

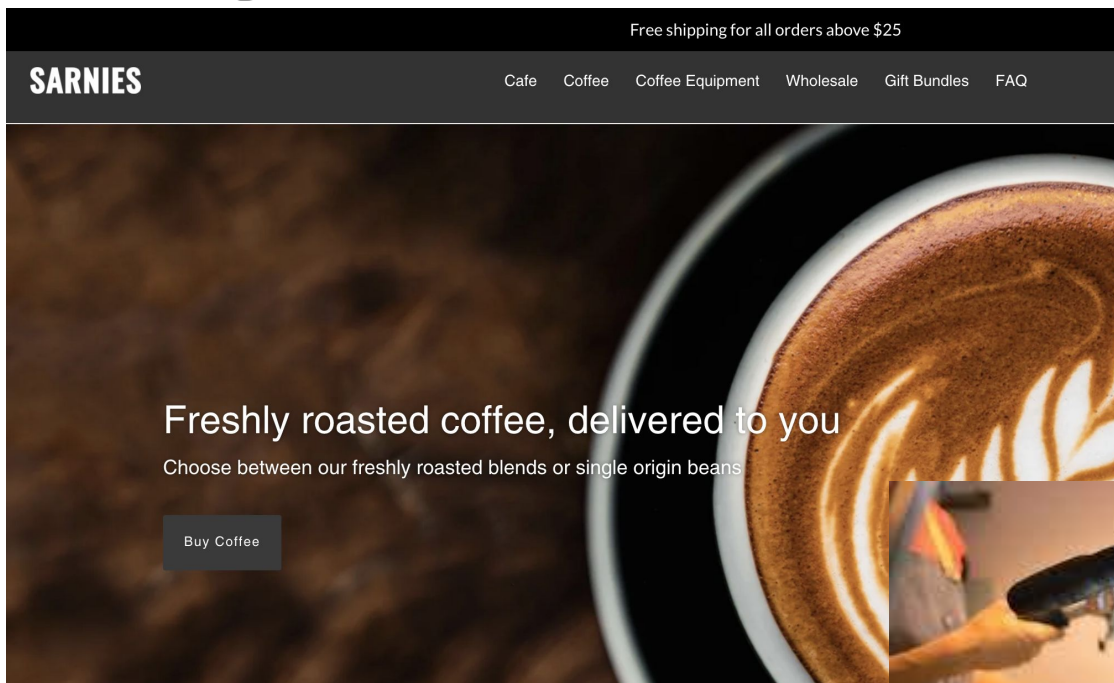


Coffee Bean Recommender

Cindy Soh



Background



Problem Statement

- Time consuming for user to search for coffee individually from different roasters website
- Restrict to one roasters subscription, lack of variety and option to many other roaster locally

List of roasters in locally:

1. Apartment Coffee (Subscription)
2. Thus Coffee (No Subscription)
3. Common Man Coffee Roasters (Online purchase)
4. Toby's Estate (Subscription)
5. Sarnies (Subscription)
6. Oriole Coffee + Bar (Online purchase)
7. Round Boy Roasters (Subscription)
8. Perk Coffee (Subscription)
9. Five Oars Coffee Roasters (Online purchase)
10. Nylon Coffee Roasters (Subscription)
11. And more



Ethiopia Guji 250gm

\$14.70

Shipping calculated at checkout.

Choose Grind Size

Whole Bean ▾

☐ One-time purchase: \$21.00

☒ Subscribe & save (30%): \$14.70

Deliver every 1 Week ▾

 Subscription details

Choose Grind Size (Individual Order)

☒ Whole Bean

Espresso

Filter

French Press

Moka Pot

Cold Brew

SIGN ME UP F

Subscriptions

Stay caffeinated - choose your preferred beans, how you want it ground and we'll take care of the rest.

Get a subscription



Problem Statement

- Subscription with one vendor
- No flexibility or change in variety

Subscriptions

Stay caffeinated - choose your preferred beans, how you want it ground and we'll take care of the rest.

Get a subscription



Ethiopia Guji 250gm

\$14.70

Shipping calculated at checkout.

Choose Grind Size

Whole Bean ▾

☐ One-time purchase: \$21.00

☒ Subscribe & save (30%): \$14.70

Deliver every 1 Week ▾

 Subscription details

Choose Grind Size (Individual Order)

✓ Whole Bean

Espresso

Filter

French Press

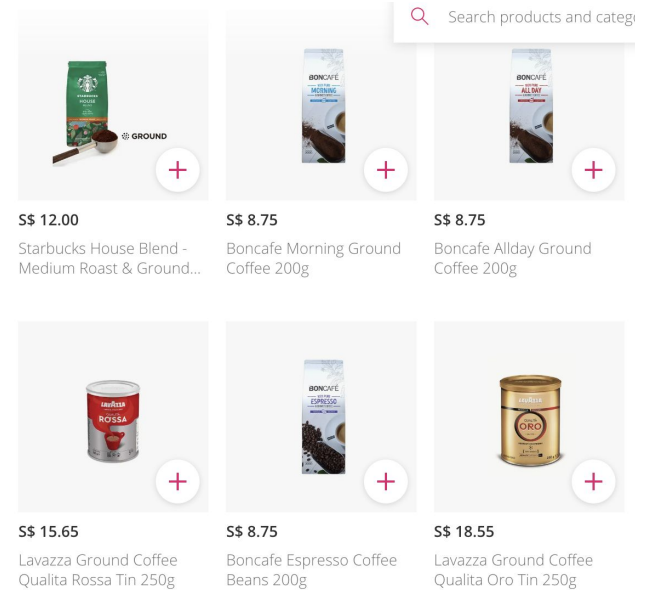
Moka Pot

Cold Brew

SIGN ME UP FOR

Business Proposal

- Platform for local roaster as vendors
- User with one subscription able to access to many local roaster and bean variety
- Monthly Surprise Box subscription for adventurous coffee brewers
- Recommender system:
 - For user selection of new weekly/monthly roast listed by roasters
 - Curation of surprise box subscription





BOKKSU
SNACK BOX

SHOP

ABOUT

BRANDS



Experience Japan From Home

Bokksu delivers to your door the experience of tasting authentic Japanese snacks, candies, and teas sourced directly from centuries-old family makers. Every box you purchase **supports these family-run businesses** in Japan and helps keep their traditions alive.

Each box ships from Japan and comes with:

Details



- **20-22 Japanese snacks**, candies, and a tea pairing curated around a fun cultural theme
- **24-Page Culture Guide Magazine** that details each product's origin, flavors, and even common allergens
- Auto-renew with the option to **pause** or **cancel** anytime
- Optional **free personalized e-gift message** that sends to your recipient upon order completion
- **Free shipping** directly from Japan!

Rating

93

Joe Bean Roasters

Finca Buenavista Edwin
Zambrano Colombia

Roaster Name

Name of coffee
bean

Roaster Location:	Rochester, New York	Review Date:	November 2022
Coffee Origin:	Hulla Department, Colombia	Aroma:	9
Roast Level:	Medium-Light	Acidity/Structure:	9
Agtron:	58/78	Body:	8
Est. Price:	\$19.00/12 ounces	Flavor:	9
		Aftertaste:	8

Main
Review

Blind Assessment

Fruit-forward, crisply tropical. Strawberry guava, baking chocolate, lemon verbena, fresh-cut oak, brown sugar in aroma and cup. Sweet in structure with balanced, lively acidity; plush, syrupy mouthfeel. Crisply cocoa-toned finish with undertones of lemon verbena.

Notes

Produced by Edwin Zambrano of Finca Buenavista from trees of the Castillo and Colombia varieties of Arabica, and processed by the natural method (dried in the whole fruit). Joe Bean Roasters is a specialty retail and wholesale coffee roaster based in Rochester, New York. Visit www.joebeanroasters.com for more information.

Information on the coffee
bean

Summary

Bottom Line

An inviting, cleanly fruit-toned, natural-processed Colombia cup with a crisp chocolate throughline.

Explore Similar Coffees

[Click here](#) for more reviews from **Joe Bean Roasters**

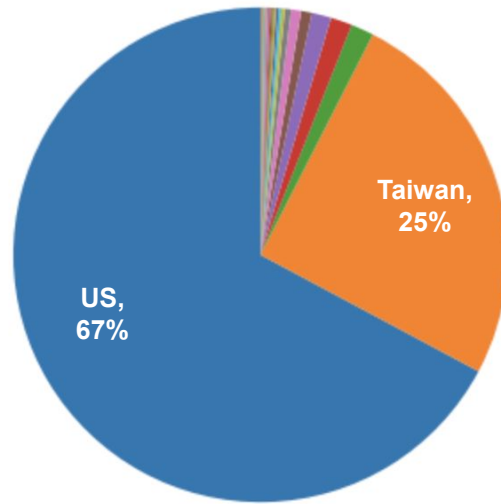
[Click here](#) for more information about coffees from **Colombia**

Data Dictionary

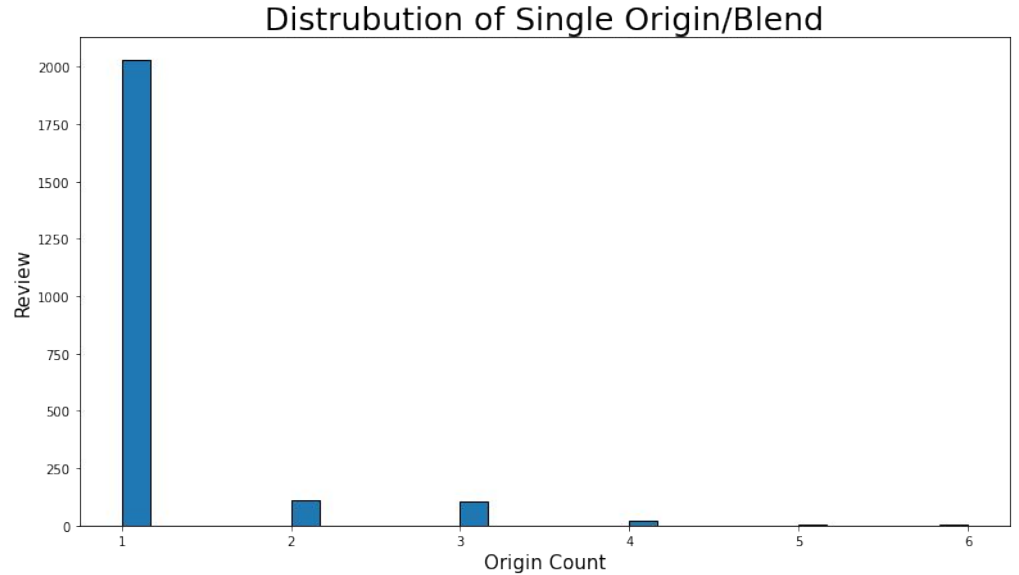
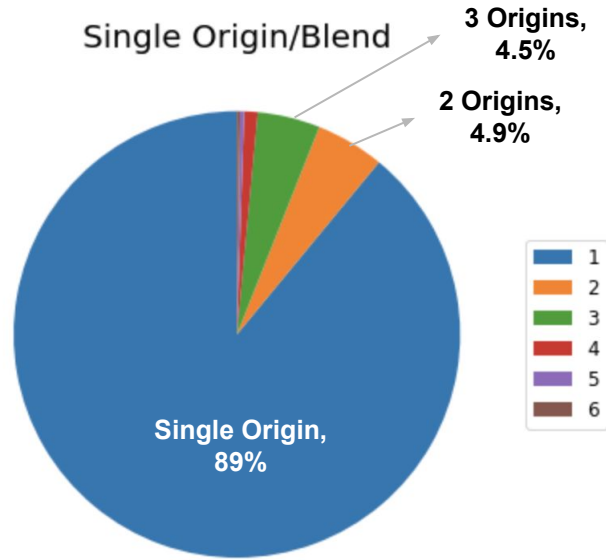
Column	Data Types	Description
rating	int64	Overall Rating of Coffee
roaster	object	Roaster Name
roaster_country_state	object	State/Country of Roaster
roaster_country	object	Country of Roaster
name	object	Name of coffee bean
roaster_city	object	City of Roaster
origin	object	Coffee Bean Origin
origin_count	int64	Coffee Bean Origin count
origin_country	object	Coffee Bean Origin Country
est_price	object	Estimated price in local currency
review_date	object	Date that review is posted
roast	object	Range from light to dark roast
agtron	object	Roast level of bean and after it is grinded
agtron_grounds	int64	Roast level of after it is grinded
agtron_bean	int64	Roast level of before it is grinded
aroma	float64	0-10 rating of coffee's aroma
acid	float64	0-10 rating of coffee's acidity
body	float64	0-10 rating of coffee's body
flavor	float64	0-10 rating of coffee's flavor
aftertaste	float64	0-10 rating of coffee's after taste
with_milk	float64	0-10 rating of coffee with milk
desc_1	object	Blind Assessment
desc_2	object	Information on coffee bean
desc_3	object	Summary

Exploratory Data Analysis

Roaster location countries

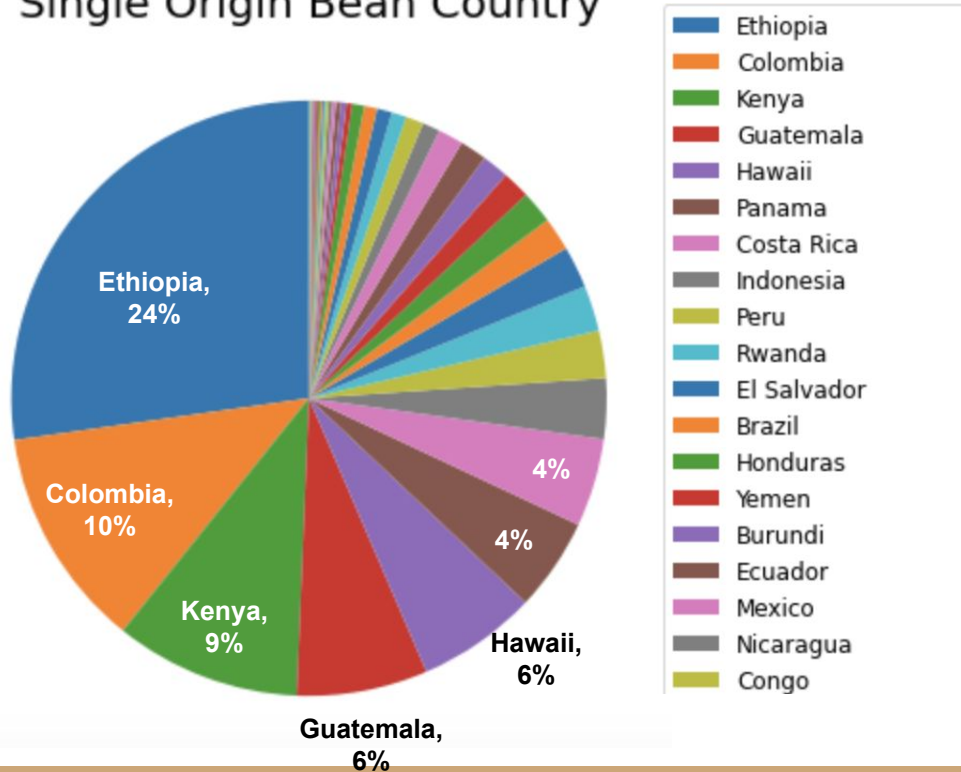


Number of country origin in each roast



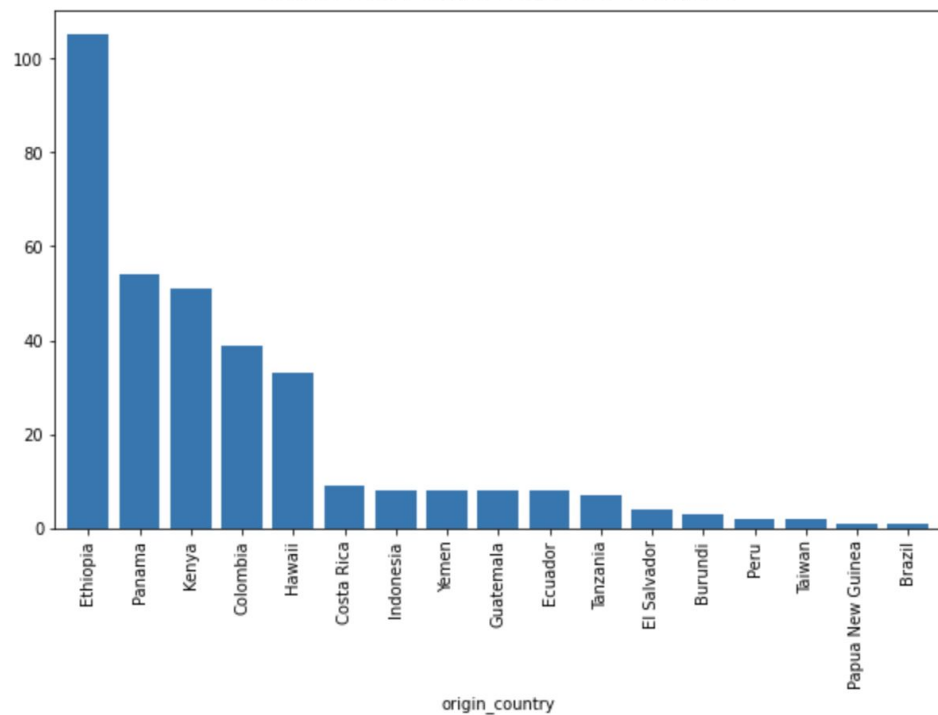
Top Country of Single Origin Coffee Bean

Single Origin Bean Country

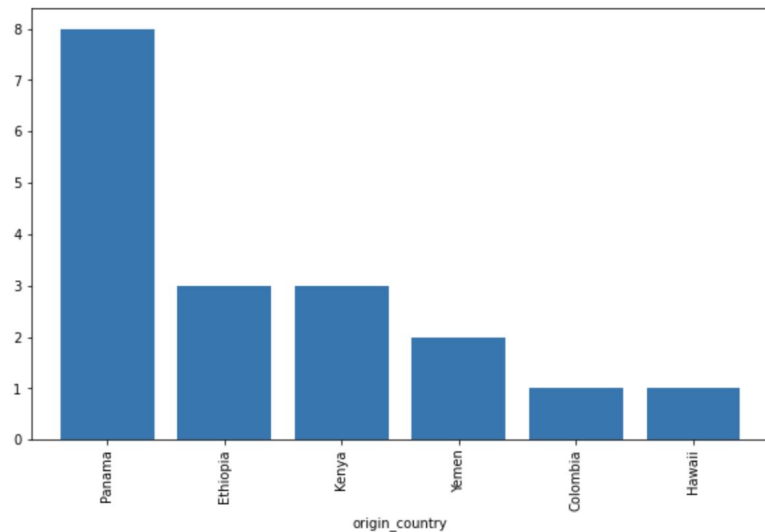


origin_country	
Ethiopia	540
Colombia	238
Kenya	200
Guatemala	141
Hawaii	129
Panama	100
Costa Rica	95
Indonesia	65
Peru	51
Rwanda	49

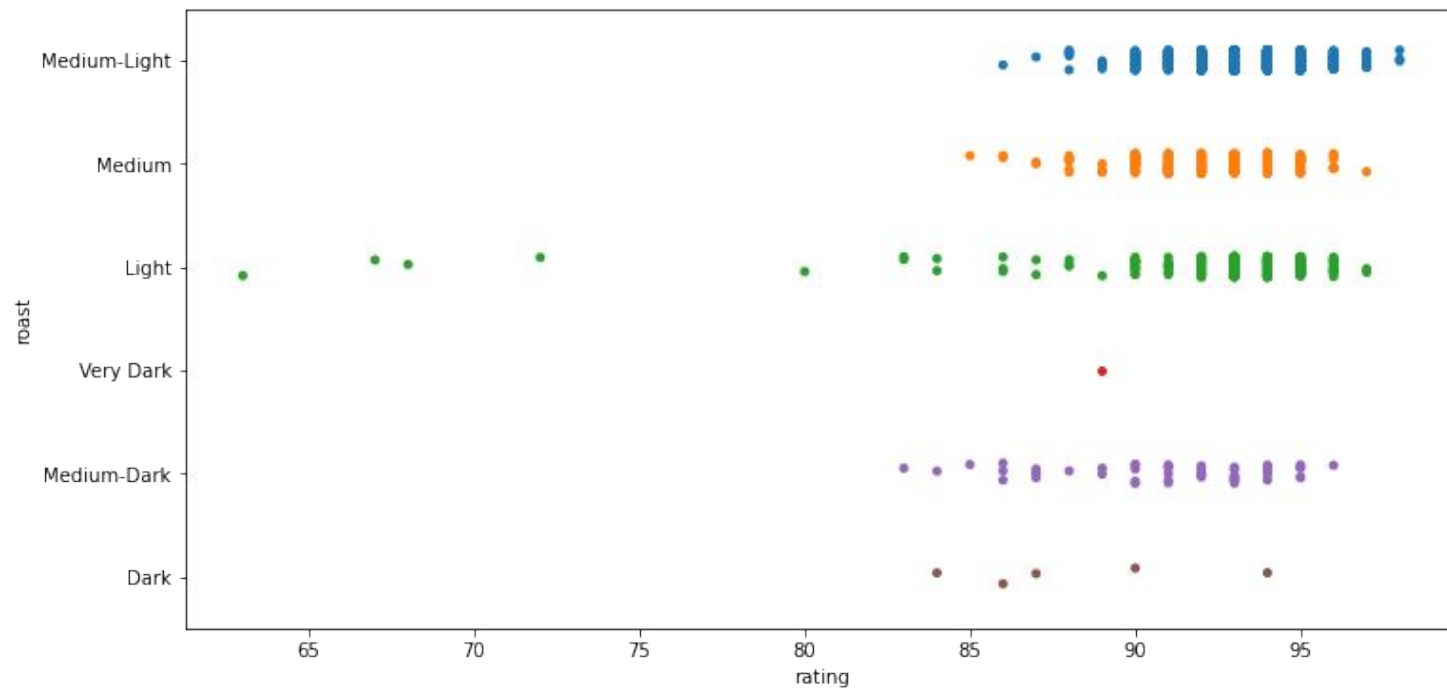
Single Bean Origin rated 95 or higher



Single Bean Origin rated 97 or higher



Roast against Rating



Preprocessing

Combine main
review and
summary

93

Joe Bean Roasters
Finca Buenavista Edwin
Zambrano Colombia

Roaster Location:	Rochester, New York	Review Date:	November 2022
Coffee Origin:	Huila Department, Colombia	Aroma:	9
Roast Level:	Medium-Light	Acidity/Structure:	9
Agtron:	58/78	Body:	8
Est. Price:	\$19.00/12 ounces	Flavor:	9
		Aftertaste:	8

Blind Assessment

Fruit-forward, crisply tropical. Strawberry guava, baking chocolate, lemon verbena, fresh-cut oak, brown sugar in aroma and cup. Sweet in structure with balanced, lively acidity; plush, syrupy mouthfeel. Crisply cocoa-toned finish with undertones of lemon verbena.

Notes

Produced by Edwin Zambrano of Finca Buenavista from trees of the Castillo and Colombia varieties of Arabica, and processed by the natural method (dried in the whole fruit). Joe Bean Roasters is a specialty retail and wholesale coffee roaster based in Rochester, New York. Visit www.joebeanroasters.com for more information.

Bottom Line

An inviting, cleanly fruit-toned, natural-processed Colombia cup with a crisp chocolate throughline.

Explore Similar Coffees

[Click here](#) for more reviews from **Joe Bean Roasters**

[Click here](#) for more information about coffees from **Colombia**

0

desc_1 Evaluated as espresso. Sweet-toned, deeply rich, chocolaty. Vanilla paste, dark chocolate, narcissus, pink grapefruit zest, black cherry in aroma and cup. Plush, syrupy mouthfeel; resonant, flavor-saturated finish. In three parts milk, rich chocolate tones intensify, along with intimations of vanilla paste and black cherry in the short finish and floral-toned citrus zest in the long.

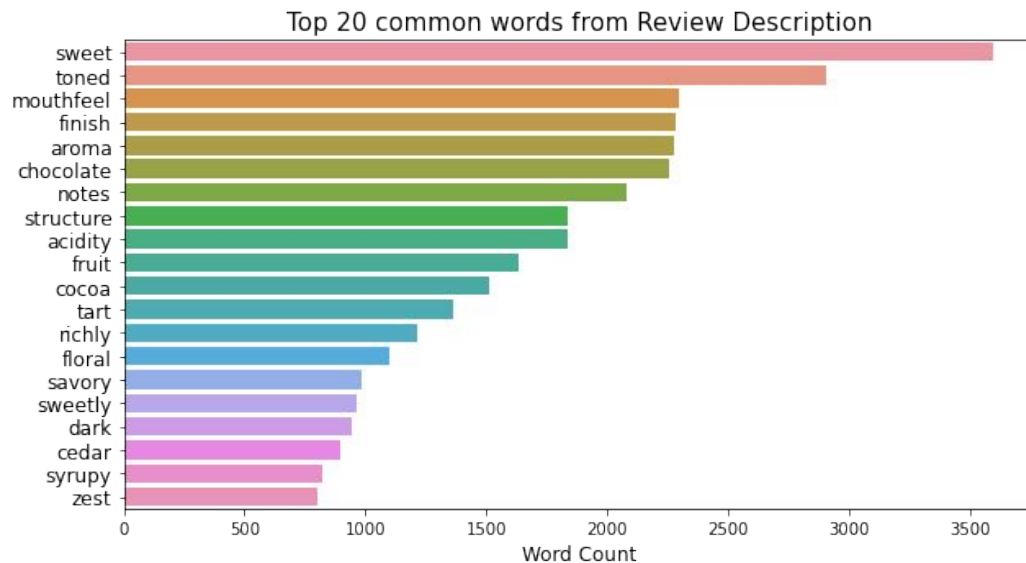
desc_2 An espresso blend comprised of coffees from Panama and Ethiopia. A.R.C., whose motto is "more than specialty," rigorously controls three vital elements: sourcing of green coffee, roasting, and, in cafe environments, water quality for brewed coffee. For more information, visit <http://www.arc.coffee/en/index.html>.

desc_3 A radiant espresso blend that shines equally in the straight shot and in milk, alive with notes of rich dark chocolate and black cherry.

desc Evaluated as espresso. Sweet-toned, deeply rich, chocolaty. Vanilla paste, dark chocolate, narcissus, pink grapefruit zest, black cherry in aroma and cup. Plush, syrupy mouthfeel; resonant, flavor-saturated finish. In three parts milk, rich chocolate tones intensify, along with intimations of vanilla paste and black cherry in the short finish and floral-toned citrus zest in the long. A radiant espresso blend that shines equally in the straight shot and in milk, alive with notes of rich dark chocolate and black cherry.

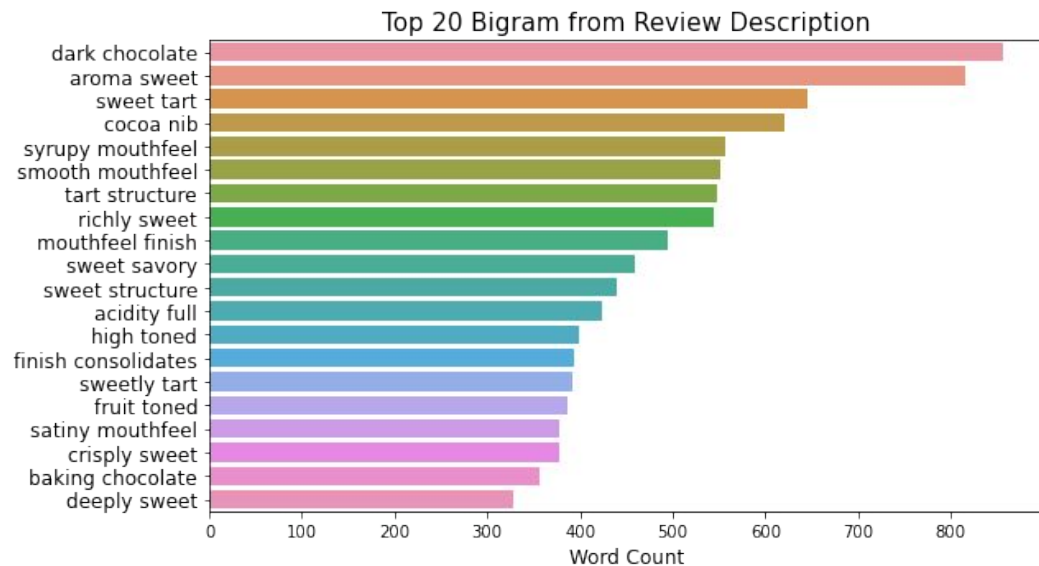
Top 20 Word

*added 'cup' to stopword



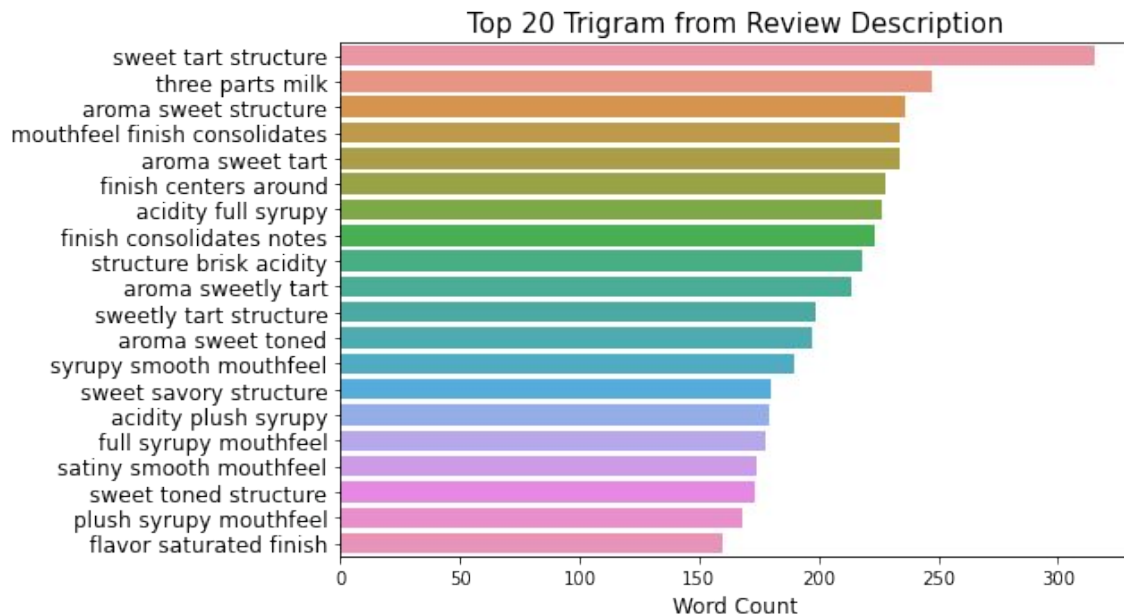
	Count
sweet	3596
toned	2903
mouthfeel	2299
finish	2283
aroma	2275
chocolate	2257
notes	2080
structure	1838
acidity	1836
fruit	1631
cocoa	1510
tart	1360
richly	1211
floral	1100
savory	984
sweetly	962
dark	941
cedar	898
syrupy	820
zest	801

Top 20 Bigram



	Count
dark chocolate	857
aroma sweet	816
sweet tart	645
cocoa nib	620
syrupy mouthfeel	556
smooth mouthfeel	551
tart structure	548
richly sweet	544
mouthfeel finish	495
sweet savory	459
sweet structure	440
acidity full	424
high toned	398
finish consolidates	393
sweetly tart	391
fruit toned	387
satiny mouthfeel	377
crisply sweet	377
baking chocolate	356
deeply sweet	328

Top 20 Trigram



	Count
sweet tart structure	315
three parts milk	247
aroma sweet structure	236
mouthfeel finish consolidates	234
aroma sweet tart	234
finish centers around	228
acidity full syrupy	226
finish consolidates notes	223
structure brisk acidity	218
aroma sweetly tart	214
sweetly tart structure	199
aroma sweet toned	197
syrupy smooth mouthfeel	190
sweet savory structure	180
acidity plush syrupy	179
full syrupy mouthfeel	178
satiny smooth mouthfeel	174
sweet toned structure	173
plush syrupy mouthfeel	168
flavor saturated finish	160

Features Selection

1. Origin
2. Taste (Scoring)
3. Tasting Notes (Key Words)

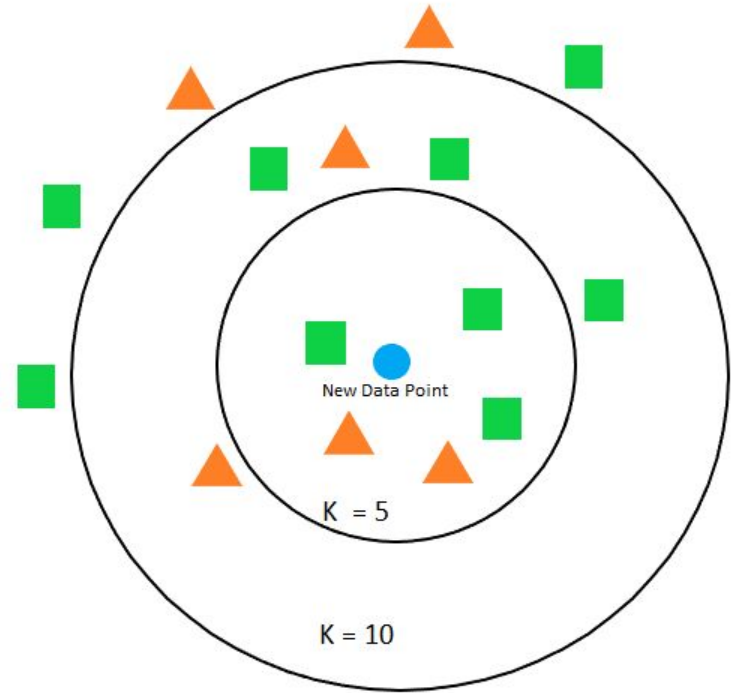
Features Selection - Origin

- Input: Columbia
- String filter in origin_country match to input

	name	roaster	roaster_country	origin_country	aroma	acid	body	flavor	aftertaste
22	Decaf Colombia Select	Bootstrap Coffee Roasters	United States	Colombia	7.0	7.0	8.0	8.0	7.0
42	Static Colombia Decaf Midnight	Espresso Republic	United States	Colombia	8.0	8.0	8.0	8.0	7.0
45	Sweet Holiday Blend	Red Rooster Coffee Roaster	United States	Colombia; Ethiopia; Kenya	9.0	8.0	9.0	9.0	8.0
60	Colombia Nariño Duende	Paradise Roasters	United States	Colombia	8.0	9.0	9.0	9.0	8.0
69	Gaitania Colombia	JBC Coffee Roasters	United States	Colombia	9.0	8.0	8.0	9.0	7.0

NearestNeighbors

- NearestNeighbors is an unsupervised technique of finding the nearest data points with respect to each data point
- It's ideal for non-linear data



Taste Rating - n_neighbour = 50

Input:

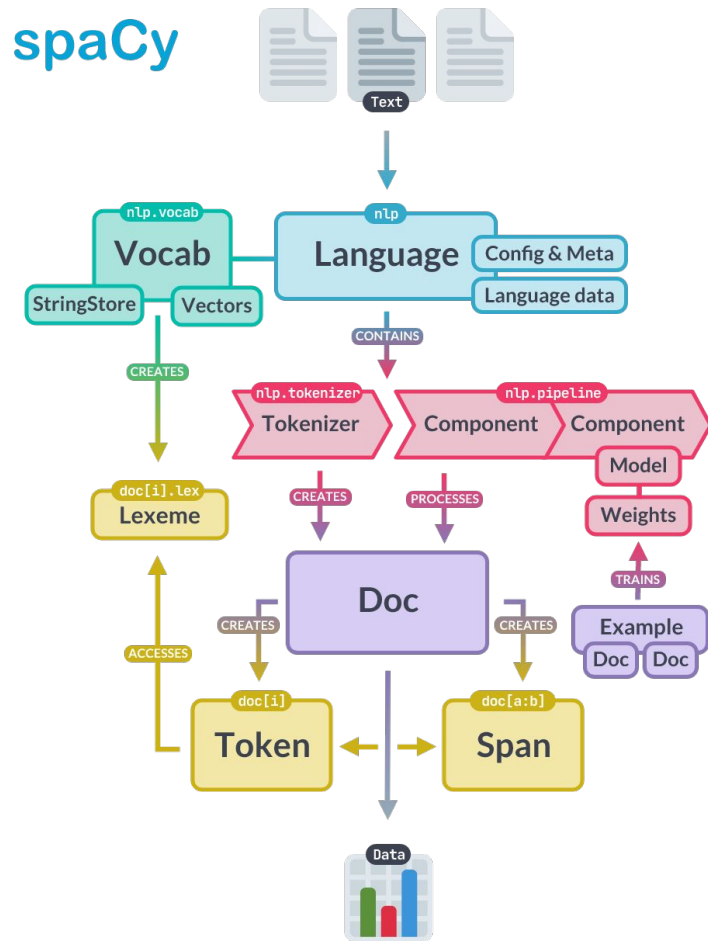
- Aroma = 6
- Acidity = 7
- Body = 8
- Flavor = 7
- After Taste = 9

	name	roaster	roaster_country	origin_country	aroma	acid	body	flavor	aftertaste
391	100% Arabica Freeze-Dried Colombian (Instant C...	Waka Coffee	United States	Colombia	7.0	7.0	8.0	8.0	8.0
983	French Riviera	Cafe Arles	Taiwan	Brazil; Guatemala; Indonesia	7.0	7.0	8.0	8.0	7.0
34	Static El Salvador Finca Maria Luisa	Espresso Republic	United States	El Salvador	8.0	7.0	8.0	8.0	8.0
22	Decaf Colombia Select	Bootstrap Coffee Roasters	United States	Colombia	7.0	7.0	8.0	8.0	7.0
889	Lover's Quarrel	Mooju Coffee	United States	Indonesia	7.0	7.0	7.0	7.0	7.0

Spacy

What is en_core_web_lg?

- en_core_web_lg is the trained pipeline for the English language.
- It contains components like ok2vec, tagger, parser, sender, ner, attribute_ruler, lemmatizer.
- the largest English model in size provided by Spacy.



Tasting Notes - Similarity Score

Description:

'Briskly sweet-savory. Pie cherry, lavender, cocoa nib, cashew butter, apple blossom in aroma and cup. Richly bittersweet structure lifted by juicy acidity; smooth, satiny mouthfeel. The crisp finish consolidates to cocoa nib and cashew. A sweetly nutty, richly bittersweet natural-processed Guatemala cup given levity by sweetly tart stone fruit notes and sweet and spicy florals.'

Input: Juicy, Sweet

	name	roaster	roaster_country	origin_country	Similarity
1581	Guatemala Finca Granada Natural	Plat Coffee Roastery	Hong Kong	Guatemala	0.473825
1742	Ethiopia Washed Sidama Bensa Bombe	Kakalove Cafe	Taiwan	Ethiopia	0.469778
2171	Guatemala Los Santos	Temple Coffee	United States	Guatemala	0.457891
2046	Philippines Sitio Belis 1017	Mostra Coffee	United States	Philippines	0.451931
2226	Symphony Costa Rica Canet Caturra Raisin Honey	Chiming Coffee	Taiwan	Costa Rica	0.449976

Tasting Notes - Similarity Score

Description:

'Briskly sweet-savory. Pie cherry, lavender, cocoa nib, cashew butter, apple blossom in aroma and cup. Richly bittersweet structure lifted by juicy acidity; smooth, satiny mouthfeel. The crisp finish consolidates to cocoa nib and cashew. A sweetly nutty, richly bittersweet natural-processed Guatemala cup given levity by sweetly tart stone fruit notes and sweet and spicy florals.'

Input: Juicy, Sweet, Bright

	name		roaster	roaster_country	origin_country	Similarity
1742	Ethiopia Washed Sidama Bensa Bombe		Kakalove Cafe	Taiwan	Ethiopia	0.614459
1581	Guatemala Finca Granada Natural		Plat Coffee Roastery	Hong Kong	Guatemala	0.599367
2171	Guatemala Los Santos		Temple Coffee	United States	Guatemala	0.597693
3	Ethiopia Suke Quto		Roast House	United States	Ethiopia	0.595554
1741	Sweet Holiday Blend	Red Rooster Coffee Roaster	United States	Colombia; Costa Rica; Ethiopia; Guatemala		0.594207

<https://beans4u.streamlit.app/>

Tasting Notes

Type in your preference. E.g. juicy, sweet, bright

sweet

Taste Rating

Aroma

0 9

Acid

0 9

Body

0 9

Flavor

0 9

Coffee Bean Recommender

Instruction

First: Type in your preferred tasting notes.

Second: Use the slider to input your preferred taste rating for each taste category

Lastly: You can select your preferred coffee bean's origin from the dropdown or None

Made with Streamlit

Next Steps

- Incorporating local coffee bean data to the model
- Setting cronjob to collect weekly/monthly to refresh data
- Develop the app to collect user data for collaborative filtering in the future