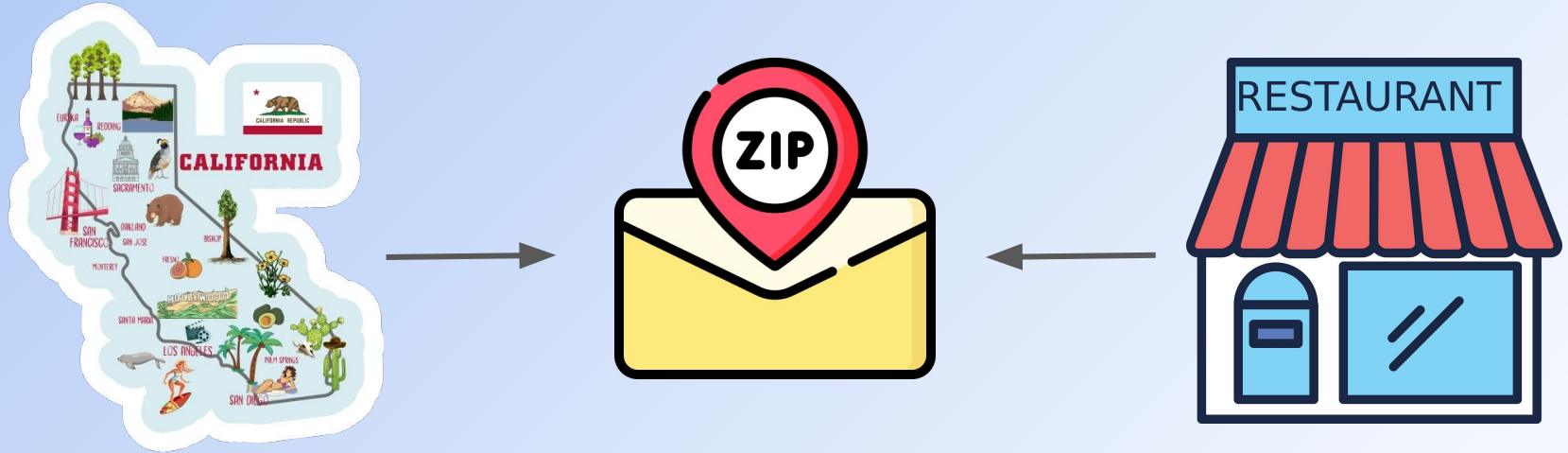
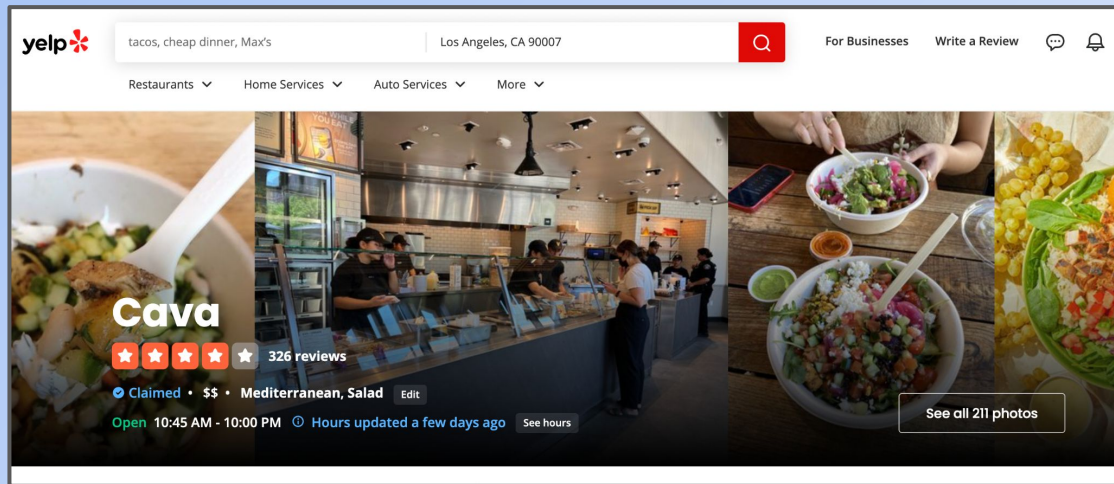


ZIP-Find


Location based Restaurant Search and Recommendation System

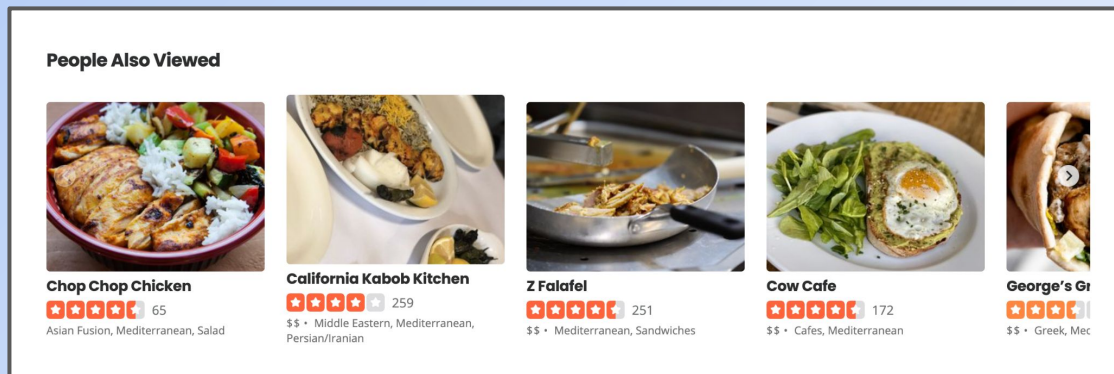


Motivation / Why KG?

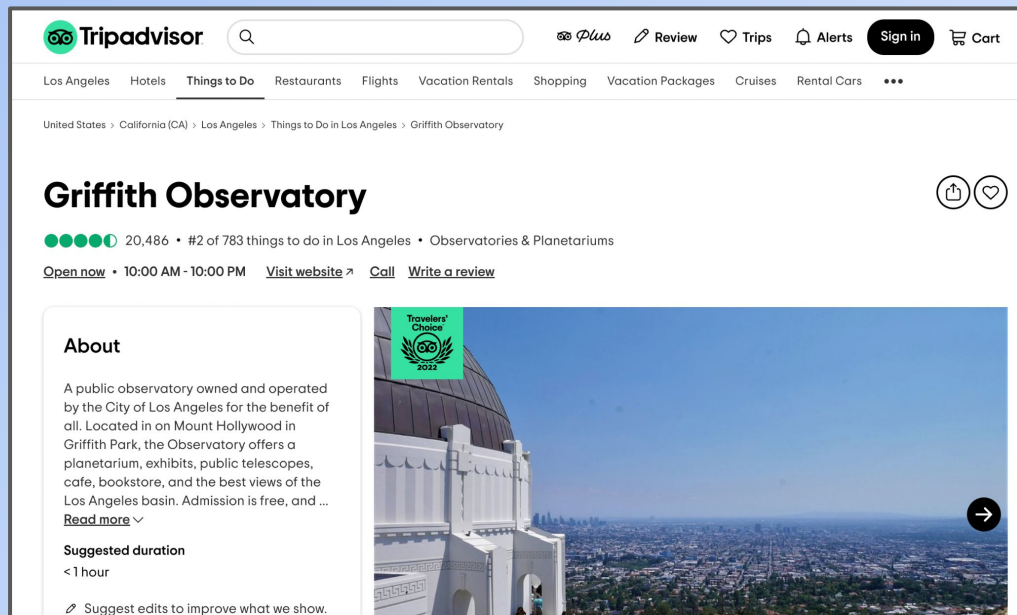


What if I
want to find
a restaurant
like Cava
near the
attraction I
like?

Not Possible on Yelp 



Motivation / Why KG?



The screenshot shows the TripAdvisor page for the Griffith Observatory. The header includes the TripAdvisor logo, a search bar, and navigation links like 'Plus', 'Review', 'Trips', 'Alerts', 'Sign in', and 'Cart'. Below the header, there's a breadcrumb trail: 'United States > California (CA) > Los Angeles > Things to Do in Los Angeles > Griffith Observatory'. The main title 'Griffith Observatory' is prominently displayed, followed by a rating of 4.5 stars (20,486 reviews) and a description: '#2 of 783 things to do in Los Angeles • Observatories & Planetariums'. There are links for 'Open now', '10:00 AM - 10:00 PM', 'Visit website', 'Call', and 'Write a review'. An 'About' section on the left describes the observatory as a public observatory owned and operated by the City of Los Angeles. A large image of the observatory's dome and the city view is shown on the right, with a 'Travelers' Choice 2022' badge in the top left corner of the image.

Griffith Observatory

20,486 • #2 of 783 things to do in Los Angeles • Observatories & Planetariums

[Open now](#) • 10:00 AM - 10:00 PM [Visit website](#) [Call](#) [Write a review](#)

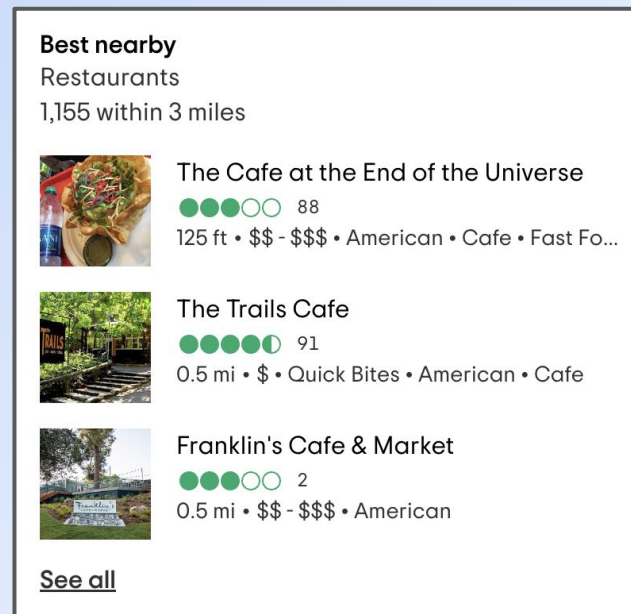
About

A public observatory owned and operated by the City of Los Angeles for the benefit of all. Located in on Mount Hollywood in Griffith Park, the Observatory offers a planetarium, exhibits, public telescopes, cafe, bookstore, and the best views of the Los Angeles basin. Admission is free, and ... [Read more](#)

Suggested duration

< 1 hour

[Suggest edits to improve what we show.](#)



The screenshot shows the 'Best nearby Restaurants' section on TripAdvisor. It lists three restaurants within 3 miles of the Griffith Observatory. Each entry includes a small image of the restaurant, its name, a rating (represented by green circles), the number of reviews, and a list of categories (e.g., American, Cafe, Fast Food).

Best nearby Restaurants

1,155 within 3 miles

The Cafe at the End of the Universe

88 reviews
125 ft • \$\$ - \$\$\$ • American • Cafe • Fast Fo...

The Trails Cafe

91 reviews
0.5 mi • \$ • Quick Bites • American • Cafe

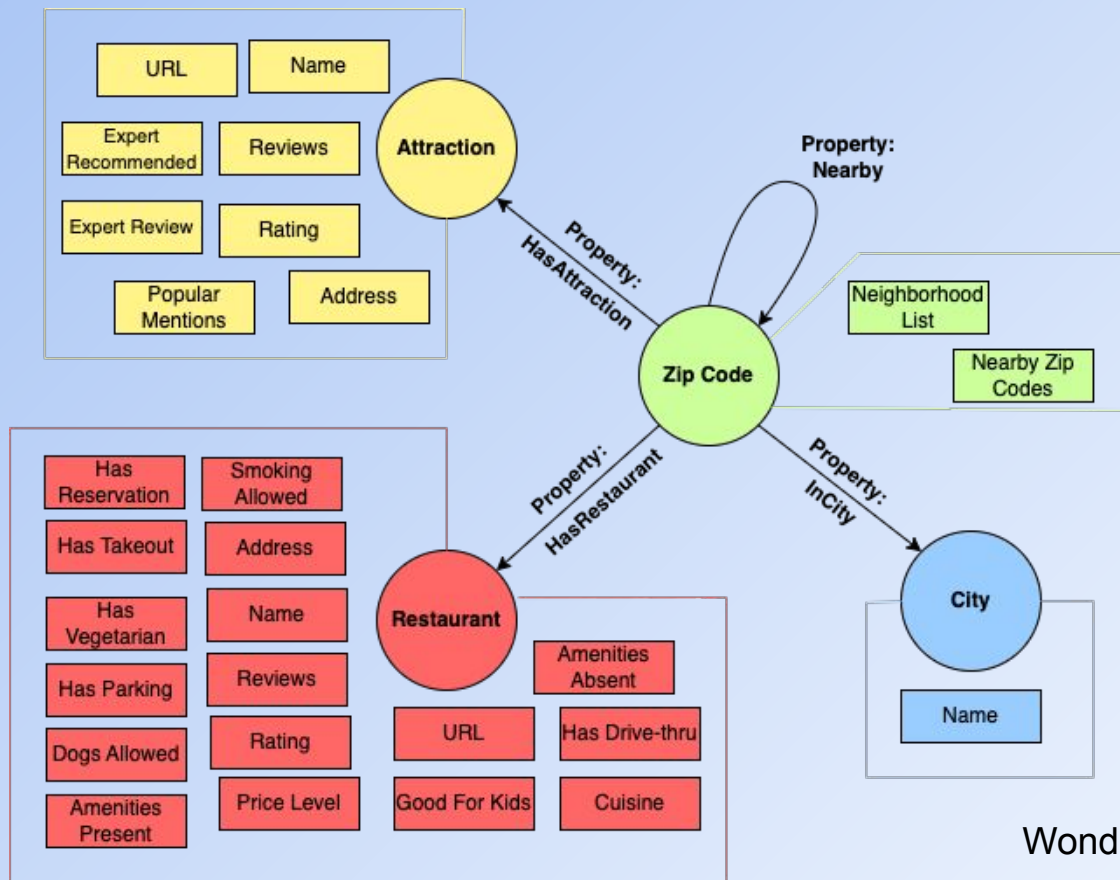
Franklin's Cafe & Market

2 reviews
0.5 mi • \$\$ - \$\$\$ • American

[See all](#)

Tripadvisor recommends restaurants near an attraction but this recommendation does not account user's preferences.

Ontology - Classes and Relations



Total Nodes: 18173

Attraction: 8182

Restaurant: 7446

Zip Code: 1553

City: 992

Total Relations: 27,344

HasAttraction: 7834

HasRestaurant: 7235

InCity: 1553

Nearby: 10722

KG Construction:

Neo4j AuraDB

Query:

Cypher

Wondering where the node for state is



Data Sources & Extraction

1) Zip Codes Data

Website: <https://www.unitedstateszipcodes.org/ca/>

ZIP Code	Type	Common Cities	County	Area Codes
90001	Standard	Los Angeles, Firestone Park, Firestone Pk	Los Angeles County	Area Code 323
90002	Standard	Los Angeles, Watts	Los Angeles County	Area Code 323 Area Code 562
90003	Standard	Los Angeles	Los Angeles County	Area Code 323
90004	Standard	Los Angeles, Oakwood	Los Angeles County	Area Code 323

ZIP Code 90014

Population

Real Estate

Employment

Schools

Post Office City: Los Angeles, CA ([View All Cities](#))

Neighborhood: Gallery Row|Fashion District|Downtown Los Angeles

County: Los Angeles County



Scrapy

Cities in ZIP code 90014

The list below includes the cities that the US Post Office accepts for ZIP code 90014. The preferred city may not be the city in which the ZIP is located. The city for 90014 is usually the name of the main post office. When mailing your package or letter, always include the preferred or acceptable cities. Using any city in the list of unacceptable cities may result in delays.

Primary/preferred city: Los Angeles, CA

Output Jsonl:

No. of lines: 1553

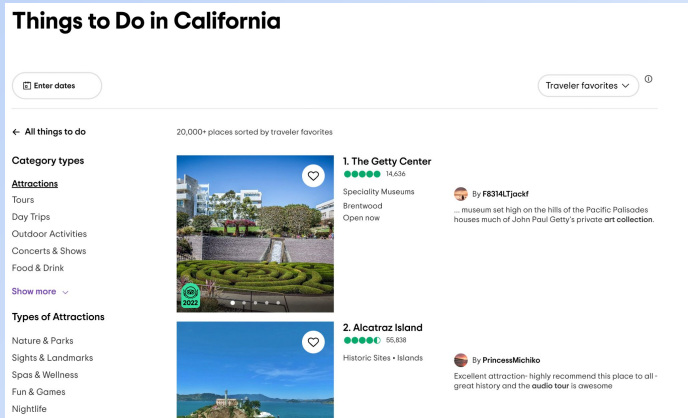
```
{"zip_code": "90039", "neighborhood_list": ["Atwater Village", "Northeast Los Angeles"], "city": "Los Angeles, CA", "nearby_zip_codes": ["90012", "90026", "90027", "90029", "90031", "90041", "90065", "91201", "91203", "91204", "91205", "91206", "91210"]}  
{"zip_code": "90038", "neighborhood_list": ["Los Angeles County"], "city": "Los Angeles, CA", "nearby_zip_codes": ["90004", "90027", "90028", "90029", "90036", "90046"]}
```


Data Sources & Extraction

2) Attraction Data - Tripadvisor

Website:

<https://www.tripadvisor.com/Attractions-g28926-Activities-allAttractions.true-California.html>



Popular mentions

hollywood sign

solar system

tesla coil

planetarium shows

rebel without a cause

james dean

great views of la

view of los angeles

greek theatre

parking spot

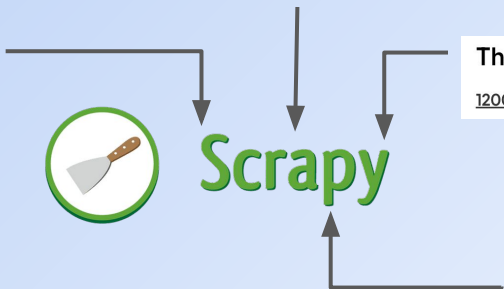
great place to visit

worth the trip

hollywood hills

downtown la

telescope



The area

1200 Getty Center Dr N Sepulveda Blvd & Getty Center Dr, Los Angeles, CA 90049-1657

Reviews Q&A

5.0 14,636 reviews

Output Jsonl:
No. of lines: 8182

```
{
  "attraction_name": "Santa Monica Pier",
  "attraction_address": "200 Santa Monica Pier, Santa Monica, CA 90401-3126",
  "zip_code": "90401",
  "attraction_rating": 4.0,
  "attraction_n_reviews": 19138,
  "attraction_url": "https://www.tripadvisor.com/Attraction_Review-g33052-d104238-Reviews-Santa_Monica_Pier-Santa_Monica_California.html",
  "popular_mentions": [
    "Filters",
    "English",
    "Detailed Reviews",
    "ferris wheel",
    "amusement park",
    "souvenir shops",
    "bubba gump",
    "carnival rides",
    "people watching",
    "merry go round",
    "fun place",
    "great place to visit",
    "pacific park",
    "nice walk",
    "few hours",
    "lots of fun",
    "street performers",
    "los angeles",
    "pier",
    "end",
    "restaurants",
    "route",
    "parking",
    "arcade",
    "games",
    "sunset",
    "ocean",
    "fishing",
    "entertainment",
    "landmark",
    "stroll",
    "sign",
    "musicians",
    "expert_review": "The Santa Monica Pier is home to rides, food, and attractions. It is also the endpoint of the famous Route 66, which makes it a great place to explore that is full of history.",
    "expert_recommended": "Yes"
  ]
}
```

Data Sources & Extraction

3) Attraction Data - Expert Traveller

Website:

<https://californiathroughmylens.com/strange-fun-attractions-list/>

<https://californiathroughmylens.com/strange-unique-northern-california/>

ABOUT



My name is Josh, and for the last 10 years, I have been exploring as much of California as I possibly can. This site is a collection of those adventures to encourage you to get out and explore California for yourself. [More about me.](#)



Scraping attraction
name and expert
review

Santa Monica Pier



The Santa Monica Pier is home to rides, food, and attractions. It is also the endpoint of the famous [Route 66](#), which makes it a great place to explore that is full of history.

Attraction data from two different sources?

RECORD LINKAGE

Entity Linking / Record Linkage



Similarity measures used for entity linking: Jaro Winkler, TF-IDF on attraction names & attraction name tokens in tripadvisor address string.

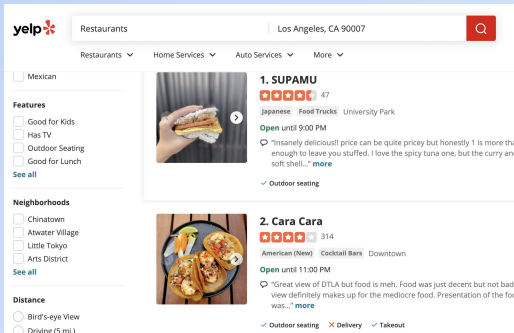
Precision = 0.9333333333333333 , Recall = 1.0 , F1 Score = 0.9655172413793104 against a manually annotated data of 28 true positive pairs.

Data Sources & Extraction

4) Restaurant Data - Yelp

Website:

<https://www.yelp.com/search?cflt=restaurants>



Located in:

USC Village

3201 S Hoover

Ste 1840

Los Angeles, CA 90089

University Park



Scrapy



Amenities and More

Health Score A

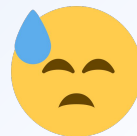
Offers Takeout

Offers Delivery

No Reservations

20 More Attributes

Extracting Amenities - CHALLENGE



1st Technical Challenge



to the rescue!



Finds & Clicks this button

20 More Attributes

Scrapes all these attributes

- **16 Hours** to run the whole scraping
- 7446 Restaurants scraped
- Amenities present and absent determined using color distinction used for measuring similarities between restaurants.

```
{"restaurant_name": "Cava", "n_reviews": "326", "average_rating": "4", "restaurant_address":  
"3201 S Hoover, Ste 1840, Los Angeles, CA 90089", "zip_code": "90089", "amenities_present":  
["Health Score", "Offers Delivery", "Offers Takeout", "Masks required", "Staff wears masks",  
"Vegan Options", "Accepts Credit Cards", "Accepts Apple Pay", "Outdoor Seating", "Trendy",  
"Casual", "Moderate Noise", "Good for Groups", "Good For Kids", "Good for Lunch", "Dinner",  
"Street Parking", "Validated Parking", "TV", "Gender-neutral restrooms", "Bike Parking"],  
"amenities_absent": ["No Reservations", "Accepts Android Pay", "Offers Catering", "No Waiter  
Service", "No Wi-Fi", "No Alcohol"], "price_level": 2, "cuisine_list": ["Mediterranean",  
"Salad"], "restaurant_url": "https://www.yelp.com/biz/cava-los-angeles-4?osq=Restaurants",  
"has_vegetarian": 1, "has_parking": 1, "has_takeout": 1, "has_drivethru": 0, "has_reservation":  
-1, "is_good_for_kids": 1, "dogs_allowed": 0, "smoking_allowed": 0}
```

Health Score A	Offers Delivery
Offers Takeout	No Reservations
Masks required	Staff wears masks
Vegan Options	Accepts Credit Cards
Accepts Apple Pay	Accepts Android Pay
Outdoor Seating	Trendy, Casual
Moderate Noise	Good for Groups
Good For Kids	Good for Lunch, Dinner
Street Parking, Validated Parking	TV
Gender-neutral restrooms	Offers Catering
No Waiter Service	No Wi-Fi
No Alcohol	Bike Parking

Show Less

City-based Attraction Search (UI)

Expert curated attraction displayed first.

Expert curated attraction also has expert's review.

Other attractions sorted by rating first then number of reviews.

Enter city you want to visit

San Francisco, CA

What do you want to explore?

☒ Attractions ☐ Restaurants

1. [Lombard Street](#)

Expert Recommended!

Average rating: 4.5 , Number of reviews: 18313

Address: 1070 Lombard Street between Jones St and Hyde St, San Francisco, CA 94109

Expert Review: I am sure you have seen this famous San Francisco attraction on TV or in a movie, but it is still fun to visit in real life.

2. [Marrakech Magic Theater](#)

Average rating: 5.0 , Number of reviews: 1288

Address: 419 O'Farrell St, San Francisco, CA 94102-2009

3. [The Speakeasy](#)

Average rating: 5.0 , Number of reviews: 563

Address: 644 Broadway, San Francisco, CA 94102

4. [Anchor Brewing Company](#)

Average rating: 5.0 , Number of reviews: 516

Address: 1705 Mariposa St, San Francisco, CA 94107-2334

5. [The Escape Game San Francisco](#)

Average rating: 5.0 , Number of reviews: 451

Address: 150 Kearny St, San Francisco, CA 94108-4802

City-based Restaurant Search (UI)

Based on 132 unique amenities we used the most important ones to create a filter-based search for restaurants.

Note: These 8 filters are fed as 8 different attributes to the restaurant node in the KG.

Restaurants sorted by rating first then number of reviews.

Enter city you want to visit

San Diego, CA

What do you want to explore?

☐ Attractions ☒ Restaurants

Are you looking for vegetarian options?

☒ Yes ☐ No

Are you looking for parking?

☐ Yes ☒ No

Are you looking for takeout?

☐ Yes ☒ No

Are you looking for drive-thru?

☐ Yes ☒ No

Are you looking for reservations?

☒ Yes ☐ No

Are you looking for a place which is good for kids?

☒ Yes ☐ No

Are you looking for a place which is good for dogs?

☒ Yes ☐ No

Are you looking for a place which allows smoking?

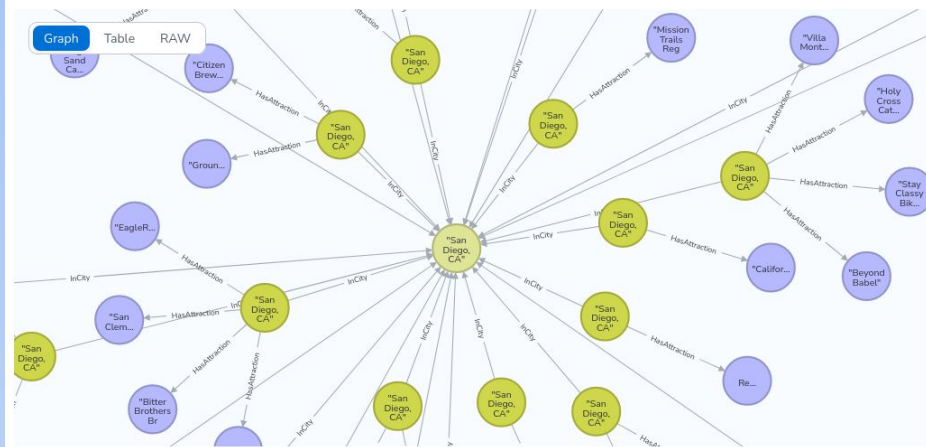
☐ Yes ☒ No

Go!

1. [Moms Pizza and Pasta](#)
Average rating: 5 , Number of reviews: 295
Address: 3422 Tripp Ct, Ste A, San Diego, CA 92121
2. [Werewolf](#)
Average rating: 4.5 , Number of reviews: 4659
Address: 627 Fourth Ave. San Diego. CA 92101

Graphs & Query: City-based search

```
$ MATCH c1 = (z)-[:InCity]-(c:CityNode{city_name:"San Diego, CA"}) MATCH c2 = (z)-[:HasAttraction
(a:AttractionNode) RETURN c1, c2
```

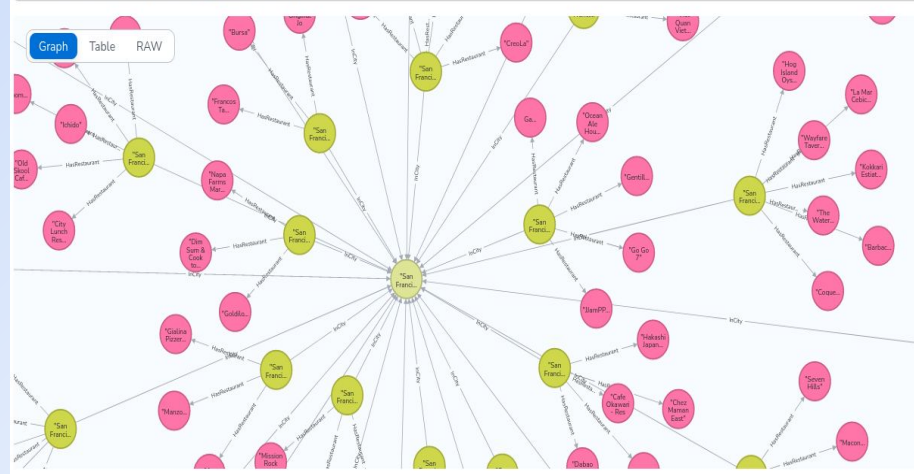


Graph for city-based attraction search.

Graph for city-based restaurant search including filters.



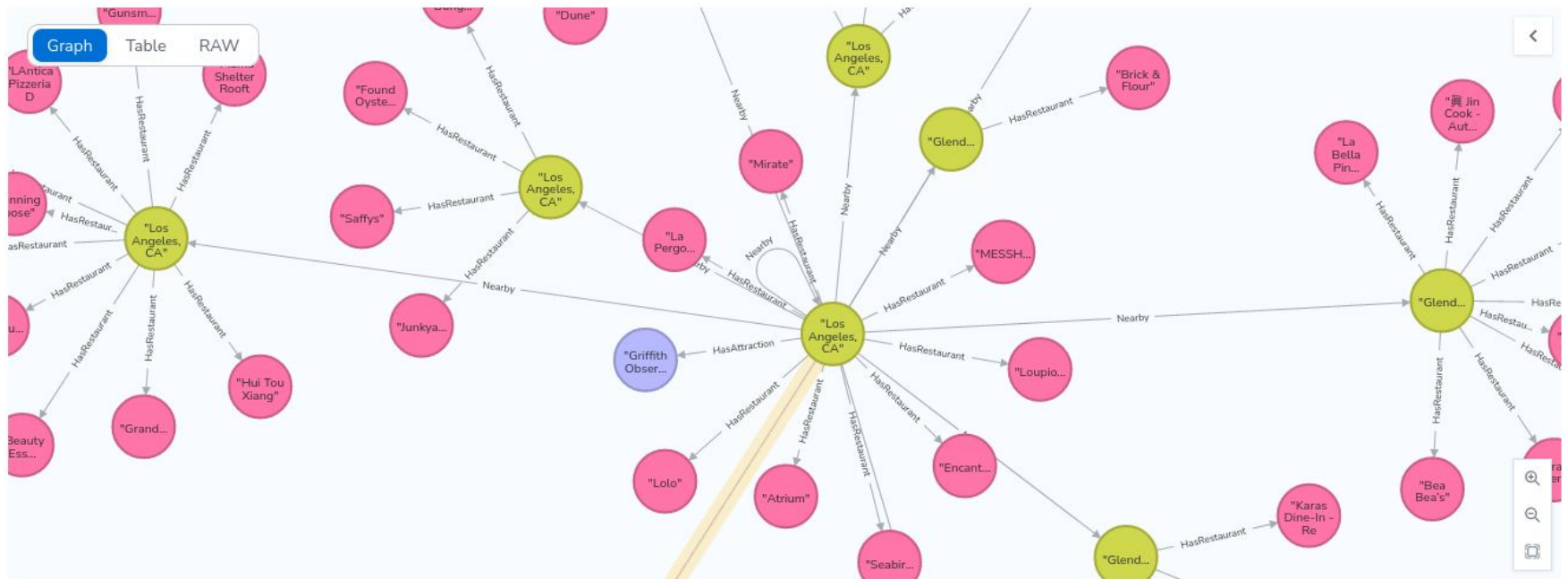
```
$ MATCH c1 = (z)-[:InCity]-(c:CityNode{city_name:"San Francisco, CA"}) MATCH c2 = (z)-[:HasRestaurant]→(r:RestaurantNode) RETURN c1, c2
```



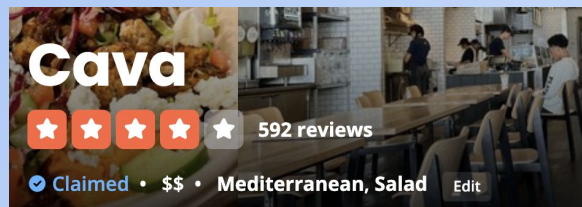
User Interface / Application

Similar Restaurants Recommendation

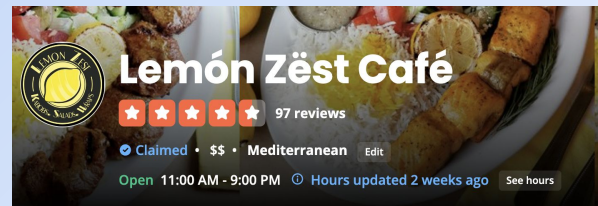
```
$ MATCH c1 = (z)-[:HasAttraction]→(:AttractionNode{attraction_name: "Griffith Observatory"}) MATCH c2 = (z)-[:Nearby]→(nz) MATCH c3 = (nz)-[:HasRestaurant]→(r) RETURN c1,c2,c3
```



Similarity Measurement for Recommendation



Cuisine Similarity
↔
Jaccard



Amenities and More	
Offers Delivery	Offers Takeout
No Reservations	Vegan Options
Many Vegetarian Options	Masks required
Staff wears masks	Accepts Credit Cards
Outdoor Seating	Touristy, Trendy
Quiet	Offers Catering
Good for Groups	Garage Parking, Street Parking, Validated Parking
Free Wi-Fi	Happy Hour Specials
Wheelchair Accessible	Gender-neutral restrooms
Heated outdoor seating	Covered outdoor seating
No Waiter Service	No Alcohol
No Drive-Thru	No TV
Private dining	Bike Parking

Amenities Similarity
↔
Jaccard

Amenities and More	
Health Score A	Offers Delivery
Offers Takeout	No Reservations
Masks required	Vegan Options
Accepts Credit Cards	Accepts Android Pay
Accepts Apple Pay	Outdoor Seating
Casual, Classy	Moderate Noise
Good for Groups	Good For Kids
Good for Lunch, Dinner	Garage Parking, Street Parking, Private Lot Parking
Gender-neutral restrooms	Offers Catering
No Waiter Service	No Wi-Fi
No Alcohol	Dogs Not Allowed
No TV	Bike Parking

$$\text{Similarity Score} = 0.67 * \text{Cuisine Similarity} + 0.33 * \text{Amenities Similarity}$$

Recommendation System (UI)

Evaluation of Recommendation:

We asked 20 friends to enter inputs for different sections of UI and collected their feedback on whether they found the results to be satisfactory or not. Out of 20 searches, 17 of them were satisfied.

Common reasons for non-satisfactory results:

- Sometimes cuisines might not exactly match. For example: 'Middle Eastern' and 'Falafel'.
- Lack of restaurant data in a given location.

Sample image of recommendation UI

The screenshot shows a web interface for a recommendation system. It has two input fields: 'Enter Attraction you will be visiting' with a dropdown menu showing 'The Getty Center', and 'Enter Restaurant you like' with a dropdown menu showing 'Cava'. Below these is a 'Go!' button. The results are a list of five restaurants, each with its name, average rating, number of reviews, and address. A large black curly bracket on the right side of the results list points to the text 'Restaurants sorted in order of similarity to the input restaurant.'

Rank	Restaurant Name	Average rating	Number of reviews	Address
1.	Lemón Zést Café	5	95	14622 Ventura Blvd, Unit 103, Sherman Oaks, CA 91403
2.	PSY Street Kitchen	4.5	501	15030 Ventura Blvd, Los Angeles, CA 91403
3.	Mizlala	4	696	4515 Sepulveda Blvd, Los Angeles, CA 91403
4.	Armav	4.5	23	970 Monument St, Ste 119, Los Angeles, CA 90272
5.	The Pastor's Grill	4.5	7	4850 Van Nuys Blvd, Sherman Oaks, CA 91403

Restaurants sorted in order of similarity to the input restaurant.

2nd Technical Challenge

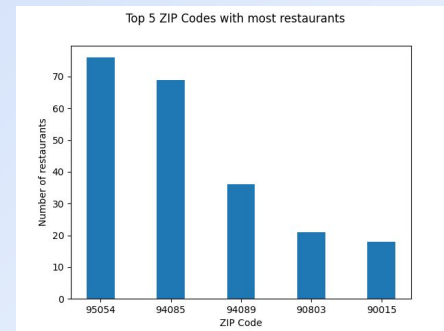
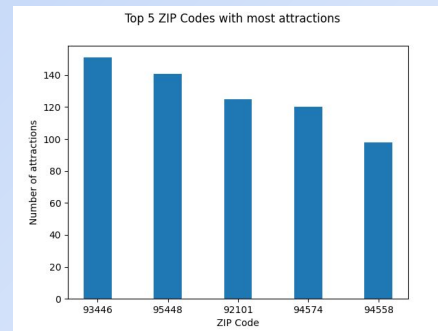
Problem: Evaluating and assessing the quality of our Knowledge Graph was a big challenge.

Solution: We tackled the challenge using following measure covered in the lecture:

1. Completeness:
 - a. Restaurants and attractions data had to be dropped due to missing critical information. (eg: Zip Code, Reviews, Rating)
 - b. 43 out of 1596 standard ZIP codes removed due to absent nearby ZIP codes data.
2. Timeliness: Data is up to date.
3. Reputation: Data comes from reputed/trusted websites like Tripadvisor, Yelp and official US website for zip codes.

Data Quality / Assessment

4. Objectivity:



5. Accuracy:

- Sometimes Selenium failed to extract complete amenities list that resulted in some data loss.
- ZIP codes from nearby states (AZ , NV, OR) appeared in dataset and had to be dropped.

Evaluation / KG Quality

Functional Requirements	Satisfied	Scope for Improvement
Accurate, Complete & Topical Information	✓	
Correct and up-to-date data	✓	
Accessible and Performant	✓	

Non-Functional Requirements	Satisfied	Scope for Improvement
Include documentation for modeling approach		✓
Ontologies, Constraints and Declarations		✓
Internally Consistent	✓	
Based on mature technologies	✓	

Learnings

1. Combining structured (list of attractions and restaurants, list of zip codes) and unstructured data (address, ratings in the form of stars, price in the form of \$\$\$) becomes easier with a KG.
2. Querying and retrieving information (nearby restaurants or attractions using zip code) is faster using KG than other conventional data structures in python.

