SpaceY Landing Analysis

Ryan Desfosses

10 Oct 22

OUTLINE


Executive Summary
• (Start)
Introduction
Methodology
• Data Collection
• Visual Analytics
• Predictive Analysis
Results
• Visualization
• SQL
• Interactive Map
• Dashboard
• Predictive Analysis
Conclusion
Appendix
• (Finish)

EXECUTIVE SUMMARY

This presentation is a culmination of the material covered in IBM Data Science Professional Certification course in Coursera.

The object of the presentation is to come to a conclusion if the imaginary company SpaceY can predict the landing of the SpaceX's first stage rocket.

INTRODUCTION

- SpaceY needs to determine if the if a well know competitor will be able to land the first stage of a rocket.
- What factors determine if the rocket will successfully land?
- Are the factors reliable enough to predict if the rocket will land?

Data collection Methodology

Data Collection - METHODOLOGY

Falcon 9 and Falcon Heavy Launches Records where collected from Wikipedia and SpaceX API

Requests

Download HTML

Pandas

Structure and Manipulate Data

Beautifulsoup4

Parse HTML

IBM_DB_SA

interface to IBM Data Servers

EDA and Visual analytics Methodology

Visual Analytics - METHODOLOGY

Libraries Used:

Folium

Source: folium Docs

Seaborn

Source: cmdlinetips.com

- Perform Exploratory Data Analysis and Feature Engineering
- Compared
  - Flight Number vs. Launch Site
  - Payload vs. Launch Site
  - success rate vs. orbit type
  - Flight Number vs. Orbit type
  - Payload vs. Orbit type
- Feature Engineering

Predictive Analysis Methodology

Supervised Machine Learning Methods

Source: blog.capterra.com

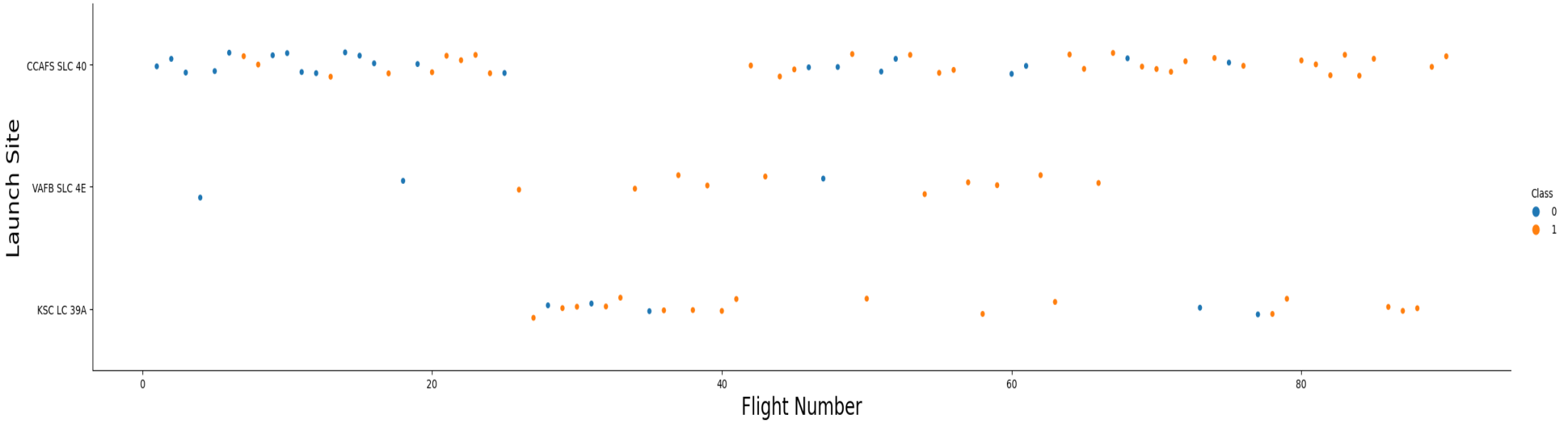
- Categorize launch outcomes as successful or failed
- Train models using multiple methods to predict the outcome
- Supervised Machine Learning Methods used:
  - K nearest neighbors (KNN)
  - Support Vector Machines (SVMs)
  - Decision Trees (DTs)
  - Logistic Regression

Predictive Analysis Results

Visualization Results

Flight Number and Launch Site

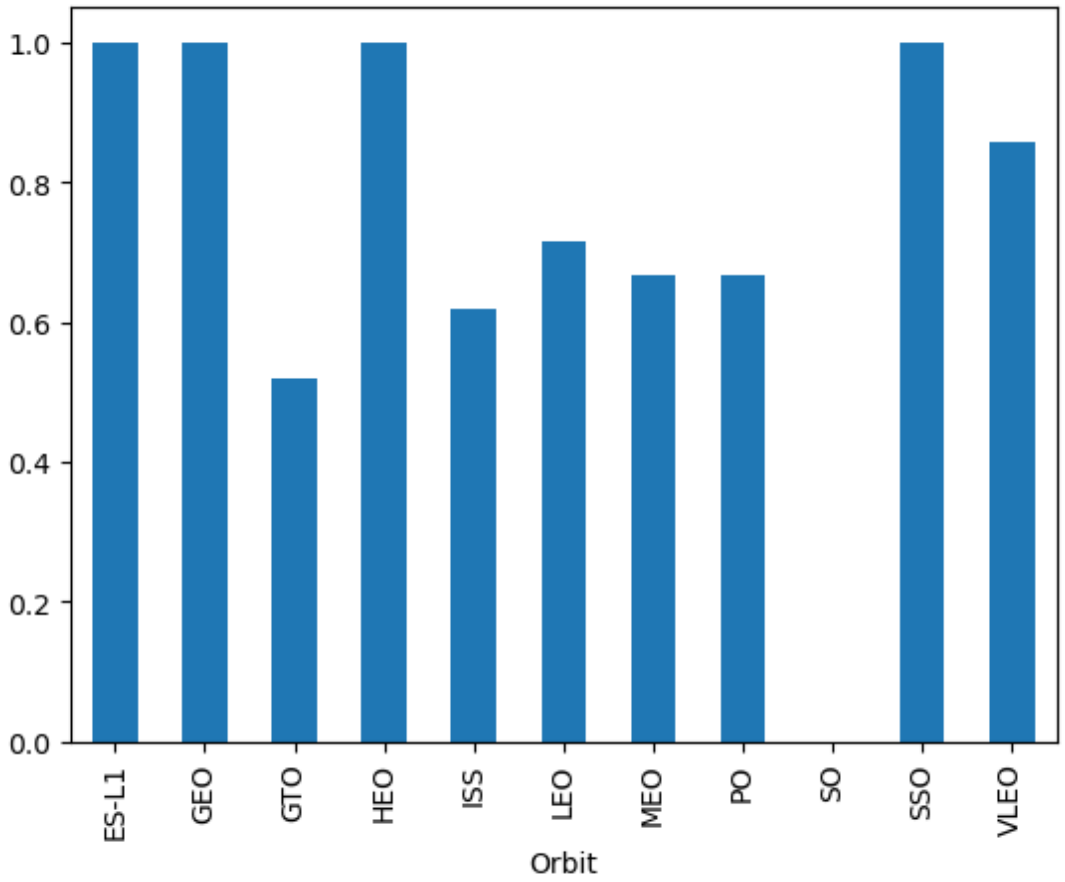
The blue dots represent a failed landing are more concentrated in the earlier flight numbers. CCAFS SLC 40 launch site is also utilized more often than both VAFB SLC 4E and KSC LC 39A.



Visualization Results

- the more massive the payload, the less likely the first stage will return.
- We see that different launch sites have different success rates. CCAFS LC-40, has a success rate of 60 %, while KSC LC-39A and VAFB SLC 4E has a success rate of 77%.
- VAFB-SLC launch site there are no rockets launched for heavy payload mass(greater than 10000). (TASK 2)

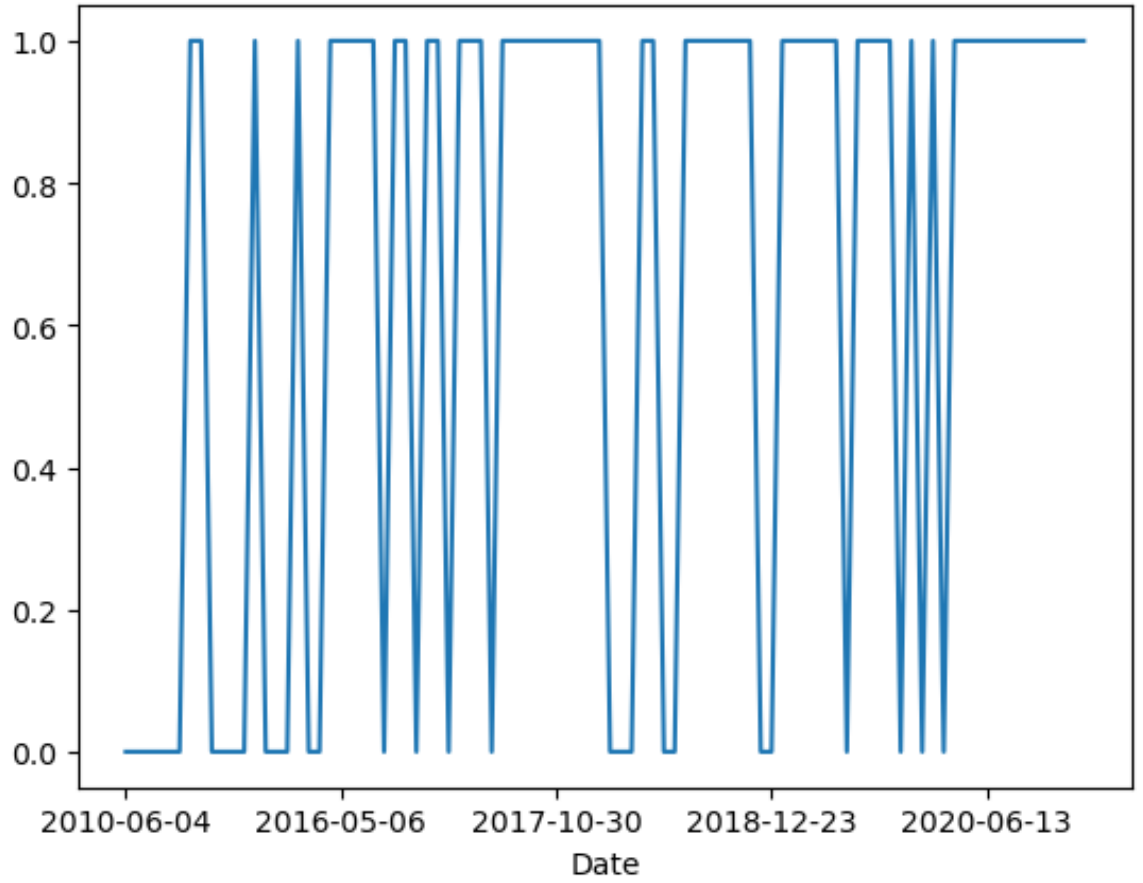
Orbits Success Rates



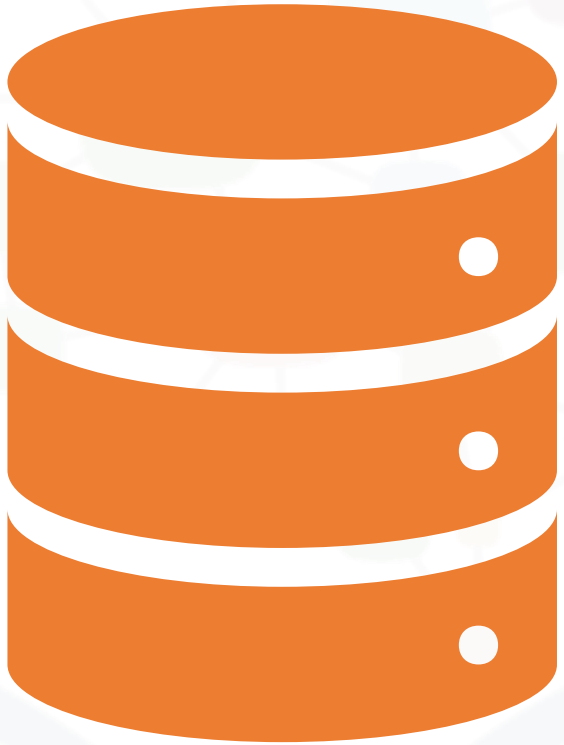
- ES-L1, GEO, and HEO orbits all had a perfect records
- GTO, ISS, LEO, MEO, and PO orbits had a success rate between 40 and 80 percent.
- SO had zero successful orbits.

Success Rates Increase over Time

- The Success rate since 2013 kept increasing till 2020
- The bottom of the chart illustrates the outcome change as well with the lines from multiple failure become points as the failures become less prevalent.



SQL Results



SQL Results

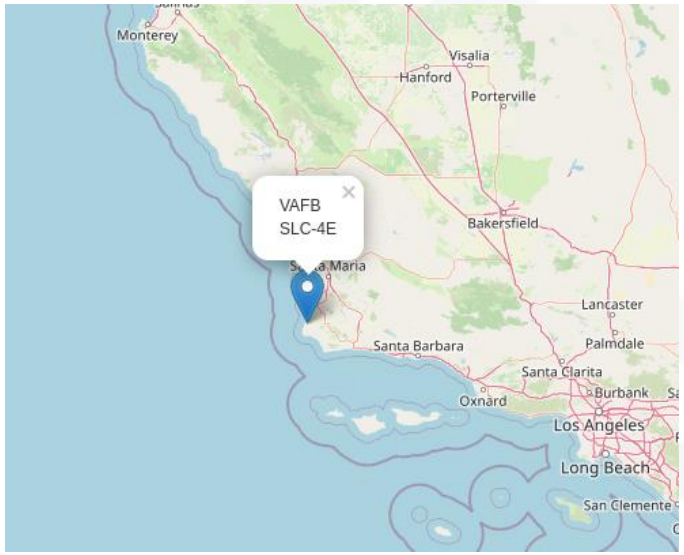
first successful landing outcome in ground pad was achieved. : 2015-12-22

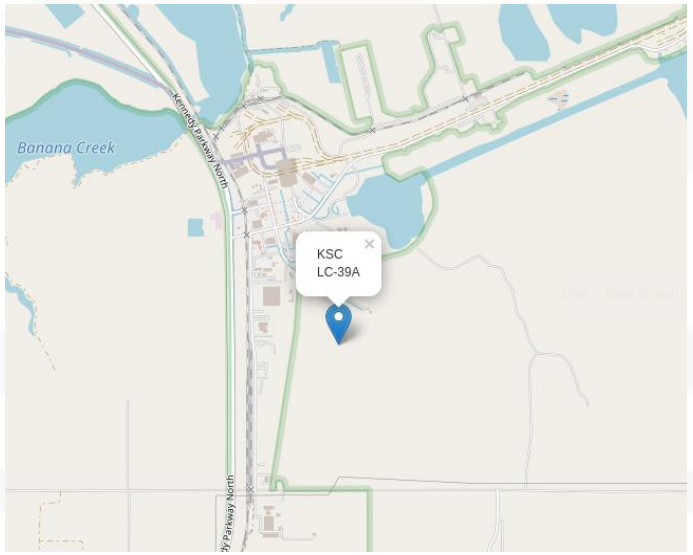
Successful landings out number failed

Identified the boosters capable of caring the max load

Names Launch Sites

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





Maximum Payload Mass

Booster Version
F9 B5 B1048.4
F9 B5 B1048.5
F9 B5 B1049.4
F9 B5 B1049.5
F9 B5 B1049.7
F9 B5 B1051.3
F9 B5 B1051.4
F9 B5 B1051.6
F9 B5 B1056.4
F9 B5 B1058.3
F9 B5 B1060.2
F9 B5 B1060.3

The list on the left are the Booster Version that successfully handled the max payload size.

Drone Ship Outcomes

The table **below** shows the failed drone ship outcomes in 2015

Booster Version
F9 FT B1022
F9 FT B1026
F9 FT B1021.2
F9 FT B1031.2

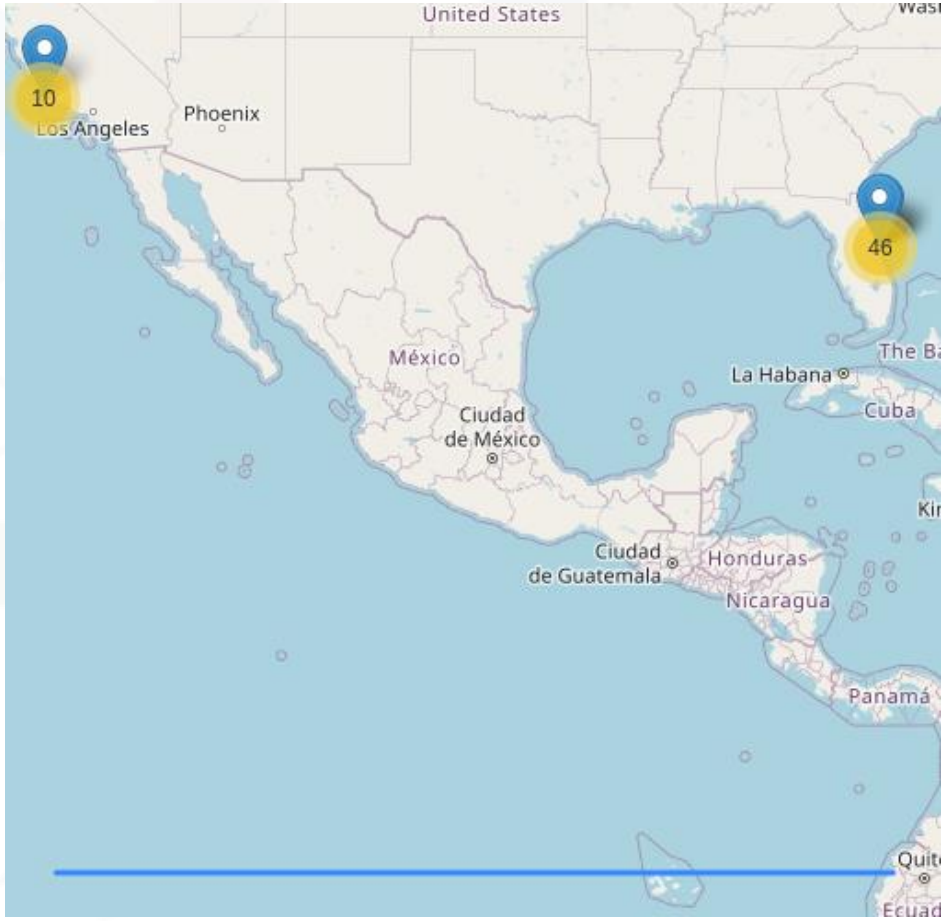
Landing Outcome	Booster Version	Launch Site
Failed	F9 v1.1 B1012	CCAFS LC-40
Failed	F9 v1.1 B1015	CCAFS LC-40

The table **above** shows the success in drone ship outcomes that have payload mass greater than 4000kg, but less than 6000kg.

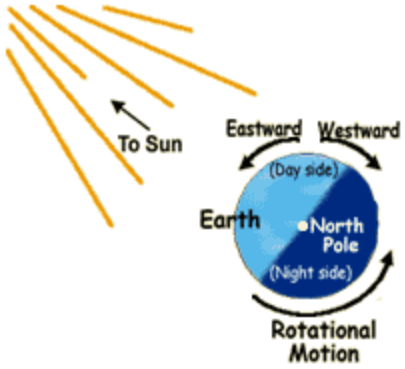
Interactive Map Results

Launch Site Locations

- All four launch sites are Southern region of the USA and close to the coast.
- KSC LC-39A, CCAFS SLC-40, and CCAFS LC-40 very close proximity to each other and make up the majority of launches (46).



Launching to the East



Also, Earth rotates eastward on its axis, one complete turn each day. At the equator, Earth's surface is rotating at 1675 kilometers per hour (1041 miles per hour)!So if we launch the rocket toward the east, it will get another big boost from Earth's rotational motion.

Source: spaceplace.nasa.gov

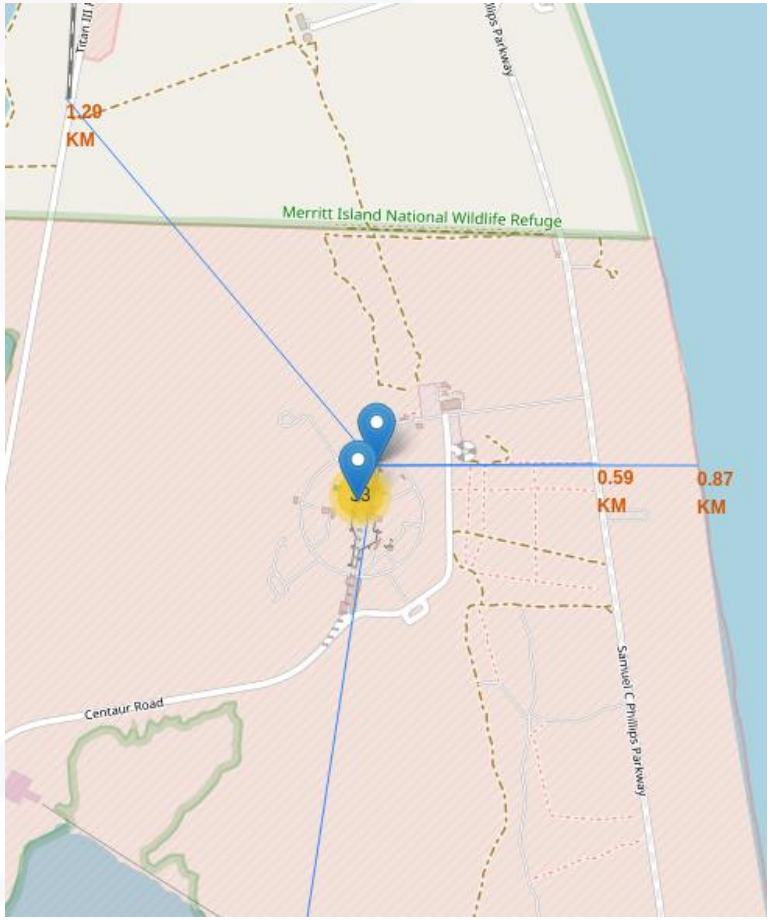
- Launch sites in Florida allow the rockets to go over the ocean rather than land
- Since the sites in Florida are farther south than VAFB SLC-4E they should benefit more from the earth's rotational motion.

Launch Sites and Land marks

The launch sites where kept farther away from cities then highways and railways.

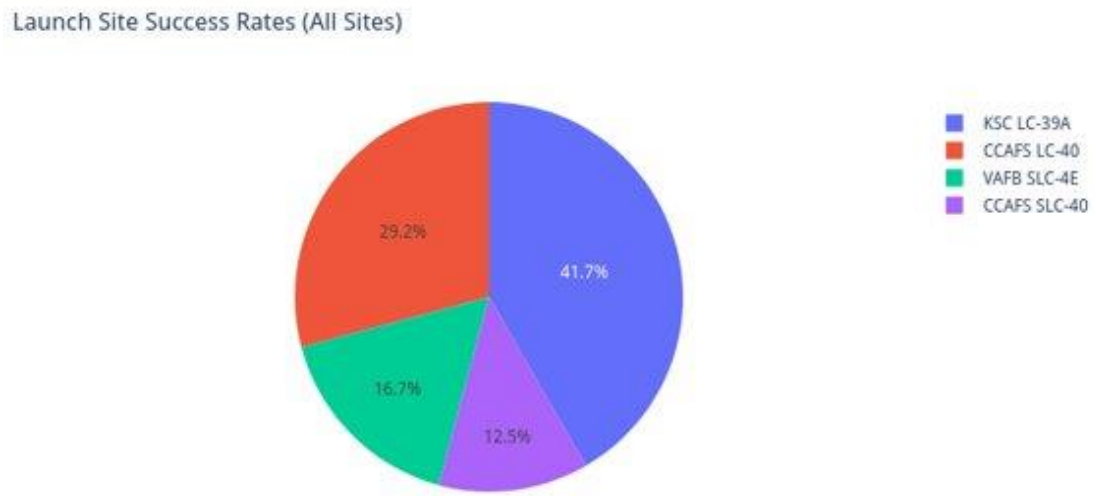
Left: shows CCAFS SLC-40 launch site is 18.21 KM from Cape Canaveral

Right: shows CCAFS SLC-40 launch site only 0.59 KM from a highway and 1.29 KM from a railway.



Dashboard Results

Launch site Success Rates

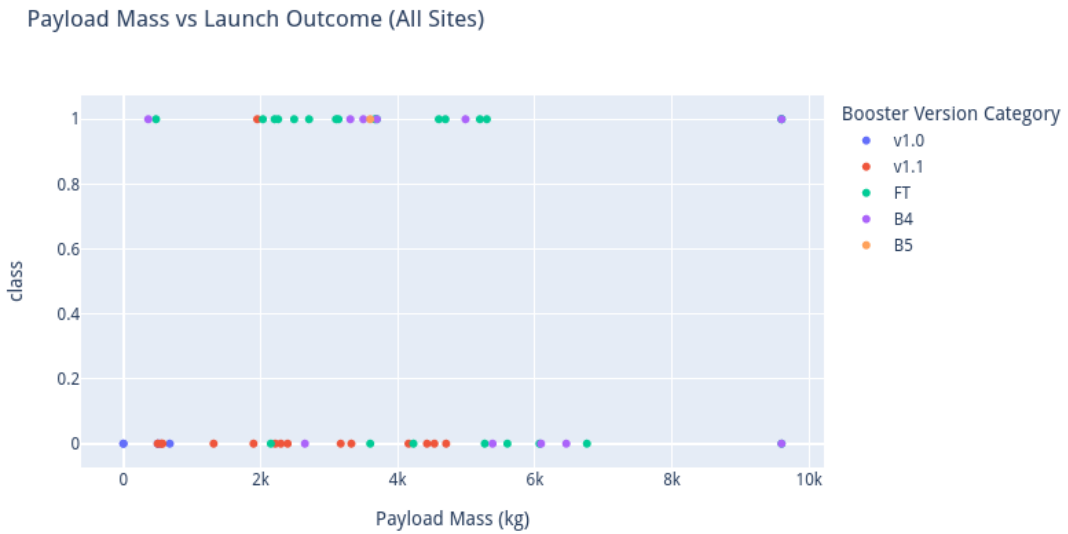


The KSC LC-39A clearly had the highest success rate out for all the launch sites.

CCAFS SSL-40 had the lowest success rate.

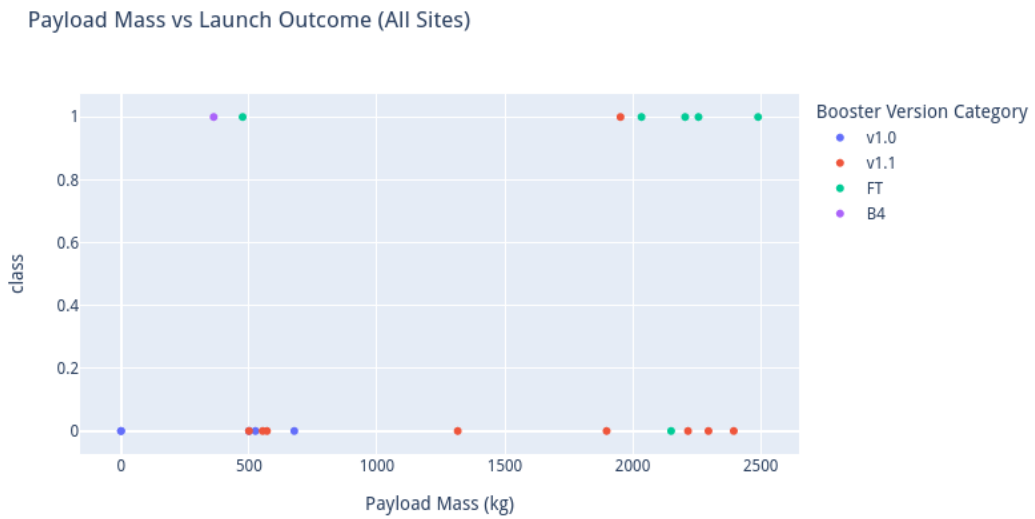
Success vs Payload Size

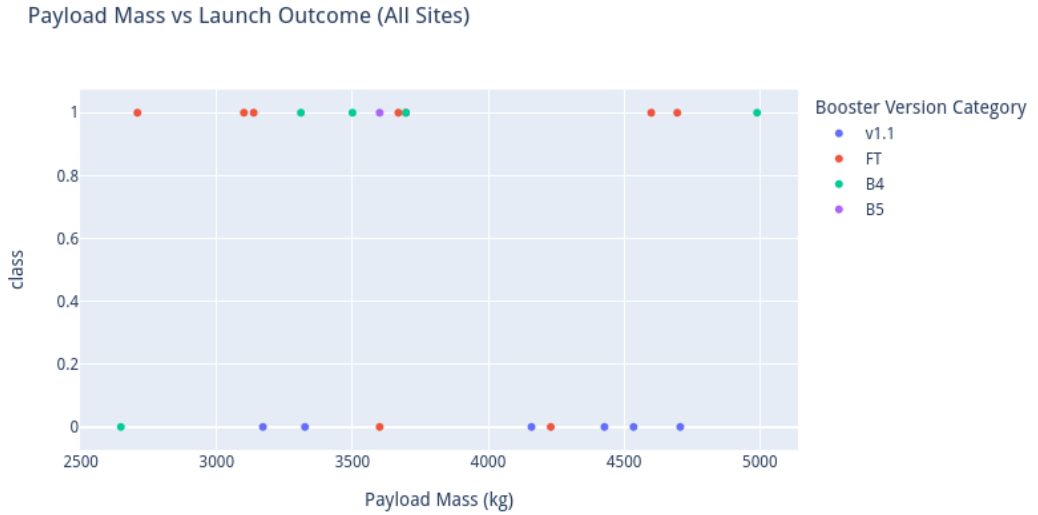
After 2000 kg, the likelihood of a successful launch outcome decreases as the payload mass increases.



Below 5000 payload mass

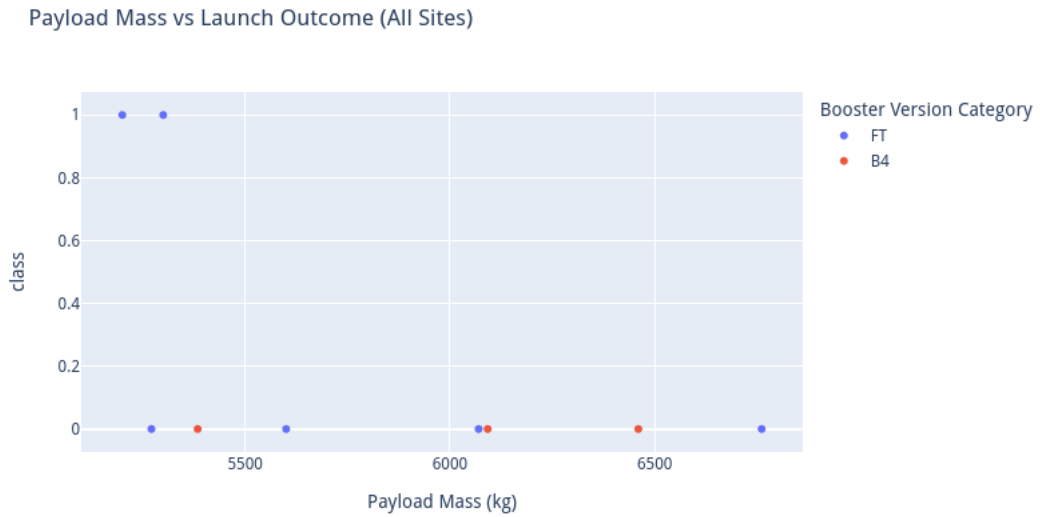
The bottom right diagram shows that 7 successful launch outcomes occurred in between the payload mass of 0 and 2500. While 11 successful launch outcomes occurred in the 2500 to 5000 range, which can be seen on the right.

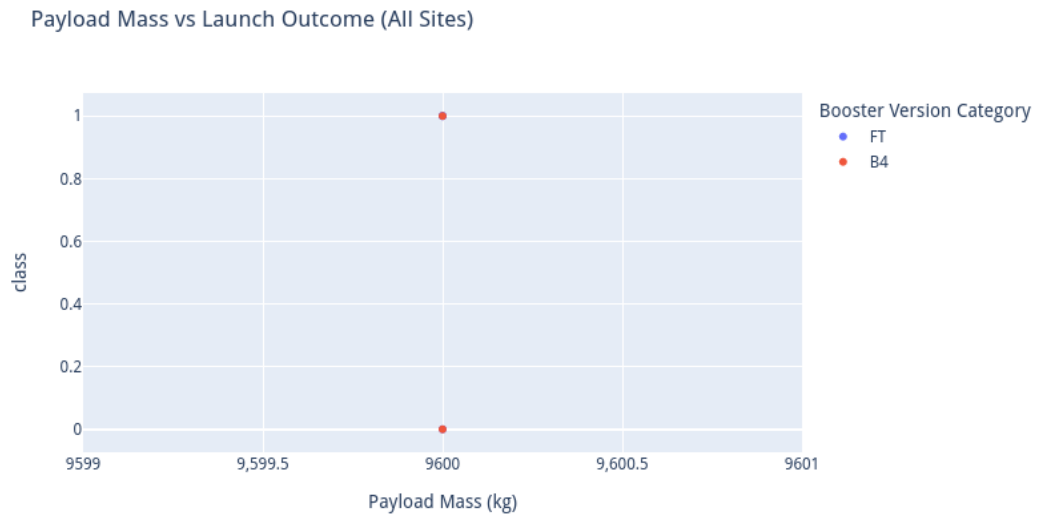




Above 5000 Payload Size

The chart on the lower left shows that only 2 of the launches between 5000kg and 7500kg had successful outcomes. The chart on the right shows that the only one attempt has successfully landed.





Booster Version Category Count

Booster Version Category	Landing Outcome	Landing Outcome Count
B4	Success	6
	Failure	5
B5	Success	1
	Failure	0
FT	Success	16
	Failure	8
V1.0	Success	0
	Failure	5
v1.1	Success	1
	Failure	14

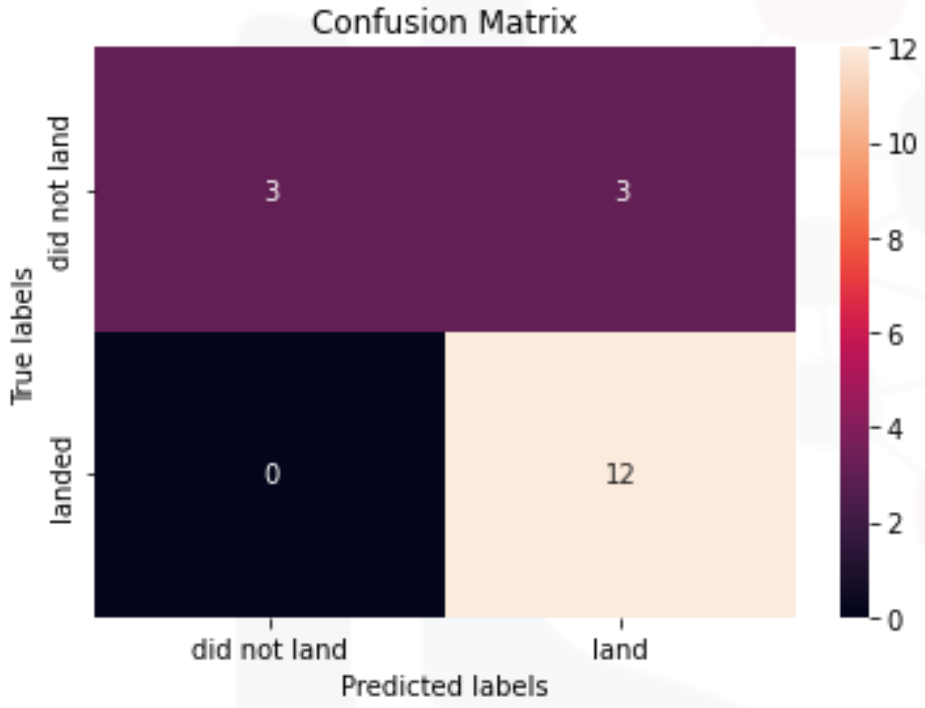
The FT version has the highest number of successful outcomes/landings.

The v1.X categories collectively only have one successful landing

B5 only had one attempt and it was successful

Predictive Analysis Results

Predictive Analysis



All four methods performed roughly the same returning a score of 0.833.

The Confusion Matrix compares the predicted outcomes with the true values.

All four methods produced the same Confusion Matrix

Even the model does a decent job predicting the outcomes. The upper right hand corner has a problem where the model predicted 'land' 3 times when the true label was 'did not land'
(known as 'False Positive').

Conclusion

Conclusion Statement

Based on the finding covered throughout this presentation 83% success rate of our model provides a reasonable level of certainty in predicting the outcome of SpaceX launch.

The key factors used to make these predictions location, payload size, and booster version.

Therefore, SpaceY should be able to use the models to make a competitive bid against SpaceX.

Appendix

Data Collected from:

- SpaceX API <https://api.spacexdata.com>
- Wikipedia <https://en.wikipedia.org/wiki/List_of_Falcon_9_and_Falcon_Heavy_launches>

Jupyter Notebooks used for to generate charts can be found here

- https://github.com/rdesfo/IBMAppliedDataScienceCapstone

Powerpoint Template

- Data Science Capstone PowerPoint template

Scikit-learn: Machine Learning in Python, Pedregosa *et al.*, JMLR 12, pp. 2825-2830, 2011.