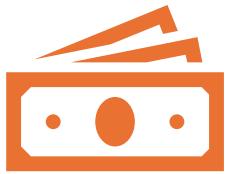


Space Y Project

IBM Data Science Capstone by Kolby Williams

Intro



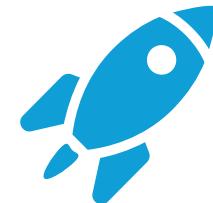
\$62 mill per launch



Predict launch Success

Three main observations (Site, Booster, Time)

Other Factors that promote success



Space Y trying to compete with Space X

We need to understand

- How Space X Launches
- What They launch and were
- Where we should replicate their processes

Methodology

Data collection
and Wrangling

Data
Organization
and Exploration

Modeling and
evaluation

Methodology (Cont')

- Data collection and Wrangling
 - Web Scrapping and API Access
 - Data Cleansing , Missing values, One-hot Encoding
 - Python Pandas Data Frame

Methodology (Cont')

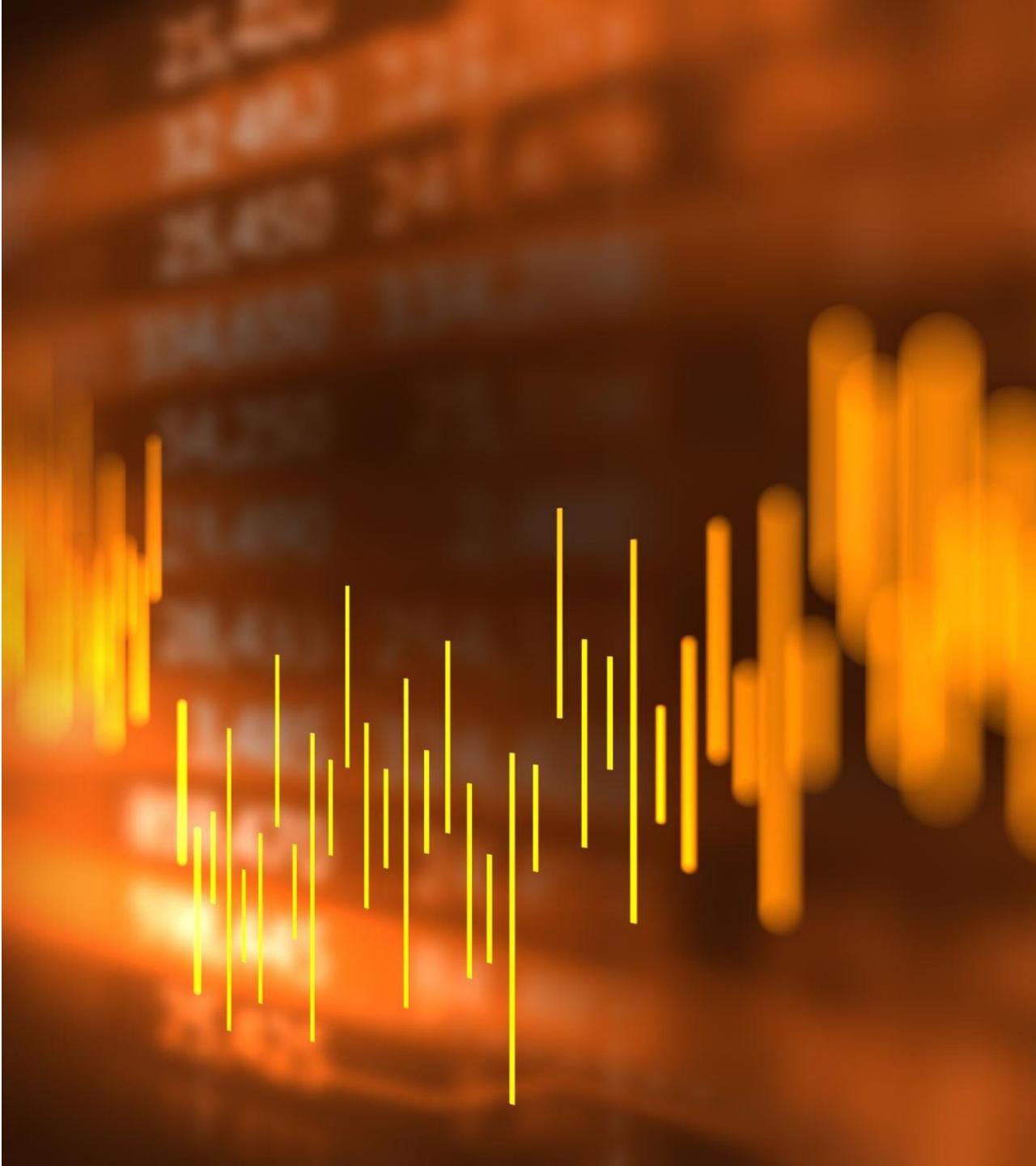
- Data Organization and Exploration
 - Numpy, Pandas and SQL Alchemy for data storing and partitioning
 - Plotly, and Matplotlib for Visualization of Booster Performance
 - Folium for Visualization of Site location and relative Proximity

Methodology(Cont')

- Modeling and evaluation
 - Python Logistic Regression, SVM , Decision tree, KNN for Class evaluation of data
 - Prediction and Confusion Matrix for Precision analysis of prediction of model

Exec Summary

- Stage Performance over time:
improved from 2010 to 2019 by 80%
- Stage Performance by Site:
 - CCAFS SLC had the highest amount of successes
 - KSC LC 39A had the second highest
 - VAFB SLC 4E Came in last
- Stage Performance By Booster:
 - SSO and VLEO had the best success rate



Results

Best Site launch
is CCAFS SLC

Best Boosters
SSO

Higher the flight
numbers the
more success
rates

Most sites are
located near
Coastlines and
railways

All methods had
an accuracy
score of 83.33%

Conclusion

Conclusion:

- Space Y needs to utilize
Similar launch sites near
coastlines and railways
with mature SSO boosters