IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

<Name> Ryuzo Hayahsi
<Date> 2023/08/04

Outline

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

Executive Summary

- Summary of methodologies
  - Webscriping
  - Data cleaning
  - Data visualization
  - Machine learning
- Summary of all results

Introduction

- determine the best launch
- What launches have the highest success rate?

Section 1

Methodology

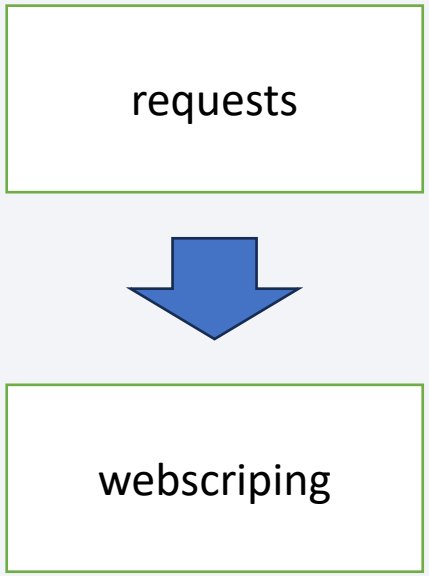
Methodology

Executive Summary

- Data collection methodology:
  - webscriping
- Perform data wrangling
  - Nan data is deleted
- Perform exploratory data analysis (EDA) using visualization and SQL
- Perform interactive visual analytics using Folium and Plotly Dash
- Perform predictive analysis using classification models
  - Logregression, KNN, SVM so on.

Data Collection

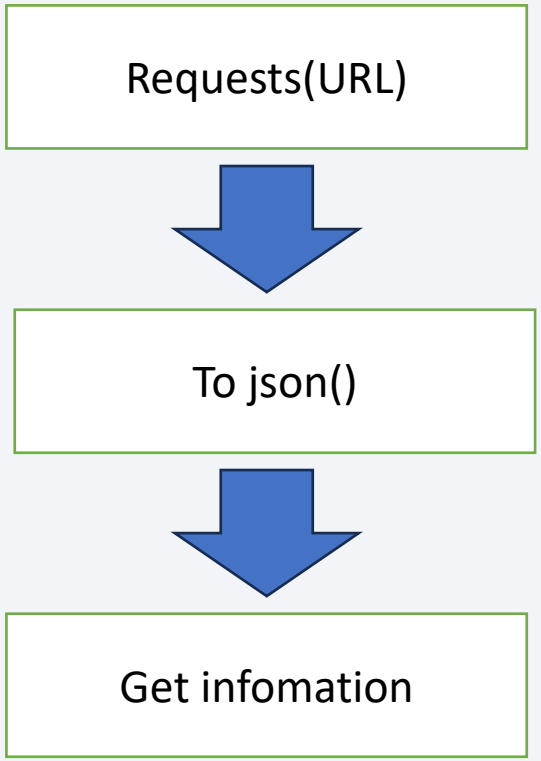
- Take a lot of information from Internet



Data Collection – SpaceX API

- Present your data collection with SpaceX REST calls using key phrases and flowcharts

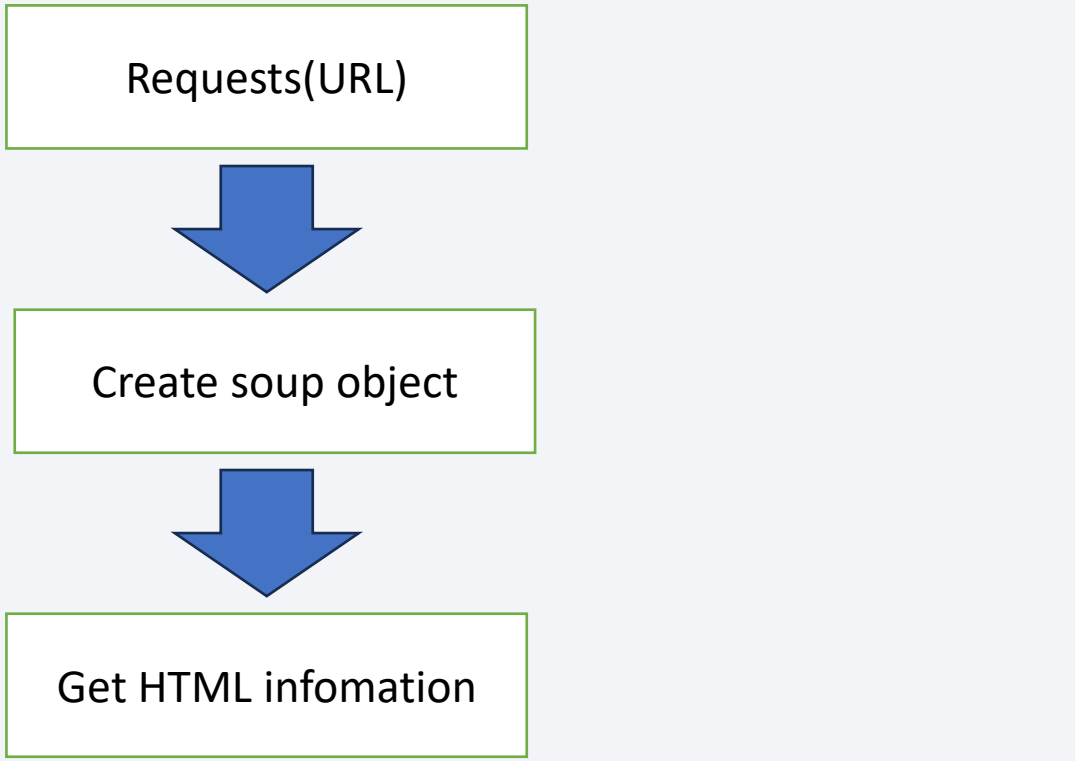
- Add the GitHub URL of the completed SpaceX API calls notebook<https://github.com/ryu2718/assingment/blob/main/jupyter-labs-spacex-data-collection-api-Copy1.ipynb>as an external reference and peer-review purpose



Data Collection - Scraping

- Present your web scraping process using key phrases and flowcharts

- Add the GitHub URL of the completed web scraping notebook,<https://github.com/ryu2718/assingment/blob/main/jupyter-labs-webscraping-Copy1.ipynb> as an external reference and peer-review purpose



Data Wrangling

- Success and failure of rocket indicated by binary value(1 or 0)
- Data cleaning
- Add the GitHub URL of your completed data wrangling related notebooks<https://github.com/ryu2718/assingment/blob/main/labs-jupyter-spacex-Data%20wrangling-Copy1.ipynb>, as an external reference and peer-review purpose

EDA with Data Visualization

- Using scatterplot, catplot, barplot, and lineplot
- To compare between two datas
- Add the GitHub URL of your completed EDA with data visualization notebook<https://github.com/ryu2718/assingment/blob/main/jupyter-labs-eda-dataviz-Copy1.ipynb>, as an external reference and peer-review purpose

EDA with SQL

- Using bullet point format, summarize the SQL queries you performed
- Add the GitHub URL of your completed EDA with SQL notebook<https://github.com/ryu2718/assingment/blob/main/jupyter-labs-eda-sql-coursera_sqllite.ipynb>, as an external reference and peer-review purpose

Build an Interactive Map with Folium

- Using a folium map, easy to understand the location on the launch site.
- Search distance between launch site and the closet coastline
- Add the GitHub URL of your completed interactive map with Folium map<https://github.com/ryu2718/assingment/blob/main/lab_jupyter_launch_site_location-Copy1.ipynb>, as an external reference and peer-review purpose

Build a Dashboard with Plotly Dash

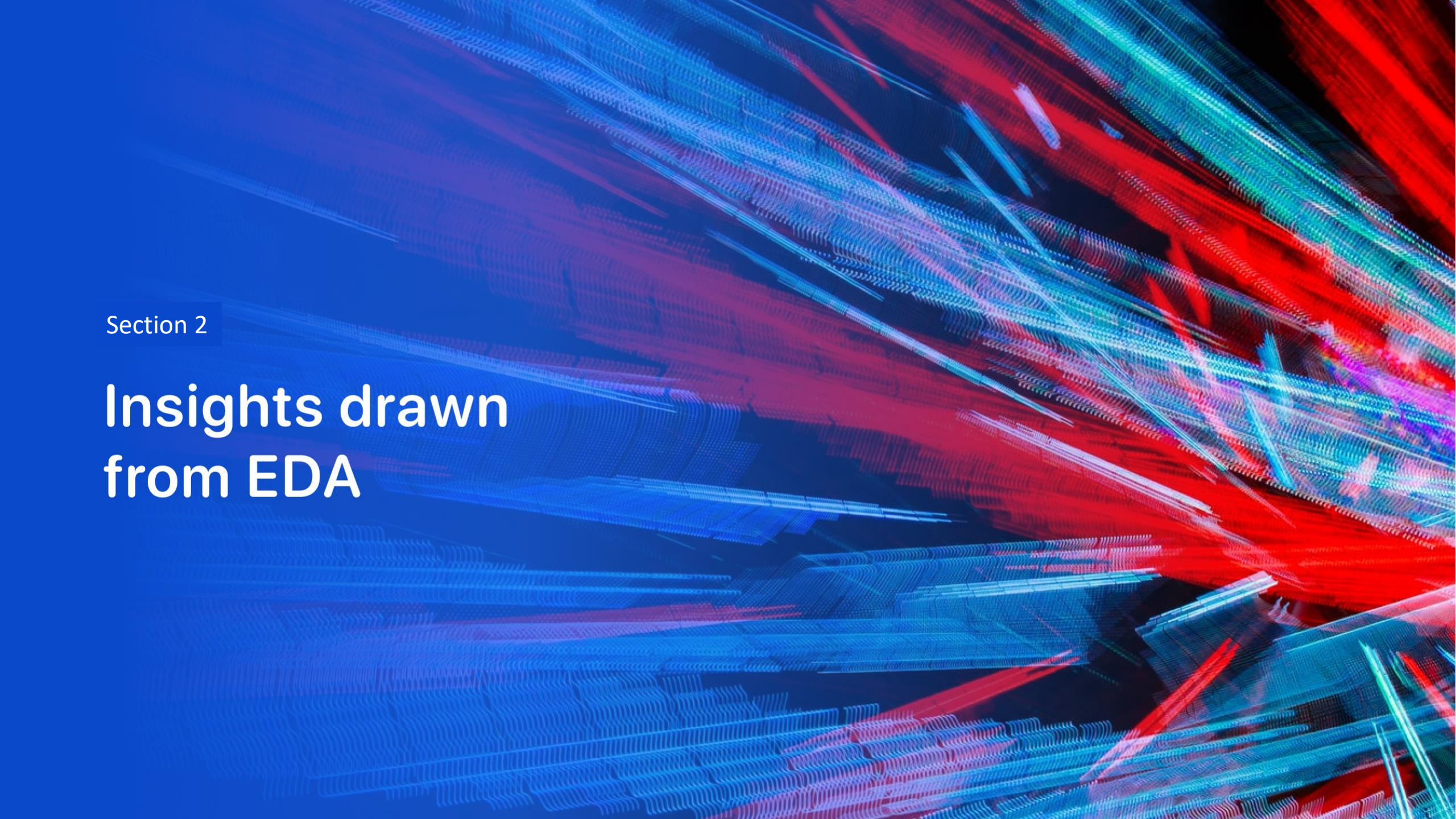
- To facilitate comparisons between graphs
- To make it easier to see the connections between graphs
- Add the GitHub URL of your completed Plotly Dash lab<https://github.com/ryu2718/assingment/blob/main/Python>, as an external reference and peer-review purpose

Predictive Analysis (Classification)

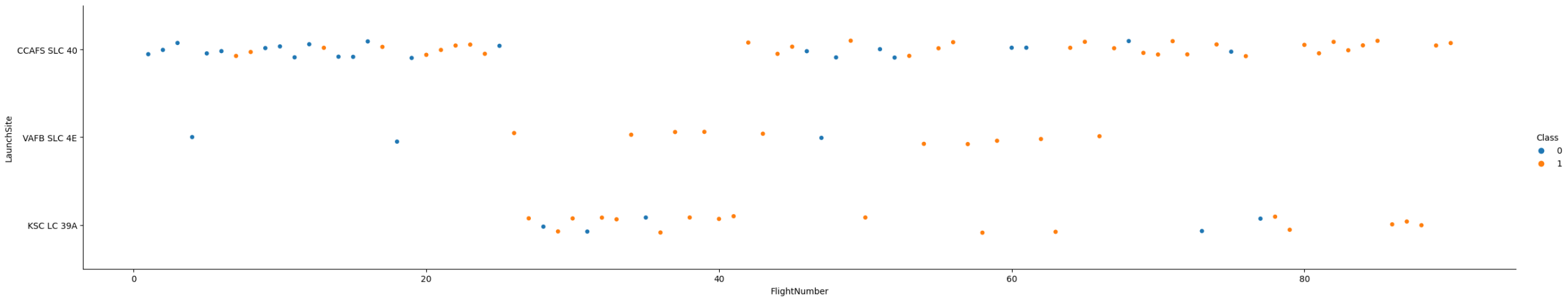
- Using Logregression, KNN, SVM, Decision tree
- Decision Tree is the most accurate.
- Add the GitHub URL of your completed predictive analysis lab<https://github.com/ryu2718/assingment/blob/main/SpaceX_Machine_Learning_Prediction_Part_5.jupyterlite-Copy1.ipynb>, as an external reference and peer-review purpose

Results

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

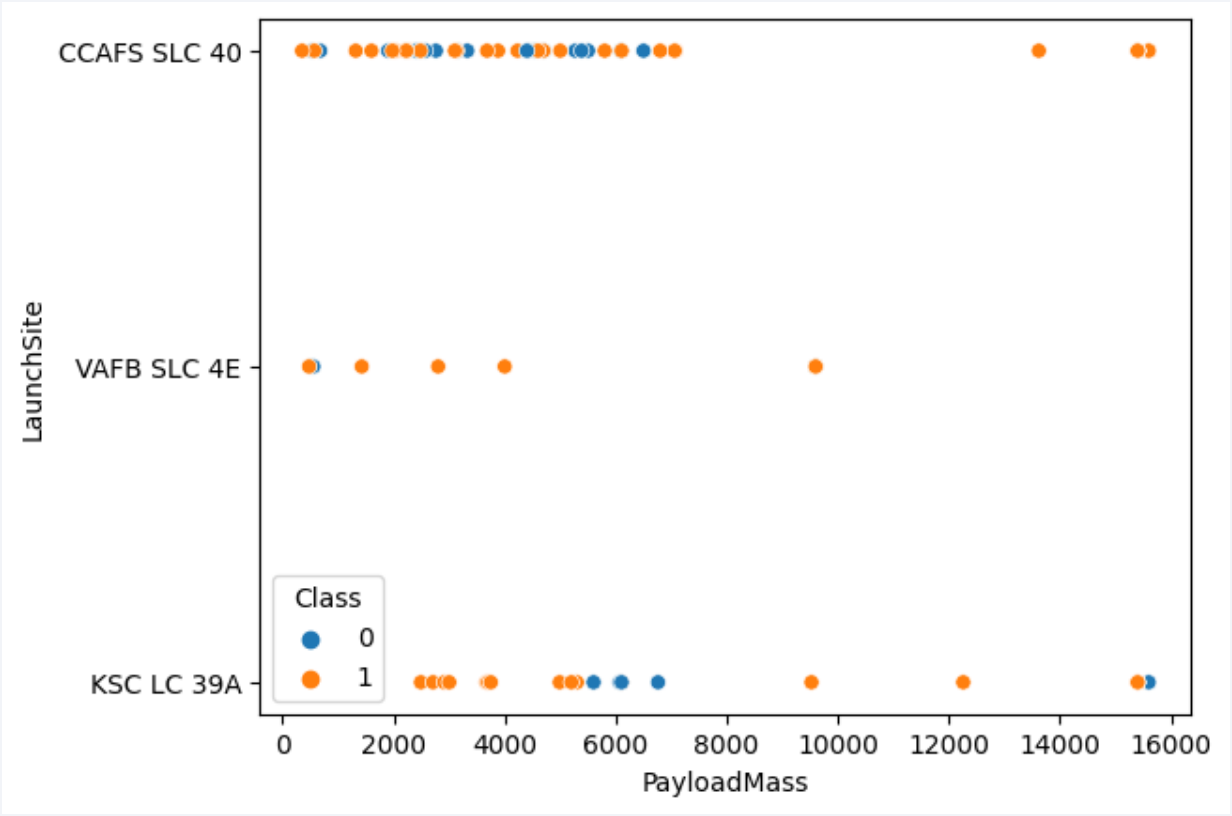

Section 2
Insights drawn from EDA

Flight Number vs. Launch Site



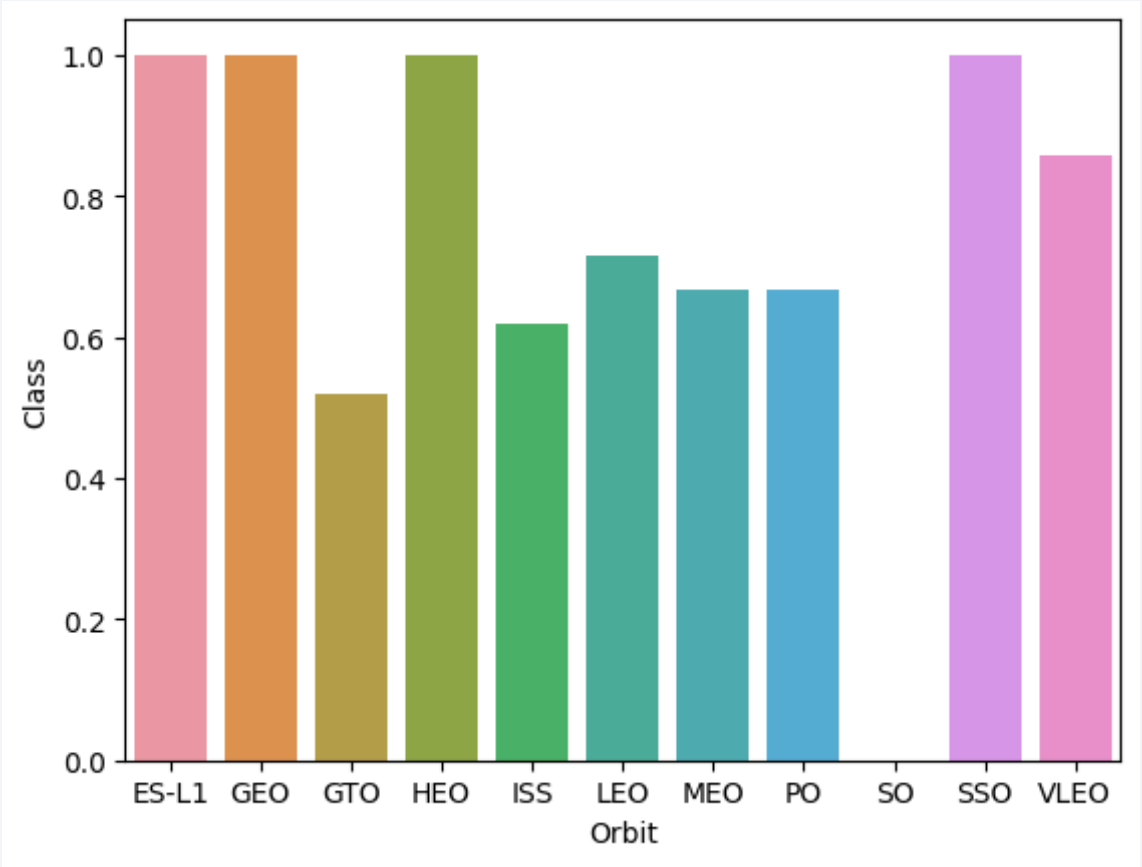
- The increasing Flight Number, the increasing success
- CCAFS SLC 40 has low Flight Number

Payload vs. Launch Site



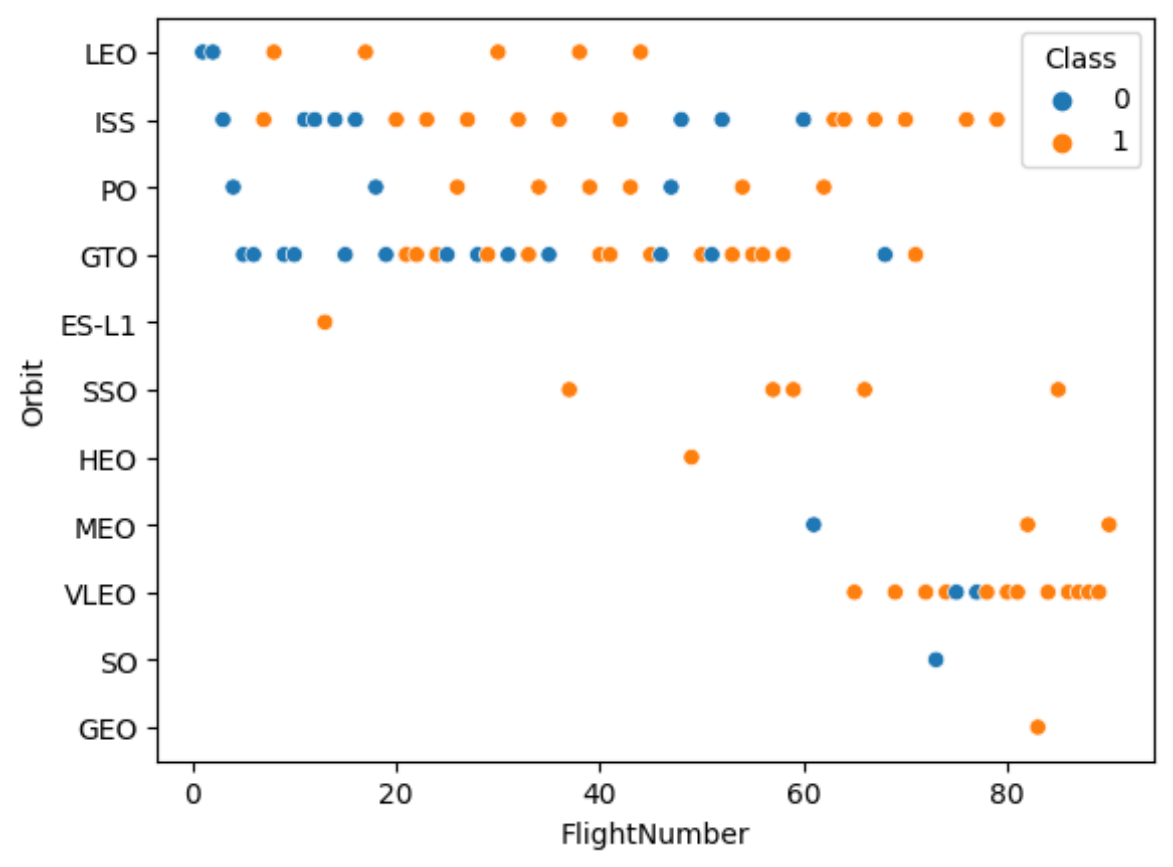
- VAFB SLC 4E has lightpayloads.

Success Rate vs. Orbit Type



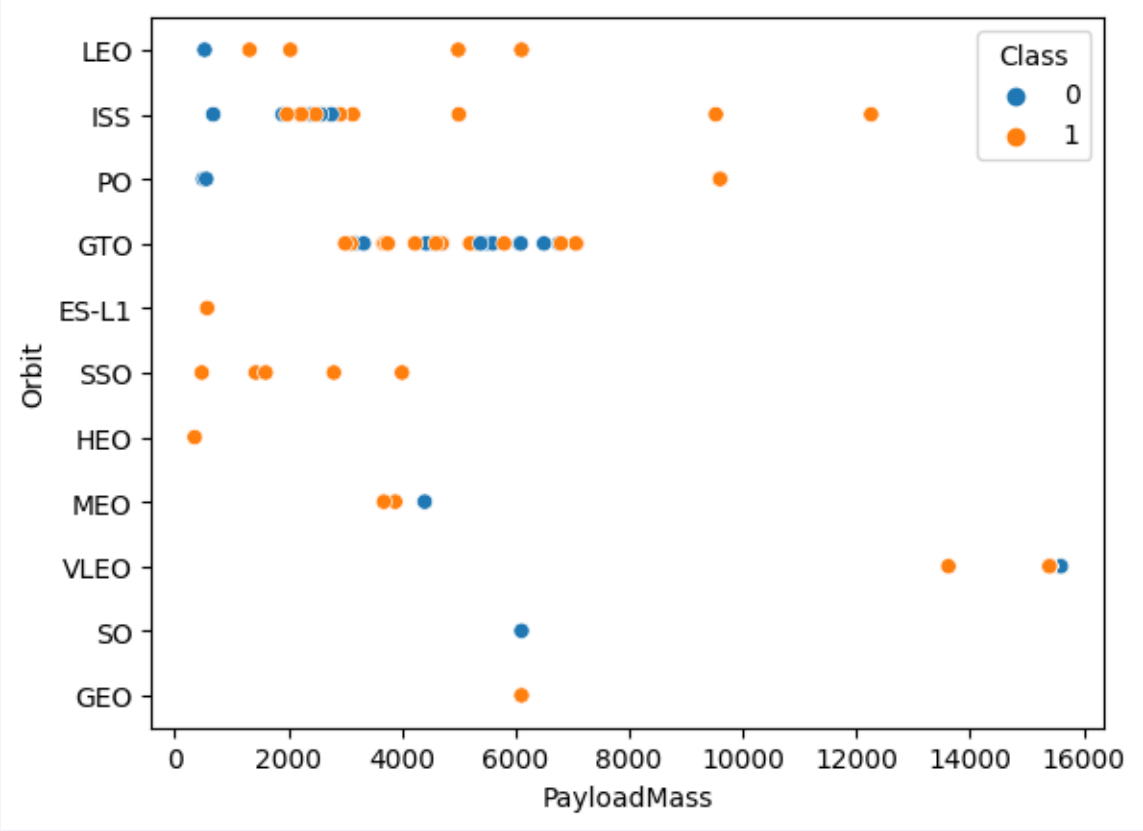
- ES-L1,GEO,HEO,SSO are all succese.

Flight Number vs. Orbit Type



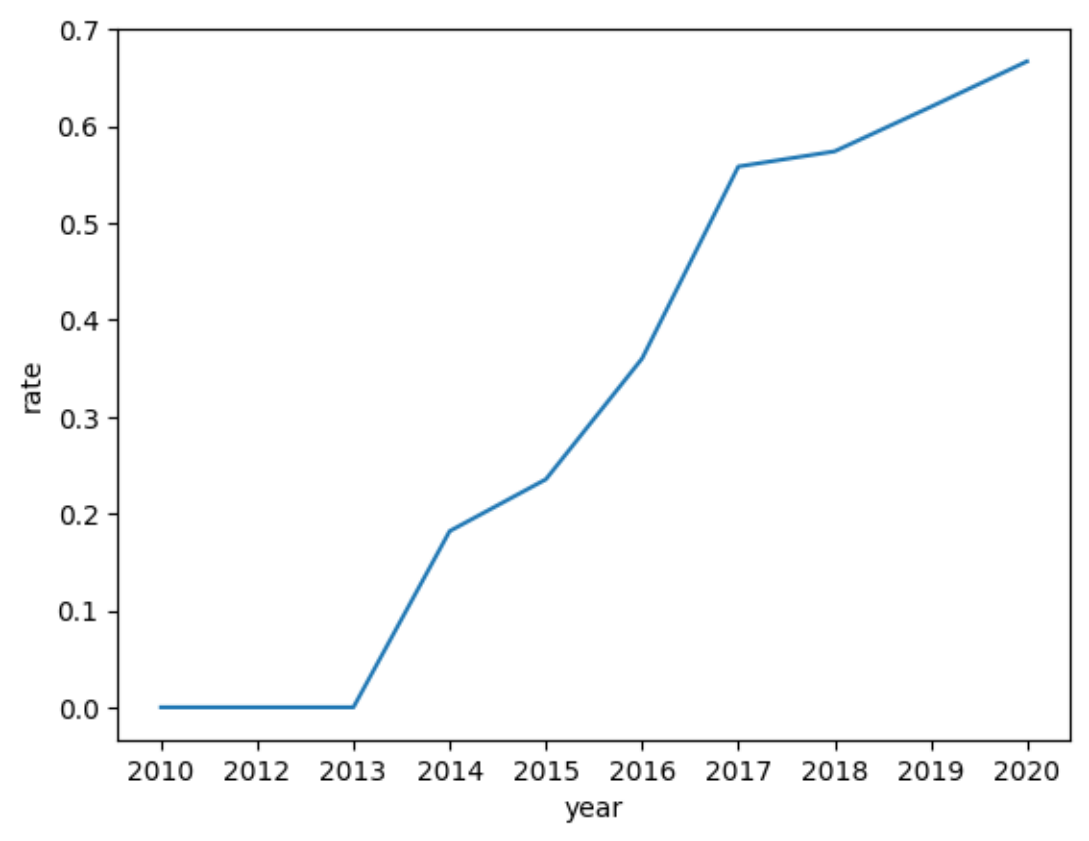
- LEO~ES-L1 have low FlightNumber.
- SSO~GEO have high FlightNumber.

Payload vs. Orbit Type



- Most orbits are light payload.

Launch Success Yearly Trend



- The succese rate has increased sharply since 2013.

All Launch Site Names

Launch site	Count
CCAFS SLC-40	34
CCAFS LC-40	26
KSC LC-39A	25
VAFB SLC-4E	16

Launch Site Names Begin with 'CCA'

	Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer	Mission_Outcome	Landing_Outcome
0	2010-04-06	18:45:00	F9 v1.0 B0003	CCAFS LC-40	Dragon Spacecraft Qualification Unit	0	LEO	SpaceX	Success	Failure (parachute)
1	2010-08-12	15:43:00	F9 v1.0 B0004	CCAFS LC-40	Dragon demo flight C1, two CubeSats, barrel of...	0	LEO (ISS)	NASA (COTS) NRO	Success	Failure (parachute)
2	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
3	2012-08-10	00:35:00	F9 v1.0 B0006	CCAFS LC-40	SpaceX CRS-1	500	LEO (ISS)	NASA (CRS)	Success	No attempt
4	2013-01-03	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 107010

	Customer	PAYLOAD_MASS__KG_
13	Iridium Communications GFZ , NASA	6460
16	NASA (CCD)	12055
17	NASA (CCDev)	12530
18	NASA (CCP)	12500
19	NASA (COTS)	525
20	NASA (COTS) NRO	0
21	NASA (CRS)	45596
22	NASA (CRS), Kacific 1	2617
23	NASA (CTS)	12050
24	NASA (LSP)	362
25	NASA (LSP) NOAA CNES	553
26	NASA / NOAA / ESA / EUMETSAT	1192
49	U.S. Air Force NASA NOAA	570

Average Payload Mass by F9 v1.1

- the average payload mass carried by booster version F9 v1.1 is 2928.4

	Booster_Version	PAYLOAD_MASS__KG_
86	F9 v1.1	2928.4
87	F9 v1.1 B1003	500.0
88	F9 v1.1 B1010	2216.0
89	F9 v1.1 B1011	4428.0
90	F9 v1.1 B1012	2395.0
91	F9 v1.1 B1013	570.0
92	F9 v1.1 B1014	4159.0
93	F9 v1.1 B1015	1898.0
94	F9 v1.1 B1016	4707.0
95	F9 v1.1 B1017	553.0
96	F9 v1.1 B1018	1952.0

First Successful Ground Landing Date

- the first successful landing outcome on ground pad is first row

Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer	Mission_Outcome	Landing_Outcome	
19	2015-12-22	01:29:00	F9 FT B1019	CCAFS LC-40	OG2 Mission 2 11 Orbcomm-OG2 satellites	2034	LEO	Orbcomm	Success	Success (ground pad)
26	2016-07-18	04:45:00	F9 FT B1025.1	CCAFS LC-40	SpaceX CRS-9	2257	LEO (ISS)	NASA (CRS)	Success	Success (ground pad)
29	2017-02-19	14:39:00	F9 FT B1031.1	KSC LC-39A	SpaceX CRS-10	2490	LEO (ISS)	NASA (CRS)	Success	Success (ground pad)
32	2017-01-05	11:15:00	F9 FT B1032.1	KSC LC-39A	NROL-76	5300	LEO	NRO	Success	Success (ground pad)
34	2017-03-06	21:07:00	F9 FT B1035.1	KSC LC-39A	SpaceX CRS-11	2708	LEO (ISS)	NASA (CRS)	Success	Success (ground pad)
38	2017-08-14	16:31:00	F9 B4 B1039.1	KSC LC-39A	SpaceX CRS-12	3310	LEO (ISS)	NASA (CRS)	Success	Success (ground pad)
40	2017-07-09	14:00:00	F9 B4 B1040.1	KSC LC-39A	Boeing X-37B OTV-5	4990	LEO	U.S. Air Force	Success	Success (ground pad)
44	2017-12-15	15:36:00	F9 FT B1035.2	CCAFS SLC-40	SpaceX CRS-13	2205	LEO (ISS)	NASA (CRS)	Success	Success (ground pad)
46	2018-08-01	01:00:00	F9 B4 B1043.1	CCAFS SLC-40	Zuma	5000	LEO	Northrop Grumman	Success (payload status unclear)	Success (ground pad)

Successful Drone Ship Landing with Payload between 4000 and 6000

Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer	Mission_Outcome	Landing_Outcome	
23	2016-06-05	05:21:00	F9 FT B1022	CCAFS LC-40	JCSAT-14	4696	GTO	SKY Perfect JSAT Group	Success	Success (drone ship)
27	2016-08-14	05:26:00	F9 FT B1026	CCAFS LC-40	JCSAT-16	4600	GTO	SKY Perfect JSAT Group	Success	Success (drone ship)
31	2017-03-30	22:27:00	F9 FT B1021.2	KSC LC-39A	SES-10	5300	GTO	SES	Success	Success (drone ship)
42	2017-11-10	22:53:00	F9 FT B1031.2	KSC LC-39A	SES-11 / EchoStar 105	5200	GTO	SES EchoStar	Success	Success (drone ship)

Total Number of Successful and Failure Mission Outcomes

- Success is 100
- Failure is 1
- But success distinguish “Success” and “Success(payload states unclear)”

Boosters Carried Maximum Payload

Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer	Mission_Outcome	Landing_Outcome	
74	2019-11-11	14:56:00	F9 B5 B1048.4	CCAFS SLC-40	Starlink 1 v1.0, SpaceX CRS-19	15600	LEO	SpaceX	Success	Success
77	2020-07-01	02:33:00	F9 B5 B1049.4	CCAFS SLC-40	Starlink 2 v1.0, Crew Dragon in-flight abort t...	15600	LEO	SpaceX	Success	Success
79	2020-01-29	14:07:00	F9 B5 B1051.3	CCAFS SLC-40	Starlink 3 v1.0, Starlink 4 v1.0	15600	LEO	SpaceX	Success	Success
80	2020-02-17	15:05:00	F9 B5 B1056.4	CCAFS SLC-40	Starlink 4 v1.0, SpaceX CRS-20	15600	LEO	SpaceX	Success	Failure
82	2020-03-18	12:16:00	F9 B5 B1048.5	KSC LC-39A	Starlink 5 v1.0, Starlink 6 v1.0	15600	LEO	SpaceX	Success	Failure
83	2020-04-22	19:30:00	F9 B5 B1051.4	KSC LC-39A	Starlink 6 v1.0, Crew Dragon Demo-2	15600	LEO	SpaceX	Success	Success
85	2020-04-06	01:25:00	F9 B5 B1049.5	CCAFS SLC-40	Starlink 7 v1.0, Starlink 8 v1.0	15600	LEO	SpaceX, Planet Labs	Success	Success
92	2020-03-09	12:46:14	F9 B5 B1060.2	KSC LC-39A	Starlink 11 v1.0, Starlink 12 v1.0	15600	LEO	SpaceX	Success	Success
93	2020-06-10	11:29:34	F9 B5 B1058.3	KSC LC-39A	Starlink 12 v1.0, Starlink 13 v1.0	15600	LEO	SpaceX	Success	Success
94	2020-10-18	12:25:57	F9 B5 B1051.6	KSC LC-39A	Starlink 13 v1.0, Starlink 14 v1.0	15600	LEO	SpaceX	Success	Success
95	2020-10-24	15:31:34	F9 B5 B1060.3	CCAFS SLC-40	Starlink 14 v1.0, GPS III-04	15600	LEO	SpaceX	Success	Success
99	2020-11-25	02:13:00	F9 B5 B1049.7	CCAFS SLC-40	Starlink 15 v1.0, SpaceX CRS-21	15600	LEO	SpaceX	Success	Success

2015 Launch Records

Date	Time (UTC)	Booster_Version	Launch_Site	Payload	PAYLOAD_MASS__KG_	Orbit	Customer	Mission_Outcome	Landing_Outcome	
13	2015-10-01	09:47:00	F9 v1.1 B1012	CCAFS LC-40	SpaceX CRS-5	2395	LEO (ISS)	NASA (CRS)	Success	Failure (drone ship)
16	2015-04-14	20:10:00	F9 v1.1 B1015	CCAFS LC-40	SpaceX CRS-6	1898	LEO (ISS)	NASA (CRS)	Success	Failure (drone ship)

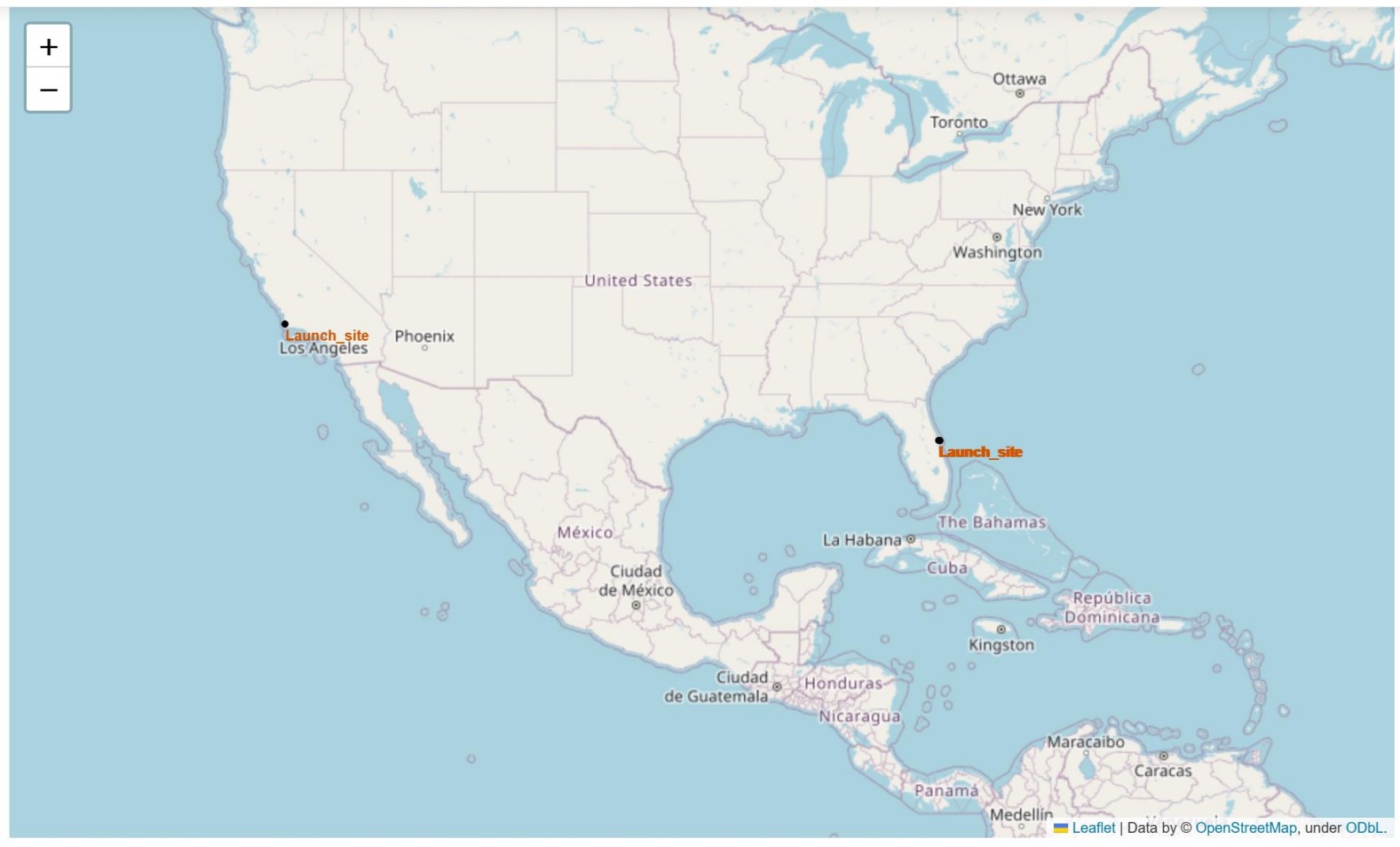
Rank Landing Outcomes Between 2010-06-04 and 2017-03-20

Landing_Outcome	count
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

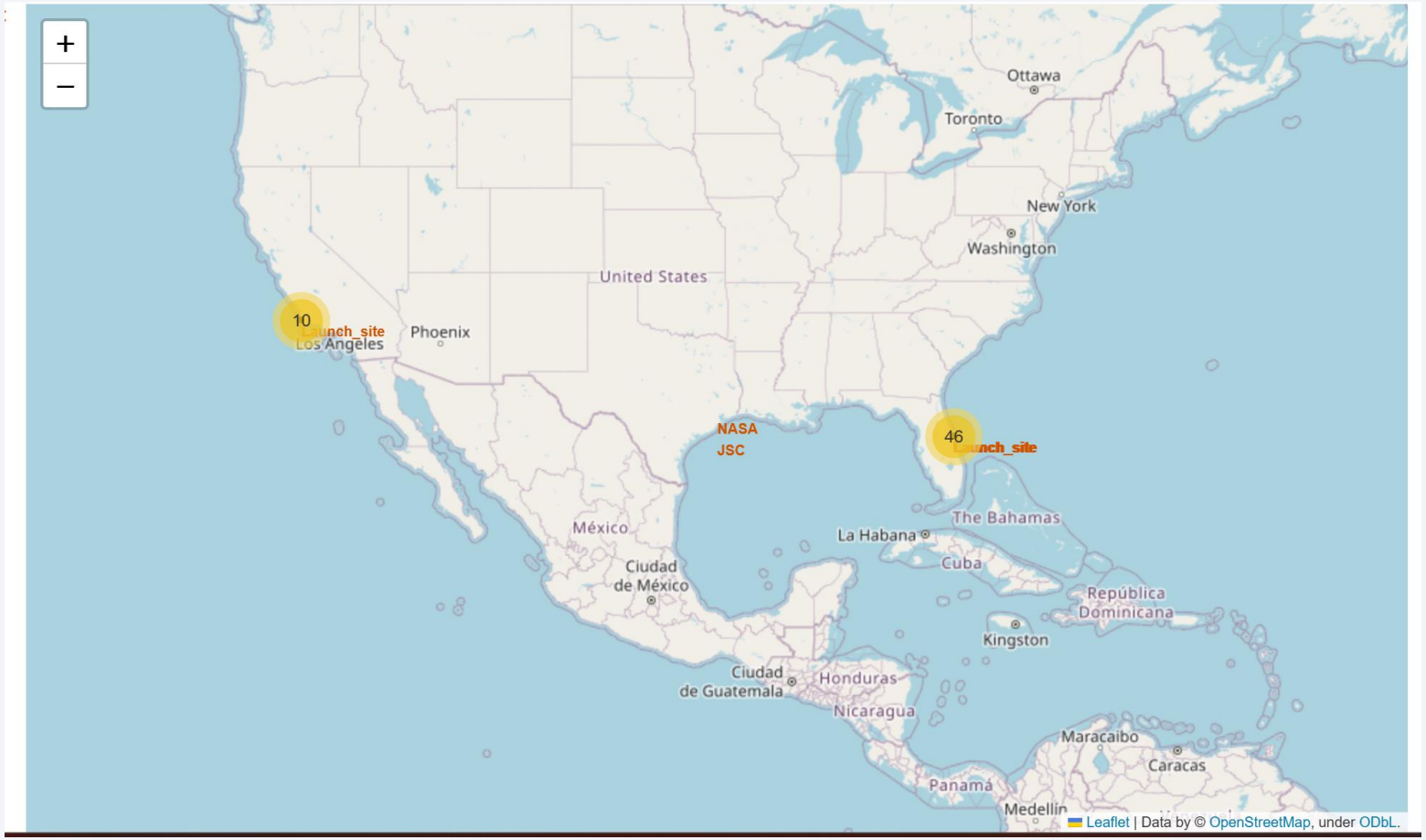
Section 3

Launch Sites Proximities Analysis

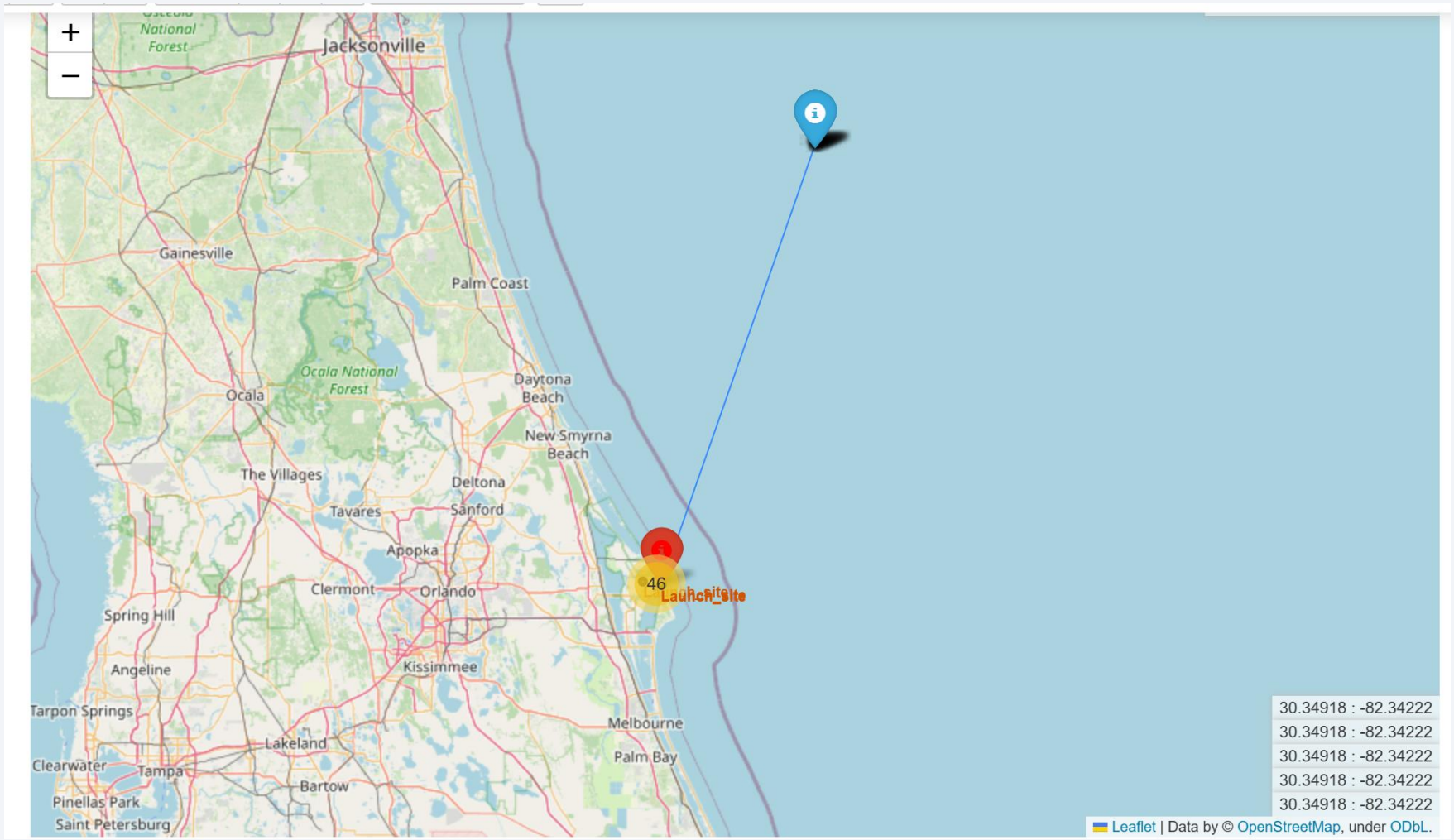
<locations about Launch_site>

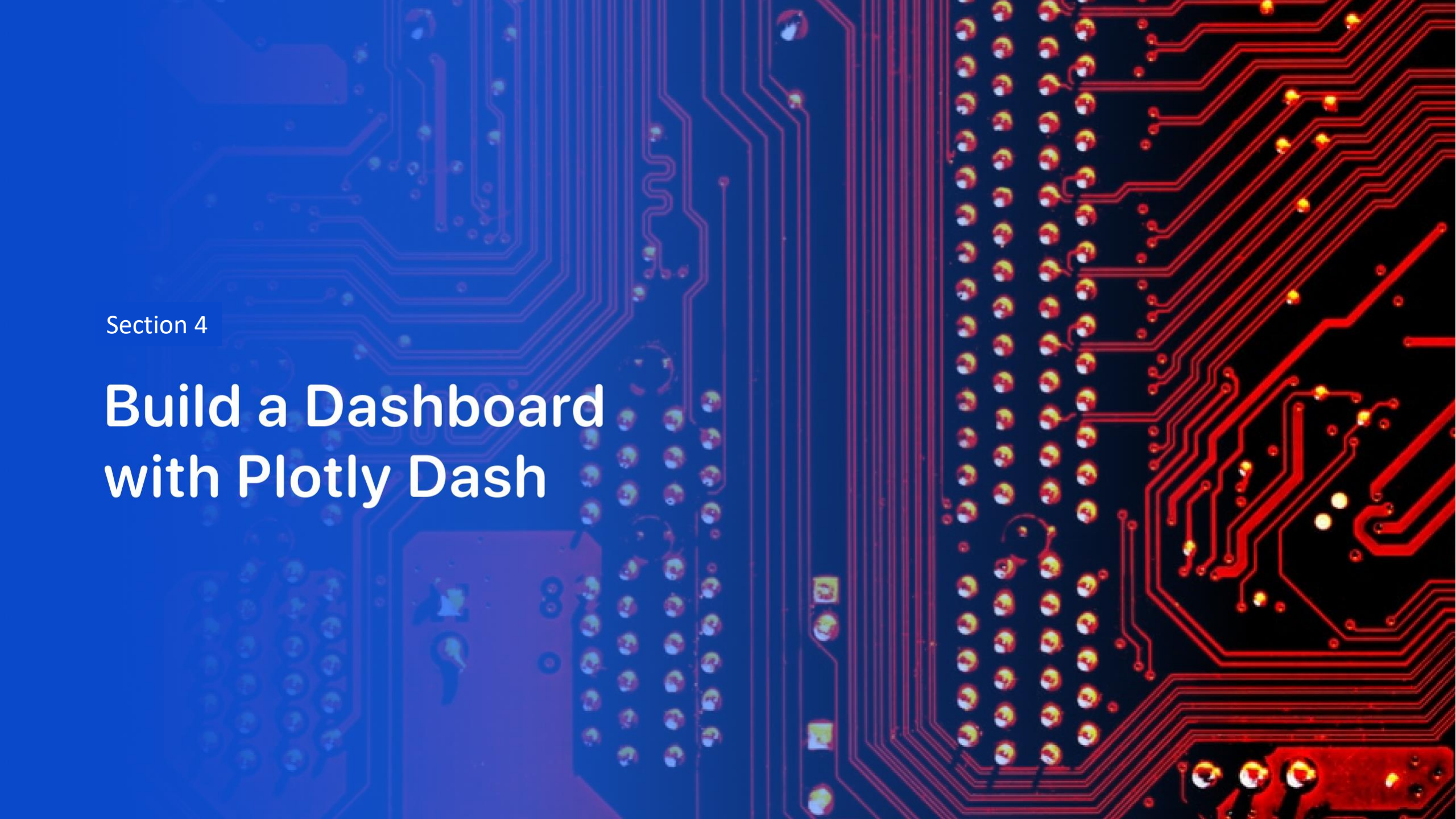


< Mark the success/failed launches for each site on the map>

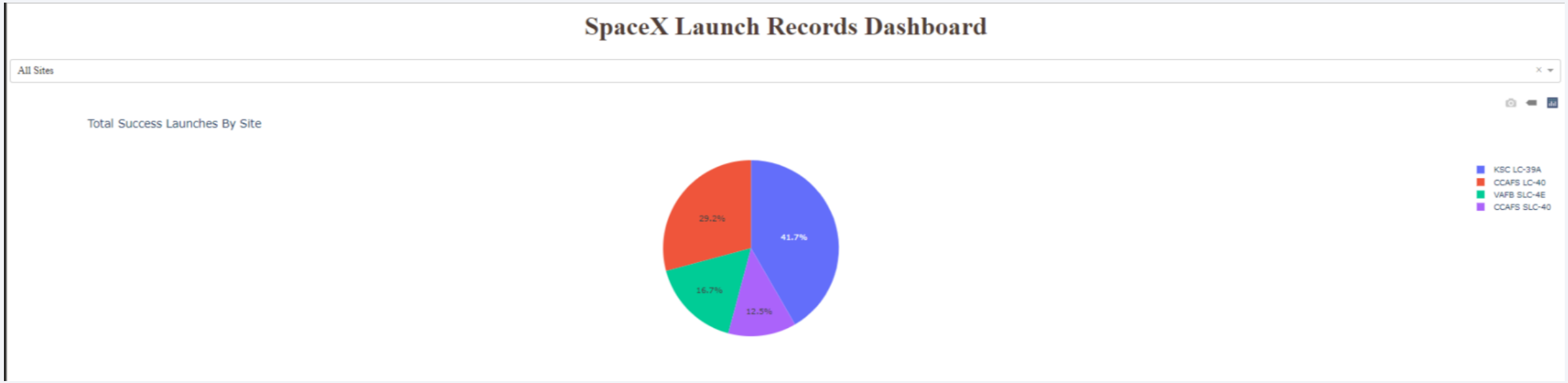


< the distances between a launch site to its proximities>

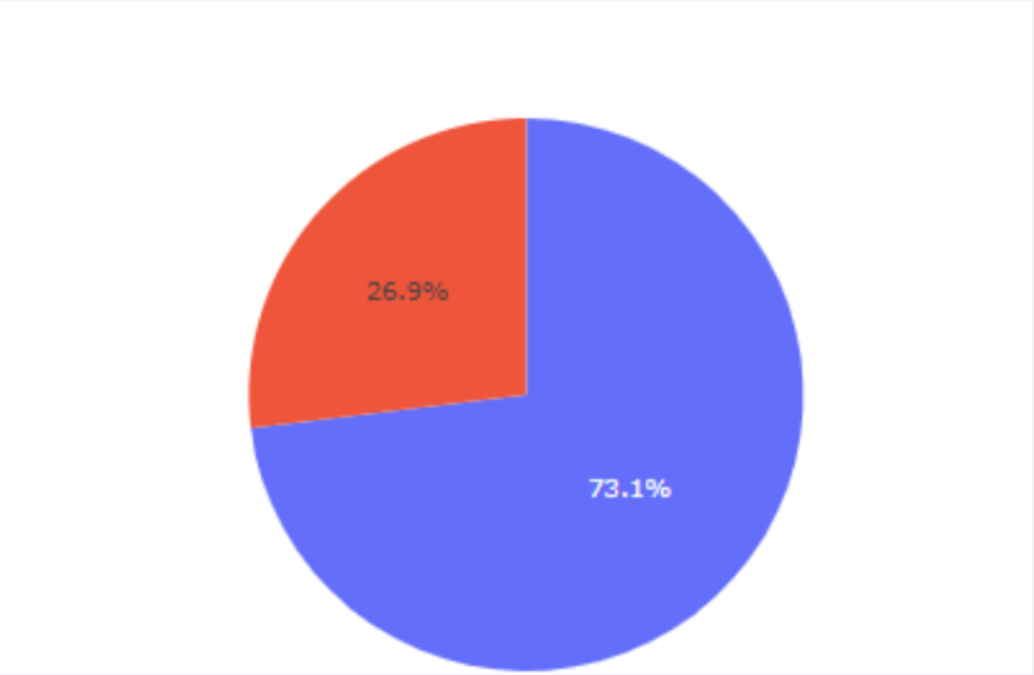



Section 4
Build a Dashboard with Plotly Dash

<Total success launches by site using pie chart>



<pie chart about CCAFS SLC-40>



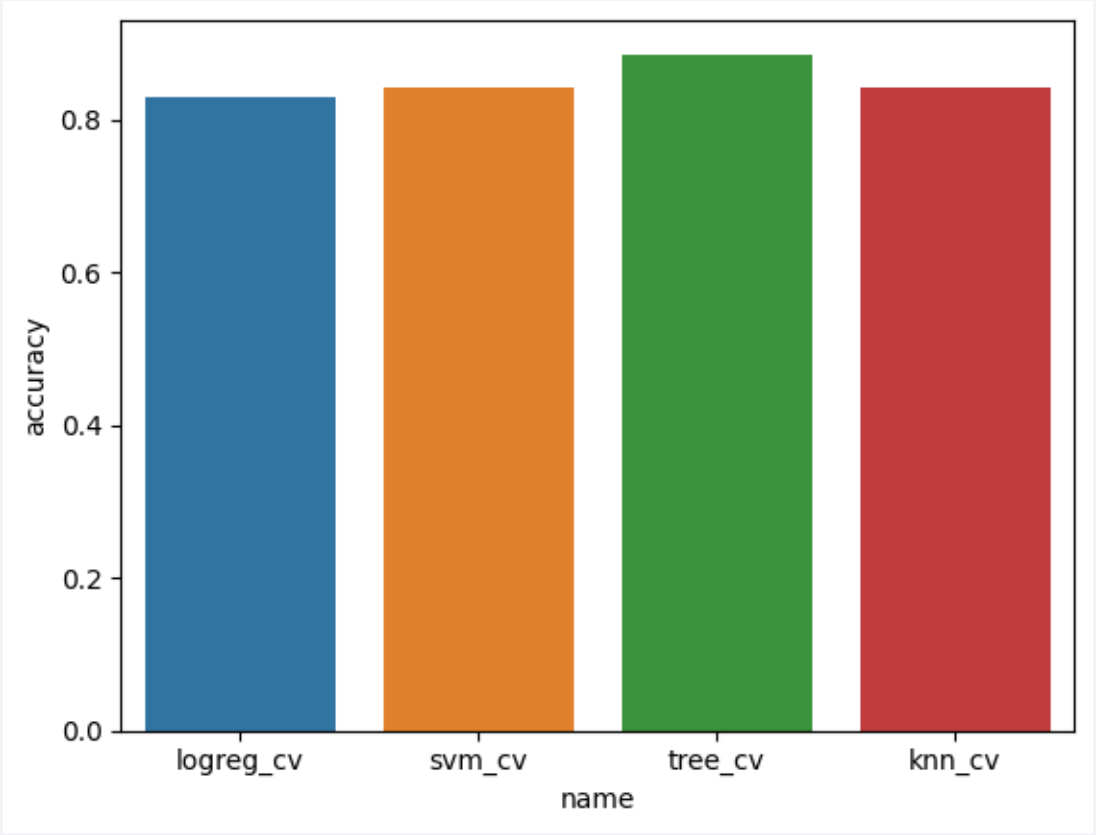
<Dashboard Screenshot 3>

- Replace <Dashboard screenshot 3> title with an appropriate title
- Show screenshots of Payload vs. Launch Outcome scatter plot for all sites, with different payload selected in the range slider
- Explain the important elements and findings on the screenshot, such as which payload range or booster version have the largest success rate, etc.

Section 5

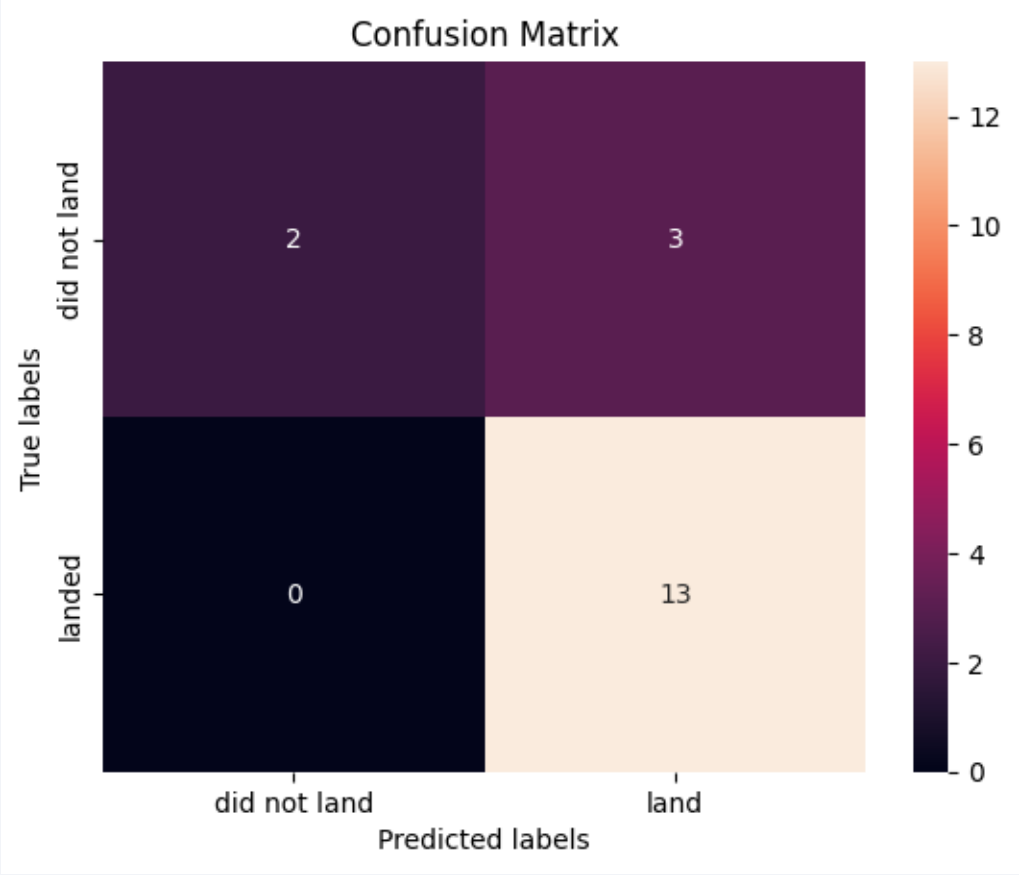
Predictive Analysis (Classification)

Classification Accuracy



- Tree model has the highest classification accuracy

Confusion Matrix



Conclusions

- The launches with the highest success rate are CCAFS SLC-40
- The best ranges are ES-L1,GEO,HEO,SSO.
- Launch technology is increasing every year.

Appendix

- Include any relevant assets like Python code snippets, SQL queries, charts, Notebook outputs, or data sets that you may have created during this project

Thank you!