Bean!

Coffee Bean Recommender

🫘 Objective

This project aims to connect local coffee producers with wholesale coffee suppliers. While there are numerous farms that are producing coffee, it might still be difficult for the wholesale business to learn about coffee producers and find the kind of coffee that the market wants. By introducing sources of high-quality coffee beans to the market, the 'Bean!' application helps improve the lives of local coffee producers and contributes to a thriving coffee community.

From the application interface, users can input their desired values of coffee bean attributes (aroma, aftertaste, acidity, sweetness and moisture), and then they will get a list of 15 coffee beans generated by the k-means algorithm with the data downloaded from Kaggle from the Coffee Quality Institute (CQI). Users can then go to CQI's webpage and look for the coffee bean owner's contact information. This application saves time for users to find out the coffee beans they want and the producers they should contact with.

🫘 Data

This Kaggle dataset was scraped from the review pages of the Coffee Quality Institute in 2018. Coffee Quality Institute (CQI) is a non-profit organization working internationally to provide the ratings for coffee bean quality since 1996. The data contains owner's information and attribute scores for 1312 arabica and 28 robusta coffee beans that were given by the CQI's trained coffee bean reviewers.
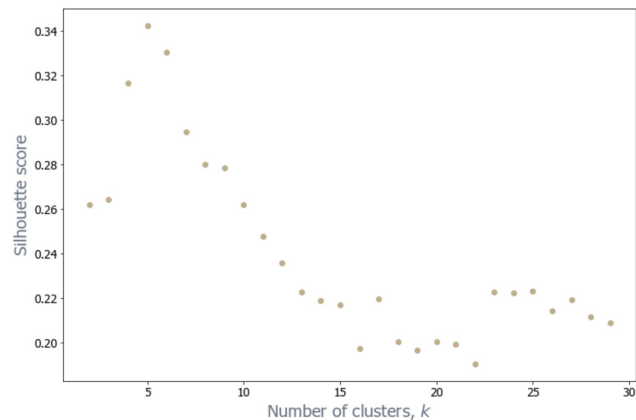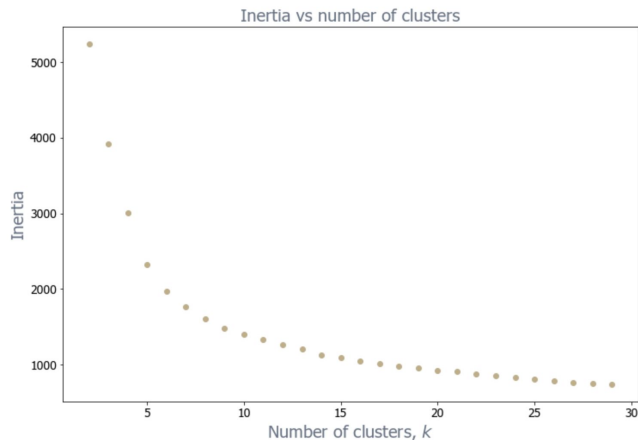
Features used in the K-means clustering analysis:

- Aroma
- Aftertaste
- Acidity
- Sweetness
- Moisture

🫘 Model

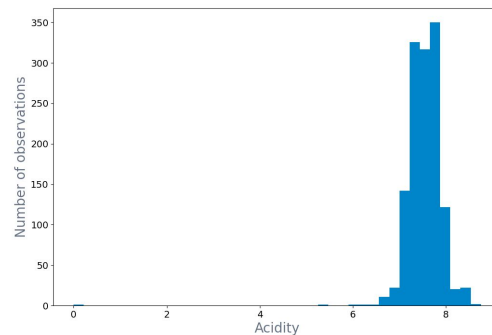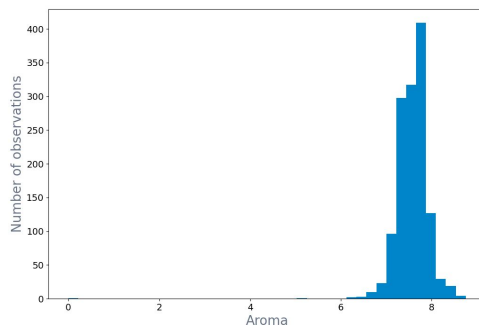The recommender was built upon K-means clustering method. The optimal number of k = 5 was chosen from the Inertia and Silhouette plots. This number of clusters should generate distinct groups that contains coffee beans with very similar selected attributes.

🫘 Insights

- The attribute ratings for different beans are very similar. The distinction between different clusters is hard to tell.
- CQI is a good platform to learn more coffee bean farms around the world. The coffee bean ratings are updated frequently. For this application, it would be nice to have database that is updated with the website.

Thank you!

Xiaohan (Aria) Wang

Northwestern MSiA 2020
Email: xiaohanwang2020@u.northwestern.edu
LinkedIn: https://www.linkedin.com/in/aria-wang/