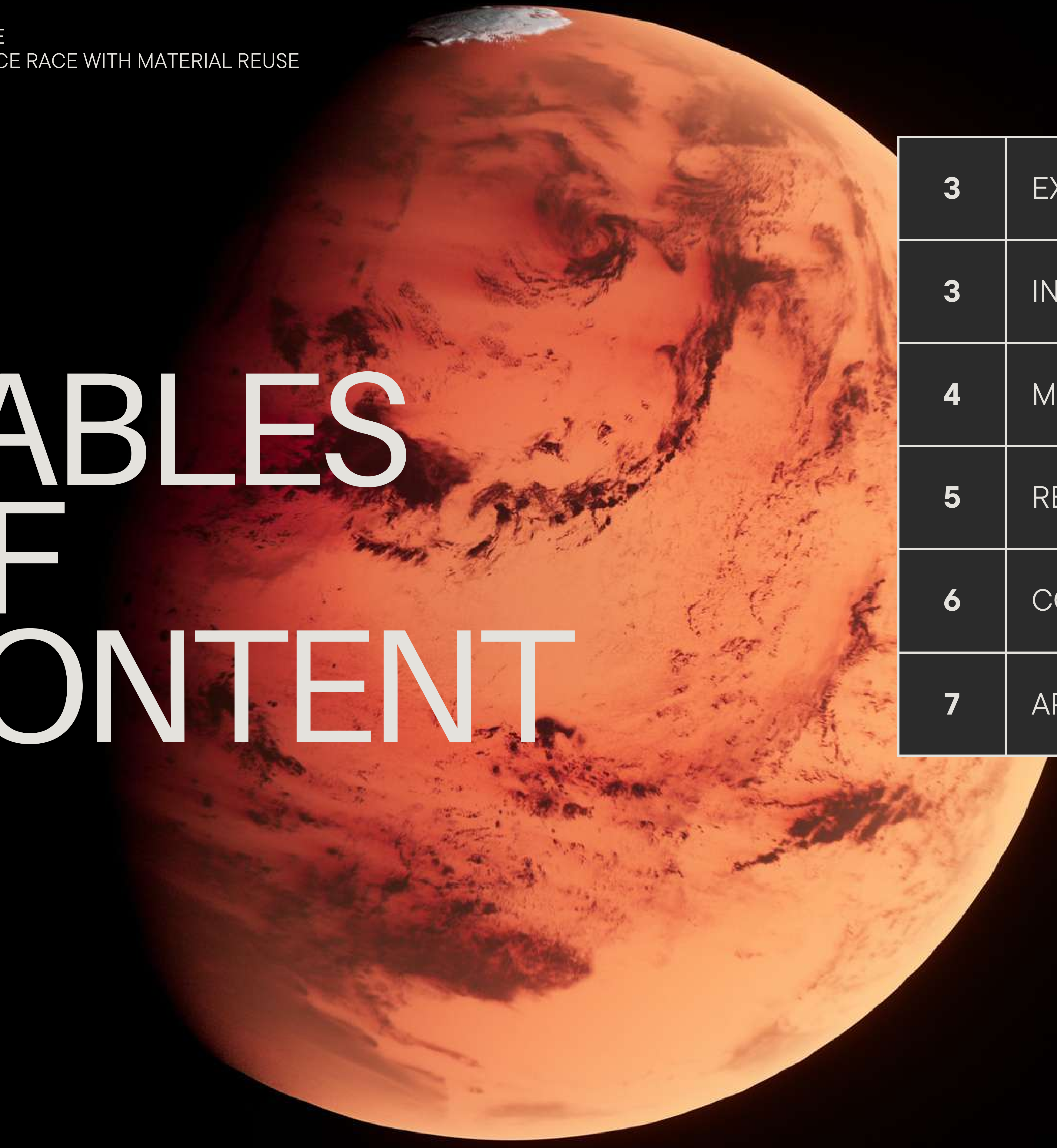
DATA SCIENCE

DATA & SPACE EXPLORATION

PRESENTED BY: F.G.J
11/09/2023



TABLES OF CONTENT

EXECUTIVE SUMMARY

COLLECT	SPACEX DATA USING REST API AND WEB SCRAPING
WRANGLE	PREPROCESSING - CLEANING, TRANSFORMING AND ORGANIZING DATA FOR FURTHER ANALYSIS
EXPLORE	EXPLORATORY DATA ANALYSIS USING SQL, VISUALIZATION AND STATISCAL PROCESSES
ANALYZE	USING STATISTICAL METHODS ON IDENTIFIED KEY VARIABLES AS WELL AS SQL AND VISUALIZATIONS
VISUALIZE	INSIGHTS FROM GEO DATA, PLOTTING, CHARTING AND CREATION OF A DYNAMIC DASHBOARD
MODELING	LEVERAGING MACHINE LEARNING (SVM, KNN, DT AND LR) TO PREDICT BOOSTERS LANDING

RESULTS

EXPLORATORY DATA ANALYSIS

- OVERALL SUCCESS RATE OF LAUNCHES IMPROVES FROM 2013
- ORBITS ES-L1, GEO, HEO, SSO HAVE 100% SUCCESS RATE
- LAUNCH SITE CCAFD SLC 40 MOST USED LAUNCH SITE
- LAUNCH SITE KSC LC-39A HAS THE HIGHEST SUCESS RATE

ANALYTICS & VISUALIZATION

- MOST LAUNCH SITES ARE LOCATED NEAR THE COAST
- ALL SITE LOCATED ABOVE CANCER TROPIC LINE
- ALL SITE LOCATED FAR FROM PUBLIC INFRASTRUCTURES

PREDICTIVE ANALYTICS

- SVM ACCURACY PERFORMED 1.75% BETTER THAN OTHER MODELS
- ALL MODELS HAVE THE SAME JACCARD SCORE
- ALL MODELS HAVE THE SAME F1_SCORE

EXECUTIVE SUMMARY

KEY FINDINGS

- SPACEX SAVES APPROX. $37M PER SUCCESSFUL BOOSTER LANDING.
- SUCCESS COMES FROM FAILURE
- PAYLOAD MASS MATTERS WITH HIGHLY CORRELATED RELATIONSHIP BETWEEN SUCCESS RATE AND PAYLOAD MASS
- ORBIT MISSION TYPE IS ALSO SHARING A POSITIVE RELATIONSHIP WITH POSITIVE SUCCESS RATE

OBJECTIVE & SCOPE

- UNDERSTAND SPACEX SUCCESS
- COLLECT, ANALYZE SPACEX AVAILABLE DATA TO HELP COMPANY MAKE DATA-DRIVE DECISIONS
- COMPANY NEEDS DYNAMIC DASHBOARD TO HAVE BETTER UNDERSTANDING OF THE DATA
- UNDERSTAND ON WHAT TO FOCUS
- SUBMIT CONCLUSIONS WITH KEY FINDINGS AND INSIGHTS

METHODOLOGY

- COLLECT, WRANGLE, EXPLORE, ANALYZE, VISUALIZE AND MODEL BUILDING.

RESULTS

EXPLORATORY DATA ANALYSIS

- OVERALL SUCCESS RATE OF LAUNCHES IMPROVES FROM 2013
- ORBITS ES-L1, GEO, HEO, SSO HAVE 100% SUCCESS RATE
- LAUNCH SITE CCAFD SLC 40 MOST USED LAUNCH SITE
- LAUNCH SITE KSC LC-39A HAS THE HIGHEST SUCESS RATE

ANALYTICS & VISUALIZATION

- MOST LAUNCH SITES ARE LOCATED NEAR THE COAST
- ALL SITE LOCATED ABOVE CANCER TROPIC LINE
- ALL SITE LOCATED FAR FROM PUBLIC INFRASTRUCTURES

PREDICTIVE ANALYTICS

- SVM ACCURACY PERFORMED 1.75% BETTER THAN OTHER MODELS
- ALL MODELS HAVE THE SAME JACCARD SCORE
- ALL MODELS HAVE THE SAME F1_SCORE

INTRODUCTION

BACKGROUND

SpaceX is a space transportation company founded by Elon Musk in 2002. SpaceX's hallmark achievements is its ability to offer more cost-effective rocket launches compared to other providers in the industry, whom charge upwards of 165 million dollars per launch. SpaceX advertises Falcon 9 rocket launches on its website with a cost of 62 million dollars. A significant part of these cost savings stems from SpaceX's innovative practice of reusing the first stage of its rockets. Consequently, the successful landing of this first stage becomes a pivotal factor in determining the overall cost of a launch. This crucial information holds significance not only for SpaceX but also for potential alternate companies looking to compete with SpaceX for rocket launch contracts.

CHALLENGES

- Investigating the impact of payload mass, launch site, number of flights, and orbits on the success of first-stage landings.
- Analyzing the trend in successful landing rates over a period of time.
- Developing the most accurate predictive model for determining the success of a first-stage landing using binary classification

METHODOLOGY

METHODOLOGY

STEPS

COLLECT

SPACEX DATA USING REST API AND WEB SCRAPING

WRANGLE

PREPROCESSING - CLEANING, TRANSFORMING AND ORGANIZING DATA FOR FURTHER ANALYSIS

EXPLORE

EXPLORATORY DATA ANALYSIS USING SQL, VISUALIZATION AND STATISCAL PROCESSES

VISUALIZE

INSIGHTS FROM GEO DATA, PLOTTING, CHARTING AND CREATION OF A DYNAMIC DASHBOARD

PREDICTIVE ANALYSIS: MODELING

MACHINE LEARNING TO PREDICT BOOSTERS LANDING

DATA COLLECTION SPACEX API

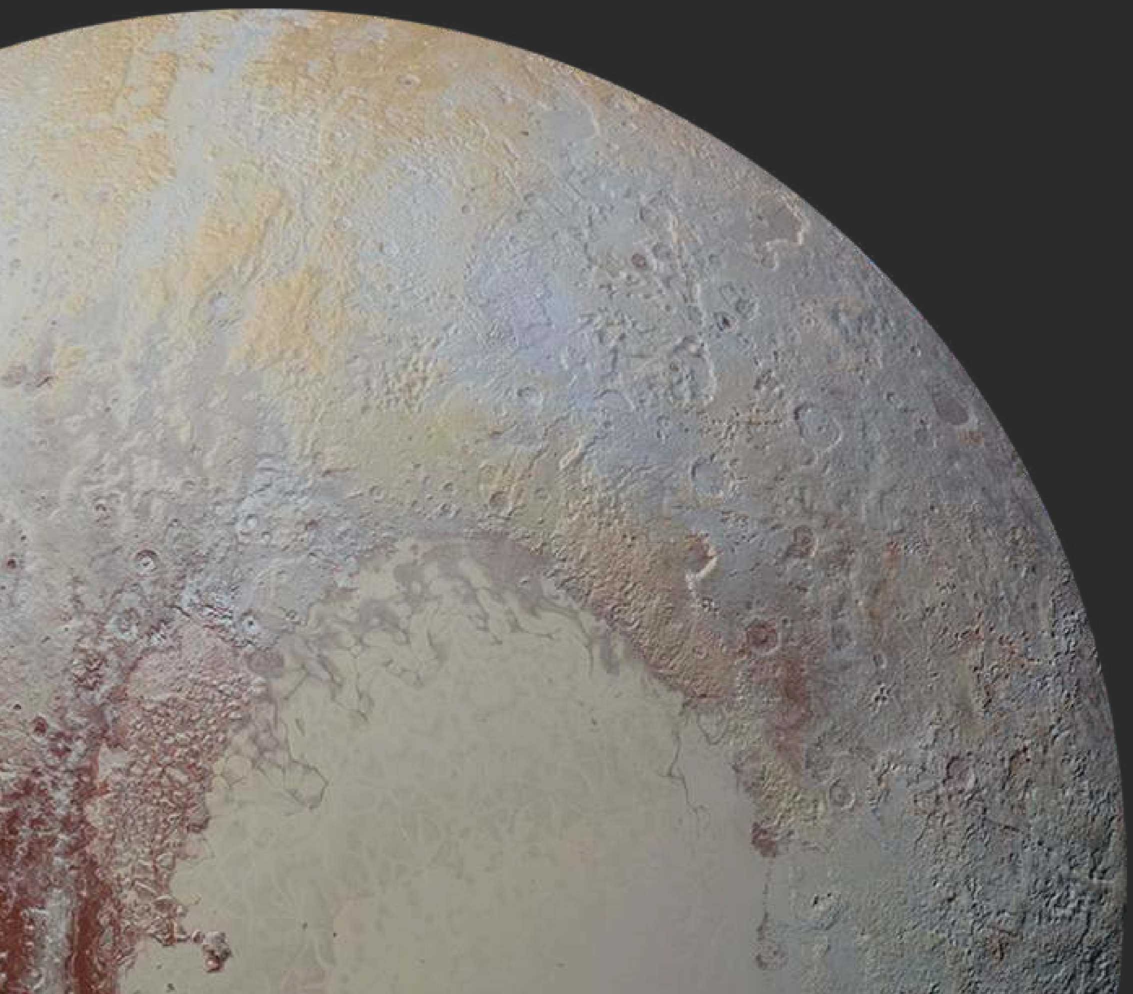
- **Requesting data through SpaceX API (get request)**
- **Decoding response** using .json() and converting it into a pd.DataFrame using .json_normalize()
- **Creating dataframes** from the requested datasets
- **Filtering dataframe** for information within the scope of the analysis
- **Cleaning dataframes** replacing missing values and checking for NaN and incorrect data points
- **Save and Export csv file**
- **GitHub URL:**
  - **github.com/fgjspaceman/SpaceX-Capstone-Data-Science-IBM/blob/main/01-SpaceX-data-collection-api.ipynb**

DATA COLLECTION
WEB SCRAPING

- Observation of the targeted website : Wikipedia “Falcon 9 Launch Record”
- Setting up scrapper to pull table data from Wikipedia
- Setting up BeautifulSoup for HTML scrapping and parsing
- Extracting data from targeted HTLM header
- Collecting and creating dataframes
- Verifying the data collected
- Save and Export csv file
- GitHub URL:
  - github.com/fgjspaceman/SpaceX-Capstone-Data-Science-IBM/blob/main/02-SpaceX-web-scraping.ipynb

DATA WRANGLING

- **Exploration Data Analysis performed**
- **Relevant features were defined, format and normalized**
- **Deeper exploration into launch sites and orbit types**
- **GitHub URL:**
  - **github.com/fgjspaceman/SpaceX-Capstone-Data-Science-IBM/blob/main/03-SpaceX-data-wrangling.ipynb**



Data Cleaning	Data Inspection	Data Transformation	Data Visualization
Data Integration	Data Encoding	Data Validation	Data Exportation

DATA EXPLORATION VISUALIZATION

CHARTING

VARIABLES OF INTEREST

Flight Number	Payload
Launch Site	Orbit Type
Boosters Outcome	Launch Outcome

Analysis Method

- **Charting / Plotting**
  - Scatter, Bar , Line, Pie charts and others were used to highlights the relationship between our variables and potential correlations.
- **Features Engineering**
  - Use of different techniques such as One-Hot Encoding and Normalization.

DATA EXPLORATION SQL

QUERIES OBJECTIVES

- F9 V1.1 AVERAGE PAYLOAD MASS
- TOTAL SUCCESSFUL AND FAILED MISSIONS
- AVG. F9 V1.1 PAYLOAD MASS
- 'CCA' STARTING LAUNCH SITES
- FIRST SUCCESSFUL GROUND PAD LANDING
- SUCCESSFUL DRONE SHIP LANDINGS (PL WITHIN 4,000-6,000)
- MAX PAYLOAD BOOSTER VERSIONS
- GITHUB URL: GITHUB.COM/FGJSPACEMAN/SPACEX-CAPSTONE-DATA-SCIENCE-IBM/BLOB/MAIN/04-SPACEX-EDA-SQL.IPYNB

DATA VISUALIZATION

DASHBOARDS PLOTLY DASH

DYNAMIC DASHBOARD

- Interactive dashboard using real data
- Python built and fully customizable
- Easily connected to live data

STAKEHOLDER FRIENDLY

- UI allows users to interact with the dashboard and charts
- Users can customized the chart for prefered-data display and visualization

INSIGHTS MACHINE

- Lightning fast to play with data and find insights
- Scatter, Pie charts, sliders and drop down menu fully working

GEO MAP FOLIUM

CIRCLE MARKERS: LAUNCH SITES

- Blue circle at NASA Space Center coordinates
- Red circles at SpaceX launch sites coordinates

COLOR MARKERS LAUNCH OUTCOMES

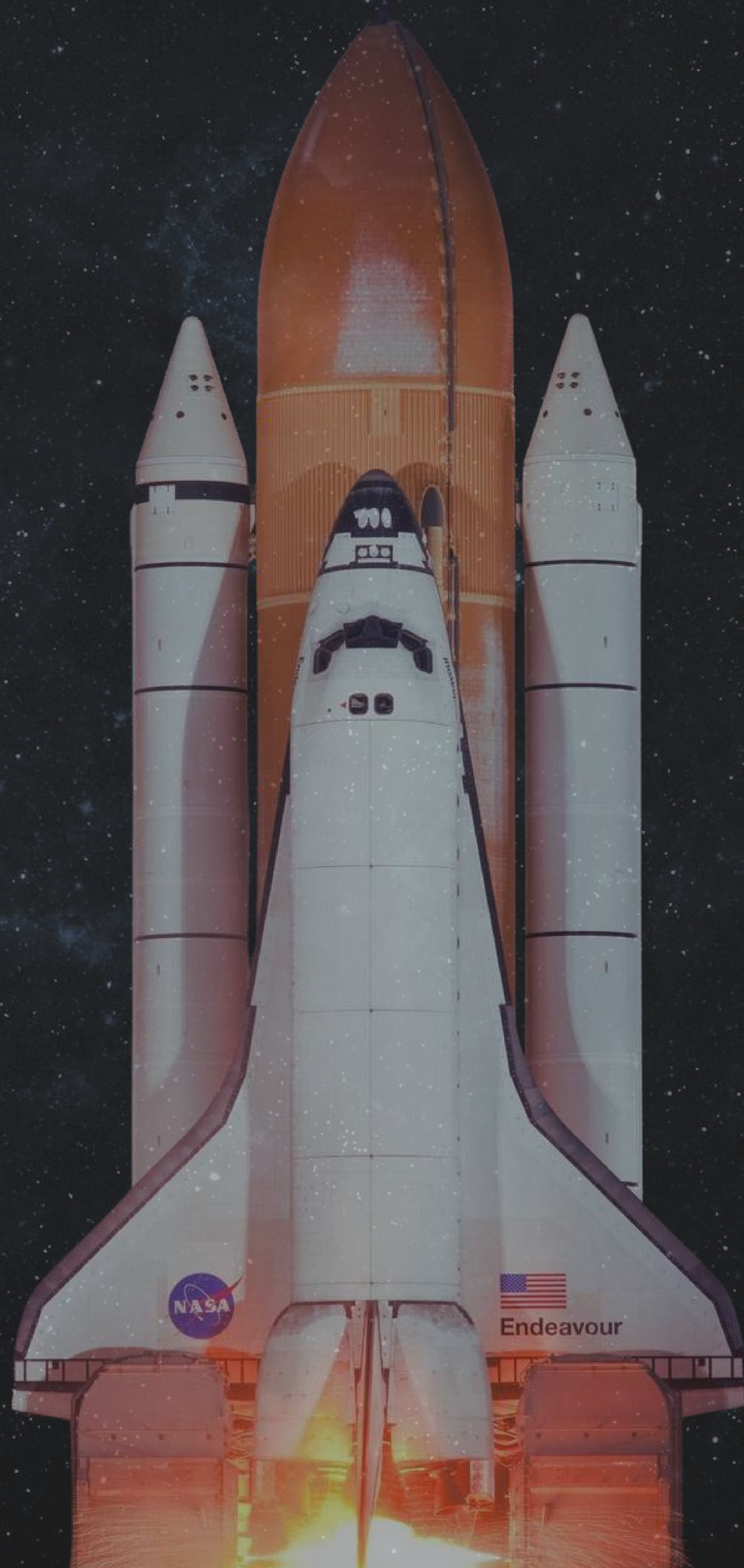
- Custom markers within a cluster of markers for specific conditions
- Green markers for successful launches
- Red markers for unsuccesful launches

COLOR PROXIMITY MARKERS

- Mouseover calculating distance between points of interests
- Line markers displaying distance between X & Y

PREDICTIVE ANALYTICS

- CONVERT THE "CLASS" COLUMN IN THE DATASET INTO A NUMPY ARRAY.
- NORMALIZE THE DATA USING STANDARDSCALER BY BOTH FITTING AND TRANSFORMING IT.
- DIVIDE THE STANDARDIZED DATA INTO TRAINING AND TESTING SETS USING TRAIN_TEST_SPLIT.
- CONSTRUCT A GRIDSEARCHCV OBJECT WITH A CROSS-VALIDATION SETTING OF 10 FOLDS FOR PARAMETER TUNING.
- APPLY GRIDSEARCHCV TO VARIOUS MACHINE LEARNING ALGORITHMS, INCLUDING LOGISTIC REGRESSION, SUPPORT VECTOR MACHINES, DECISION TREES, AND K-NEAREST NEIGHBORS.
- COMPUTE THE ACCURACY OF EACH MODEL ON THE TEST DATA USING THE .SCORE() METHOD.
- ANALYZE THE CONFUSION MATRIX FOR EACH MODEL TO ASSESS THEIR PERFORMANCE IN TERMS OF TRUE POSITIVES, TRUE NEGATIVES, FALSE POSITIVES, AND FALSE NEGATIVES.
- DETERMINE THE BEST MODEL BY EVALUATING METRICS SUCH AS JACCARD SCORE, F1 SCORE, AND ACCURACY.



DATA SCIENCE
WINNING SPACE RACE WITH MATERIAL REUSE

RESULTS

RESULTS

EXPLORATORY DATA ANALYSIS

- OVERALL SUCCESS RATE OF LAUNCHES IMPROVES FROM 2013
- ORBITS ES-L1, GEO, HEO, SSO HAVE 100% SUCCESS RATE
- LAUNCH SITE CCAFD SLC 40 MOST USED LAUNCH SITE
- LAUNCH SITE KSC LC-39A HAS THE HIGHEST SUCESS RATE

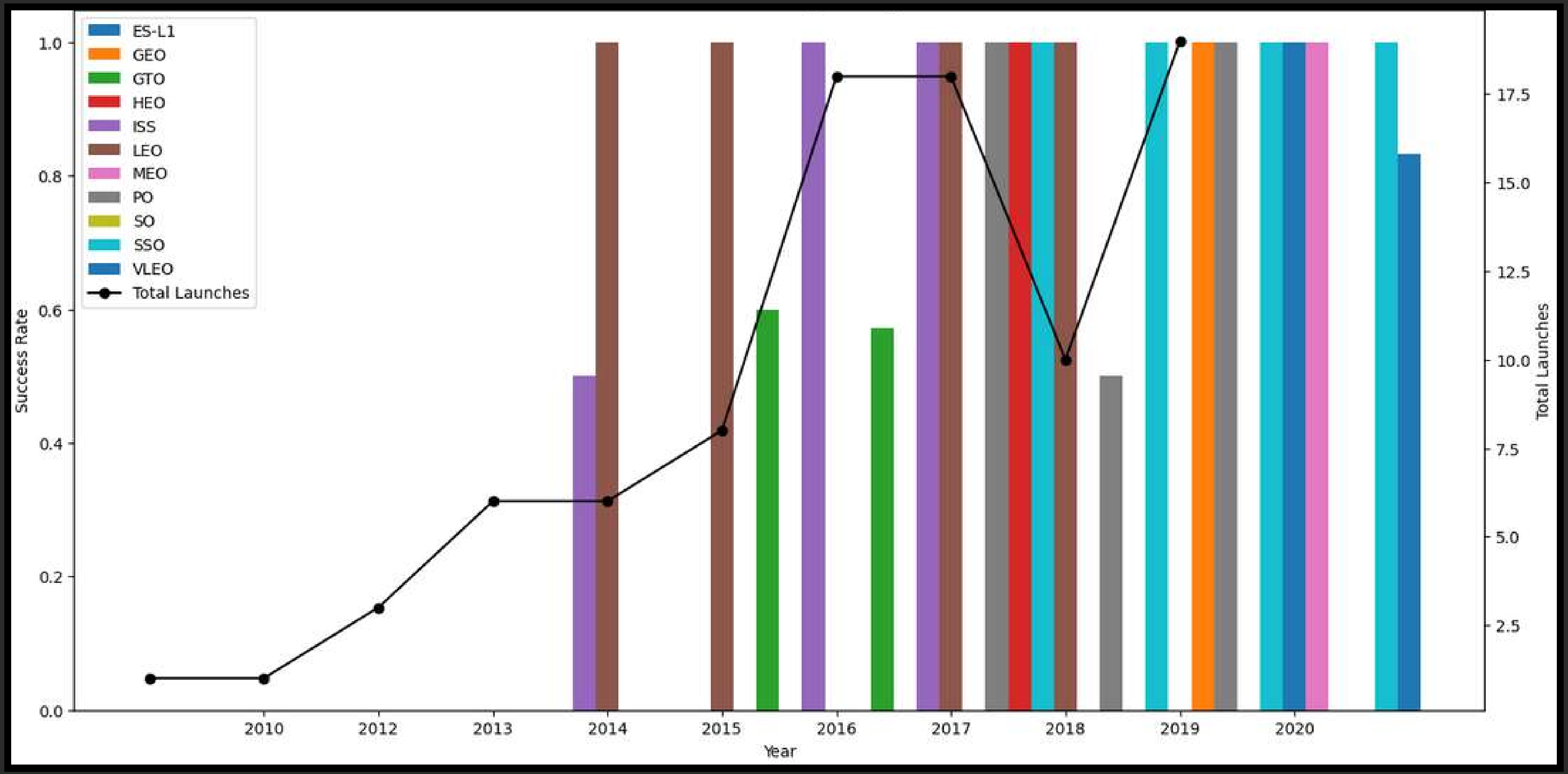
ANALYTICS & VISUALIZATION

- MOST LAUNCH SITES ARE LOCATED NEAR THE COAST
- ALL SITE LOCATED ABOVE CANCER TROPIC LINE
- ALL SITE LOCATED FAR FROM PUBLIC INFRASTRUCTURES

PREDICTIVE ANALYTICS

- SVM ACCURACY PERFORMED 1.75% BETTER THAN OTHER MODELS
- ALL MODELS HAVE THE SAME JACCARD SCORE
- ALL MODELS HAVE THE SAME F1_SCORE

SUCCESS RATE OF LAUNCHES BY ORBIT TYPE OVER TIME



OVERALL PREDICTIVE MODEL RANKING

	LogReg	SVM	Tree	KNN
Jaccard_Score	0.800000	0.800000	0.800000	0.800000
F1_Score	0.888889	0.888889	0.888889	0.888889
Accuracy	0.833333	0.848214	0.833333	0.833333

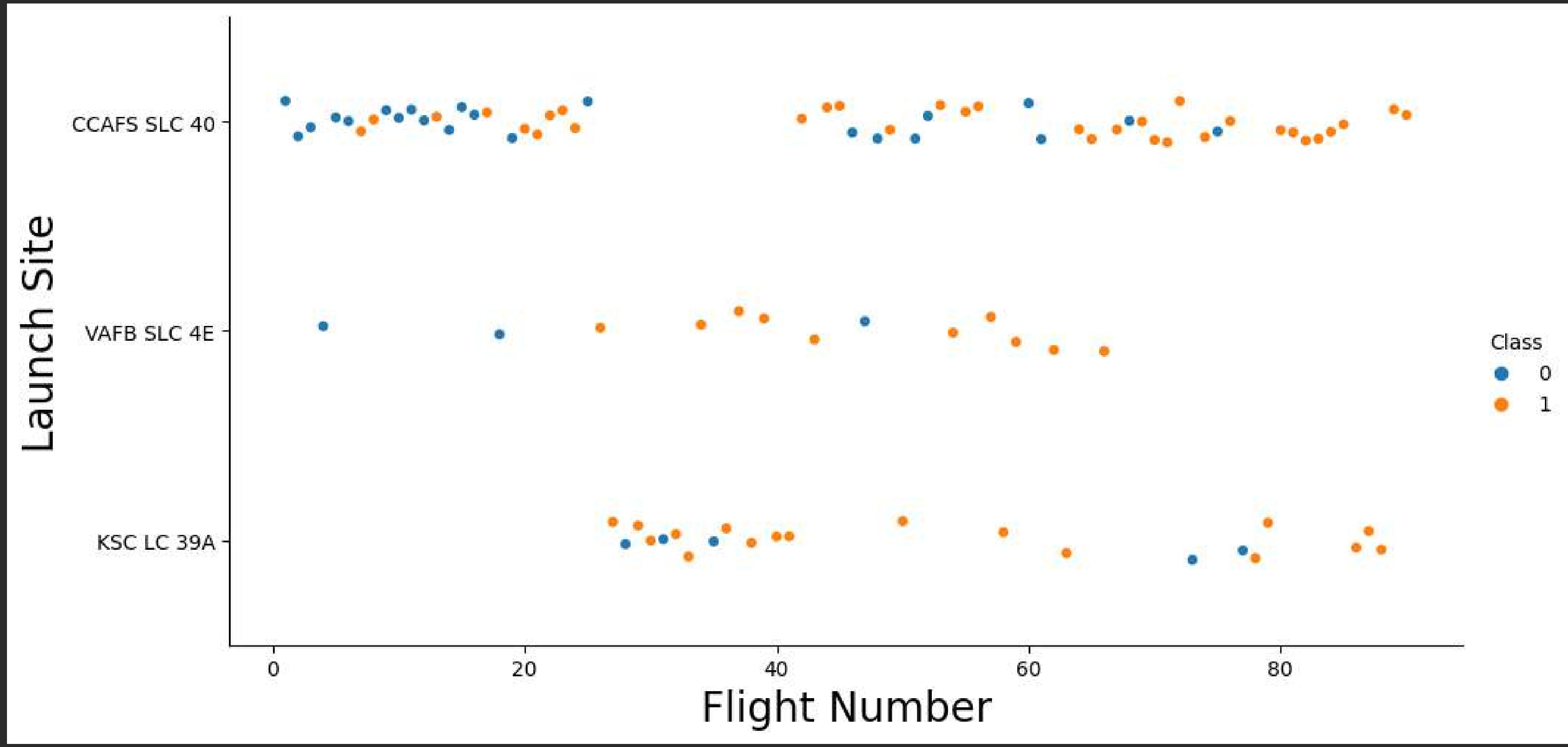
BOOSTERS LANDING VS. LAUNCH SITE

CLASS 0 = FAIL
CLASS 1 = SUCCESS

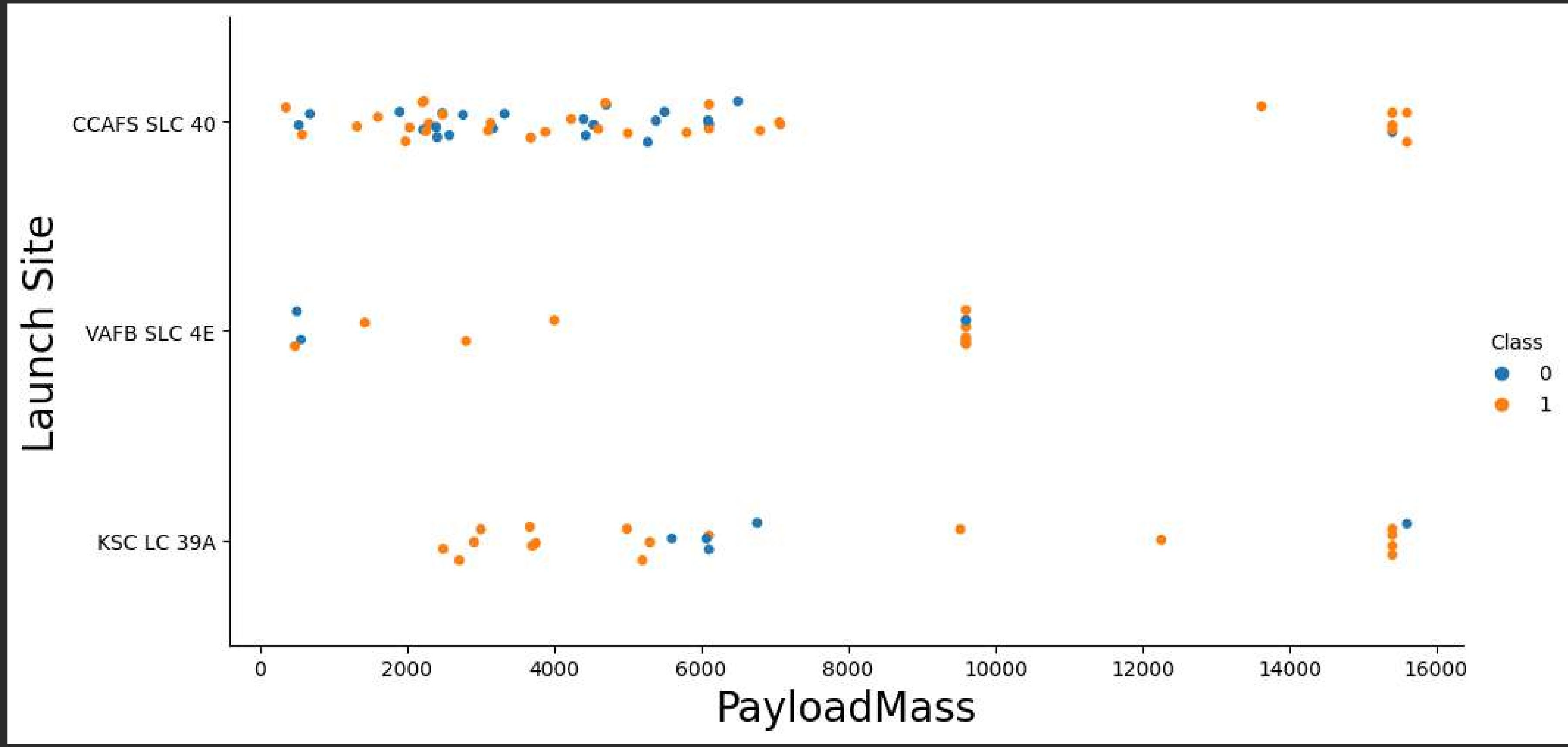
- FLIGHTS CONDUCTED IN THE PAST HAD A LOWER CHANCE OF SUCCESS.
- FLIGHTS CONDUCTED MORE RECENTLY HAD A BETTER CHANCE OF SUCCESS.
- APPROXIMATELY HALF OF THE LAUNCHES TOOK PLACE AT THE CCAFS SLC 40 LAUNCH SITE.
- THE LAUNCH SITES VAFB SLC 4E AND KSC LC 39A HAVE EXPERIENCED A HIGHER RATE OF SUCCESS.
- FROM THIS DATA, WE CAN CONCLUDE THAT RECENT LAUNCHES ARE MORE LIKELY TO SUCCEED.

- WHEN THE PAYLOAD MASS IS HIGHER, THERE IS A GREATER CHANCE OF A SUCCESS.
- 88% OF LAUNCHES WITH A PAYLOAD EXCEEDING 7,000 KILOGRAMS RESULTED IN SUCCESS.
- KSC LC 39A ACHIEVED A 100% SUCCESS RATE FOR LAUNCHES WITH PAYLOADS LESS THAN 5,500 KILOGRAMS.
- VAFB SLC 4E HAS NOT OPERATED ANY LAUNCHES WITH PAYLOADS GREATER THAN APPROXIMATELY 10,000 KILOGRAMS.

FLIGHT NUMBER VS LAUNCH SITE



PAYLOADMASS REPARTITION OVER LAUNCH SITE



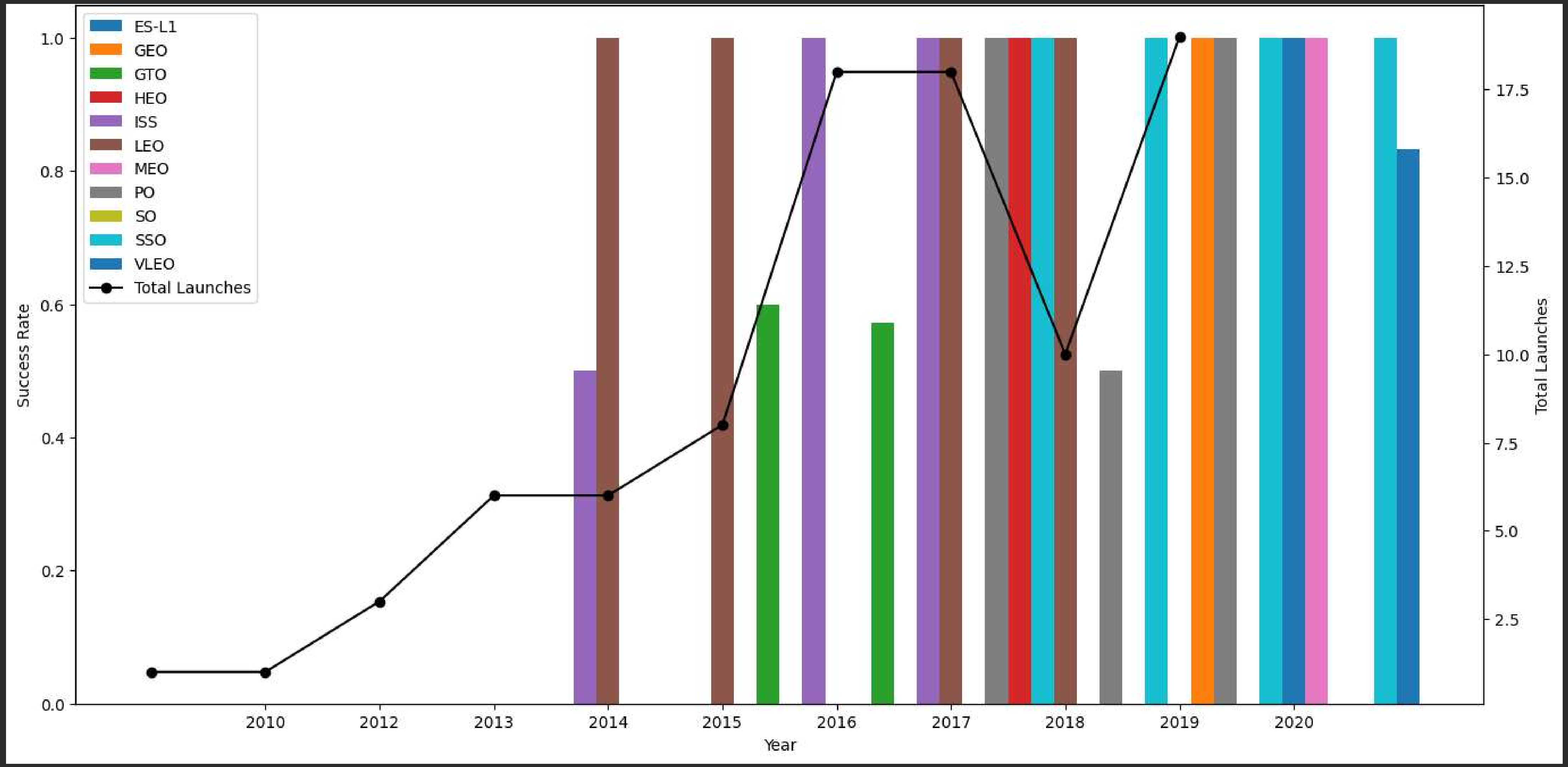
BOOSTERS LANDING VS. ORBIT TYPE

SUCCESS RATE IN %

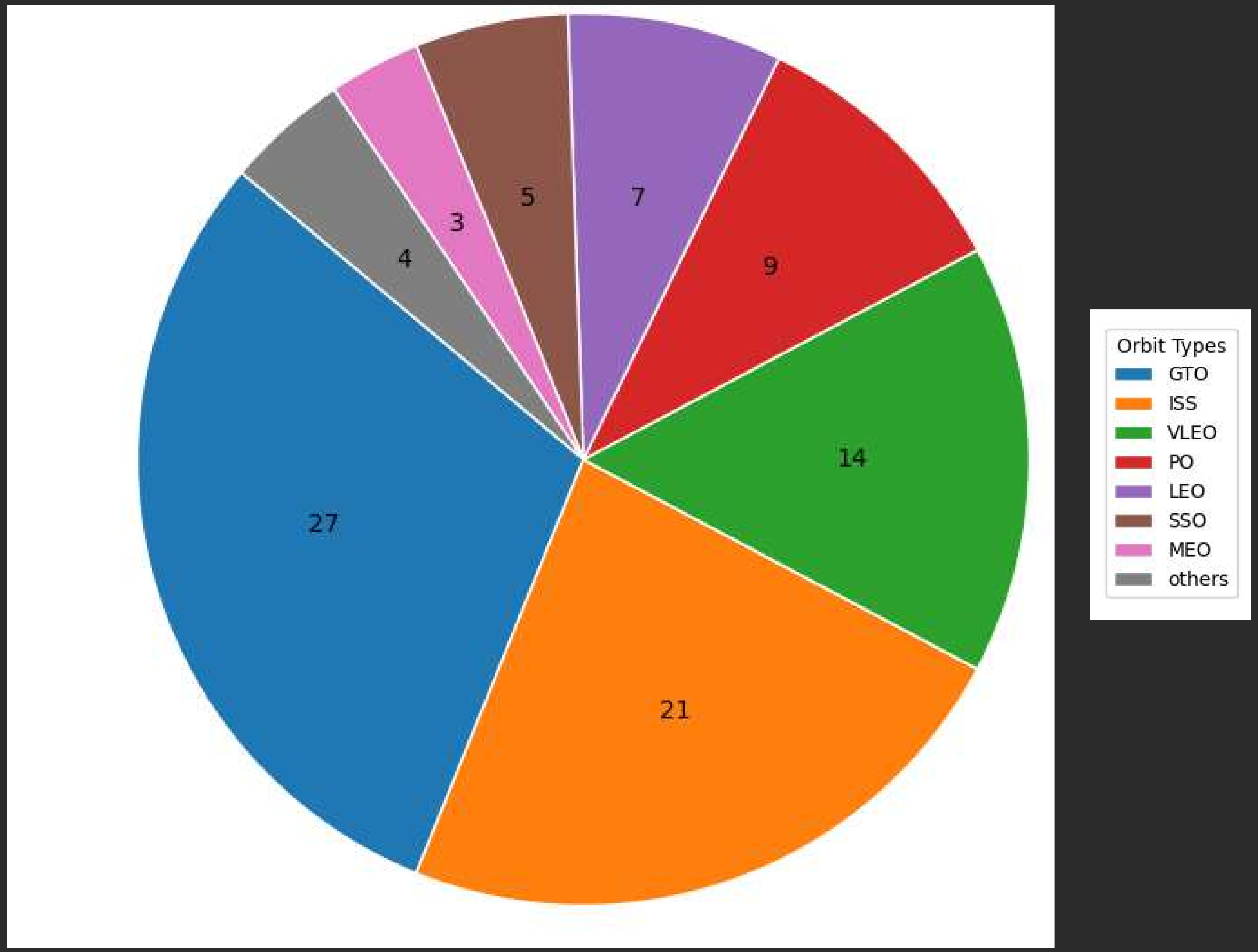
- **100% SUCCESS RATE:** ES-L1, GEO, HEO AND SSO
- **70%-99% SUCCESS RATE**: VLEO, LEO
- **0% SUCCESS RATE:** SO (ONLY 1 MISSION)

- **68.9% OF ALL MISSIONS ARE ASSOCIATED WITH JUST THREE TYPES OF ORBITS:**
  - **GTO: 30%**
  - **ISS: 23.3%**
  - **VLEO: 15.6%**

SUCCESS RATE OF LAUNCHES BY ORBIT TYPE OVER TIME



ORBIT TYPE MISSION DISTRIBUTION COUNT



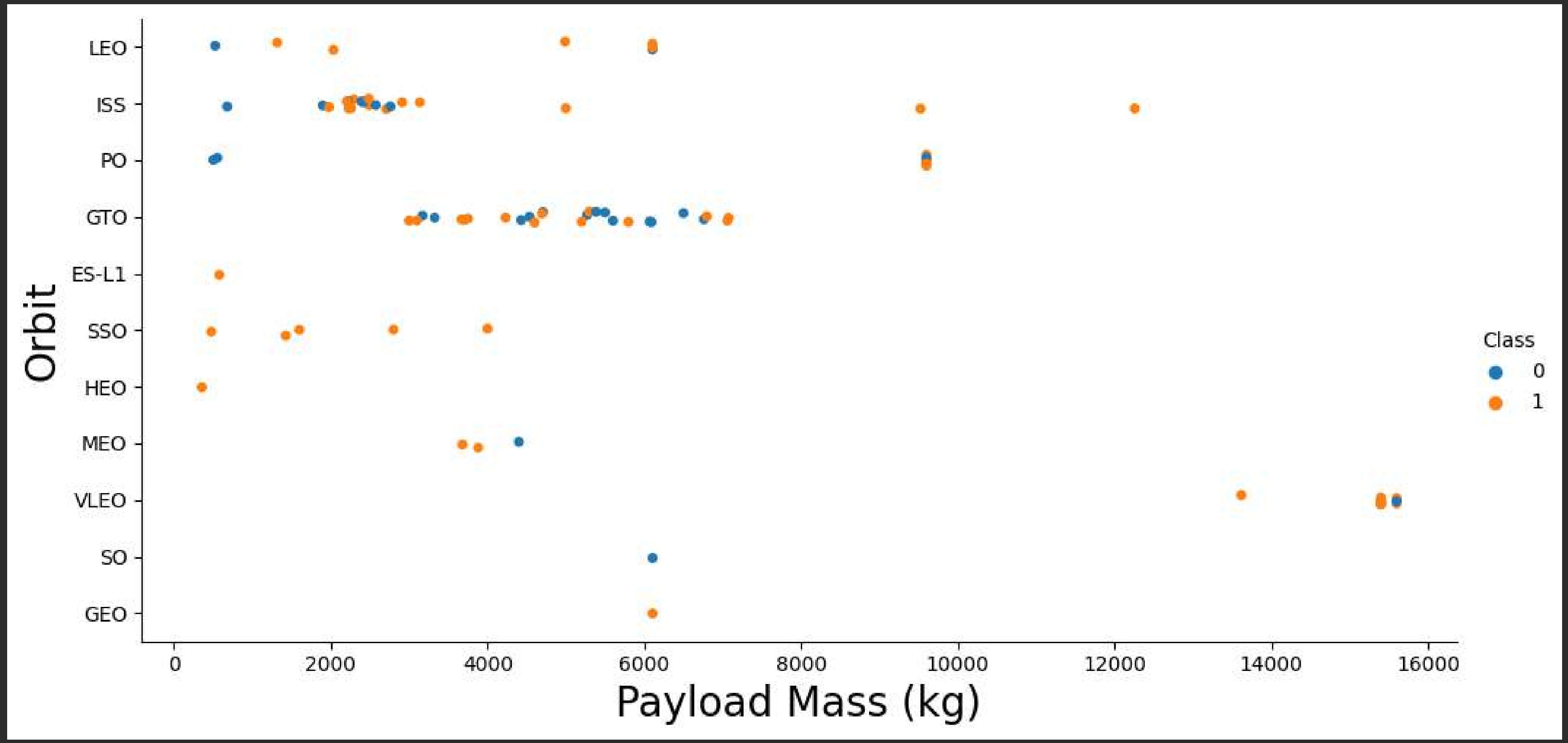
BOOSTERS LANDING VS. ORBIT TYPE

CLASS 0 = FAIL
CLASS 1 = SUCCESS

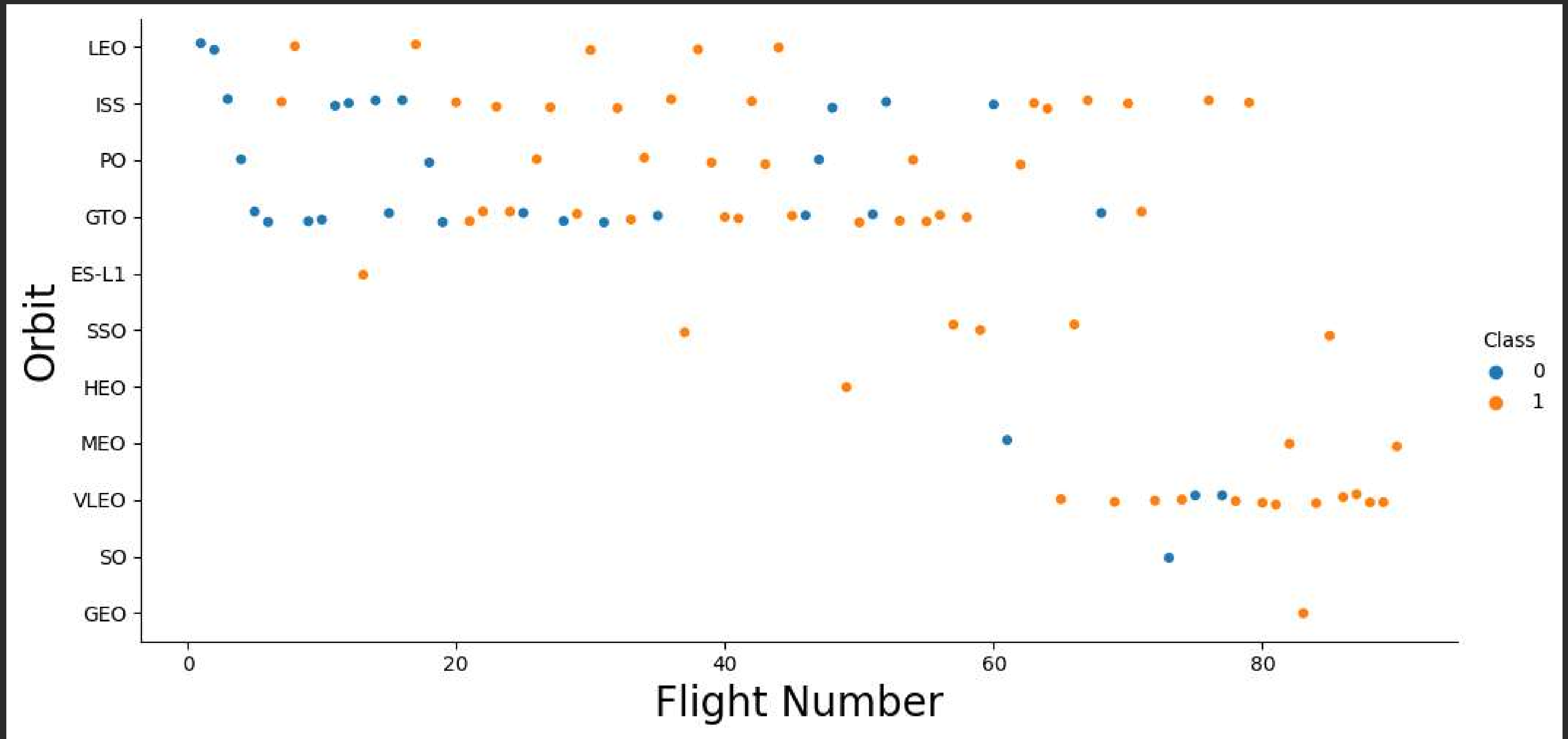
- VLEO MISSION HAS CARRIED THE HEAVIEST PAYLOADS
- VLEO HAS HEAVY PAYLOADS MISSION (ABOVE 13000KG)
- MOST ISS MISSIONS ARE BELOW 3000KG PAYLOAD
- ALL SSO MISSIONS ARE BELOW 4000KG
- GTO MISSIONS NEVER CARRIED MORE THAN 8000KG
- HEAVY PAYLOADS PERFORM WELL IN LEO, ISS, AND PO ORBITS.

- GTO ORBIT FLIGHTS WERE MORE FREQUENT IN EARLIER FLIGHTS
- VLEO ORBIT FLIGHTS ARE DOMINANT SINCE FLIGHT N°64
- 85.71% OF VLEO ORBIT FLIGHTS ARE SUCCESSFUL
- FIRST GEO ORBIT FLIGHT WAS SUCCESSFUL
- ISS ORBIT FLIGHTS ARE ON A 6 WINS-STREAK

PAYLOAD MASS (KG) VS ORBIT TYPE



FLIGHT NUMBER VS ORBIT TYPE



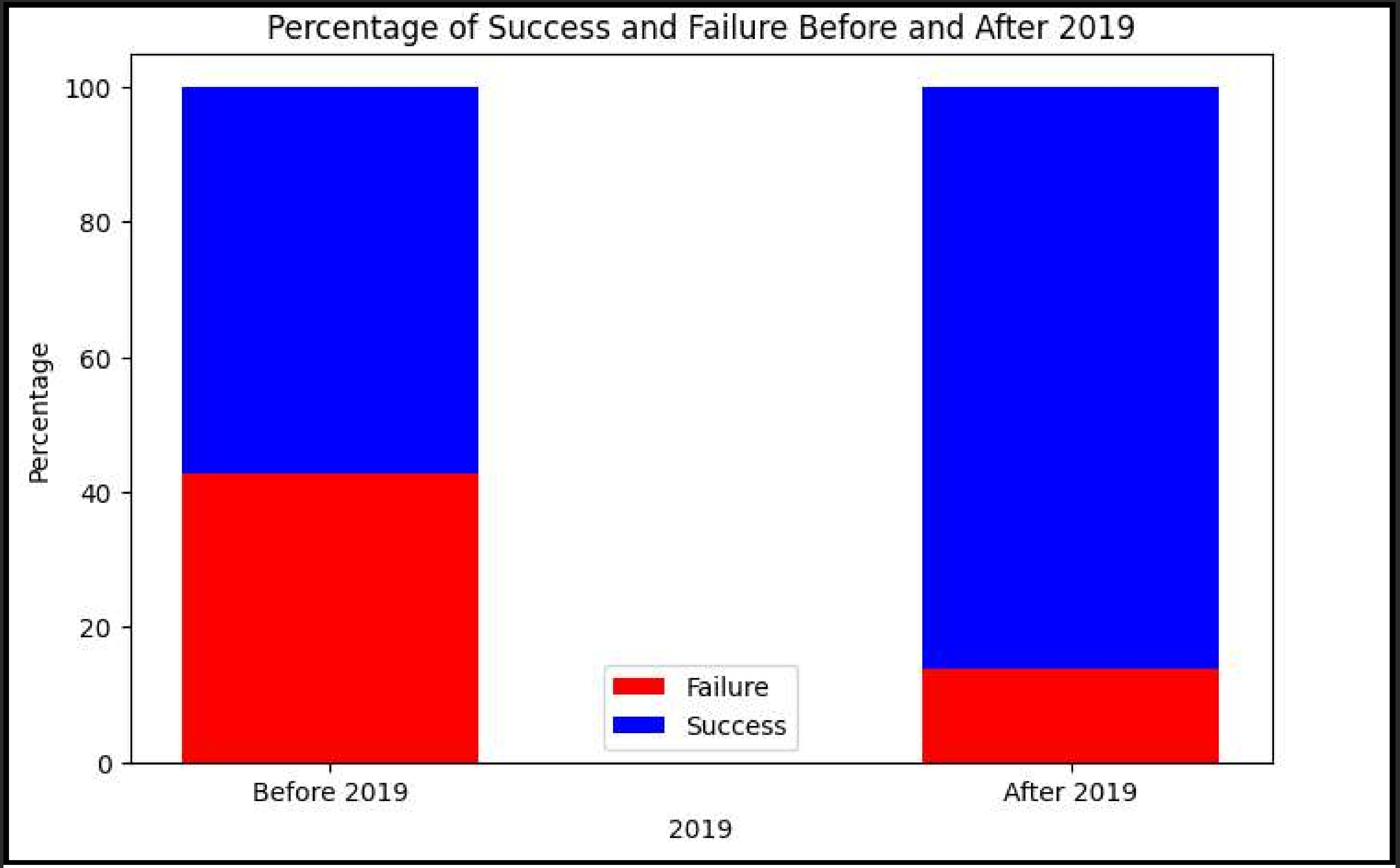
SUCCESS AND FAILURE OVERTIME

SUCCESSFUL LANDING:

- **SINCE 2019,** SUCCESS POURCENTAGE IS **76.59%**
- **PRIOR TO 2019,** SUCCESS POUCENTAGE WAS **55.81%**

FAILURE LANDING

- **SINCE 2019,** FAILURE POURCENTAGE IS **23.40%**
- **PRIOR TO 2019**, FAILURE POUCENTAGE WAS **44.19%**

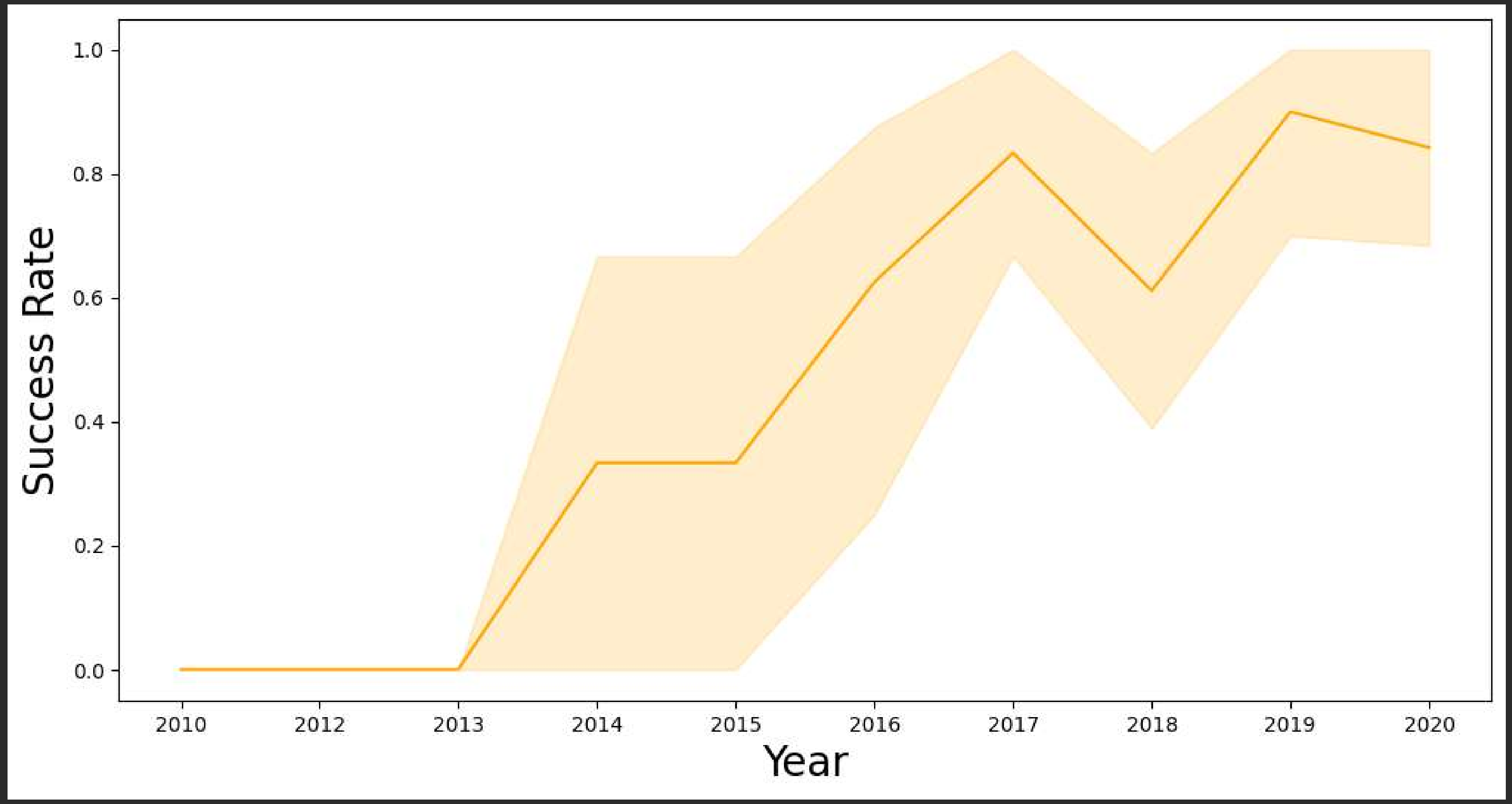



```
Percentage of '0' (Failures) before 2019: 42.6230%
Percentage of '1' (Successes) before 2019: 57.3770%
Percentage of '0' (Failures) after 2019: 13.7931%
Percentage of '1' (Successes) after 2019: 86.2069%
```

BOOSTERS LANDING SUCCESS RATE

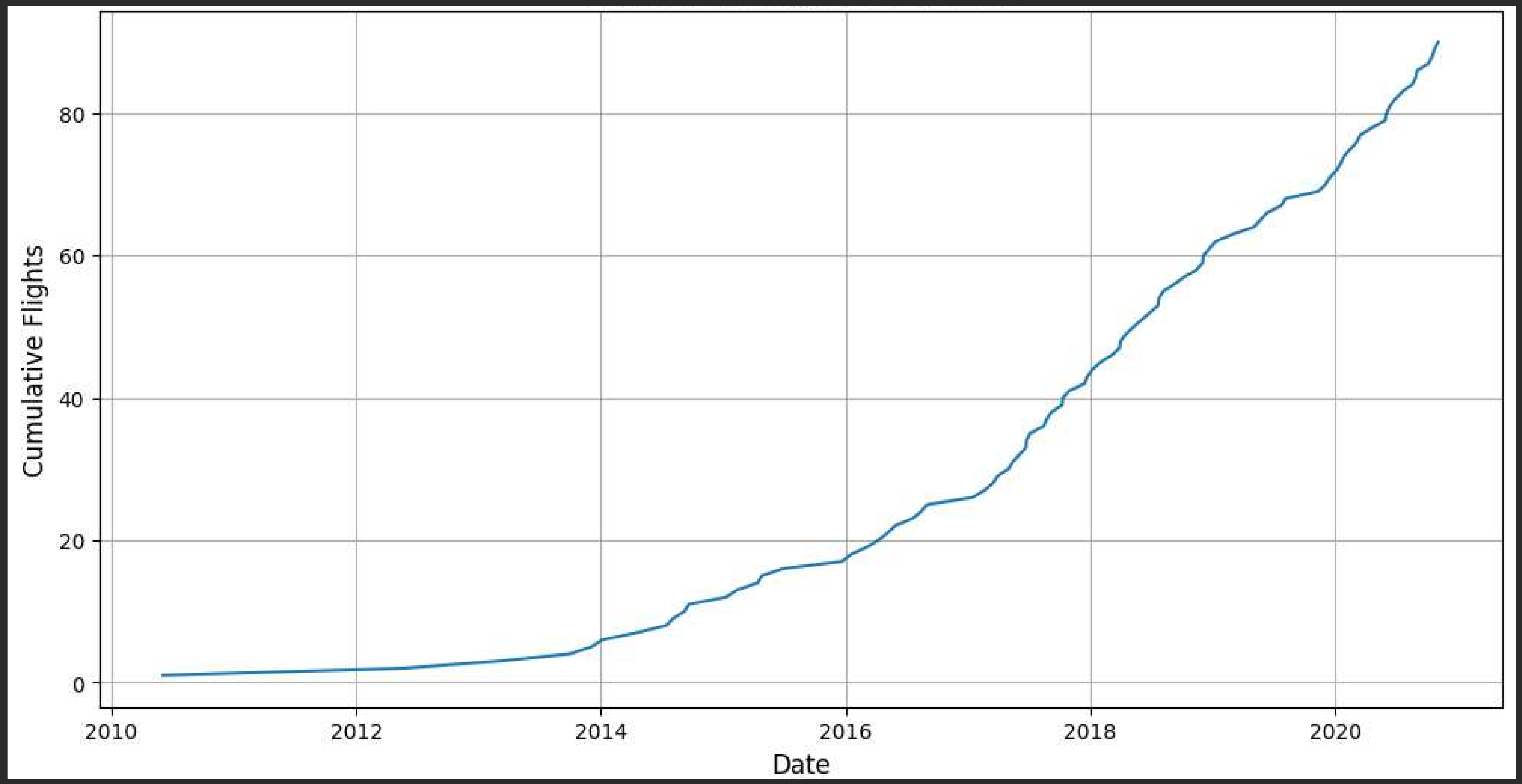
- SPACEX HAD A 0% SUCCESS RATE FROM 2010 TO 2013
- FROM 2013 SUCCESS RATE STARTED TO INCREASE
- IN 2017 SUCCESS RATE REACHED 84%
- IN 2019 IT REACHED ITS ALL TIME HIGH AT 90%
- OVERALL SUCCESS RATE HAS INCREASE OVERTIME TO SIT AT 85%

SUCCESS RATE OVER TIME



- SPACEX FALCON 9 FLIGHTS NUMBER HAS BEEN GROWING FASTER SINCE 2014
- THE LINE IS ACCELARATING FROM 2018
- OVERALL THE NUMBER OF FLIGHTS OVERTIME AND THE NUMBER OF SUCCESSFUL BOOSTERS LANDING OVERTIME ARE BOTH GROWING POSITIVELY

CUMALATIVE FLIGHTS OVER TIME



LAUNCH SITES INFORMATION

LAUNCH SITE NAMES:
- CCAFS LC-40
- CCAFS SLC-40
- KSC LC-39A
- VAFB SLC-4E

	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

LAUNCH OUTCOMES
- GREEN MARKER FOR SUCCESSFUL LAUNCH
- RED MARKER FOR UNSUCCESFUL LAUNCH

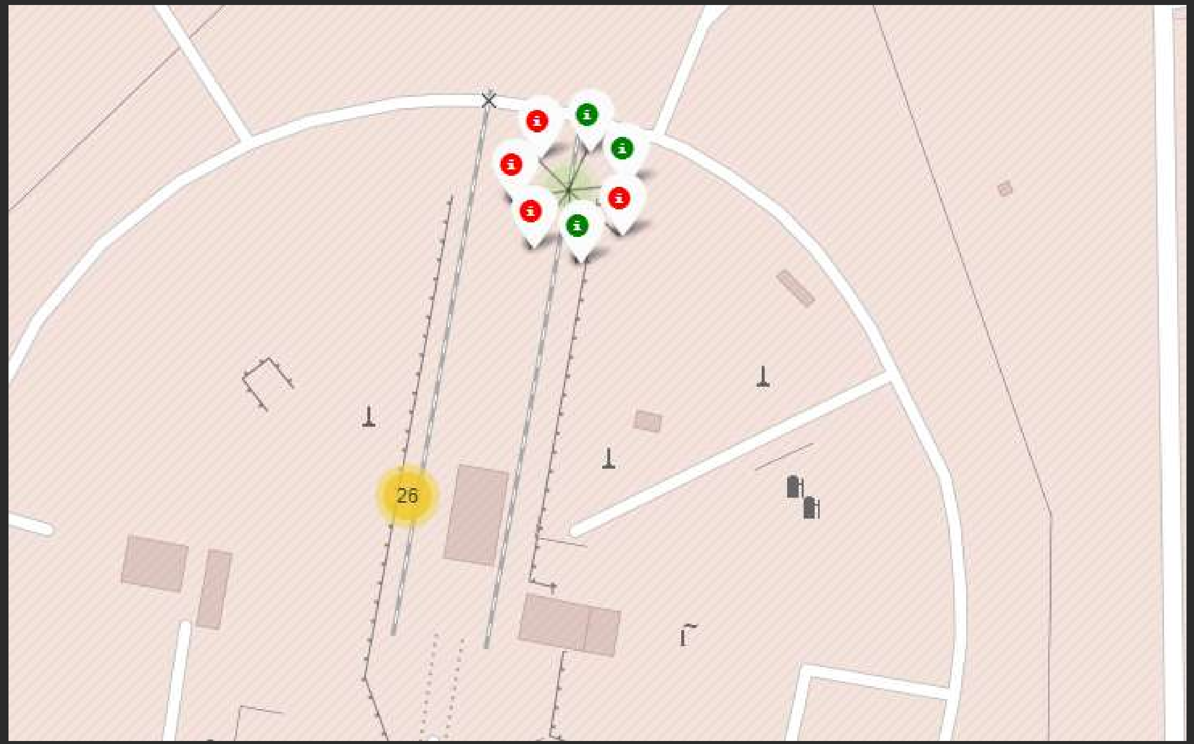
PROXIMITY CHECKS
- 0.88 KM TO THE NEAREST COASTLINE
- 22.32 KM TO THE CLOSEST RAILWAY
- 22.90 KM TO THE CLOSEST CITY
- 27.2 KM TO THE CLOSEST HIGHWAY

NO-INFRASTRUCTURE AREA
- CLOSEST PUBLIC INFRASTRUCTURE IS LOCATED 22KM AWAY FROM THE CENTER

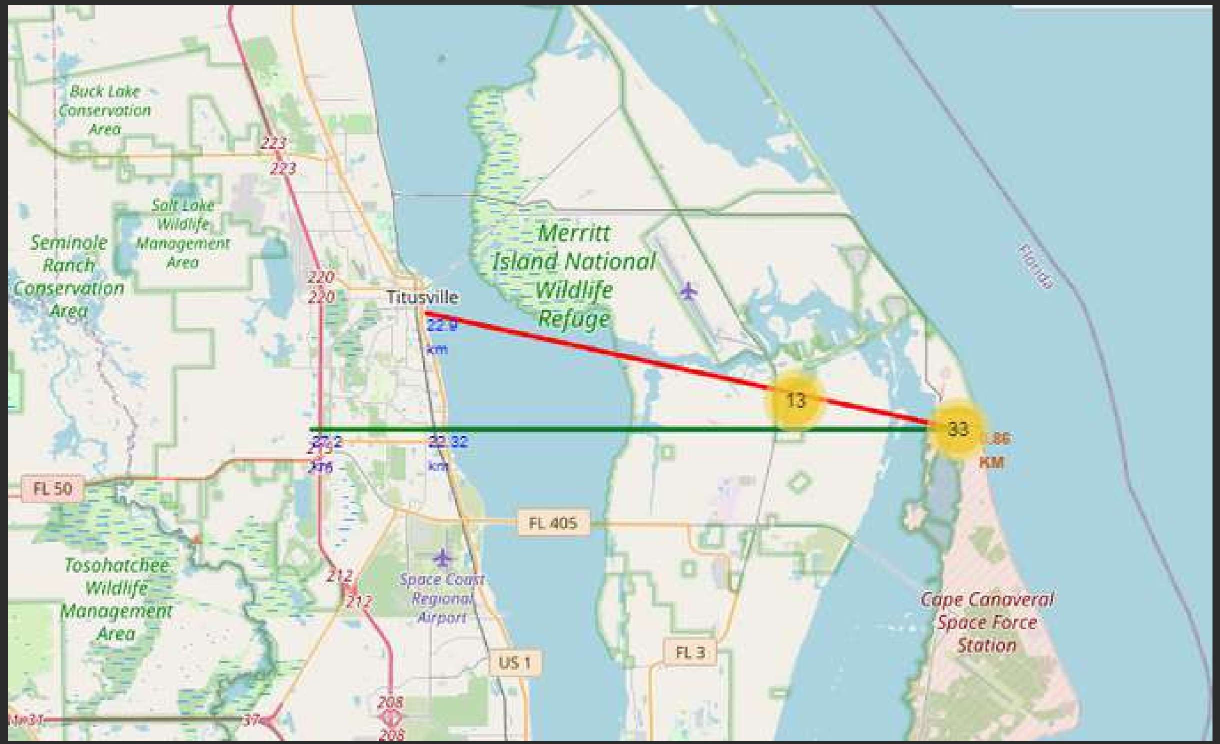
COAST PROXIMITY
- LAUNCH CENTER ARE LOCATED WITHIN A MILE OFF THE COASTS

LAUNCH SITES LOCATIONS

LAUNCH OUTCOMES



LAUNCH SITES AREA ANALYSIS



FALCON 9 BOOSTERS INFORMATION

ALL FALCON 9 VERSIONS USED:

- 0 FALCON 9 B0003
- 1 FALCON 9 B0005
- 2 FALCON 9 B0007
- 3 FALCON 9 B1003
- 4 FALCON 9 B1004
- 5 FALCON 9 B1005
- 6 FALCON 9 B1006
- 7 FALCON 9 B1007
- 8 FALCON 9 B1008
- 9 FALCON 9 B1010
- 10 FALCON 9 B1011
- 11 FALCON 9 B1012
- 12 FALCON 9 B1013
- 13 FALCON 9 B1015
- 14 FALCON 9 B1016
- 15 FALCON 9 B1017
- 16 FALCON 9 B1018
- 17 FALCON 9 B1019
- 18 FALCON 9 B1020
- 19 FALCON 9 B1021
- 20 FALCON 9 B1022
- 21 FALCON 9 B1023
- 22 FALCON 9 B1025
- 23 FALCON 9 B1026
- 24 FALCON 9 B1028
- 25 FALCON 9 B1029
- 26 FALCON 9 B1030
- 27 FALCON 9 B1031
- 28 FALCON 9 B1032
- 29 FALCON 9 B1034
- 30 FALCON 9 B1035
- 31 FALCON 9 B1036
- 32 FALCON 9 B1037
- 33 FALCON 9 B1038
- 34 FALCON 9 B1039
- 35 FALCON 9 B1040
- 36 FALCON 9 B1041
- 37 FALCON 9 B1042
- 38 FALCON 9 B1043
- 39 FALCON 9 B1044
- 40 FALCON 9 B1045
- 41 FALCON 9 B1046
- 42 FALCON 9 B1047
- 43 FALCON 9 B1048
- 44 FALCON 9 B1049
- 45 FALCON 9 B1050
- 46 FALCON 9 B1051
- 47 FALCON 9 B1054
- 48 FALCON 9 B1056
- 49 FALCON 9 B1058
- 50 FALCON 9 B1059
- 51 FALCON 9 B1060
- 52 FALCON 9 B106

FALCON 9 BOOSTERS:

- 53 VERSIONS RELEASED SO FAR
- EACH MODEL CARRIES SPECIFIC PAYLOADS
- B1049 VERSION IS THE MOST USED VERSION WITH 6 FLIGHTS

FALCON 9 BOOSTERS:

- B1049 VERSION IS ALWAYS CARRYING HEAVIER PAYLOADS.
- B1049 HAS IS THE MOST USED BOOSTER AND IS ALSO RANK 1 HEAVYLIFTER BY PAYLOAD (AVERAGED)
- B1049 FAVORITE ORBIT IS VLEO WITH 4 FLIGHTS

TOTAL N° OF FLIGHTS BY BOOSTER

	Value	Count
0	B1049	6
1	B1051	5
2	B1059	4
3	B1056	4
4	B1048	4
5	B1046	4
6	B1060	3
7	B1058	3
8	B1047	3
9	B1031	2

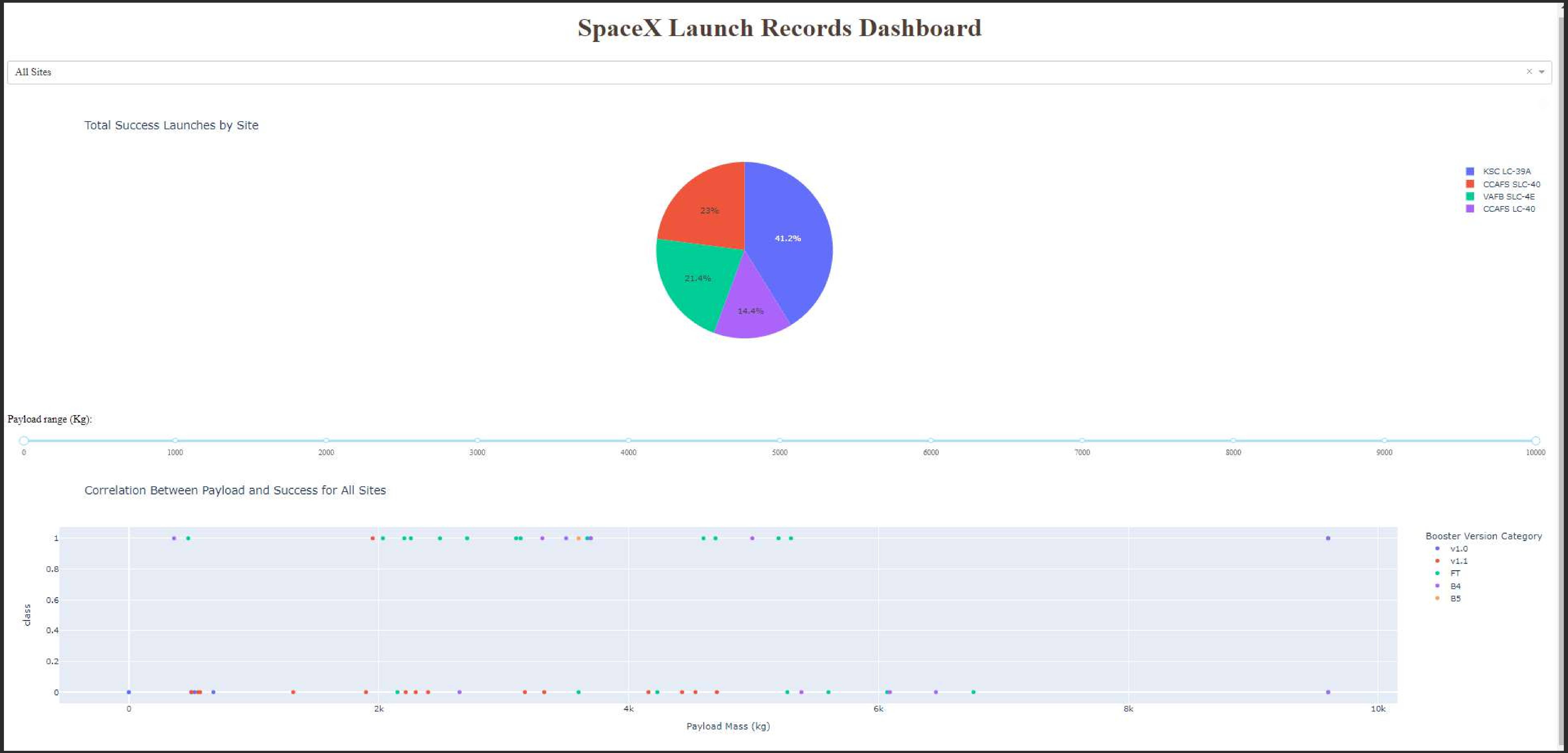
AVG. PAYLOAD MASS BY BOOSTER

	Booster_Serial	PayloadMass
0	Falcon 9 B1049	12746.666667
1	Falcon 9 B1051	12016.800000
2	Falcon 9 B1060	11560.000000
3	Falcon 9 B1048	10900.000000
4	Falcon 9 B1058	10343.319804
5	Falcon 9 B1041	9600.000000
6	Falcon 9 B1036	9600.000000
7	Falcon 9 B1037	6761.000000
8	Falcon 9 B1056	6727.425000
9	Falcon 9 B1029	6634.500000

ORBIT COUNT BY BOOSTER

	Booster_Serial	Orbit	Count
0	Falcon 9 B1049	VLEO	4
1	Falcon 9 B1047	GTO	3
2	Falcon 9 B1051	VLEO	3
3	Falcon 9 B1056	ISS	2
4	Falcon 9 B1041	PO	2
5	Falcon 9 B1039	ISS	2
6	Falcon 9 B1046	GTO	2
7	Falcon 9 B1048	VLEO	2
8	Falcon 9 B1059	ISS	2
9	Falcon 9 B1035	ISS	2

DYNAMIC & INTERACTIVE DASHBOARD DASH & PLOTLY



DATA SCIENCE
WINNING SPACE RACE WITH MATERIAL REUSE

PREDICTIVE ANALYSIS

CLASSIFICATION ACCURACY

DECISION TREE CLASSIFIER OBTAINED THE HIGEST CLASSIFICATION ACCURACY

- ALL THE MODEL PRODUCED THE SAME OUTPUTS IN TERM OF ACCURACY AND SCORING.
- DECISION TREE MODEL CLASSIFIER WAS ABLE TO GET A 0.8875 SCORE.

**DATA PROVIDED SMALL RELATIVELY SMALL TO PERFORM SUCH ML TECHNIQUES*
**WITH SOME TWISTING SVM ACHIEVED A BETTER ACCURACY (1 OUT OF 1*10^9)*

OVERALL PREDICTIVE MODEL RANKING

	LogReg	SVM	Tree	KNN
Jaccard_Score	0.800000	0.800000	0.800000	0.800000
F1_Score	0.888889	0.888889	0.888889	0.888889
Accuracy	0.833333	0.848214	0.833333	0.833333

```
Best model is DecisionTree with a score of 0.8875
Best parameters are:
criterion: entropy
max_depth: 4
max_features: auto
min_samples_leaf: 2
min_samples_split: 2
splitter: best
```

```python
# Dictionary to store the best scores for each model
# Format "Xmodel_cv.best_score_"
models = {
    'KNeighbors': knn_cv.best_score_,
    'DecisionTree': tree_cv.best_score_,
    'LogisticRegression': logreg_cv.best_score_,
    'SupportVector': svm_cv.best_score_
}

# Find the best model with its score
best_algorithm = max(models, key=models.get)
best_score = models[best_algorithm]

# Dictionary to store the best parameters for each model
# format "Xmodel_cv.best_params_"
best_params = {
    'DecisionTree': tree_cv.best_params_,
    'KNeighbors': knn_cv.best_params_,
    'LogisticRegression': logreg_cv.best_params_,
    'SupportVector': svm_cv.best_params_
}

# Winner
print(f"Best model is {best_algorithm} with a score of {best_score:.4f}")

# Best parameters of winner
if best_algorithm in best_params:
    print("Best parameters are:")
    for param, value in best_params[best_algorithm].items():
        print(f"{param}: {value}")
```

CONFUSION MATRIX

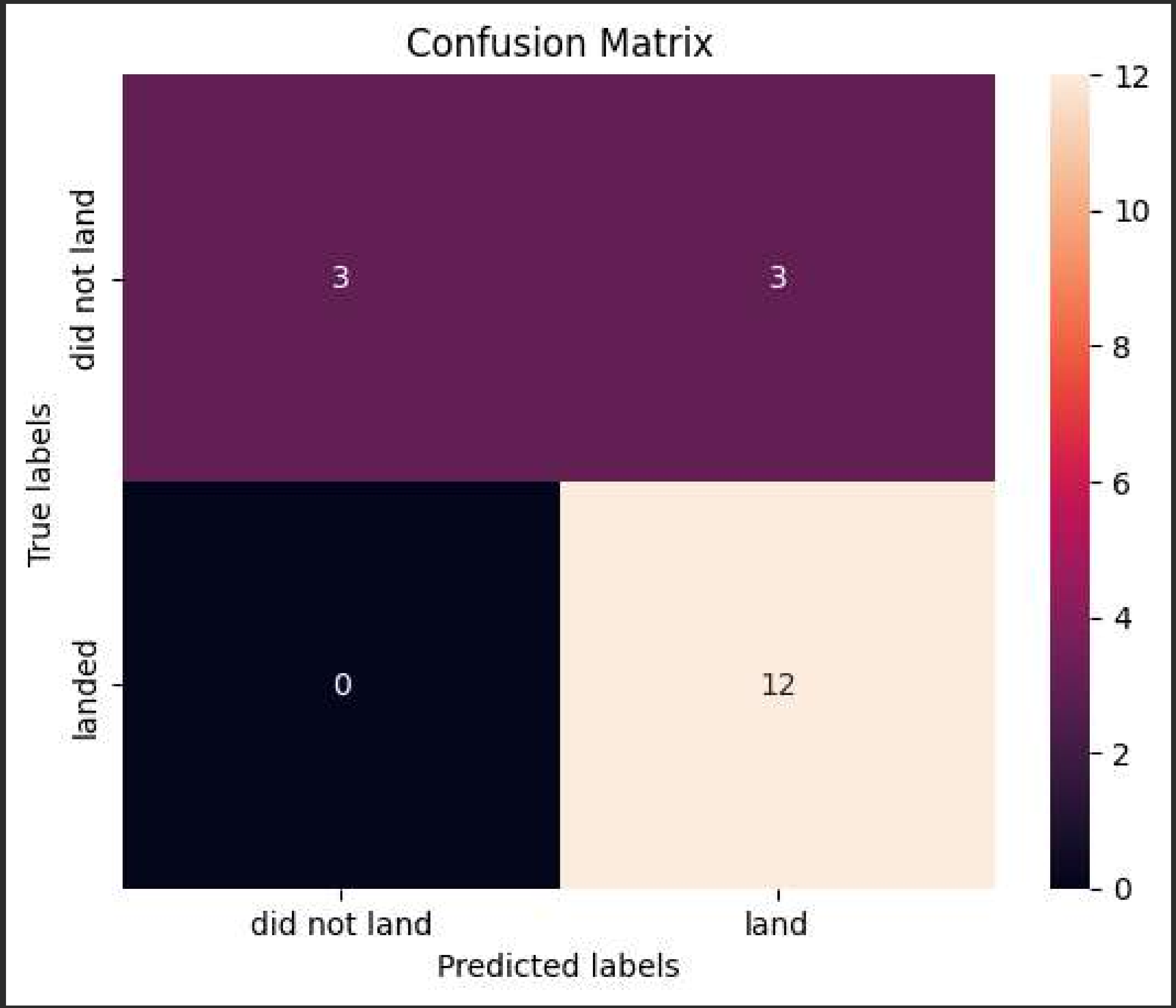
A CONFUSION MATRIX GIVES AN OVERVIEW OF PERFORMANCES FOR A GIVEN CLASSIFICATION MODEL.

OUTPUT

- 12 TRUE POSITIVE
- 3 TRUE NEGATIVE
- 3 FALSE POSITIVE
- 0 FALSE NEGATIVE

SUMMARY

- 12 INSTANCES CORRECTLY CLASSIFIED AS POSITIVE.
- 3 INSTANCES CORRECTLY CLASSIFIED AS NEGATIVE.
- 3 INSTANCES INCORRECTLY CLASSIFIED AS POSITIVE WHEN THEY WERE ACTUALLY NEGATIVE.
- 0 INSTANCES INCORRECTLY CLASSIFIED AS NEGATIVE WHEN THEY WERE ACTUALLY POSITIVE.



CONCLUSION

KEY PHRASE: SPACEX SAVES APPROXIMATIVELY $37 MILLION DOLLARS SAVED PER SUCCESSFUL LANDING.

SUCCESS COMES FROM FAILURE.

IT TOOK 3 YEARS TO SPACEX TO SEE HIS FIRST SUCCESSFULL BOOSTER LANDING BACK TO EARTH (2013). TODAY WITH 10 YEARS OF DATA TO ANALYSIS WE CAN SEE HOW THE NUMBER OF FLIGHTS AND THE SUCCESS RATE ARE CORRELATED. AS OF TODAY THE SUCCESS RATE IS AS HIGH AS 85%

PROSPERITY COMES FROM SUCCESS

SPACEX IS ABLE TO PRICE HIS ROCKET LAUNCH AT **$62M,** WHICH IS ON AVERAGE **$103M** LESS THAN ITS COMPETITORS. THIS PRICE IS BASED ON THE ABILITY TO BRING BACK THEIR BOOSTERS SUCCESSFULLY.
CNBC* PUBLISHED THAT THE AVERAGE PRICE FOR A BOOSTER LIES AROUND $37M. CONSIDERING ONE OF THE LATEST FALCON 9 BOOSTERS, THE **B1058** WHICH HAS **16 LAUNCHES** AS OF TODAY, THAT IS **$592M** SAVED, MORE THAN HALF OF BILLION WITH THIS ONE VERSION.

PAYLOAD MASS MATTERS.

THROUGHOUT THE ANALYSIS WE DEMONSTRATED THAT THE HIGHER THE PAYLOAD MASS, THE HIGHER THE SUCCESS RATE. THE MODEL PRODUCED THE SAME OUTPUTS IN TERM OF ACCURACY AND

SUCCESSFUL LANDINGS ARE STRIVING. FAILURE BELONGS TO THE PAST.

WE DEMONSTRATED THAT MOST FAILURES HAPPENED PRIOR TO 2019, AND AS OF TODAY **SUCCESSFUL LANDINGS** REPRESENT **85% OF ALL LAUNCHES**.

APPENDIX

RESSOURCES:

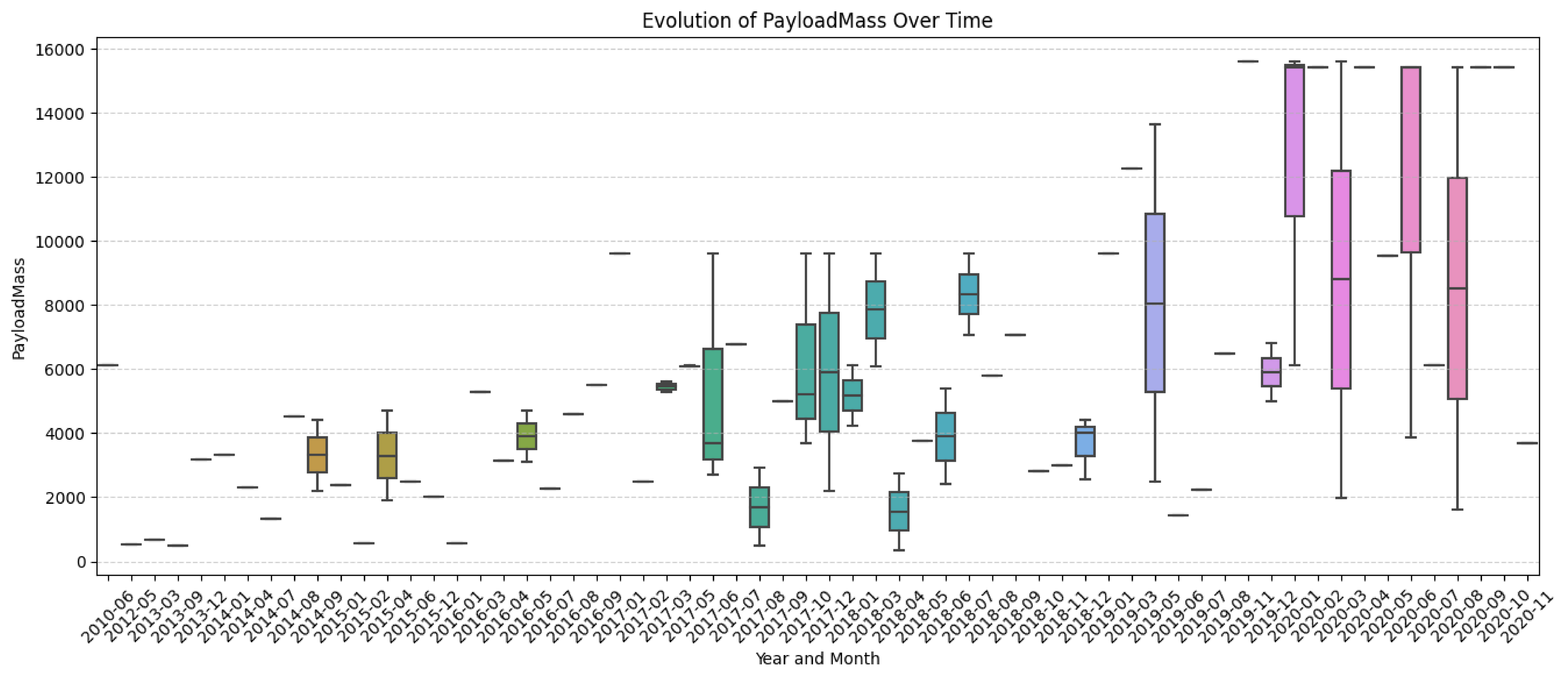
-FALCON 9 BOOSTERS INFORMATIONS: HTTPS://EN.WIKIPEDIA.ORG/WIKI/LIST_OF_FALCON_9_FIRST-STAGE_BOOSTERS
-**GENERAL INFORMATIONS ON SPACEX**: HTTPS://WWW.SPACEX.COM/LAUNCHES/MISSION/?MISSIONID=SL-7-2
-BOOSTER FALCON 9 COST: HTTPS://WWW.CNBC.COM/2018/05/11/FULL-ELON-MUSK-TRANSCRIPT-ABOUT-SPACEX-FALCON-9-BLOCK-5.HTML
-SPACEX GENERAL INFORMATIONS: HTTPS://WWW.BRITANNICA.COM/TOPIC/SPACEX
-DASH-PLOTLY DOCUMENTATION: HTTPS://DASH.PLOTLY.COM/
-PYTHON VISUALIZATION IDEAS: HTTPS://INFORMATIONISBEAUTIFUL.NET/WDVP/GALLERY-2019/
-SQL CHEAT SHEET QUERIES: HTTPS://WWW.SQLTUTORIAL.ORG/SQL-CHEAT-SHEET/
-PYTHON VISUALIZATION GALLERY & CODES: HTTPS://PYTHON-GRAPH-GALLERY.COM/

GITHUB PROJECT REPOSITORY:

- SPACEX-CAPSTONE-DATA-SCIENCE-IBM
- HTTPS://GITHUB.COM/FGJSPACEMAN/SPACEX-CAPSTONE-DATA-SCIENCE-IBM

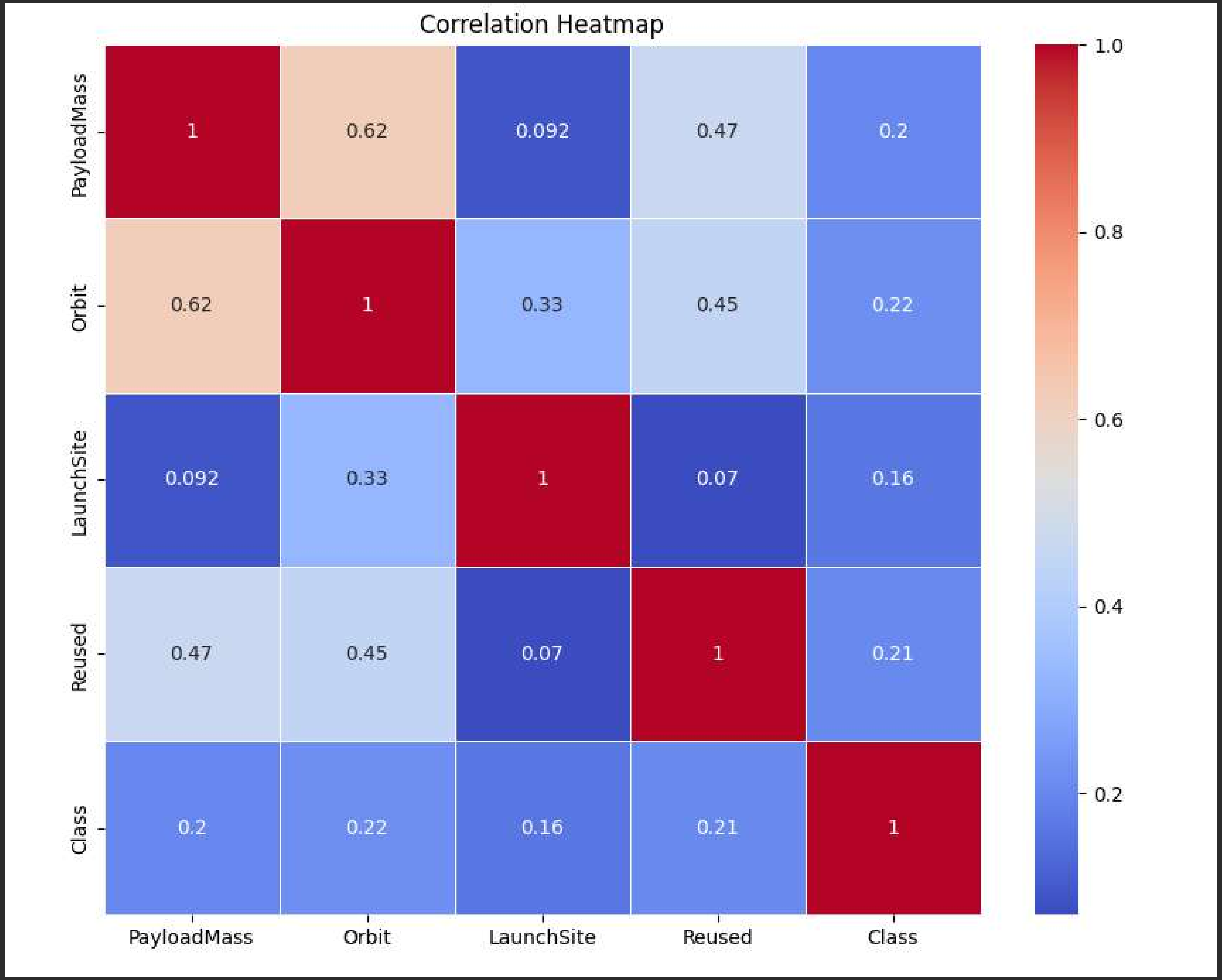
EXTRA CHARTS TO EXPLORE

09_SPACEX_VISUALIZATION_NOTEBOOK



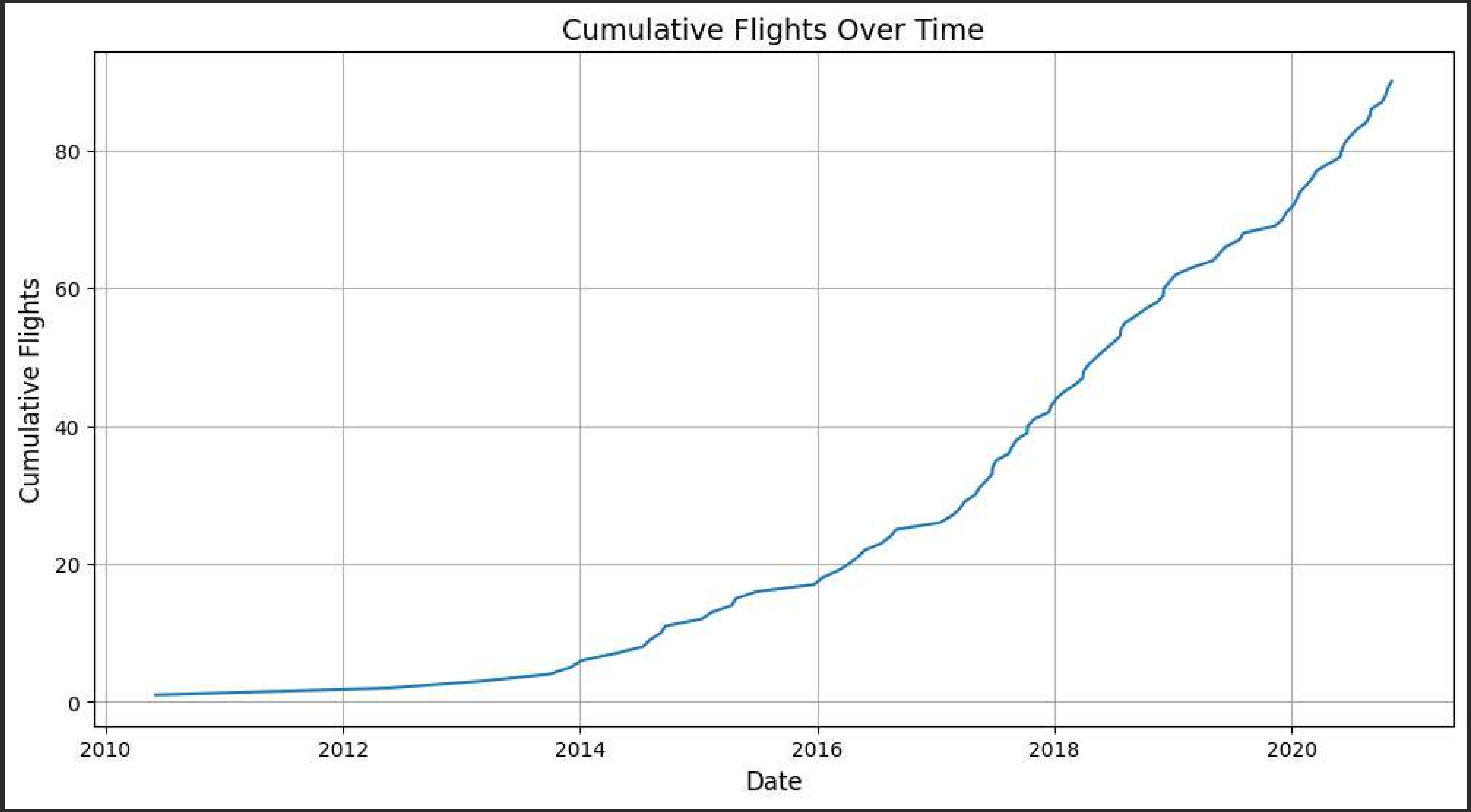
EXTRA CHARTS TO EXPLORE

09_SPACEX_VISUALIZATION_NOTEBOOK



EXTRA CHARTS TO EXPLORE

09_SPACEX_VISUALIZATION_NOTEBOOK



EXTRA CHARTS TO EXPLORE

09_SPACEX_VISUALIZATION_NOTEBOOK

	Orbit	Launches_2010	Launches_2012	Launches_2013	Launches_2014	Launches_2015	Launches_2016	Launches_2017	Launches_2018	Launches_2019
0	ES-L1	0	0	0	0	1	0	0	0	0
1	GEO	0	0	0	0	0	0	0	0	0
2	GTO	0	0	1	3	1	5	7	8	2
3	HEO	0	0	0	0	0	0	0	1	0
4	ISS	0	0	1	2	3	2	4	3	4
5	LEO	1	1	0	1	1	0	2	1	0
6	MEO	0	0	0	0	0	0	0	1	0
7	PO	0	0	1	0	0	1	4	2	1
8	SO	0	0	0	0	0	0	0	0	0
9	SSO	0	0	0	0	0	0	1	2	1
10	VLEO	0	0	0	0	0	0	0	0	2


THANK YOU