

“Global Trends in Coffee: A Comprehensive Analysis of Imports, Exports, and Consumption Patterns across Countries (1990-2020)”

IST 652 Final Project

Ceara Stewart

Debbie Landowski | Syracuse University | 6/9/2023

# Introduction

Coffee is one of the world’s most beloved and widely consumed beverages, I can’t even start my day before brewing myself a cup or two. But not many people think about where their coffee comes from, or just how much its consumption has changed over the years, most people only care about what new and fun flavor is available at our closest Starbucks or Dunkin. Coffee holds a prominent position in global trade and understanding the dynamics of coffee imports, exports, and consumption across different countries is important to not only big businesses but the global economy as well.

This paper provides a comprehensive analysis of coffee imports, exports, and consumption across different countries from 1990 to 2020. By examining long-term trends and patterns in the global coffee market, one can gain insights into the changing consumer preferences, economic interdependencies, and the impact of global events. Throughout this paper, I will be presenting findings related to three posed questions:

* Do countries consume more of the coffee they specialize in producing?
* Is there a difference in coffee amounts for countries who produce, consume and export as well as countries who import, consume and re-export?
* Is there a pattern of changing coffee consumption from 1990 to 2000 and does green coffee investments follow this pattern?

## Data and its Source

For this paper, the data is pulled from Kaggle, a popular site containing compiled data sets for public use. The seven data sets I downloaded from Kaggle come from the International Coffee Organization’s site. The seven data sets were presented in Excel csv files and are listed as follows:

* Coffee\_domestic\_consumption.csv
* Coffee\_exports.csv
* Coffee\_green\_coffee\_inventorie.csv
* Coffee\_import.csv
* Coffee\_importers\_consumption.csv
* Coffee\_production.csv
* Coffee\_re\_export.csv

Each data set contains the country’s name, years from 1990 to 2000 and coffee amounts in pounds. The coffee\_production.csv and the domestic\_coffee\_consumption have an additional column containing coffee type produced/consumed. You can find the data sets using this link <https://www.kaggle.com/datasets/michals22/coffee-dataset>.

## Program Description

The code was written and executed in Python Jupyter Notebook; an open-source software designed for interactive computing. Jupyter Notebook enables code execution in separate blocks and facilitates manipulation without the necessity to rerun the entire code. This allows for one to set specific, cleaned code such as data frames to blocks as well as compute large numbers of individual analysis without the longer waits associated with having to rerun hours of code. Utilizing Jupyter Notebook allowed for code writing, execution, and the inclusion of comments that provide insight into the purpose of each code block, while also obtaining the desired outputs.

Following the execution of the code in Jupyter Notebook, both the code and its corresponding output were transferred to Visual Studio Code, or VS Code for short. VS Code is a source-code editor. It serves as the platform for submitting the code as a Jupyter source file and offered a range of functionalities that include code debugging, syntax highlighting, and code refactoring. These features are instrumental in guaranteeing the quality and accuracy of the final code that is submitted.

## Data Exploration and data cleansing

To pull in the Excel csv files, the csv and pandas packages were called, along with a matplotlib package needed to create graphs and charts using the data. The csv files did not read into Jupyter Notebook in the same way it can be viewed from the Excel file directly, so each header of the csv file, which contained the years, needed to be skipped over and later extracted into an empty years list that is entered back into the data frames.

The coffee amounts in pounds were converted into integers to run mathematical calculations without running into errors. And the country names needed to also be assigned into an empty list and appended to ensure each country was represented in its own row. All this data had to be added together under a new data frame where the country had to be set as an index to ensure that when the years were added back as column headers, the data was still correctly matched to the right country. Some data frames also needed the last column deleted so that “total coffee amounts per country” weren’t skewing the data.

When running some analysis, the years header had to be changed so that a streamline match or join could be made between two data sets. No data values were dropped if a zero occurred as some countries didn’t start participating in global coffee production, exportation and importation until later years.

# Analysis

To save space in the actual body of this paper, only a small snippet of the entire output is included in the paper as a reference. The entire output is presented via an html attachment. This is due to the length of the output as well as the number of output snippets. Each block of code is named and references the question the code seeks to answer.

## Question #1

A screenshot of a computer

Description automatically generated with medium confidence The first question I seeks to answer is “Do countries consume more of the coffee they specialize in producing?” This question utilizes the coffee\_production\_df and the domestic\_consumption\_df and looks to see if production countries who produce either Arabica or Robusta, also consume the coffee they specialize in producing. To answer this question, the two data frames were joined together using an inner join, which combines records from two tables whenever there are matching values in a common field to both tables. The common field used in this case was “Country”. Once the tables were joined, a comparison was run using a Boolean operator, which returns a “true” if the values in the column are not equal, and a “false” if the values in the columns are equal. The results are then printed, where any “false” value is printed as “Match” and any “true” value is printed as “Mismatch”. A “Match” means that the country that specializes in either Arabic or Robusta production also consumes that same coffee type domestically.

The code output provides the knowledge that all production countries consume the same coffee type that they produce. This makes sense as domestically produced products are cheaper than imported products. Buying domestically allows consumers to save money and use it to consume other products, boosting the country’s economy. If any of the production countries has resulted in a “Mismatch”, it would have meant that the coffee type they choose to produce is more expensive than that is produced elsewhere, leading to an import of cheaper coffee for consumers to drink. This economic theory usually is only found among “rich” nations, and since most coffee producing countries are not as economically superior, the results of the data make economic sense.

## Question #2

The second question I would like to answer is “Is there a difference in coffee amounts between what is produced and or imported versus what is consumed? And is this amount equal to what is exported or re-exported?” This question involves the coffee\_imports\_df, the coffee\_re\_exports\_df, the importer\_consumption\_df, the coffee\_production\_df, and the domestic\_consumption\_df.

I started with the countries that produce the coffee and found the difference, for each year, between what they produced and what they consumed. This ended up being what I call “remaining coffee” as it is what is left over for the countries to export.

A picture containing text, screenshot, font, number

Description automatically generated

I then joined this new “remaining coffee” data frame with the exporters data frame and ran a comparison. Any countries that printed a “Match” exported exactly what was left over after consumption, any countries that printed a “Mismatch” meant that the coffee amounts that were exported were different from what was left over after consumption. For production countries, every output was a “Mismatch”.

A screenshot of a computer

Description automatically generated with medium confidence

A picture containing text, screenshot, font, line

Description automatically generated Next, I looked at countries who import their coffee. I ran the same steps of code as I ran for the producer countries. I first found the difference between what was imported and what was consumed. Then I joined that “remaining coffee” data set with the re-exportation data set and ran a comparison. Any countries that printed a “Match” re-exported exactly what was left over after consumption, any countries that printed a “Mismatch” meant that the coffee amounts that were re-exported were different from what was left over after consumption. Like the producer countries, the importer countries all had outputs that read “Mismatch”.

A picture containing text, screenshot, font, number

Description automatically generated

This means that there are amounts of coffee for both producer and importer countries left over each year after consumption and exportation. But what is done with that left over coffee? I researched and found that ground coffee is only good for 3-5 months while whole beans are good for 6-9 months. This means that most left over coffee ends up in landfills and is a contributing factor to greenhouse gas emissions as it breaks down. Some countries, like Moldova, are trying to combat this climate and waste issue by turning left over, unused coffee into raw materials and eco products, <https://www.undp.org/moldova/news/what-can-we-do-leftover-coffee-moldovan-and-polish-companies-partner-upcycle-spent-coffee-grounds>. While some small businesses are using left over coffee and turning them candles, eco-friendly dish cleaners, and garden fertilizers, <https://timesofindia.indiatimes.com/life-style/food-news/5-smart-ways-to-use-your-leftover-coffee-grounds/photostory/93970103.cms?picid=93970431>. But this doesn’t solve the fact that countries are producing more than what is able to be consumed, and there should be the discussion of measures to reduce coffee production that doesn’t harm the countries who make it.

## Question #3

The third question to be answered is “Is there a pattern of an increase or decrease in coffee consumption over the years? And does coffee investment follow this pattern?” To answer this question, I built four different graphs that plot yearly coffee consumption amounts and green coffee inventory. I used thedomestic\_consumption\_df, the importer\_consumption\_df, and the green\_coffee\_df. I utilized the matplotlib.pyplot package to produce the graphs.

The first graph looks at domestic consumption, or what producing countries consume each year from 1990 to 2020. The consumption amounts were summed by country across each year. The graph shows that coffee consumption increased each year for producing countries until we see a small dip in 2020. With more data we might have been able to see if the decrease continues for the next three years are jumps back up to show more of an increasing trend. The logistic nature of the domestic consumption graph suggests that producing countries were able to gain access to more resources to produce more coffee, and once they reached a peak productivity level, started to level out in production. As more coffee was produced, it became cheaper to consumer domestically and the countries moved away from importing more expensive coffee. This led to more consumption as domestic demand increased due to the new affordability of coffee.

A picture containing text, screenshot, plot, line

Description automatically generated

The second graph looks at total importer consumption from 1990 to 2019. The consumption amounts were summed by country across each year, like the domestic consumption graph. The importer consumption graph shows that coffee consumption overall increased over time, but certain years had a spike of consumption compared to others. If you look at which years, you can see a small jump in consumption in 1991, when the cold war ended. There is a decrease in consumption in 1994, when the North American Free Trade Agreement (NAFTA) went into effect. We see another small drop in 2000, which aligns with Y2K. We see another drop in 2009, which follows the September 15th housing market crash of 2008. These consumption trends aligning with historical markers might be a coincidence, but like any other heavily consumed product, historical markers can either slow or increase consumer spending, and this impact coffee consumption.

A picture containing text, screenshot, plot, line

Description automatically generated

The third graph looks at green coffee inventory for countries who invest in green coffee production from 1990 to 2019. The coffee amounts are summed by country across each year, like the other two graphs. Also like the other two graphs, there is an overall increase from 1990 to 2019, but with two major dips. After doing some research, I found this article that could explain these dips, [Learning Lessons From the 1990s: Long-Term Growth Prospects for the U.S. | Economic Policy Institute (epi.org)](https://www.epi.org/publication/webfeatures_viewpoints_l-t_growth_lessons/#:~:text=When%20consumption%20was%20slow%20in%20the%201990s%2C%20so,suit%20and%20declined%20at%20the%20end%20of%202000.). In the mid 1990’s and early 2010’s, consumption of goods across the board slowed due to various factors, one being the repeated oil price shock, which subsequently slowed investments. The avoid high interest rates, investments stayed low until consumer consumption accelerated again. For those who don’t know what green coffee is, it is unroasted, raw coffee beans. Their extract is popular in many dietary supplements that started to increase in popularity in the 1990’s. The drop of the green coffee market amounts aligns with investment drops during these two time periods. With overall consumption slowing, especially of items not deemed as a necessity, countries who had stakes in green coffee production slowed investment and thus had less inventory.

A picture containing text, screenshot, plot, diagram

Description automatically generated

The final graph is a layered graph containing the data from the domestic and importer consumption as well as the green coffee amounts. As shown on the graph, importer consumption of coffee is double that of domestic/producer country consumption, while green coffee remains at the lowest level throughout the years. Overall, there is a set pattern of consumption increase from 1990 to 2019. This could be due to the increase in the population, the widespread expansion of popular coffee chains like Starbucks and Dunkin, and the new fad around drinking coffee. Importer consumption remains high as countries who import coffee can afford more to consume as well as having the consumption of coffee present as a cultural entity. The increase in domestic consumption shows logistic trends like an economic increase in funds and resources, allowing for more production, thus a surplus in coffee that can be sold cheaply domestically. And for green coffee, at a larger scale, and compared to consumption, does not show as drastic of an increase, and has larger drops. This is due to green coffee being used for dietary supplements, which aren’t cheap nor are readily consumed by the public in the same way standard coffee is. It also makes it more susceptible to changes in economic markets as it is a riskier investment than the more readily and reliably consumed typical roasted coffee. But overall, each graph shows an increase and shows susceptibility to economic changes and historical events.

A picture containing text, screenshot, line, plot

Description automatically generated

# Conclusion

Behind tea, coffee is the most consumed beverage globally. And after compiling the output in the data analysis section, three things were found; countries consume more of the coffee type they produce, after production, exportation, and consumption there is still coffee left over, and coffee consumption and green coffee inventory has increased from 1990 – 2019 and can be helped or hindered by the rest of the global economy.

Economically, it is cheaper to consumer what is produced domestically. And this economic theory is proven true as the countries who produced either Robusta or Arabica coffee also consumed those coffee types. Economics also play a role in consumption, as drips or sharp increases in global consumption followed major historical events, such as Y2K and the end of the Cold War. And economic slowing of consumption results in decreases in investment and thus less inventory in green coffee.

But to me, the most important finding was that even after domestic consumption, exportation, importer consumption, and re-exportation, there is still a leftover supply of coffee. This leftover coffee supply has led to increases in greenhouse gas emissions and is a contributing factor to global warming. But small businesses and country-wide initiatives are thinking of ways to use that leftover coffee. But at the end of the day, coffee has become such a large and important part of our life that we are producing more than is demanded, subsequentially harming our planet. Although recycling coffee into raw materials and eco-friendly products does help some, an economic change in the production of coffee needs to be looked at to decrease the waste. And the ending suggestion that I have is that since coffee is such an important aspect of our daily lives, its price could be increased. This would allow producing countries to still make an economic profit off coffee exportation while simultaneously decreasing the amount that is produced, thus decreasing waste. This, along with other ideas, is something that the International Coffee Organization should invest in researching to not only benefit the global economy, but to also combat waste and greenhouse gas emissions.