



Peer Assignment Applied Data Science Capstone

Hengkie Wirawijaya

April 04, 2023

<https://github.com/qprod2000/IBM-Data-Science-Professional-Certification>

OUTLINE



- Executive Summary
- Introduction
- Methodology
- Results
 - Visualization – Charts
 - Dashboard
- Discussion
 - Findings & Implications
- Conclusion
- Appendix

EXECUTIVE SUMMARY



- SpaceX advertises Falcon 9 rocket launches on its website, with a cost of 62 million dollars; other providers cost upward of 165 million dollars each, much of the savings is because SpaceX can reuse the first stage
- If we can determine if the first stage will land, we can determine the cost of a launch
- This information can be used if an alternate company wants to bid against SpaceX for a rocket launch
- This report provided with an overview of the problem and the tools to complete the project

INTRODUCTION



- The commercial space age is here, companies are making space travel affordable for everyone
- SpaceX advertises Falcon 9 rocket launches on its website, with a cost of 62 million dollars; other providers cost upward of 165 million dollars each
- Much of the savings is because SpaceX can reuse the first stage
- This stage is quite large and expensive
- This project will determine whether SpaceX will reuse the first stage
- Using machine learning model and use public information to predict if SpaceX will reuse the first stage

METHODOLOGY (Data Collection & Data Wrangling)



- Collecting SpaceX launch data that is gathered from an API, specifically the SpaceX REST API
- Perform a get request using the requests library to obtain the launch data, response will be in the form of a JSON
- The column Outcome indicates if the first stage successfully landed

METHODOLOGY (EDA & Interactive Visual Analytics)



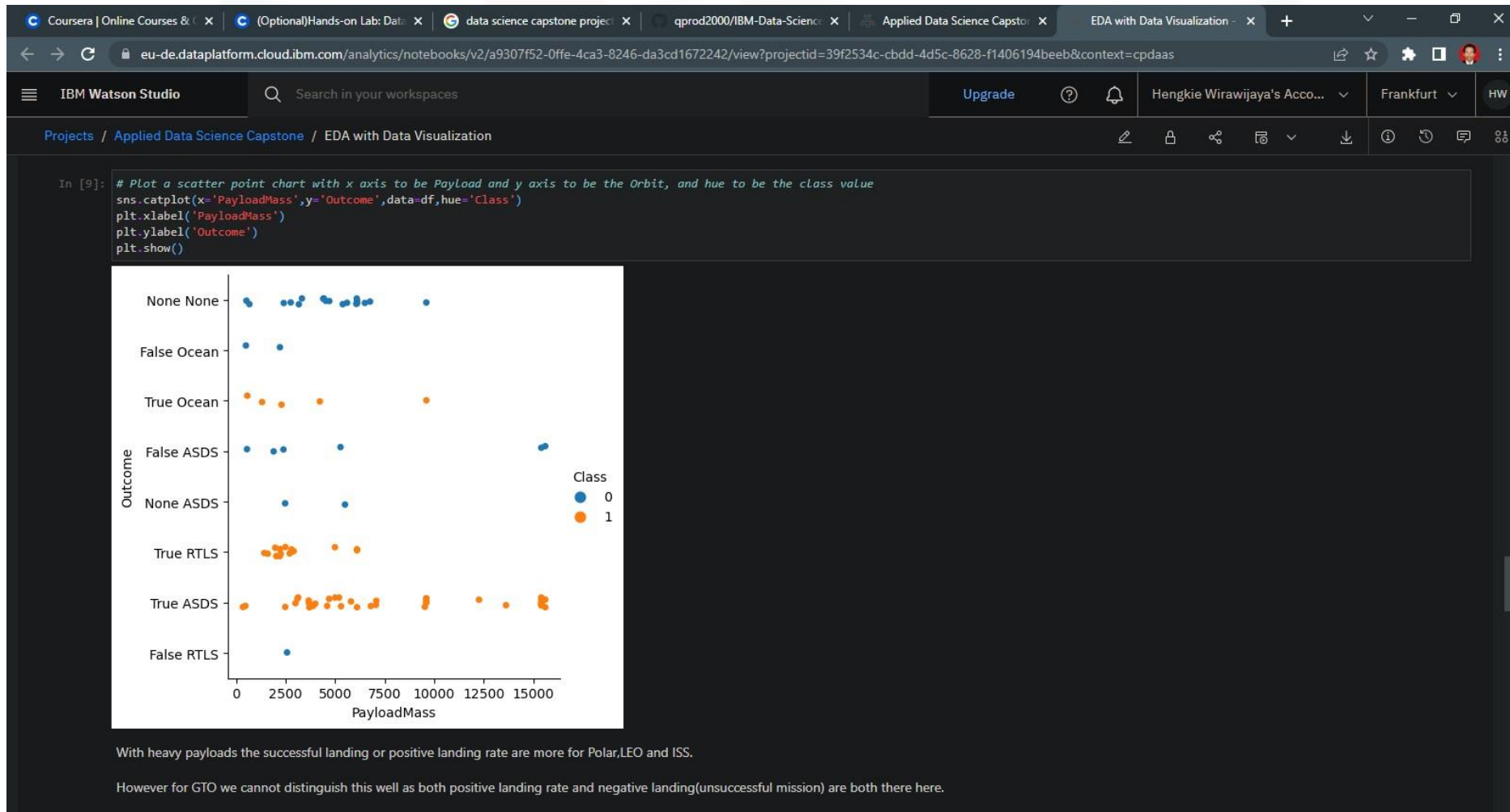
- Exploratory Data Analysis is the first step of any data science project
- Combine multiple features to determine what attributes are correlated with successful landings
- Interactive visual analytics enables users to explore and manipulate data in an interactive and real-time way
- After the dashboard is built, it can use to find more insights from the SpaceX dataset more easily than with static graphs

METHODOLOGY (Predictive Analysis)

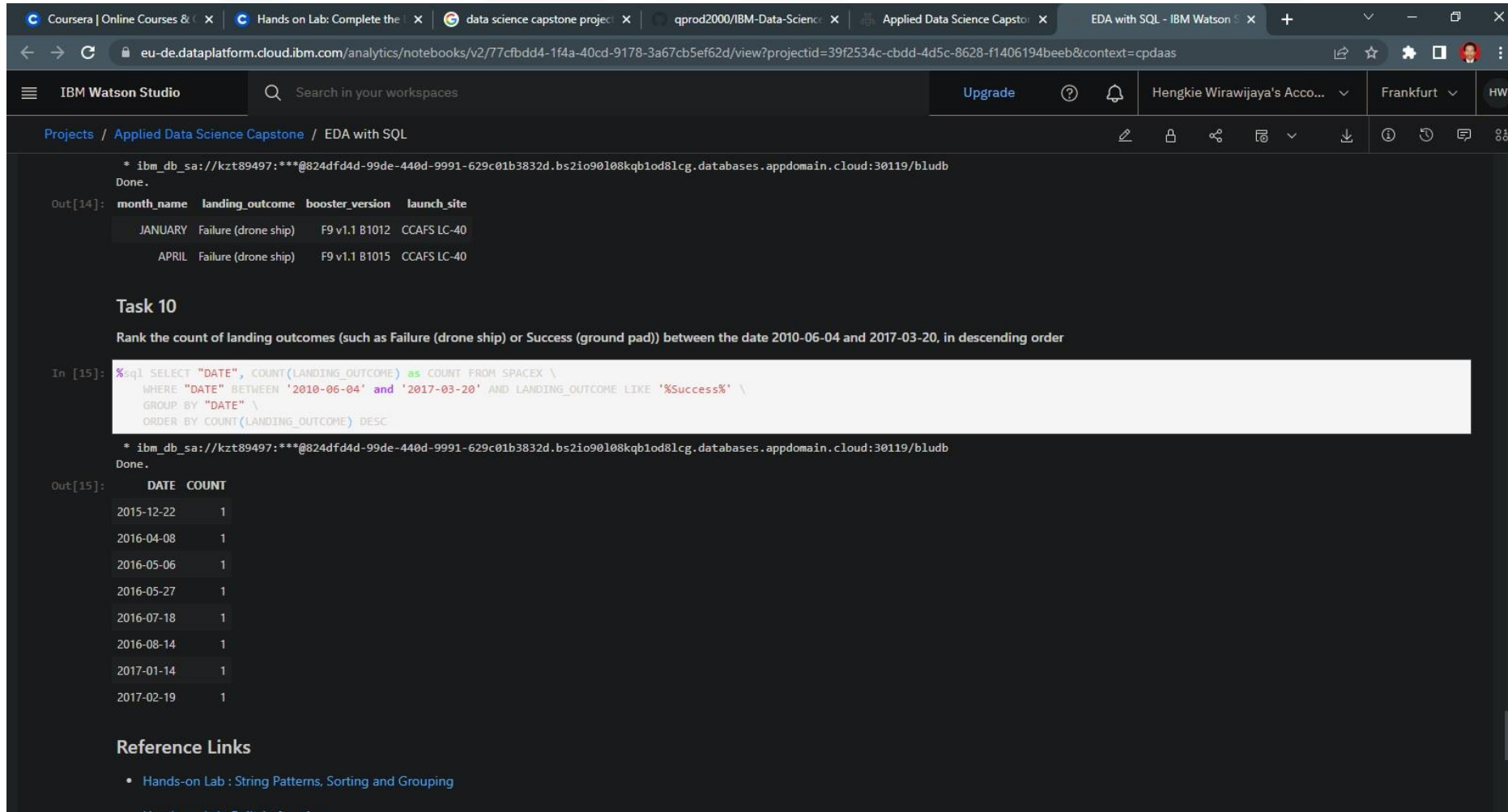


- Interactive visual analytics enables users to explore and manipulate data in an interactive and real-time way
- With interactive visual analytics, users could find visual patterns faster and more effectively
- Using Folium and Plotly Dash to build an interactive map and dashboard to perform interactive visual analytics
- Dashboard can be use to find more insights from the SpaceX dataset more easily than with static graphs

RESULTS (EDA with Visualization)



RESULTS (EDA with SQL)



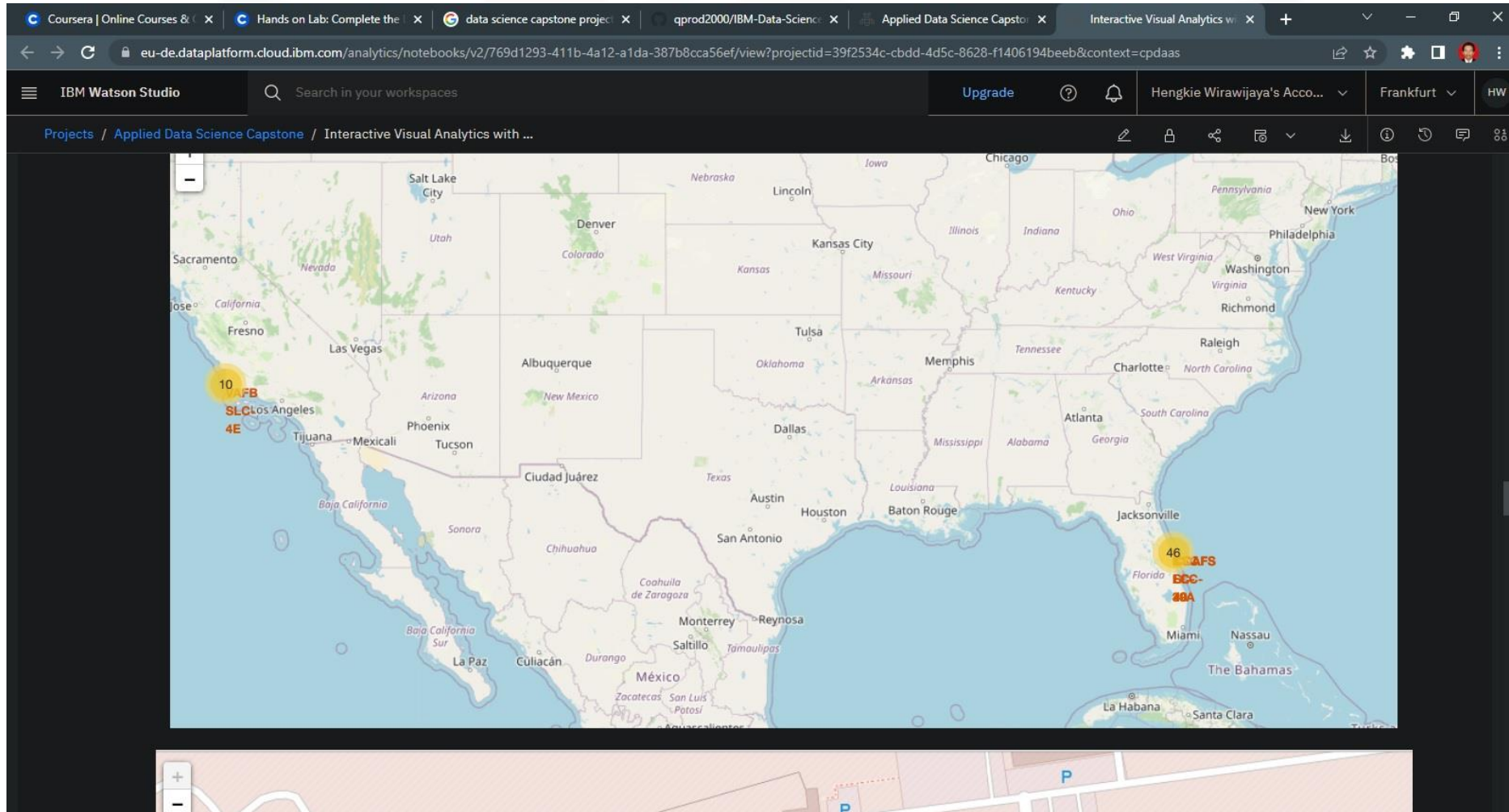
The screenshot shows the IBM Watson Studio interface. The browser address bar displays the URL: `eu-de.dataplatform.cloud.ibm.com/analytics/notebooks/v2/77cfbdd4-1f4a-40cd-9178-3a67cb5ef62d/view?projectId=39f2534c-cbdd-4d5c-8628-f1406194beeb&context=cpdaas`. The page title is "EDA with SQL". The notebook content shows a previous output (Out[14]) with a table of landing outcomes. Below that, "Task 10" is defined: "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". The current input (In [15]) is a SQL query: `%sql SELECT "DATE", COUNT(LANDING_OUTCOME) as COUNT FROM SPACEX \nWHERE "DATE" BETWEEN '2010-06-04' and '2017-03-20' AND LANDING_OUTCOME LIKE '%Success%' \nGROUP BY "DATE" \nORDER BY COUNT(LANDING_OUTCOME) DESC`. The output (Out[15]) shows the results of the query as a table with two columns: DATE and COUNT.

DATE	COUNT
2015-12-22	1
2016-04-08	1
2016-05-06	1
2016-05-27	1
2016-07-18	1
2016-08-14	1
2017-01-14	1
2017-02-19	1

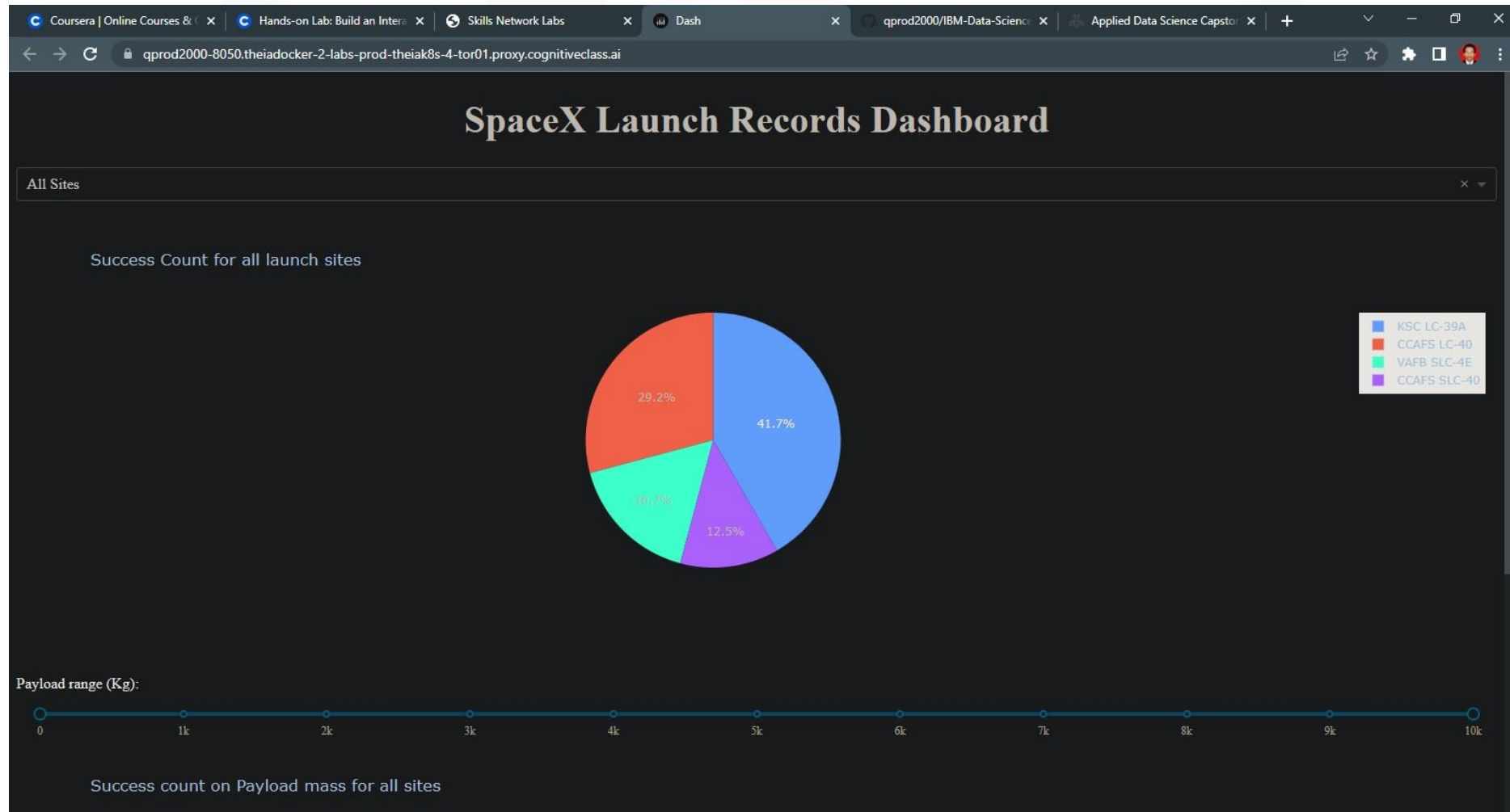
Reference Links

- Hands-on Lab : String Patterns, Sorting and Grouping
- Hands-on Lab : Built-in functions

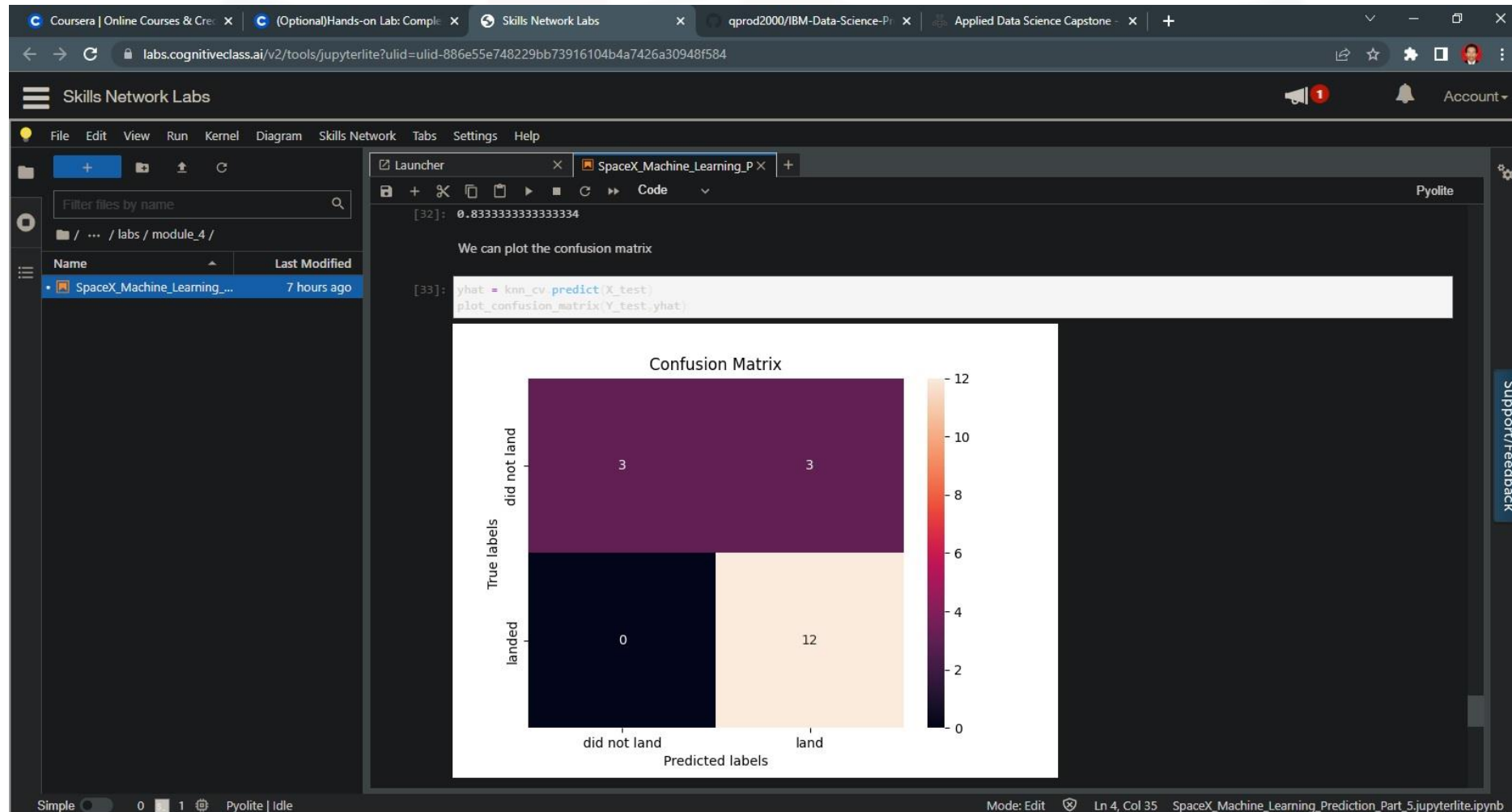
RESULTS (Interactive Map with Folium)



RESULTS (Plotly Dashboard)



RESULTS (Predictive Analysis)



CONCLUSION



- Allon Mask of SpaceY can use this model to predict with relatively high accuracy whether a launch will have a successful Stage 1 landing before launch to determine whether the launch should be made or not
- If possible more data should be collected to better determine the best machine learning model and improve accuracy

APPENDIX



- <https://github.com/qprod2000/IBM-Data-Science-Professional-Certification>