

# FOOD HABIT ANALYSIS OF NORTH AMERICANS

**Name: Md Rashel Uzzaman**

**Student ID: 23099317**

**M.Sc. in Artificial Intelligence**

**Friedrich-Alexander-Universität Erlangen-Nürnberg**

# Research Question:

- Is there a pattern in food habits in different regions of North America?

- Objective:

To identify regional food preferences and trends across the USA and Canada, aiding businesses in market entry strategies.

# Data Overview

- Dataset Source: Kaggle
- Dataset Details: 1,500 restaurant entries
- Attributes: Name Cuisine(s), Rating, City, State, and more
- Challenges Identified: Multiple cuisine names with inconsistencies (e.g., "burger" vs. "burgers"), Right-skewed restaurant ratings.

# Data Pipeline

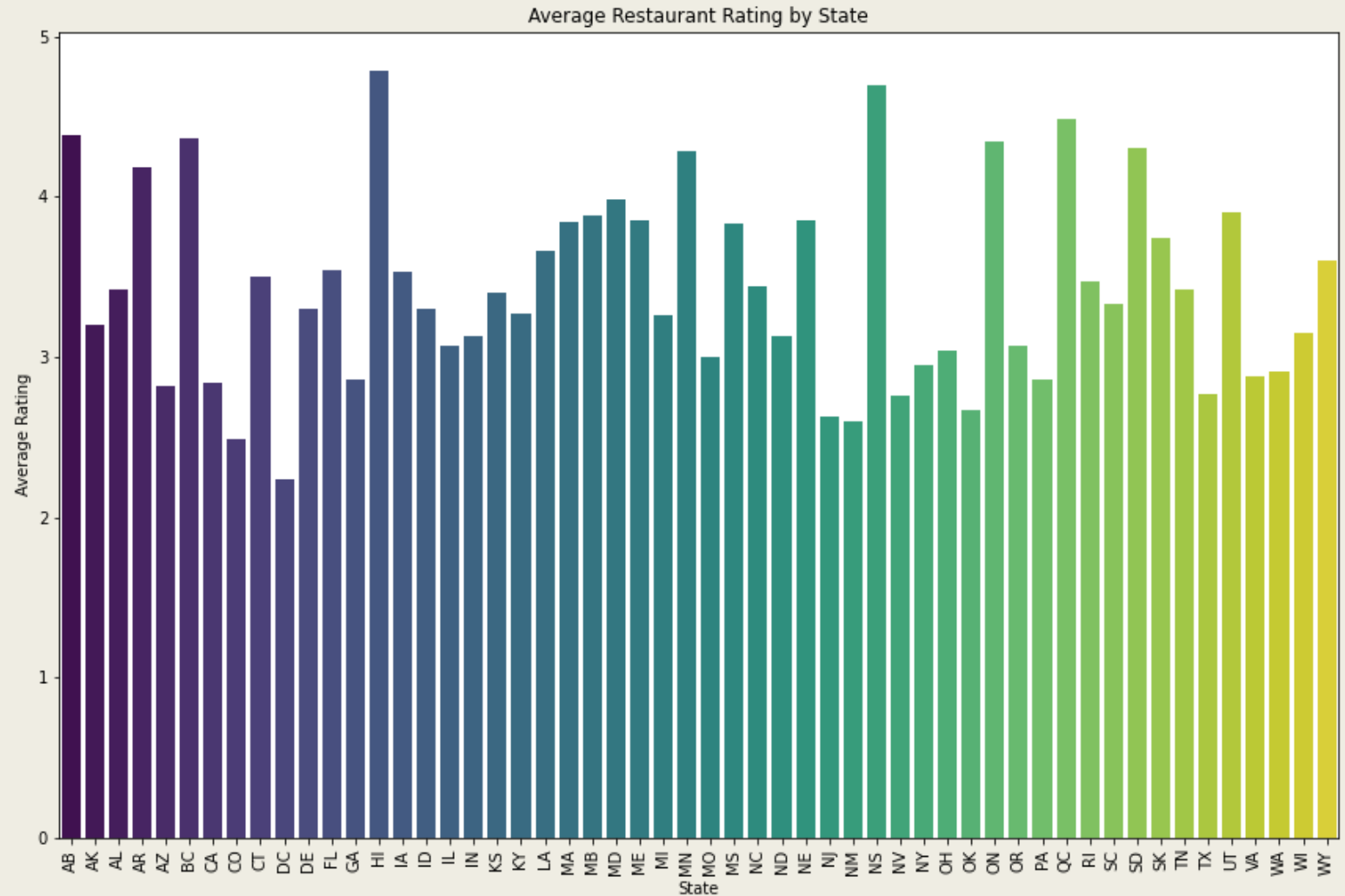
- Step 1: Download Data : Used httpx library to fetch data from Kaggle.
- Step 2: Data Extraction: Utilized zip file library to extract CSV from ZIP archive.
- Step 3: Data Cleaning and Normalization : Standardized cuisine names using Gensim's glove2word2vec model to reduce unique cuisine entries from 21,521 to 572.
- Step 4: Data Storage : Saved cleaned data in SQLite database for easy querying using SQL Alchemy.

# Data Cleaning

- Duplicate and inconsistent entries in the cuisine column, such as: "burgers" vs. "burger", "Coffee and Tea" vs. "Coffee & Tea".
- ❖ **Result:**
  - Reduced unique cuisine names from 21,521 to 572.

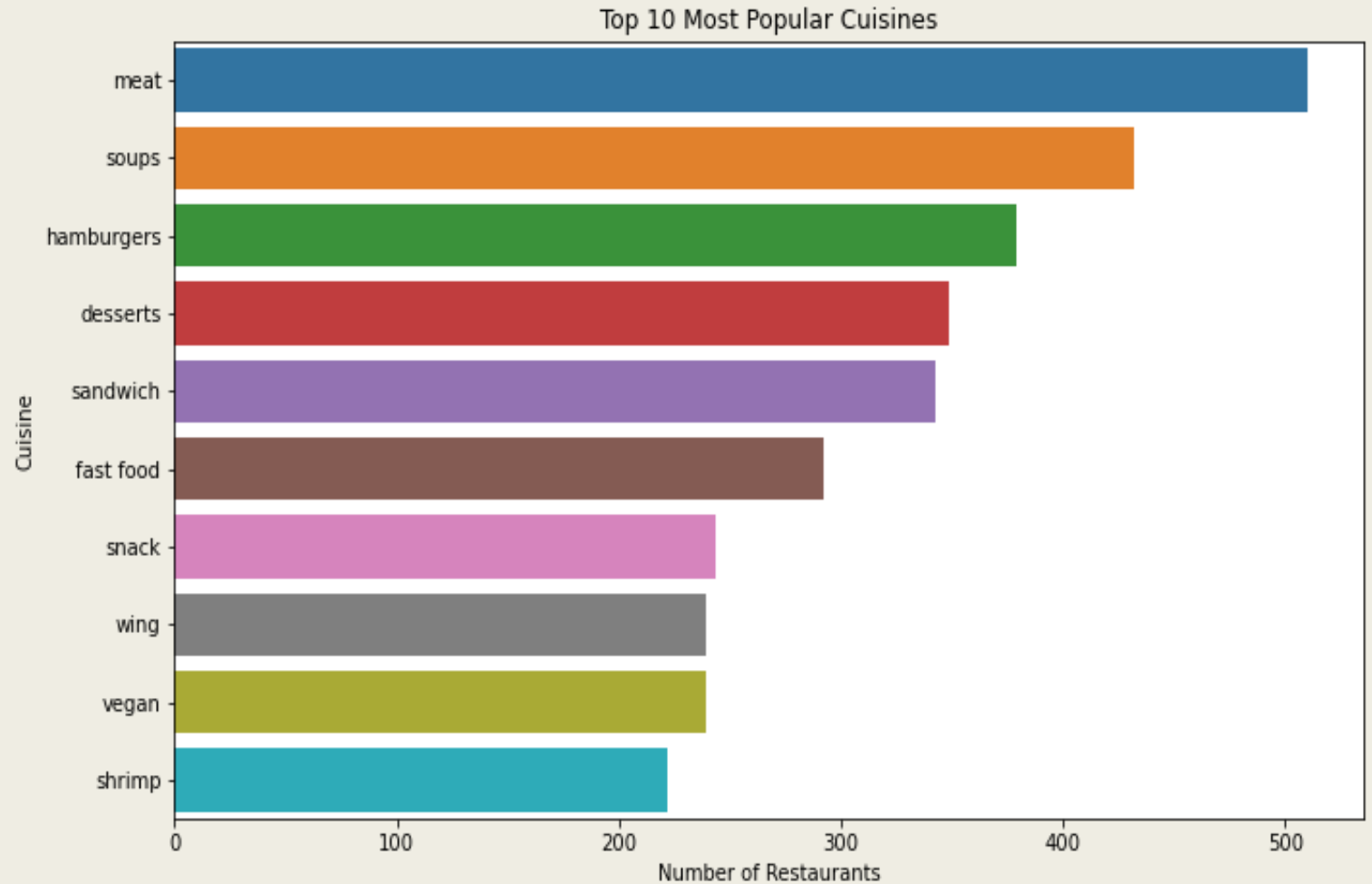
# Findings

- States like Hawaii (HI) and Quebec (QC) have higher average ratings.
- States such as Connecticut (CT), California (CA) and New Jersey (NJ) have lower ratings.
- Indicates potential variation in restaurant quality.



# Top 10 Most Popular Cuisines

- "Meat," "Soup," and "Hamburgers" are the top three most frequent cuisines.
- "Fast Food" and "Vegan" also feature prominently, reflecting evolving dietary preferences.



# Heatmap of Average Cuisine Ratings

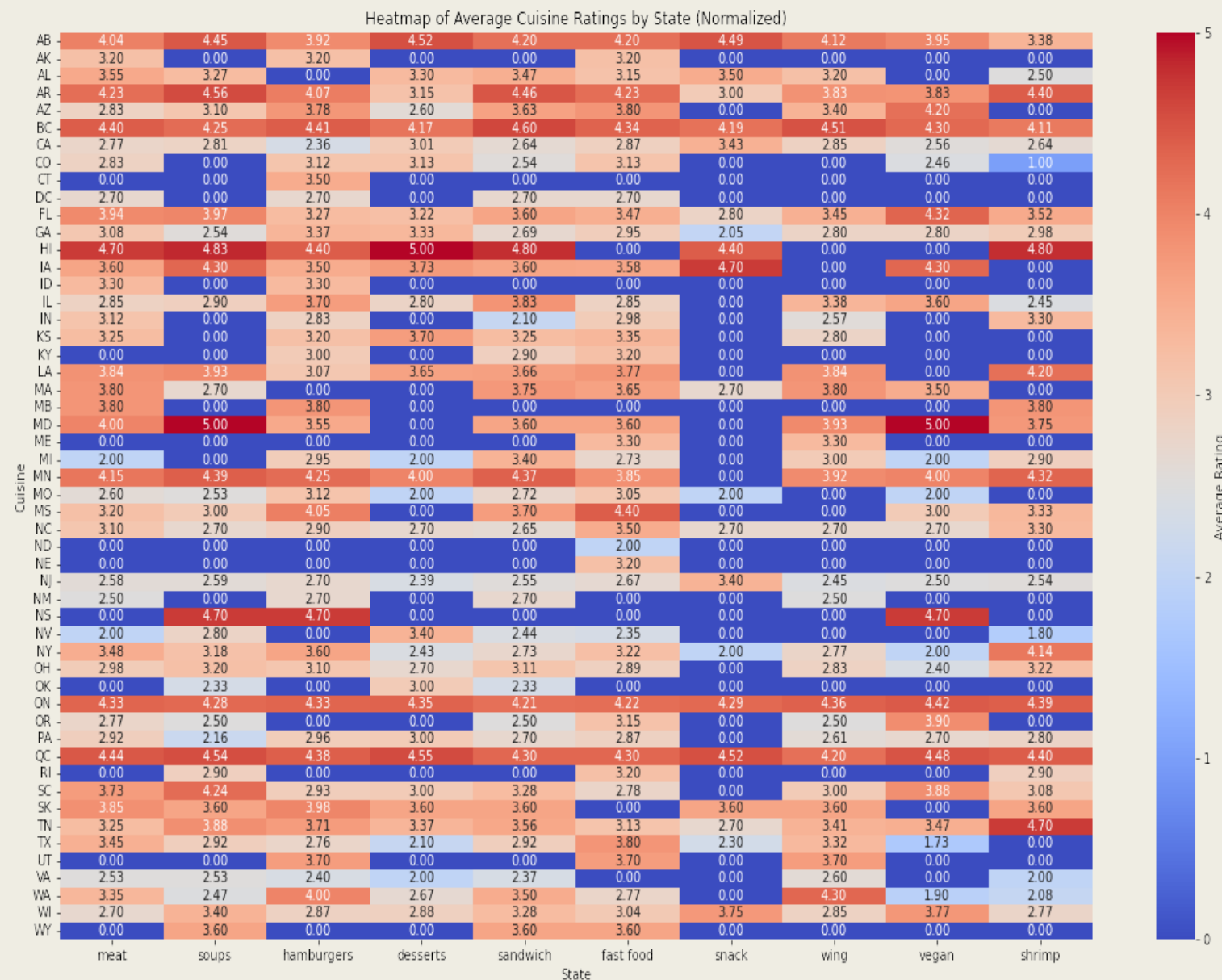
- Heatmap Showing Cuisine Ratings by State.

- Maryland (MD): High ratings for "Soup" and "Vegan."

- Georgia (GA): Low ratings for "Soup" and "Vegan."

- Michigan (MI): Low ratings across top 10 cuisines, indicating unique regional preferences.

- Consistent high ratings for "Hamburgers" and "Fast Food".





# Challenges and Limitations

- Only 1,500 restaurants.
- Simplification of cuisine names may overlook nuanced distinctions.
- Certain states show no data for some cuisines due to data sparsity.

# Conclusion

- Clear regional patterns in food preferences.
- High ratings for "Soup" and "Vegan" in some regions; consistent preference for "Hamburgers" and "Fast Food." Variability in restaurant quality across states.

## ❖ Applications:

- Insights can guide businesses in menu optimization and regional marketing strategies.

# Future Work

- Expand dataset to include more restaurants.
- Incorporate additional variables such as price range and customer demographics.
- Use advanced machine learning models for deeper insights.

Thank You!

