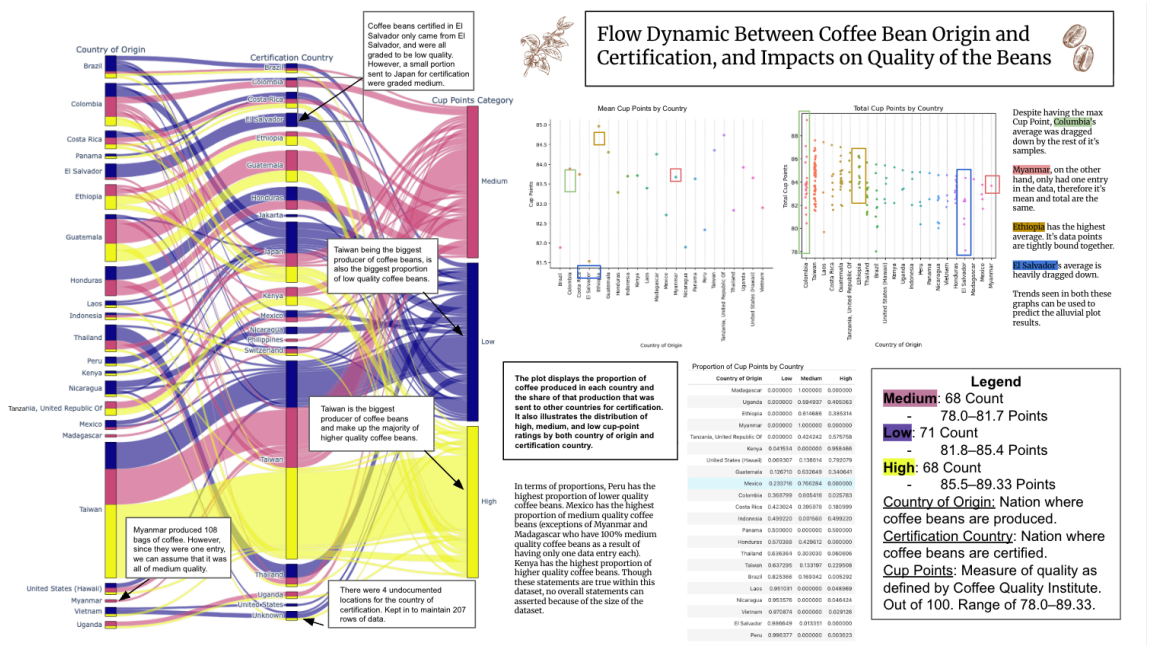
Final Product - Flow Dynamic Between Coffee Bean Origin and Certification, and Impacts on Quality of the Beans

- NAME: Chloe Wang
- ID: Qtr7bs



This is an alluvial plot showing the flow relationship between the origin country (of which coffee beans were produced), the certification country (of which coffee beans were brought to and certified), and lastly, the quality of the coffee beans as defined by the original data set. While four columns (categories) were chosen to be explored in depth, these three were chosen in the end because the flow made **sense**. From the country where the coffee beans were produced, to the nation where they were certified, it made sense to see how that flow impacted the quality of the coffee beans. However, it is important for me to mention here that while the Cup Points Category is divided into low, middle, and high, that is only in relation to the original range of cup points these bags of coffee beans were assigned to. In other words, from the original range of 78.0-89.33, the range was divided into thirds (lowest 33%, middle 33%, and highest 33%. A low score according to the alluvial plot above may actually still be considered middle or high since the difference in range from the original range is just around 10--low considering the highest cup points a bag of beans can get is 100.

You can see from the alluvial plot that Taiwan is a major producer of the coffee, at least in the case of this study. In addition, you can see that Taiwan is also a major certifier of

coffee beans, with virtually all of it's coffee produced being funneled into it's own certification program. This is in stark contrast to some other countries, such as Nicaragua or Myanmar, who don't have their own certification centers and can only rely on other countries. It is also worth noting that the hue of this plot is based on the cup point rating to allow readers to see the proportion of medium, low, and high quality coffees that come out of each nation as well as the certification nation.

All Resources

Resource Name	Description	Link / File Path
df_arabica_clean.csv	Original Coffee Quality Institute May 2023 dataset	Chloe/df_arabica_clean.csv
Rashed Sumon's Kaggle Post	Description of Coffee Quality Data from May 2023	https://www.kaggle.com/code/rashedsumon/coffee-quality-data-cqi-may-2023/input
Establishing_Data.ipynb	Notebook for establishing the data	Chloe/Establishing_Data
Exploring_Data.ipynb	Exploratory data analysis notebook	Chloe/Exploring_Data
Coffee_Alluvialplot.png	Exported final alluvial diagram used in report	Chloe/Coffee_Alluvialplot.png
CQI Website	Background information about certification and scoring	https://coffeeinstitute.org
ChatGPT	For help in coding	https://chatgpt.com
Design HW07	Alluvial Plot Code	Chloe/HW07 (1) (1).ipynb