



# **INSIGHTFUL REVIEWS**

## **REVEALING HOSPITALITY REVIEW INSIGHTS**

---

# CONTENT INDEX

1. INTRODUCTION

2. METHODOLOGY

3. DATA PROCESSING

4. MODELING

5. ERROR ANALYSIS

6. CONCLUSION

7. RECOMMENDATION

8. REFERENCES



# INTRODUCTION

IN A WORLD WHERE CUSTOMER FEEDBACK CAN MAKE OR BREAK A HOSPITALITY BUSINESS,  
HOW DO WE SIFT THROUGH THOUSANDS OF REVIEWS TO FIND WHAT TRULY MATTERS?

# VALUES

## VISION

Empowering individuals in hospitality with informed decisions for crafting exceptional guest experiences and elevating service standards

## MISSION

To transform customer feedback into actionable insights, driving service excellence in the hospitality sector



# PROBLEM STATEMENT

In the hospitality sector, businesses often struggle to navigate through a sea of unstructured customer feedback from AIRBNB platforms. This feedback, while rich in insights, remains largely untapped due to its volume and complexity.

The inability to effectively analyze this feedback can lead to missed opportunities for service improvement, a lack of timely response to customer needs, and, ultimately, a potential decline in guest satisfaction and business profitability

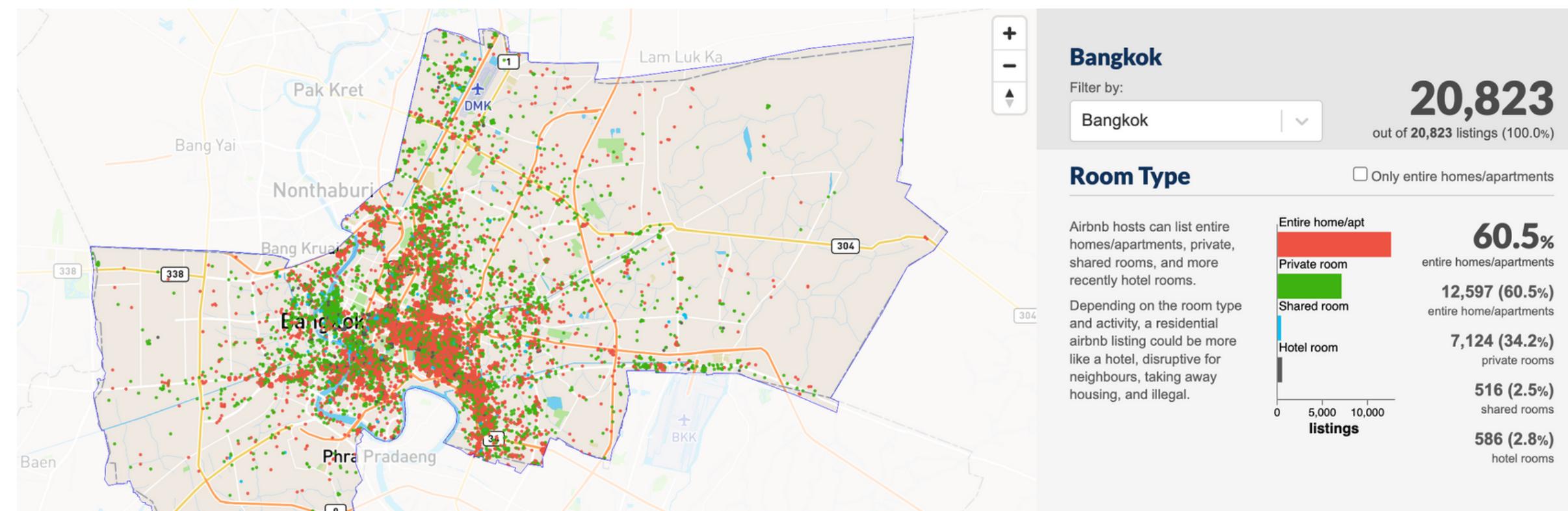
# TARGET AUDIENCE

- **Airbnb Corporate Team** : Seeking actionable insights from customer feedback to enhance guest satisfaction and make informed decisions.
- **Prospective Guests** : Looking for personalized recommendations tailored to their preferences and needs when choosing accommodation.

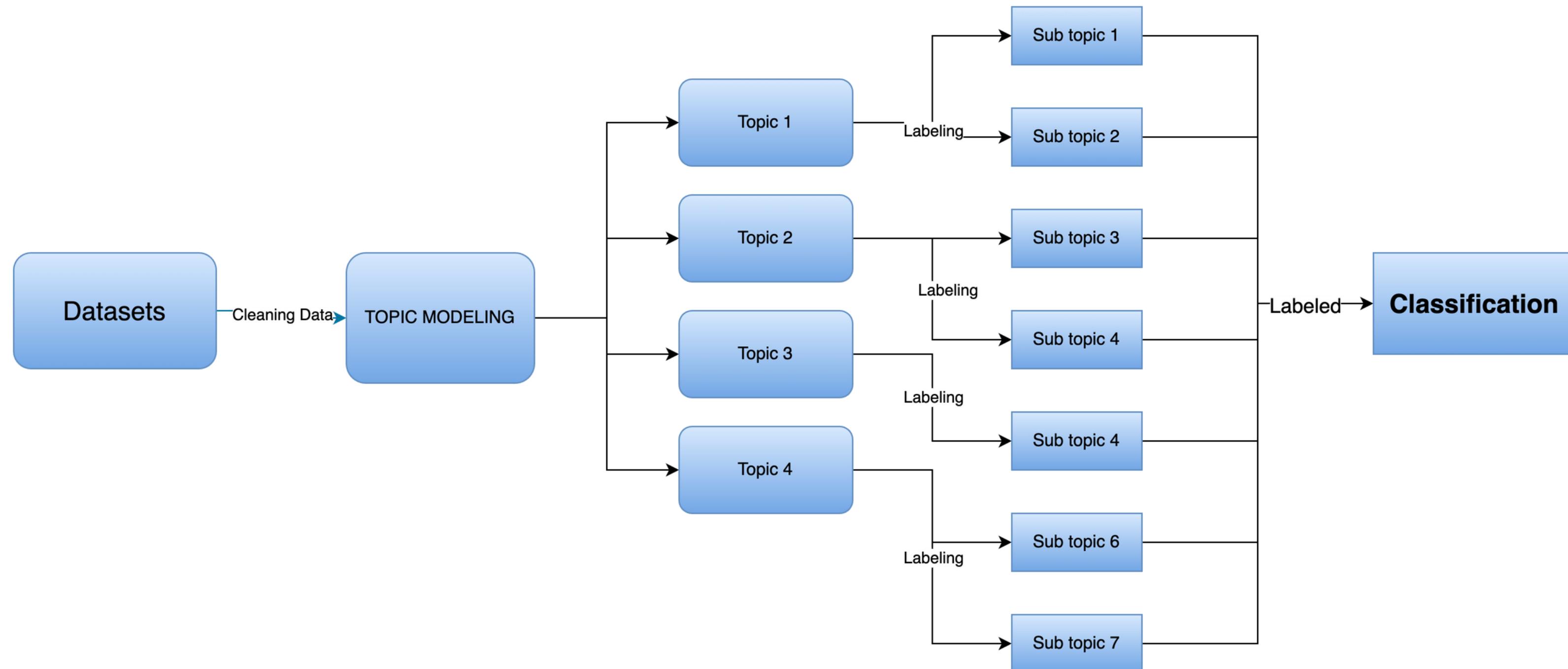
# METHODOLOGY

## Data Collection

- Insight AIRBNB provides a rich dataset of guest feedback, encompassing diverse aspects of guest experiences in various types of accommodations globally (**BANGKOK**)
- Data from 07/04/2012-18/09/2023 , 300,000 reviews



# METHODOLOGY



# DATA PROCESSING

1

## NORMALIZE TEXT

- Replace/remove special characters (/,<,>, etc.)
- Keep only English letters and numbers

2

## REMOVE STOPWORDS AND COMMON WORDS

- Remove English stopwords. (a, is, a, etc)
- Filter out frequent but less informative words (e.g., 'good', 'great').

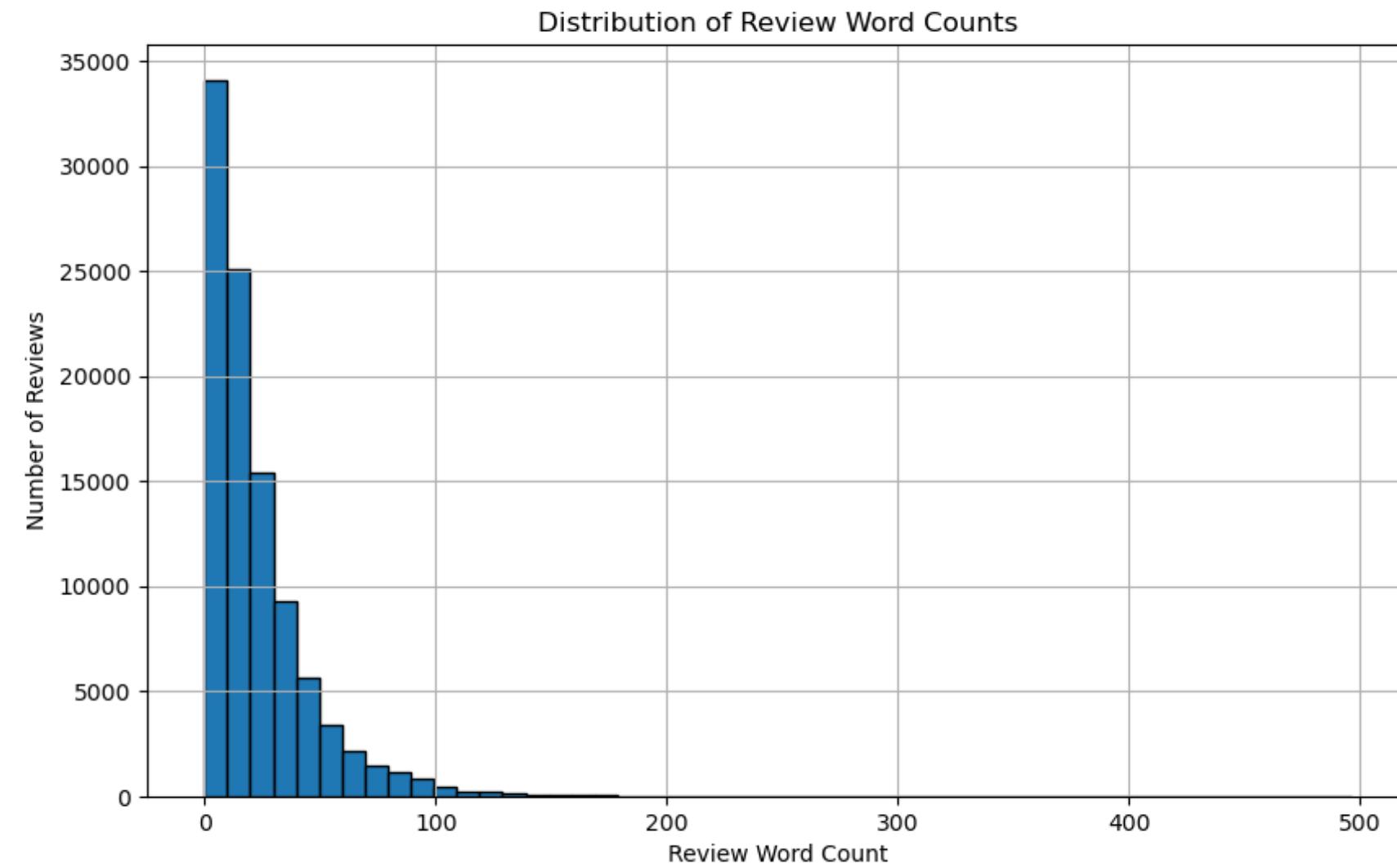
3

## DATA READY FOR ANALYSIS

- Proceed to topic modeling and further analysis.

# EXPLORATORY DATA ANALYSIS

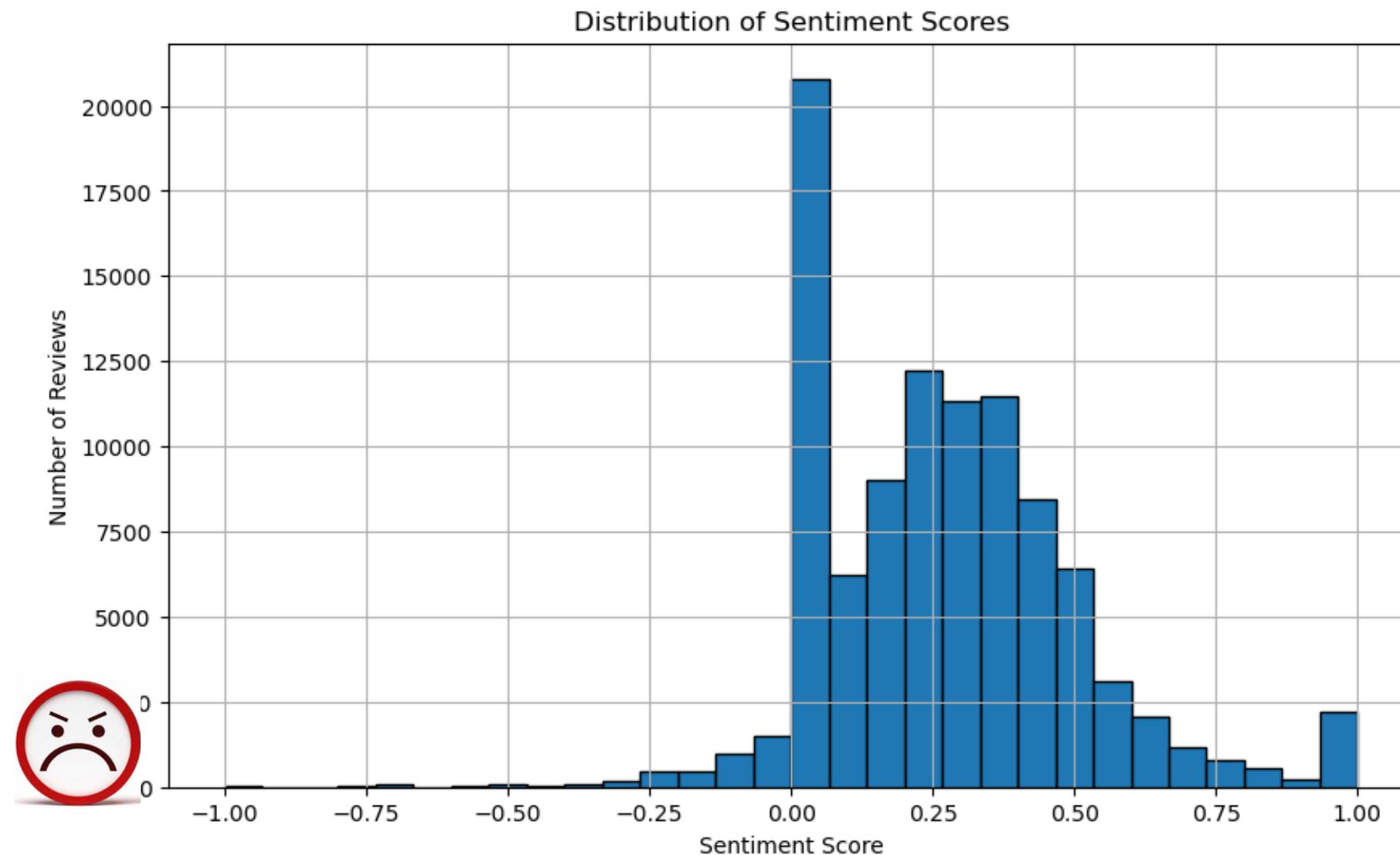
## Distribution Patterns in Textual Data



**Common Review Length**  
**30-50 Words**

# EXPLORATORY DATA ANALYSIS

## Sentiment Of Reviews



★★★★★ · June 2023

Very friendly hosts and will answer any questions you may have. There is a 7-Eleven in the neighborhood, very convenient, there is a mall next door.



★★★★★ · September 2023

If you want to earn this price, do it at this price point. Tidy up or renovate the room.



# EXPLORATORY DATA ANALYSIS

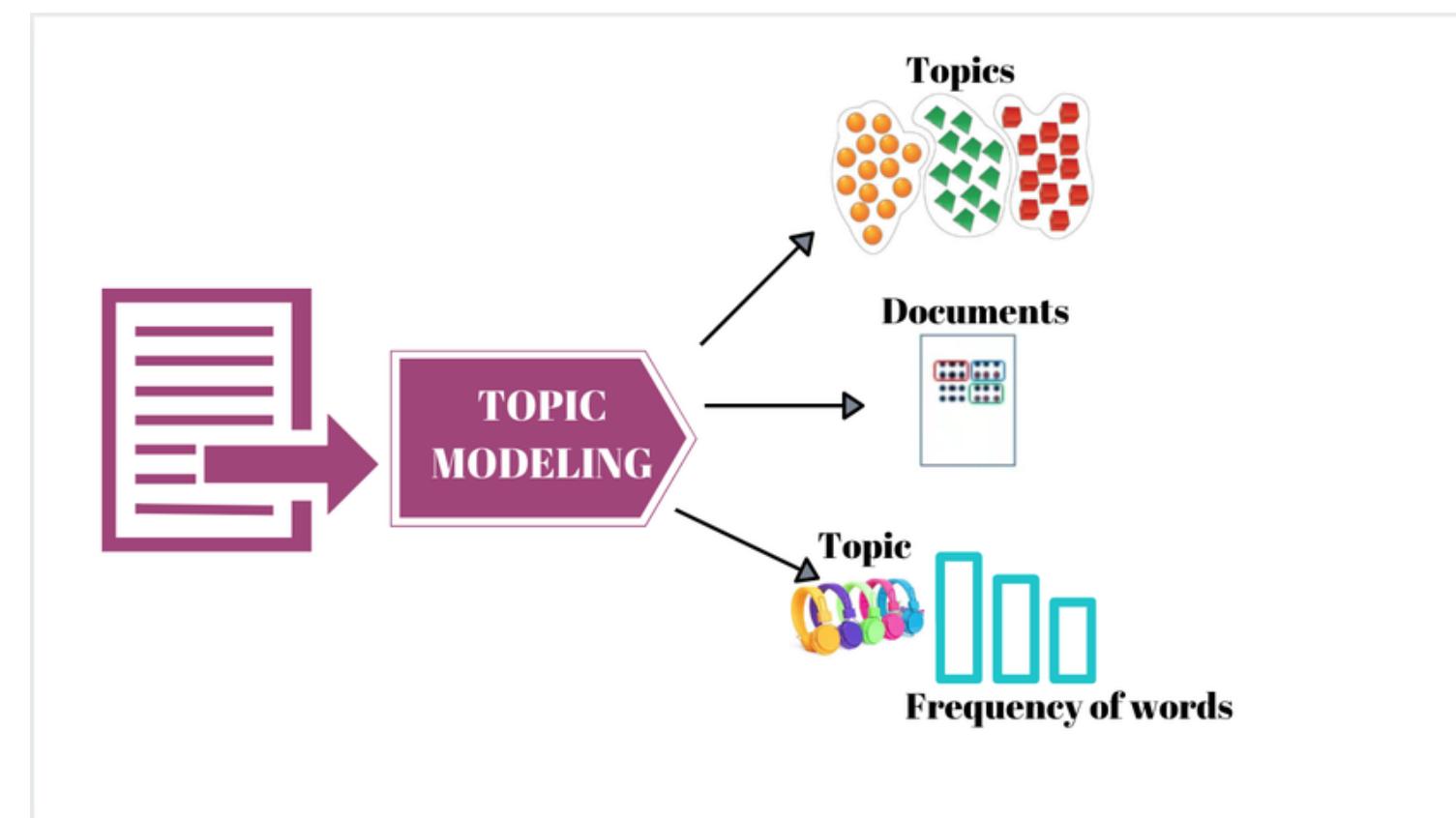
# WORDCLOUD



WORD	FREQUENCY
Place	45874
Location	23121
Clean	21923
Host	21122
Apartment	18113

# TOPIC MODEING

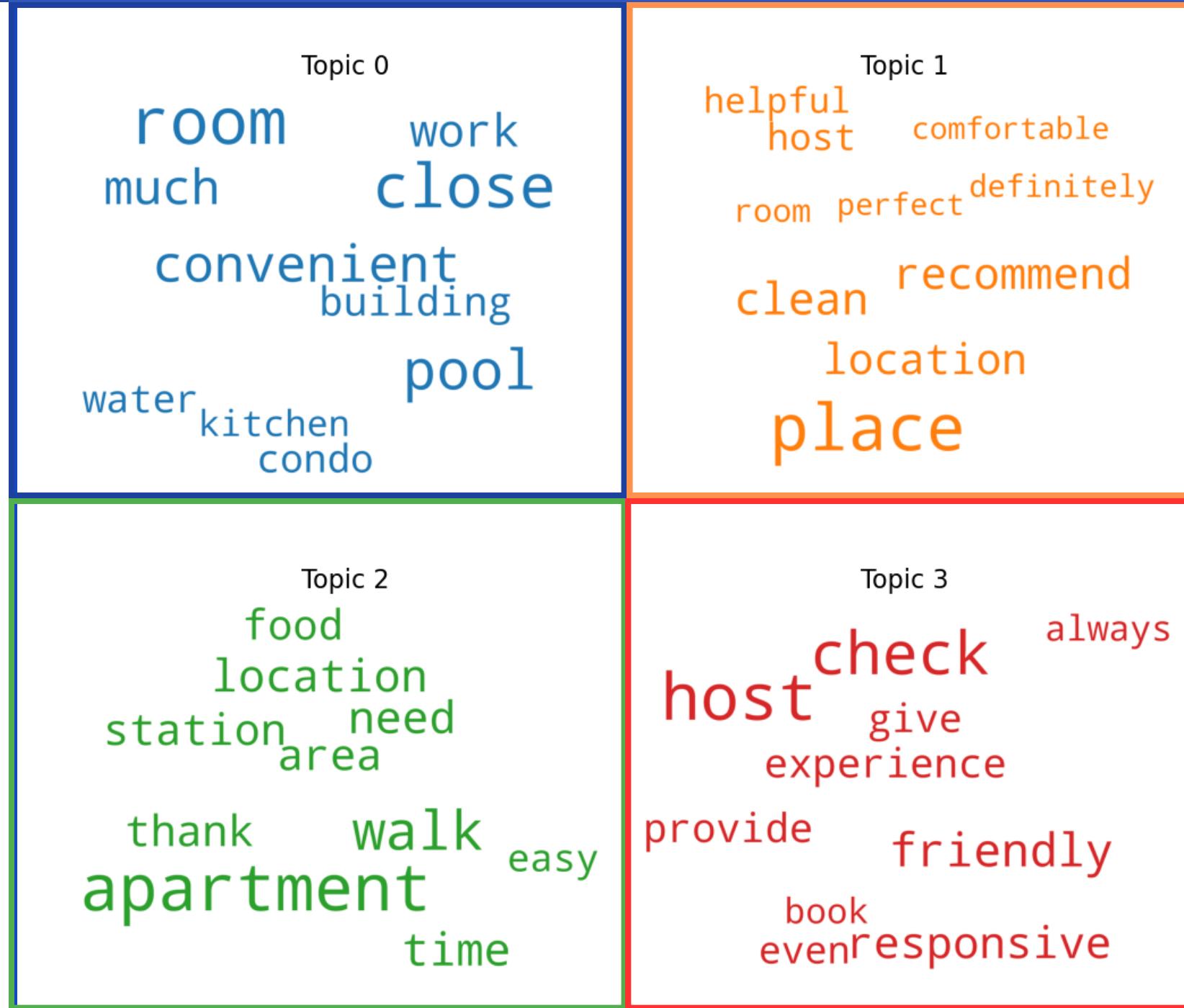
In our project, we're using a special technique called Latent Dirichlet Allocation, or **LDA**. It goes through all the customer reviews and sorts them into different groups, each group representing a particular theme or topic that guests often talk about.



# TOPIC MODELING

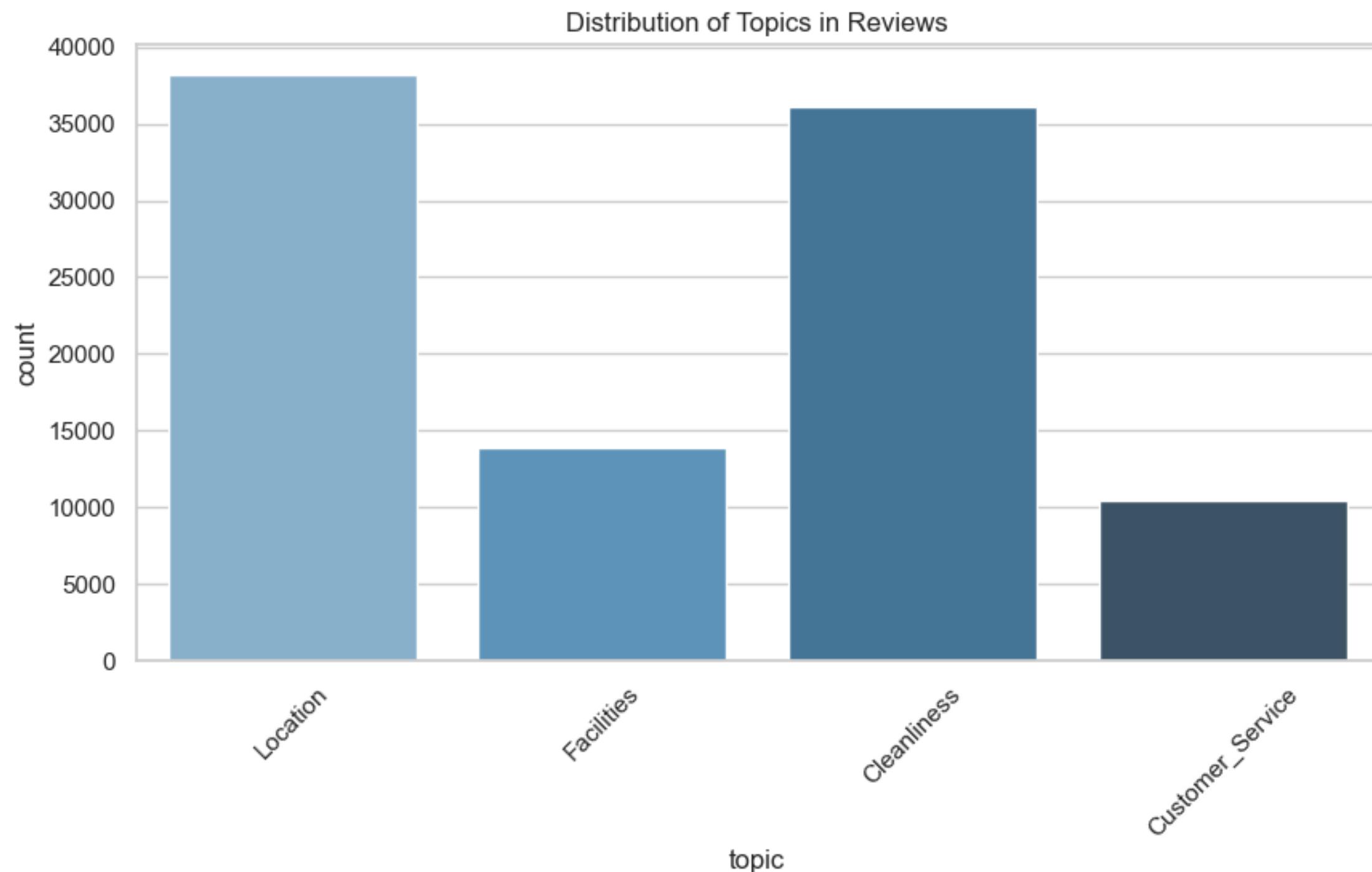


# TOPIC MODELING



- **Topic 0: Facilities**
- **Topic 1: Cleanliness**
- **Topic 2: Location**
- **Topic 3: Customer Service**

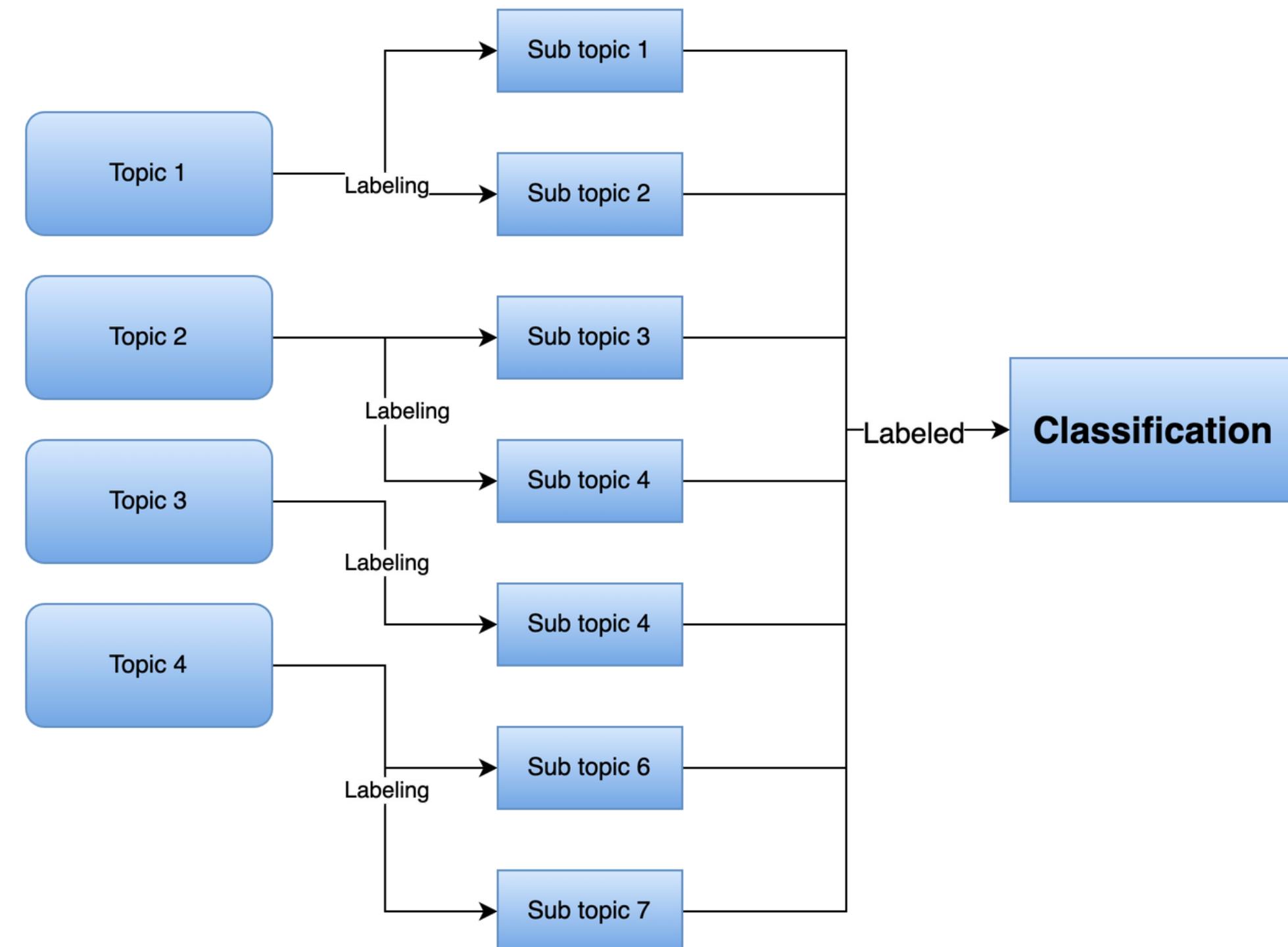
# TOPIC MODELING



# LABELING

## Few Shot

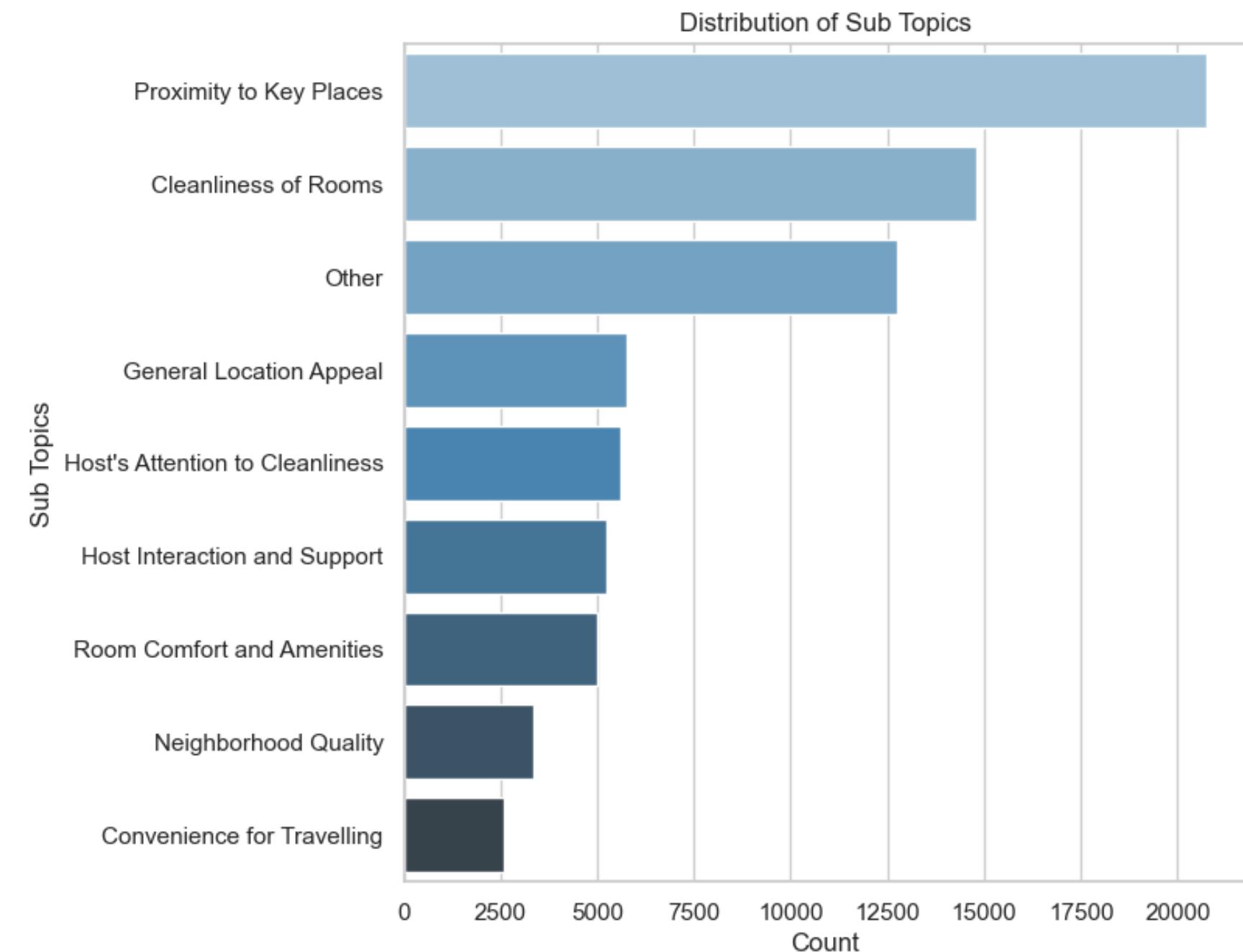
Few-shot learning (FSL) aims to generate a classifier using limited labeled examples.



# LABELING

## Creation of Sub-Topics

1. Proximity to Key Places
2. Cleanliness of Rooms
3. General Location Appeal
4. Host's Attention to Cleanliness
5. Host Interaction and Support
6. Room Comfort and Amenities
7. Neighborhood Quality
8. Convenience for Travelling
9. Other



# MODEL PROCESSING

## PURPOSE OF CLASSIFICATION MODEL

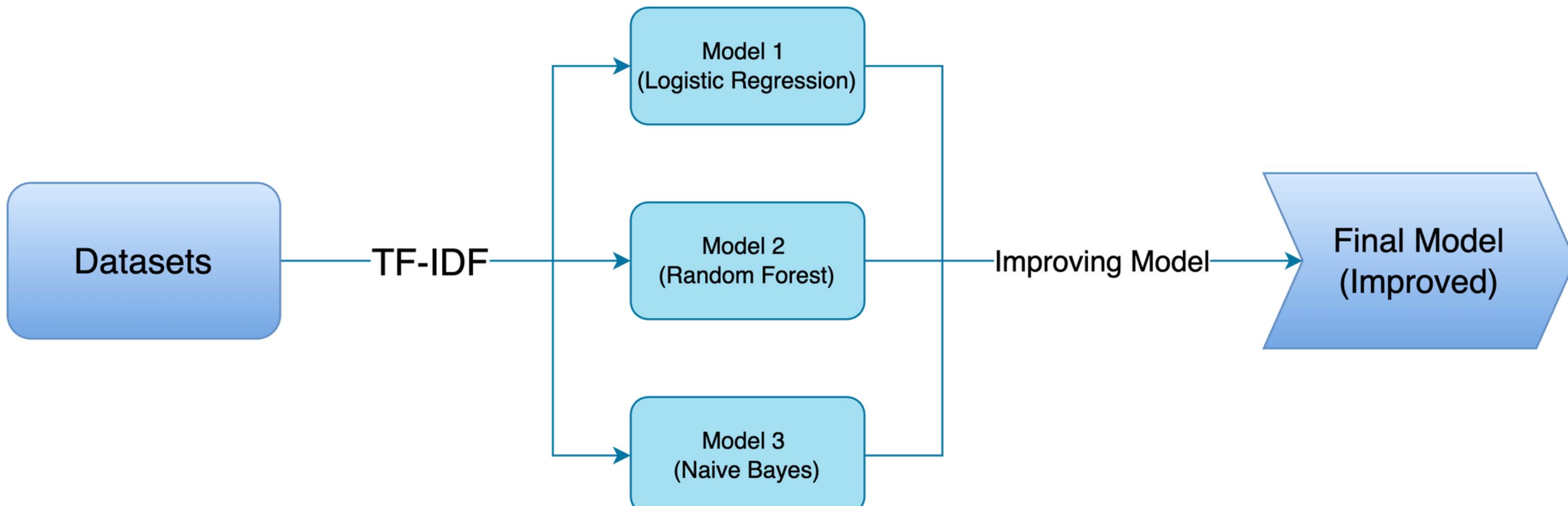
- Using a special tool, a classification model, to quickly sort customer reviews into specific groups
- This approach helps us analyze feedback more efficiently and accurately, pinpointing exactly what guests are talking about in their reviews.

## MODEL OVERVIEW

- We picked a model (like Logistic Regression, Random Forest, Naive Bayes ) for train model
- Choose The best one of model

# MODEL PROCESSING

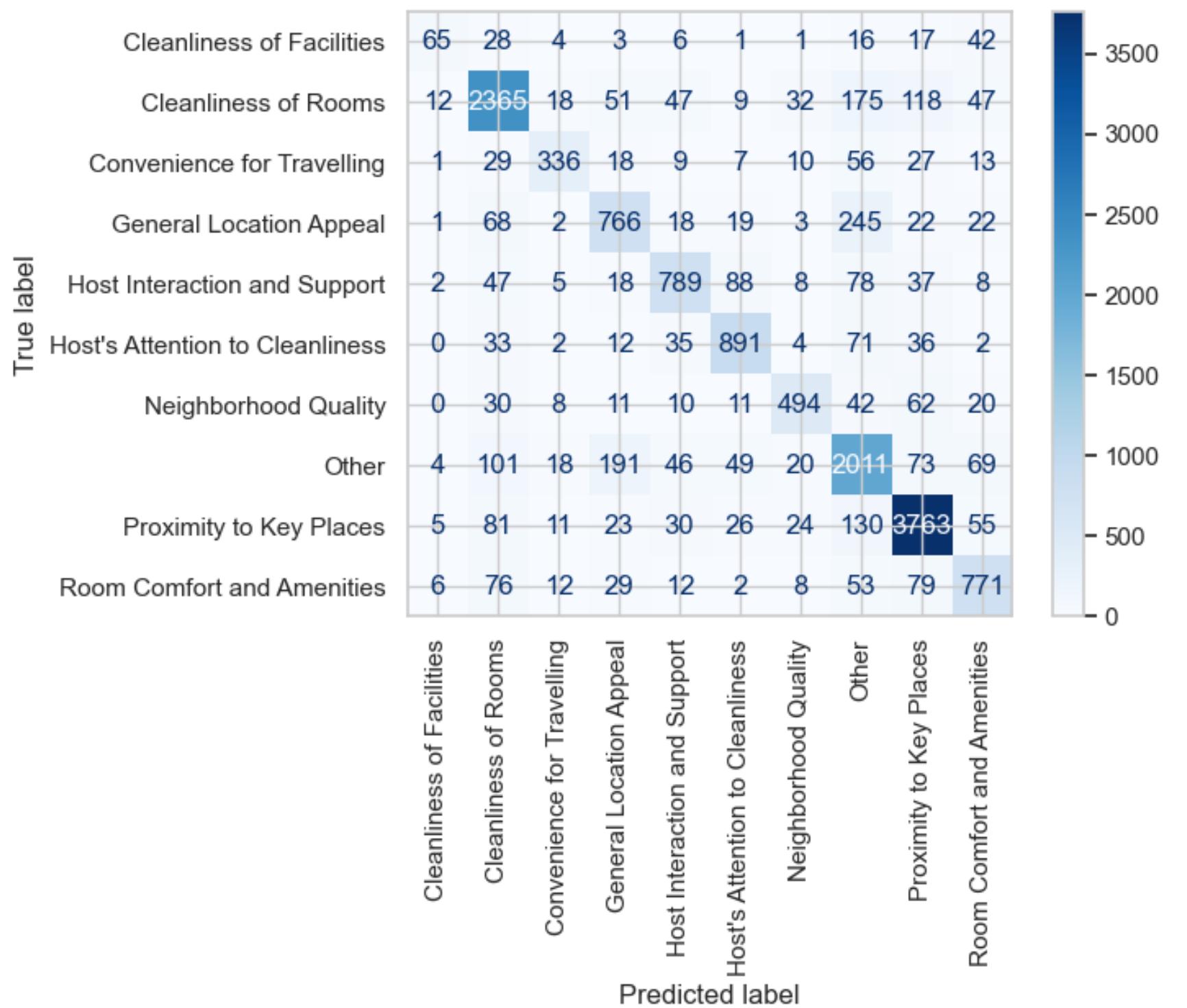
## MODEL WORKFLOW



# MODEL COMPARISON

MODEL	F1 (Test Data)	RECALL (Test Data)	Precision (Test Data)
Logistic Regression	0.70	0.68	0.77
Random Forest	0.63	0.61	0.70
Naive Bayes	0.50	0.55	0.51

# MODEL IMPROVEMENT



## Logistic Regression (Improved)

Metrics

Score  
(Test Data)

F1

0.78

Recall

0.77

Precision

0.80

# ERROR ANALYSIS

## ACTUAL CLASS

Proximity to Key Places

Proximity to Key Places

Proximity to Key Places

## PREDICT CLASS

Cleanliness of Rooms

Cleanliness of Rooms

Room Comfort and Amenities

## REVIEWS

clean place, exactly photo! pleasant bit , close local market street food also BTS

first time bangkok place location clean peaceful close everything

apartment location, close shopping center, complete supporting facilities, endless swimming pool gym comfortable

# CONCLUSION AND RECOMMENDATION

- **Topic Modeling Success:** The project successfully implemented topic modeling (LDA) to categorize user reviews into distinct topics, enhancing the review classification system.
- **Insights from Data:** The analysis revealed various sub-topics within the reviews, such as Facilities, Cleanliness, Location, and Host Interaction, offering deep insights into customer preferences and experiences.
- **Model Performance:** Logistic Regression was found to be the most effective model with improved scores in F1 is **78%**

# CONCLUSION AND RECOMMENDATION

- **Implement Review Filtering on Airbnb**
- **Continuous Model Improvement**
- **Using Sub-labels for In-depth Customer Surveys**
- **Expansion to Other Languages:** Given Airbnb's global presence, recommend expanding the model to include multiple languages, thereby making the review filtering feature universally applicable and useful.

Filter Reviews

Room Cleanliness

Location close to  
key place

Neighborhood Quality

Room Comfort and  
Amenities

Host Interaction and  
Support

Host's Attention to  
Cleanliness

General Location  
Appeal

Other

# LIMITATION

- **Language Limitation:** The project currently only processes reviews in English
- **Computer Resource**



Thank You