黃佩筠和清大衛

Sammi and David

**Comparison of Coffee Production in South America to the Prevalence of Locations of Starbucks**

Our group proposes to use two CSV files, both acquired from Kaggle, one of which contains data on the production, consumption, and exportations of coffee, the other of which contains data on the locations of Starbucks in various countries. We will link the two datasets on the category of country, but since the file with information on Starbucks uses a code for country (AE for United Arab Emirates, for example), while the file with information on production uses the actual name of the country, we will locate a third dataset that links countries and codes for countries. We will then join the dataset for production to the dataset with conversions, and then use the combined dataset in a join with the dataset for Starbucks. We will focus our investigation on countries of South America that have the highest levels of production and exports for coffee in order to see if Starbucks has located stores in countries that might serve as their most significant suppliers.