IBM Developer
SKILLS NETWORK

Winning Space Race with Data Science

Taoufik ben lhosni
10-08-24

Outline

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

Executive Summary

- Summary of methodologies
  - Data Collection
  - EDA
  - Map with Folium
  - Plotly
  - Predictive analysis

Introduction

- Project background and context

  The goal of this project is to predict whether the Falcon 9 first stage will successfully land after a launch. SpaceX reports that a Falcon 9 launch costs $62 million, significantly less than other providers who charge upwards of $165 million per launch. The cost difference is largely due to SpaceX's ability to reuse the first stage of the rocket. By predicting the likelihood of a successful landing, we can better understand the true cost of a launch, which is valuable information for companies looking to compete with SpaceX in the rocket launch market.

- Problems to Address
  - Key Characteristics: What are the primary factors that distinguish successful landings from failed ones?
  - Impact of Rocket Variables: How does each rocket variable influence the likelihood of a successful or failed landing?
  - Optimal Conditions: What conditions must be met for SpaceX to achieve the highest possible landing success rate?

Section 1

Methodology

Executive Summary

Data Collection

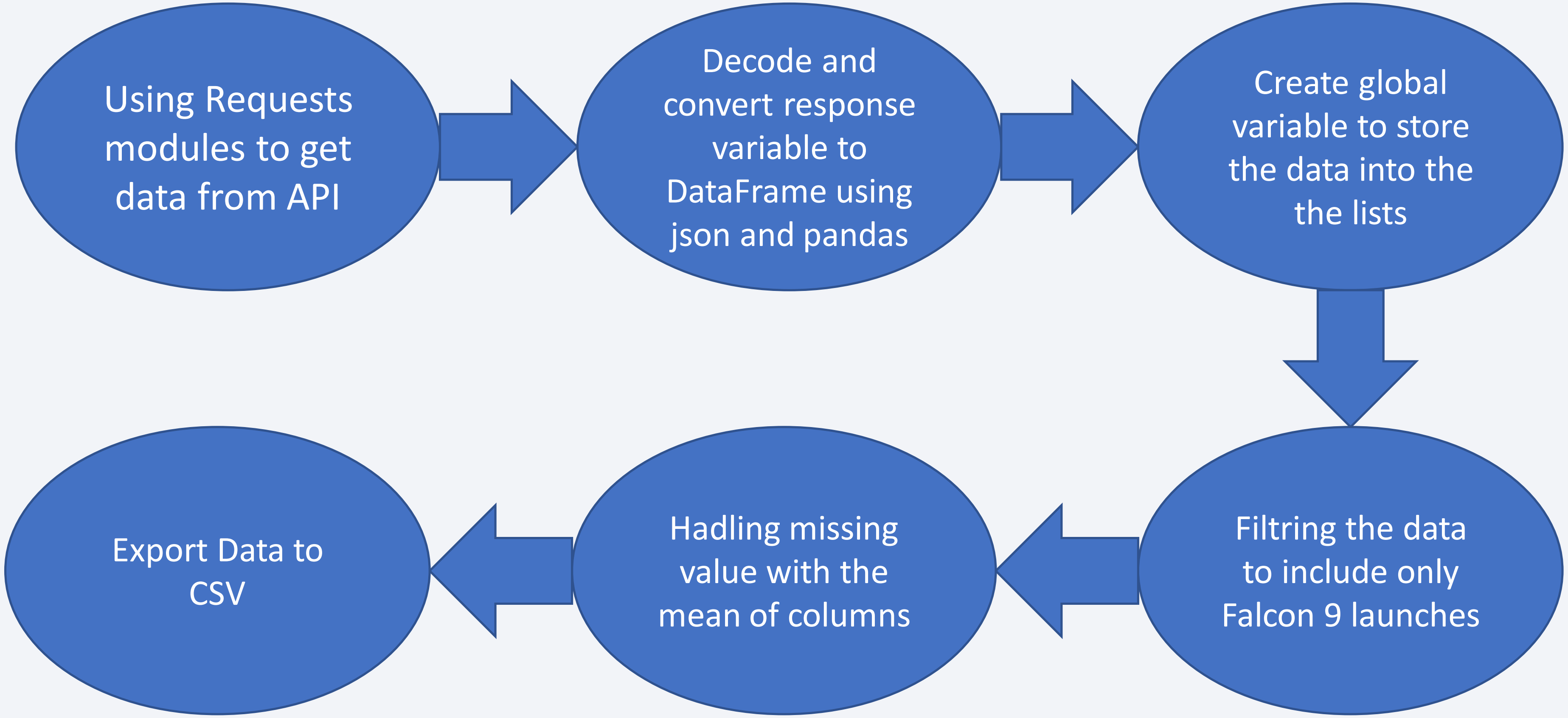
Data Wrangling

exploratory data analysis

Foluim and Plotly

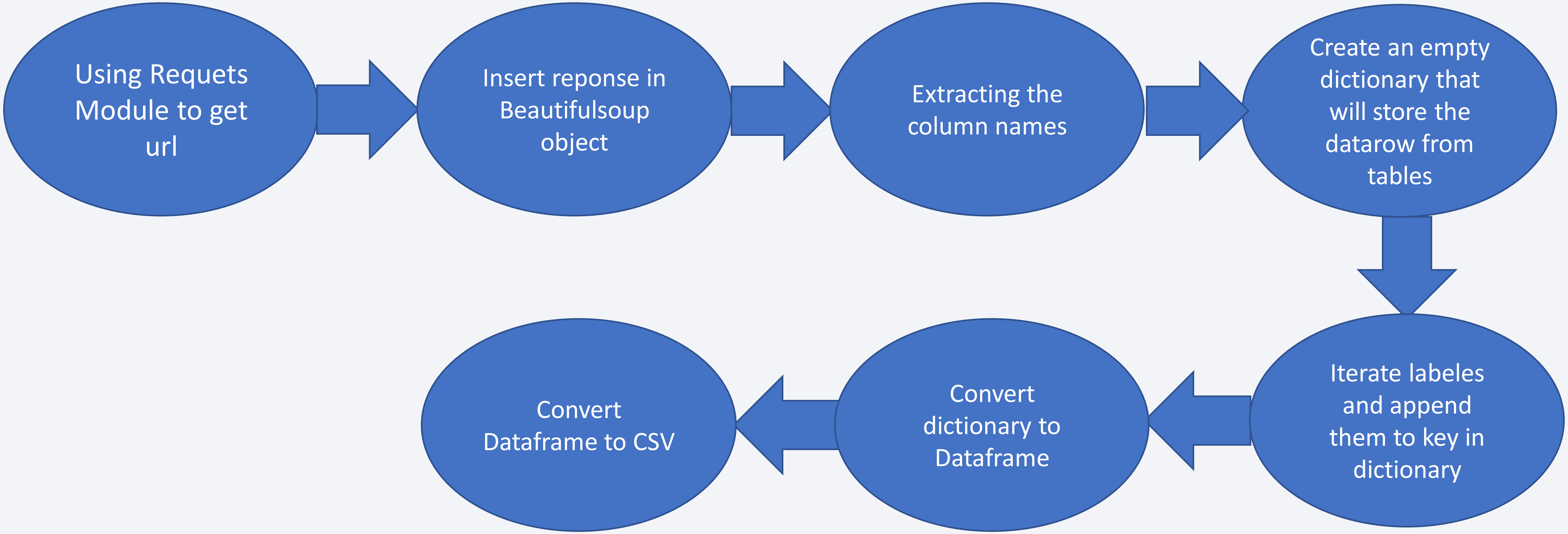
Predictive analysis

Data Collection – SpaceX API



- Source code in Github

Data Collection - Scraping



- [Source code in Github]()

Data Wrangling

Determine the number of launches for each

Calculate the count and frequency of each orbit type.

Analyze the count and frequency of mission outcomes based on orbit type.

Generate outcome labels using the data from the "outcome" column.

- [Source code in Github]

EDA with Data Visualization

- Summarize what charts were plotted and why you used those charts:
  - Scatter Plot: Visualizes the correction between :
    - ❑Between Payload and Launch Site
    - ❑Between Flight Number and Launch Site
    - ❑Between FlightNumber and PayloadMass
  - Bar chart: represents the magnitude of the Orbite type related to Success Rate
  - Line Plot: compare Success Rate between 2010 and 2020

- [Source code in Github]()

EDA with SQL

- Summarization the SQL queries you performed:
  - Display the names of the unique launch sites in the space mission
  - Display the total payload mass carried by boosters launched by NASA (CRS)
  - Display average payload mass carried by booster version F9 v1.1
  - List the date when the first succesful landing outcome in ground pad was acheived.
  - List the names of the boosters which have success in drone ship and have payload mass greater than 4000 but less than 6000
  - List the names of the booster_versions which have carried the maximum payload mass. Use a subquery
  - List the failed landing_outcomes in drone ship, their booster versions, and launch site names for the in year 2015
  - Rank the count of landing outcomes (such as Failure (drone ship) or Success (ground pad)) between the date 2010-06-04 and 2017-03-20, in descending order

- [Source code in Github]()

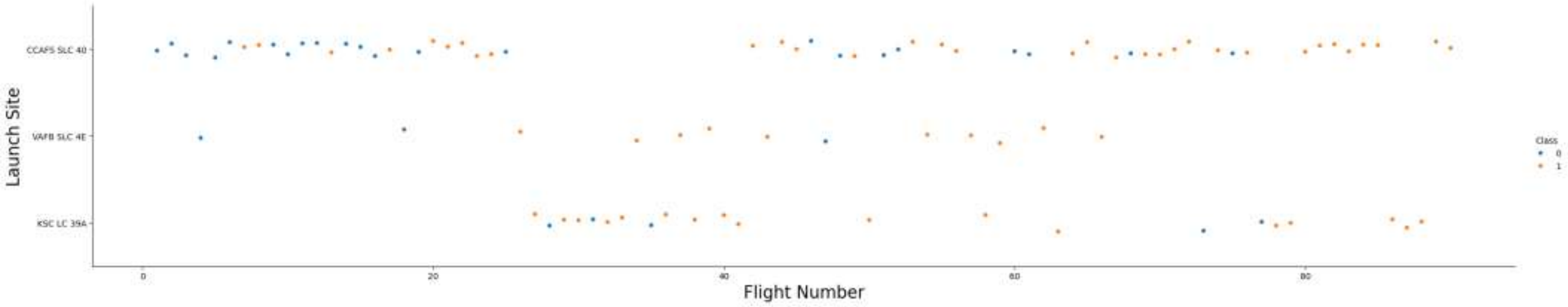
Build an Interactive Map with Folium

- Create a folium Map object
- Add folium.Circle and folium.Marker for each launch site on the site map
- Mark the success/failed launches for each site on the map
- New column in launch_sites dataframe called marker_color to store the marker colors based on the class value
- Calculate the distances between a launch site to its proximities
- [Source code in Github]()

Section 2

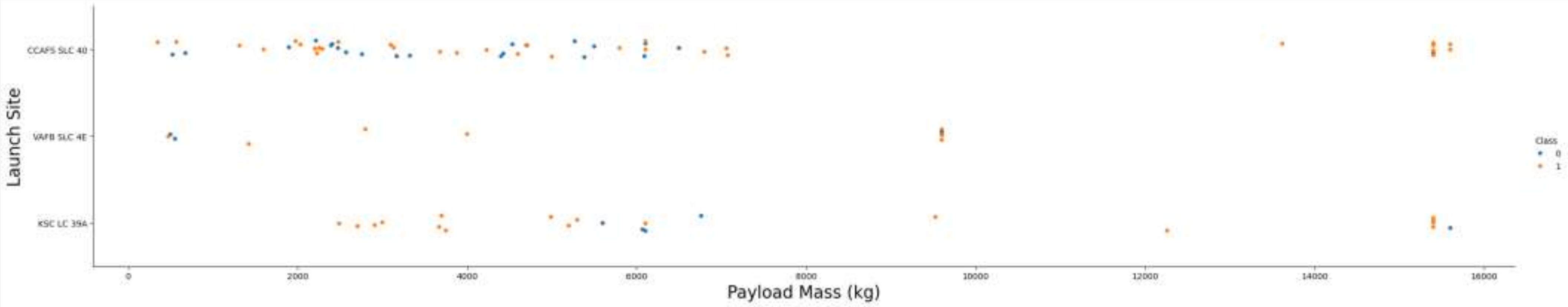
Insights drawn from EDA

Flight Number vs. Launch Site



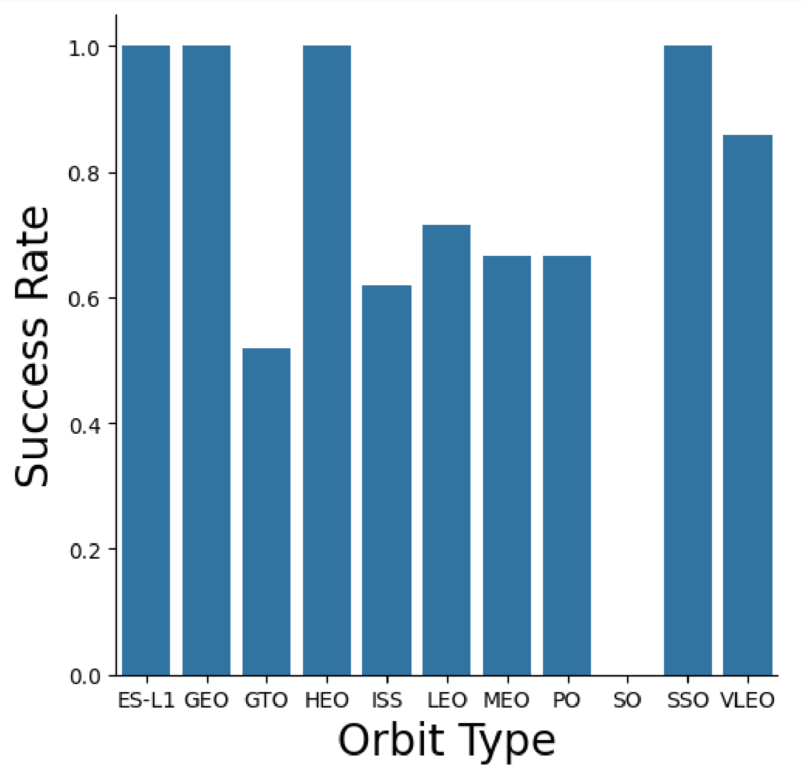
- More than half flight are for CCAFS SLC 40 launch site
- The range of 80 flights and above is achieved successfully

Payload vs. Launch Site



- Almost payload mass between 2000 kg and 7000 kg
- In KSC LC 39A launch site, almost payload masses between 2000 kg and 7000 kg are achieved successfully

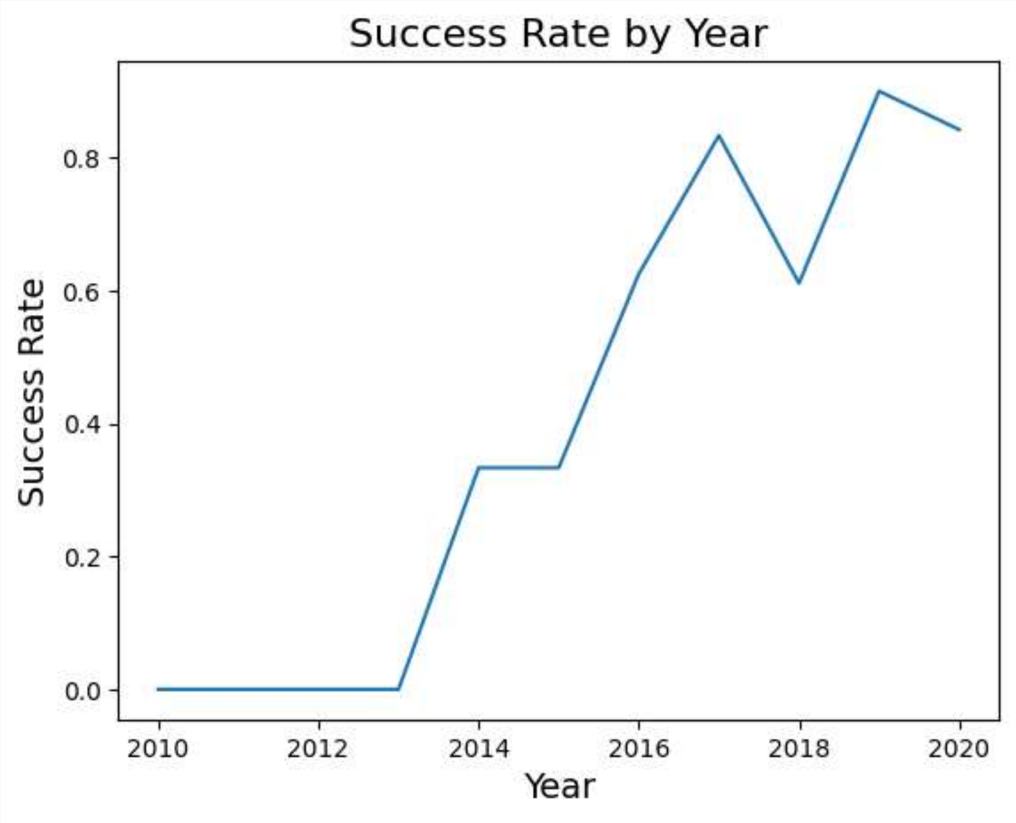
Success Rate vs. Orbit Type



- ES-L1, GTO, HEO, SSO orbit types are completely successed
- SO orbit type is failed

Launch Success Yearly Trend

- The increase of success Rate started from 2013
- The trend still increased until 2020



All Launch Site Names

Display the names of the unique launch sites in the space mission

In [4]:
```
%sql select distinct launch_site from SPACEXDATASET;
```

```
 * ibm_db_sa://wzf08322:***@0c77d6f2-5da9-48a9-81f8-86b520b87518.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:31198/bludb
Done.
```

Out[4]:

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

Launch Site Names Begin with 'CCA'

Task 2

Display 5 records where launch sites begin with the string 'CCA'

In [5]:
```
%sql select * from SPACEXDATASET where launch_site like 'CCA%' limit 5;
```

```
* ibm_db_sa://wzf08322:***@0c77d6f2-5da9-48a9-81f8-86b520b87518.bs2io90108kqb1od8lcg.databases.appdomain.cloud:31198/bludb
Done.
```

Out[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

Task 3

Display the total payload mass carried by boosters launched by NASA (CRS)

In [6]:
```
%sql select sum(payload_mass__kg_) as total_payload_mass from SPACEXDATASET where customer = 'NASA (CRS)';
```

```
 * ibm_db_sa://wzf08322:***@0c77d6f2-5da9-48a9-81f8-86b520b87518.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:31198/bludb
Done.
```

Out[6]:

total_payload_mass
45596

Average Payload Mass by F9 v1.1

Display average payload mass carried by booster version F9 v1.1

In [7]:
```
%sql select avg(payload_mass__kg_) as average_payload_mass from SPACEXDATASET where booster_version like '%F9 v1.1%';
```

```
 * ibm_db_sa://wzf08322:***@0c77d6f2-5da9-48a9-81f8-86b520b87518.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:31198/bludb
Done.
```

Out[7]:

average_payload_mass
2534

First Successful Ground Landing Date

List the date when the first succesful landing outcome in ground pad was acheived.

Hint:Use min function

In [8]:
```
%sql select min(date) as first_successful_landing from SPACEXDATASET where landing__outcome = 'Success (ground pad)';
```

```
 * ibm_db_sa://wzf08322:***@0c77d6f2-5da9-48a9-81f8-86b520b87518.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:31198/bludb
Done.
```

Out[8]:

first_successful_landing
2015-12-22

Successful Drone Ship Landing with Payload between 4000 and 6000

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

In [9]:
```
%sql select booster_version from SPACEXDATASET where landing__outcome = 'Success (drone ship)' and payload_mass__kg_ betweer
```

```
 * ibm_db_sa://wzf08322:***@0c77d6f2-5da9-48a9-81f8-86b520b87518.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:31198/bludb
Done.
```

Out[9]:

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

Total Number of Successful and Failure Mission Outcomes

List the total number of successful and failure mission outcomes

In [10]:

```
%sql select mission_outcome, count(*) as total_number from SPACEXDATASET group by mission_outcome;
```

```
 * ibm_db_sa://wzf08322:***@0c77d6f2-5da9-48a9-81f8-86b520b87518.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:31198/bludb
Done.
```

Out[10]:

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

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

Rank the count of landing outcomes (such as Failure (drone ship) or Success (ground pad)) between the date 2010-06-04 and 2017-03-20, in descending order

In [13]:
```
%%sql select landing__outcome, count(*) as count_outcomes from SPACEXDATASET
      where date between '2010-06-04' and '2017-03-20'
      group by landing__outcome
      order by count_outcomes desc;
```

```
 * ibm_db_sa://wzf08322:***@0c77d6f2-5da9-48a9-81f8-86b520b87518.bs2io90l08kqb1od8lcg.databases.appdomain.cloud:31198/bludb
Done.
```

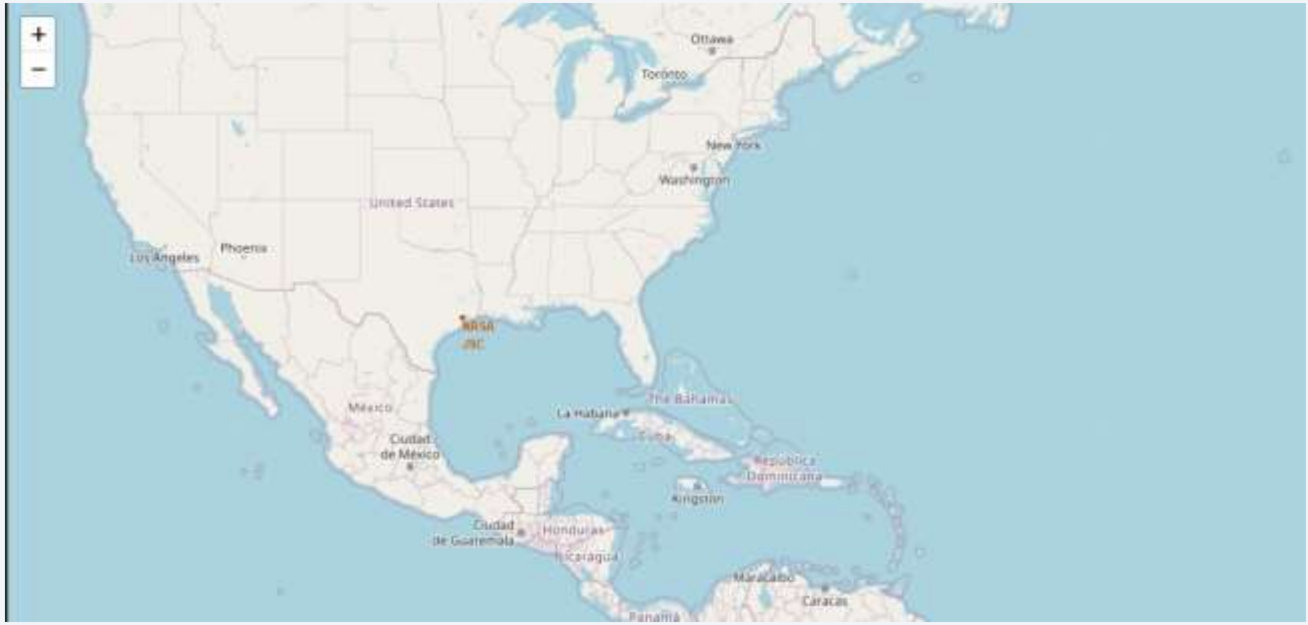
Out[13]:

landing__outcome	count_outcomes
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

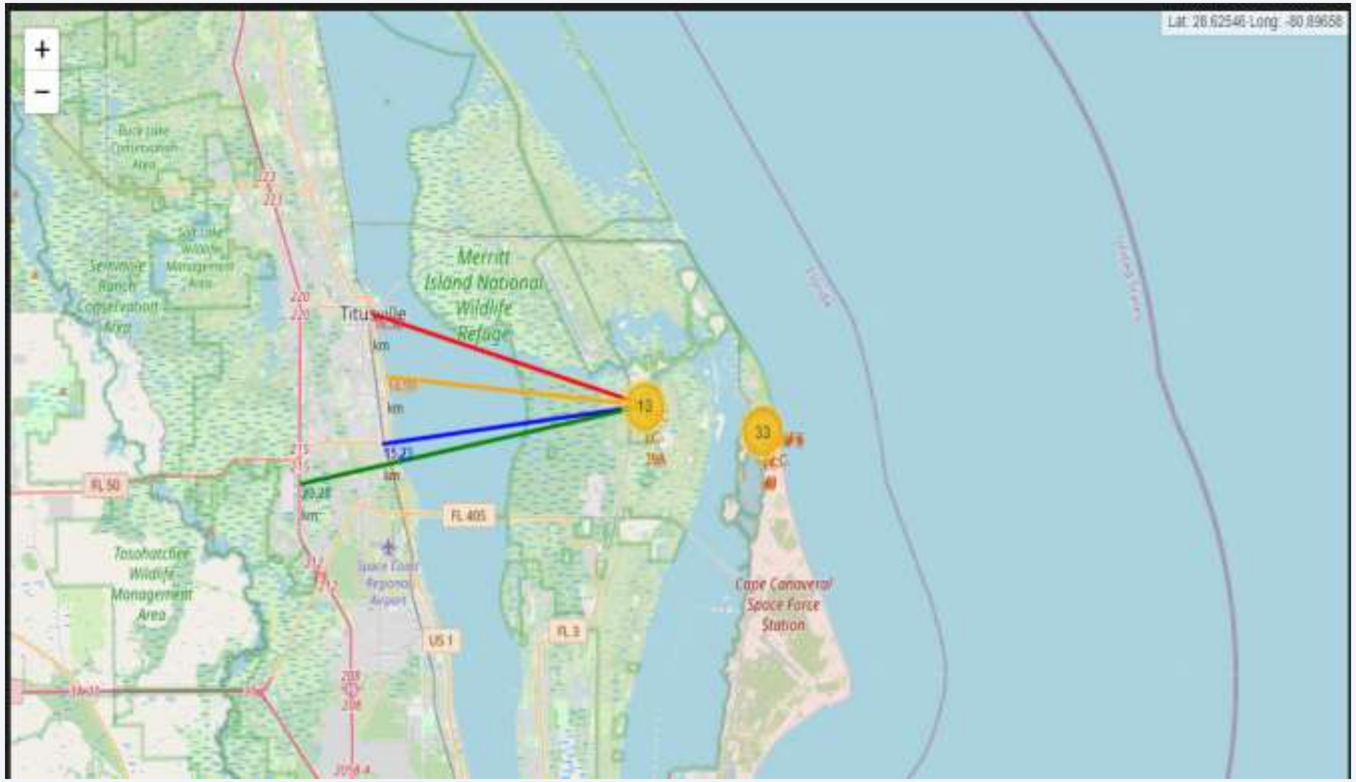
<Folium Map Screenshot 1>



Most launch sites are strategically located near the equator, where the Earth's rotation speed is fastest at 1,670 km/hour. Launching from the equator allows a spacecraft to carry this speed into space, thanks to inertia, which helps the spacecraft maintain the necessary velocity to stay in orbit.

Additionally, launch sites are often situated close to the coast. This coastal proximity allows rockets to be launched over the ocean, reducing the risk of debris falling or explosions occurring near populated areas.

<Folium Map Screenshot 2>



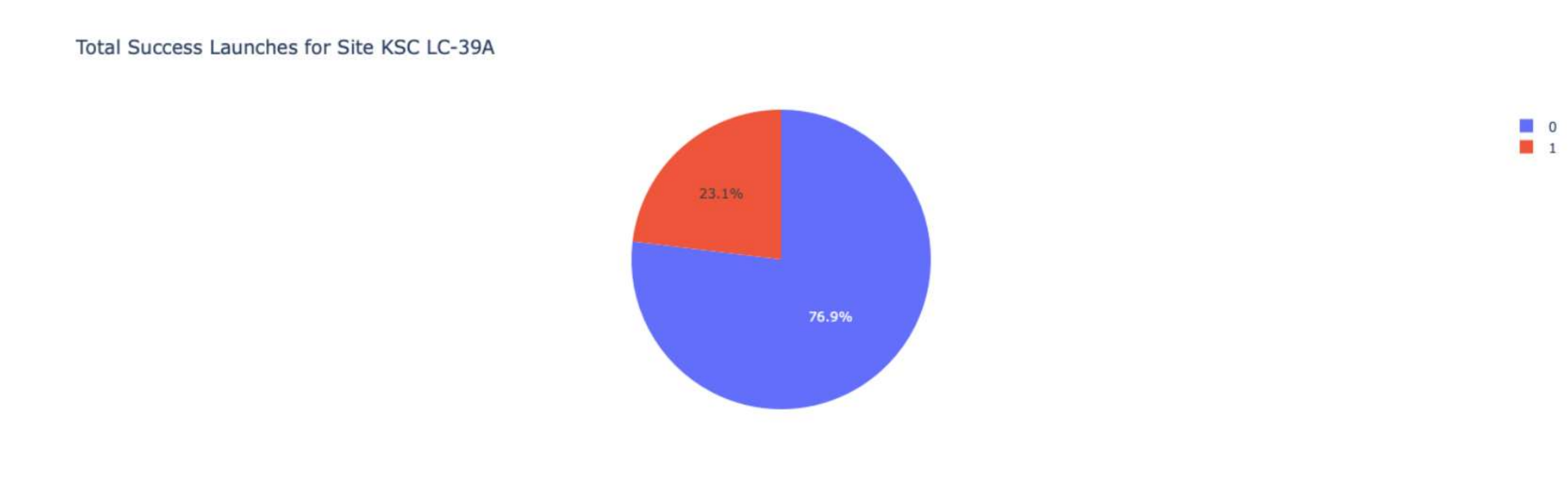
From the visual analysis of the launch site KSC LC-39A, it is evident that the site is:

- Relatively close to a railway, with a distance of 15.23 km.
- Relatively close to a highway, at 20.28 km.
- Relatively close to the coastline, at 14.99 km.

Section 4

Build a Dashboard with Plotly Dash

<Dashboard Screenshot 1>



- This pie-chart shows that 76.9% of launches from site KSC LC-39A were successefuls

Thank you!