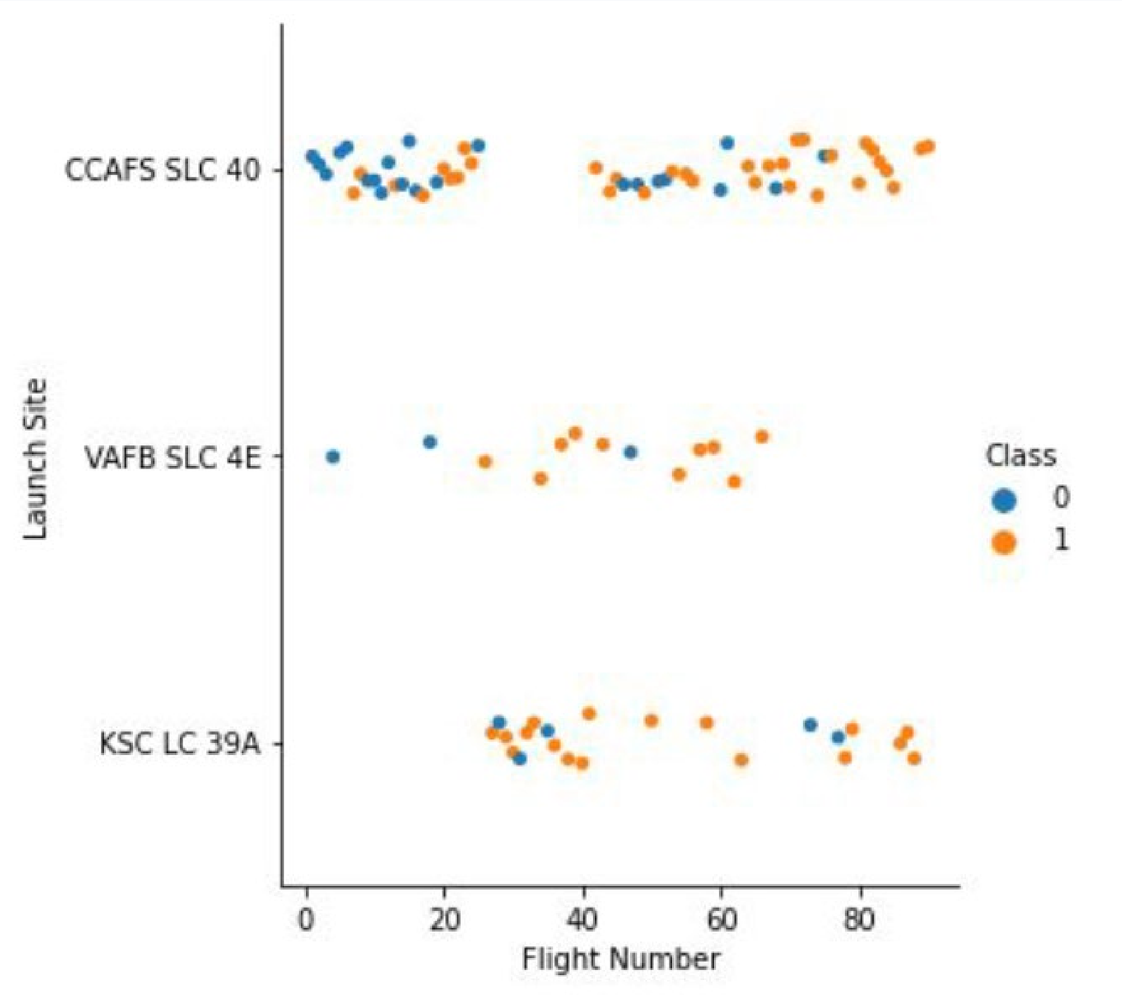
Section 2

Insights drawn from EDA

Flight Number vs. Launch Site

- Class 0 → Failed launch
- Class 1 → Successful launch

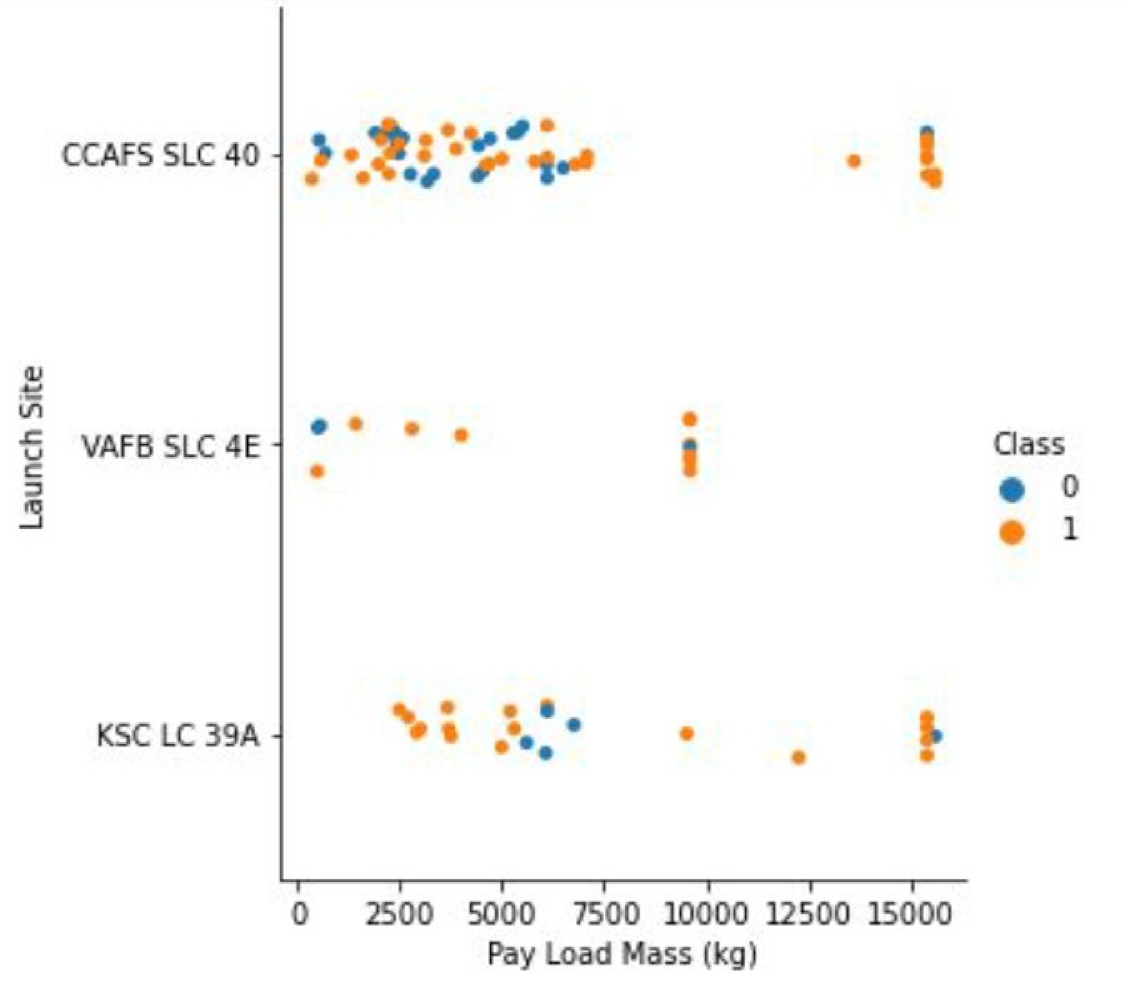
- Success rate and flight number are positively related to each other
- As the success rate has increased significantly since the 20th flight, this point seems to be a major breakthrough.



Payload vs. Launch Site

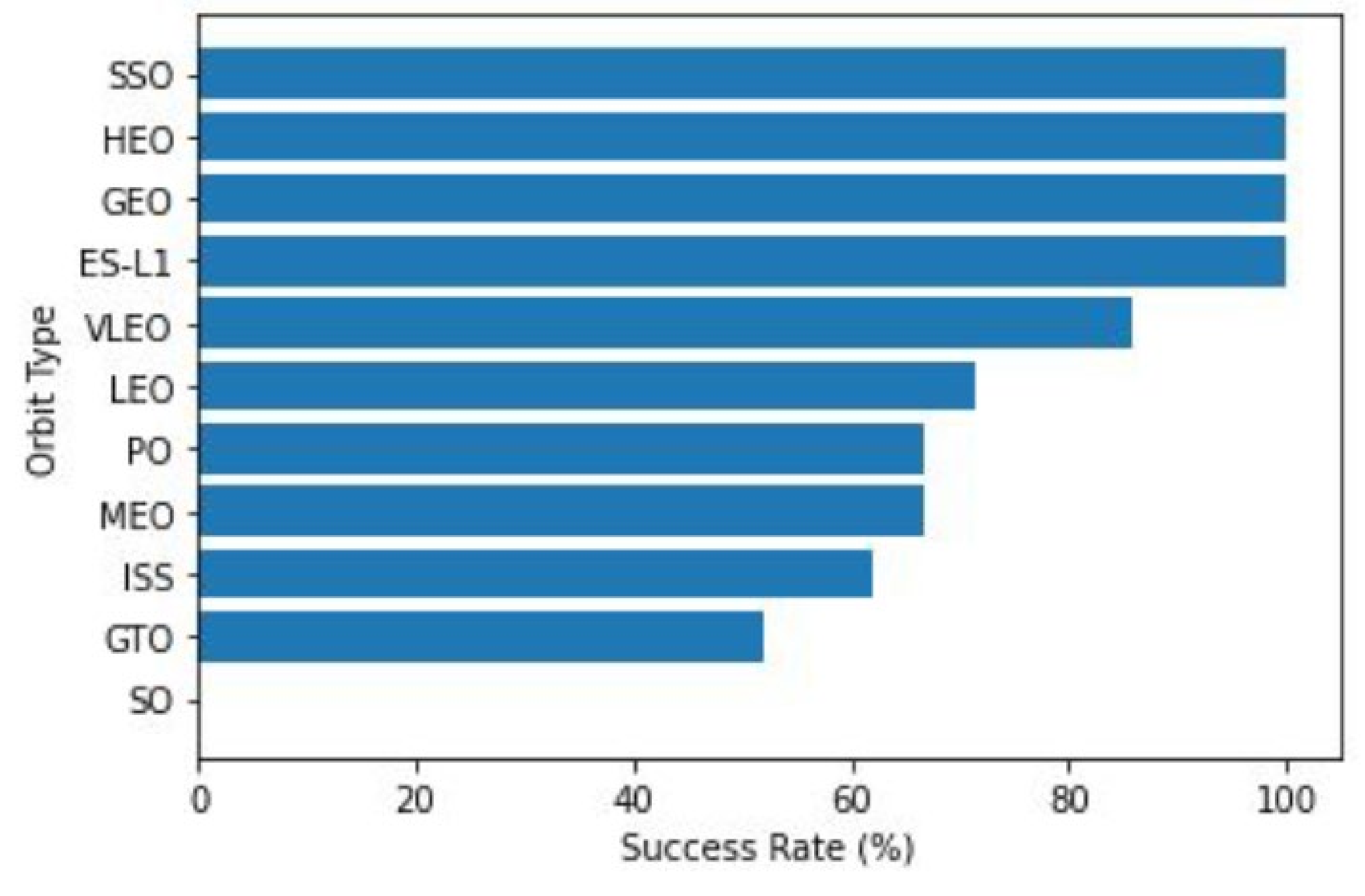
- Class 0 → Failed launch
- Class 1 → Successful launch

- Although larger pay load mass means higher success rate, there is no clear pattern between the successful launch and pay load mass



Success Rate vs. Orbit Type

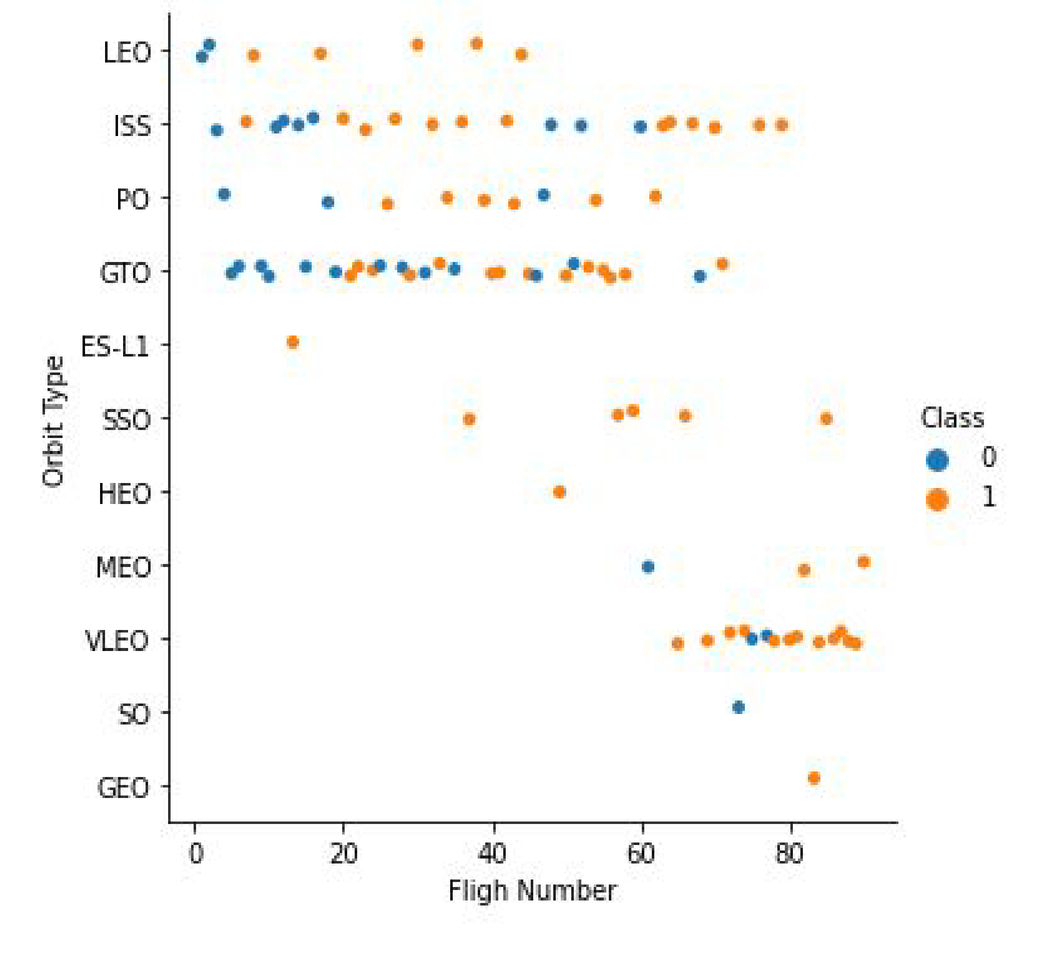
- SSO, HEO, GEO and ES-L1 type orbits have the best success rate which is 100%
- GTO orbits success rate is the lowest one which is around 50%
- SO orbit has no success rate for single attempt



Flight Number vs. Orbit Type

- Class 0 → Failed launch
- Class 1 → Successful launch

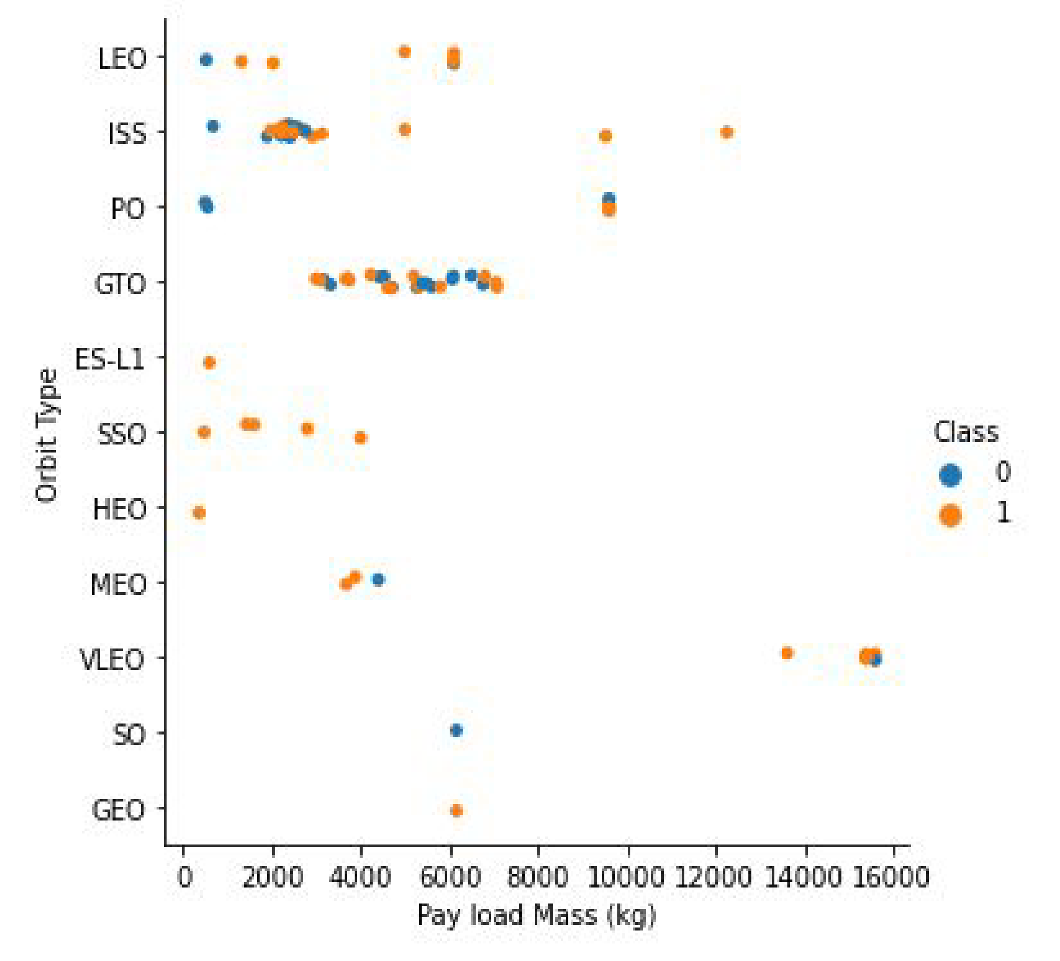
- General pattern is that the launch outcome has a positive relationship with the flight number
- However, for the GTO orbit, the situation is the opposite.
- SpaceX starts from the LEO orbit with an average success rate and chooses the VLEO orbit with the highest success rate towards final flights.



Payload vs. Orbit Type

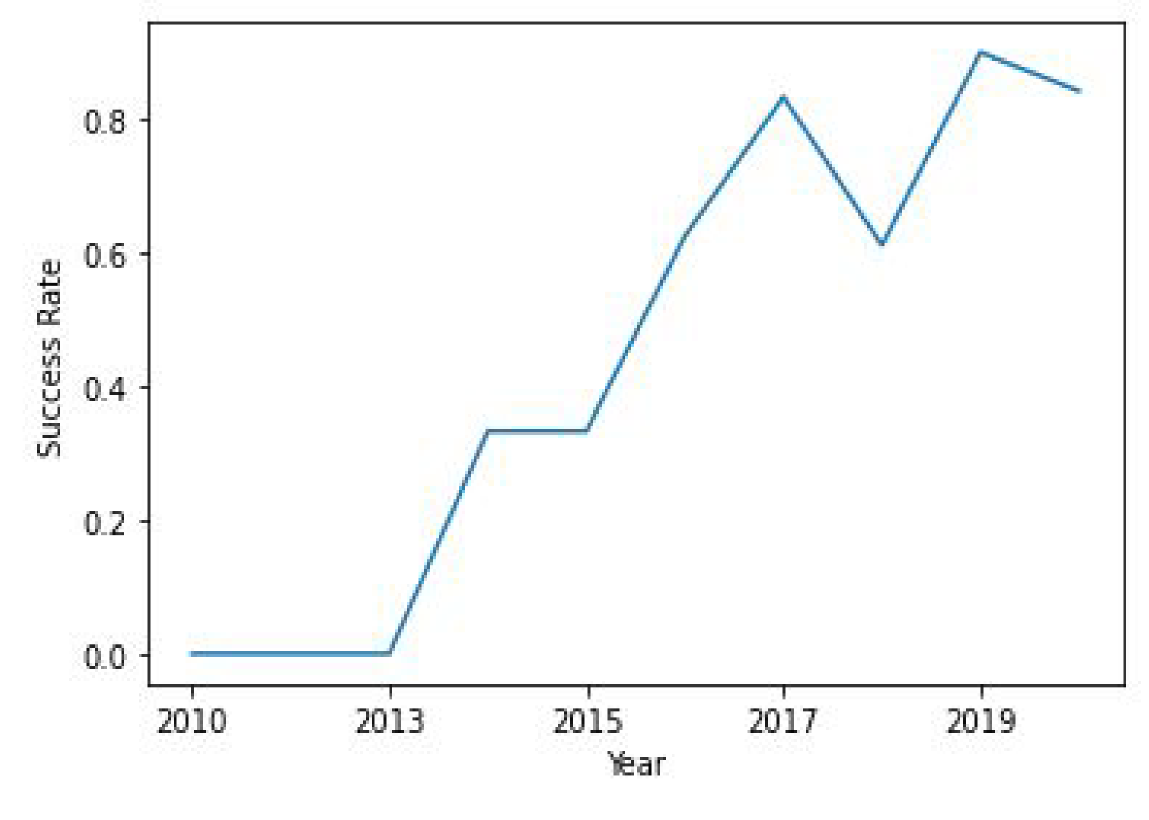
- Class 0 → Failed launch
- Class 1 → Successful launch

- It is seen that the success rate increases as the payload mass increases in the LEO and ISS orbits.
- However, it is difficult to predict whether the landing will be successful in the GTO orbit. There is no clear pattern.
- SSO orbit looks perfect for the low pay load mass rate.



Launch Success Yearly Trend

- Between 2010 and 2013, the trend is stable.
- From 2013 to 2017, the general trend is in the direction of increase.
- There is a decrease in the trend towards 2018.
- In 2019, the highest success rate, approximately 90%, was reached.



All Launch Site Names

- To the find the names of the unique launch sites, DISTINCT clause used in SQL query

SELECT DISTINCT LAUNCH_SITE FROM SPACEXTBL

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

- Here, 'SPACEXTBL' is the table name and LAUNCH_SITE is the column name in the table

Launch Site Names Begin with 'CCA'

- To the find 5 records where launch sites begin with `CCA`, LIKE operator and LIMIT clause used in the SQL query

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

DATE	time__utc_	booster_version	launch_site	payload	payload_mass__kg_	orbit	customer	mission_outcome	landing__outcome
2010-06-04	18:45:00	F9 v1.0 B0003	CCAFS LC-40	Dragon Spacecraft Qualification Unit	0	LEO	SpaceX	Success	Failure (parachute)
2010-12-08	15:43:00	F9 v1.0 B0004	CCAFS LC-40	Dragon demo flight C1, two CubeSats, barrel of Brouere cheese	0	LEO (ISS)	NASA (COTS) NRO	Success	Failure (parachute)
2012-05-22	07:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525	LEO (ISS)	NASA (COTS)	Success	No attempt
2012-10-08	00:35:00	F9 v1.0 B0006	CCAFS LC-40	SpaceX CRS-1	500	LEO (ISS)	NASA (CRS)	Success	No attempt
2013-03-01	15:10:00	F9 v1.0 B0007	CCAFS LC-40	SpaceX CRS-2	677	LEO (ISS)	NASA (CRS)	Success	No attempt

Total Payload Mass

- To the calculate the total payload carried by boosters from NASA, SUM() function used for the PAYLOAD_MASS_KG_ column and WHERE clause to filter boosters from CUSTOMER column.

```
SELECT SUM(PAYLOAD_MASS__KG_)
AS total_payload_mass_kg
FROM SPACEXTBL
WHERE CUSTOMER = 'NASA (CRS)'
```

- 45596

Average Payload Mass by F9 v1.1

- AVG() function used to calculate the average value of column PAYLOAD_MASS__KG_.
- To the filter dataset to perform calculations only if booster_version is F9 v1.1 WHERE clause is used.

```
SELECT AVG(PAYLOAD_MASS__KG_)
AS avg_payload_mass_kg
FROM SPACEXTBL
WHERE BOOSTER_VERSION = 'F9 v1.1'
```

- 2928

First Successful Ground Landing Date

- MIN() function used to find out the earliest date in the column DATE.
- To filter the dataset to perform a search only if landing_outcome is ‘Success’ WHERE clause is used.

```
SELECT MIN(DATE)
AS first_successful_landing_date
FROM SPACEXTBL
WHERE LANDING__OUTCOME
= 'Success (ground pad)'
```

- 2015-12-22

Successful Drone Ship Landing with Payload between 4000 and 6000

- To filter the dataset to perform a search if landing__outcome is 'Success (drone ship)' WHERE clause is used with the AND operator is that display record which PAYLOAD_MASS__KG_ between 4000 and 6000.

```
SELECT BOOSTER_VERSION
FROM SPACEXTBL
WHERE LANDING__OUTCOME = 'Success (drone ship)'
AND (PAYLOAD_MASS__KG_ BETWEEN 4000 AND 6000)
```

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

Total Number of Successful and Failure Mission Outcomes

- To calculate the total number of successful and failure mission outcomes COUNT() function used with GROUP BY statement which is groups rows that have the same values into summary rows to find the total number in each mission_outcome.
- Nearly 99% missions are successfully completed.

```
SELECT
MISSION_OUTCOME,
COUNT(*) AS total_number
FROM SPACEXTBL
GROUP BY
MISSION_OUTCOME
```

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

Boosters Carried Maximum Payload

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

F9 B5 B10xx boosters could carried the maximum payload.

```
SELECT DISTINCT BOOSTER_VERSION,
PAYLOAD_MASS__KG_
FROM SPACEXTBL
WHERE PAYLOAD_MASS__KG_ = (
SELECT MAX(PAYLOAD_MASS__KG_)
FROM SPACEXTBL)
```

- To list the names of the booster which have carried the maximum payload mass subquery is used.
- Firstly using the MAX(), maximum value of payload mass calculated and after that to perform a search if PAYLOAD_MASS_KG__ is the maximum value of payload mass the dataset is filtered.

2015 Launch Records

- To filter the dataset to perform a search if landing_outcome is ‘Failure (drone ship)’, WHERE clause is used with AND operator to display records in 2015.

```
SELECT LANDING__OUTCOME,
BOOSTER_VERSION,
LAUNCH_SITE
FROM SPACEXTBL
WHERE LANDING__OUTCOME
= 'Failure (drone ship)'
AND YEAR(DATE) = '2015'
```

landing__outcome	booster_version	launch_site
Failure (drone ship)	F9 v1.1 B1012	CCAFS LC-40
Failure (drone ship)	F9 v1.1 B1015	CCAFS LC-40

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

- WHERE clause filter the dataset to search if date is between 2010-06-04 and 2017-03-20.
- Using ORDER BY method with DESC, sort the records in descending order.
- Results shows the number of successes and failures between defined date was similar.

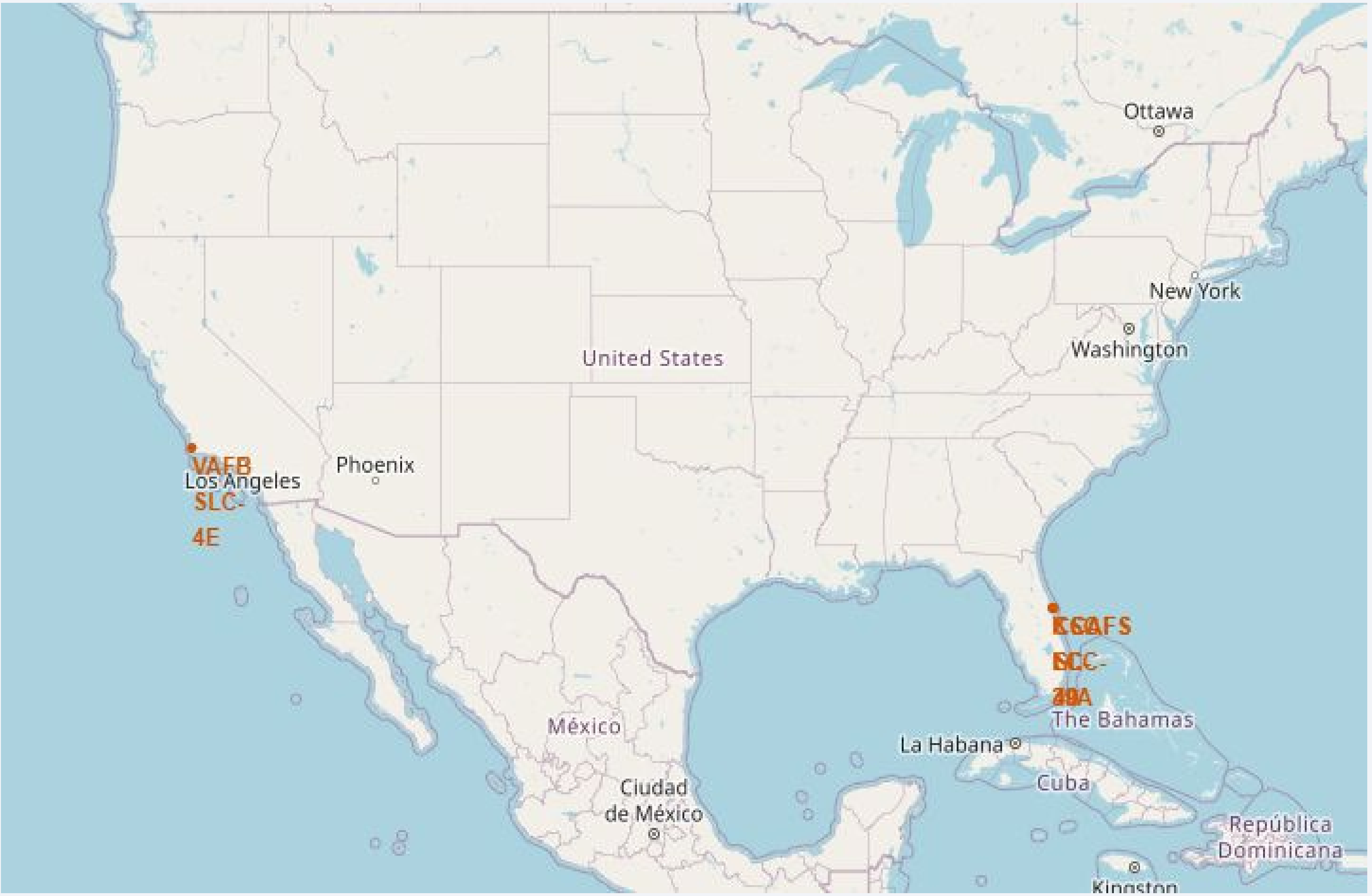
```
SELECT LANDING__OUTCOME,
COUNT(LANDING__OUTCOME) AS total_number
FROM SPACEXTBL
WHERE DATE BETWEEN '2010-06-04' AND '2017-03-20'
GROUP BY LANDING__OUTCOME
ORDER BY total_number DESC
```

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

Section 3

Launch Sites Proximities Analysis

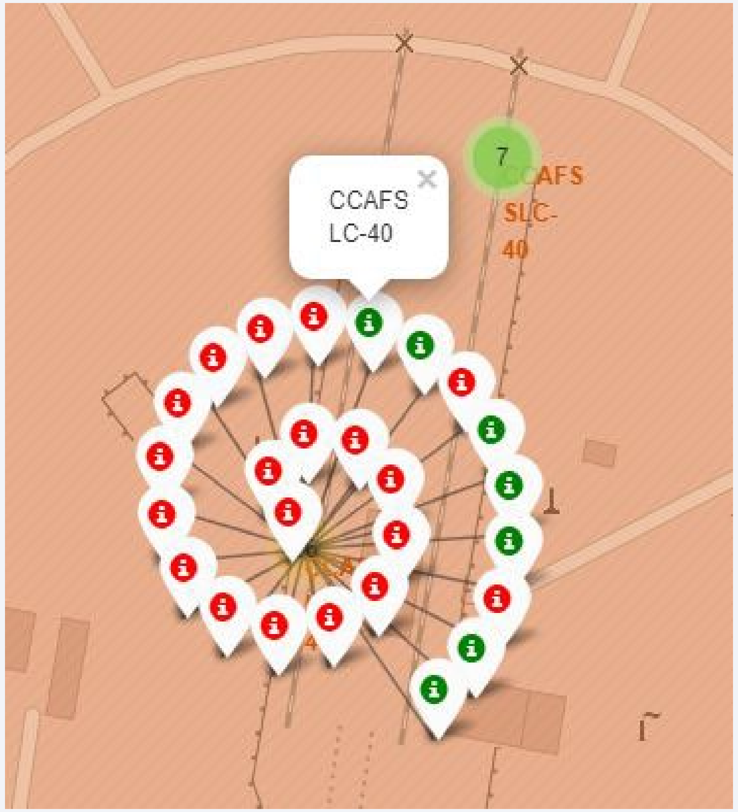
Locations of Launch Sites

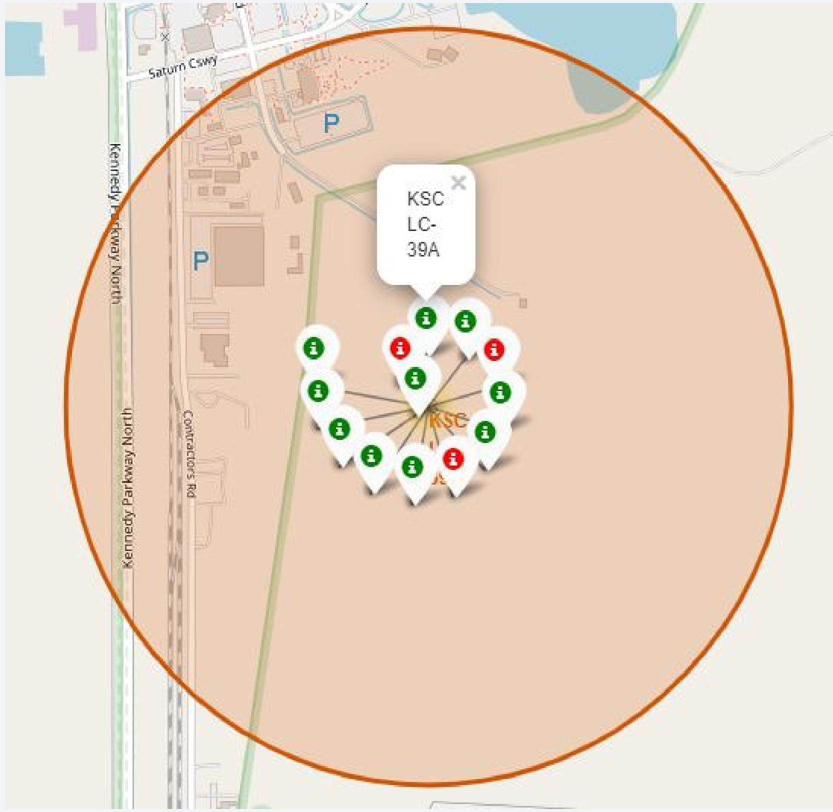


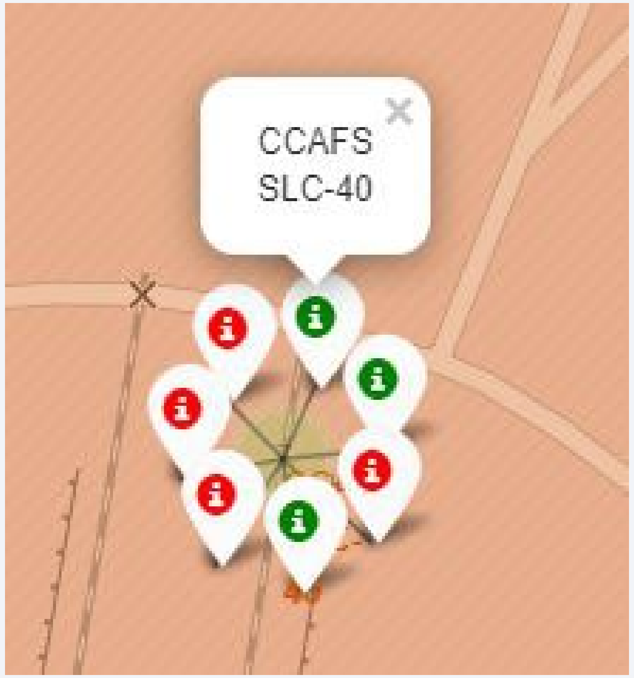
- Launch sites are located in United States.
- Launch sites are near the coast.

Launch Outcomes with Color Label

- Green color represents successful landing
- Red color represents failed landing



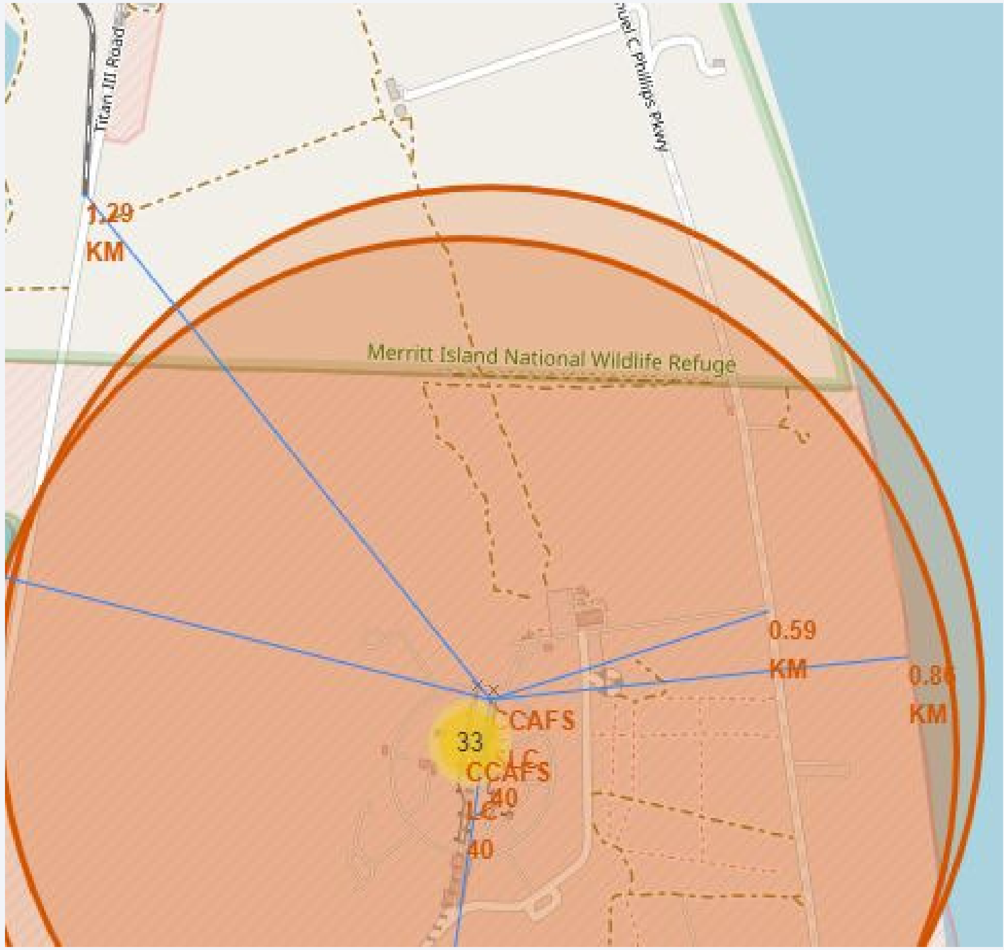




Proximities of Launch Sites

- Three questions are investigated:
  - Is it close proximity to railways? (For the transportation)
  - Is it close proximity to highways? (For the transportation)
  - Is it close proximity to coastline? (For the safety)

- For the right map, all three questions have same answer which is Yes.
  - So, it is <u>good</u> choice for the launch sites.

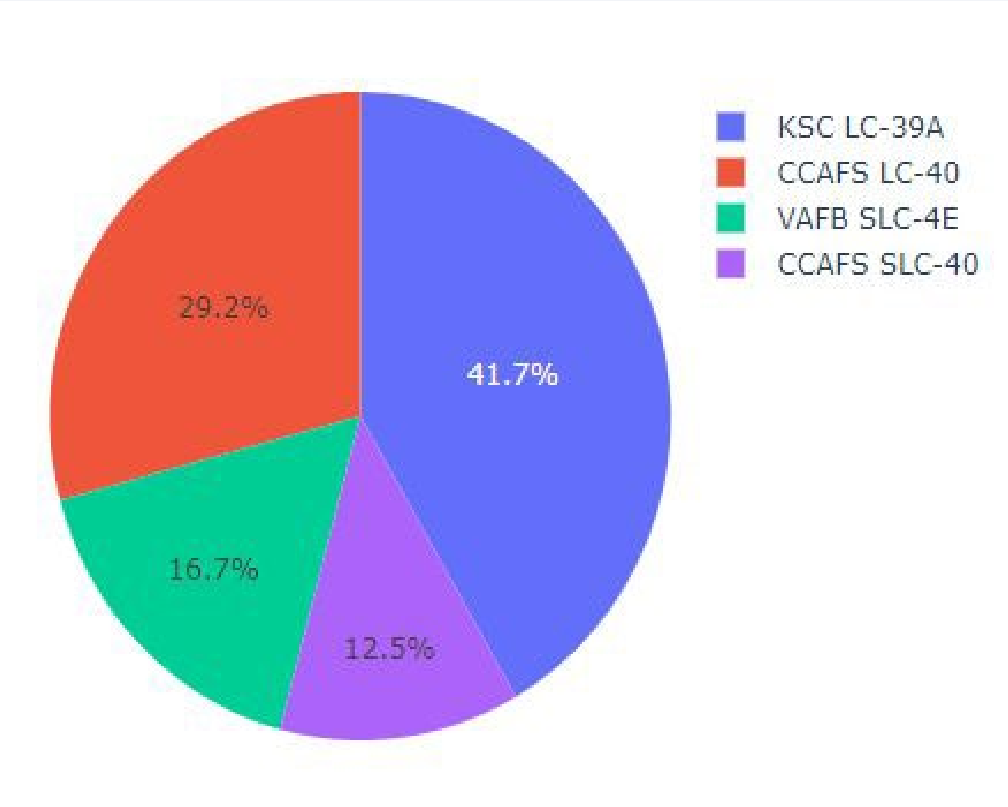


Section 4

Build a Dashboard with Plotly Dash

Success of Launches by All Sites

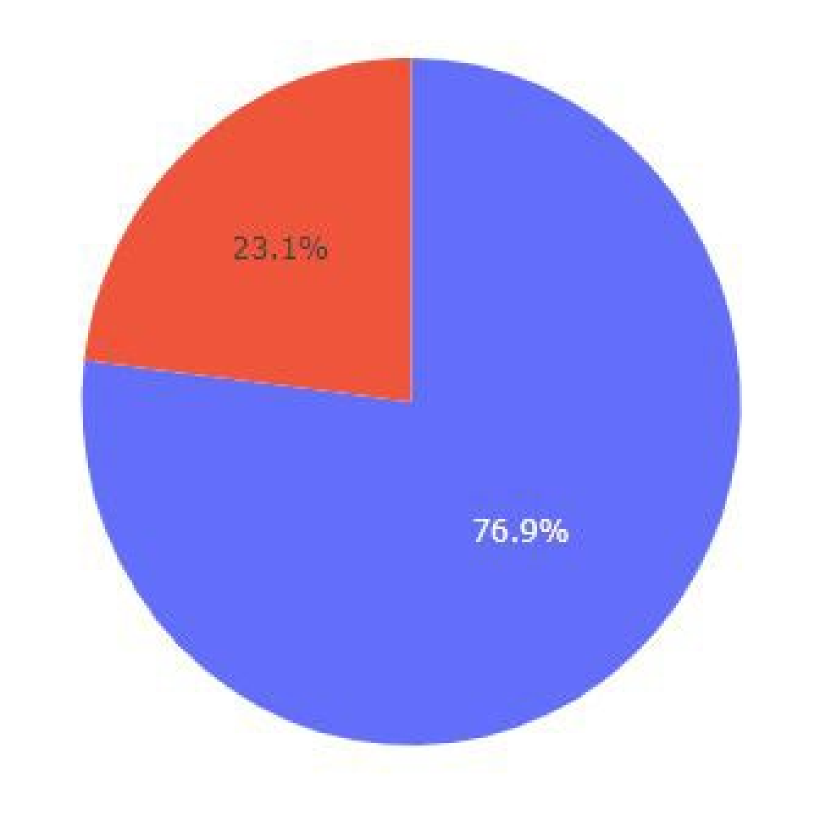
- Although KSLC-39A is the most successful launch, CCAFS SLC-40 is the most unsuccessful launch among all sites.
- The VAFB SLC-4E has the somewhat successful launch, possibly due to
  - This launch is located only California. Possibly, we can interpret from here west coast may be higher difficulty than east coast.



Launch Site with Highest Launch Success Ratio

- Blue represents 'Success'
- Red represents 'Failure'

- Pie charts shows the success rate for KSLC-39A site.
- Total landing number is 13.
  - 10 of them are successful landings
  - 3 of them are failed landings



Payload vs Launch Outcome

- Let's investigate the situation into the two part
  - For low weighted payloads (0 – 5000 kg)
  - For high weighted payloads (5000 kg – 10000 kg)

For the low part

Payload vs Launch Outcome

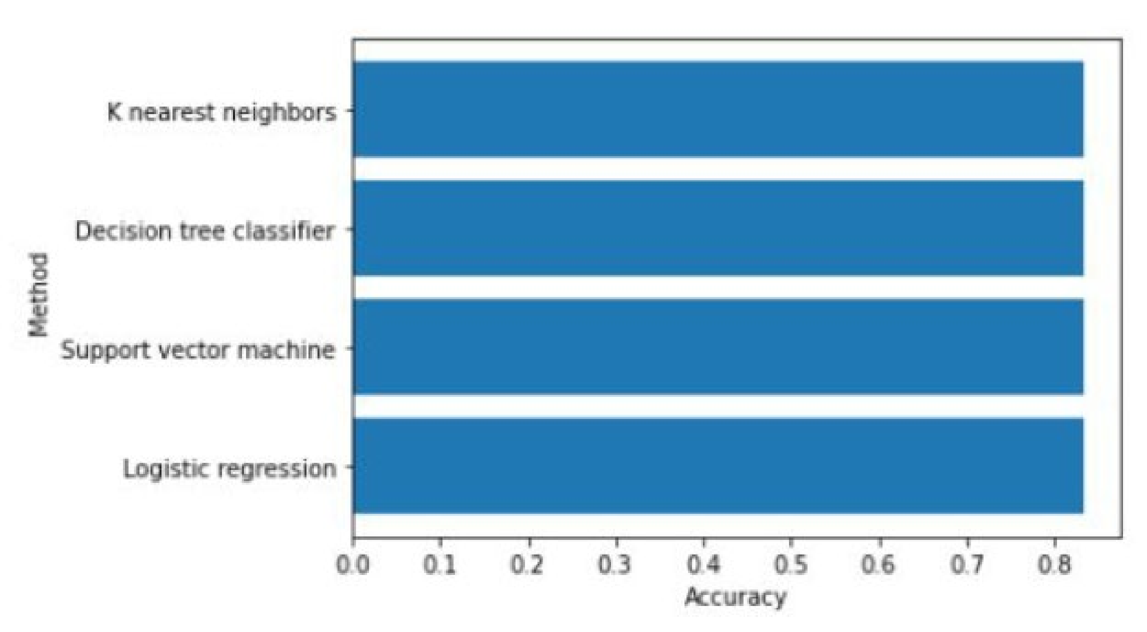
For the high part

Section 5

Predictive Analysis (Classification)

Classification Accuracy

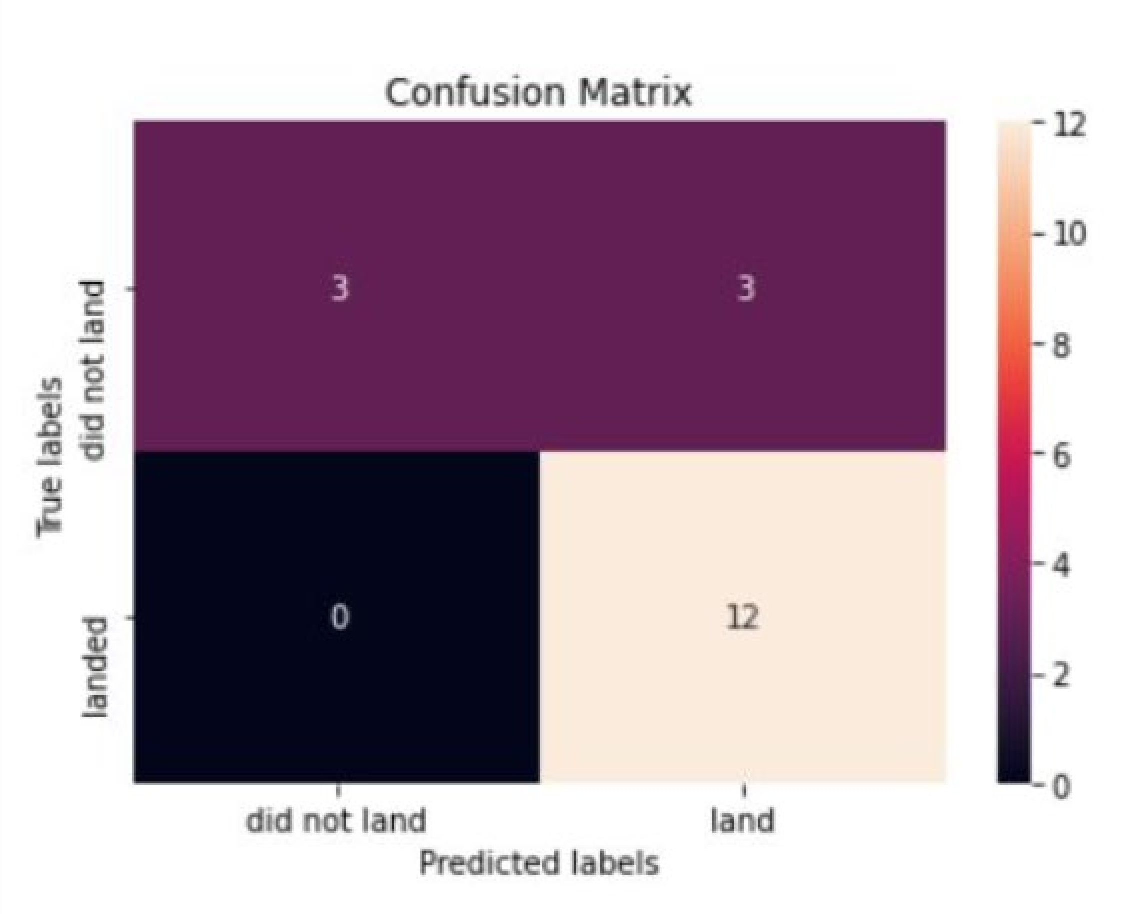
- For the test set all classification models have the same accuracy which is 83.3%
- Therefore there is no best model for this data set
  - Because dataset are very small
  - To choice the best model, mora data is needed.



	Method	Accuracy
0	Logistic regression	0.833333
1	Support vector machine	0.833333
2	Decision tree classifier	0.833333
3	K nearest neighbors	0.833333

Confusion Matrix

- Confusion matrix is same for all classification models because of the they have same accuracy rate.
- The models predict;
  - 12 successful landing when the true label was successful
  - 3 failed landing when the true label was successful
  - 3 successful landing when the true label was failure
- Finally, the success rate is 15/18 which is 83.3%



Conclusions

1
- It was determined that the success rate increased as the number of flights increased. In fact, this rate has exceeded 80% in 2019.

2
- The orbital types with the maximum success rate are ES-L1, GEO, HEO and SSO.

3
- The launch with the highest success rate among all sites is KSLC-39A.

4
- The launch sites are close to highways, railways and coastline but far from cities due to the safety of city.

5
- Low weighted payloads are higher success rate than high weighted payloads.

6
- Defined dataset, all classification models are the same accuracy which is 83.3%

Appendix

- Coursera Course URL

Thank you!