# Falcon9 First-Stage Success

Melisa Ates

14-Sep-2020

# OUTLINE

- Executive Summary

- Introduction

- Methodology

- Results
  - EDA with Visualization
  - EDA with SQL
  - Interactive Visual Analysis with Folium
  - Interactive Visual Analysis with Ploty Dash
  - Predictive Analysis

- Conclusion

IBM **Dev**oper

SKILLS NETWORK

# EXECUTIVE SUMMARY

- Summary of Methodologies
  - Data Collection API
  - Data Collection w/ Web Scraping
  - Data Wrangling
  - Exploratory Data Analysis w/ SQL
  - Exploratory Data Analysis w/ Visualization
  - Interactive Visual Analytics w/ Folium
  - Interactive Visual Analytics w/ Ploty Dash

- Summary of Results
  - Exploratory Data Analysis
  - Interactive Visual Analytics
  - Predictive Analysis

# INTRODUCTION

- SpaceX advertises Falcon 9 rocket launches on its website with a cost of **62** million dollars. While other providers cost upwards of **165** million dollars each. This is due to SpaceX's capability of *reusing the first stage*.

- The **project goal** is to predict if the first stage will land successfully.
  - if we can determine whether the first stage will land or not, we can determine the **cost of a launch.**

# METHODOLOGY

1. Data Collection
2. Data Wrangling
3. Exploratory Data Analysis (EDA)
4. Interactive Visual Analysis
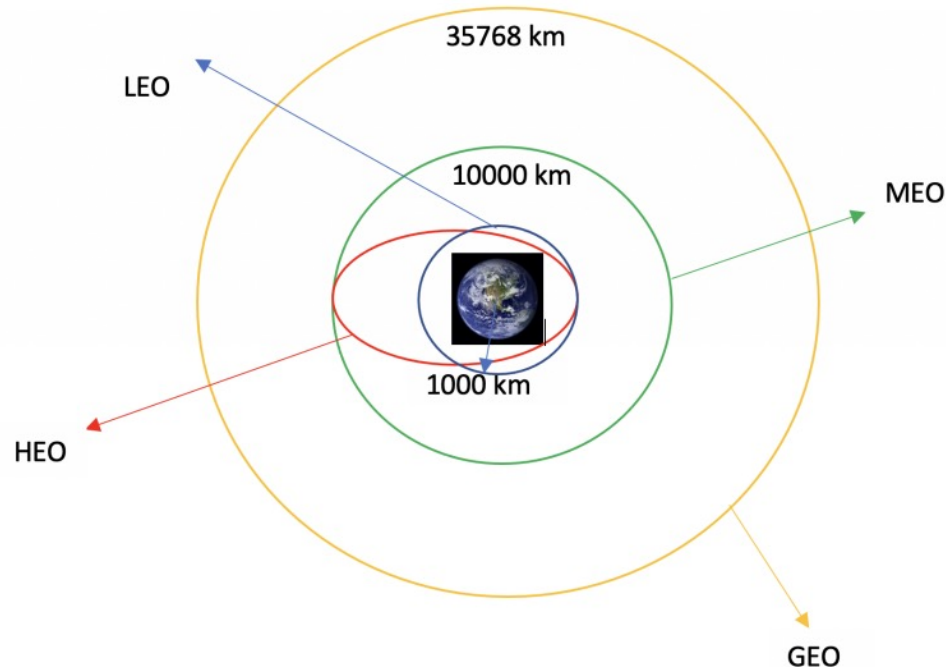
# DATA COLLECTION

- **SpaceX REST API**

The API gives us data about launches including:

➤ Info of the rocket used

➤ Payload delivered

➤ Launch Specifications

➤ Landing Outcome

- **Web Scraping**

➤ Wikipedia

➤ Using *BeautifulSoup*

# DATA WRANGLING



- Checked null values
- Calculated the number of launches on each site
- Calculated the number and occurrence of each orbit
- Calculated the number and occurrence of mission outcome per orbit type
- Created a landing outcome label from Outcome column

IBM Developer

SKILLS NETWORK

# EXPOLORATORY DATA ANALYSIS(EDA)

## EDA W/ VISUALIZATION

1. Visualize the relationship between Flight Number and Launch Site

2. Visualize the relationship between Payload and Launch Site

3. Visualize the relationship between success rate of each orbit type

4. Visualize the relationship between Payload and Orbit type

5. Visualize the launch success yearly trend

# EDA W/ VISUALIZATION

- **Features Engineering**

1. Create dummy variables to categorical columns

2. Cast all numeric columns to float64

| | FlightNumber | PayloadMass | Flights | GridFins | Reused | Legs | Block | ReusedCount | Orbit_ES-L1 | Orbit_GEO | ... | Serial_B1048 | Serial_B1049 | Serial_B1050 | Serial_B1051 | Serial_B1054 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 1.0 | 6104.959412 | 1.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | ... | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| **1** | 2.0 | 525.000000 | 1.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | ... | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| **2** | 3.0 | 677.000000 | 1.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | ... | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| **3** | 4.0 | 500.000000 | 1.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | ... | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| **4** | 5.0 | 3170.000000 | 1.0 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | ... | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

IBM Developer

SKILLS NETWORK

# EDA W/ SQL

- Displayed average payload mass carried by different booster versions

- Displayed the names of the booster versions which have carried the maximum payload mass

- Ranked the count of landing outcomes (such as Failure (drone ship) or Success (ground pad)) between specified dates

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

IBM Developer

SKILLS NETWORK

# INTERACTIVE VISUAL ANALYSIS
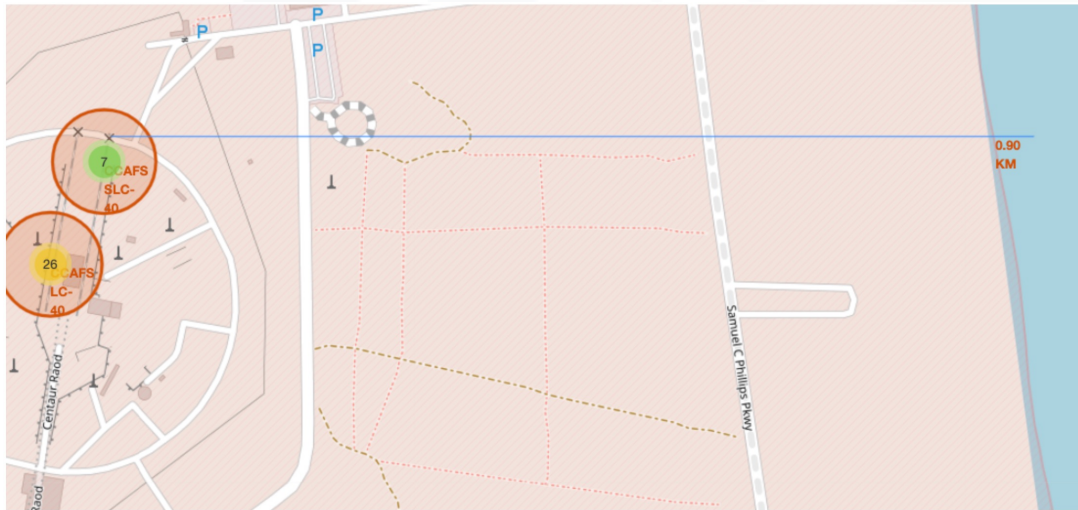


**FOLIUM**

1. Marked all launch sites on a map

2. Marked the success/failed launches for each site on the map

3. Calculated the distances between a launch site to its proximities

IBM Developer

SKILLS NETWORK

# INTERACTIVE VISUAL ANALYSIS

## PLOTY DASH

Built a Plotly Dash application for users to perform interactive visual analytics on SpaceX launch data in *real-time.*

➢The dashboard application contains input components such as a *dropdown list and a range slider to interact with a pie chart and a scatter point chart*.

# RESULTS

IBM Developer

SKILLS NETWORK
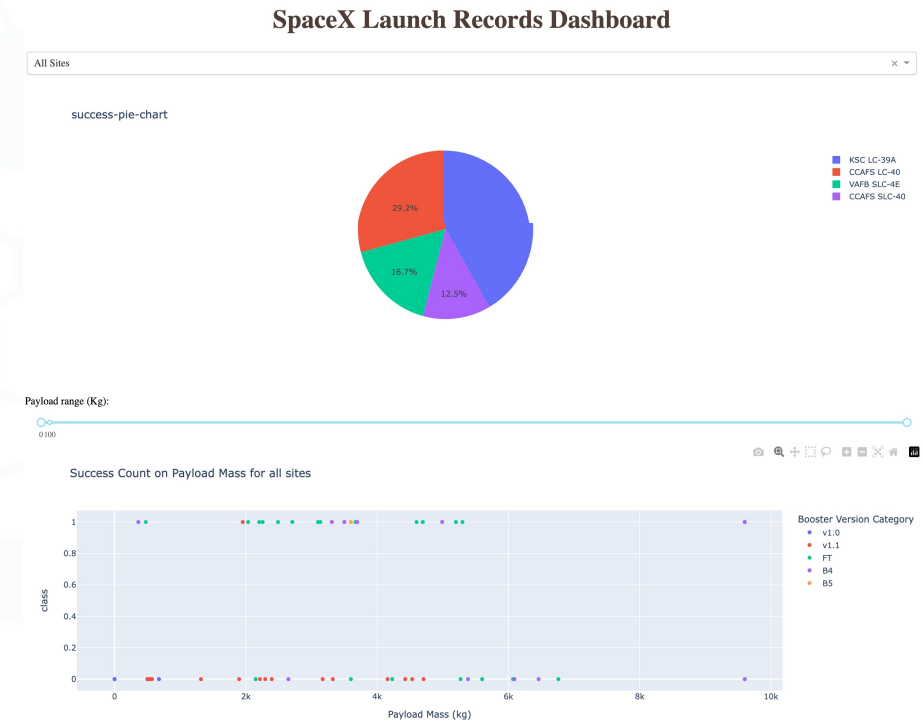
# Exploratory Data Analysis Results

## EDA w/ Visualization showed that:

- Launches from the site of CCAFS SLC 40 are significantly higher than other sites.

- The orbit types ES-L1, GEO, HEO, SSO are among the highest success rate.

- Launch success rate has increased significantly since 2013.

# Exploratory Data Analysis Results

## EDA w/ SQL showed that:

- The total PayLoad mass is 45596
- The average Payload mass by F9.v1.1 is 2928.400000
- First successful ground pad date is 2015-12-22
- The booster versions of Successful Drone ship landing w/ payload between 4000 and 6000 :
  - F9 FT B1022
  - F9 FT B1026
  - F9 FT B1021.2
  - F9 FT B1031.2

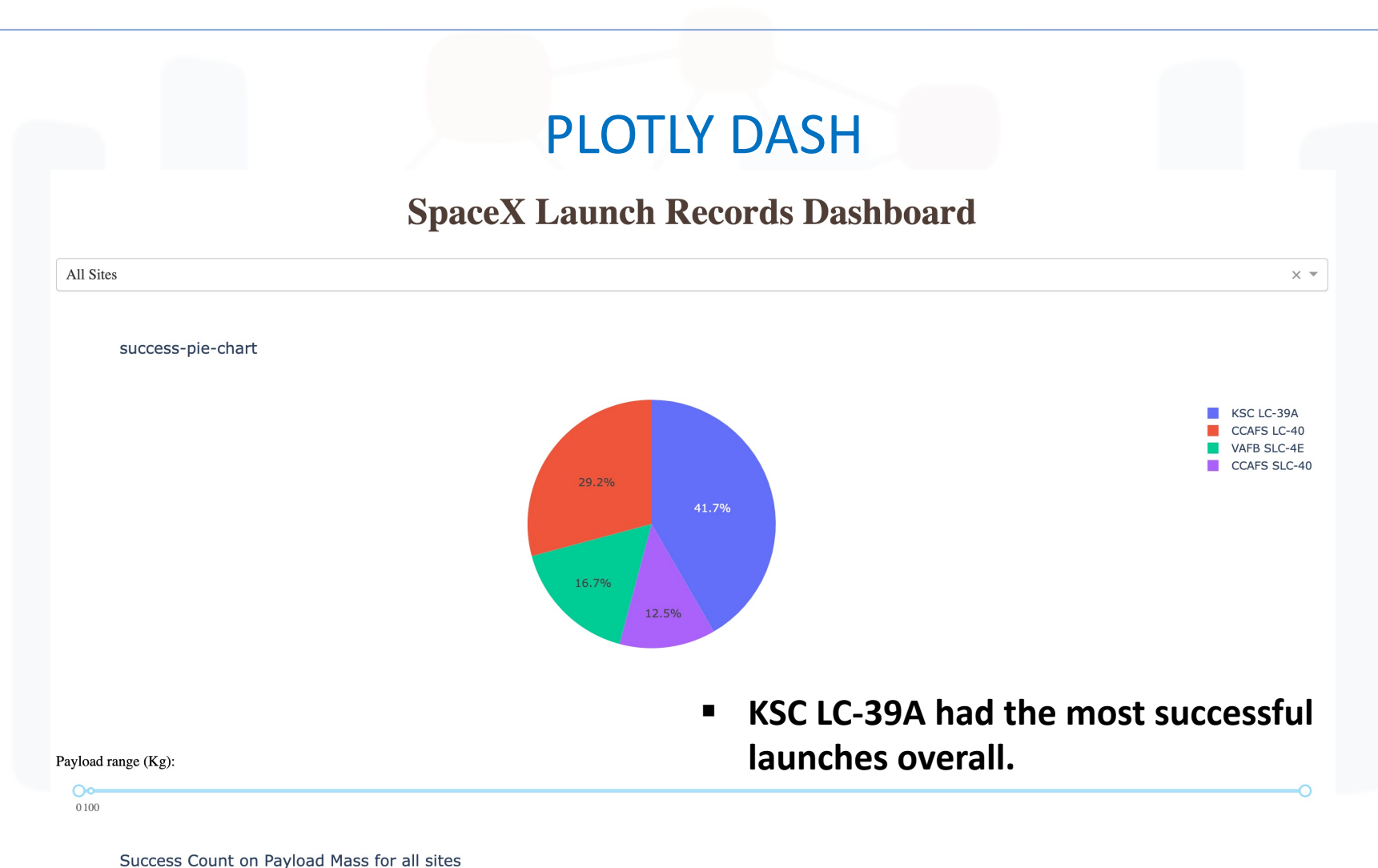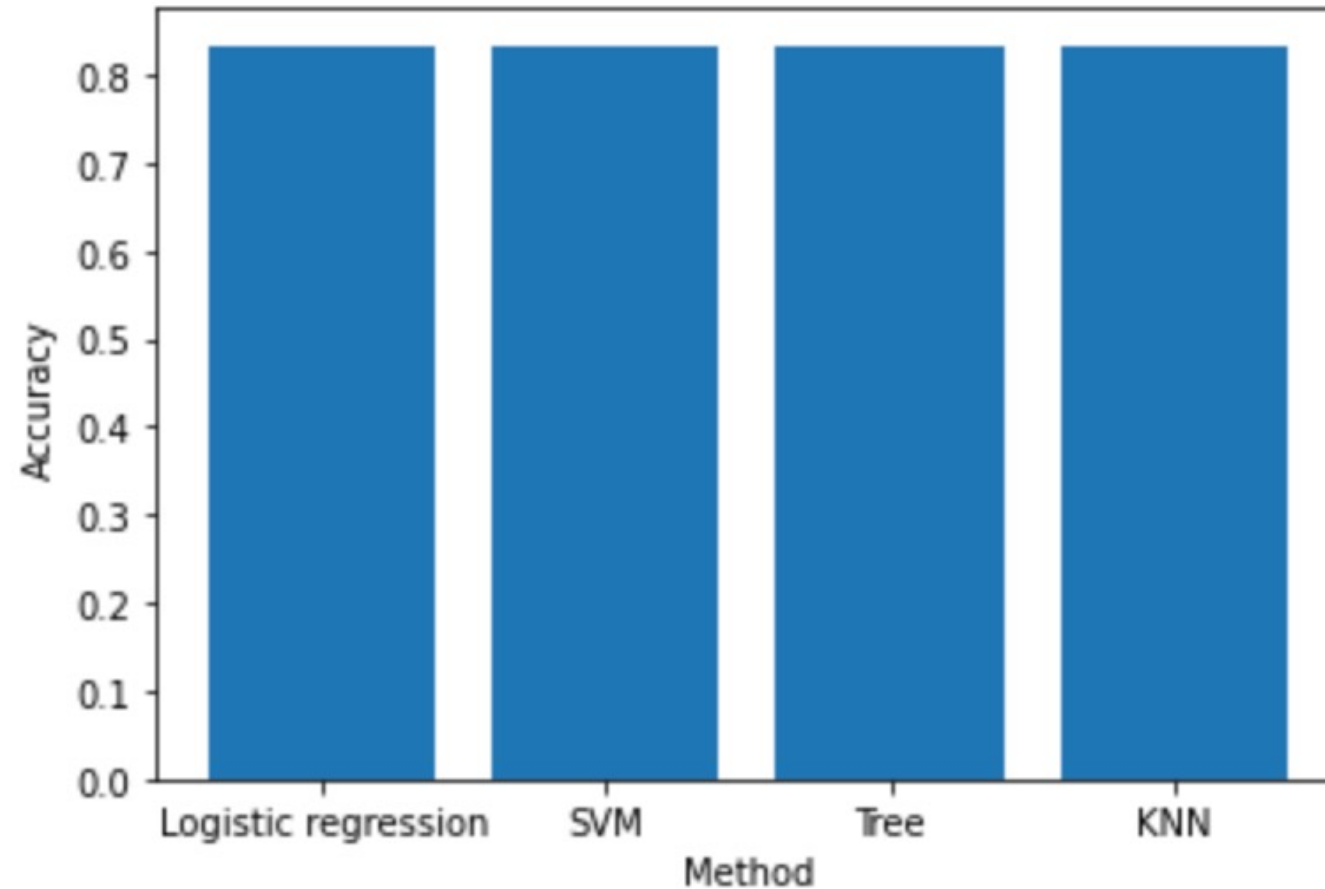IBM Developer

SKILLS NETWORK

# Interactive Visual Analytics Results

## FOLIUM
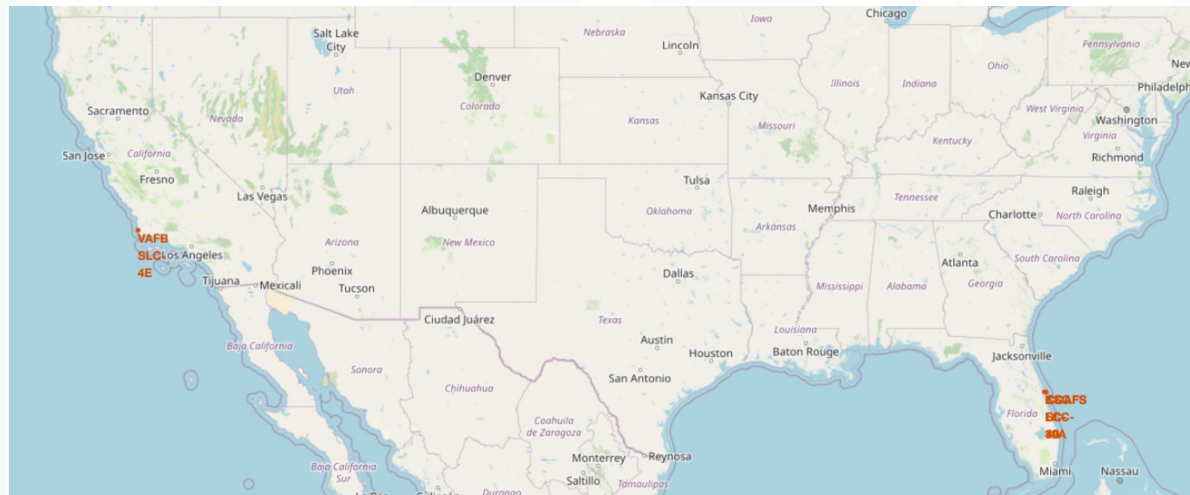


- all launch sites are in proximity to the Equator line
- all launch sites are in very close proximity to the coast

IBM Developer

SKILLS NETWORK

# Interactive Visual Analytics Results

## PLOTLY DASH

### SpaceX Launch Records Dashboard

All Sites

success-pie-chart



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

41.7%
29.2%
16.7%
12.5%

- **KSC LC-39A had the most successful launches overall.**

Payload range (Kg):

0 100

Success Count on Payload Mass for all sites

# Predictive Analysis Results



- **All models had the accuracy of 83.3% on the test data**

# CONCLUSION

All predictive models yielded to the same outcome in terms of the accuracy on test data (83.3%).

Launches from the site of CCAFS SLC 40 are significantly higher than other sites

The orbit types ES-L1, GEO, HEO, SSO are among the highest success rate.

KSC LC-39A had the most successful launches overall.

# Thank you!