SPACE Y LAUNCH PREDICTIONS

Tommy Mazal

https://github.com/temazal/Capstone

4/21/23

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Space Y is an emerging company looking to compete in the space race to launch satellites and other equipment into orbit. SpaceX is a leading competitor that is capable of reusing its first stage rocket boosters saving on costs for redeployments. A successful return usually costs around $62MM while a new rocket build can cost upwards of $165MM.

Utilizing SpaceX's launch histories and success rates, an estimate of what it would cost to compete can be achieved. With web scraping, data analytics, visualizations, and predictive analysis an accurate model of successful launches can be calculated resulting in an estimate of launch costs.

This assessment determined a few things. Space Y would need to develop a rocket that can launch with an average cost of $86MM to compete with SpaceX, that or develop a reusable rocket booster akin to SpaceX. It would also be recommended to launch from the KSC LC-39A in Florida to have the highest percentage of successful launches. Focus on launching into the ES-L1, GEO, HEO, and SSO orbits is also recommended as these had the highest percent of success.

METHODOLOGY

METHODOLOGY

Data Collection

- Data collection API
- Data collection with Web Scraping
- Data Wrangling

Exploratory Data Analysis

- Exploratory Data Analysis with SQL
- Exploratory Data Analysis with Visualization

Interactive Visual Analytics and Dashboard

- Interactive Visual Analytics with Folium
- Interactive Dashboard with Ploty Dash

Predictive Analysis

- Machine Learning Prediction

DATA COLLECTION API

b'[{"fairings":{"reused":false,"recovery_attempt":false,"recovered":false,"ships":[]},"links":{"patch":{"small":"https://images2.imgbox.com/94/f2/NN6Ph45r_o.png","large":"https://images2.imgbox.com/5b/02/QcxHUb5V_o.png"},"reddit":{"campaign":null,"launch":null,"media":null,"recovery":null},"flickr":{"small":[],"original":[]},"presskit":null,"webcast":"https://www.youtube.com/watch?v=0a_00nJ_Y88","youtube_id":"0a_00nJ_Y88","article":"https://www.space.com/2196-spacex-inaugural-falcon-1-rocket-lost-launch.html","wikipedia":"https://en.wikipedia.org/wiki/DemoSat"},"static_fire_date_utc":"2006-03-17T00:00:00.000Z","static_fire_date_unix":1142553600,"net":false,"window":0,"rocket":"5e9d0d95eda69955f709d1eb","success":false,"failures":[{"time":33,"altitude":null,"reason":"merlin engine failure"}],"details":"Engine failure at 33 seconds and loss of vehicle","crew":[],"ships":[],"capsules":[],"payloads":["5eb0e4b5b6c3bb0006eeb1e1"],"launchpad":"5e9e4502f5090995de566f86","flight_number":1,"name":"FalconSat","date_utc":"2006-03-24T22:30:00.000Z","date_unix":1143239400,"date_local":"2006-03-25T10:30:00+12:00","date_precision":"hour","upcoming":false,"cores":[{"core":"5e9e289df35918033d3b2623","flight":1,"gridfins":false,"legs":false,"reused":false,"landing_attempt":false,"landing_success":null,"landing_type":null,"landpad":null}],"auto_update":true,"tbd":false,"launch_library_id":null,"id":"5eb87cd9ffd86e000604b32a"},{"fairings":{"reused":false,"recovery_attempt":false,"re[illegible]overed":false,"ships":[]},"links":{"patch":{"small":"https://images2.imgbox.com/f9/4a/ZboXReNb_o.png","large":"https:[illegible]es2.imgbox.com/80/a2/bkWotCIS_o.png"},"reddit":{"campaign":null,"launch":null,"media":null,"recovery":null},"flic[illegible]{"small":[],"original":[]},"presskit":null,"webcast":"https://www.youtube.com/watch?v=Lk4zQ2wP-Nc","youtube_id":"Lk4z[illegible]-1-rocket-fails-reach-or

Converted raw data into a more useable format

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

90 rows × 17 columns

- Utilized data from SpaceX's API from: https://api.spacexdata.com/v4/launches/past
- Pandas used to create a dataframe table
- Filtered out unnecessary columns and selected only Falcon 9 boosters
- Replaced missing payload masses with average values to keep consistent

DATA COLLECTION WITH WEB SCRAPING

[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 (Dr... 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 ... 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 p... 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.								
83	18 March 2020, 12:16[510]	F9 B5 ♺ B1048.5	KSC, LC-39A	S... v1				Success	Failure (drone ship)
	Fifth operational launch of Starlink satellites. It was the first time a first ... ge burn, the booster suffered premature shut down of an engine, the first of a Merlin 1D variant and first since the CRS-1 ... ing failure in a row, later revealed to be caused by residual cleaning fluid trapped inside a sensor.[513]								
84	22 April 2020, 19:30[514]	F9 B5 ♺ B1051.4	KSC, LC-39A	S... v1				Success	Success (drone ship)

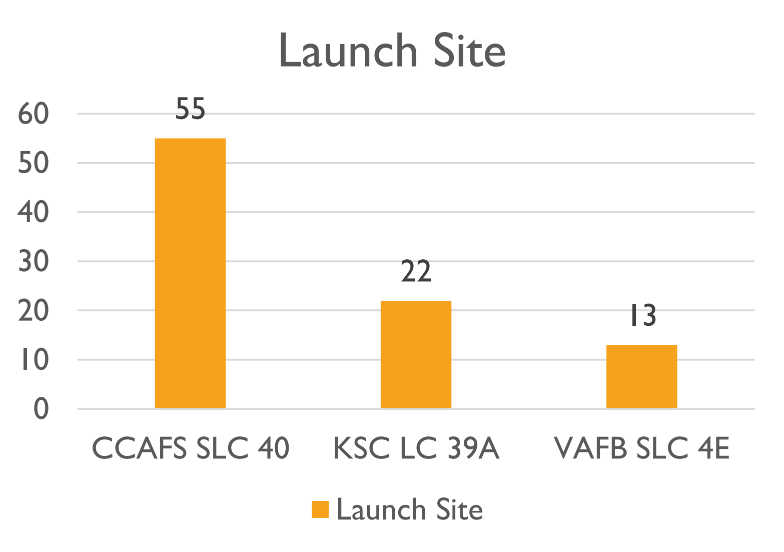
Converted web page data into a more useable format

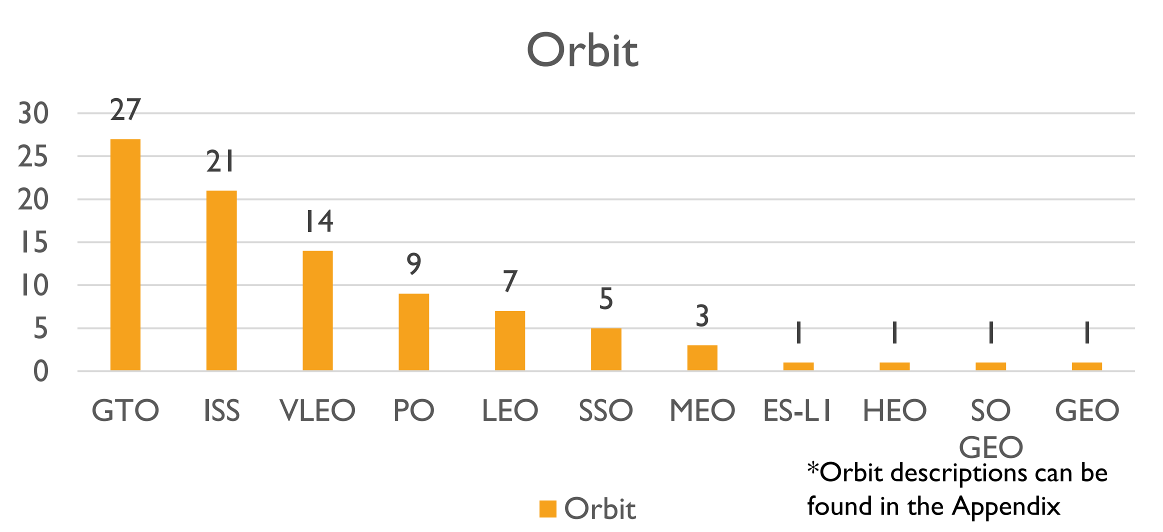
	Flight No.	Launch site	Payload	Payload mass	Orbit	Customer	Launch outcome	Version Booster	Booster landing	Date	Time
0	1	CCAFS	Dragon Spacecraft Qualification Unit	0	LEO	SpaceX	Success\n	F9 v1.0B0003.1	Failure	4 June 2010	18:45
1	2	CCAFS	Dragon	0	LEO	NASA (COTS)\nNRO	Success	F9 v1.0B0004.1	Failure	8 December 2010	15:43
2	3	CCAFS	Dragon	525 kg	LEO	NASA (COTS)	Success	F9 v1.0B0005.1	No attempt\n	22 May 2012	07:44
3	4	CCAFS	SpaceX CRS-1	4,700 kg	LEO	NASA (CRS)	Success\n	F9 v1.0B0006.1	No attempt	8 October 2012	00:35
4	5	CCAFS	SpaceX CRS-2	4,877 kg	LEO	NASA (CRS)	Success\n	F9 v1.0B0007.1	No attempt\n	1 March 2013	15:10
...	...	...	...	...	...	...	...	...	...	...	...
116	117	CCSFS	Starlink	15,600 kg	LEO	SpaceX	Success\n	F9 B5B1051.10	Success	9 May 2021	06:42
117	118	KSC	Starlink	~14,000 kg	LEO	SpaceX Capella Space and Tyvak	Success\n	F9 B5B1058.8	Success	15 May 2021	22:56
118	119	CCSFS	Starlink	15,600 kg	LEO	SpaceX	Success\n	F9 B5B1063.2	Success	26 May 2021	18:59
119	120	KSC	SpaceX CRS-22	3,328 kg	LEO	NASA (CRS)	Success\n	F9 B5B1067.1	Success	3 June 2021	17:29
120	121	CCSFS	SXM-8	7,000 kg	GTO	Sirius XM	Success\n	F9 B5	Success	6 June 2021	04:26

121 rows × 11 columns

- Utilized data from SpaceX's Wikipedia page: https://en.wikipedia.org/w/index.php?title=List_of_Falcon_9_and_Falcon_Heavy_launches&oldid=1027686922
- BeautifulSoup used to create an object with the Wikipedia's html data
- Parsed the data and used Pandas to create a new dataframe of the Falcon 9 launch and recovery results

DATA WRANGLING





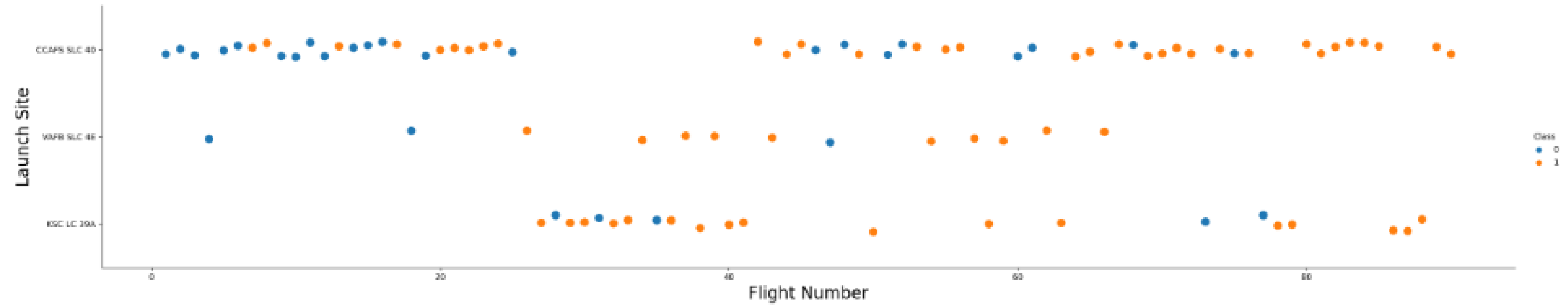
*Orbit descriptions can be found in the Appendix

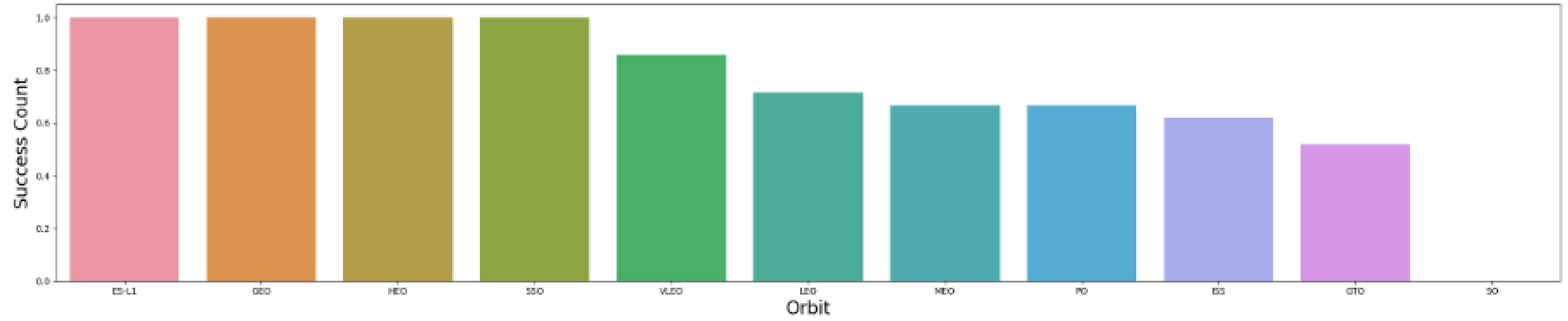
- After gathering data, exploratory data analysis could be initiated
- Using Pandas, the launch site count was determined with CCAFS SLC 40 having the most at 55 launches
- The orbit occurrence was also determined with GTO (geosynchronous) orbit having the highest amount at 27
- Landing outcomes were also converted from False/True (unsuccessful and successful) to numeric values of 0/1 respectively
- The success rate was found to be about 67%

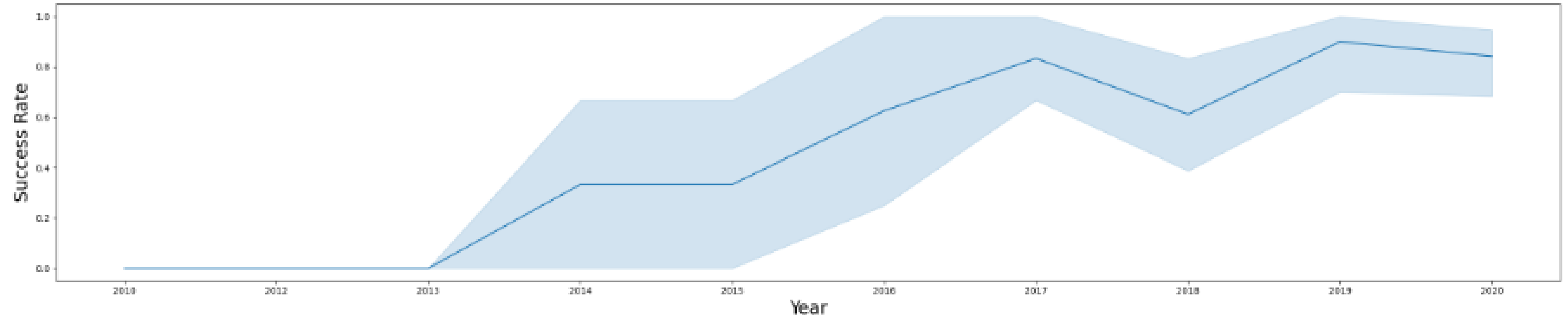
EXPLORATORY DATA ANALYSIS WITH SQL

- The objective of this portion was to store the data gathered in a SQL database
- Basic calculations were applied to show the functionality of SQL
- Some interesting outcomes include things like:
  - The first successful landing occurred on 6/5/16
  - 100 out of the 101 missions were deemed successful
  - More successful landings occurred on drone ships rather than ground pads, although the percentages were not determined

EXPLORATORY DATA ANALYSIS WITH VISUALIZATIONS

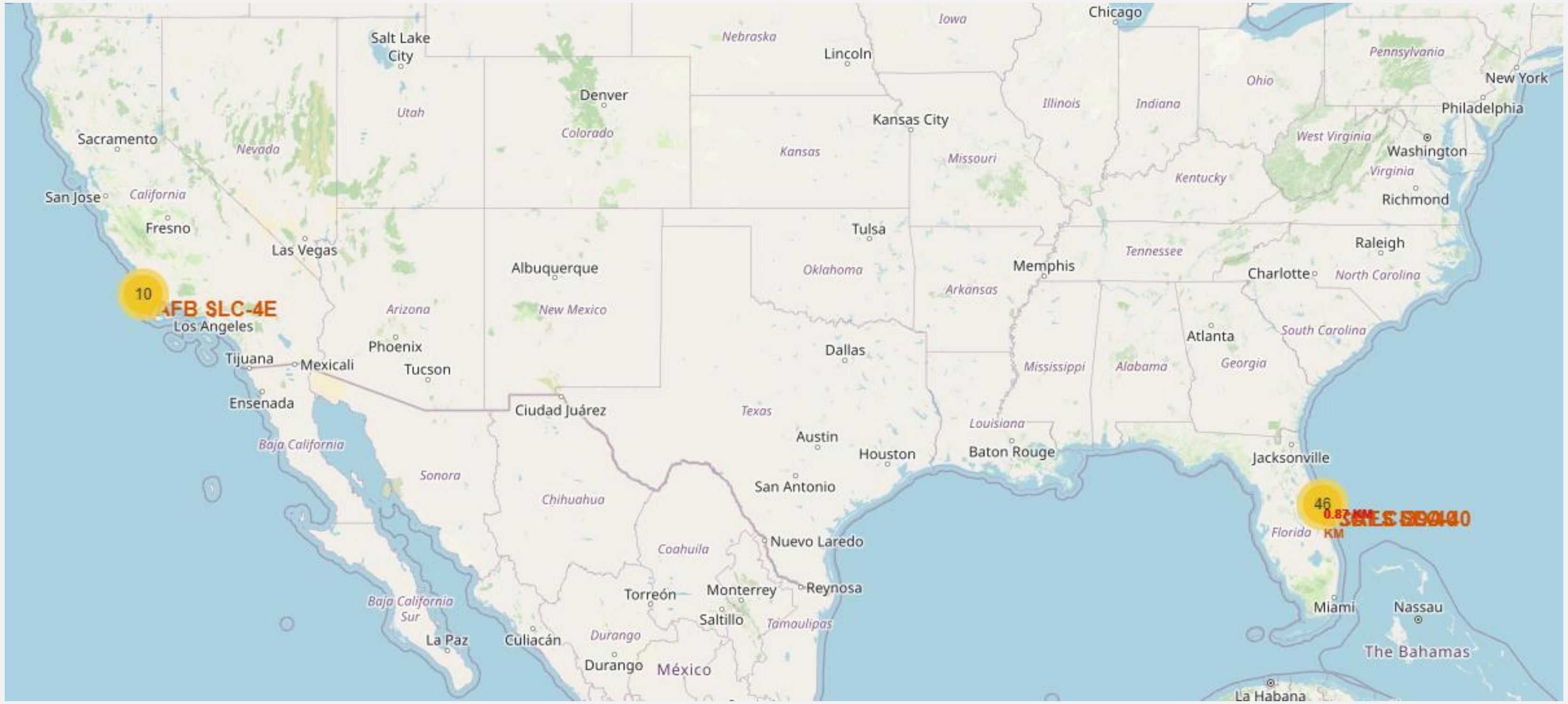






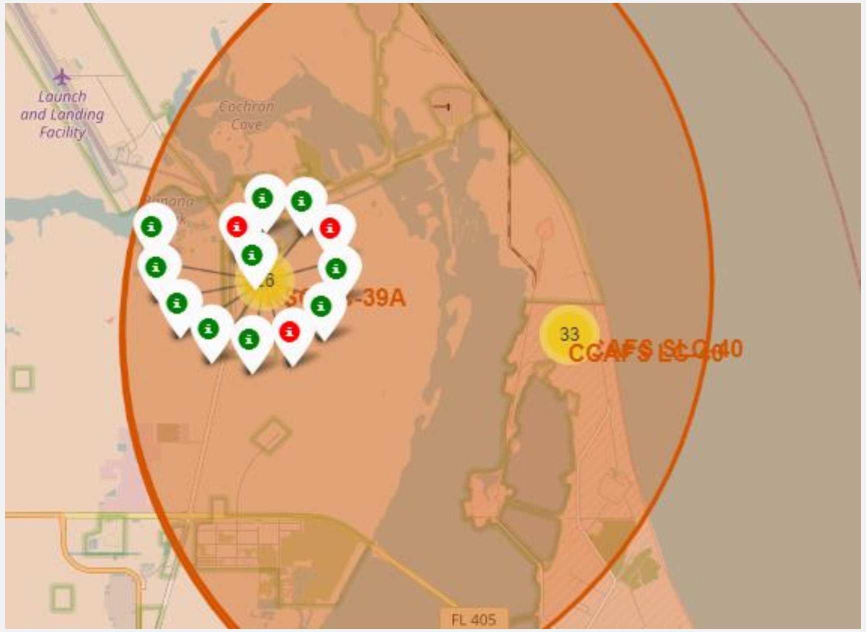
- Once the data was captured, additional visualizations were created to help describe the info
- In the top plot the number of successful landings (orange) increases with the total flight count
- The second bar chart shows the top four orbits (ES-L1, GEO, HEO, and SSO) for successful launches with 100% success rates
- SO orbit did not have any successful launches
- The final line chart again shows how success rate typically increased over time with 2018 being the exception

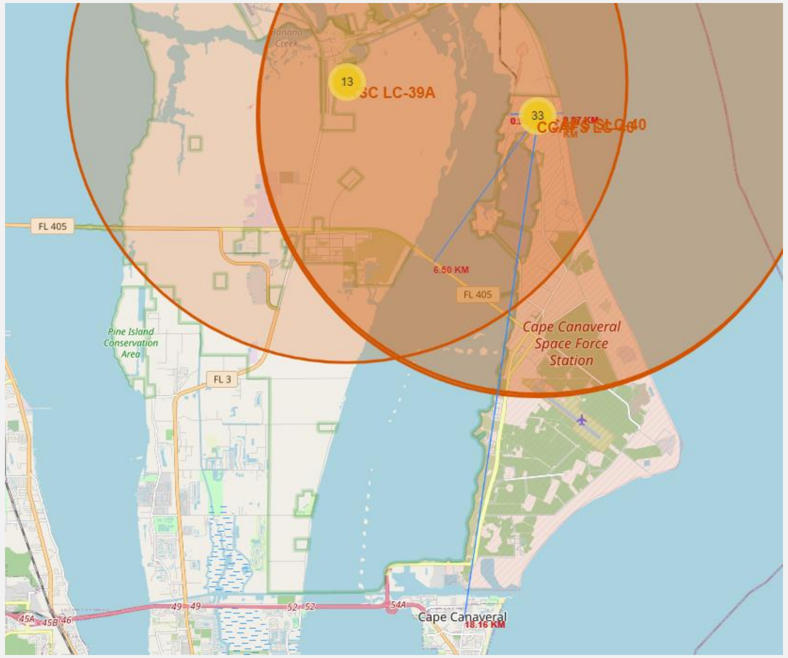
INTERACTIVE VISUAL ANALYTICS WITH FOLIUM



- Folium was then used as another way to view data interactively
- From the map, there are two primary launch points
  - Vandenberg Space Force Base
  - Cape Canaveral Space Force Station
- The map labels and identifies the number of launches and allows the user to zoom in and select each launch site to view additional information
- It can be noted that each site is located, within reason, near the equator and also near an ocean

INTERACTIVE VISUAL ANALYTICS WITH FOLIUM

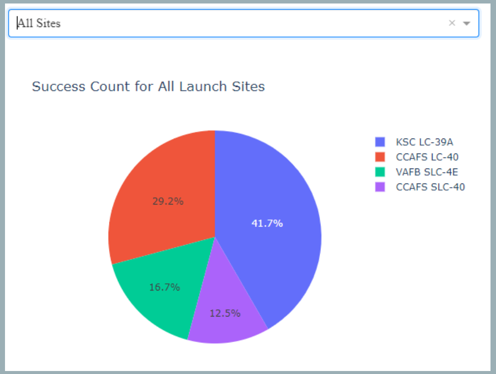


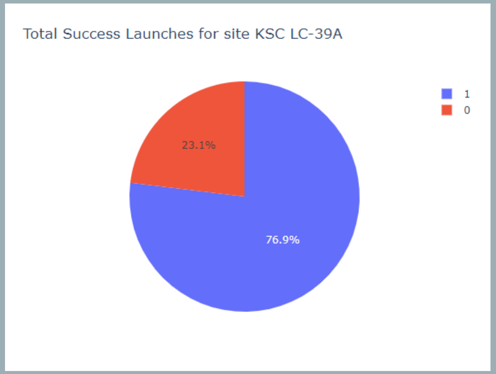


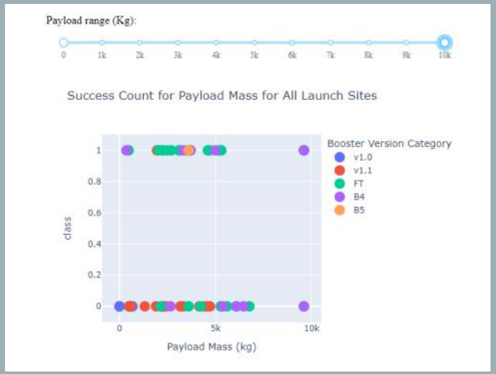
- Once a site is selected a marker cluster appears with a visual representation of the successful launches
- KSC LC-39A led the pack with a success rate of about 77%
- Interesting notes is the KSC LC-39A is the closest to roads and railways but furthest from the ocean

INTERACTIVE DASHBOARD WITH PLOTY DASH

- Ploty allowed for an interactive dashboard to quickly change visualizations
- The dropdown selects different launch sites and displays their success and failure rates in the pie chart and slider plot underneath
- KSC LC-39A could be quickly analyzed as the site with the highest percentage of success
- The slider plot allows for quick analysis of payload mass and success rate
- Typically the B5 booster version was the most successful

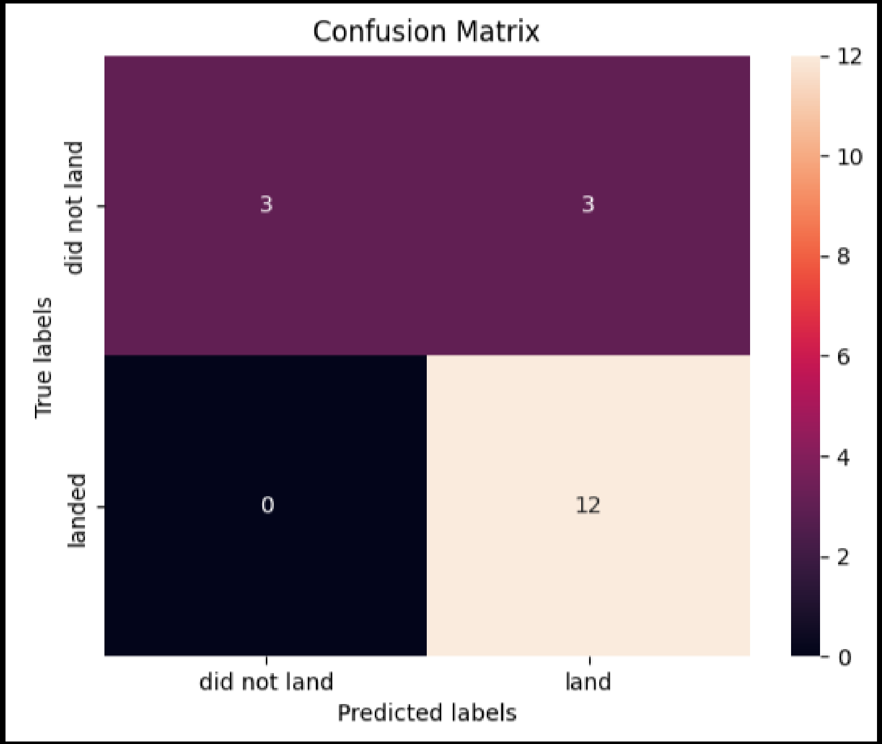






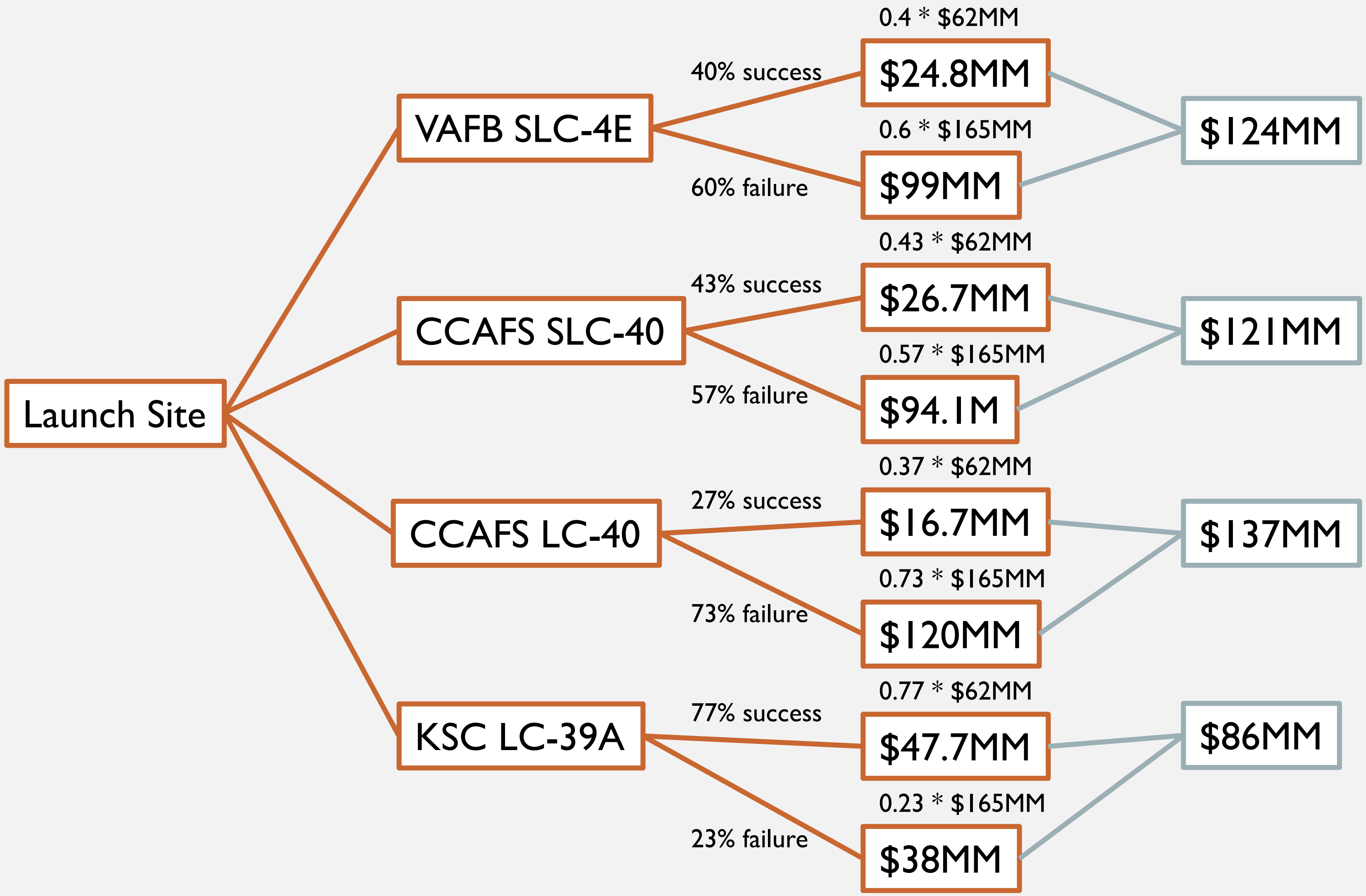
MACHINE LEARNING PREDICTION

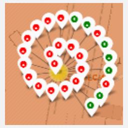
- Predictive analysis was used to determine the most accurate predictive method for launch results
- A sample size of 80% of the data was used to test against the other 20% to determine this accuracy
- The biggest error in prediction was a false positive where the prediction claimed the booster landed successfully but actually failed
- The most accurate prediction method was the decision tree at about 89% accuracy



Prediction Method	Accuracy
Decision Tree	89%
Logistic Regression	83%
Support Vector Machine	83%
K Nearest Neighbors	83%

DECISION TREE COST ESTIMATES





- This is a method of quantifying success rates based on launch costs
- Assuming a successful return would cost $62MM to launch and an unsuccessful return would require a new build of $165MM, KSC LC-39A is the cheapest costs on average
- It should be noted that CCAFS LC-40 likely had many of the initial launches so it's chance of success may be skewed; 7/9 most recent launches were successful

CONCLUSIONS

- Percentage of successful landings increased as the number of flights increased
- ES-L1, GEO, HEO, and SSO orbit launches were the most successful with 100% success rates
- KSC LC-39A launch site had the best success rate of about 77%
- Space Y would need to develop a rocket that can launch with an average cost of $86MM to compete with SpaceX, that or develop a reusable rocket booster

APPENDIX

ORBIT DESCRIPTIONS

- **LEO**: Low Earth orbit (LEO)is an Earth-centred orbit with an altitude of 2,000 km (1,200 mi) or less (approximately one-third of the radius of Earth),[1] or with at least 11.25 periods per day (an orbital period of 128 minutes or less) and an eccentricity less than 0.25.[2] Most of the manmade objects in outer space are in LEO [1].
- **VLEO**: Very Low Earth Orbits (VLEO) can be defined as the orbits with a mean altitude below 450 km. Operating in these orbits can provide a number of benefits to Earth observation spacecraft as the spacecraft operates closer to the observation[2].
- **GTO** A geosynchronous orbit is a high Earth orbit that allows satellites to match Earth's rotation. Located at 22,236 miles (35,786 kilometers) above Earth's equator, this position is a valuable spot for monitoring weather, communications and surveillance. Because the satellite orbits at the same speed that the Earth is turning, the satellite seems to stay in place over a single longitude, though it may drift north to south," NASA wrote on its Earth Observatory website [3] .
- **SSO (or SO)**: It is a Sun-synchronous orbit also called a heliosynchronous orbit is a nearly polar orbit around a planet, in which the satellite passes over any given point of the planet's surface at the same local mean solar time [4] .
- **ES-L1** :At the Lagrange points the gravitational forces of the two large bodies cancel out in such a way that a small object placed in orbit there is in equilibrium relative to the center of mass of the large bodies. L1 is one such point between the sun and the earth [5] .
- **HEO** A highly elliptical orbit, is an elliptic orbit with high eccentricity, usually referring to one around Earth [6].
- **ISS** A modular space station (habitable artificial satellite) in low Earth orbit. It is a multinational collaborative project between five participating space agencies: NASA (United States), Roscosmos (Russia), JAXA (Japan), ESA (Europe), and CSA (Canada) [7]
- **MEO** Geocentric orbits ranging in altitude from 2,000 km (1,200 mi) to just below geosynchronous orbit at 35,786 kilometers (22,236 mi). Also known as an intermediate circular orbit. These are "most commonly at 20,200 kilometers (12,600 mi), or 20,650 kilometers (12,830 mi), with an orbital period of 12 hours [8]
- **HEO** Geocentric orbits above the altitude of geosynchronous orbit (35,786 km or 22,236 mi) [9]
- **GEO** It is a circular geosynchronous orbit 35,786 kilometres (22,236 miles) above Earth's equator and following the direction of Earth's rotation [10]
- **PO** It is one type of satellites in which a satellite passes above or nearly above both poles of the body being orbited (usually a planet such as the Earth [11]

LINK TO GITHUB REPOSITORY

https://github.com/temazal/Capstone