

GASTRONOMIE



AI

by Ajay Mahendra Surve, Chitvan Nagaich, Maurya Sasanka Bhima, Prakriti Sharma, & Sreeya Guha



when there's a
million dining
options....

GASTRONOME_{AI}

GASTRONOME

AI

G.

OUR PICKS for you

Sort By ▾



5 POINTS

📍 23 mi



INCA'S
PERUVIAN CUISINE
📍 23 mi



TERESA'S
MOSAIC CAFÉ

📍 24 mi



THE CORONET
📍 25 mi



1912
BREWING COMPANY

📍 28 mi



AGUSTÍN
KITCHEN
📍 28 mi

G.

OUR PICKS
for you

Sort By ↘



5 POINTS

📍 2.3 mi



INCA'S
PERUVIAN CUISINE



5 POINTS

📍 2.3 mi

INCA'S PERUVIAN CUISINE

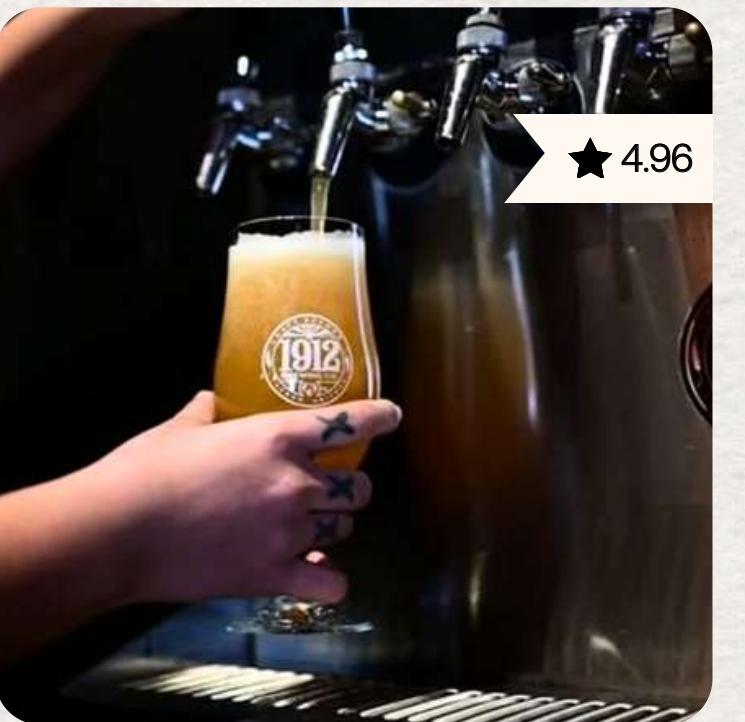


★ 4.97

28 mi

TERESA'S MOSAIC CAFÉ

📍 2.4 mi



★ 4.96

26 mi

1912 BREWING COMPANY

📍 2.8 mi



★ 4.94

28 mi

AGUSTÍN KITCHEN

ABOUT US

- AI-powered dining discovery platform that provides personalized recommendations to its users
- Leveraging Yelp reviews, Sentiment Analysis, and Machine Learning
- Highlight any hidden gems or treasures



How are we proceeding?

Yelp Datasets

- Restaurant metadata (name, location, category, ratings)
- User reviews (text data for sentiment & topic modeling)
- Check-in & user preferences



Data Preprocessing

- Remove missing values, clean text for NLP, Tokenization and Lemmatization, and Merging Datasets
- Ensures clean and structured data for better NLP performance
- Reduces noise in restaurant descriptions & reviews



NLP: Topic Modeling

- Techniques Used:
 - BERTopic → Identifies hidden themes in user reviews.
 - Zero-Shot Classification → Assigns meaningful restaurant categories (e.g., Fine Dining, Fast Food).
- Helps in understanding user sentiment & restaurant specialties.
- Improves category-based filtering for recommendations.

Cosine Similarity Model



- Convert restaurant categories & topics into numerical vectors (TF-IDF / Embeddings)
- Compute similarity scores between restaurants
- Rank restaurants based on similarity to user preferences

Recommendation UI

- Users can search based on cuisine, location, and preferences
- The system ranks the top-N most relevant restaurants

EXPLORATORY DATA ANALYSIS

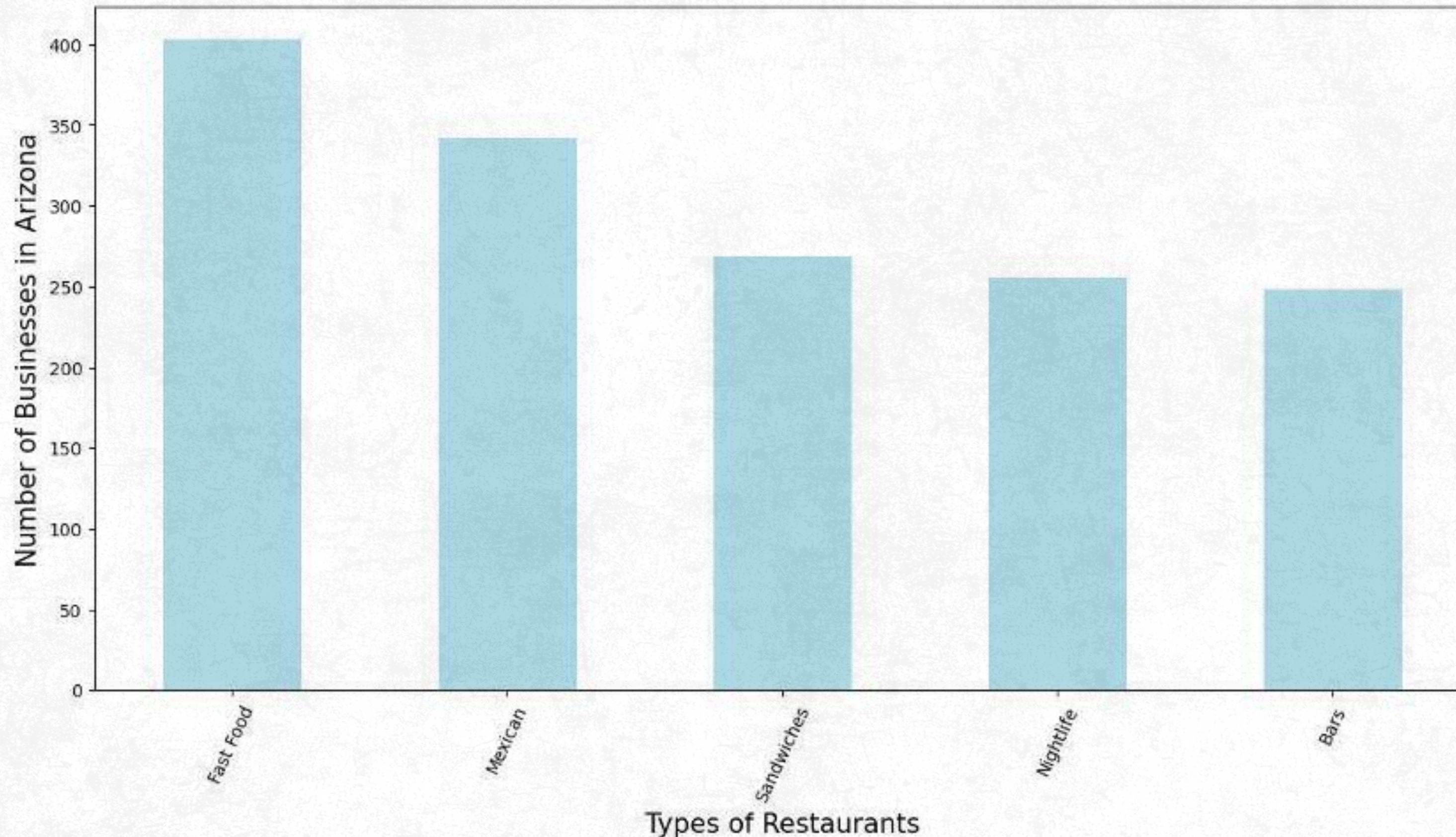
- EDA will allow us to gain insight into the structure, trends, and patterns in the dataset
- Will help identify key themes in customer sentiment, rating distribution, and review characteristics
- Lays the groundwork for NLP analysis and business insights

Data Filtration

- Yelp data had an exhaustive amount of information related to Business, Reviews, Users, Check-in, and Tips.
- We narrowed down our analysis to food businesses based in Arizona to create a personalized recommendation system for Gastronome users.
- The preprocessed data was fed to NLP techniques and recommendation models for the final output.

Top 5 Restaurant Categories

Top 5 Restaurant Categories in Arizona

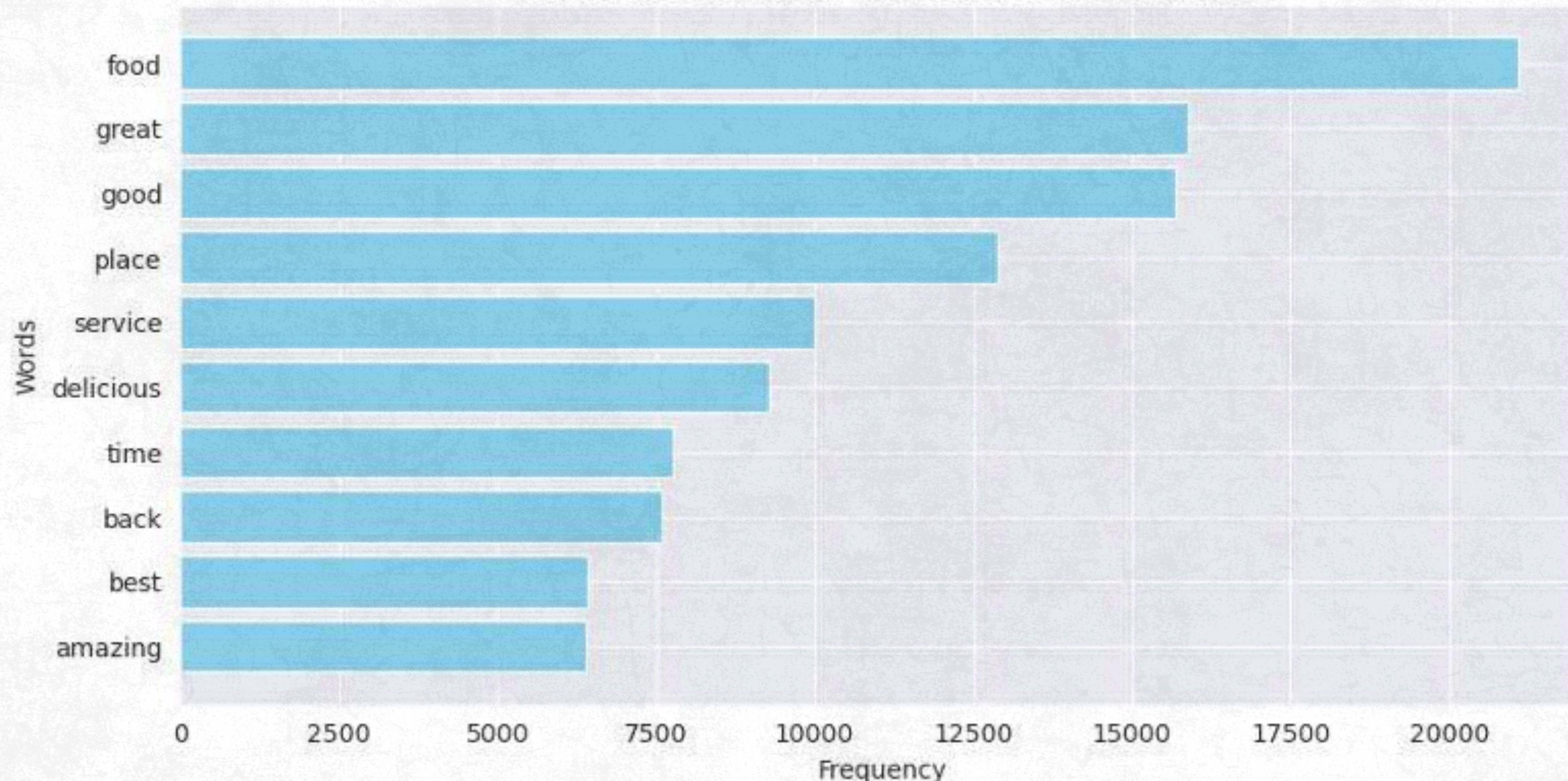


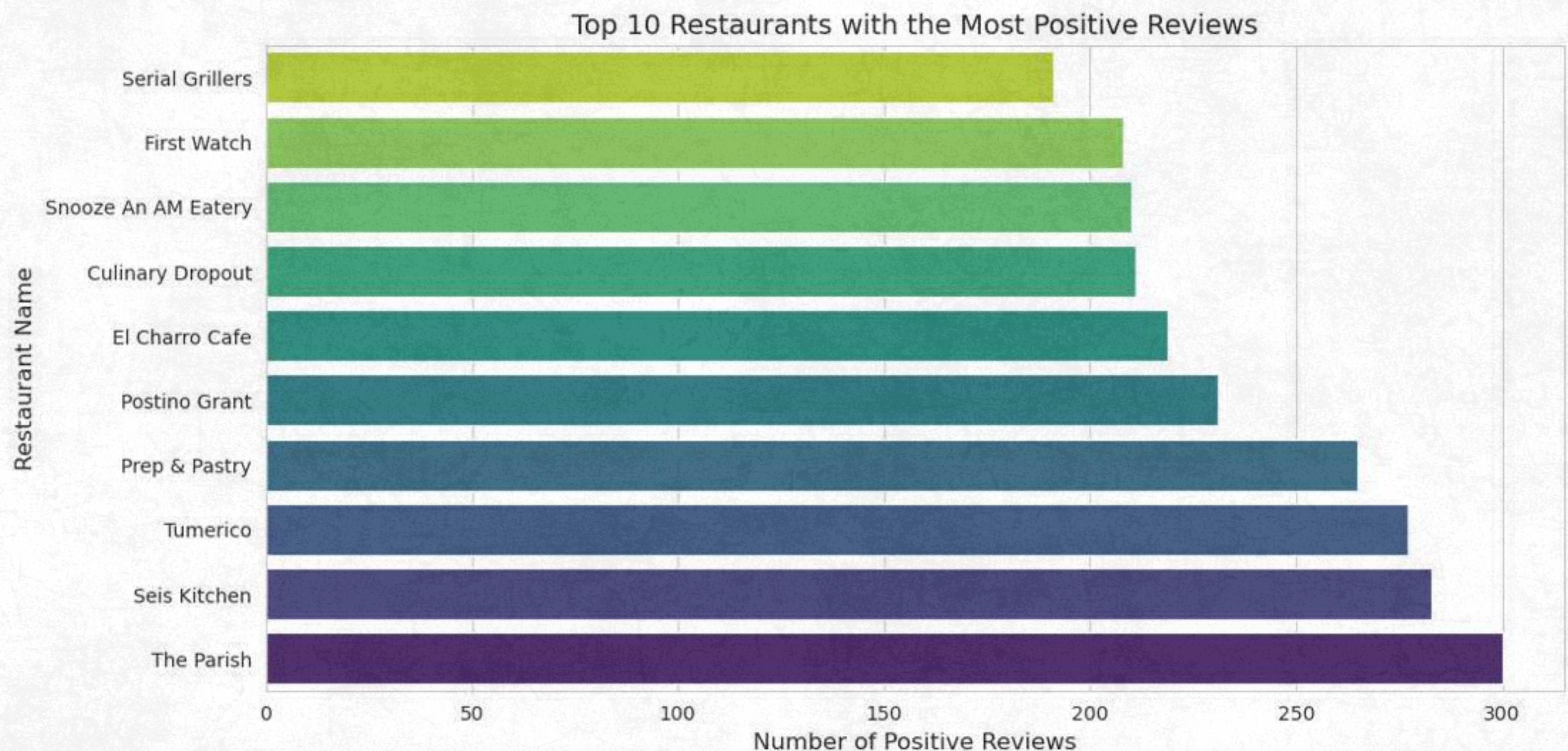
Positive Reviews

place fresh little
chicken
friendly us
came
super friendly
favorite
back service
always
best
im pizza
order
well time
delicious
ordered
tucson
good
one amazing
got even
made wait
first
definitely
nice try
love menu
recommend

Positive Reviews

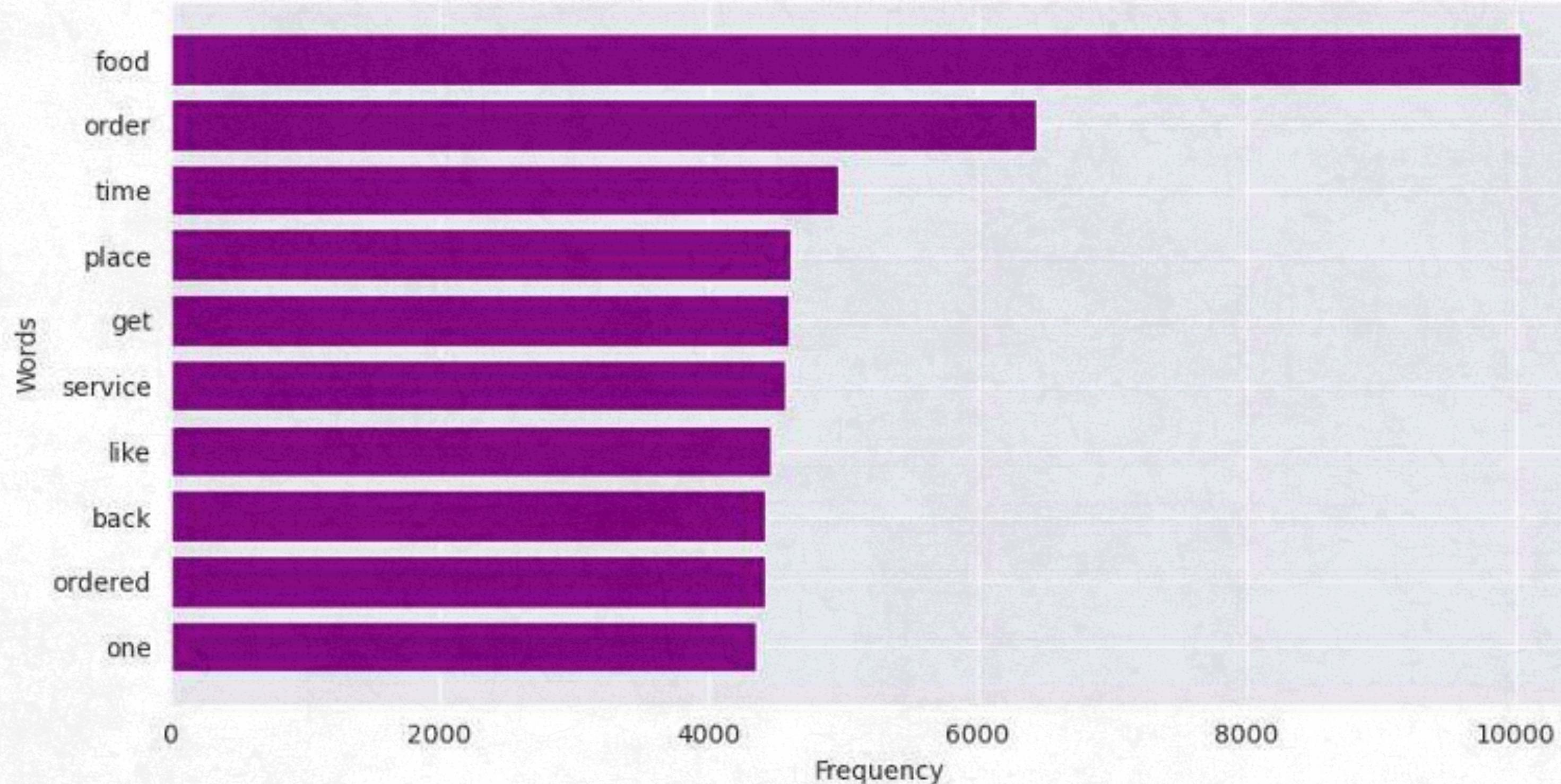
Top 10 Most Common Words in Positive Reviews





Negative Reviews

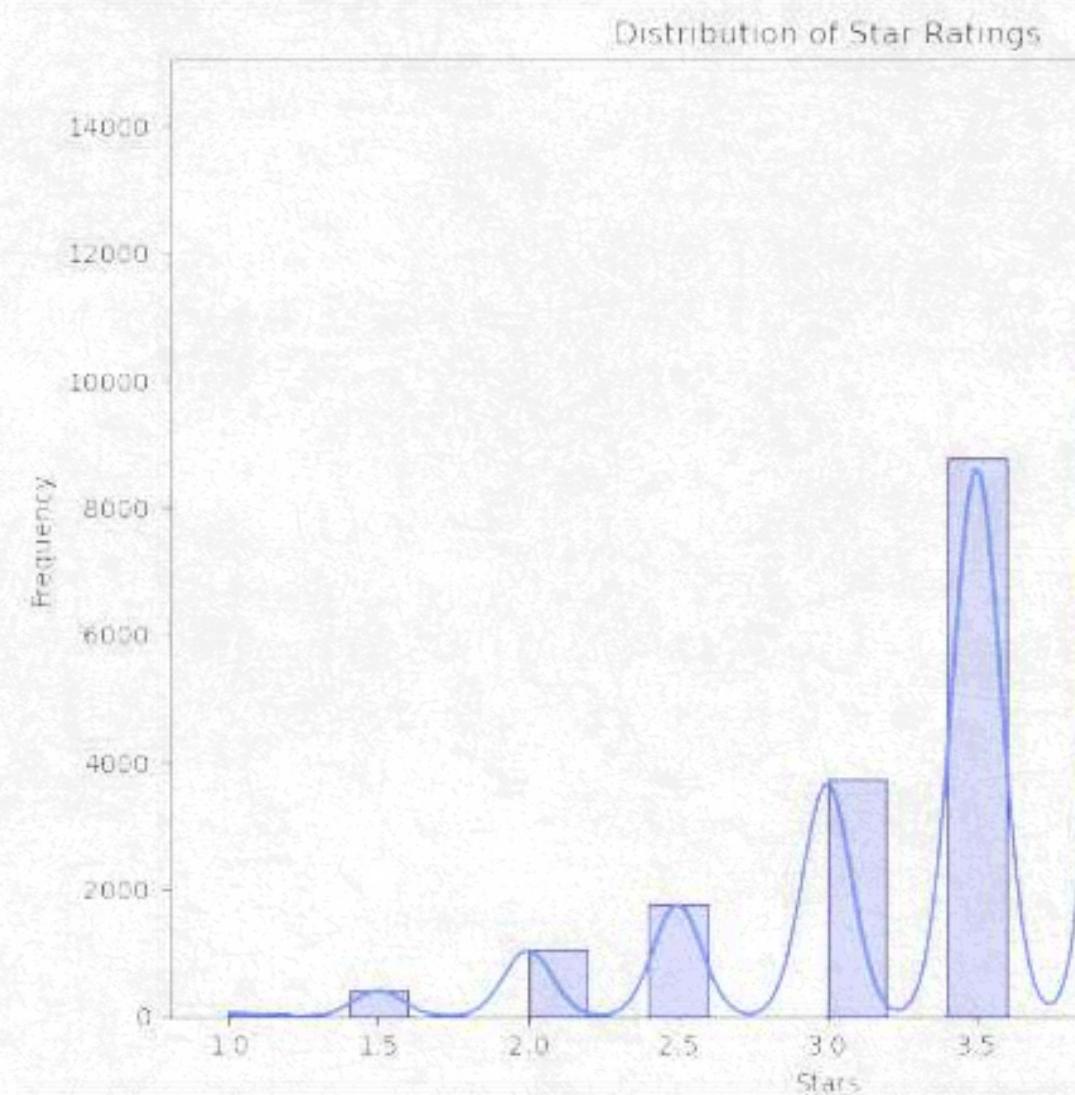
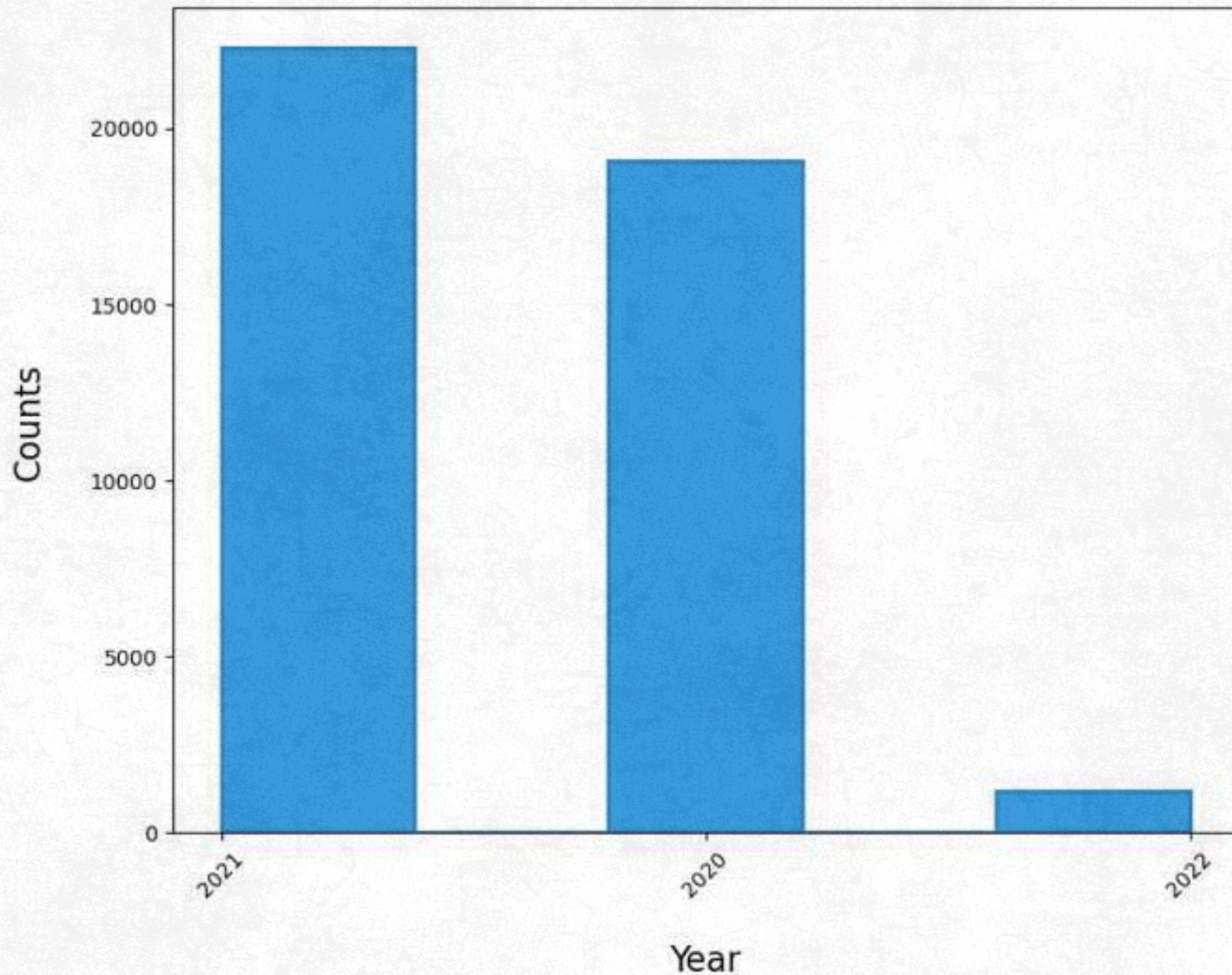
Top 10 Most Common Words in Negative Reviews

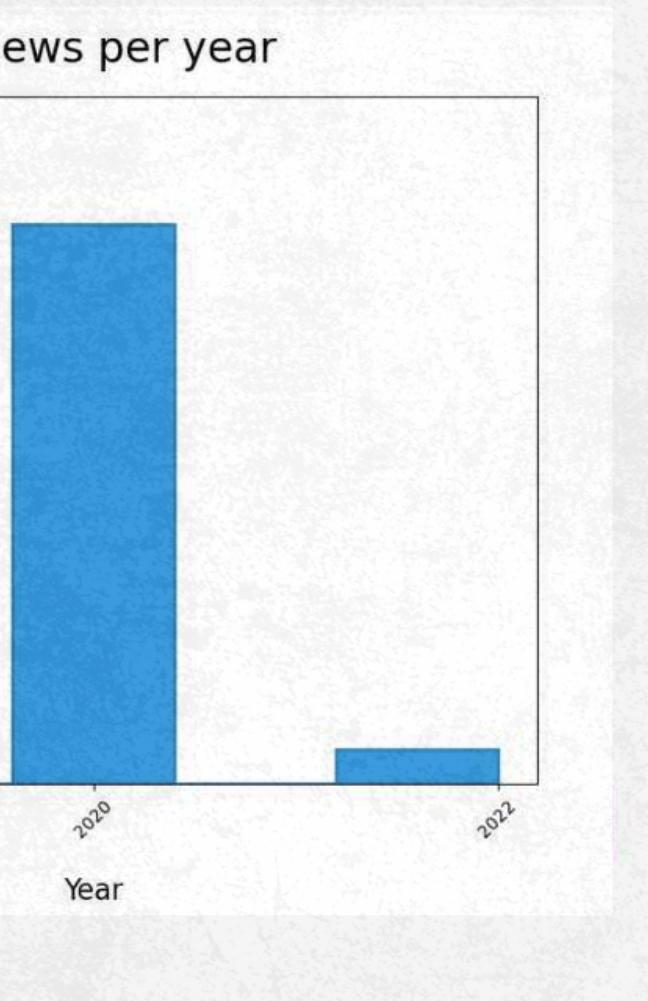


Negative Reviews

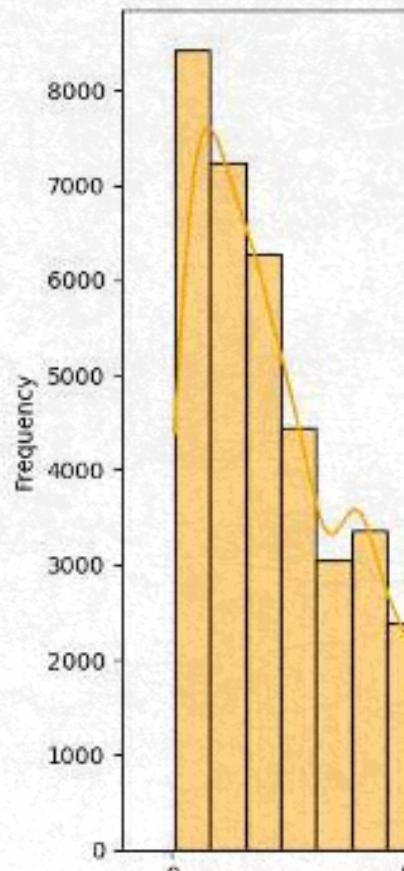
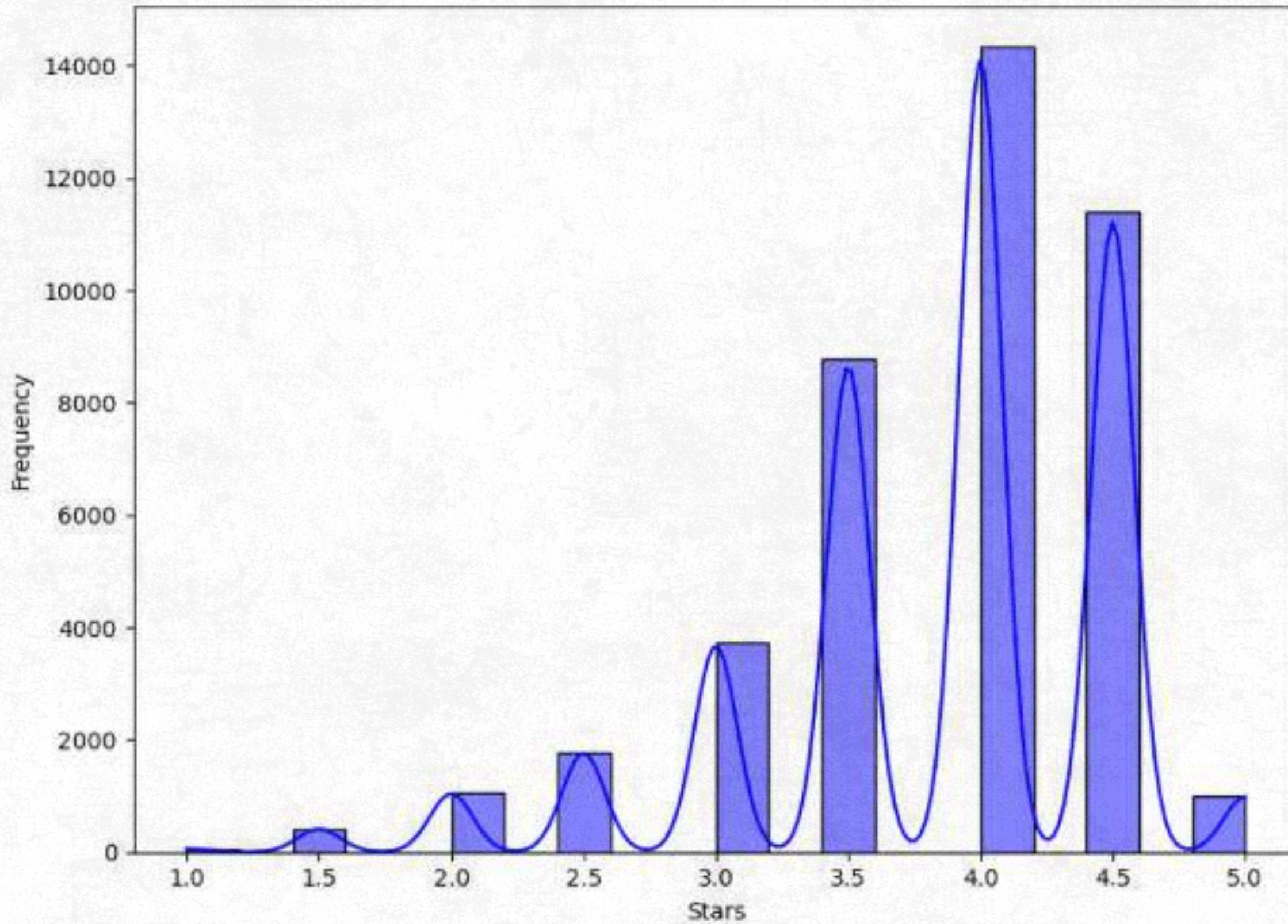
ordered us got told back even took went place bad didnt really come said chicken f food minutes im table came one way dont service never customer go wait restaurant first time manager experience pizza

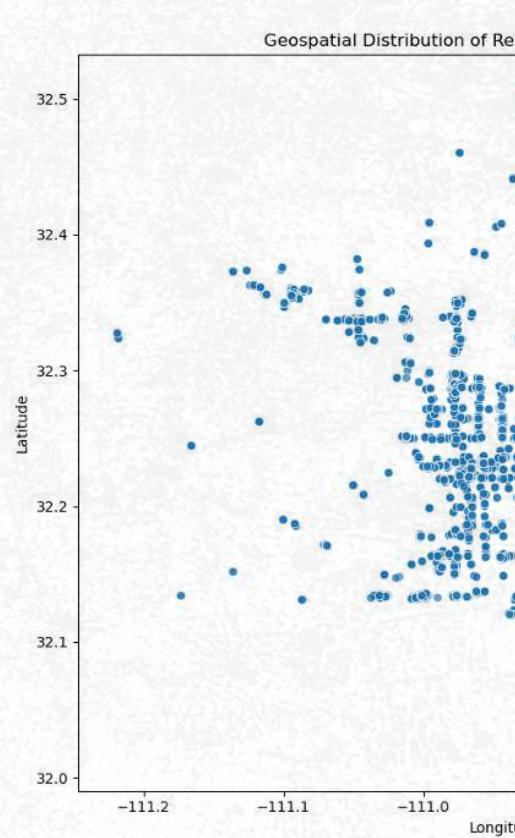
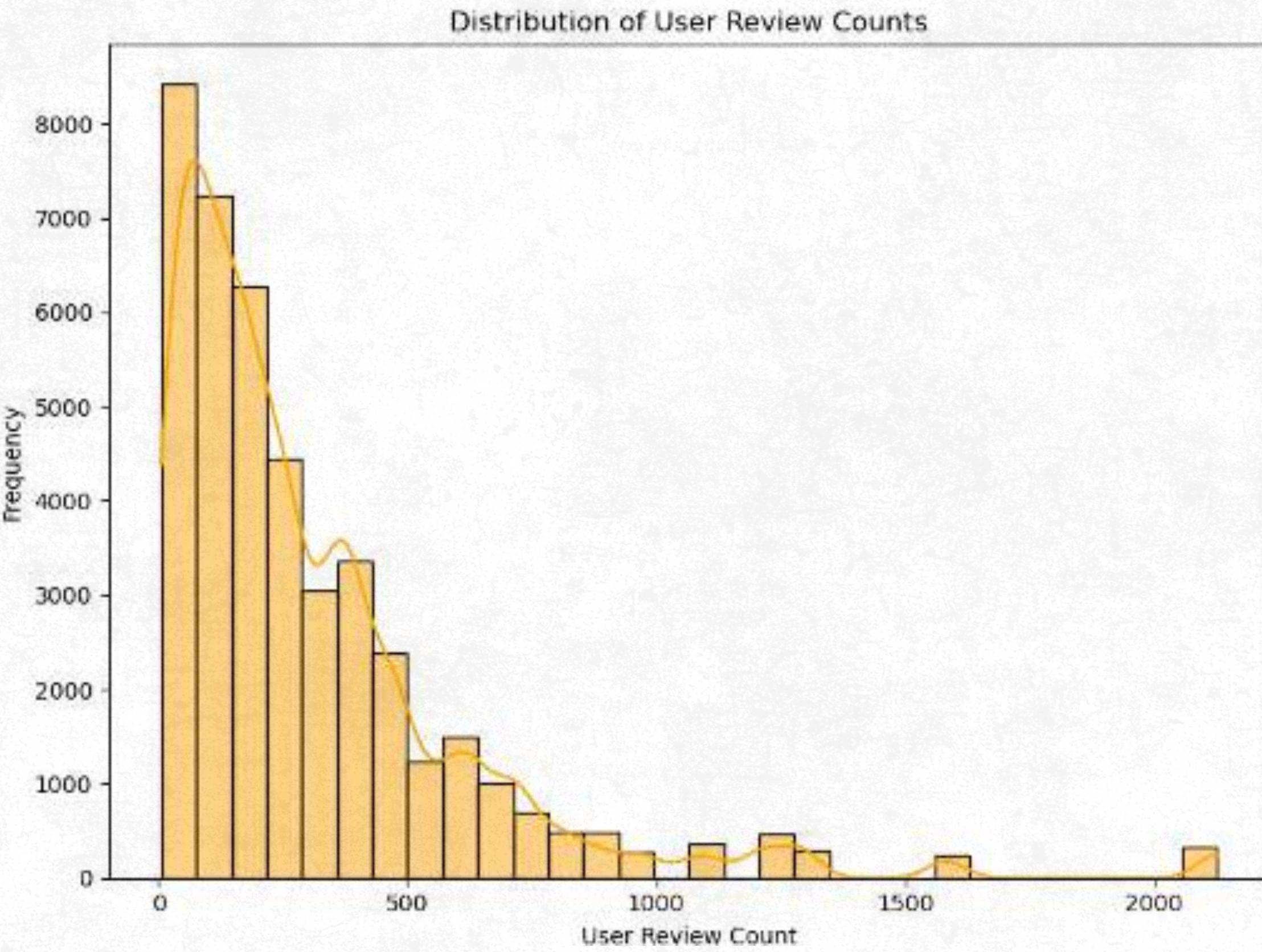
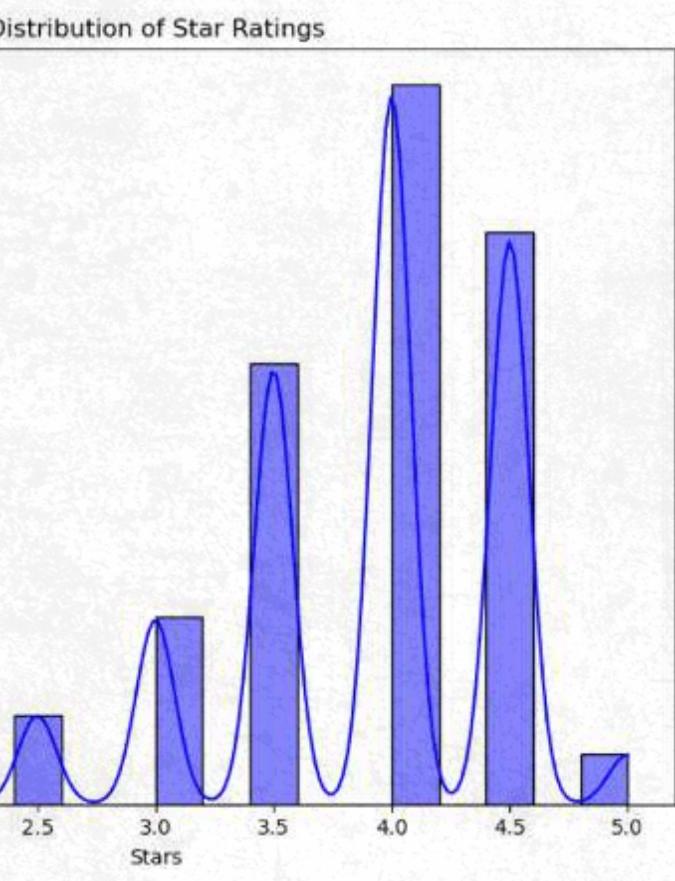
Reviews per year

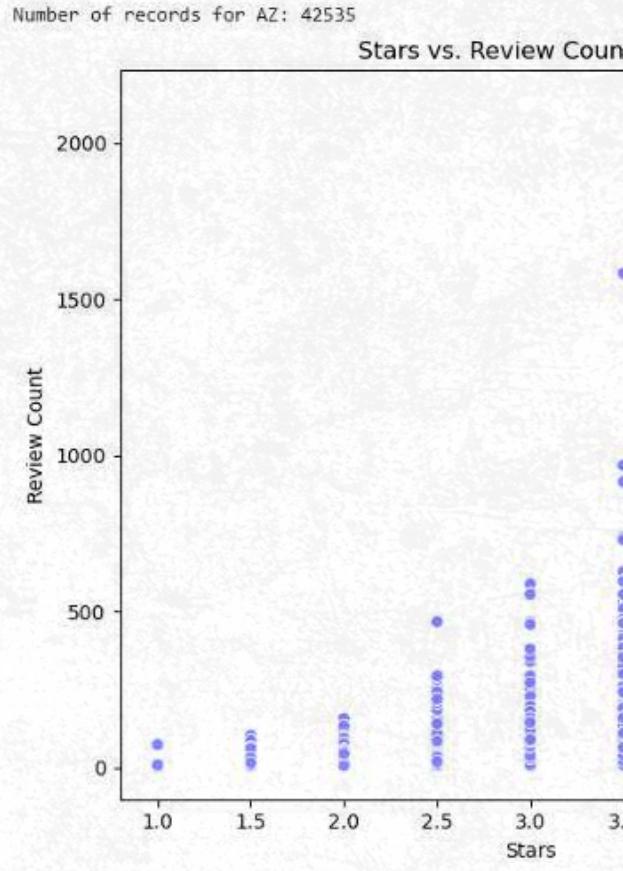
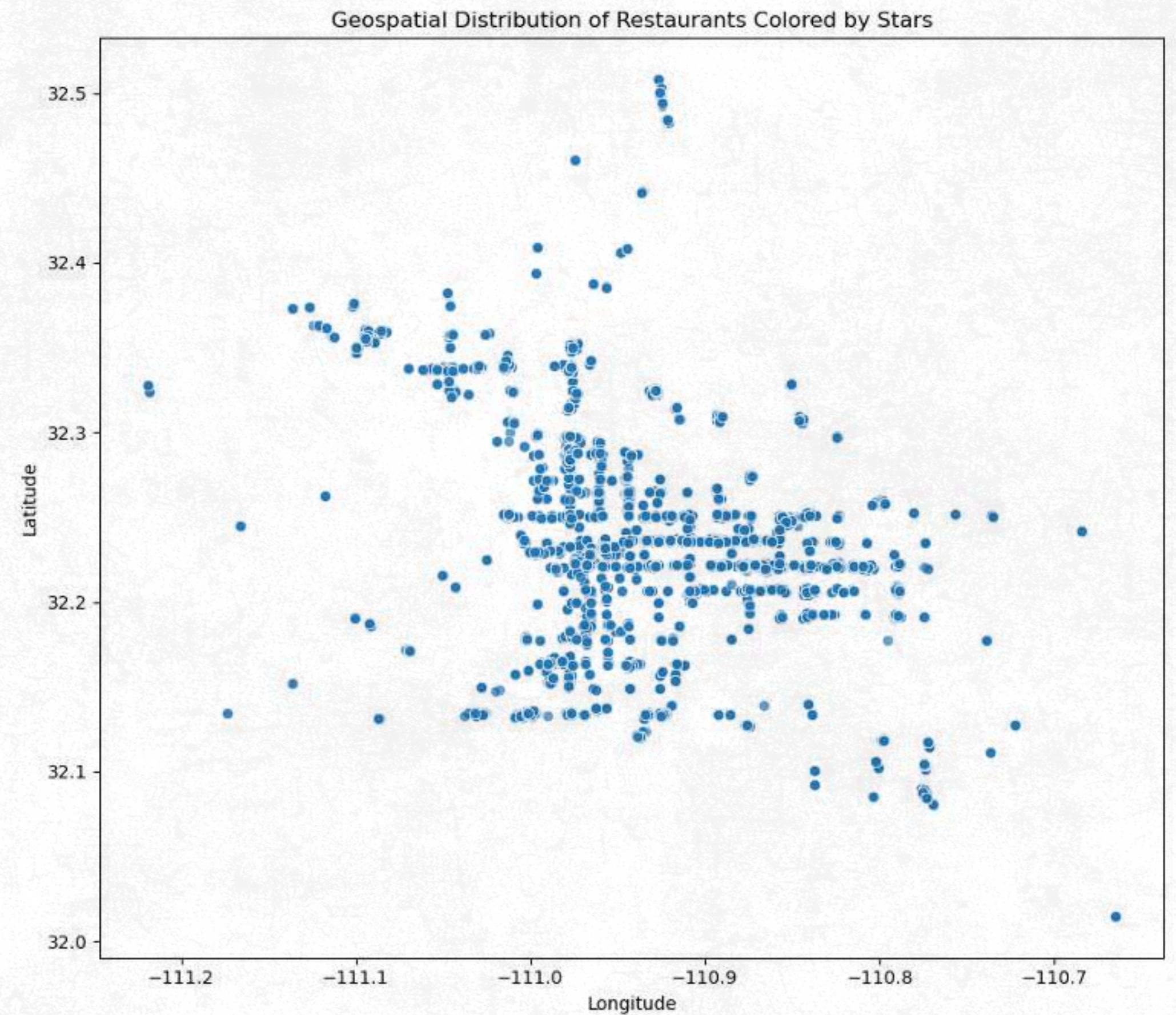




Distribution of Star Ratings

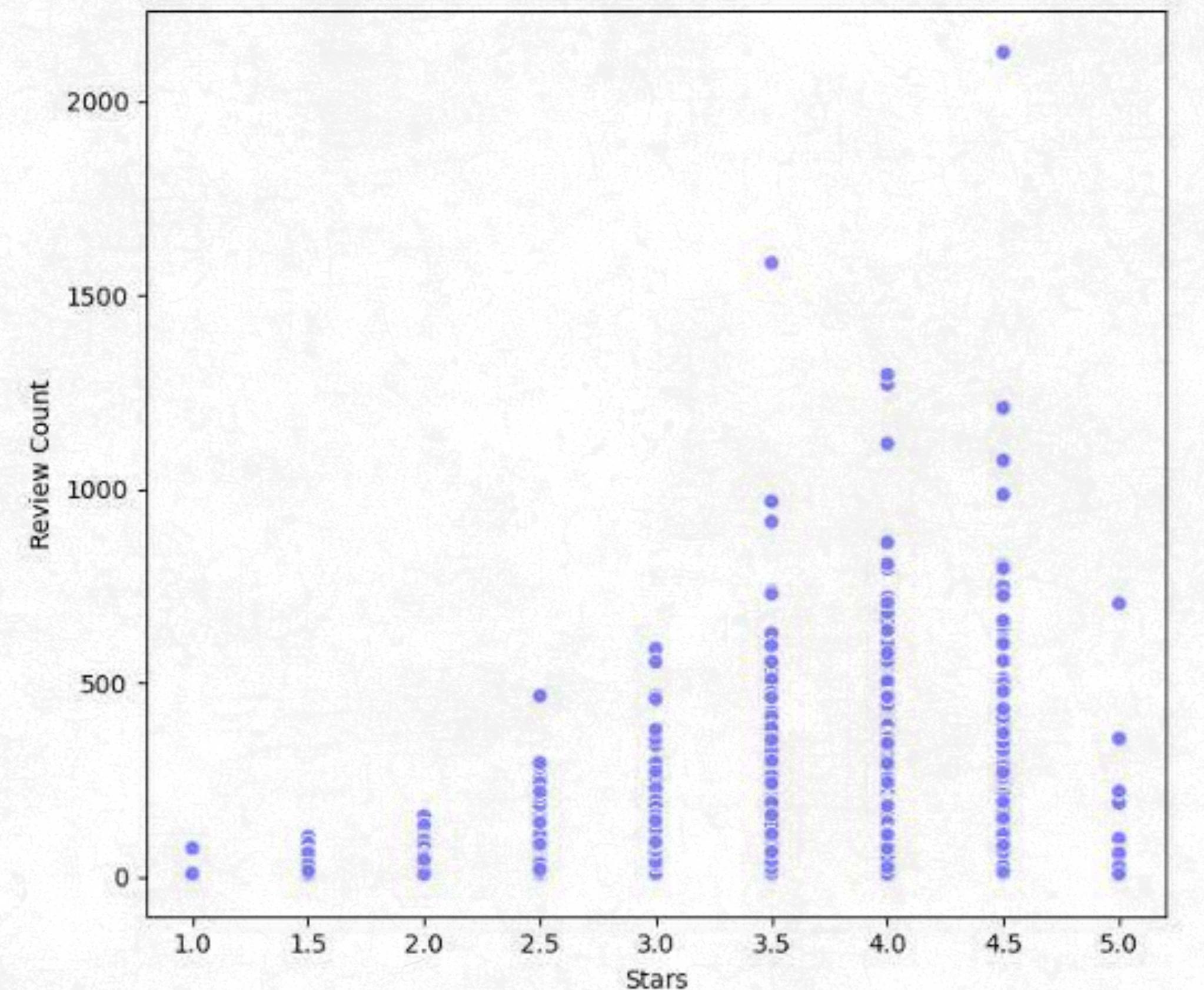
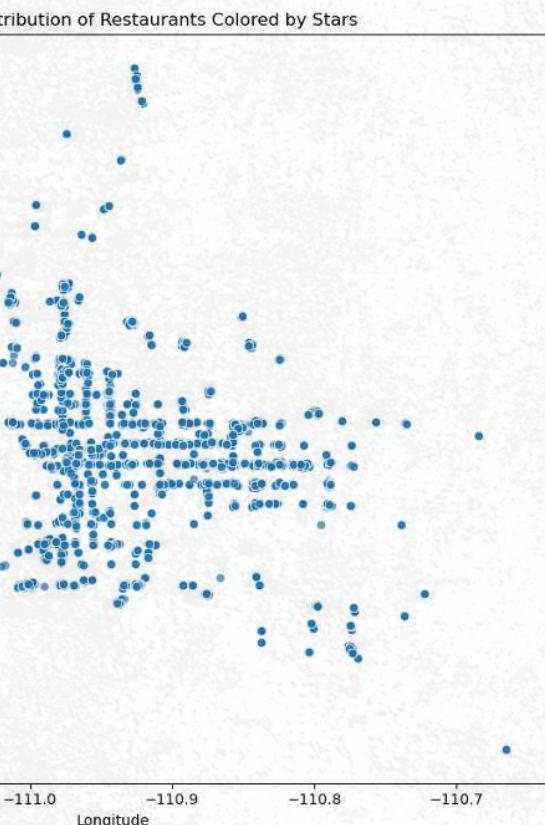




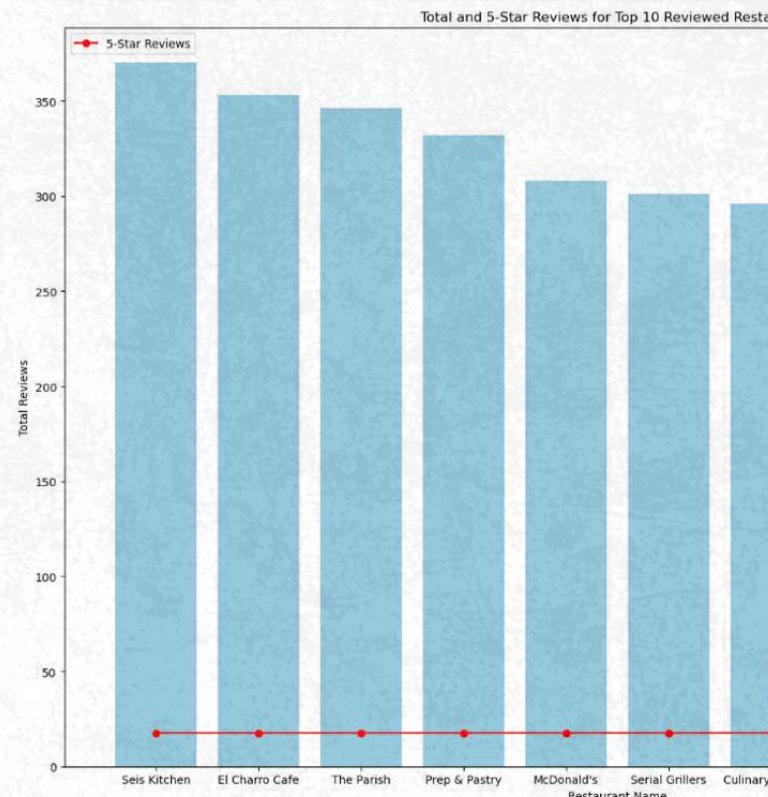


Number of records for AZ: 42535

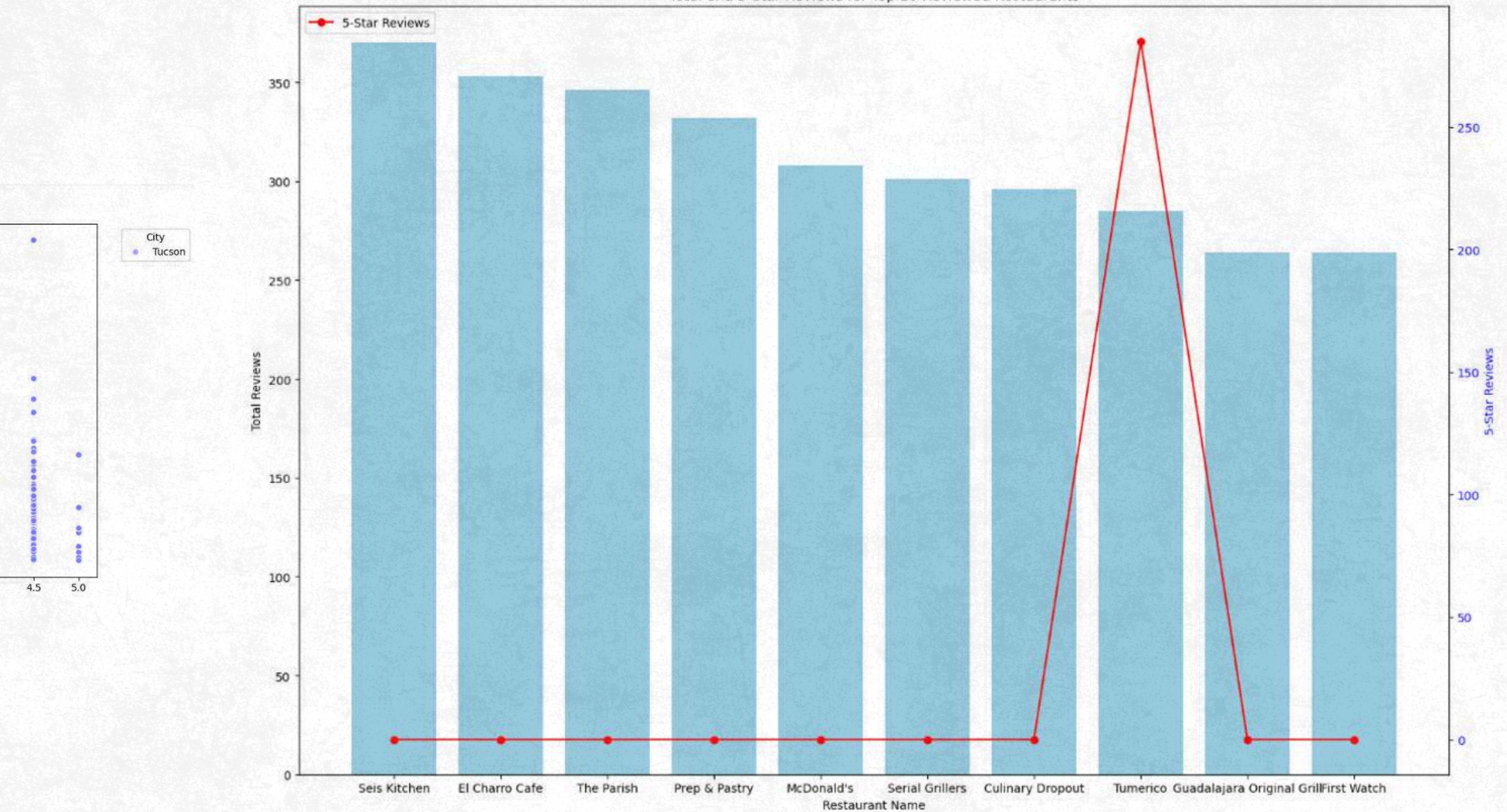
Stars vs. Review Count in AZ



City
Tucson



Total and 5-Star Reviews for Top 10 Reviewed Restaurants



NATURAL LANGUAGE PROCESSING

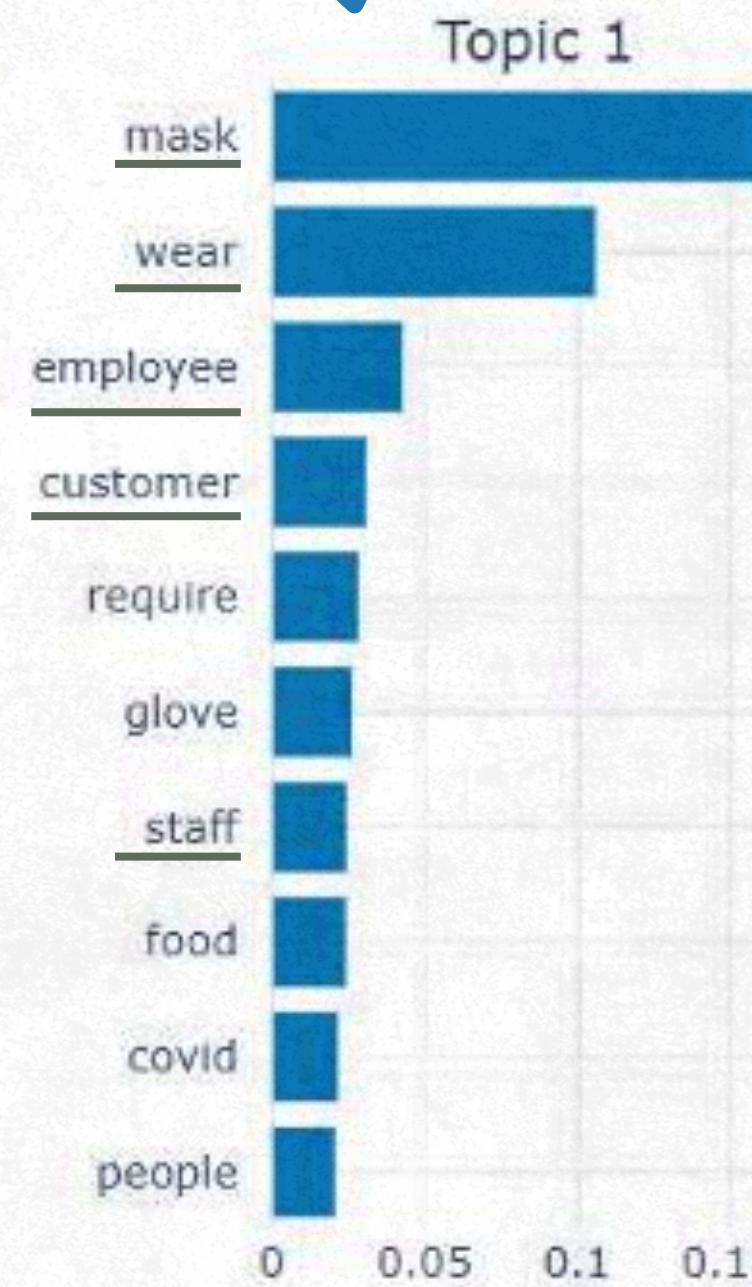
- NLP uncovers insights from text reviews to enhance business decisions through:
 - Sentiment Analysis: Detect positive, neutral, negative reviews
 - BERTopic: Extract dominant topics from reviews
 - Keyword Extraction: Identify commonly mentioned food and service elements

BERTopic

Indicates that reviews discuss taste, food quality, and customer satisfaction



Food Quality



Service Quality

Relates to service experience, staff behavior, and safety measures

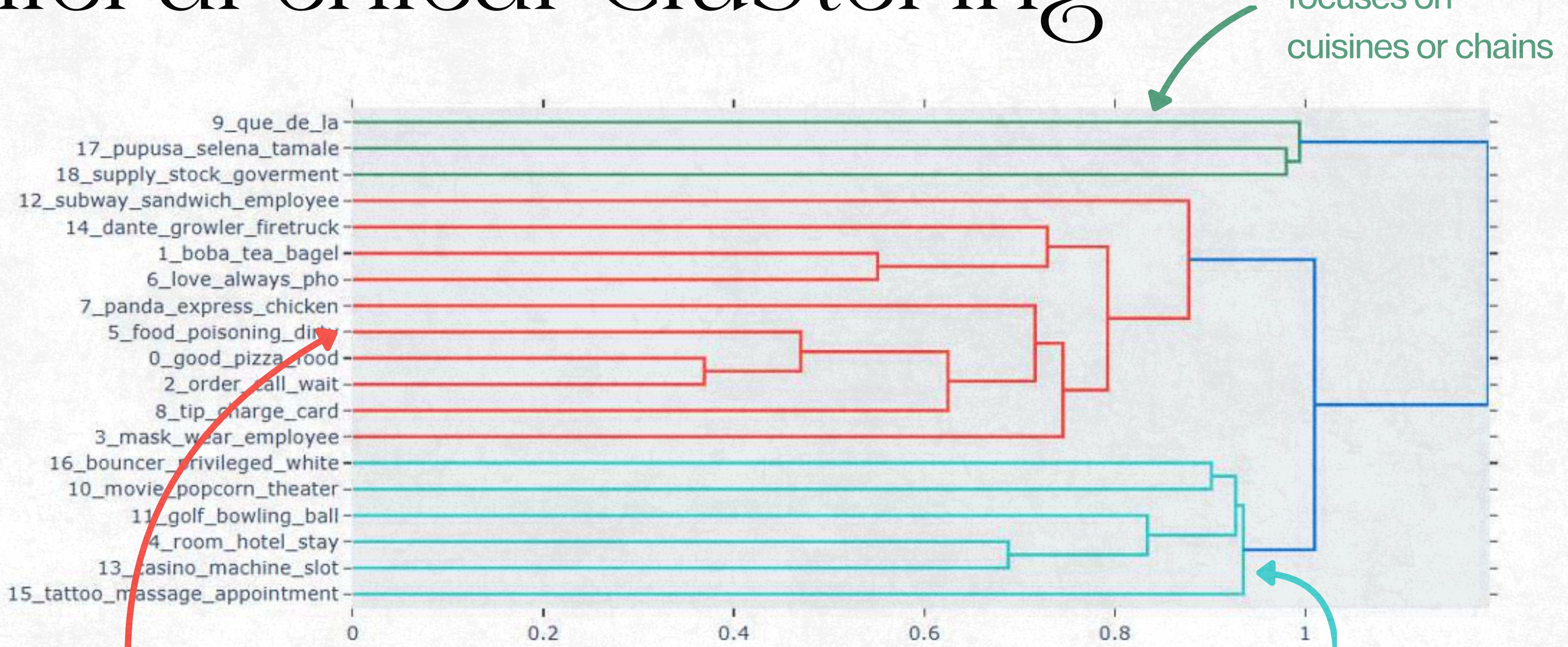
Topic Word Scores



Ambiance

Captures discussions about the environment, comfort, and overall ambiance of a restaurant or dining place

Hierarchical Clustering

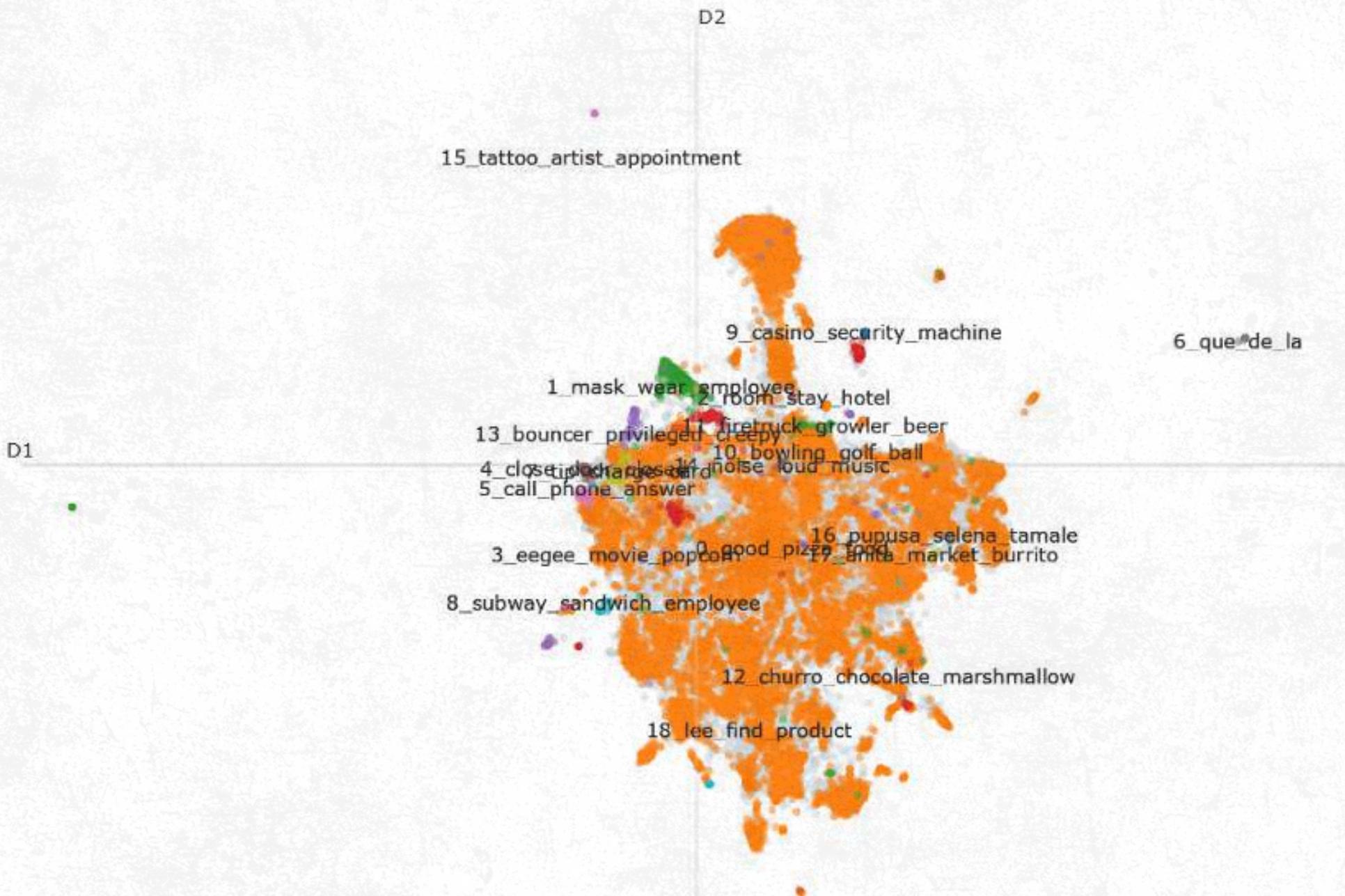


Red Cluster focuses on
service-related or
hygiene issues

Green Cluster
focuses on
cuisines or chains

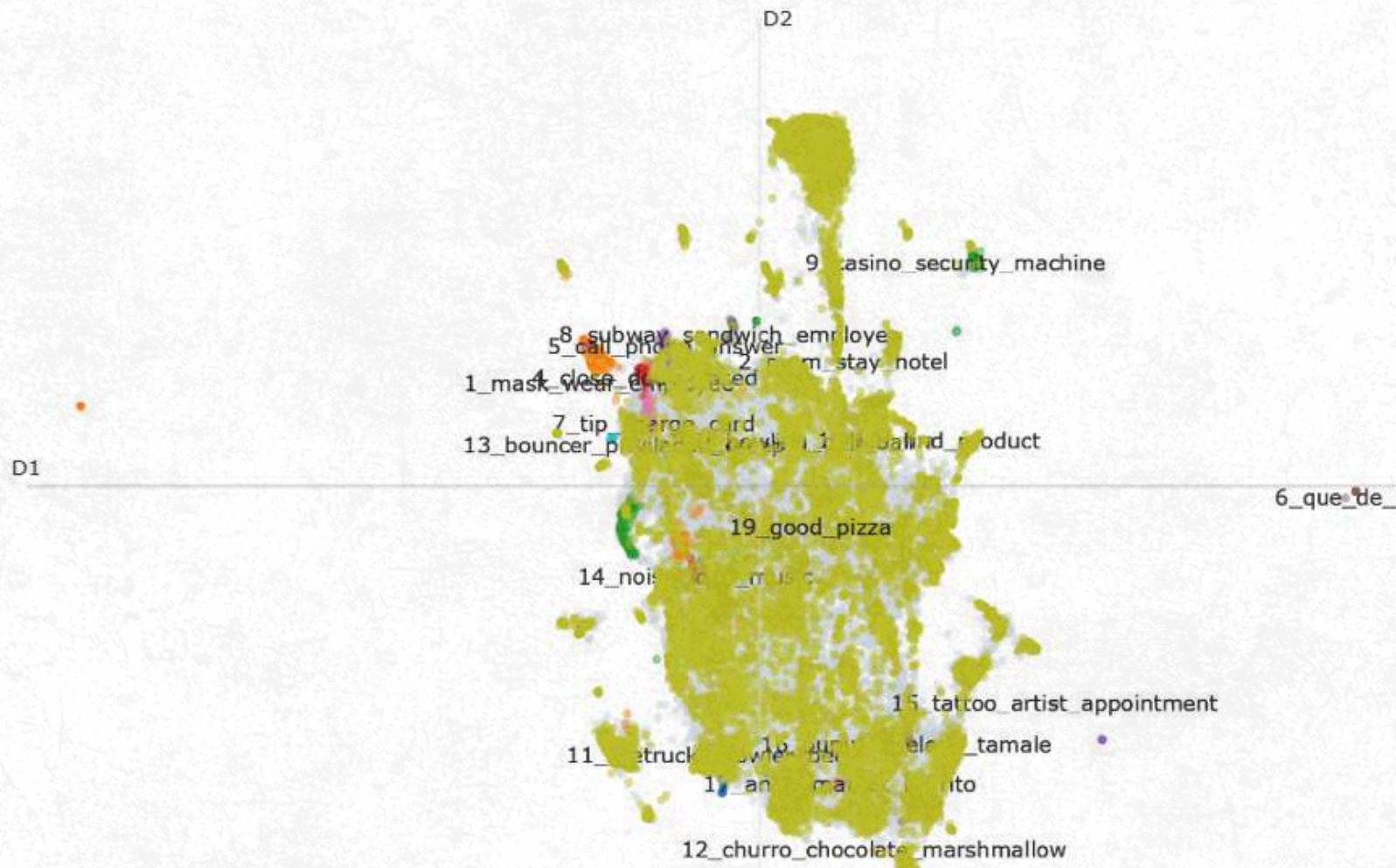
Blue cluster focuses on entertainment
venues with food services, such as
movie theaters, hotels, and casinos

Documents & Topics



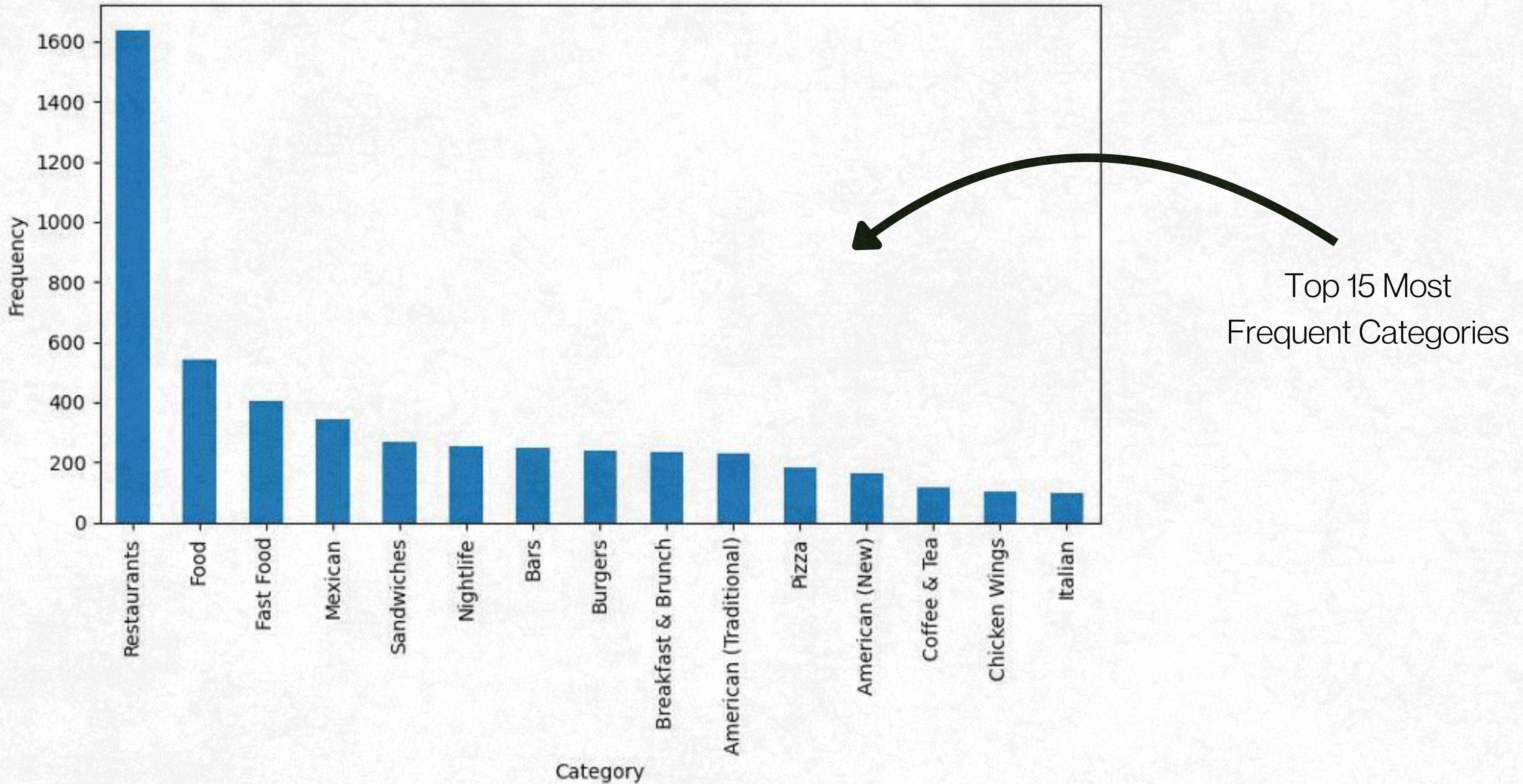
- **Dense Center Cluster** → Most topics are overlapping, meaning many reviews share similar discussions
- **Outlier Topics** (like "tattoo_artist_appointment" and "casino_security_machine") → These outlier topics are specific and do not overlap much with others
- **Well-Separated Topics** → Topics like "mask_wear_employee" and "good_pizza_food" are clearly separated, indicating strong distinct themes

Hierarchical Documents & Topics



- Topic "good_pizza_food_order_place" is the largest cluster → This means a significant number of reviews are focused on food quality and ordering experience
- Distant Topics (e.g., "casino_security_machine") → Indicates that some topics are very distinct and not related to general restaurant themes
- Hierarchical Structure Helps Group Similar Reviews → Topics related to food ordering, customer experience, and ambiance are naturally grouped.

Category Segregation Using Huggingface



Category Segregation Using Huggingface



- "Restaurants," "Fast Food," "Breakfast & Brunch," "Nightlife," and "Burgers" are among the **most discussed** categories
- Diverse cuisine types (Mexican, Italian, Chinese, Sushi) appear, indicating a variety of restaurant options

RESULTS & BUSINESS INSIGHTS

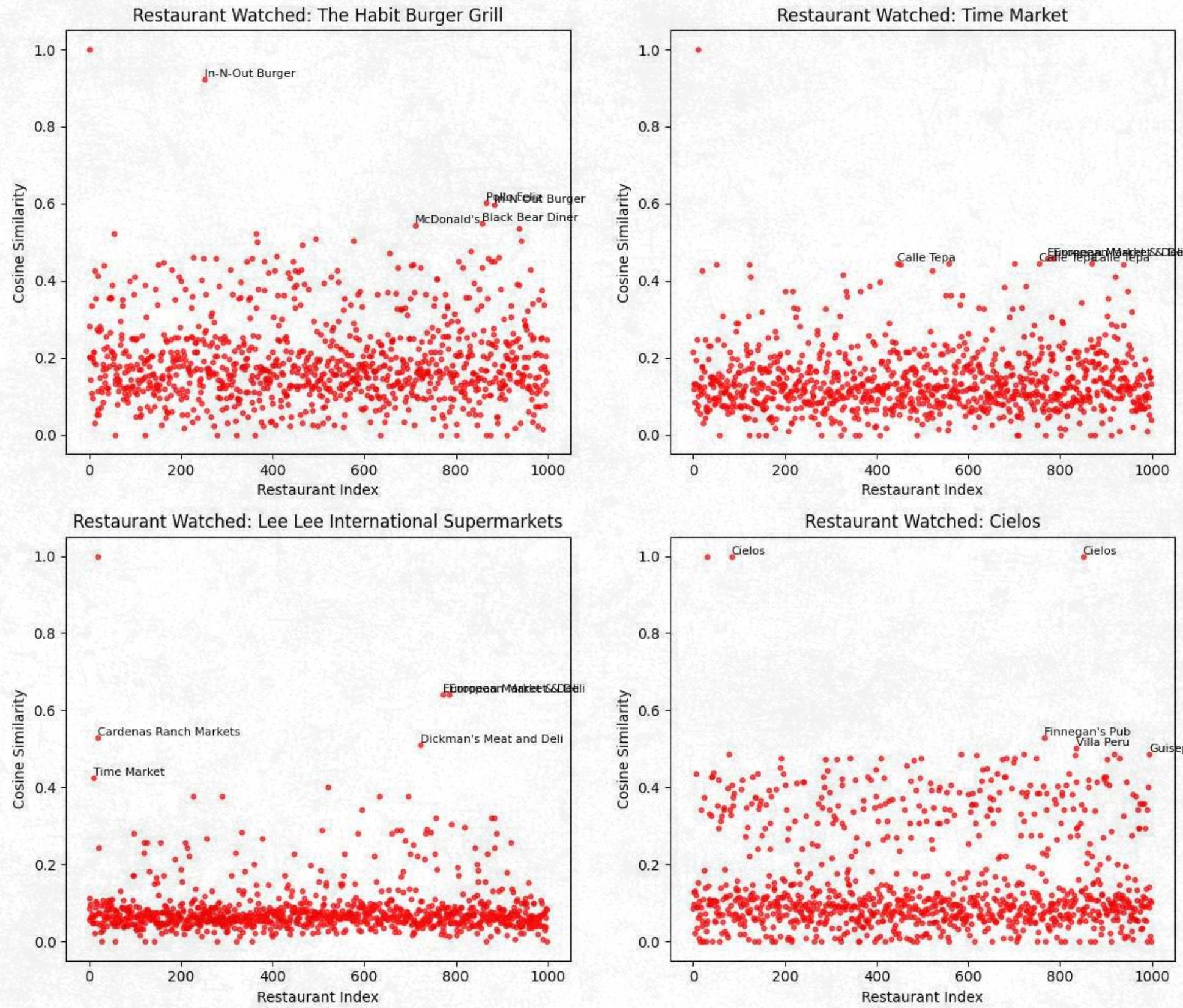
TF-IDF Transformation

```
→ TF-IDF matrix shape: (42600, 1770)
   TF-IDF computation time: 8.71 seconds
```

Converts each restaurant into a vector of term weights based on term importance

- Matrix Shape: Example: 42,000 rows × 1,707 columns (restaurants × unique terms after filtering)
- Purpose: Highlights key terms (e.g., categories, attributes) that define a restaurant relative to others.
- Computation Time: Example: 3.31 seconds to generate TF-IDF vectors
- Usage: These vectors enable restaurant comparisons by measuring similarity in term distributions

Cosine Similarity Plots



- Each of these plots shows how similar every restaurant in our dataset is to a specific “watched” restaurant, based on cosine similarity scores
 - The x-axis represents the index of each restaurant in the dataset.
 - The y-axis (Cosine Similarity) measures how closely each restaurant matches the “watched” restaurant’s features. A higher value means a closer match.
- Each red dot is a restaurant; the few labeled are the top-N closest matches.

Index 2

```
→ The watched restaurant is this one: Barista Del Barrio
The number 1 recommended restaurant is this one: Barista Del Barrio
The number 2 recommended restaurant is this one: Caffe Luce
The number 3 recommended restaurant is this one: Cartel Roasting
The number 4 recommended restaurant is this one: Coffee X Change
The number 5 recommended restaurant is this one: El Minuto Cafe
The categories of the watched restaurant is:
Restaurants, Food, Breakfast & Brunch, Coffee & Tea, Cafes
The categories of the number 1 recommended restaurant is:
Restaurants, Food, Breakfast & Brunch, Coffee & Tea, Cafes
The categories of the number 2 recommended restaurant is:
Food, Restaurants, Coffee & Tea, Cafes
The categories of the number 3 recommended restaurant is:
Restaurants, Food, Coffee & Tea, Cafes
The categories of the number 4 recommended restaurant is:
Restaurants, Sandwiches, Food, Coffee & Tea, Burgers
The categories of the number 5 recommended restaurant is:
Food, Restaurants, Mexican
```

Configured settings to generate restaurant recommendations.

- Target Restaurant: Index 2 is used as the reference for similarity comparisons
- Top_N: Set to 5 to retrieve the five most similar restaurants
- Print Recommendations: Enabled (`print_recommendation=True`) to display restaurant names

Index 60

```
→ The watched restaurant is this one: Sachiko Sushi
The number 1 recommended restaurant is this one: Sachiko Sushi
The number 2 recommended restaurant is this one: Shogun Japanese Restaurant & Sushi Bar
The number 3 recommended restaurant is this one: Kazoku
The number 4 recommended restaurant is this one: O Sushi Restaurant
The number 5 recommended restaurant is this one: Sushi Cortaro On River
The 'plot' of the watched restaurant is:
Japanese, Restaurants, Sushi Bars 3210 E Valencia Rd Tucson {'RestaurantsPriceRange2': '2', 'BusinessAcceptsCreditCards': 'True', 'RestaurantsGoodForGroups': 'True', 'HasTV': 'True', 'RestaurantsAttire': 'Casual', 'RestaurantsType': 'Sushi Bars', 'RestaurantsName': 'Sachiko Sushi', 'RestaurantsAddress': '3210 E Valencia Rd Tucson', 'RestaurantsPhone': '(520) 721-0000', 'RestaurantsCuisine': 'Japanese', 'RestaurantsPriceRange1': '1', 'RestaurantsPriceRange3': '3', 'RestaurantsHasParking': 'False', 'RestaurantsHasDishwasher': 'False', 'RestaurantsHasSmoking': 'False', 'RestaurantsHasOutdoorSeating': 'False', 'RestaurantsHasBikeParking': 'False', 'RestaurantsHasPetFriendly': 'False', 'RestaurantsHasWheelchairAccess': 'False', 'RestaurantsHasParking': 'False', 'RestaurantsHasDishwasher': 'False', 'RestaurantsHasSmoking': 'False', 'RestaurantsHasOutdoorSeating': 'False', 'RestaurantsHasBikeParking': 'False', 'RestaurantsHasPetFriendly': 'False', 'RestaurantsHasWheelchairAccess': 'False'}
The 'plot' of the number 1 recommended restaurant is:
Japanese, Restaurants, Sushi Bars 3210 E Valencia Rd Tucson {'RestaurantsPriceRange2': '2', 'BusinessAcceptsCreditCards': 'True', 'RestaurantsGoodForGroups': 'True', 'HasTV': 'True', 'RestaurantsAttire': 'Casual', 'RestaurantsType': 'Sushi Bars', 'RestaurantsName': 'Sachiko Sushi', 'RestaurantsAddress': '3210 E Valencia Rd Tucson', 'RestaurantsPhone': '(520) 721-0000', 'RestaurantsCuisine': 'Japanese', 'RestaurantsPriceRange1': '1', 'RestaurantsPriceRange3': '3', 'RestaurantsHasParking': 'False', 'RestaurantsHasDishwasher': 'False', 'RestaurantsHasSmoking': 'False', 'RestaurantsHasOutdoorSeating': 'False', 'RestaurantsHasBikeParking': 'False', 'RestaurantsHasPetFriendly': 'False', 'RestaurantsHasWheelchairAccess': 'False', 'RestaurantsHasParking': 'False', 'RestaurantsHasDishwasher': 'False', 'RestaurantsHasSmoking': 'False', 'RestaurantsHasOutdoorSeating': 'False', 'RestaurantsHasBikeParking': 'False', 'RestaurantsHasPetFriendly': 'False', 'RestaurantsHasWheelchairAccess': 'False'}
The 'plot' of the number 2 recommended restaurant is:
Japanese, Sushi Bars, Restaurants 5036 N Oracle Rd Tucson {'BikeParking': 'True', 'OutdoorSeating': 'False', 'RestaurantsGoodForGroups': 'True', 'HasTV': 'True', 'RestaurantsAttire': 'Casual', 'RestaurantsType': 'Sushi Bars', 'RestaurantsName': 'Shogun Japanese Restaurant & Sushi Bar', 'RestaurantsAddress': '5036 N Oracle Rd Tucson', 'RestaurantsPhone': '(520) 721-0000', 'RestaurantsCuisine': 'Japanese', 'RestaurantsPriceRange1': '1', 'RestaurantsPriceRange3': '3', 'RestaurantsHasParking': 'True', 'RestaurantsHasDishwasher': 'False', 'RestaurantsHasSmoking': 'False', 'RestaurantsHasOutdoorSeating': 'False', 'RestaurantsHasBikeParking': 'True', 'RestaurantsHasPetFriendly': 'False', 'RestaurantsHasWheelchairAccess': 'False', 'RestaurantsHasParking': 'True', 'RestaurantsHasDishwasher': 'False', 'RestaurantsHasSmoking': 'False', 'RestaurantsHasOutdoorSeating': 'False', 'RestaurantsHasBikeParking': 'True', 'RestaurantsHasPetFriendly': 'False', 'RestaurantsHasWheelchairAccess': 'False'}
The 'plot' of the number 3 recommended restaurant is:
Sushi Bars, Japanese, Restaurants 4210 E Speedway Blvd Tucson {'RestaurantsGoodForGroups': 'True', 'HasTV': 'True', 'RestaurantsAttire': 'Casual', 'RestaurantsType': 'Sushi Bars', 'RestaurantsName': 'Kazoku', 'RestaurantsAddress': '4210 E Speedway Blvd Tucson', 'RestaurantsPhone': '(520) 721-0000', 'RestaurantsCuisine': 'Japanese', 'RestaurantsPriceRange1': '1', 'RestaurantsPriceRange3': '3', 'RestaurantsHasParking': 'False', 'RestaurantsHasDishwasher': 'False', 'RestaurantsHasSmoking': 'False', 'RestaurantsHasOutdoorSeating': 'False', 'RestaurantsHasBikeParking': 'False', 'RestaurantsHasPetFriendly': 'False', 'RestaurantsHasWheelchairAccess': 'False', 'RestaurantsHasParking': 'False', 'RestaurantsHasDishwasher': 'False', 'RestaurantsHasSmoking': 'False', 'RestaurantsHasOutdoorSeating': 'False', 'RestaurantsHasBikeParking': 'False', 'RestaurantsHasPetFriendly': 'False', 'RestaurantsHasWheelchairAccess': 'False'}
The 'plot' of the number 4 recommended restaurant is:
Sushi Bars, Restaurants, Japanese 4689 E Speedway Blvd Tucson {'RestaurantsGoodForGroups': 'True', 'RestaurantsAttire': 'Casual', 'RestaurantsType': 'Sushi Bars', 'RestaurantsName': 'O Sushi Restaurant', 'RestaurantsAddress': '4689 E Speedway Blvd Tucson', 'RestaurantsPhone': '(520) 721-0000', 'RestaurantsCuisine': 'Japanese', 'RestaurantsPriceRange1': '1', 'RestaurantsPriceRange3': '3', 'RestaurantsHasParking': 'False', 'RestaurantsHasDishwasher': 'False', 'RestaurantsHasSmoking': 'False', 'RestaurantsHasOutdoorSeating': 'False', 'RestaurantsHasBikeParking': 'False', 'RestaurantsHasPetFriendly': 'False', 'RestaurantsHasWheelchairAccess': 'False', 'RestaurantsHasParking': 'False', 'RestaurantsHasDishwasher': 'False', 'RestaurantsHasSmoking': 'False', 'RestaurantsHasOutdoorSeating': 'False', 'RestaurantsHasBikeParking': 'False', 'RestaurantsHasPetFriendly': 'False', 'RestaurantsHasWheelchairAccess': 'False'}
The 'plot' of the number 5 recommended restaurant is:
Restaurants, Sushi Bars, Japanese 75 W River Rd, Ste 181 Tucson {'RestaurantsGoodForGroups': 'True', 'HasTV': 'True', 'RestaurantsAttire': 'Casual', 'RestaurantsType': 'Sushi Bars', 'RestaurantsName': 'Sushi Cortaro On River', 'RestaurantsAddress': '75 W River Rd, Ste 181 Tucson', 'RestaurantsPhone': '(520) 721-0000', 'RestaurantsCuisine': 'Japanese', 'RestaurantsPriceRange1': '1', 'RestaurantsPriceRange3': '3', 'RestaurantsHasParking': 'False', 'RestaurantsHasDishwasher': 'False', 'RestaurantsHasSmoking': 'False', 'RestaurantsHasOutdoorSeating': 'False', 'RestaurantsHasBikeParking': 'False', 'RestaurantsHasPetFriendly': 'False', 'RestaurantsHasWheelchairAccess': 'False', 'RestaurantsHasParking': 'False', 'RestaurantsHasDishwasher': 'False', 'RestaurantsHasSmoking': 'False', 'RestaurantsHasOutdoorSeating': 'False', 'RestaurantsHasBikeParking': 'False', 'RestaurantsHasPetFriendly': 'False', 'RestaurantsHasWheelchairAccess': 'False'}
```

Target Restaurant: Index 60, representing Sachiko Sushi.

- Recommendation Generation: Based on similarity metrics.
- Top_N: Retrieved the five most similar restaurants.
- Watched Restaurant: Sachiko Sushi used as the reference.
- Output: System displayed five recommended restaurants.

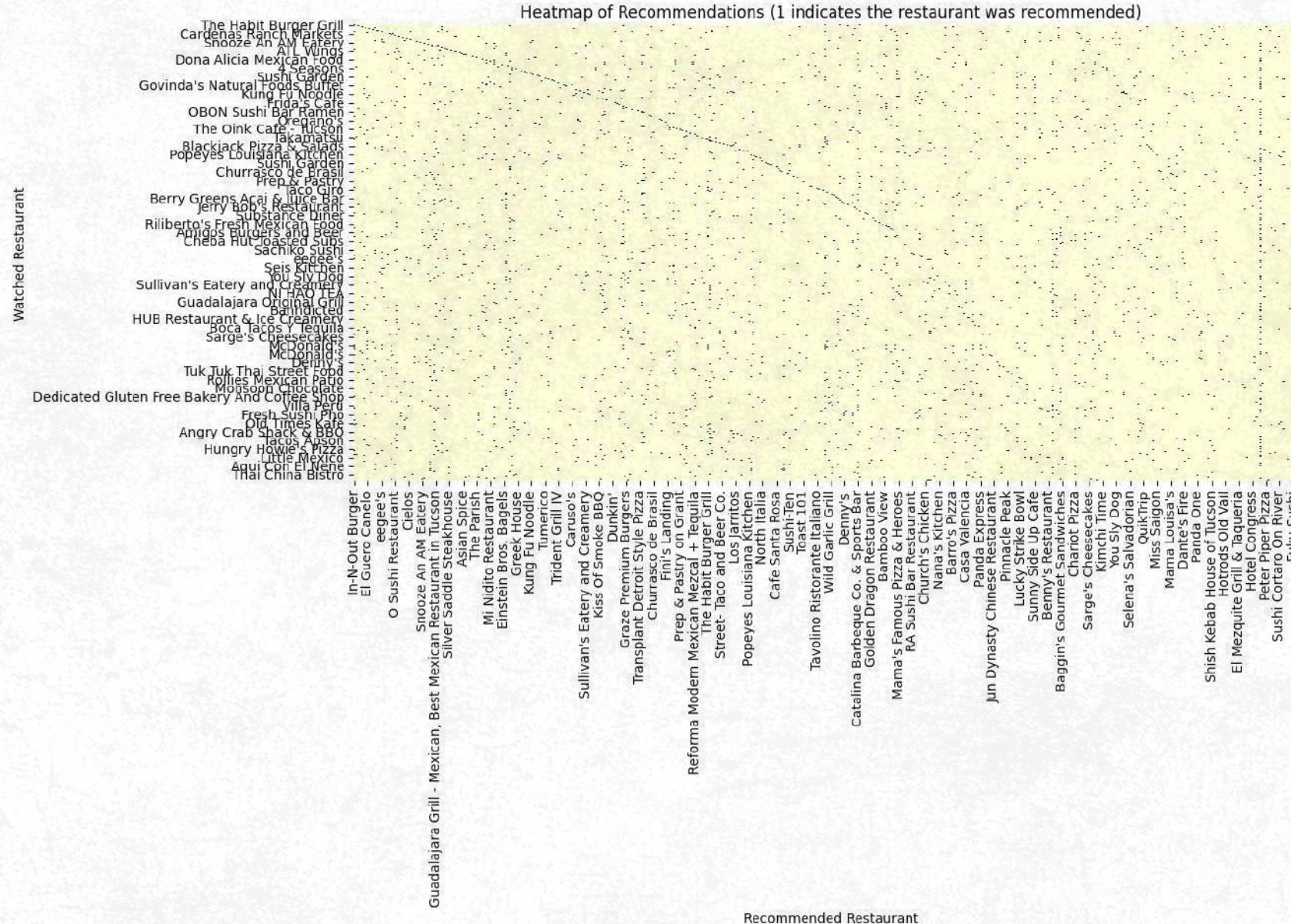
Recommendation Compilation

	Watched Restaurant	First Recommendation \
0	The Habit Burger Grill	In-N-Out Burger
1	Pat's Drive In	Kenney D's
2	Barista Del Barrio	Barista Del Barrio
3	Aqui Con el Nene	Aqui Con el Nene
4	Domino's Pizza	Maria Bonita Mexican Kitchen
5	Union Public House	Union Public House
6	Fresh Sushi Pho	Fresh Sushi Pho
7	El Guero Canelo	El Guero Canelo
8	Med Cuisine	Med Cuisine
9	Dickey's Barbecue Pit	Dickey's Barbecue Pit
	Second Recommendation	Third Recommendation \
0	Pollo Feliz	Black Bear Diner
1	Blake's Lotaburger	Wings Over Broadway
2	Caffe Luce	Cartel Roasting
3	Waffle House	Los Betos
4	Baggin's Gourmet	Hungry Howie's Pizza
5	Proof Artisanal Pizza & Pasta	Reforma Modern Mexican Mezcal + Tequila
6	Kazoku	Sushi-Cho
7	BK Tacos	Tacos Apson
8	Sinbad's	Caravan Grill
9	Monsoons Tap & Grill	Cheddar's Scratch Kitchen
	Fourth Recommendation	Fifth Recommendation
0	McDonald's	C.I. Chu's Mongolian Barbeque
1	Truland Burgers & Greens	You Sly Dog
2	Coffee X Change	El Minuto Cafe
3	Jerry Bob's Restaurant	Baggin's Gourmet Sandwiches
4	Domino's Pizza	Jimmy John's
5	None	None
6	Takamatsu	Shogun Japanese Restaurant & Sushi Bar
7	Los Jarritos	Delicias Mexican Grill
8	Shish Kebab House of Tucson	CharroVida
9	Chick-fil-A	KFC

Recommendation Compilation: Created a structured DataFrame

- **Similarity Metric:** Used cosine similarity to find similar restaurants.
- **Top_N:** Compiled top 5 recommended restaurants for each restaurant.
- **Watched Restaurant:** Listed alongside its recommendations.
- **Output:** Clear DataFrame showing restaurant-recommendation pairs.

Heatmap with Cosine Similarity



Visualization: Heat map representing cosine similarity scores.

- Purpose: Identifies clusters of similar restaurants.
- Interpretation: Higher similarity scores indicate closely related dining options.
- Usage: Helps in analyzing restaurant groupings based on feature similarities.
- Insight: Enables quick identification of patterns and relationships among restaurants.

THANK YOU