

### ABSTRACT

In this data-driven study, we explore the global coffee industry by examining the relationship between coffee flavor profiles and their geographical origins. Using various visualizations, such as choropleth plots, line chart, bubble plots, radar plots, and chord diagrams, we identify leading coffee-producing countries, distinctive flavor attributes, and intricate connections between coffee varieties and their countries of origin. This comprehensive analysis enables coffee enthusiasts to expand their knowledge and enrich their tasting experiences. The visualizations and insights generated by this research provide valuable guidance for both experienced coffee connoisseurs and newcomers alike, enhancing their understanding of the diverse world of coffee and its fascinating interplay between geography and flavor.

### INTRODUCTION

In recent years, the global coffee industry has experienced a surge in consumer interest, with a heightened focus on understanding the intricate flavors and diverse origins of coffee beans. This report delves into the complexities of coffee production, examining the relationships between the geographical origins of coffee beans, their flavor profiles, and the varieties they produce. By analyzing data from a worldwide coffee tasting competition, we aim to offer a comprehensive understanding of the coffee landscape, enabling consumers to make informed decisions about their coffee preferences.

### RESEARCH QUESTIONS

- What are the key flavor attributes that distinguish the highest-rated coffee-producing countries, and how do these attributes relate to the specific geographical regions where the coffee beans are grown?
- How do the different coffee varieties contribute to the flavor profiles of coffee beans, and what is the distribution of these varieties across various coffee-producing countries?
- Can we identify clusters or patterns in coffee flavors and their geographical origins, which may provide insights into the regional characteristics of coffee cultivation and processing techniques?

### METHODOLOGIES

Figure 1 showcases a choropleth map that vividly illustrates the geographic distribution of coffee production across continents and years, offering a comprehensive view of the global coffee landscape. Created using Plotly library, the engaging color scheme represents various production levels, enabling users to quickly comprehend trends, regional differences, and the impact of local factors on coffee production. Interactive features like sliders and buttons let users explore data at their own pace, allowing them to focus on specific years, continents, or even individual countries, resulting in a personalized experience that enhances their understanding of the complex coffee industry.<sup>[1]</sup>

### METHODOLOGIES

For Figure 2, the line chart unravels the story of coffee production over time. It highlights the nuances of each continent's production level for each year. The synergy of lines and marks creates a visually engaging narrative, succinctly capturing the evolution of production across years. Designed with the user in mind, the line chart offers an accessible and captivating insight into the historical trends of coffee production.<sup>[1]</sup>

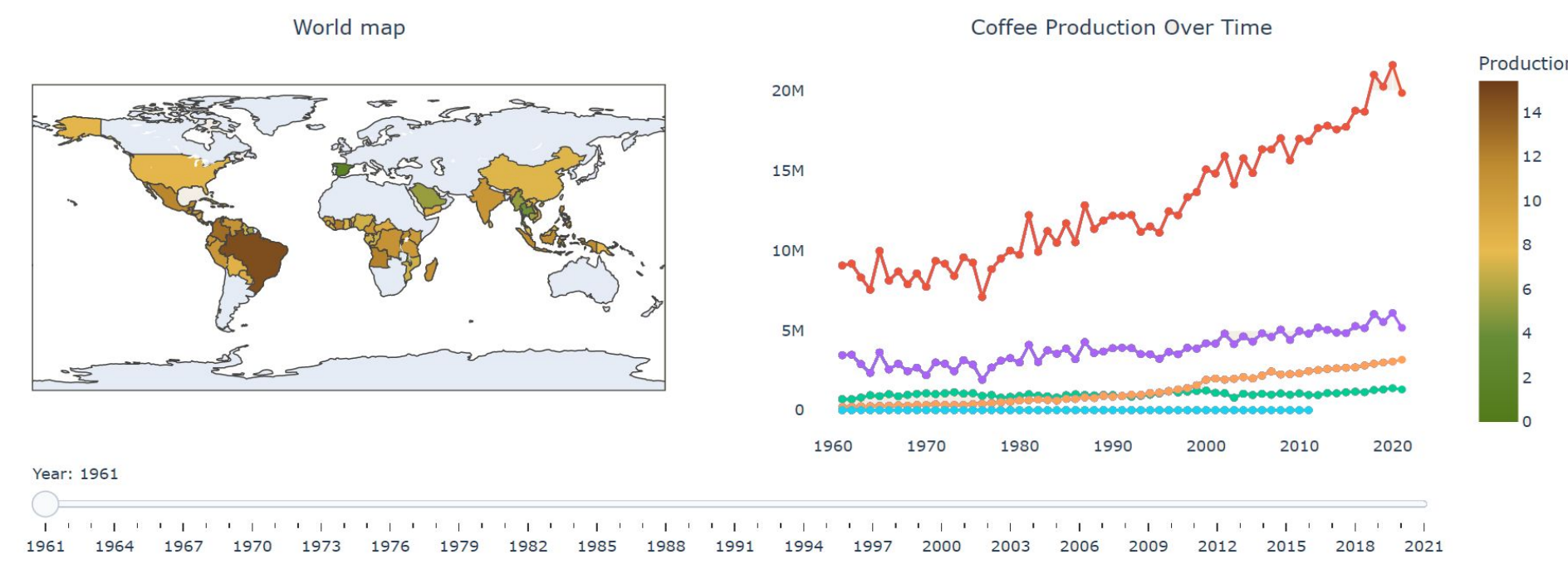


Figure 5 presents an interactive chord diagram that visualizes the relationships between countries and coffee varieties based on total coffee bean reviews. Created using the D3.js library, this dynamic diagram displays outer arcs representing countries and varieties, while inner chords signify the connections between them. Interactive features like highlighting related arcs and chords on mouseover events enhance user engagement. The diagram offers a unique perspective on the interplay between coffee-producing countries and their corresponding bean varieties, helping users gain insights into their preferences.<sup>[2]</sup>

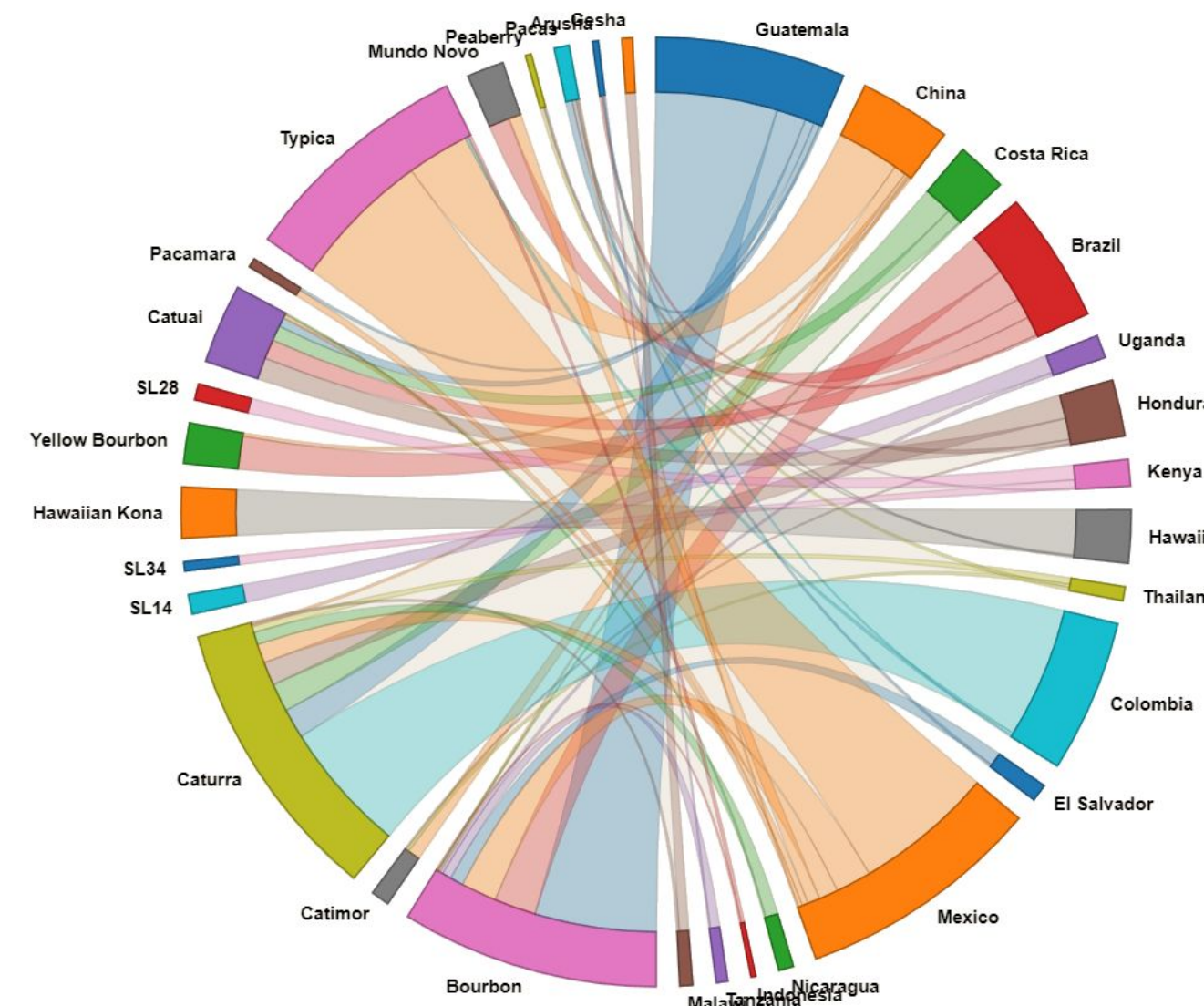


Figure 5: The interactive chord diagram offers a fascinating lens to explore the intriguing relationships between coffee varieties and their countries of origin. This captivating representation waves vast swaths of information together and transforming raw data into a visual narrative that invites the viewer on an enthralling journey through the world of coffee.

### METHODOLOGIES

Figure 3 presents an interactive bubble chart, depicting the coffee quality of various countries based on total cup points. Crafted using the D3.js library, this dynamic visualization features bubbles of different sizes, representing each country's total number of cups, with national flags as backgrounds. Hovering over a bubble reveals the country's ID and cup points. The force simulation ensures that bubbles move according to attraction towards the axes and collision avoidance.



Figure 3: Shows information about different countries' coffee quality based on their total cup points (calculated by summing all favor tasting measurements)

Figure 4 showcases a radar plot, created using Python's Plotly library, to illustrate data across categories like Aroma, Flavor, Aftertaste, Acidity, Body, Balance, Uniformity, and Cup Cleanliness for specific countries. The input includes a dataframe and a value 'k' that determines the row to be used. The radar plot utilizes Plotly's go.Scatterpolar function, setting data, colors, and labels accordingly. The update\_layout function establishes polar coordinates, radial and angular axis ranges and labels, background color, and plot title, resulting in an informative and visually appealing chart.<sup>[3]</sup>

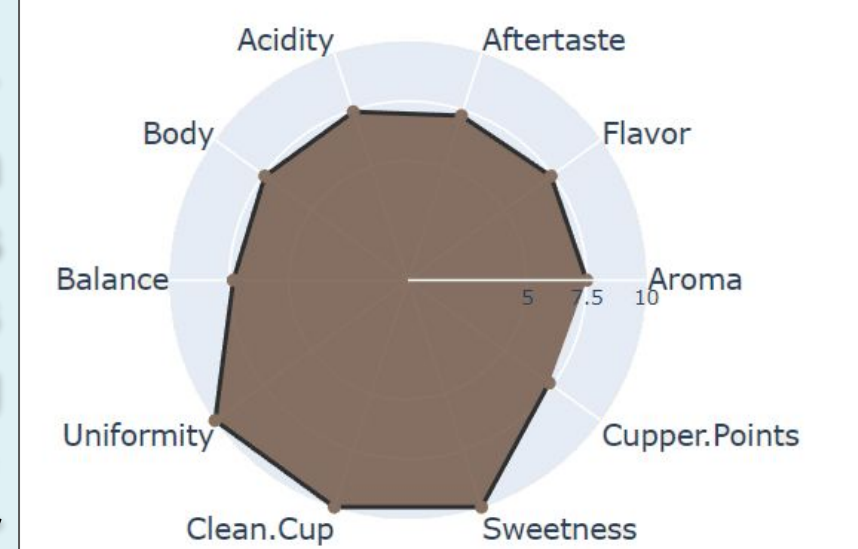


Figure 4: The radar plot is used to explore the flavor profile of different coffee varieties. The area indicates the overall intensity of the flavor.

### CONCLUSIONS

In conclusion, this research project has successfully delved into the diversity of coffee production by utilizing interactive visualizations to examine the relationships between geographical origins, flavor profiles, and coffee bean varieties. Our findings have unveiled key patterns in coffee flavors, their geographical origins, and regional characteristics of coffee cultivation and processing techniques. These insights, derived from engaging visualizations like choropleth maps, line charts, chord diagrams, bubble charts, and radar plots, can significantly enhance consumer understanding of coffee origins and flavors. This, in turn, empowers them to make more informed decisions regarding their coffee preferences. Furthermore, our findings can serve as valuable information for coffee producers and industry stakeholders, allowing them to improve their products and capitalize on emerging trends. Overall, this research project contributes meaningfully to the ongoing conversation around coffee production and its intricate landscape, paving the way for further exploration and development within the global coffee industry.

### Contact Information

Jinsheng Luo: [j12912@georgetown.edu](mailto:j12912@georgetown.edu)  
Yihan Bian: [yb214@georgetown.edu](mailto:yb214@georgetown.edu)  
Chang Xie: [cx89@georgetown.edu](mailto:cx89@georgetown.edu)

### References

1. Food and Agriculture Organization of the United Nations, "Coffee, green- Production", Our World in Data, 2023-02-22, <https://ourworldindata.org/grapher/coffee-bean-production?tab=table&time=earliest..2020>
2. Coffee Quality database from CQI, Kaggle, <https://www.kaggle.com/datasets/volpato/coffee-quality-database-from-cqi>
3. Coffee Reviews Dataset, Kaggle, <https://www.kaggle.com/datasets/schmoyote/coffee-reviews-dataset?resource=download>