



Understanding Customer Perceptions: Topic Modeling Analysis of Toronto Specialty Coffee Shop Online Reviews

Diego Mauricio Mora Garzon ¹, Shadi Ebrahimi Mehrabani ²

MBA candidate, International Business University (IBU), Toronto, Canada¹

Faculty member, International Business University (IBU), Toronto, Canada²

Abstract: The specialty coffee shop market in Toronto has become increasingly competitive, making it essential for business owners to understand the factors that drive customer satisfaction and differentiation. This study aims to identify the main themes expressed in Google Maps reviews of Toronto's specialty coffee shops over the past year, providing actionable insights for entrepreneurs and industry stakeholders. Over 5000 customer reviews were analyzed using BERTopic (Bidirectional Encoder Representations from Transformers Topic), a state-of-the-art topic modeling approach that leverages contextual language understanding to extract clear and meaningful topics from large volumes of text. The analysis revealed distinct positive themes, such as cozy atmospheres and high-quality coffee, as well as negative aspects like unfriendly service and poor value for money. By correlating these topics with review ratings, the study highlights specific opportunities for improvement and differentiation in the market. These findings offer practical value for business planning, enabling coffee shop owners to make data-driven decisions and enhance customer experiences in a crowded urban landscape. Beyond its local insights, this research introduces a scalable analytical framework that can be applied to market research, business planning, and feasibility studies in diverse sectors, empowering others to extract actionable intelligence from large volumes of unstructured textual data.

Keywords: Topic modeling, BERTopic, Google reviews, Specialty Coffee

I. INTRODUCTION

Specialty coffee shops—known for their high-quality, artisanal coffee—play a significant role in modern society and have become cultural landmarks. These businesses are found in almost every major city worldwide, attracting many customers. However, the specialty coffee industry is highly competitive, especially in large North American cities, where the market is saturated. To succeed, coffee shop owners must differentiate themselves and maintain high quality [1].

Even with many online customer reviews, there is little understanding of what makes customers happy or unhappy in this sector. This research addresses this gap by identifying specialty coffee shops' key strengths and weaknesses in Toronto, Canada, by analyzing customer reviews from Google Maps. The goal is to support business planning by highlighting opportunities to improve negative experiences and build positive ones. All this combined will substantially impact mitigating the risk for new companies entering this market.

As a first approach, this research has the following general objectives:

- Using topic modeling to identify and categorize the top five positive and negative topics in customer experiences as expressed in Google Maps reviews of specialty coffee shops in Toronto.
- Develop actionable business planning and differentiation strategies recommendations based on the identified customer experience topics.

II. RESEARCH CONTEXT

This part provides an overview of the existing relevant literature, particularly focusing on the factors that contribute to success and value creation in both new and existing businesses. Following the literature review, the methodology outlines the approach that has been used in the research.

A. Background

The global specialty coffee market was estimated at USD 101.6 billion in 2024 and is expected to grow 10.4% from 2025 to 2030. This growth is primarily driven by changing consumer behavior, where ethical and sustainable production is



becoming increasingly important. Over the past decade, there has been a noticeable shift from low-quality to high-quality coffee consumption, a trend expected to continue over the next five years[1].

Additionally, coffee shops are not only places to get a product but also an experience around coffee, in some cases people use coffee shops as the preferred location for social encounters, reading alone or working in a more relaxed way. In cities today, coffee shops are becoming more common, and their purpose is expanding. People don't just go there to relax or work—they also use them as social spaces where they can connect with others and even form communities[2].

North America is a major leader when it comes to both drinking coffee and the growth of the specialty coffee market. In 2024, the region made up 32% of the global revenue from single-origin coffee, showing a strong local interest in high-quality and ethically sourced products. The United States alone brings in about 25% of all coffee imports worldwide, highlighting its key role in the global coffee industry [3].

B. Literature Review

Finding the correct value proposition is a challenge that every new business needs to deal with when entering in an already competitive industry, to support the relevance and to create a reliable methodology for this research, a literature review is structured to cover the importance of the specialty coffee industry, the value that is possible to extract from customers reviews using text processing and topic modeling, the importance of analyzing the external environment for an entrepreneur and also touching some general points on the legal and ethical concerns by using automatic extraction of data from web sites commonly known as scrapping.

Customers might have different opinions about the services and products consumed, but this perception is only available after the users have interacted with the service or the product. One place with a high data richness to investigate the customers opinions is online reviews, which from the business perspective is a free tool with data directly connected to their users' experience. These online data serve as powerful and reliable tools for measuring consumer perspectives [2]. Research has shown that positive online customer reviews increase the chances of people buying a product more than negative ones, which ultimately is correlated to a possible increase in sales. When reviews are negative, how believable and useful they seem plays an important role in the decision to buy [4].

Business insights are often the result of data analysis, and the source to achieve this process is having access to reliable data. As mentioned, customer-facing companies, like coffee shops, have access to UGC (User-generated Content) collected via online interactions, in particular for this case through Google reviews. From Xu et al. [4] research, this type of data has an open structure and is hard to analyze. The type of data is human written text that might have different grammatical structures and tones, in general every text is different, which creates a challenge when analyzing and interpreting the data. Additionally, companies tend to focus on other most common and easy to generate metrics related to customer service such as response times or financial information.

An ideal scenario for an entrepreneur is to understand what the market wants and what to avoid when planning a business. This, of course, is not entirely possible, in a research Mcadoo [5] concluded that 73% of the coffee shops owners who used a combination of customer loyalty strategies, including social media engagement, store experience, and inimitable products and services-controlled costs and improved revenues, however there is no practical immediate insights for the market in Toronto.

The positive side in this case is that there is data available regarding what the customers think, but there is no existing analysis of this kind for the specialty coffee industry in Toronto. To fill this gap, this research aims to implement topic modelling to extract value from the review text transcripts and contribute to building up that ideal scenario mentioned before.

The strategy to analyze this type of data and extract value from the insights is possible to do using topic modeling. This is an NLP (Natural Language Processing) method that generates a straightforward and easy to interpret overview of the different topics or issues that are present in a set of documents on in this case online reviews. Previous research has been done using other text processing methods like LDA (Latent Dirichlet Allocation), but the limitation is the lack of contextual information in those models; context is relevant for this study because in a customer review, we are dealing with natural human language [6]. In that order of ideas, this research employs the BERTopic model. Different studies have demonstrated that BERTopic extracts are meaningful and close to reality topics from a variety of documents [7].

Multiple studies have been performed using topic modelling on top of online customer reviews, however most of them use models based on the frequency of words found in the text to extract the topic but not including the context, for online reviews the context is relevant given that the value of the comment resides in the sentiment expressed by the user, this



raises a problem when interpreting the results. BERTopic has the advantage of generating easy-to-interpret sets of results, which is desirable when the goal is also to extract business insights that are not only clear but also actionable [8].

Lechaudé et al. [8] used BERTopic focused on the online retail industry to understand the customer returns drivers. Krishnan [6] used different techniques, including BERTopic, for more general purposes, getting clear results from their analysis. There is no evidence of using a state-of-the-art topic model like BERTopic in the coffee industry and none with the final purpose of performing a competitor's analysis to inform future business plans, taking advantage of the UGC.

As Calheiros et al. [9]. acknowledge, their study has limitations that warrant cautious interpretation of the results. One key issue is that it didn't identify any topics related to negative feedback, which makes it harder to find areas where customer satisfaction could be improved—even though there was less negative feedback overall. In the future, researchers could investigate using fully automated methods, since this study used a mix of manual and computer-based techniques. Ideally, both parts would be combined into one complete technology system, and that is specifically what this research intends to achieve.

Finally, there is an important consideration regarding the extraction of data from Google. This process, commonly known as data scraping, is a widely used and legally recognized method for collecting information that is already publicly accessible on the internet. While the technical act of scraping public data is legal, the key issue lies in how the collected data is used. In this case, the purpose of scraping is strictly limited to supporting business research and informing future planning efforts. The data will not be used for commercial resale, marketing, or any activity that could potentially harm individuals' privacy. Furthermore, any personal or sensitive information that might be gathered unintentionally have been carefully removed to ensure full compliance with ethical standards and data protection laws [9]. Table I shows a summary of the relevant literature in the past 8 years.

TABLE I. LITERATURE REVIEW SUMMARY

Author	Year	Main Topic	Relevance to This Research
Bartolomeo	2024	Effectiveness of BERTopic in topic extraction	Provides evidence for the accuracy and usefulness of BERTopic in extracting meaningful topics.
Lechaudé et al.	2024	BERTopic in the online retail industry	Demonstrates BERTopic application in understanding customer behavior in other industries.
Is Web Scraping Legal?	2024	Legal and ethical considerations of data scraping	Ensures the research complies with legal and ethical standards in data collection and usage.
Kim & Lee	2023	Impact of online reviews on consumer decisions [10]	Supports the use of online reviews as a reliable source for understanding customer perceptions.
Krishnan	2023	Comparison of topic modeling techniques	Validates the choice of BERTopic for capturing context in customer reviews.
Lechaudé et al.	2021	BERTopic for business insights from reviews	Shows BERTopic ability to generate actionable business insights from online review data.
Guo et al.	2020	Influence of positive/negative reviews on purchasing	Highlights the importance of review sentiment in shaping business outcomes.
Mcadoo	2020	Strategies for coffee shop success	Demonstrates the value of customer feedback and market analysis for business planning in the coffee sector.
Xu et al.	2017	Challenges in analyzing user-generated content	Justifies the need for advanced text analysis methods like topic modeling for unstructured review data.
Calheiros et al.	2017	Limitations in topic modeling of reviews	Informs methodological improvements and the need for automation in this research.



III. METHODOLOGY

A. Scope

This qualitative research is limited to specialty coffee shops in Toronto, Canada and specifically to a well-defined polygon representing the area where most of the coffee shops are in terms of density of population. This is relevant to guarantee access to a high number of potential customers' reviews.

B. Data Source and Timeframe

The data source is secondary data retrieved from customers' online reviews in Google, which provides unsolicited opinions and feedback from January 2024 to May 2025. Reviews were randomly selected.

C. Data Extraction

This research collected data using the Apify Google Maps Reviews Scraper [11], filtering by location and type of business. The relevant data points are the reviews text transcript, publishing date, and star rating. This data has been extracted in the form of CSV (Comma-separated Values) files.

D. Data Processing

After data extraction, text cleaning and pre-processing have been conducted using Python. Using the CSV files exported from Apify, the data loaded to a Google Colab notebook. General text processing techniques applied to prepare the text transcripts for the topic extraction.

E. Topic Modeling

After cleaning up the data, the topics within the reviews are identified by the BERTopic model—a lightweight and easy-to-interpret model. Topic modeling is a method used in language analysis to find and group common subjects in a large set of texts. Instead of reading each text one by one, topic modeling helps to automatically discover patterns in the way words are used. This makes it easier to understand what customers are talking about, especially when dealing with a big amount of written content like online reviews. This process saves time and gives a more organized view of the main ideas present in the text.

BERTopic is a state-of-the-art tool for topic modeling that uses machine learning and language understanding to find clearer and more useful topics. It is built on advanced language models that can understand the meaning of words based on the context they appear in. BERTopic groups similar texts together and gives a list of important keywords to describe each group. Unlike older topic modeling methods, BERTopic works better with short or casual texts and gives easy-to-understand visual summaries. This makes it a helpful choice for researchers who need to study and explain large text collections [12].

IV. ANALYSIS

With the topics identified, an insights generation process has been conducted. This process correlates the identified topics with the star review to segment the positive and negative issues.

To achieve the objective of this research, a comprehensive list of opportunities extracted from the insights presented as a set of improvement or differentiation opportunities to be used in any business plan.

Figure 1 is a general schema of the methodology proposed for this research:

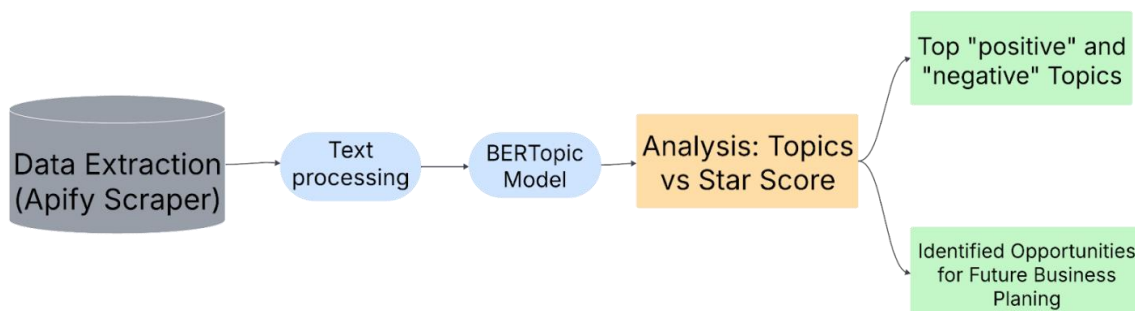


Figure 1. Proposed Research Methodology



V. ETHICAL CONSIDERATIONS

As mentioned in the literature review, the only consideration is the legal concerns of scraping data from the internet. Since this research is intended to provide some information for future business planning, and all personally identifiable data has been removed, the proposed methodology does not involve any unethical practices.

VI. RESULTS

The data extraction process was performed using the Scrapper mentioned in the methodology. This research extracted more than 9000 reviews to analyze.



Figure 2. Coffee Shop Location for Google Maps Scrapped Reviews

The raw data needed for topic modelling is the name of the coffee shop (title), the star rating, and the text transcript of the comments. Additionally, the comment's publish date was extracted to have the ability to track the topics in time. The map in Figure 2 shows the location of the coffee shops with Scrapped comments and in Table II, there is an example of 5 random rows of raw data. To ensure confidentiality, the actual names of the coffee shops are not disclosed on any tables.

TABLE II. RAW DATA EXAMPLE

title	publishedAtDate	stars	text
Coffee shop A	2025-03-06	5	Wonderful new coffee place. Exceptional drinks...
Coffee shop B	2024-11-06	4	Great service, loved the vegan bagel
Coffee shop C	2025-05-28	1	Worst place to be... they have no idea what co...
Coffee shop D	2025-04-04	5	The barista is nice to talk with. Love their c...
Coffee shop E	2025-01-11	5	oh this place was just lovely i'll be honest i...

The next step is cleaning the data. The extraction results included two coffee shops that are not particularly in the specialty coffee shop industry, and they are big coffee chains. For this research, the chain coffee shops were not considered, and the focus was only on small or medium-size coffee shops. After excluding them the data set contained 202 different coffee shops with an average of 35 comments per shop. On the other hand, in the Scrapping process filtering for only coffee shops category businesses was used, but in the resulting data there are reviews for other types of businesses like bagel shops; these are non-valuable reviews to filter out. After removing those reviews, the data obtained focuses on 7226 coffee shops reviews.



TABLE III. LESS THAN 10 CHARACTERS COMMENTS EXAMPLE

	title	publishedAtDate	stars	text	text_length
	Coffee shop M	2025-04-21	5	Amazing 🍷	9
	Coffee shop N	2024-12-12	5	Good	4
	Coffee shop O	2024-06-29	5	Excellent	9
	Coffee shop U	2024-12-15	4	Nice	4
	Coffee shop Y	2024-07-07	5	Excellent	9

The length of the comments is highly relevant for this analysis, too short comments have a lack of context to interpret a result, for example in Table 3, there are some comments with less than 10 characters, those comments should be removed from the analysis. After some testing, the conclusion is that at least 30 characters are needed to get meaningful comments, as shown in Table IV. Following the data cleaning process, a total of 6,457 reviews remained.

TABLE IV. EXAMPLE OF COMMENTS BETWEEN 30 AND 35 CHARACTERS LONG

	title	publishedAtDate	stars	text	text_length
	Coffee shop S	2024-04-06	5	Best macchiato I've had in years	32
	Coffee shop T	2025-03-23	5	I was lucky. I was with a hot date	34
	Coffee shop V	2025-05-11	5	Colombian and Flat White... Wooww..	33
	Coffee shop W	2024-05-11	1	Dry bread, drink was liquidated.	32
	Coffee shop Z	2024-12-12	5	Great coffee and friendly baristas	34

There is also an imbalance in the number of positive reviews (4 and 5 stars) compared to the negative (1 and 2). Figure 3 shows the proportion of reviews by number of stars. To avoid any bias in the model, it is necessary to split the dataset in two, one for the positive and the other for the negative reviews and then run the topic model on top of each dataset, to mitigate the effect of the imbalance.

Proportion of Reviews by Star Rating

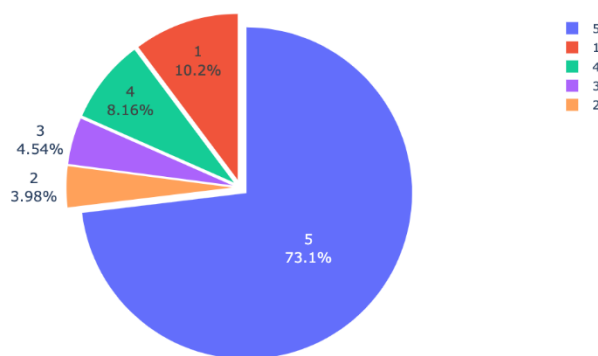


Figure 3. Reviews Volume Distribution by Star Score

VII. TOPIC MODELLING

One advantage of using BERTopic model is that there is no need to do too much data preprocessing; the model already includes some of the most common text transformations to start inferring the results. This model transforms every comment and then uses the context to represent the meaning of the words and creates a topic using the top meaningful words.



In Figure 4, there is an example of some of the topics detected by the model for positive reviews. In this context of customers reviews, a topic is a group of comments that talk about the same concept even if they use different words.

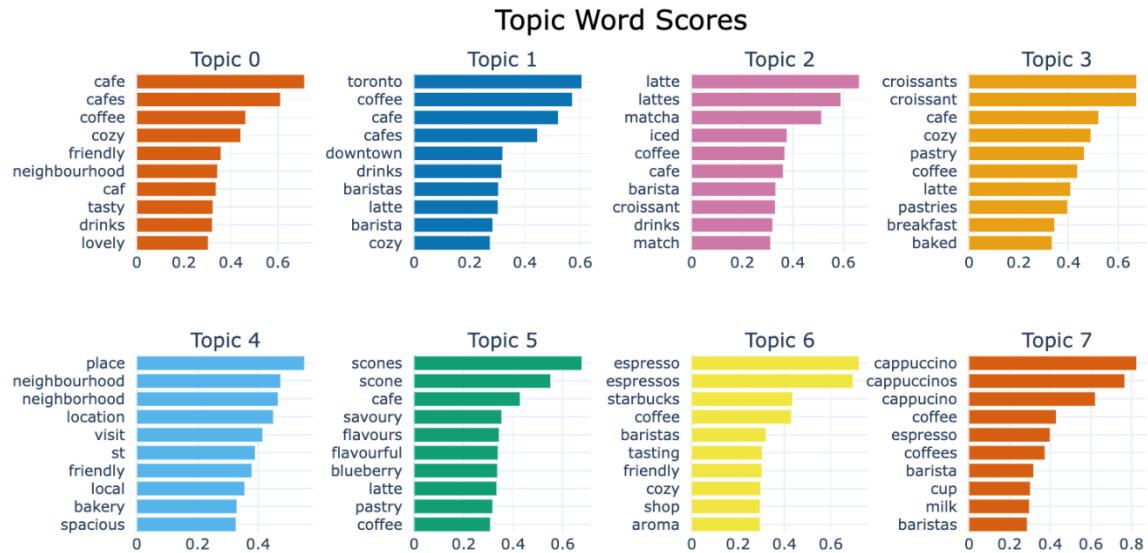


Figure 4. Top 8 topics and word value

All the comments get an assigned topic because the model creates different clusters to group every transcript. Visually this looks like every document is a dot in a two-dimensional space, and then depending on the context, each dot is assigned to a particular group or topic as shown in Figure 5.

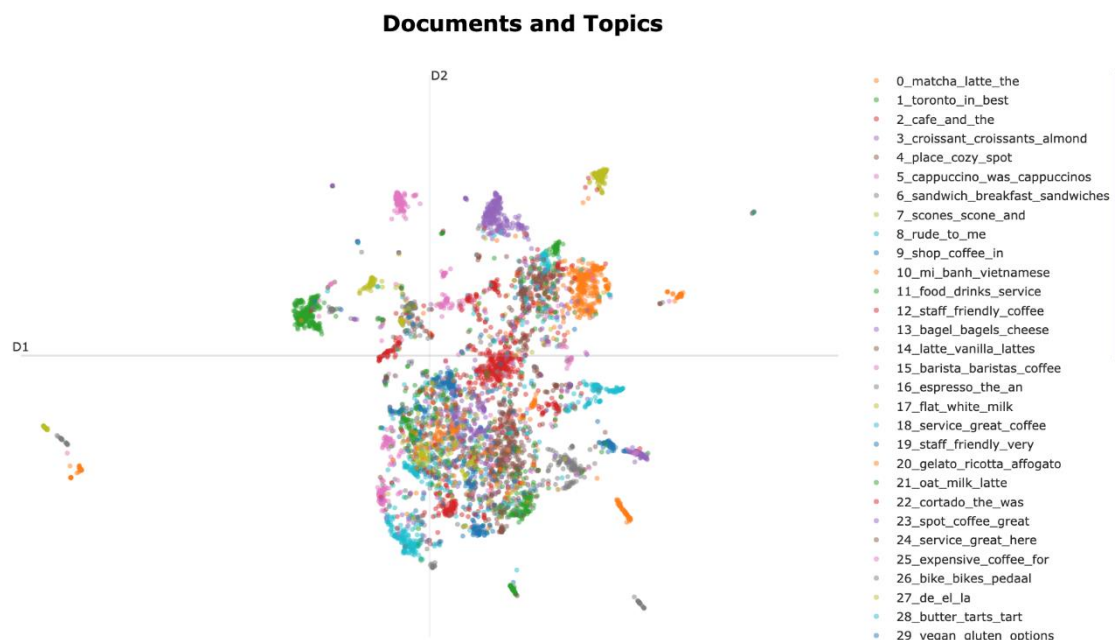


Figure 5. Topic assignment for positive reviews

VIII. TOPICS AND STARS SCORE

A. Positive Reviews

The topic modeling process produces a consolidated table, with each conversation assigned a topic and accompanied by three different interpretations of it. These interpretations are only a set of the most meaningful words within a topic. For example one of the topics is “*cafe_and_the_this*” which is not very easy to interpret, in order to gain more insightful



information from this topic, it is possible to include additional interpretations to the trained model to get “cafe,coffee,cozy,visit,lovely,neighbourhood,patio,drinks,decent,vibes” which is a way more easy to understand.

This research considered the approach defined by Elkouri [13] where the positive reviews correspond to scores equal to 4 and 5 and negative to 1 and 2. Starting with positive comments, the top topics are listed in Table V. Finally, a prompt was designed to interpret the topics and get more actionable insights. The prompt was then added to an OpenAI module to get an interpretation. This is the prompt used:

You are an expert interpreter on topic modelling from BERTopic results. Give insights from these results using the Aspect1 column and the Representative_Docs. These topics are the top from positive customer online reviews for coffee shops.

TABLE V. TOP FIVE TOPICS FOR POSITIVE REVIEWS

Representation	% of total positive reviews
['cafe', 'cafes', 'coffee', 'cozy', 'neighbourhood', 'drinks', 'patio', 'welcoming', 'vibes', 'interior']	7.76%
['toronto', 'coffee', 'cafe', 'downtown', 'baristas', 'latte', 'cozy', 'visit', 'places', 'recommend']	7.26%
['lattes', 'matcha', 'iced', 'cafe', 'croissant', 'tasty', 'chai', 'vanilla', 'bowls', 'ordered']	6.40%
['croissants', 'cafe', 'cozy', 'pastry', 'latte', 'breakfast', 'pie', 'almond', 'staff', 'enjoyed']	4.78%
['neighborhood', 'visit', 'st', 'bakery', 'spacious', 'welcoming', 'parking', 'spots', 'lunch', 'washrooms']	4.70%
['scones', 'cafe', 'savory', 'flavourful', 'blueberry', 'latte', 'pastry', 'caramel', 'basque', 'bacon']	3.36%
['espressos', 'starbucks', 'coffee', 'baristas', 'tasting', 'cozy', 'customers', 'wonderful', 'cups', 'welcoming']	3.19%
['baristas', 'coffee', 'starbucks', 'friendly', 'incredible', 'review', 'met', 'bar', 'serve', 'pastries']	3.06%
['cappuccino', 'espresso', 'coffees', 'barista', 'muffin', 'roast', 'delicious', 'croissants', 'raspberry', 'cozy']	2.94%
['shops', 'downtown', 'cocoa', 'enjoy', 'corporate', 'welcoming', 'university', 'drinks', 'interior', 'patio']	2.66%

Basically, what was asked the GPT model was to use the results from BERTopic and then create an interpretation using the most representative comment from each of the top five topics. Table VI is the result of the interpretations made using the prompt.

TABLE VI. INTERPRETATION RESULTS FOR POSITIVE REVIEWS TOPICS

Main Theme	Key Aspects (Aspect1 & 2)	Customer Focus	Representative Insights
Cozy Neighborhood Cafés	Cozy, cute atmosphere; seating; friendly service; welcoming vibe; interior	Ambience, comfort, and friendly staff	Customers love the cozy, welcoming vibe and return for the relaxing atmosphere.
Best Coffee in Toronto	Best coffee; favorite shop; downtown; baristas; recommend	High-quality coffee and strong reputation	Cafés described as “the best” in Toronto draw loyal and enthusiastic customers.
Matcha & Specialty Lattes	Matcha, lattes, iced drinks, unique flavors (strawberry, coconut, vanilla)	Creative, tasty, and Instagram-worthy drink options	Matcha and flavored lattes are standout items that enhance customer delight.
Croissants & Pastries	Croissants (almond, pistachio), flaky, buttery, delicious; pastries with latte	Quality baked goods paired with drinks	Delicious croissants elevate the café experience and are frequently praised.
Social & Community Hangout Spots	Lovely place; great people; food and drinks; spacious; location; work-friendly; amenities	Socializing, working, and relaxing in a comfortable setting	Customers appreciate cafés as great places to work, chat, and feel welcome.



B. Negative Reviews

For the negative reviews the process is the same, the result is shown in the following tables and in general the insight describes that negative reviews center mostly around bad customer service, with customers feeling ignored and disrespected by staff. Another strong negative aspect identified is poor value for money, where expensive drinks don't meet expectations in terms of taste, preparation, or portion size. Customers are also frustrated when shops are closed outside posted hours, which hurts trust and convenience and when they are not allowed to use their laptops inside the coffee shop. Overall, the most damaging experiences involve both interpersonal issues and inconsistency in product or service.

Following the same logic, Table VII shows the top topics for negative reviews with their word representation and the percentage of the total of reviews. Table VIII includes the list of interpretations given by the prompt for negative topics.

TABLE VII. TOPICS FOR NEGATIVE REVIEWS

Words Representation	% of total negative reviews
['cafe', 'coffee', 'customers', 'baristas', 'orders', 'rude', 'shop', 'terrible', 'staff', 'employee']	50.00%
['latte', 'lattes', 'coffee', 'iced', 'espresso', 'syrup', 'barista', 'vanilla', 'pistachio', 'flavour']	13.89%
['latte', 'matcha', 'coffee', 'iced', 'strawberry', 'drinks', 'barista', 'flavour', 'blueberry', 'ordered']	7.55%
['breakfast', 'sandwiches', 'bread', 'pistachio', 'lattes', 'bland', 'tasteless', 'eggs', 'cookie', 'rice']	6.46%
['espresso', 'coffee', 'tasting', 'flavourless', 'baristas', 'cortado', 'milk', 'vanilla', 'roasted', 'drip']	5.36%
['cappuccino', 'coffee', 'cafe', 'espresso', 'barista', 'milk', 'pistachio', 'tasted', 'cream', 'sour']	4.05%
['croissant', 'croissants', 'latte', 'baked', 'coffee', 'stale', 'meal', 'almond', 'warmed', 'impatient']	4.27%
['laptop', 'cafe', 'wifi', 'ipad', 'customers', 'sit', 'weird', 'policy', 'outlet', 'outside']	3.61%
['closing', 'closed', 'opening', '5pm', 'sunday', 'close', 'downtown', 'morning', 'restaurant', 'checked']	3.17%
['cafe', 'service', 'cappuccino', 'bar', 'clients', 'ordered', '20', 'drinks']	1.64%

TABLE VIII. INTERPRETATION RESULTS FOR POSITIVE REVIEWS TOPICS

Main Theme	Key Aspects (Aspect1 & 2)	Customer Focus	Representative Insights
Rude Staff & Poor Customer Service	<i>coffee, rude, service, order, customer, cafe, customers, baristas, shop, terrible</i>	How customers were treated	Customers felt disrespected and unwelcome due to staff attitude (e.g., rude tone, lack of engagement). Service was described as careless and cold, even when the atmosphere was good.
Bad Quality Lattes and Pricing Issues	<i>latte, iced, milk, taste, vanilla lattes, syrup, pistachio, flavour, espresso</i>	Drink quality and value for money	Lattes were described as bland, watered down, or made without proper ingredients. Overpriced drinks with confusing portion sizes frustrated customers.
Poor Matcha Experience	<i>matcha, latte, milk, powder, quality, iced, strawberry, blueberry, drinks</i>	Expectations around specialty drinks	Matcha lattes were gritty, weak, or improperly mixed. Customers expected ceremonial-grade quality but received low-effort drinks that felt overpriced.
Disappointing Food	<i>sandwich, breakfast, egg, dry, food, bland,</i>	Food preparation and consistency	Breakfast items were dry, unevenly cooked, or disappointing in taste. Some customers



(Sandwiches & Breakfast)	<i>pistachio, tasteless, cookie, rice</i>		experienced issues with availability and order accuracy.
Weak Espresso & Underwhelming Coffee Flavor	<i>espresso, beans, shot, taste, cup, cortado, milk, flavourless, roasted, drip</i>	Core coffee quality and preparation	Coffee was weak, watery, or sour. Espresso shots lacked strength and proper extraction. Customers expected better from specialty cafés.
Cappuccino Quality & Order Mistakes	<i>cappuccino, milk, cream, cup, bad barista, tasted, sour, espresso</i>	Accuracy and execution of drinks	Issues included incorrect milk used, cold cappuccinos, or too much foam. Attempts to fix mistakes were poorly handled by staff.
Stale & Incorrect Croissants	<i>croissant, fresh, stale, almond, dry latte, baked, impatient, warmed</i>	Food freshness and order accuracy	Croissants were not fresh or wrong items were served. Staff responded dismissively or with attitude when asked to correct the mistake.
No-Laptop Policy & Misleading Workspaces	<i>laptop, wifi, policy, space, rules, seating, outlet, ipad, customers, shop</i>	Work-friendly environment and café policies	Customers expected to use laptops but were surprised by restrictive policies. Felt misled by listings that suggested work-friendly spaces.
Incorrect or Inconsistent Opening Hours	<i>closed, open, hours, google, location, door, opening, sign, checked, downtown</i>	Reliability of store operations	Cafés were closed during listed open hours or didn't match posted schedules. This caused inconvenience and wasted time for customers trying to visit.

IX. DISCUSSION, KEY SUCCESS FACTORS, AND LIMITATION

Using this interpretation, a possible conclusion from the positive reviews is that customers consistently value coffee shops that offer more than just a good cup of coffee. What really matters to them is a welcoming, cozy environment, friendly staff, and high-quality food and drinks. Whether it's the charm of a neighborhood patio, the warmth of the service, or the consistency of a great cup of coffee, the overall experience is what drives positive feedback and loyalty. In other words, a coffee shop that wants to enter the market to compete needs to focus on a whole coffee experience including the service, space and compliments rather than just a good cup of coffee.

Based on this analysis, key success factor areas for generating value in both the planning of new coffee shops and the improvement of existing ones should include:

- Comfort and Ambience: Cozy place that generates good vibes and comfort
- Friendly Staff: Guarantee the best possible service choosing the right staff and baristas
- Coffee Quality: Focus on differentiation by quality and not price
- Complementary Goods: Baking goods should follow the same high quality of the coffee
- Creativity: Implement temporary different and creative coffee drinks with new flavors

On the negative side, as mentioned, the improvement points are around the quality of the products, service and experience as an integral part of the business. Customers want to have a good time in the coffee shop and not just have a grab-and-go experience. This segment of the market values quality over quantity and even over fast delivery of the product. The main pain points to avoid for future and current entrepreneurs in this industry are:

- Service: Baristas are facing the customer; good barista skills are not the only requirement for the staff.
- Quality vs Price: the problem is not to charge a higher market price, the goal is to support the value with quality
- Meet expectations: Communicate the expectations of each product properly. If you promise the best coffee in town, it should be.
- Laptop and Wi-Fi policy: Find solutions to long-staying customers but don't prohibit the use of laptops.

The limitations of this research primarily stem from the data source. Focusing on Google reviews may introduce bias due to the restriction of the sample to only Google Maps active users.

X. CONCLUSION

The research clearly shows that both good and bad customer experiences in specialty coffee shops go beyond just the taste of the coffee. Positive experiences are related to different factors like friendly staff, relaxing spaces, quality drinks, and creative food options. Coffee shops that offer these factors together are more likely to build strong customer loyalty.



New business owners should focus on delivering full experiences, not just products, to stand out in a competitive market like Toronto's.

On the other hand, the most common reasons for negative reviews include poor service, drinks that don't match their price, and unclear shop policies. These issues make customers feel frustrated and ignored, and they damage trust. A good coffee shop must avoid these problems by training staff well, making sure the quality matches the price, and keeping communication clear. Overall, listening to what customers say online is a valuable and low-cost way to learn, improve, and succeed in the specialty coffee business.

XI. RECOMMENDATIONS

Future studies should take a closer look at the clear imbalance between positive and negative reviews found in this research. Since most of the feedback on Google Reviews is positive, it might hide problems or areas that need improvement which people mention less often. Looking into why this happens, like possible platform biases or common customer habits, could help businesses better understand customer satisfaction and spot hidden issues.

The analysis method used in this study can also be useful outside of the specialty coffee market. By using automated topic modeling and sentiment analysis, this approach can be applied to customer feedback in different industries to find helpful insights. Companies can use it to guide business decisions, improve their products, and create better customer experience based on actual feedback.

One of the main benefits of this method is that it's highly automated, so it doesn't need much manual work. This makes the process faster and easier to manage, and it can be used as a long-term tool to keep track of market trends and support decision-making. As more businesses turn to data to stay competitive and meet customer needs, tools like this will become even more important.

REFERENCES

- [1]. "Specialty Coffee Market Size, Share & Growth Report, 2030," Grand View Research, GVR-4-68040-060-1, 2030.
- [2]. J. Ferreira, C. Ferreira, and E. Bos, "Spaces of consumption, connection, and community: Exploring the role of the coffee shop in urban lives," *Geoforum*, vol. 119, pp. 21–29, Feb. 2021, doi: 10.1016/j.geoforum.2020.12.024.
- [3]. V. Swar, "Single-Origin Coffee Market Size, Share and Trends 2025 to 2034," Precedence Research, 6106, May 2025. Accessed: Jul. 01, 2025. [Online]. Available: <https://www.precedenceresearch.com/single-origin-coffee-market>
- [4]. J. Guo, X. Wang, and Y. Wu, "Positive emotion bias: Role of emotional content from online customer reviews in purchase decisions," *J. Retail. Consum. Serv.*, vol. 52, p. 101891, Jan. 2020, doi: 10.1016/j.jretconser.2019.101891.
- [5]. J. H. Mcadoo, "DOCTOR OF BUSINESS ADMINISTRATION".
- [6]. A. Krishnan, "Exploring the Power of Topic Modeling Techniques in Analyzing Customer Reviews: A Comparative Analysis," Aug. 19, 2023, arXiv: arXiv:2308.11520. doi: 10.48550/arXiv.2308.11520.
- [7]. G. Bartolomeo, "Revocable Anonymous Credentials from Attribute-Based Encryption," Dec. 12, 2024, arXiv: arXiv:2308.06797. doi: 10.48550/arXiv.2308.06797.
- [8]. R. M. P. Lechaude, "Decoding Customer Feedback: Employing Topic Modelling (BERTopic) to Understand Product Return Drivers from Online Reviews".
- [9]. A. C. Calheiros, S. Moro, and P. Rita, "Sentiment Classification of Consumer-Generated Online Reviews Using Topic Modeling," *J. Hosp. Mark. Manag.*, vol. 26, no. 7, Art. no. 7, Oct. 2017, doi: 10.1080/19368623.2017.1310075.
- [10]. J. M. Kim and E. Lee, "The effect of perceived threat on online service reviews," *J. Serv. Mark.*, vol. 37, no. 3, Art. no. 3, Feb. 2023, doi: 10.1108/JSM-06-2021-0220.
- [11]. Compass, "Google Maps Scraper · Apify," Apify. Accessed: May 23, 2025. [Online]. Available: <https://apify.com/compass/crawler-google-places>
- [12]. M. Grootendorst, "BERTopic: Neural topic modeling with a class-based TF-IDF procedure," 2022, arXiv. doi: 10.48550/ARXIV.2203.05794.
- [13]. A. Elkouri, "Predicting the Sentiment Polarity and Rating of Yelp Reviews," 2015, arXiv. doi: 10.48550/ARXIV.1512.06303.