

A close-up photograph showing a person's hands sorting through a large pile of coffee beans. The beans are spread out on a wooden tray, and the person's hands are visible, with one hand resting on the beans. The beans are a mix of light and dark colors, indicating different stages of roasting or processing. The background is slightly blurred, focusing attention on the hands and the beans.

W200 Final Project | Arabica Coffee Exploration

Bailey Kuehl | Elias Tavaréz | Michael Thottam

December 5th, 2022

The Impact of Coffee on Society

Coffee /'kôfē,'käfē/

A drink produced from the roasted and ground seeds of the coffea plant

40%

Estimated percentage of the world's population that drinks coffee every day

2,000,000,000

The estimated number of coffee cups drank per day

Dataset



Coffee Harvested from **2008-2018**

	owner	country_of_origin	processing_method	flavor	aftertaste	acidity	body	balance	uniformity	clean_cup	sweetness	total_cup_points	mean_altitude_feet
0	metad plc	Ethiopia	Washed / Wet	8.83	8.67	8.75	8.50	8.42	10.00	10.00	10.00	90.58	6807.743000
1	metad plc	Ethiopia	Washed / Wet	8.67	8.50	8.58	8.42	8.42	10.00	10.00	10.00	89.92	6807.743000
2	yidnekachew dabessa	Ethiopia	Natural / Dry	8.58	8.42	8.42	8.50	8.25	10.00	10.00	10.00	89.00	6561.680000
3	metad plc	Ethiopia	Washed / Wet	8.50	8.25	8.50	8.42	8.33	10.00	10.00	10.00	88.83	6807.743000
4	ji-ae ahn	Brazil	Natural / Dry	8.42	8.42	8.50	8.25	8.33	10.00	10.00	10.00	88.83	2407.042947
...
1236	cadexsa	Honduras	Washed / Wet	6.33	6.17	6.50	6.67	6.17	8.00	8.00	8.00	69.17	4757.218000
1237	juan carlos garcia lopez	Mexico	Washed / Wet	6.83	6.25	7.42	7.25	6.75	10.00	0.00	10.00	68.33	2952.756000
1238	myriam kaplan-pasternak	Haiti	Natural / Dry	6.58	6.42	6.67	7.08	6.67	9.33	6.00	6.00	67.92	1148.294000
1239	exportadora atlantic, s.a.	Nicaragua	Other	6.58	6.33	6.25	6.42	6.08	6.00	6.00	6.00	63.08	3608.924000
1240	juan luis alvarado romero	Guatemala	Washed / Wet	6.67	6.67	7.67	7.33	6.67	8.00	1.33	1.33	59.83	1417.320000

Variables

Coffee Grade Variables

Variable	Aroma	Flavor	After-taste	Acidity	Balance	Body	Uniformity	Clean Cup	Cup Points	Sweetness
Desc.	Flavor-related properties about the coffee, graded for each review									

Farming / Metadata Variables

Variable	Farm name	Processing Method	Bag Weight	Grading Date	Country	Mean Altitude
Desc.	Production and ownership				Environment	

Cleaning the Data

Remove
Certification IDs,
Duplicate Columns

Original Dataset
43 columns
1311 observations



Transform
Altitude, Bag Weights,
Harvest Year



Outlier Removal
Coffees with
unsalvageable
scores/features



Final Dataset
36 columns
1241 observations

Impute
Processing Methods,
Farms, Colors, etc.

Questions

1

Are there **correlations** between how reviewers perceive coffee's various properties and how they score the coffee overall?

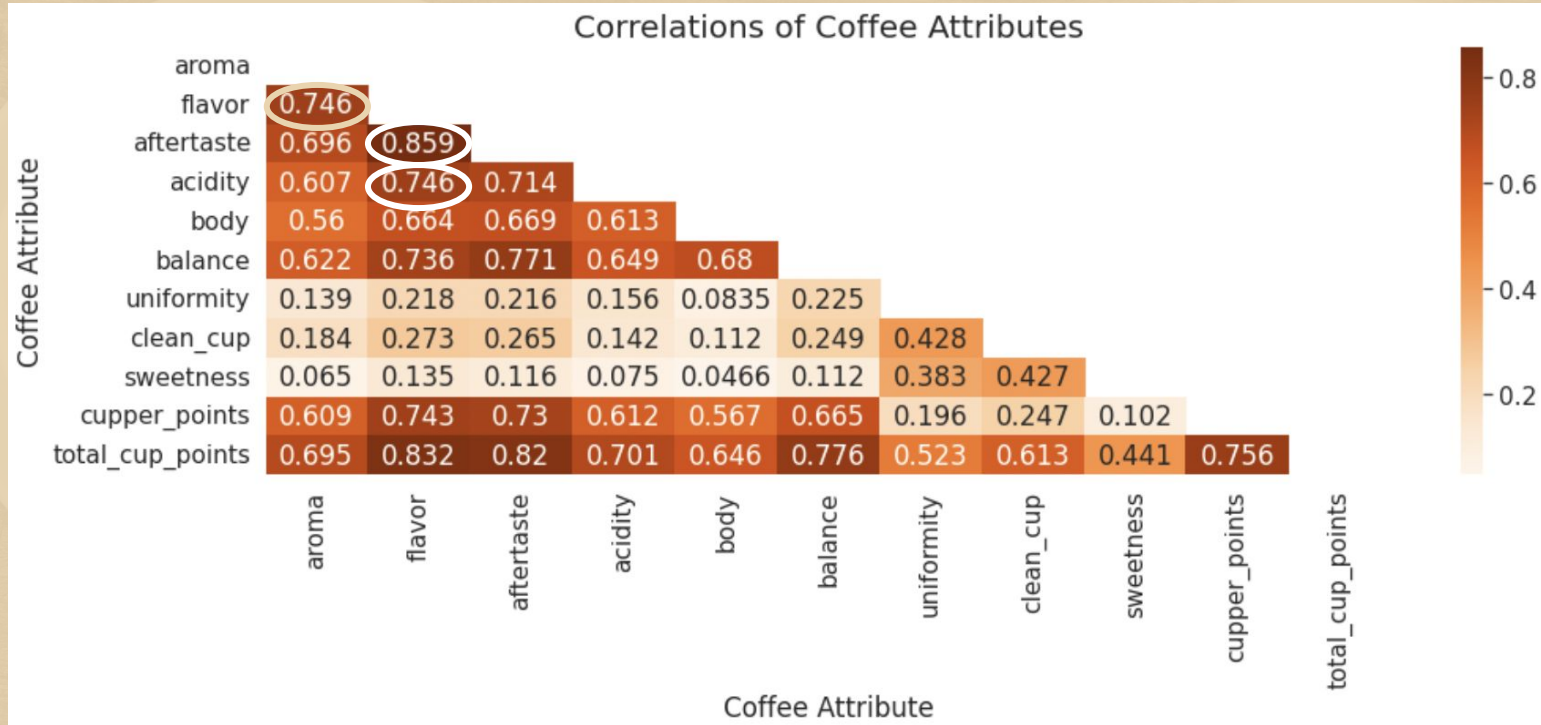
2

Does the **geographic location** (region, country) in which the coffee is grown and harvested affect the perceived quality of the coffee?

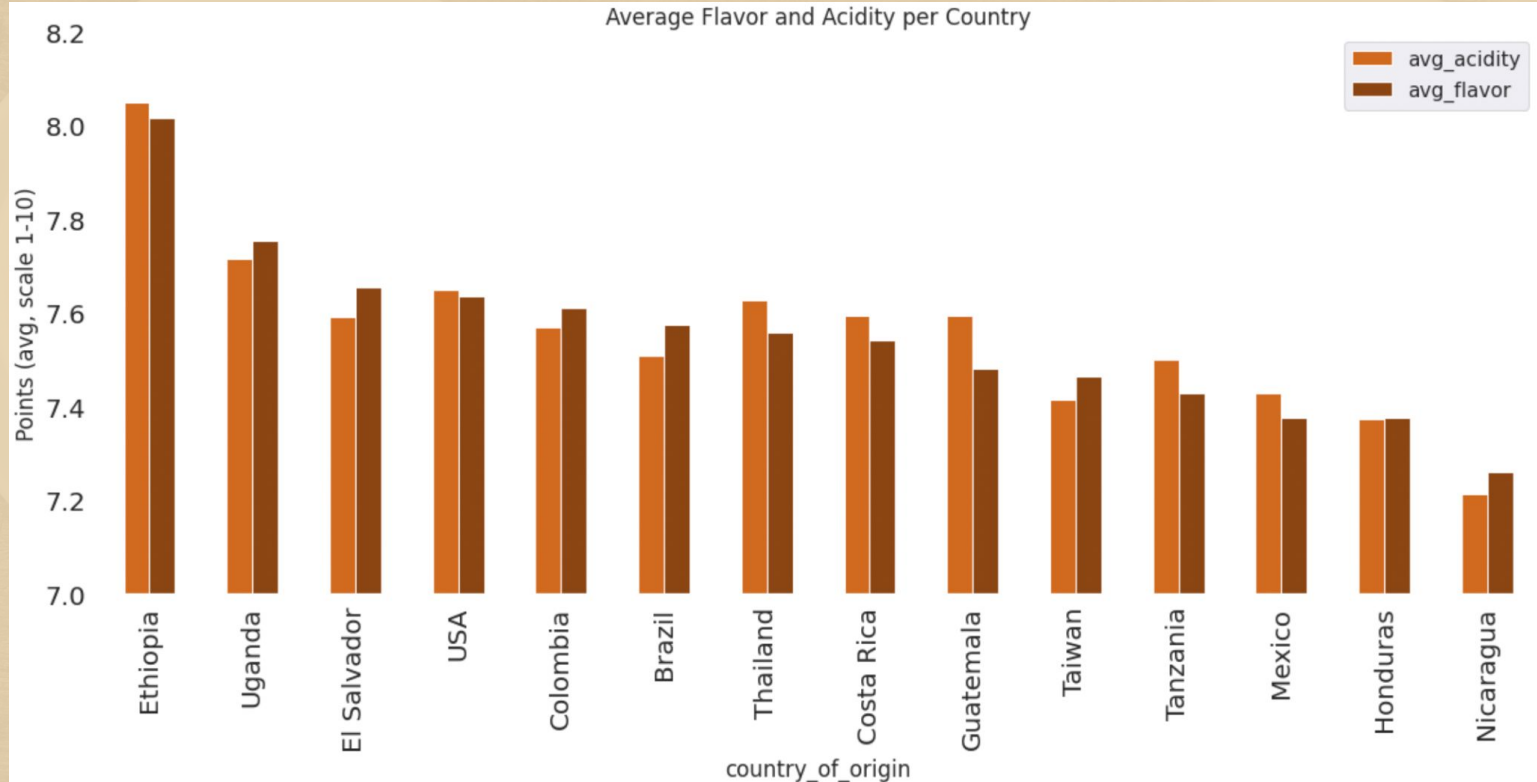
3

How does the **farm's owner and processing methods** affect the coffee's perceived flavor?

Coffee Characteristic Correlation: Key Figures

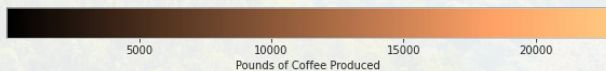


Coffee Characteristic Correlation: Key Figures

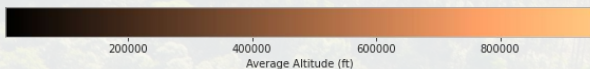


Geographical Features Key Plots/Figures

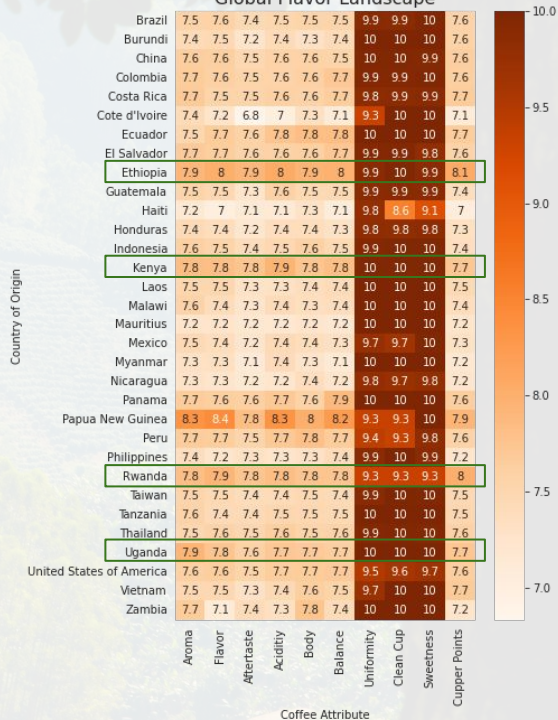
Arabica Coffee Production by Country



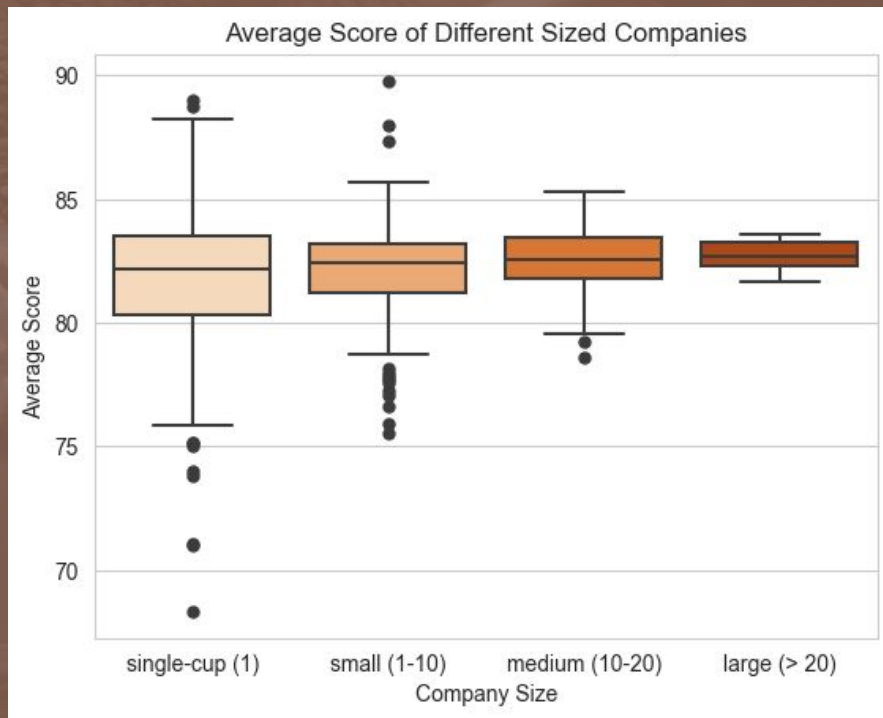
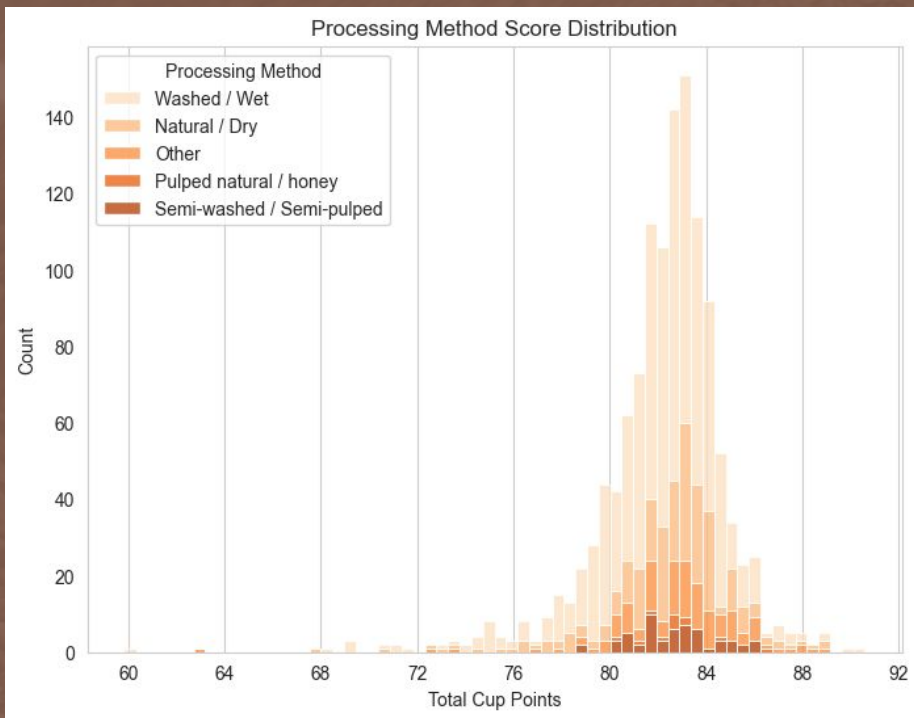
Average Altitude of Coffee Harvested



Global Flavor Landscape



Processing Methods, Farm Conditions: Key Figures



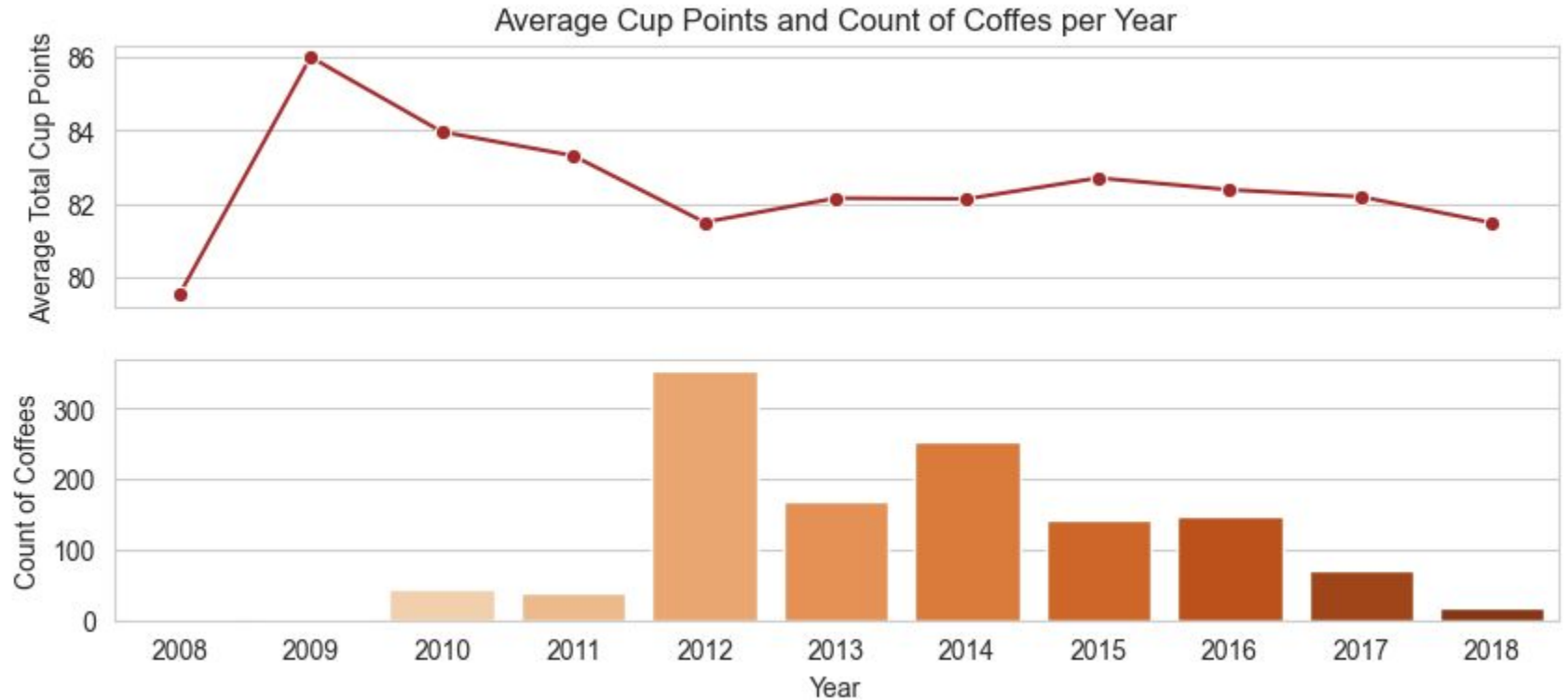
Conclusions

1. Several of the different coffee attributes reviewed likely overlapped with each other (**flavor & aroma**, as well as **flavor & aftertaste**)
2. **Ethiopia** has the **most flavorful** (and **highly acidic**) **coffee**, according to reviewers
3. While the **Americas** produce the **most coffee**, **Africa** tends to produce more uniquely **flavorful** coffee
4. There was no best **processing method**, they all centered around a similar score
5. **Number of coffee** cups a **company** had graded was not indicative of the average score they received

Questions?



Appendix Graph



Appendix Graph

	aroma	flavor	aftertaste	acidity	body	balance	uniformity	clean_cup	sweetness	cupper_points
processing_method										
Natural / Dry	7.59	7.58	7.46	7.55	7.59	7.59	9.79	9.80	9.82	7.58
Other	7.61	7.61	7.49	7.58	7.55	7.58	9.88	9.88	9.92	7.66
Pulped natural / honey	7.54	7.51	7.53	7.55	7.62	7.52	10.00	10.00	10.00	7.54
Semi-washed / Semi-pulped	7.60	7.59	7.46	7.54	7.55	7.58	9.87	9.95	9.95	7.54
Washed / Wet	7.55	7.48	7.36	7.53	7.49	7.48	9.85	9.83	9.93	7.44

	aroma	flavor	aftertaste	acidity	body	balance	uniformity	clean_cup	sweetness	cupper_points
processing_method										
Natural / Dry	8.58	8.67	8.50	8.50	8.50	8.58	10.00	10.00	10.00	8.67
Other	8.67	8.67	8.58	8.42	8.58	8.75	10.00	10.00	10.00	10.00
Pulped natural / honey	8.00	8.00	8.00	8.25	8.00	8.17	10.00	10.00	10.00	8.17
Semi-washed / Semi-pulped	8.50	8.17	8.00	8.08	8.33	8.25	10.00	10.00	10.00	8.42
Washed / Wet	8.75	8.83	8.67	8.75	8.50	8.58	10.00	10.00	10.00	8.75