Cafe Dataset Analysis – Kelvin Kumar

Introduction

The purpose of this project was to analyse customer feedback data from multiple café locations and present key insights through a clear and actionable dashboard. Using Excel for data cleaning and preparation, I processed the dataset to ensure consistency, accuracy, and usability. I then created two visualisations: one showing the average customer rating for each location, and another highlighting the most frequently used words in customer comments.

Together, these outputs provide both quantitative and qualitative insights into customer experience, helping identify performance trends across locations and recurring themes in written feedback. The aim of this walkthrough is to outline the steps taken, explain the reasoning behind each decision, and present the insights generated from the data.

Dataset Overview

The dataset contains customer feedback from multiple café locations. Key fields include:

- **Location** – the café where the feedback was submitted
- **Rating (1–5)** – the customer's score for their experience
- **Comment** – written feedback describing the visit
- **Transaction Date and Time** – when the feedback was recorded
- **Transaction Value** – the amount spent on the transaction

This combination of structured and unstructured data enables both quantitative analysis (ratings, trends, comparisons) and qualitative analysis (common themes in comments). The dataset contains customer feedback from multiple café locations, including both numerical ratings and written comments. It includes fields for location, customer rating, customer comment, transaction timestamp, and transaction value. This combination of structured and unstructured data supports quantitative trend analysis as well as qualitative insights into customer experience.

Data Cleaning & Preparation

Tools Used: Excel, Tableau

One issue in the dataset was that some transaction dates were incorrectly stored in the *transaction value* column. Although this affected only 0.16% of rows, it needed to be corrected to ensure accurate analysis. I created a helper column ("transaction_date_and_time_helper") to reconstruct the proper timestamp and then moved the corrected values into the appropriate field. The transformation used a date–time reconstruction formula based on the original numeric pattern:

=UPPER(TEXT(DATE(2000+LEFT(x,2), MID(x,3,2), MID(x,5,2)) + TIME(MID(x,7,2), MID(x,9,2), MID(x,11,2)), "dd/mm/yyyy h:mm:ss AM/PM"))

Additionally, a small number of stray feedback IDs were found outside the main table. These were removed as they were not part of the dataset and were not required for analysis.

Another step in preparing the dataset involved creating a separate sheet to analyse the frequency of specific words and phrases within customer comments. This helped identify the topics customers mentioned most often after visiting the cafés. The sheet included two columns: "Word/Phrase" and "Appearances."

To calculate how many times a given word or phrase appeared across all comments, I used the following formula:

=SUMPRODUCT((LEN(edited_data!$C$2:$C$2408)

- LEN(SUBSTITUTE(LOWER(edited_data!$C$2:$C$2408), LOWER(x), "")))

/ LEN(x))

Here, **x** represents the word or phrase being analysed. This method allowed for consistent, case-insensitive counting and provided a clear overview of the most common themes in customer feedback.

Methodology

After cleaning the dataset in Excel, I structured it for analysis and visualisation. This included correcting data types, reconstructing timestamps, and ensuring each field aligned with its intended meaning.

I then selected two key visuals to highlight both numerical trends and comment themes:

- **Average Rating by Location** to compare performance across cafés.
- **Most Common Words in Customer Comments** using the word-frequency calculations created in Excel.

These visuals were chosen to combine structured data (ratings) with unstructured insights (comment themes), providing a more complete picture of customer experience.

Finally, I arranged the dashboard in Tableau with a clear, simple layout that allows users to quickly identify trends, compare locations, and understand the main topics customers mention in their feedback.
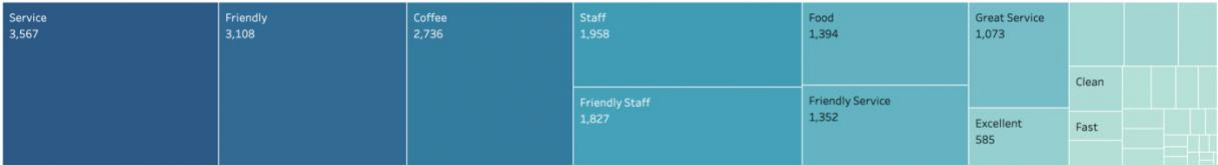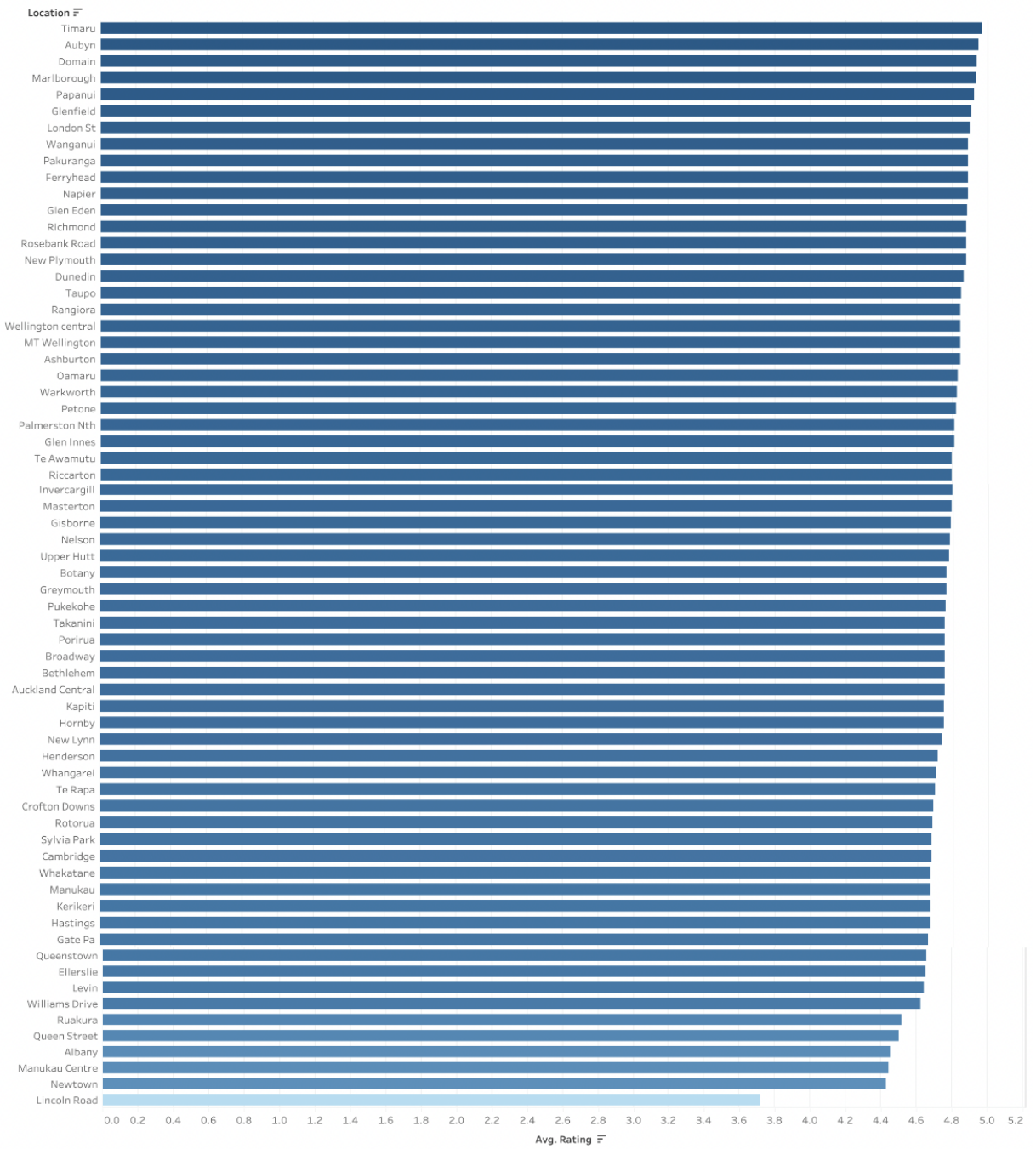
Dashboard Overview

Link to Dashboard:

https://public.tableau.com/app/profile/kelvin.kumar/viz/cafe_graphs/Dashboard1?publish=yes

Screenshots:

Appearances of Most Common Words in Customer Feedback

Average Rating Across Location

AI-Assisted Analysis

AI-Assisted Analysis of Customer Ratings and Location Performance

The customer feedback graph highlights the most frequently used words and phrases found in the comments submitted across all café locations. This analysis was carried out to identify the topics customers mention the most, which helps reveal what aspects of the café experience matter most to them. It also provides insight into both positive experiences and potential areas where improvements could be made.

The results show that several key themes consistently appear throughout the comments. Words relating to staff and customer service are among the most common, which indicates that interactions with employees play a central role in how customers evaluate their visit. Frequent appearance of terms related to friendliness, speed and helpfulness suggests that customers value efficient and approachable service.

Product-related words also appear often, particularly those referring to coffee, food quality and freshness. This shows that customers tend to comment on the core offerings of the café and that these items strongly influence satisfaction levels. Repeated references to quality or consistency may reflect what customers expect from the brand on a regular basis.

Another pattern visible in the graph is the use of location-specific or environment-focused terms, such as cleanliness, space and atmosphere. This suggests that the physical environment contributes to the overall experience and affects how customers feel during their visit.

The high frequency of positive terms aligns with the rating distribution, which shows that most customers leave favourable reviews. However, the word analysis also highlights recurring concerns in lower-rated comments, such as long waits, order mistakes or temperature issues. These themes appear less often but still provide valuable guidance for operational improvements.

Overall, the feedback graph provides a summary of what customers talk about the most. It reinforces the strengths shown in the ratings data and offers direction for consistent improvement across locations. By monitoring these key themes over time, management can identify shifts in customer expectations and respond more effectively.

Key Insights and Recommendations

Key Insights

- Customer feedback is strongly centred around service and staff interactions. Words such as "service," "friendly," "staff" and "great service" appear most frequently, showing that customer experience is shaped primarily by the quality of personal interactions.
- Product-related terms such as "coffee" and "food" also appear often, confirming that core offerings continue to influence overall satisfaction.

- Average ratings across locations are consistently high, with most cafés scoring between 4.3 and 4.8. Only a small number of locations show ratings below this range, indicating generally strong performance across the network.
- Locations with slightly lower ratings tend to have a higher volume of comments that reference wait times, order accuracy or inconsistent service, suggesting a link between operational issues and lower satisfaction.

Recommendations

- Continue investing in staff training to maintain a friendly and consistent level of service, as this is the most influential factor in customer feedback.
- Review processes at the lower-rated locations to identify the cause of longer wait times or order errors, as addressing these issues will likely increase ratings.
- Encourage regular product quality checks to ensure coffee and food remain consistent across all sites.
- Use ongoing keyword tracking to monitor any shifts in customer concerns and to identify emerging trends early.
- Consider sharing positive customer themes across locations to reinforce best practices and strengthen the overall customer experience.

Conclusion

This project delivered a cleaned dataset and visual insights that highlight customer satisfaction levels, and the key themes customers mention in their feedback. The analysis adds value by showing which locations perform strongly, where improvements are needed and what customers prioritise most during their visits. If more time were available, the next steps would include deeper sentiment analysis, trend analysis over time and linking feedback with operational data for stronger decision making.