APPLYING NLP THROUGH TOPIC MODELING

- Presented by Poonam Vyas

Table of Contents

1. Introduction

Understanding customer sentiment at scale is critical for organisations operating in highly competitive, experience-driven markets such as the fitness industry. With thousands of customer reviews published across digital platforms, the partner organisation has access to a rich but complex source of insight that, if effectively analysed, can highlight both systemic issues and opportunities to strengthen customer experience.

This project was undertaken to support the organisation's leadership in **translating unstructured customer feedback into actionable insight**. By analysing customer reviews from **Google** and **Trustpilot**, the objective was to identify recurring patterns in dissatisfaction, assess the emotional intensity of customer experiences, and determine whether specific locations or operational areas warrant deeper investigation.

Rather than focusing on individual complaints, the analysis adopts a **data-led, thematic approach** to uncover broader drivers shaping customer perceptions. A combination of frequency-based analysis, topic modelling, emotion detection, and large language models was used to progressively refine insights—from surface-level patterns in review language to deeper understanding of emotionally charged feedback and its underlying causes.

The findings of this report are intended to inform strategic decision-making by highlighting where operational processes, service delivery, or customer interactions may be misaligned with expectations. By connecting customer sentiment to recurring themes and locations, the analysis aims to provide the organisation with clear direction on where targeted interventions could deliver the greatest impact on customer satisfaction and brand perception.

2. Executive Summary

This project analysed large volumes of customer reviews from **Google** and **Trustpilot** to identify recurring issues, emotional drivers of dissatisfaction, and opportunities for improvement across the partner organisation's locations. Using a combination of **text preprocessing**, **topic modelling**, and **emotion analysis**, the investigation uncovered clear, consistent patterns behind negative customer experiences.

The results show that **customer dissatisfaction is primarily driven by operational and service-related issues**, rather than isolated incidents. Problems around **customer service quality**, **membership management**, **overcrowding**, **equipment availability**, and **staff behaviour** consistently appear across both platforms and locations. Emotion analysis further confirms that **anger is strongly associated with these themes**, particularly where customers feel ignored, unfairly charged, or poorly treated.

Crucially, the findings provide **actionable insights**: addressing service responsiveness, improving staff training, managing capacity, and increasing transparency around membership and fees would likely have the greatest impact on reducing negative sentiment.

3. Data Overview & Preparation

The analysis began with a large real-world dataset:

- **Google reviews:** 23,250 records
- **Trustpilot reviews:** 16,673 records

After removing non-English reviews and records with missing values, the refined datasets consisted of:
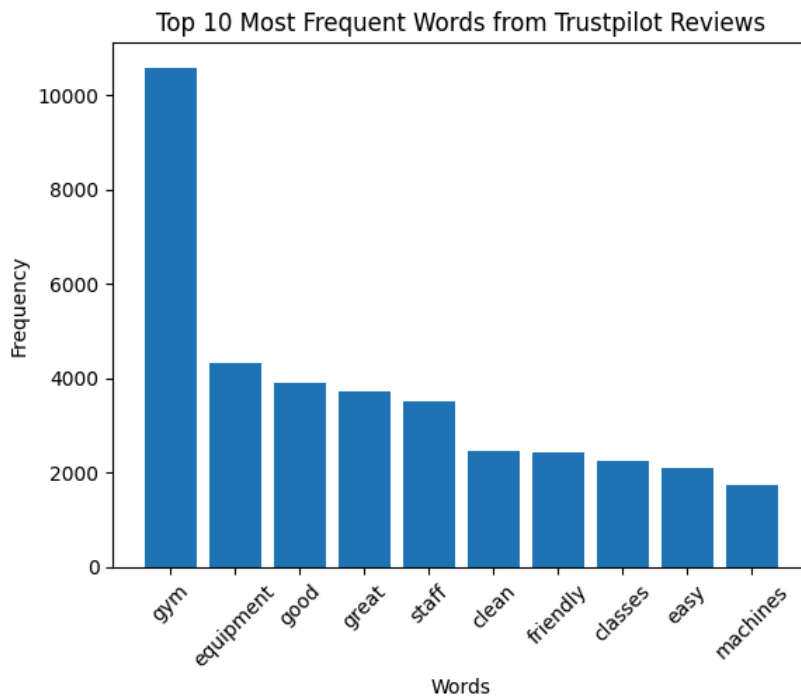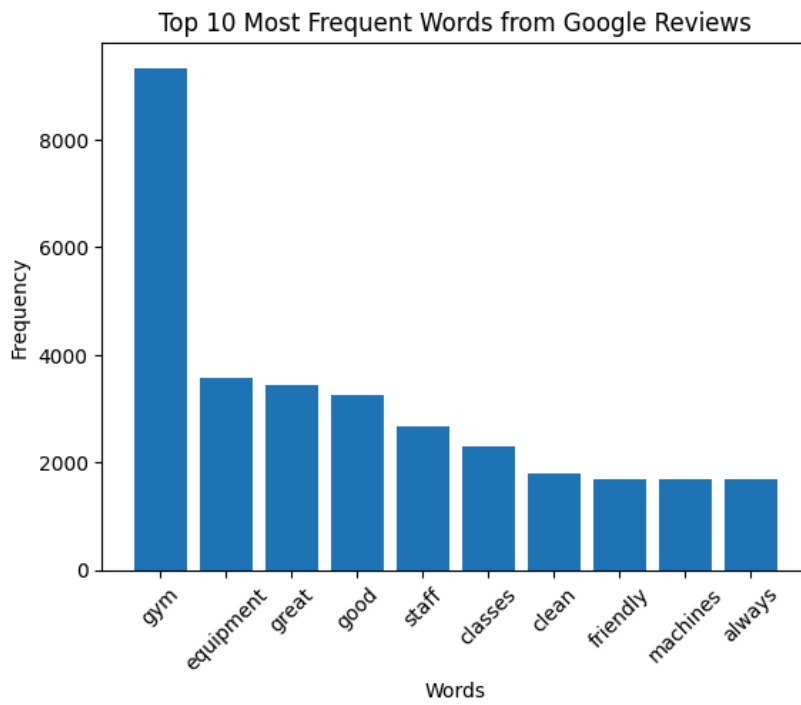
- **Google:** 11,879 reviews
- **Trustpilot:** 15,813 reviews

This preprocessing step significantly improved data quality and ensured that downstream analysis reflected genuine customer feedback rather than noise.

Across both datasets, reviews covered a wide geographical footprint:

- **455 unique locations** in Google
- **376 unique locations** in Trustpilot
- **310 locations** common to both platforms

This overlap enabled meaningful cross-platform comparison and trend validation.

Top 10 Most Frequent Words from Google Reviews

Top 10 Most Frequent Words from Trustpilot Reviews

Word Cloud from Google Reviews

Word Cloud from Trustpilot Reviews

3a. Analytical Approach

To ensure reliable and meaningful topic identification, a combination of complementary techniques was applied.

- **BERTopic** was used as the primary method to capture semantically rich themes from customer reviews.
- **LLM-based topic extraction using the Phi-4-mini model via Ollama** was employed to generate human-readable, review-level topics aligned with business interpretation.
  - While this approach successfully produced actionable recommendations for the majority of negative reviews, a subset of highly emotional or low-context comments did not yield structured suggestions across all fields.
  - A minority of highly emotional or low-context comments did not yield structured suggestions, highlighting that emotional intensity does not always translate into clear, actionable feedback.
- **LDA** was applied at the final stage as a validation step, enabling comparison with earlier techniques and assessing whether any additional or overlooked insights emerged.

Together, this layered approach strengthened confidence in the findings and confirmed the consistency of key themes.
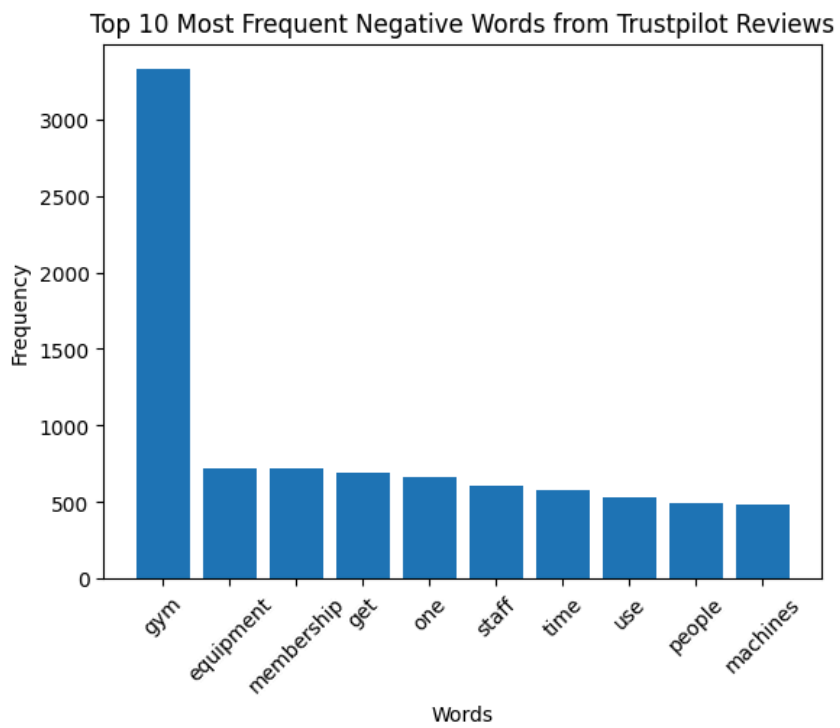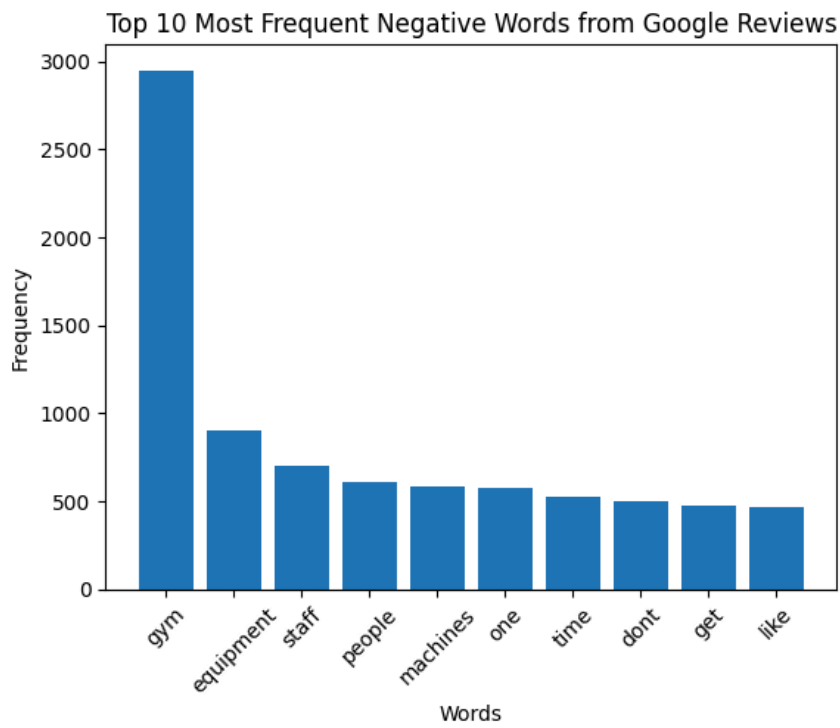
4. Identifying Negative Feedback Hotspots

Using review scores to identify dissatisfaction, the analysis isolated:
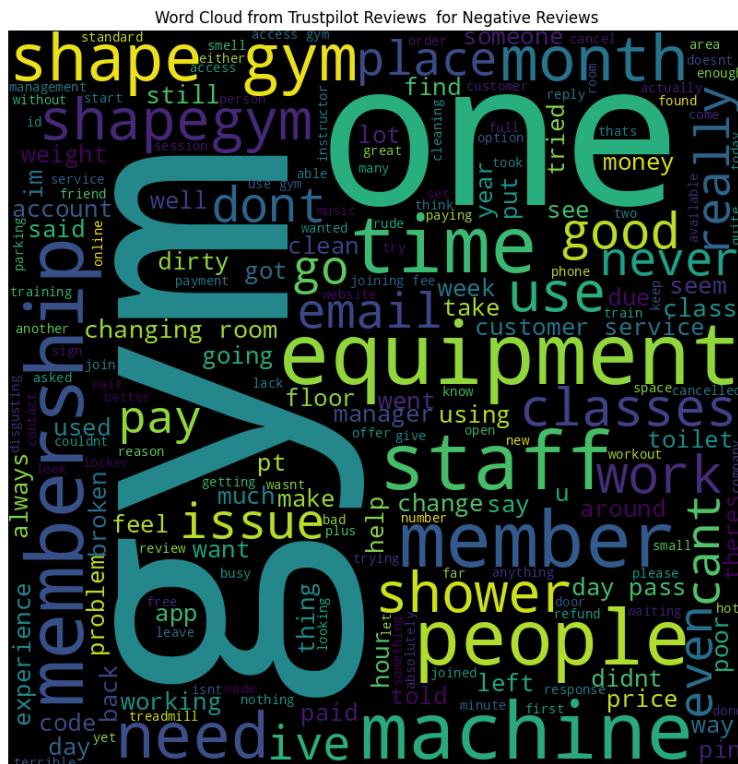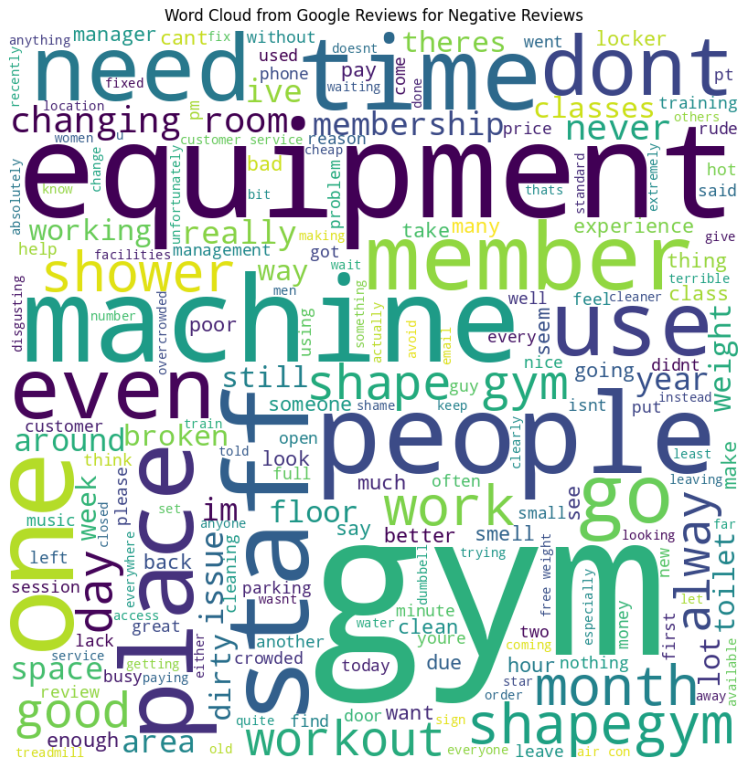
- **2,423 negative Google reviews**
- **3,403 negative Trustpilot reviews**

Several locations consistently appeared among the **top 20 sites with the highest number of negative reviews** on both platforms, including:

- London Stratford
- London Enfield
- Birmingham City Centre
- Bradford Thornbury
- New Barnet

The presence of the same locations across platforms indicates that issues are **location-specific rather than platform-specific**, suggesting operational challenges at certain gyms rather than bias in review channels.

Top 10 Most Frequent Negative Words from Google Reviews

Top 10 Most Frequent Negative Words from Trustpilot Reviews

Word Cloud from Google Reviews for Negative Reviews

Word Cloud from Trustpilot Reviews for Negative Reviews

5. Topic Modelling: What Customers Are Complaining About

Multiple topic modelling techniques were applied, including **BERTopic**, **LLM-assisted topic extraction**, and **LDA**, to ensure robustness. Across methods, the same core themes repeatedly emerged.

The most frequent and impactful topics include:

1. **Customer dissatisfaction and service experience**
   This was the most prominent theme across both datasets. Customers frequently describe feeling ignored, frustrated by poor communication, or dissatisfied with how issues were handled.
2. **Membership cancellation, upgrades, and billing issues**
   Complaints around cancellation difficulty, unexpected charges, refunds, and unclear policies were common and emotionally charged.
3. **Overcrowding and limited equipment availability**
   Many reviews mention gyms being too busy, insufficient machines, or long waits, particularly during peak hours.
4. **Staff behaviour and attitude**
   Negative interactions with staff—perceived rudeness, lack of empathy, or unhelpfulness—frequently escalated dissatisfaction.
5. **Cleanliness, hygiene, and facilities**
   Concerns around cleanliness, broken equipment, showers, air conditioning, and general maintenance also contributed to poor experiences.

Importantly, the latest BERTopic output shows **improved granularity** compared to earlier runs. Previously broad themes such as "service issues" are now clearly separated into **staff behaviour**, **response times**, **technical issues**, and **policy-related frustrations**, allowing more targeted insight.

6. Emotion Analysis: Understanding Customer Anger

Emotion analysis using a BERT-based model was applied to identify emotionally charged reviews. The results revealed **anger as the most prevalent emotional response** across customer reviews on both Google and Trustpilot:
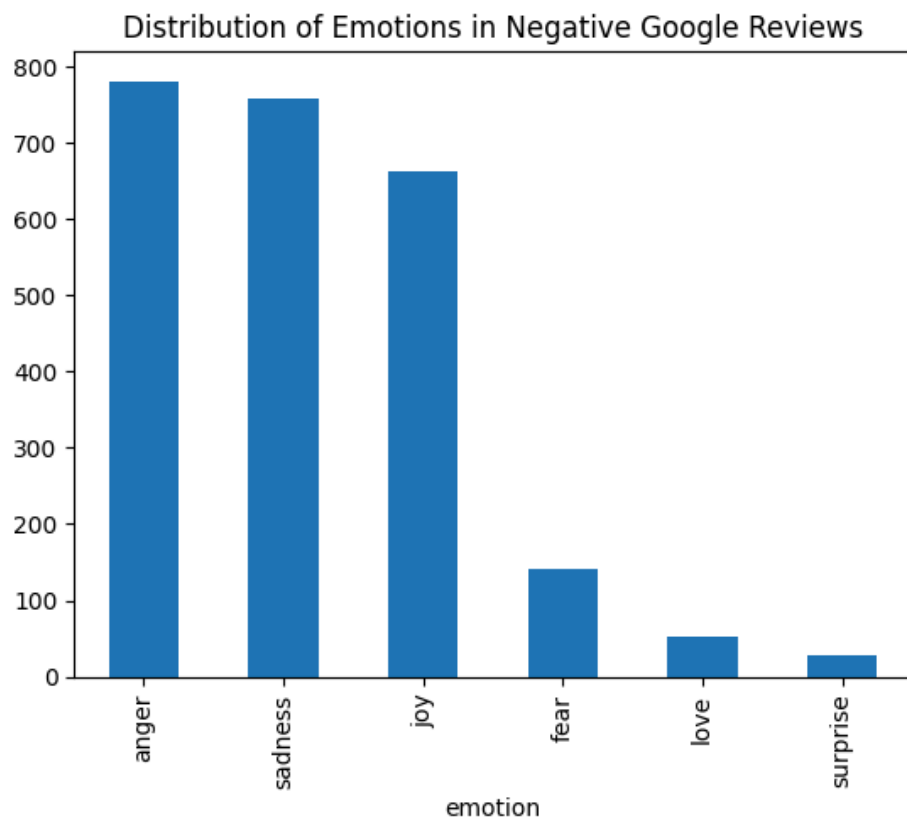
- **781 Google reviews** expressing anger
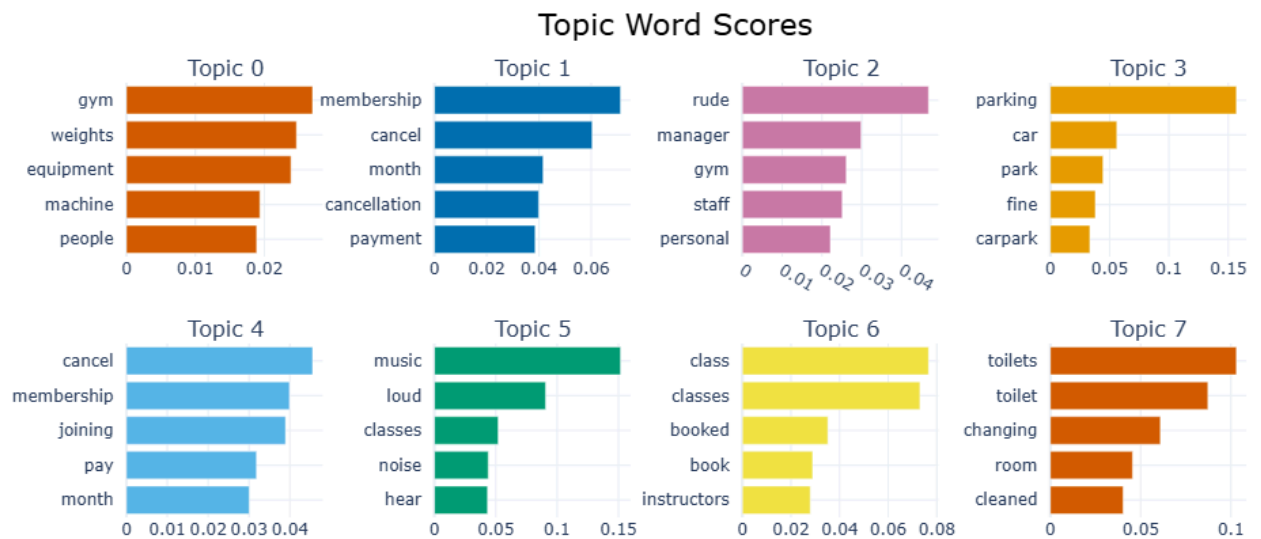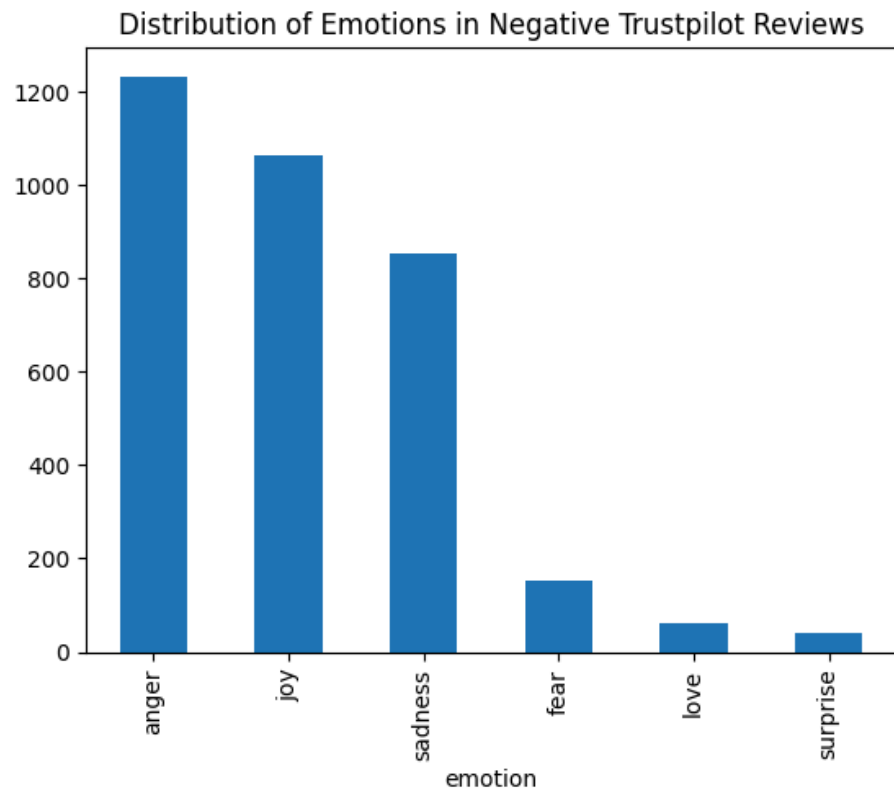- **1,232 Trustpilot reviews** expressing anger

These results indicate a **sustained and widespread level of customer dissatisfaction**, highlighting issues that extend beyond isolated incidents and warrant strategic attention. These reviews were treated as a priority segment for analysis, as anger often signals breakdowns in service delivery, unresolved operational issues, or perceived lack of fairness—factors that pose the greatest risk to customer trust and brand reputation.

Anger was strongly associated with:

- Poor customer service responses
- Membership and billing disputes
- Repeated access issues (PINs, passes, app problems)
- Perceived unfairness or lack of accountability

This confirms that **anger is not random**; it is consistently linked to **process failures and human interactions**, rather than facilities alone.

Distribution of Emotions in Negative Google Reviews

Distribution of Emotions in Negative Trustpilot Reviews

Topic Word Scores

7. Key Insights & Trends

Several high-level insights emerge from the combined analysis:

- **Service issues outweigh facility issues**
  While equipment and cleanliness matter, customers are more likely to leave angry reviews when service problems are handled poorly.

- **Operational friction drives emotional escalation**
  Access failures, cancellations, and billing errors are tolerable initially, but become major sources of anger when resolution is slow or unclear.

- **Staff behaviour acts as an emotional multiplier**
  Negative staff interactions significantly intensify dissatisfaction, even when the original issue is minor.

- **Problems are recurring and systemic**
  The repetition of topics across locations and platforms suggests structural issues rather than isolated incidents.

	topic	total_count	google_count	trustpilot_count
178	customer dissatisfaction	8	1	7
59	customer service experience	3	3	0
12	gym overcrowding	3	2	1
3	hygiene and cleanliness	2	2	0
119	customer service issues	2	1	1

8. Actionable Recommendations

Based on the insights obtained, the following actions are likely to have the greatest impact:

1. **Strengthen customer service responsiveness**
   Faster response times, clearer escalation paths, and consistent communication would directly address the largest source of dissatisfaction.

2. **Improve staff training and accountability**
   Training focused on empathy, conflict handling, and service recovery could reduce emotionally charged interactions.

3. **Increase transparency around memberships and billing**
   Clearer cancellation processes, fee explanations, and proactive communication would reduce frustration and anger.

4. **Address overcrowding and equipment availability**
   Location-specific capacity management and equipment planning would improve perceived value for money.

5. **Actively monitor high-risk locations**
   Locations repeatedly appearing in negative review rankings should be prioritised for operational review.