Leveraging Data Analysis to Boost Coffee Chains' Performance
By: Valentina Torres

Coffee holds a vital place in our daily lives and global economy. With over 2.25 billion cups consumed daily, coffee's popularity creates employment opportunities and contributes significantly to coffee-producing countries. However, in today's competitive market, coffee chains like Starbucks and Dunkin' Donuts need to understand consumer preferences and behavior to improve their performance.

Problem and Data used

To address this challenge, this research paper focuses on leveraging data analysis to provide valuable recommendations to big coffee chains. By analyzing consumer reviews and behavior, we aim to uncover key insights that will empower these chains to enhance customer satisfaction, optimize pricing strategies, and tailor advertising campaigns effectively. Utilizing various data analytics techniques, such as Exploratory Data Analysis (EDA) to identify patterns, Pricing Analysis to understand pricing trends, and Sentiment Analysis to extract valuable feedback from reviews, we strive to offer actionable insights to boost the overall performance of coffee chains in the dynamic and evolving coffee industry.

The selected dataset for analysis originates from coffeereview.com, a website that offers thousands of coffee and espresso reviews. The dataset includes essential columns like coffee origin, price per 100 grams in USD, rating on a scale of 0 to 100, review date, and the actual review itself. The dataset was sourced from Kaggle, providing valuable insights into customer preferences and market trends. The source of the dataset is Kaggle and can be found in the references section.

Steps and Results

Exploratory Data Analysis (EDA)

Through descriptive statistics, visualizations, and correlation analysis, EDA helps us understand patterns, trends, and factors influencing coffee sales and customer preferences. Some of the results evidence from the data shows the analysis on coffee roasters, origins and their relationships to price and ratings.

The most frequent roasters in the dataset are Kakalove Café, JBC Coffee Roasters, and Paradise Roasters, which collectively make up a substantial portion of the reviewed coffees. This popularity among consumers can be attributed to the consistent delivery of high-quality coffee products and positive word-of-mouth recommendations. These roasters have likely built strong reputations over time, leading to increased customer loyalty and trust. Coffee enthusiasts often seek out these well-established brands for a reliable and enjoyable coffee experience, contributing to their prominence in the dataset.

Among the coffee origins reviewed, Ethiopia, Colombia, and Kenya emerge as the most common countries of coffee production. This prominence is not only a reflection of these countries' coffee-growing traditions but also indicative of the distinctive taste profiles associated with their unique coffee-growing conditions. For instance, Ethiopia is renowned for its diverse coffee varieties and natural processing methods, leading to a wide range of flavor notes. Similarly, Colombia's high-altitude coffee plantations and meticulous processing techniques contribute to its reputation for producing well-balanced and flavorful coffees. The popularity of Kenya as a coffee origin is often linked to its bright acidity and complex fruity flavors.

Consumers' preference for these origins demonstrates a growing interest in exploring coffee varieties with specific regional characteristics.
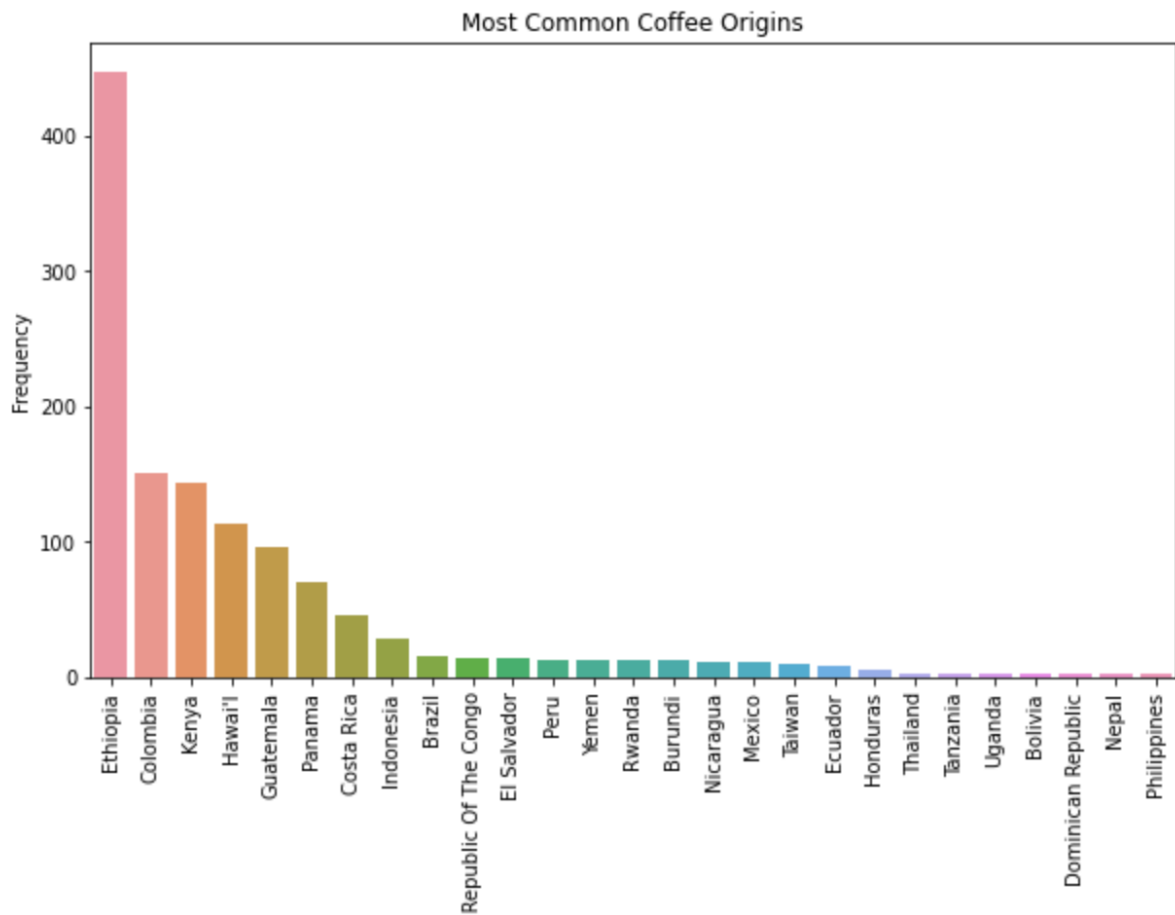
The relationship between origin and price of the coffee highlights significant variations in pricing. For example, countries such as Panama, often command premium prices for their coffee offerings. This could be attributed to rare and unique coffee varieties found in these regions, contributing to occasional high-priced coffees. On the other hand, countries like the Democratic Republic Of The Congo and Thailand, exhibit more consistent pricing. This consistency may be due to their focus on producing coffees with a stable and reliable taste profile.

On the other hand, the relationship between origin and rating shows variations in coffee quality and consumer preferences across different countries. Countries such as Colombia, Indonesia, and Ethiopia, demonstrate a wider range of coffee qualities and flavors, leading to diverse ratings. This could be due to factors such as varying coffee processing methods, growing conditions, and roasting techniques. In contrast, countries with consistent ratings, like Tanzania, may have more uniform coffee quality and flavor profiles, making them reliable choices for consumers seeking consistent taste experiences. These findings provide valuable insights into the diversity and uniqueness of coffee qualities worldwide, helping coffee chains curate their offerings to match regional preferences and cater to diverse consumer tastes.

Pricing Analysis

To start with, I analyzed the distribution of coffee prices in the dataset where I could evidence that the majority of coffee prices fall within the lower price range. This finding indicates that affordability is a significant factor for most consumers, as they tend to prefer reasonably priced options for their daily coffee consumption. On the other hand, the smaller number of higher-priced coffees suggests that premium-priced coffees form a smaller market segment. Consumers in this segment are often willing to pay more for unique and exclusive coffee experiences, such as specialty and single-origin coffees.

Additionally, the analysis on the relationship between ratings and price shows a positive correlation between coffee ratings and prices, indicating that higher-rated coffees tend to have higher prices. This correlation can be attributed to consumers' perception of better quality associated with higher-rated coffees, leading to increased demand and willingness to pay a premium type. As mentioned before, coffee enthusiasts seeking exceptional and memorable coffee experiences are often willing to pay higher prices for coffees with higher ratings, viewing them as more valuable and satisfying choices.

I also wanted to see how price behaved before, during and after the COVID-19 pandemic and evidence that during this period of time prices varied the most. This fluctuation can be attributed to several factors. During this period, coffee shops faced significant challenges, such as decreased sales and increased operational costs, leading to adjustments in pricing strategies. Supply chain disruptions and uncertainties faced by producers and retailers in predicting consumer demand and adjusting their pricing accordingly also affected coffee availability. The presence of more pricing outliers suggests that coffee chains experimented with different pricing approaches to stay competitive during the challenging economic conditions.


Comparison of Prices Before, During, and After COVID-19
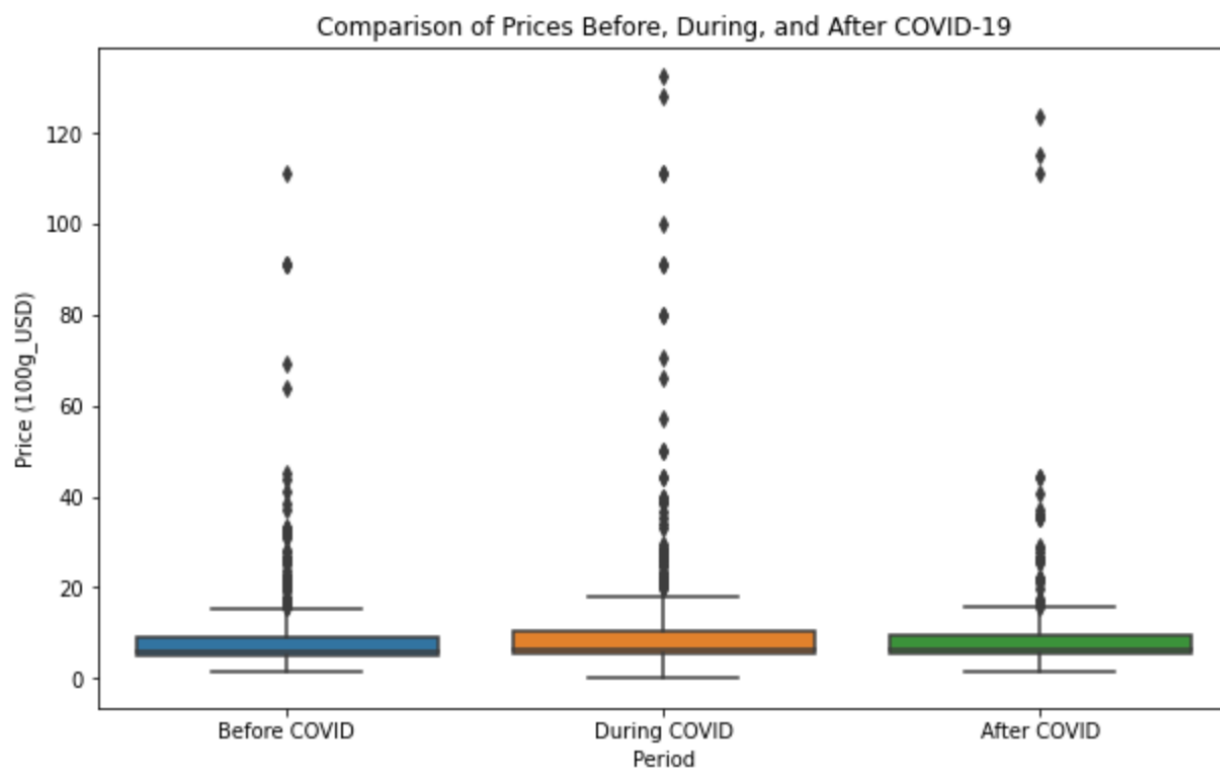
Reviews and Sentiment Analysis

Initially, I counted the number of reviews over time to see if there was any seasonal patterns and trends in coffee consumption and review activity. Periodic spikes and dips in review counts suggest seasonal variations in coffee consumption, which may be influenced by factors such as harvest seasons, temperature, and cultural habits. For example, review counts increased during the colder months which could be because consumers seek more warm beverages, including coffee. The analysis reveals fluctuations in the number of reviews and mean ratings across different months, suggesting that seasonal factors may influence consumer perceptions of coffee quality. Coffee enthusiasts may have varying taste preferences during different times of the year, leading to fluctuations in coffee ratings. For example, consumers may prefer lighter and fruitier coffee profiles during warmer months, while opting for bolder and richer flavors in colder seasons.

When looking at the different roast types, the analysis reveals that light and medium-light roasts generally receive higher ratings compared to dark and medium-dark roasts. This preference for lighter roasts may be due to their milder flavors, which appeal to a broader range of coffee enthusiasts. Lighter roasts often preserve the unique flavors and aromas of the coffee beans, allowing consumers to appreciate the distinctive characteristics of different coffee origins. On the other hand, darker roasts may have a more pronounced smoky or bitter taste, which may appeal to specific consumers but not resonate with a broader audience.

Furthermore, I wanted to conduct a deeper analysis for the COVID-19 time period, which showed reduced rating variation and more critical reviews. The reduction in rating variation can be attributed to limited coffee exploration opportunities during lockdowns. As consumers had fewer chances to try new coffee types or visit different coffee shops, they may have become more conservative in their ratings, leading to a narrower range of scores. The increase in critical reviews may be a consequence of consumers spending more time at home, which allowed them to engage in more detailed evaluations. With additional time for reflection and assessment, consumers may have become more discerning in their reviews, leading to a higher number of critical assessments.
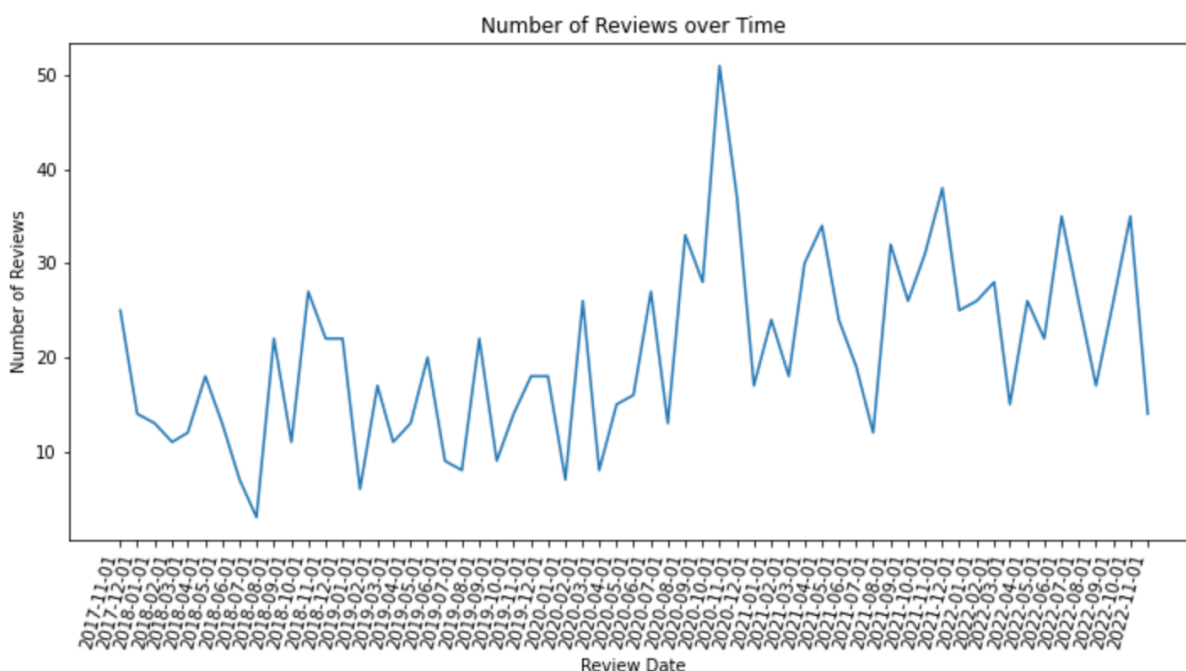
Finally, I wanted to analyze the text from the review's column. The study revealed frequently mentioned words in customer reviews, such as aroma, mouthfeel, acidity, and chocolate. These specific keywords indicate that consumers focus on various sensory aspects of the coffee experience beyond just taste. The emphasis on aroma suggests that consumers value the fragrance and aroma that coffee imparts, making it an essential element of the coffee-drinking experience. Additionally, words like mouthfeel and acidity indicate that consumers pay attention to the physical sensations and taste attributes of the coffee, seeking well-balanced and enjoyable drinking experiences. The mention of chocolate highlights that consumers appreciate coffee with chocolatey undertones, a flavor profile often associated with certain coffee origins or roast levels. Furthermore, the presence of words like syrup, almond, lemon, baking, and tart indicates that consumers enjoy pairing coffee with various food items, such as desserts or baked goods, which enriches their overall coffee-drinking experience.

Recommendations

To boost their performance, big coffee chains should capitalize on the popularity of top roasters such as Kakalove Café, JBC Coffee Roasters, and Paradise Roasters. Collaborating with these well-established brands will improve the chains' reputation, trust and loyalty among customers, ultimately driving sales. Additionally, showcasing coffee origins from Ethiopia, Colombia, and Kenya can attract coffee enthusiasts seeking unique and diverse flavor experiences. Chains can introduce limited-time offerings or specialty coffee menus featuring coffee from these regions, providing consumers with a taste journey around the world and tailoring marketing efforts to specific customer segments.

Moreover, Pricing Analysis insights can help optimize the chains' pricing strategies. By offering a range of prices, they can cater to a wider customer base, including budget-conscious consumers and those willing to pay more for exceptional experiences. Smart pricing decisions will attract both value-seeking customers and enthusiasts seeking exclusive and highly-rated coffee varieties. To enhance the customer experience, incorporating consumer feedback from Sentiment Analysis is key. Big coffee chains can use this data to improve their offerings based on specific customer preferences and expectations. Investing in training baristas to focus on aroma, mouthfeel, acidity, and other sensory aspects will improve the customer´s perception.

To stay competitive, coffee chains should also introduce seasonal promotions and limited-edition coffee blends that align with consumer preferences. By offering a variety of roast profiles, including light and medium-light roasts that receive higher ratings, they can reach a broader audience. Addressing customer concerns promptly and prioritizing customer satisfaction, especially during challenging times like COVID-19, will foster loyalty and trust. Emphasizing aromas and sensory experiences in marketing campaigns can differentiate the chains. Furthermore, coffee chains can add desserts to their menus based on frequently mentioned words in reviews.

To remain relevant and successful, big coffee chains must continuously monitor changing trends and consumer preferences through reviews and seasonal patterns. Adapting to evolving tastes and preferences will enable chains to offer unique and personalized experiences, ensuring long-term success in the dynamic coffee industry. By implementing these recommendations, coffee chains like Starbucks and Dunkin' Donuts can gain a competitive edge, increase customer satisfaction, and drive revenue growth, ultimately solidifying their position as leaders in the coffee market.

Future work

In the pursuit of continuous improvement and refining the recommendations for big coffee chains, there are several strategies to explore with more time and resources. One key area to focus on is cost optimization recommendations. While I analyzed pricing trends, a more detailed cost analysis could provide valuable insights into cost structures, production efficiencies, and supply chain management.

Unfortunately, the original dataset lacked the necessary information for an in-depth cost analysis, and no additional datasets were available for supplementation. However, with access to more comprehensive and descriptive data related to production costs, labor, and supply chain logistics, I could provide detailed cost optimization strategies to maximize profitability while delivering quality products to consumers.

Additionally, future work could be done to customer segmentation and behavior analysis to tailor marketing and advertising campaigns to specific customer preferences and demographics, further increasing customer engagement and loyalty. By exploring these avenues, I can continue to fine-tune our recommendations and offer coffee chains comprehensive strategies for sustained success in the competitive coffee market.

References

Schmoyote. (2021). Coffee Reviews Dataset [Data set]. Kaggle. URL: https://www.kaggle.com/datasets/schmoyote/coffee-reviews-dataset

Castellana, F., De Nucci, S., De Pergola, G., Di Chito, M., Lisco, G., Triggiani, V., Sardone, R., & Zupo, R. (2021). Trends in Coffee and Tea Consumption during the COVID-19 Pandemic. National Library of Medecine. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8535644/

Cadwalader Z. (2022, November 7). Scientists Discover How Coffee May Be Inhibiting COVID-19 Infection. Sprudge. Retrieved from https://sprudge.com/scientists-discover-how-coffee-may-be-inhibiting-covid-19-infection-193595.html

Smolokoff, A. (2021, June 09). Pandemic Shakes Coffee Market. Food Beverage Insider. Retrieved from https://www.foodbeverageinsider.com/beverages/pandemic-shakes-coffee-market

Ting, H., & Thurasamy, R. (2016, May 17). What matters to infrequent customers: a pragmatic approach to understanding perceived value and intention to revisit trendy coffee café. SpringerPlus. Retrieved from https://springerplus.springeropen.com/articles/10.1186/s40064-016-2259-5