"Which Country Provides the Best Coffee?"

Problem Description

We are a coffee company based in Southampton, UK, and are looking to improve where we source our coffee from. The quality of coffee sourced from our existing suppliers has significantly reduced in recent years, so we want to change suppliers whilst maintaining our reputation as the best artisanal coffee place in town.

We have obtained a wealth of coffee supplier data from the independent Coffee Quality Institute, but we lack the skills required to analyse the data to help us decide where to send our coffee buyers to negotiate a new supply.

We would like an answer to the following question: **"Which country should we send our buyers to?"**

- Provide evidence for your choice, including statistics and visualisations of all relevant variables that have impacted your choice
- The modified code, supplied in a GitHub repository, with short documentation that provides instructions for us to easily re-run your analysis code to reproduce your findings

Please bear in mind the following information about the coffee preferences of our business and customers:

- A survey found that our customers care most about the flavour, aroma, body and uniformity of their coffee.
- Our customers prefer coffee with a lower caffeine content.
- We would prefer to send our buyers to a country with lots of coffee producers so that they can visit multiple producers during their visit.
- We have a preference for washed/wet processing of beans because it results in a more consistent and predictable flavour profile.

Dataset

The data comes from the Coffee Quality Institute and was assembled by Data Scientist James LeDoux into the Coffee Quality Database GitHub Repository. The data contains reviews of 1312 arabica and 28 robusta coffee beans from the Coffee Quality Institute's trained reviewers. The features include:

- Quality measures: aroma, flavour, aftertaste, acidity, body, balance, uniformity, cup cleanliness, sweetness, moisture, defects
- Bean Metadata: processing method, colour, species (arabica / robusta)
- Farm Metadata: owner, country of origin, farm name, lot number, mill, company, altitude, region

Existing Code

We have some existing code that was developed by a part-time barista who was studying computer science at the University. We would like to build on this existing code, whilst ensuring we're able to run the updated code ourselves in the future.

This code generates a visual plot of United States coffee suppliers in terms of flavor and cupper points, and can be found at https://github.com/softwaresaved/coffee-analysis. It is not currently documented, but can be run using the following commands:

```
python3 -m venv venv
source venv/bin/activate
pip install -r requirements.txt
python analyse-coffee.py
```

The code produces the following plot. The plot is quite confusing and doesn't help us, but please use the code as a starting point if it's helpful to you at all.