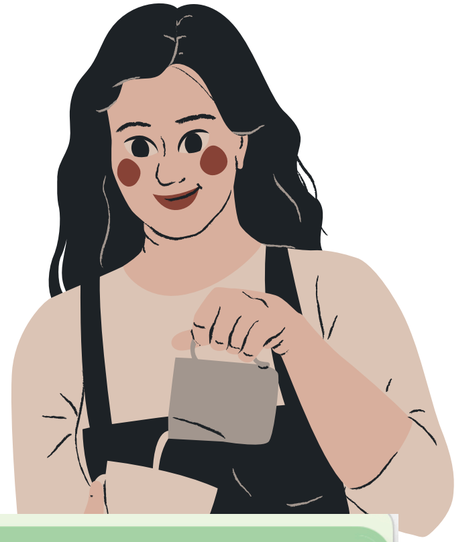

CAFE
RECOMMENDER
KIMBERLY CUI

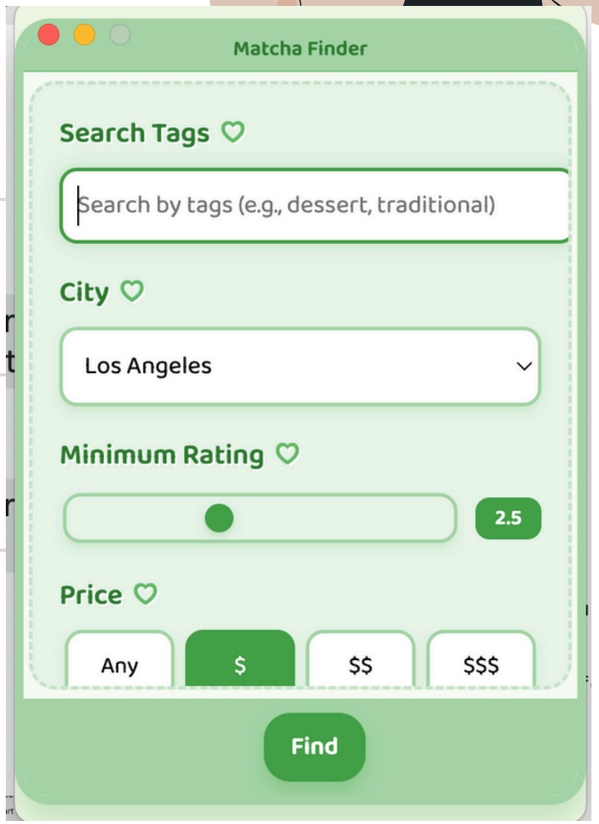
PURPOSE

NEED TO FIND A NEW CAFE?
OVERWHELMED BY TOO MANY CHOICES?
STRICT CRITERIA?

NO FEAR!

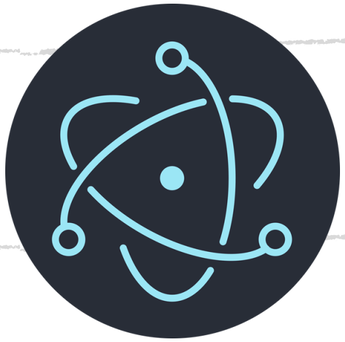


- Find your VIBE with this cafe recommender!
- matcha-focused spots only 🍵
- Price, rating, and city filters
- Reviews-based vibe categorization
- 🎵 cute iPod-style UI



WHAT I USED

yelp fusion



YAKE (yet another keyword extractor)

DATASET

script.py 2 | matcha_analysis.py 9+ | review_analysis.py 2

Users > kimberlycui > Desktop > script.py > ...

```python
import requests
import pandas as pd
import time

SEARCH_URL = "https://api.yelp.com/v3/businesses/search"
REVIEW_URL = "https://api.yelp.com/v3/businesses/{}/reviews"

# Define a list of locations (you can expand this)
locations = ["Los Angeles, CA", "San Francisco, CA", "New York, NY", "Seattle, WA"]

# Function to fetch data
def get_yelp_data(location, term="matcha cafe", limit=50):
    businesses = []
    for offset in range(0, 250-limit, limit):  # Yelp allows up to 1000 results
        print(offset)
        params = {
            "term": term,
            "location": location,
            "limit": limit,
            "offset": offset,
            "categories": "cafes,coffee",
            "sort_by": "rating",
            "price": "1,2,3,4",
        }
        response = requests.get(SEARCH_URL, headers=HEADERS, params=params)

        if response.status_code == 200:
            data = response.json()
            businesses.extend(data.get("businesses", []))
        else:
            print(f"Error {response.status_code}: {response.json()}")
            break  # Stop if there's an error
```

yelp fusion

name	rating	price	review_count	address	city	latitude	longitude	url
Tea Master Matcha Cafe & Green Tea Shop	4.8	$$	2046	450 E 2nd St, Los Angeles, CA 90012	Los Angeles	34.0471823589909	-118.238639755523	https://www.yelp.com/biz/tea-master-r
Matcha Village	4.2	$	1191	2202 S Sawtelle Blvd, Los Angeles, CA 90064	Los Angeles	34.03804	-118.44089	https://www.yelp.com/biz/matcha-villa
Midori Matcha Cafe	4.2	$$	1125	123 Astronaut Ellison Onizuka St, Ste 101- C, Los Angeles, CA 90012	Los Angeles	34.0499578055129	-118.24206790276727	https://www.yelp.com/biz/midori-matcl
Cha Cha Matcha	4.0	$$	444	510 N Robertson Blvd, West Hollywood, CA 90048	West Hollywood	34.0804007872686	-118.38497646153	https://www.yelp.com/biz/cha-cha-mat
Chitchat Coffee and Matcha	4.2	$$	474	1854 Sawtelle Blvd, Los Angeles, CA 90025	Los Angeles	34.04276	-118.4448	https://www.yelp.com/biz/chitchat-coff
Matcha Cafe Maiko- Rowland Heights	4.1	$$	588	18180A Colima Rd, Rowland Heights, CA 91748	Rowland Heights	33.9883115	-117.9072417	https://www.yelp.com/biz/matcha-cafe
Fika Fika Coffee & Creamery	4.9	$	140	7267 Melrose Ave, Los Angeles, CA 90046	Los Angeles	34.08365	-118.34807	https://www.yelp.com/biz/fika-fika-coff
Ondo Coffee	4.8	$$	59	710 Heliotrope Dr, Los Angeles, CA 90029	Los Angeles	34.08399820185051	-118.2948207	https://www.yelp.com/biz/ondo-coffee-
Makisupa Coffee Bar	4.8	$$	34	2894 Rowena Ave, Silver Lake, Los Angeles, CA 90039	Los Angeles	34.107908	-118.267837	https://www.yelp.com/biz/makisupa-co
Cafe Ruisseau	4.8	$$	180	12150 Millennium Dr, Los Angeles, CA 90094	Los Angeles	33.979614868581	-118.408223949385	https://www.yelp.com/biz/cafe-ruisseau
Tilt Coffee Bar	4.8	$$	789	334 S Main St, Kiosk 1, Los Angeles, CA 90013	Los Angeles	34.048162858309	-118.246295170034	https://www.yelp.com/biz/tilt-coffee-ba
Airoma Cafe	4.7	$$	1578	10246 Westminster Ave, Garden Grove, CA 92843	Garden Grove	33.758857	-117.950252	https://www.yelp.com/biz/airoma-cafe-
Yeems Coffee	4.7	$	771	3033 W 6th St, Ste 107, Los Angeles, CA 90020	Los Angeles	34.064007712098366	-118.28828741104604	https://www.yelp.com/biz/yeems-coffe
Faregrounds	4.7	$$	67	459 N Roxbury Dr, Beverly Hills, CA 90210	Beverly Hills	34.068171115781274	-118.4071518	https://www.yelp.com/biz/faregrounds-
Okayama Kobo Bakery & Cafe	4.7	$$	1610	155 W Center St Promenade, Anaheim, CA 92805	Anaheir	review_1		
Nook Coffee Bar	4.6	$$	439	980 W 190th St, Ste B1, Torrance, CA 90502	Torranc	The service was great as the staff will give detail on the products they were selling. The ambiance was good as well for photos. I order		
Paradocs Coffee and Tea	4.6	$	465	1032 S Fairfax Ave, Los Angeles, CA 90019	Los Ang	Delightful matcha shop with a good selection of both matcha and non-matcha drinks. There's also a food menu that includes snacks lik		
SACHI.LA	4.6	$$	284	4574 S Centinela Ave, Los Angeles, CA 90066	Los Ang	Was craving some matcha soft serve and this place delivered! I got the swirl, which was matcha and hojicha. Got it in a cone (extra $1		
Stereoscope Coffee	4.6	$$	300	4925 E 2nd St, Long Beach, CA 90803	Long Be	I love the staff, every time I come in Aaron always remembers my drink. It's the small things :)		
DAMO	4.6	$$	812	3510 W 8th St, Los Angeles, CA 90005	Los Ang	The Baseline: For a coffee shop the baseline involves the quality of the roast, the taste of the coffee, and the preparation of specialty		
Maru Espresso Bar	4.6	$$	151	201 N Canon Dr, Beverly Hills, CA 90210	Beverly	I love Cafe Maiko!! It's always a great place to come after dinner in the plaza. My boyfriend and I love coming for BCD Tofu House and		
Stagger Coffee	4.6	$$	294	3069 W 8th St, Los Angeles, CA 90005	Los Ang	I really liked their honeycomb waffle and Laureen was very helpful in recommending me to try the earl grey latte. Definitely stopping b		
Yeems Coffee	4.5	$$	94	15420 S Western Ave, Gardena, CA 90249	Garden	Ondo Coffee is by far one of the hippest and coolest coffee shops I have visited in a while! This specialty coffee shop located in East H		
Rny Coffee Studio	4.5	$	348	3829 W 6th St, Los Angeles, CA 90020	Los Ang	Street Parking only. Staff is very welcoming and helpful. This business is a safe place to all. They offer beverages, snacks, and food.		
6XS COFFEE	4.5	$$	287	3907 W 6th St, Los Angeles, CA 90020	Los Ang	Friendly neighborhood jaunt. Staff is friendly. Drinks are delicious. They have lots of Gluten free goodies.		
Meccha Matcha	4.5	$$	714	11816 Artesia Blvd, Artesia, CA 90701	Artesia	i've always been fond of their strawberry cloud latte, this is easy on my stomach and it s not super sweet! i like their strawberry milk c		
Paderia Bakehouse	4.5	$$	312	3032 Wilshire Blvd, Ste E, Santa Monica, CA 90403	Santa M	Matcha einspanner with 50% since I don't like matcha that tastes too sweet so the sugar level is good for me.		
3THYME Coffee	4.5	$$	298	600 S Harvard Blvd, Ste 100, Los Angeles, CA 90005	Los Ang	AHHH I've been wanting to try Yeems for so long and I'm glad to say it did not disappoint! I came here on a Sunday afternoon ~ 2:30		
Coffee MCO	4.5	$	776	2580 W Olympic Blvd, Los Angeles, CA 90006	Los Ang	Love this place! The staff is so warm and welcoming. The food is amazing. I highly recommend getting the spicy don. I will definitely b		
						Don't be intimidated, the line goes by fast!! You can also call ahead to order! Here's a list of my bread recs: -salt and butter/cookie sa		
						Taking a spot in The Enclave, a complex of restaurants and cafes with plenty of seats right under the open sky! I really really love this		
						I'd rate this as a top ten coffee shop in all of LA. It's non-descript. Almost hidden. Hole-in-the-wall vibe. Taka is the owner and he is th		
						DRINK wait stop why is this place like a hidden gem.. the matcha was so good hello?? SERVICE service is fair ~ i think there's a little b		

extracted: NAME, RATING, PRICE, CITY, URL, 3 REVIEWS

YAKE (YET ANOTHER KEYWORD EXTRACTOR)

FIRST COMBINE ALL 3 REVIEWS!

CONFIGURE THE YAKE KEYWORD EXTRACTOR

- lan="en": Language of the text (English).
- n=2: Extract keywords that consist of two-word phrases (bigrams).
- dedupLim=0.5: controls how aggressively YAKE avoids returning duplicated keywords.
- top=5: Extracts the five most important keyword phrases per combined review.

DEFINE A KEYWORD EXTRACTION FUNCTION

- function takes the text of combined reviews and applies YAKE
- returns a list of the top 5 keyword phrases extracted from that text

APPLY KEYWORD EXTRACTION TO DATASET

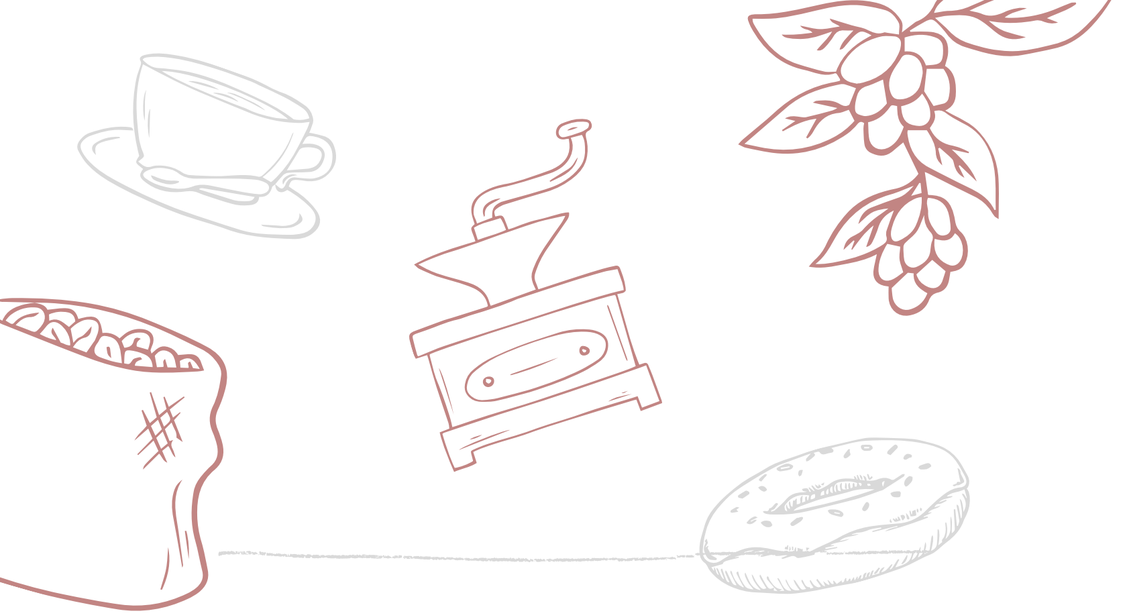
- Applies function (extract_yake_keywords) to each combined review.
- Stores the resulting list of keywords in a new column (tags)

```python
# YAKE keyword extractor configuration
kw_extractor = yake.KeywordExtractor(
    lan="en",
    n=2,  # Bigrams (2-word phrases), adjust if needed
    dedupLim=0.5,
    top=5,  # Top 5 keywords per document
    features=None
)

# Extract keywords using YAKE
def extract_yake_keywords(text):
    keywords = kw_extractor.extract_keywords(text)
    return [kw[0] for kw in keywords]

# Apply keyword extraction
df["tags"] = df["combined_reviews"].apply(extract_yake_key

# Save the updated dataframe
output_file = "matcha_cafes_with_tags.csv"
df.to_csv(output_file, index=False)
```



HOW DOES YAKE EXTRACT KEYWORDS?

1. TERM FREQUENCY
2. TEMR POSITION
    A. PHRASES OCCURRING EARLIER IN THE TEXT CAN HAVE HIGHER RELEVANCE.
3. TERM RELATEDNESS TO CONTEXT
4. CASE (UPPERCASE, TITLE-CASE)
5. STOP WORDS AND COMMON WORDS
    A. COMMON, LOW-VALUE WORDS (E.G., "AND," "THE," "WITH") ARE TYPICALLY IGNORED
6. PHRASE UNIQUENESS
7. YAKE PREFERS PHRASES THAT ARE LESS GENERIC, GIVING PRIORITY TO TERMS UNIQUELY CHARACTERIZING A DOCUMENT OR REVIEW.

KEYWORDS EXAMPLE

UPSIDE DOWN

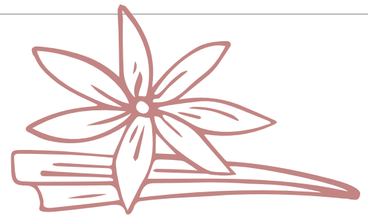
donation based friendly people and atmosphere great events to meet people great selection to drinks to enjoy	Just had an amazing first experience here! I attended one of their coffee tastings. Due to the recent tragedies here in LA they did it a little differently,...	If you're looking for a coffee experience that transcends the ordinary, look no further than Cold Brew. This delightful beverage is the epitome of smooth,...

['donation based', 'based friendly', 'friendly people', 'atmosphere great', 'great events']

ANKO

Pretty good! Got the mango bingsoo and blueberry latte. The bingsoo was smaller than usual but the ice was a lot more creamy and flavorful. The latte was...	Anko still has the best taiyaki in Koreatown (I always have to stop by and order 2 red bean and 2 custard)! The outside is always crispy, while the inside...	Cozy and quaint cafe, super sweet and clean aesthetic with chill music ambiance. Got the green grape refresher, which I am a sucker for anything grape...

['Pretty good', 'good', 'mango bingsoo', 'blueberry latte', 'latte']



Percentage Match Breakdown

Tag Matching: 50%

- Still uses the same exact/partial matching system
- Most important factor after location filtering

Rating Matching: 35%

- Higher minimum ratings filter out lower-rated cafes
- Partial credit for cafes that almost meet your minimum

Price Matching: 15%

- Exact price matches get full points
- Similar prices get partial credit

Acurrúcame Café 85%

Los Angeles | Rating: 4.9 | $

friendly space · space · calm · ...

Matcha Village 50%

Los Angeles | Rating: 4.2 | $

good selection · matcha · non-matcha drinks · ...

Fika Fika Coffee & Creamery 50%

Los Angeles | Rating: 4.9 | $

grey latte · honeycomb waffle · earl grey · ...

DEMO

THANK YOU KAYLIN & JACOB!