

Team 7:

Moez Hudda
Mohammed Elzubeir
Pablo Aganza
Sonia Song

Github Repo:

https://github.com/UC-Berkeley-I-School/Project2_Team_7 (public)

Primary Data Set:

<https://www.kaggle.com/datasets/parasrupani/coffee-distribution-across-94-counties>

Column Names to explore:

Countries, Year, Arabica Production, Bean Exports, Domestic Consumption, Export and Imports

Supplemental Datasets:

Coffee Quality Data (CQI) - example from May 2023:

<https://www.kaggle.com/datasets/fatihb/coffee-quality-data-cqi>

(Join on country)

Coffee Ranking:

<https://www.kaggle.com/datasets/schmoyote/coffee-reviews-dataset>

(Join on country)

Winning coffee by region:

<https://allianceforcoffeeexcellence.org/bolivia-2009/#1513038058770-76a416aa-73df>

(Regions within countries)

Reviews of 1312 arabica and 28 robusta coffee beans from the Coffee Quality Institute's trained reviewers:

<https://github.com/jldbc/coffee-quality-database/tree/master/data>

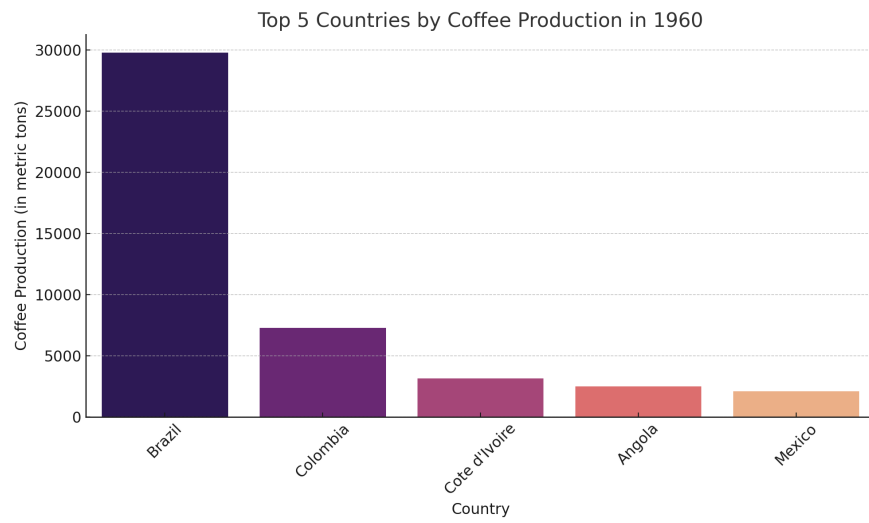
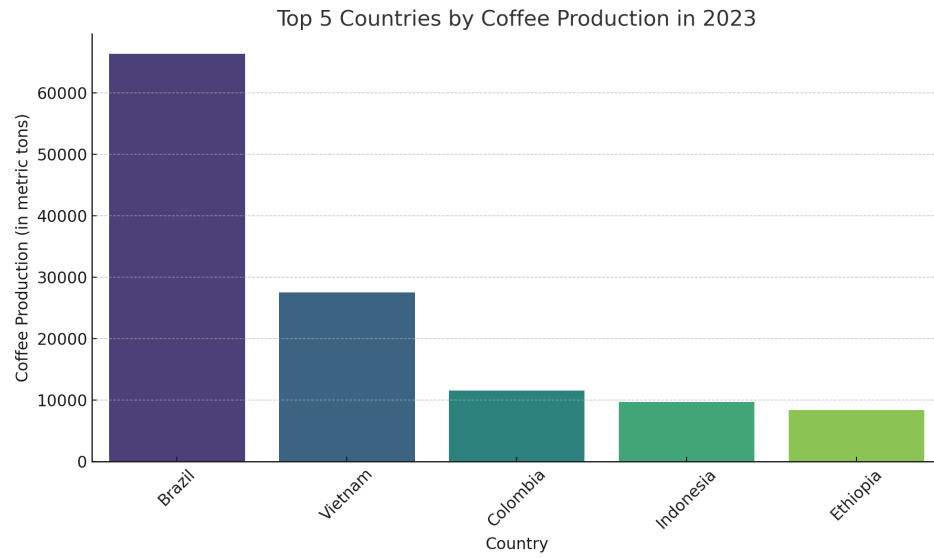
GDP

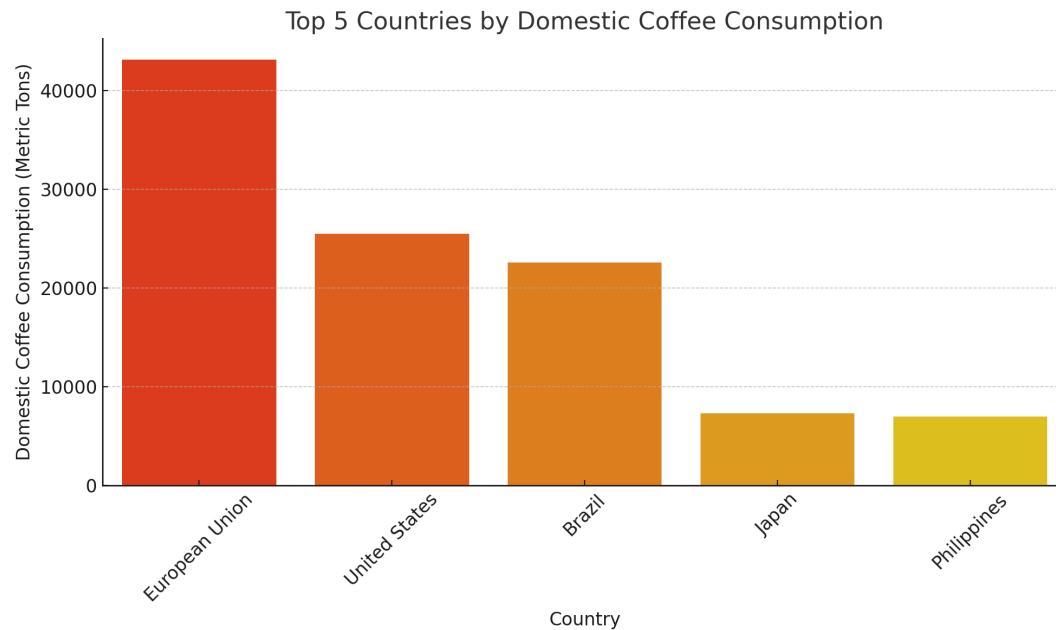
<https://data.worldbank.org/indicator/NY.GDP.MKTP.CD>

Country Land Area

<https://data.worldbank.org/indicator/AG.LND.TOTL.K2>

Plots, figures and tables:





Country by Quality

Consumption of country by farmed of country

Altitude by quality

Altitude by volume

Plan to Cover & Organize

Which countries grow the best coffee beans?

Consumption of coffee beans by the country in which it was farmed.

Relationships between altitude, quality, volume produced.

What percentage of a country's GDP is coffee production?

Size of country to production volume ratios.

Why is it the best quality coffee?

We plan to display this data within either graphs, plots, tables or charts and present it with explanations of how we were able to use the main data set + supplemental data sets to come to these predictions.