Aspect-Based Restaurant Recommendation System

Presented by-
Avraham Reissberg
Shivani Tayade
Sravya Bhaskara

Agenda

- INTRODUCTION
- KNOW YOUR DATA
- PROJECT OVERVIEW
- ASPECT SENTIMENT ANALYSIS
- RECOMMENDATION SYSTEM
- CHALLENGES

Introduction

Goal: Enhance user experience by providing personalized restaurant recommendations

Problem Statement: Traditional rating systems fail to capture aspect-specific preferences (food, service, ambiance)

📂 Dataset: Yelp Academic Dataset

📊 Models Used: Aspect-based sentiment analysis & recommendation systems

📝 Input: User text-based reviews

📢 Output: Restaurant recommendations

Example

- User's Preference: High value placed on food quality
- Recommendation Output: Businesses with high sentiment scores in the food aspect

Know your data

Review Dataset

Contains user-written reviews with ratings, and text data for sentiment analysis

User

Includes user profiles helping in personalized recommendations.

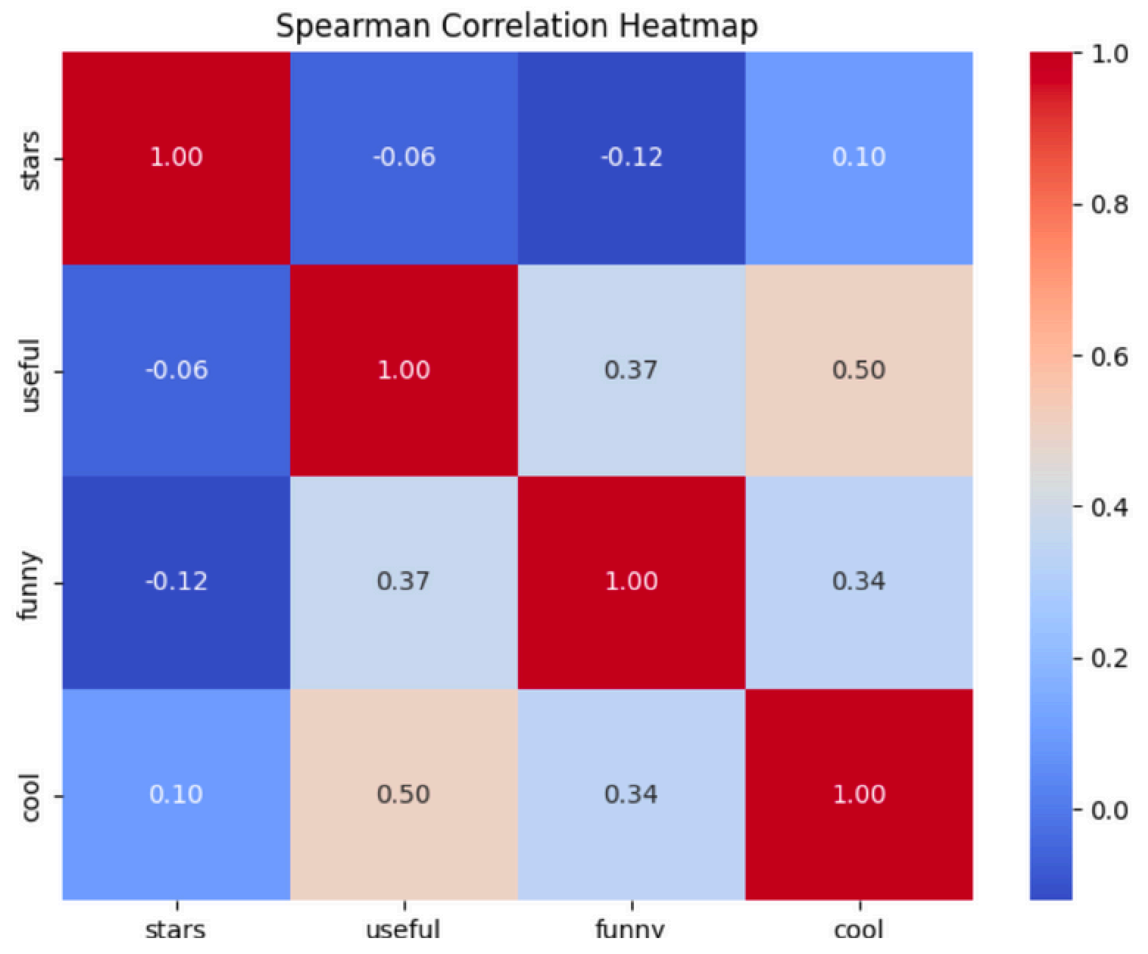
Business

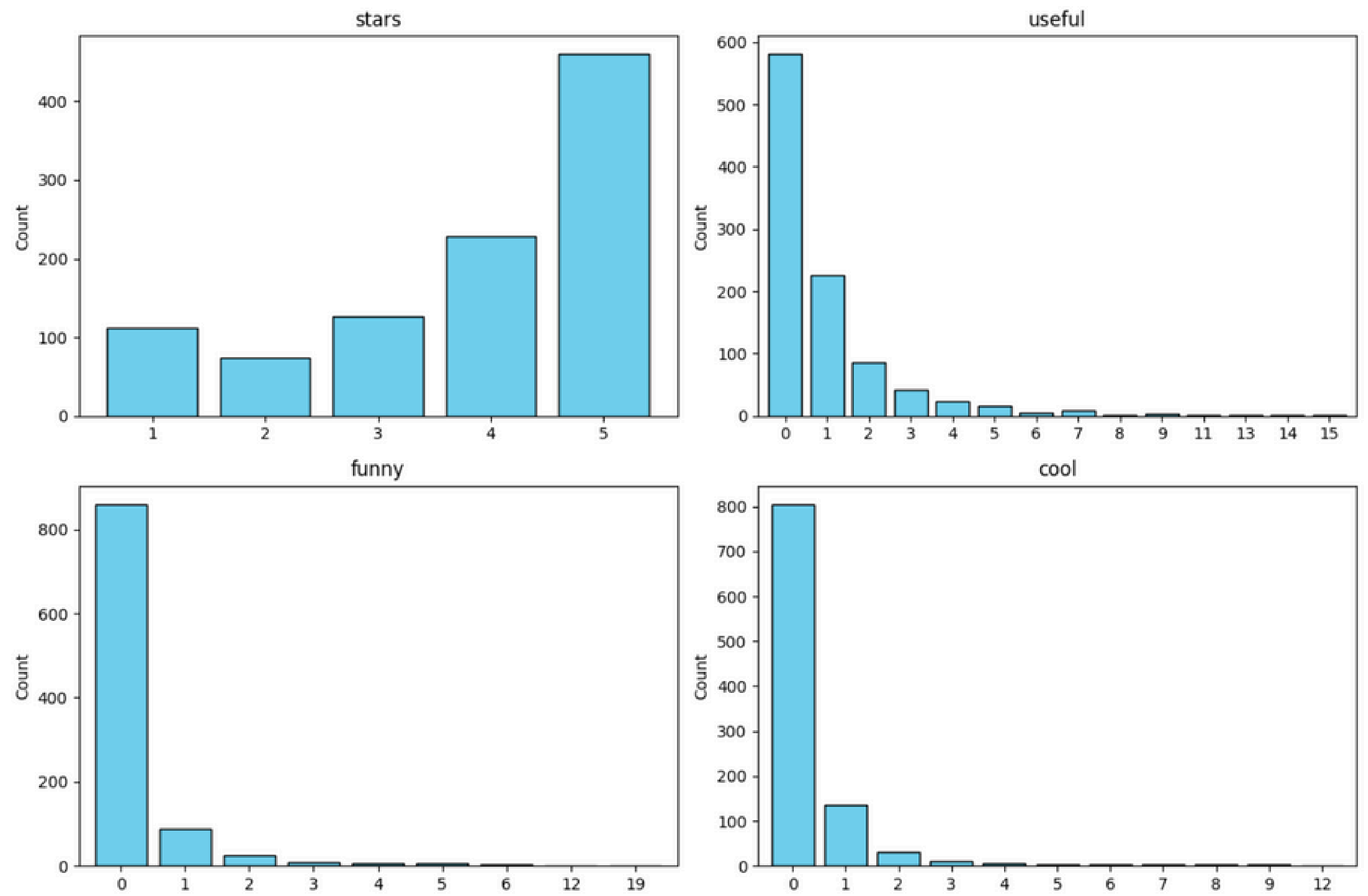
Provides details about businesses, including location, cusine, and aggregate ratings.

Key Highlights from EDA-

Review Data

- Strong correlations between review metrics (useful, funny, cool) and user engagement indicate influential reviewers drive interactions.
- Highly skewed distributions suggest a small number of users dominate engagement.



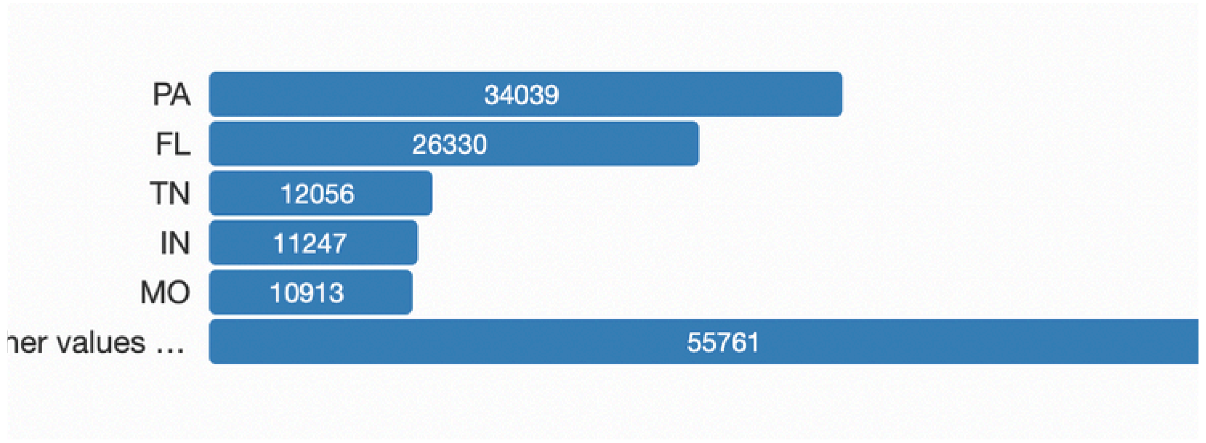


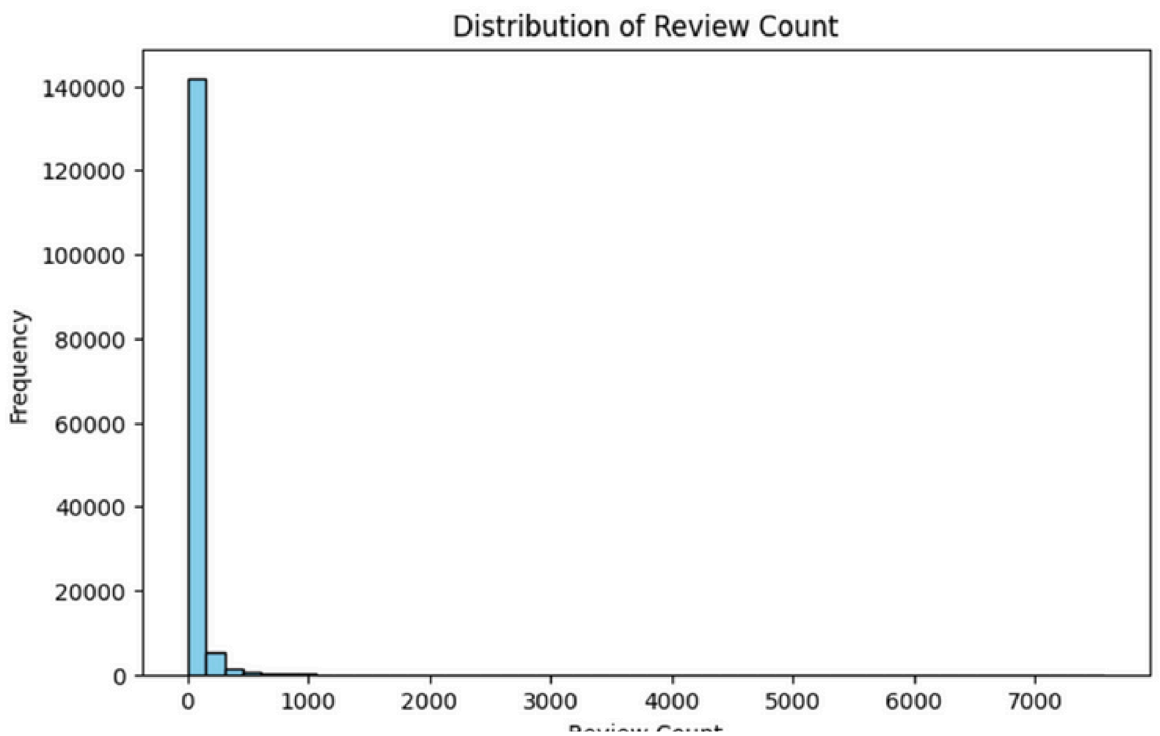
Key Highlights from EDA-

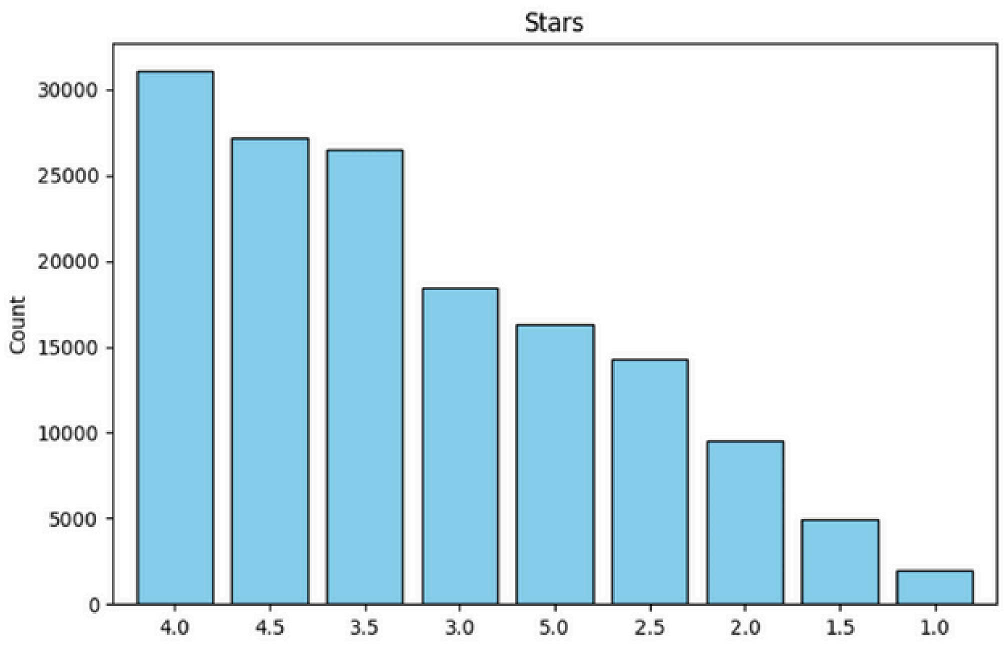
Business Data

- Businesses concentrated in major cities, ideal for targeted marketing.
- Wide variance in star ratings and review counts, reflecting diverse customer experiences.
- Incomplete attribute and operational hour data highlight areas businesses can improve.

Distribution of Business across States



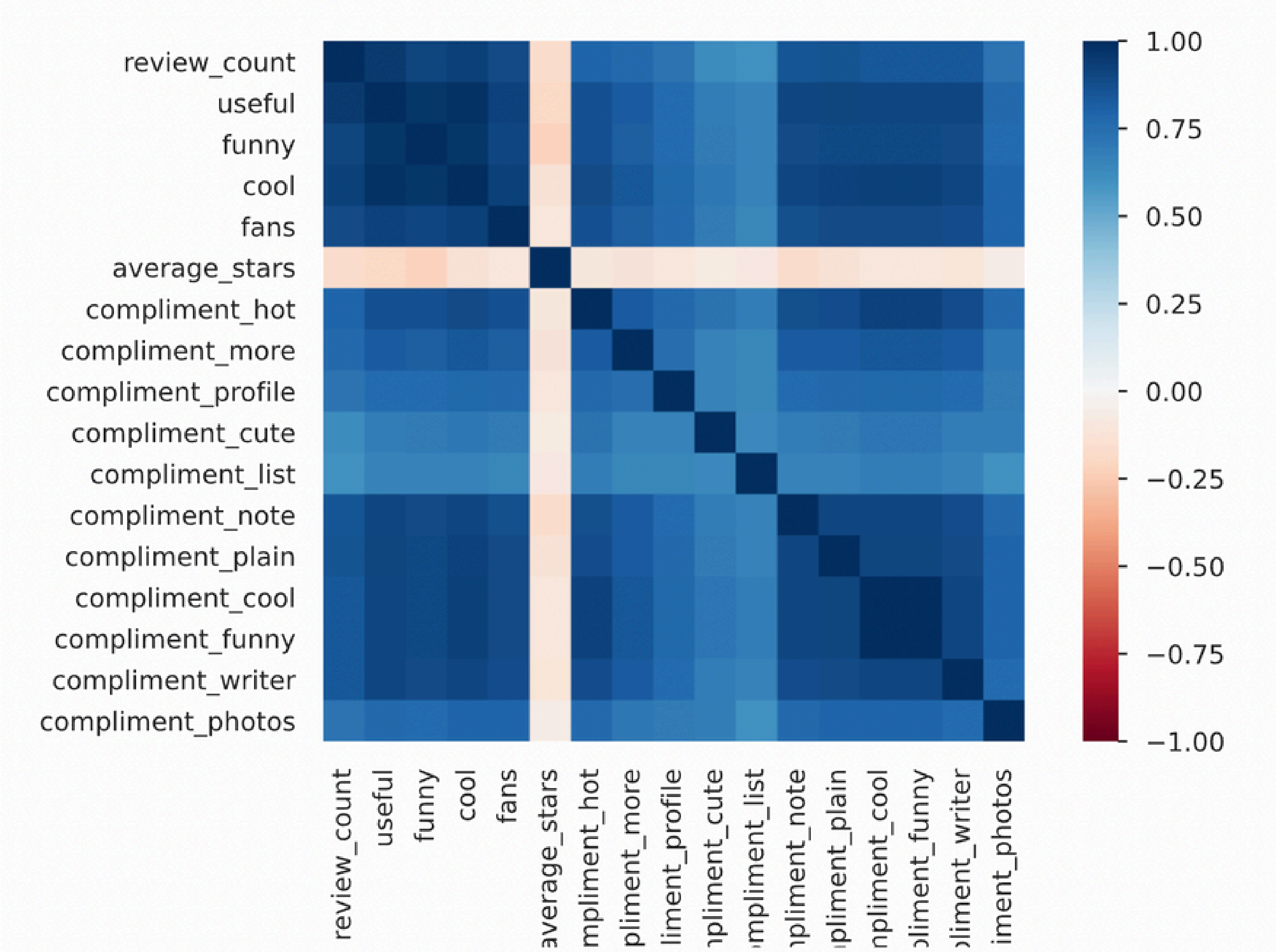


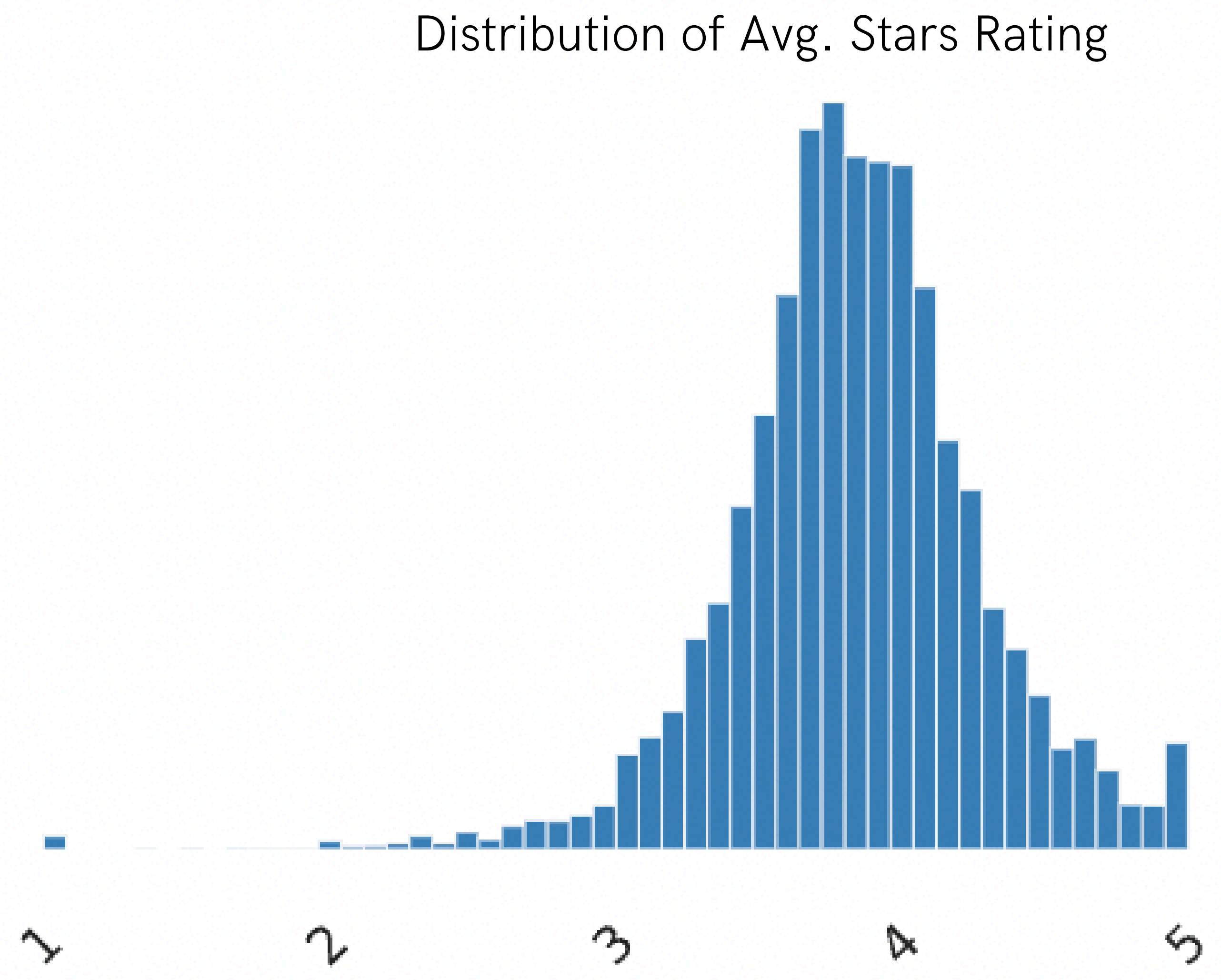


Key Highlights from EDA-

User Data

- Strong correlations between review metrics (useful, funny, cool) and user engagement indicate influential reviewers drive interactions.
- Mostly positive star ratings- skewed distributions





Project Overview

01

Text Extraction

Extract key aspects (e.g., food, service, ambiance) and their context from user reviews.
Identify which aspects contribute most to user satisfaction

02

Sentiment Analysis

Assign sentiment scores (1 to 5) to each aspect based on review content.
Compare aspect-level sentiment with overall ratings

03

Personalized Recommendation

Match users with places based on their aspect preferences.
Recommend restaurants that align with users' most valued aspects.

Aspect Based Sentiment Analysis

Review → Our favorite is the lamb curry and korma
Aspect → Food
Context → "Our favorite"

1 WORD COUNT

Combine and Preprocess reviews by removing noise, stopwords, and irrelevant characters

Identify and extract relevant aspects (food, service)

2 GET CONTEXT

We need to match the aspects with our reviews and get contect for each of the aspect

3 SENTIMENTS

Assign sentiment scores (scale: 1 to 5) for each aspect.

4 EVALUATION

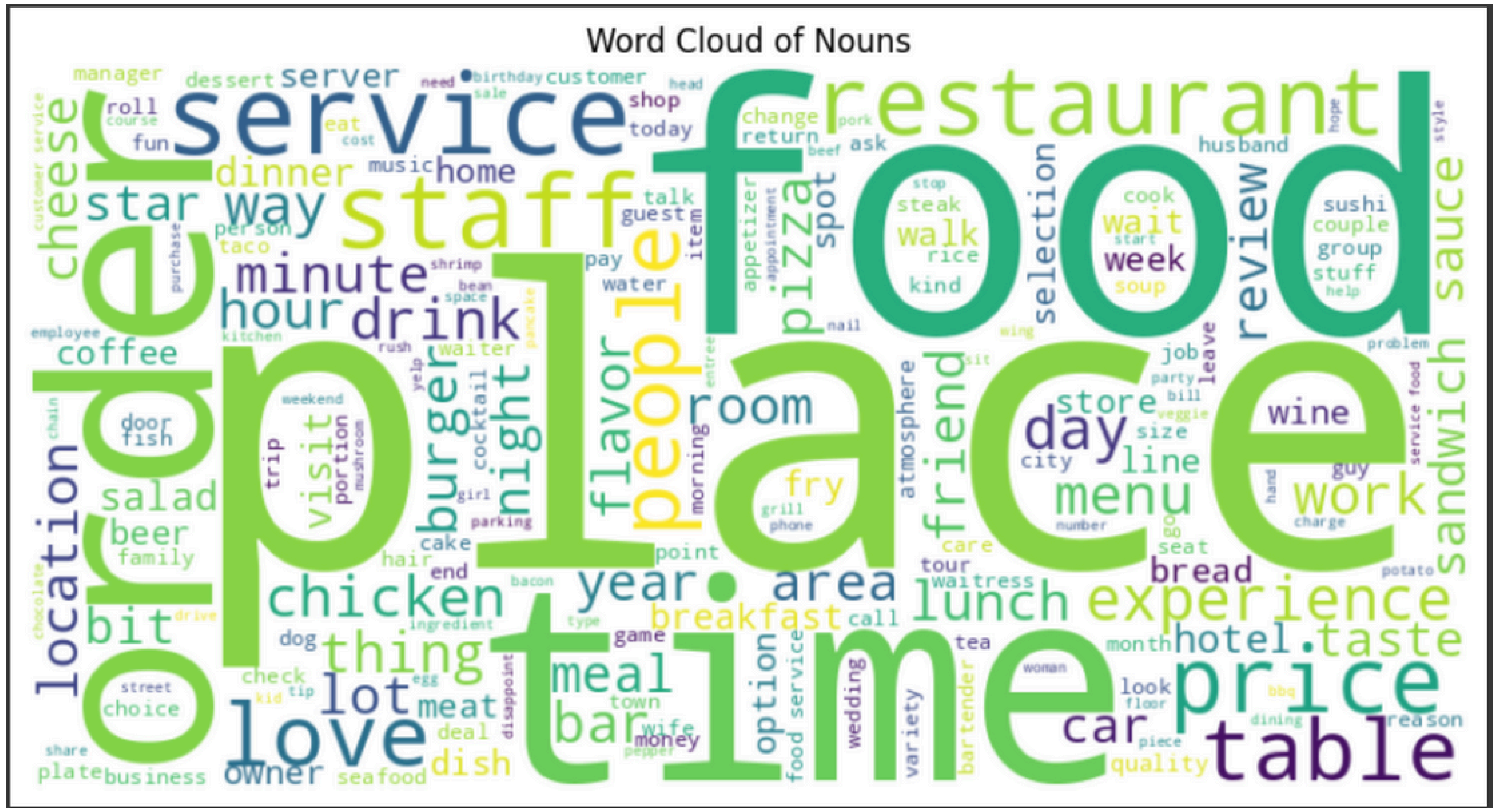
Evaluate aspect sentiment scores

Select the best-performing sentiment analysis approach based on accuracy.

Word Count

Bi, Tri Gram

Word Cloud




```
Top 10 Most Important Words by TF-IDF Score:
good        40.777769
food        36.418233
place       35.288902
great       34.092139
time        26.499395
service     25.935970
come        22.277575
order       22.039387
love        21.482448
like        20.836829
dtype: float64
```

TF-IDF

```
Top 10 Most Frequent Bigrams:
food service: 39
customer service: 28
service food: 26
place food: 23
ice cream: 21
love place: 21
time time: 18
order food: 16
food drink: 14
place order: 13

Top 10 Most Frequent Trigrams:
place food service: 4
food love place: 4
hair hair hair: 4
parking lot staff: 3
ice cream cone: 3
food service food: 3
roast beef roast: 3
order fish taco: 3
prix fixe menu: 3
service beer selection: 3
```

Final Aspects

Food

Service

Place

Time

Get Contexts

Approach 1- Check each aspect for keyword matches

```
Text: Amazingly amazing wings and homemade bleu cheese. Had the ribeye: tender, perfectly prepared, delicious. Nice selection of craft beers. Would DEFINITELY recommend checking out this hidden gem.

place
['No context found']
food
['No context found']
time
['No context found']
service
['No context found']
{'place': 2.0, 'food': 2.0, 'time': 2.0, 'service': 2.0}
```

Approach 2 - Compute similarity between review sentences and aspect categories

```
Text: Amazingly amazing wings and homemade bleu cheese. Had the ribeye: tender, perfectly prepared, delicious. Nice selection of craft beers. Would DEFINITELY recommend checking out this hidden gem.

Food
['Amazingly amazing wings and homemade bleu cheese.', 'Had the ribeye: tender, perfectly prepared, delicious.']
Service
['No context found']
Place
['Nice selection of craft beers.']
Time
['Would DEFINITELY recommend checking out this hidden gem.']
{'Food': 5.0, 'Service': 2.0, 'Place': 5.0, 'Time': 5.0}
```

Sentiments and Evaluation

Sentiment on scale of 1 to 5

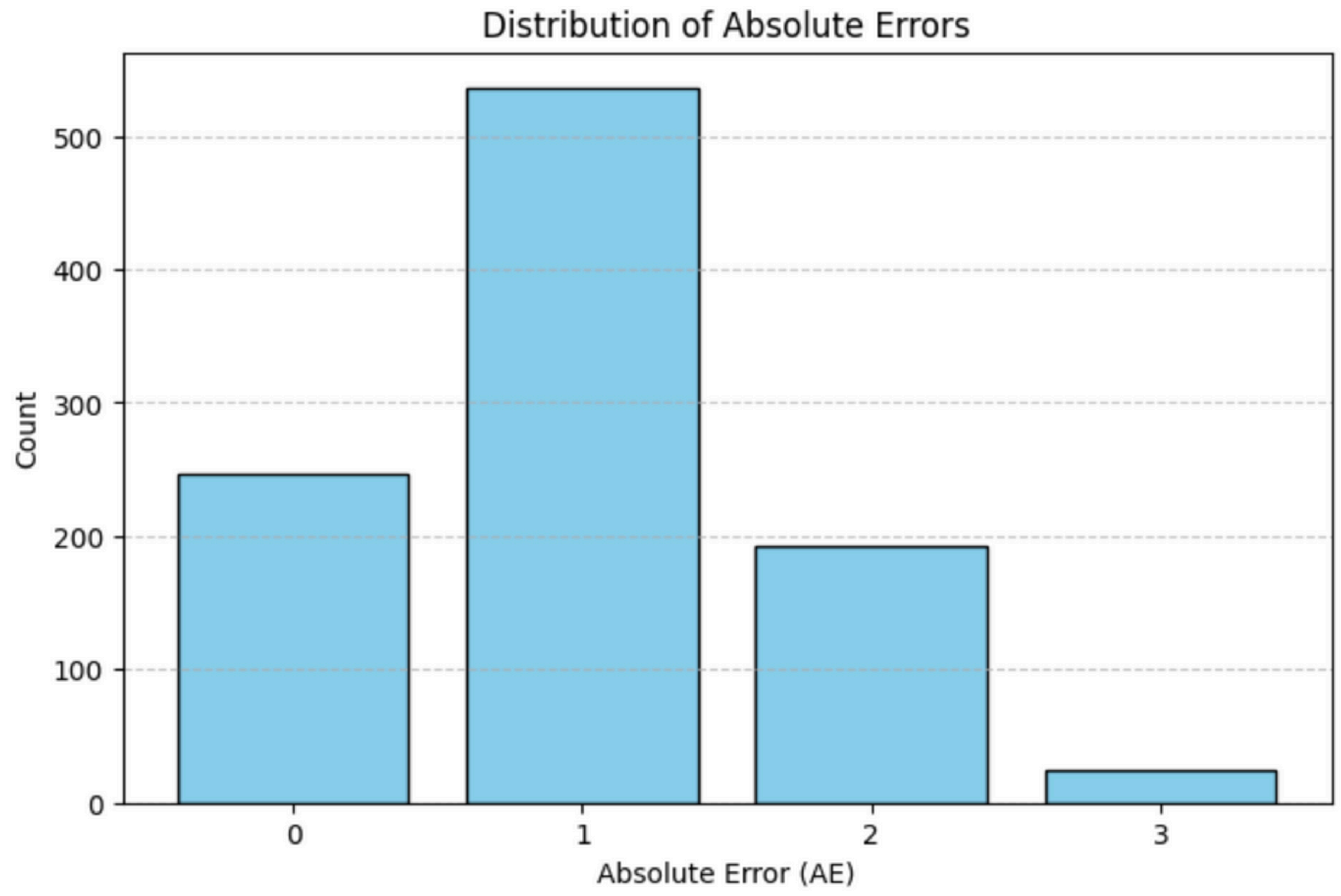
Model 1- Text Blob

Model 2 - Hugging Face sentiment analysis model

What if there is no sentiment allocated to the aspect - Using 2 as default

Overall rating (Combing the sentiment using weighted averaged)

Compare Overall Sentiment (Predicted) with Stars (Actual)




```
Weights: {'Place': 0.3, 'Food': 0.3, 'Time': 0.2, 'Service': 0.2}
 MSE: 1.52, RMSE: 1.23
```

Recommendation Systems

Content Based Filtering

- Recommends restaurants similar based on user preferences (Given as prior input)
- Uses features like (food quality, service) and city to match preferences
- Works well for new users (Cold Start Problem)

Collaborative Filtering

- Recommends restaurants based on similar users' likes and dislikes.
- Finds users with similar tastes and suggests places liked by users with similar interests.
- Limitation: Struggles with new users due to cold start problem.

Content Based Filtering

City: Philadelphia
Aspect 1: Place
Aspect 2: Food

```
Top 20 restaurants in Philadelphia based on Place and Food, sorted by Place, Food, and Final Sentiment:
```

	business_name	Place	Food	final_sentiment
261	Franzone's Pizzeria & Restaurant	5.0	5.0	5.0
485	El Camino Real	5.0	5.0	5.0
932	Mac's Tavern	5.0	5.0	5.0
451	Gennaro's Tomato PIE	5.0	5.0	4.0
482	Octopus Falafel Truck	5.0	5.0	4.0
568	Tria Cafe Wash West	5.0	5.0	4.0

- For new users, we will take input of City and the aspects they care the most (preferences)
- We find the best restaurants for the aspects based on our aspect based sentiment analysis and share it
- The results are sorted based on two preferences given by user

Collaborative Based Filtering

- Input - Aspect Based Ratings per User per Business
- For existing users, we calculate the similarity with other users based on which aspects they care most about
- Once we find similar users, restaurants they rate highly are shared as recommendations
- Tecniques used: cosine_similarity

- Combine User, Business and Ratings data to get User * Business Matrix

```
Top 5 recommended restaurants for user John:

Smiths Restaurant and Bar
Taqueria La Hacienda
Rittenhouse Grill
Jasmine Rice - Rittenhouse
Bridesburg Pizza
```

How to Identify users with similar aspect-based preferences?

- Users are similar if they prioritize the same aspect (place)
- Determine weightage based on how aspects influence overall ratings
- Example: A user praises place, finds food okay, but still gives 5 stars → Place matters most

Challenges

- Context Extraction - Mapping the aspect to relevant context in a review (Used Semantic Mapping)
- Large Dataset & Processing Time – Handling and processing massive review data efficiently. (Used GPUs)
- Weighted Averaging for Sentiment – Finding the best approach to aggregate aspect sentiments into an overall rating.
- User Preference Variability – Different users may weigh aspects differently, making personalized recommendations complex.

Thank You