Final Data Analysis Report

2206-03 Capstone Project

A picture containing icon

Description automatically generated

MealMap

An analysis on reviews and ratings of various restaurants using SVM model.

**Team- 10**

**Rahul Sapparapu:** Project Manager

**Jayarani Rajesh:** Data Analyst

**Arunkumar Senthilkumar:** Business Analyst

Table of Contents

[Executive summary: 2](#_Toc132021862)

[Problem description: 3](#_Toc132021863)

[Business goal: 3](#_Toc132021864)

[Data analysis goals: 3](#_Toc132021865)

[Dataset description: 4](#_Toc132021866)

[Data preparation details: 4](#_Toc132021867)

[Data analysis solution: 4](#_Toc132021868)

[Recommendations: 5](#_Toc132021869)

[Limitations: 5](#_Toc132021870)

[Conclusion: 5](#_Toc132021871)

[Appendix: 5](#_Toc132021872)

[Findings: 5](#_Toc132021873)

# Executive summary:

MealMap is a Startup company founded in 2022 in response to the challenges faced by the restaurant industry during the COVID-19 pandemic. With a significant increase in takeout and delivery orders, coupled with consumers' reliance on online reviews for dining decisions, there was a clear need for a platform that could facilitate a more convenient and personalized way for foodies to connect with local restaurants. MealMap aims to fulfill this need by providing a user-friendly platform that simplifies the process of discovering and ordering food, while promoting local restaurants and their unique offerings.

The objective of this report is to describe MealMap's data analysis solution and provide recommendations for operational improvements. The data analysis goal is to analyze user reviews and ratings to identify patterns and insights that can help MealMap, and its restaurant partners improve the user experience.

The dataset used for the analysis includes reviews and ratings for seven different cuisine types of restaurants: Korean, Japanese, Chinese, Vietnamese, Thai, French, and Italian. The data includes information on the restaurant name, cuisine type, location, ratings, and reviews.

Through the data analysis, MealMap can provide personalized recommendations to users based on their preferences and past behavior, help restaurants improve their service, quality, and overall customer experience, and identify trends and patterns in user behavior that can inform marketing strategies and product development. Preliminary recommendations include collecting additional data on user behavior, conducting surveys and focus groups to gain deeper insights into user preferences and needs, and partnering with restaurants to offer exclusive deals and promotions to MealMap users.

Overall, MealMap's data analysis solution has the potential to significantly improve the user experience for foodies and restaurants alike, and we believe that the company is well-positioned for success in the highly competitive online food ordering and delivery market.

Problem description:

MealMap is a platform for ordering meals that connects foodies with nearby eateries. Due to the rise in takeout and delivery orders and the increased reliance of customers on internet evaluations when choosing where to eat, the COVID-19 epidemic has had a huge influence on the restaurant business. MealMap strives to offer a user-friendly platform that makes it easier to find and order food while highlighting nearby eateries and their distinctive features.

Business goal:

MealMap's main commercial objective is to raise the profile and clientele of nearby eateries. MealMap seeks to assist establishments stand out in a fiercely competitive market and draw in new customers by offering a quick and personalized solution for foodies to interact with nearby eateries.

Data analysis goals:

Analyzing user feedback and ratings with the aim of finding trends and insights that can assist MealMap and its restaurant partners enhance the user experience. Finding trends and patterns in user behavior, such as preferences for cuisine kinds, price points, and locations, is necessary for this. MealMap may help restaurants improve their service, quality, and overall customer experience by identifying these patterns and using that knowledge to make personalized recommendations to users.

# Dataset description:

MealMap received a Yelp dataset with reviews and ratings for restaurants serving seven distinct cuisines: Korean, Japanese, Chinese, Vietnamese, Thai, French, and Italian. The dataset is a broad and varied resource for examining customer behaviour and restaurant performance because it includes 5.2 million reviews and 174,000 companies.

The dataset's input variables include user-generated reviews and ratings as well as details about the restaurant, such as its name, location, and cuisine type. The restaurant's overall score, the individual ratings for certain features like the calibre of the food, the level of service, and the ambiance, as well as the review's language, are among the output variables. Because they offer extensive and specific information about user preferences, experiences, and expectations, text reviews are extremely significant. Raw data pre-processing procedures were used to clean the data and eliminate any duplicates or unnecessary information to get it ready for analysis. Key charts were used to depict the resultant data to give a general overview of the data distribution and spot any potential outliers or abnormalities.

## Data preparation details:

Before analyzing the Yelp dataset, MealMap applied raw data pre-processing techniques to ensure the accuracy and reliability of the data. The following steps were taken:

1. Data Cleaning: Duplicate entries and irrelevant information were removed from the dataset. This was done to ensure that the analysis was based on unique and relevant data.
2. Data Formatting: The data was formatted to ensure consistency and ease of analysis. For example, all restaurant names were standardized to ensure that they were consistent throughout the dataset.
3. Data Integration: The Yelp dataset was integrated with other relevant datasets, such as location data and demographic data, to provide additional insights and context.
4. Data Sampling: A random sample of the data was taken to ensure that the analysis was manageable and representative of the overall dataset.
5. Data Transformation: The data was transformed into a format that was suitable for analysis. This included converting text reviews into numerical data using natural language processing techniques.
6. Data Validation: The data was validated to ensure that it was accurate and free from errors. This was done by cross-checking the data with external sources and using statistical methods to identify any potential outliers or anomalies.

Overall, the data preparation process ensured that the data was accurate, consistent, and suitable for analysis.

# Data analysis solution:

Exploratory Data Analysis:

During exploratory data analysis, we counted the number of restaurants in each cuisine type and the number of reviews for each cuisine type. We also visualized the distribution of restaurants based on their ratings and cuisine types using charts and graphs.

Review analysis:

To analyze user reviews and identify patterns and insights, we performed the following steps:

* Cleaned the category column in the business table to identify different cuisine types of restaurants.
* Linked the business ID in the business table to the review table to collect all the reviews for each cuisine type of restaurant.
* Conducted sentiment analysis on the reviews to analyze the most frequent words in positive and negative reviews.
* Implemented an SVM model to get relatively positive and negative words and their scores.
* Identified the top 10 positive and negative words for each cuisine type of restaurant to understand the reasons behind high and low scores.
* Compared the advantages and disadvantages of different types of restaurants to generate a series of recommendations for future development.

# Recommendations:

Based on our analysis, we generated a series of recommendations for restaurants to improve the user experience. The recommendations cover a range of topics, including service, food quality, and decor. We analyzed review words such as "rude", "overpriced", and "slow" to identify specific areas where restaurants could improve. By implementing these recommendations, restaurants can enhance their overall customer experience, increase their ratings and reviews, and attract more customers to their establishments.

# Limitations:

The quantity and scope of the dataset, which might not fully represent the population of foodies and restaurants, is one of the constraints of our analysis. The reliability of the reviews and ratings, which might not always be accurate or objective, also places constraints on our analyses. To resolve these limitations, additional study and analysis may be required.

# Conclusion:

Data analysis provided insightful information about user preferences and behaviour. Users are prepared to spend extra for restaurants that provide distinctive and sustainable options because they appreciate authenticity and quality, according to our research. Additionally, we discovered that people choose tailored recommendations that are based on their previous choices and behaviour. Our suggestions include gathering more information on user behavior, running polls and focus groups, and working with eateries to provide special offers and promotions. We advise MealMap to keep placing a high priority on supporting regional companies and sustainability.

Appendix:

Findings:

As a result of our analysis, we have made the following recommendations to different restaurant cuisines:

**Korean**

1. Taste should be emphasized, and fresh ingredients should be used.
2. Make sure that your waitstaff is always friendly and professional.
3. It is important to advertise on social media, adjust prices if necessary, and create a menu to suit your audience.
4. Keeping the environment clean is one of the most important things to do.

Chart, bar chart

Description automatically generated

**Japanese**

1. Make your menu creations using fresh ingredients and use fresh ingredients.
2. Make sure your waitstaff is friendly and avoids making mistakes by training them.
3. Make sure raw ingredients are kept fresh by investing in technology.
4. Keep the environment clean and adjust based on customer feedback.

Chart, bar chart

Description automatically generated

**Chinese**

1. The importance of delicious and authentic food cannot be overstated.
2. Prepare your wait staff so that they are professional and fast in their service.
3. Price your products and services in a reasonable manner.
4. Improve your cooking techniques by using fresh ingredients and using fresh ingredients.

Chart, bar chart

Description automatically generated

**Thai**

1. Ensure that the food you serve is tasty, fresh, and of a wide variety.
2. Make sure your wait staff is friendly, attentive, and quick in their service.
3. Price should be kept at a reasonable level.
4. Ensure that the environment is always kept clean and comfortable.

Chart, bar chart

Description automatically generated

**Vietnamese**

1. Make sure that you serve your food in a tender and generous manner.
2. It is important to maintain a clean and attractive environment.
3. Price your products at a reasonable level.
4. Make sure that waitstaff are trained to be professional, nice, and to avoid always making mistakes.

Chart

Description automatically generated

**French**

1. Make sure that the environment is always clean and fancy.
2. Provide tender and generous servings of food to your patients as a matter of priority.
3. Educate your wait staff on the importance of being polite, professional, and avoiding mistakes.

Chart, bar chart

Description automatically generated

**Italian**

1. Make sure that the environment is always clean and fancy.
2. Provide tender and generous servings of food to your patients as a matter of priority.
3. Make sure that waitstaff are trained to be professional, nice, and to avoid always making mistakes.

Chart, histogram

Description automatically generated

Data cleaning:

Graphical user interface, text, application

Description automatically generated

Merge Two Datasets:  
Text

Description automatically generated

A picture containing text

Description automatically generated