



UNIVERSITY OF
TORONTO
MISSISSAUGA

CCT416 Final Research Project - Mapping Public Sentiment On Canada's Housing

AUTHORS

Abdul (1007984736) & Uzair (1007958337)

Instructor: Mircea Gherghina



April 2, 2024

Abstract

The escalating housing costs in Canada have sparked widespread concern, impacting both homeowners and renters alike. This study investigates public sentiment regarding housing costs in Canada by analyzing social media data from two prominent Reddit communities: r/CanadaHousing and r/Canada. The research question guiding this study is: "What is the opinion of social media users regarding housing costs in Canada?" To address this question, sentiment analysis techniques were applied to categorize posts into positive, negative, or neutral sentiments, while topic modelling was used to identify common themes within the discussions. The findings reveal a trending sense of dissatisfaction and concern among social media users regarding the challenges of housing affordability in Canada. Negative sentiments dominate the discussions, with common themes including pricing, government policies, and regional disparities. These findings align with existing literature on housing affordability challenges in Canada and provide meaningful observations that can inform policy decisions and public ideas on this critical issue.

Introduction

The escalating housing costs in Canada have become a topic of national concern. With home prices up by more than 35% in just four years, the affordability of housing is a significant issue affecting a large portion of the population (Armstrong, 2024). This issue is not only confined to homeowners but also impacts renters, with rental prices continually hitting record highs. Given the widespread impact of this issue, it is crucial to understand the public sentiment towards housing costs in Canada. Social media platforms, where people freely express their opinions and concerns, provide a rich source of data for this analysis. This motivated the selection of the case for this study.

The research question for this study is: "What is the opinion of social media users regarding housing costs in Canada?" This question aims to contribute to the ongoing debates about housing affordability in Canada by providing insights into public sentiment derived from social media data. The study will analyze social media posts related to housing costs in Canada, using sentiment analysis techniques to categorize the posts into positive, negative, or neutral sentiments. The findings from this research could potentially inform policy decisions and stimulate further academic discussions on this topic.

Literature Review

The existing literature pertinent to the research question concerning social media users' perspectives on housing costs in Canada offers a nuanced understanding of the multifaceted challenges surrounding housing affordability and its societal ramifications. Farhan's (2023) examination of the Canadian housing affordability crisis illuminates the intricate interplay between economic, political, and social factors, showcasing how both government and bankers benefit from the existing shortage of residential property, impeding effective solutions to the housing dilemma. Farhan (2023) also found that young Canadians now face the dilemma of renting or owning a home, and thus suggests that Canadian leadership should facilitate the supply of residential property by removing barriers and restrictions and offering affordable land for developers. This aligns with the findings of Ramage et al. (2021), who delve into the experiences of tenants in affordable housing units in Calgary, emphasizing that there is a need for more holistic and

tailored interventions that address the economic, social, and health inequities of affordable housing tenants, as well as more robust and participatory policy-making that involves the voices of those with lived experience. Their work underscores the importance of considering broader economic and social disparities in crafting housing policies.

Kipfer and Sotomayor's (2024) critique of market housing solutionism highlights systemic capitalism issues that worsen housing problems, with homelessness rates doubling since the 1980s and public housing waiting lists growing. Many households spend a large income portion on housing, affecting tenants disproportionately. These issues are rooted in capitalism's systemic problems, where profit motives often prioritize financial gain over housing stability. While solutions often call for urban planning reforms, it's crucial to recognize capitalism's hidden contradictions that worsen these challenges, leading to increased social crises and housing access inequalities. Kipfer and Sotomayor emphasize the need for comprehensive solutions that prioritize housing stability, aligning with the research question on social media users' opinions on Canadian housing costs. Thus, stressing the importance of addressing housing insecurity and inequality.

Moreover, Revington and August (2019) delve into the concept of financialization, revealing how the commodification of real estate assets prioritizes investor profits over tenant well-being, thereby exacerbating inequalities in access to stable housing. Revington and August (2019) emphasize how finance has come to dominate various aspects of economies, firms, states, and households, leading to a structural transformation. Within this context, real estate assets have become subject to financial logics and practices, with properties being treated primarily as financial assets rather than as homes or community spaces.

Furthermore, Smith's (1983) historical analysis offers valuable context by tracing the evolution of Canada's housing landscape, from a historically well-housed state to contemporary challenges exacerbated by policy shifts and economic restructuring. Smith (1983) provides a detailed account of the transition from a thriving private rental sector to government-supported homeownership and public rental sectors, and the restructuring of the tax system in 1972, both of which have had long-term implications on housing affordability.

Interestingly enough, Garcia and Chen (2022) investigate how social media platforms influence the perceptions and attitudes of Canadian millennials towards housing affordability. Through analyzing discussions, opinions, and narratives on platforms like Twitter and Reddit, they provide insights into how social media discourse reflects and shapes public sentiment on housing costs.

Wang and Li (2023) also employ advanced computational methods to explore public discourse on housing costs in Canada across various social media platforms. Using sentiment analysis and topic modelling, they uncover prevalent themes, concerns, and attitudes among social media users regarding housing affordability, highlighting the diverse range of opinions and perspectives within the online community.

Finally, Nguyen and Patel (2024) examine the role of social media as a tool for advocacy and mobilization around housing affordability in Canada. Through qualitative interviews and quantitative analysis of social media engagement metrics, they assess the effectiveness of social media campaigns in

raising awareness, influencing policy discourse, and mobilizing public support for affordable housing initiatives.

In conclusion, our research question, “What is the opinion of social media users regarding housing costs in Canada?” extends existing literature by incorporating public sentiment perspectives, which has been largely unexplored. The works of Farhan (2023), Ramage et al. (2021), Kipfer and Sotomayor (2024), Revington and August (2019), and Smith (1983) provide a comprehensive understanding of the multifaceted challenges surrounding housing affordability in Canada. However, they primarily focus on the perspectives of policymakers, economists, and tenants. By exploring the views of social media users, this research can uncover new dimensions of housing affordability, potentially revealing how public sentiment aligns and diverges from the findings of these studies. This approach doesn’t refute existing findings, but enriches them, offering a more holistic understanding of housing affordability in Canada.

Data and Methods

Data Collection

The data for this analysis was collected from two subreddits on Reddit: [CanadaHousing](#) and [Canada](#). The data was collected using the Python Reddit API Wrapper (PRAW), which allows for easy access to Reddit’s API for data scraping purposes. Additionally, the data collection process involved fetching the top 100 posts from each subreddit. The data was collected on March 28-30, ensuring that the analysis is based on the most recent and relevant posts. **NOTE:** All Code Used In Methods 1 & 2 Are Provided In Separate Jupyter Submission Files (Without Our Reddit API Keys Of Course).

Method 1: Sentiment Analysis

Summary

For this analysis, we looked at 2 subreddits and collected multiple posts and their titles in order to analyze their overall sentiment. We first collected the first 100 posts from r/canadahousing and turned them into a dataset with categories including post title, score, content, number of comments, URL, and date. We then generated a word cloud (Figure 1) to get an understanding of key terms and phrases that are evident across the titles of these posts, and tried to interpret them.

After this, we used a sentiment intensity analyzer on the dataframe, creating a new column that categorizes the titles based on their score. We then plotted these results on a scatterplot (Figure 2), using appropriate colours to represent the sentiment classification. Red for a negative classification, blue for a neutral one, and green for a positive one. We decided to graph the score and the number of comments, as posts with more of each represent a more active post, so the sentiment classification of that becomes more significant.

To remove any sort of bias or inaccuracy, we analyzed another community as well, r/Canada, which has over 2.4 million members. We did the same thing, creating a dataframe with all the same topics, but this time we filtered the posts to only include the ones that talked about housing and also increased the post

Method 2: Topic Modeling

Summary

For this method, we focused on the same two subreddits, r/CanadaHousing and r/Canada, and collected a significant number of posts to perform topic modelling. We initially collected the top 100 posts within the year from r/CanadaHousing, but had to increase to 1000 for r/Canada since the initial number of posts did not provide enough data for a comprehensive topic modelling analysis. Additionally, for r/CanadaHousing, we didn't have to do any filtering since the whole subreddit is dedicated to Canada's housing, but we had to filter the posts in r/Canada for those that contained the keyword 'housing' in the title or content. The variables included in our analysis for both subreddits were the post's title, score, content, number of comments, post URL, and full URL.

The first step in our analysis was to preprocess the text content of the posts. We tokenized the text, converted tokens to lowercase, removed non-alphanumeric characters, removed stopwords, and stemmed the tokens using the Porter Stemmer. This preprocessing step was crucial to reduce the dimensionality of our data and to focus on the most meaningful words.

Next, we used the Latent Dirichlet Allocation (LDA) model for topic modeling. We chose LDA because it is a widely used topic modeling technique that can effectively identify hidden topics in large volumes of text. We set the number of topics to 10 and created a dictionary and a document-term matrix from the preprocessed text. We then ran the LDA model on this corpus.

To visualize the topics identified by the LDA model, we created bar plots of the top words for each topic and their probabilities (Figure 5 and Figure 9). We also calculated and plotted the average topic proportions for each topic (Figure 6 and Figure 10). These visualizations helped us understand the most prevalent topics and their relevance in the dataset.

In addition, we identified and printed the top 2 posts associated with topic 5 for r/CanadaHousing and topic 3 for r/Canada, as they had the highest average proportion in comparison to the other topics. This allowed us to see which posts were most representative of the most prevalent topic, and what exactly they discussed. We found this particularly useful for understanding the context and content of the topics. However, it's important to note that for r/CanadaHousing, we were able to print the 2 posts without problems (Figure 7), but we had to print the full URL for r/Canada instead, and then go to the posts directly (Figure 11), as the 2 top posts weren't texts, but rather links to articles.

Finally, we used 'pyLDAvis' to create an interactive visualization of the topics (Figure 8 and Figure 12). This provided a more intuitive understanding of the topics and their relationships with each other.

Overall, the topic modeling method allowed us to uncover the underlying topics in the Reddit posts and understand their relevance and relationships. This method was particularly useful for analyzing large volumes of text data and provided valuable insights into the discussions happening in the r/CanadaHousing and r/Canada subreddits.

Produced Graphs For r/Canada Subreddit

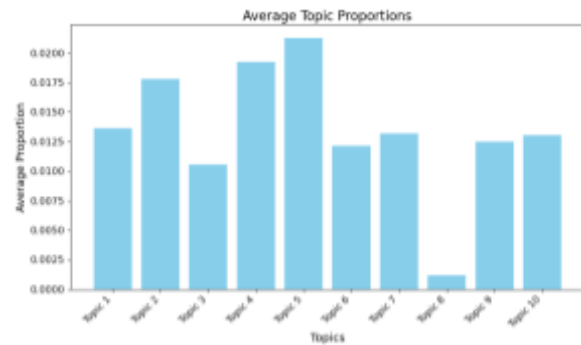


Figure 6. Average Proportions of Topics From LDA Analysis For r/CanadaHousing

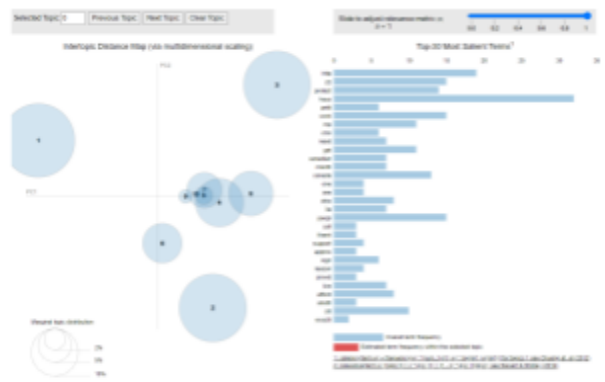


Figure 8. Interactive Visualization of Topic Model From LDA Analysis For nCanadaHousing

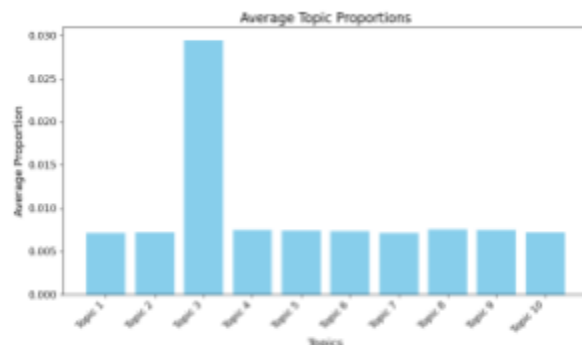
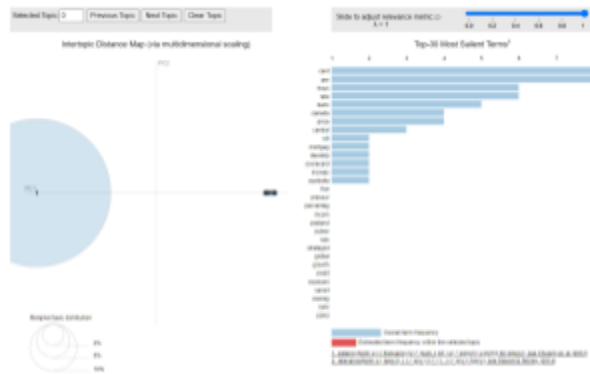


Figure 10. Average Proportions of Topics From LDA Analysis For r/Canada



Results

Method 1: Sentiment Analysis

Analyzing Graphs For r/CanadaHousing Subreddit

The word cloud (Figure 1) from r/CanadaHousing prominently displayed words related to housing such as “housing”, “house”, and “condo”, indicating frequent discussions on these topics. The geographic terms like “Canada,” “Ontario,” and “Vancouver” also highlight the regional focus of these discussions. Other terms like “crisis”, “affordable”, and “market” suggest ongoing dialogues about housing affordability and market conditions



Analyzing the results of Figure 2, we can see that the two most dominant classifications are negative and neutral, with there not being a single positive post higher on the y-axis. In fact, the post with the most comments was negative, indicating passion and strong negative feelings from users about the topic. Even throughout the rest of the graph, posts with higher scores are generally negative, and only about 9/100 posts were positive, which is less than 10%.

Glancing over this word cloud from r/Canada (Figure 3), we can see that words like “Canada,” “housing,” and “government” are larger and more prominent, indicating they are common topics. Other visible words include “Ottawa,” “federal,” “affordable,” and “program,” among others, giving insights into sub-topics of interest. Infact, the prominence of words related to housing and government indicates a strong focus on these issues within the subreddit. The presence of words like “affordable” and “program” further suggests that

[illegible]

Relationship between Number of Comments and Score (Colored by Sentiment)

Score

Number of Comments

Positive
Negative
Neutral

Number of Comments	Score	Sentiment
0	0	Neutral
10	50	Positive
20	50	Positive
30	50	Positive
40	50	Positive
50	50	Positive
60	50	Positive
70	50	Positive
80	50	Positive
90	50	Positive
100	50	Positive
110	50	Positive
120	50	Positive
130	50	Positive
140	50	Positive
150	50	Positive
160	50	Positive
170	50	Positive
180	50	Positive
190	50	Positive
200	50	Positive
210	50	Positive
220	50	Positive
230	50	Positive
240	50	Positive
250	50	Positive
260	50	Positive
270	50	Positive
280	50	Positive
290	50	Positive
300	50	Positive
310	50	Positive
320	50	Positive
330	50	Positive
340	50	Positive
350	50	Positive
360	50	Positive
370	50	Positive
380	50	Positive
390	50	Positive
400	50	Positive
410	50	Positive
420	50	Positive
430	50	Positive
440	50	Positive
450	50	Positive
460	50	Positive
470	50	Positive
480	50	Positive
490	50	Positive
500	50	Positive
510	50	Positive
520	50	Positive
530	50	Positive
540	50	Positive
550	50	Positive
560	50	Positive
570	50	Positive
580	50	Positive
590	50	Positive
600	50	Positive
610	50	Positive
620	50	Positive
630	50	Positive
640	50	Positive
650	50	Positive
660	50	Positive
670	50	Positive
680	50	Positive
690	50	Positive
700	50	Positive
710	50	Positive
720	50	Positive
730	50	Positive
740	50	Positive
750	50	Positive
760	50	Positive
770	50	Positive
780	50	Positive
790	50	Positive
800	50	Positive
810	50	Positive
820	50	Positive
830	50	Positive
840	50	Positive
850	50	Positive
860	50	Positive
870	50	Positive
880	50	Positive
890	50	Positive
900	50	Positive
10	100	Negative
20	100	Negative
30	100	Negative
40	100	Negative
50	100	Negative
60	100	Negative
70	100	Negative
80	100	Negative
90	100	Negative
100	100	Negative
110	100	Negative
120	100	Negative
130	100	Negative
140	100	Negative
150	100	Negative
160	100	Negative
170	100	Negative
180	100	Negative
190	100	Negative
200	100	Negative
210	100	Negative
220	100	Negative
230	100	Negative
240	100	Negative
250	100	Negative
260	100	Negative
270	100	Negative
280	100	Negative
290	100	Negative
300	100	Negative
310	100	Negative
320	100	Negative
330	100	Negative
340	100	Negative
350	100	Negative
360	100	Negative
370	100	Negative
380	100	Negative
390	100	Negative
400	100	Negative
410	100	Negative
420	100	Negative
430	100	Negative
440	100	Negative
450	100	Negative
460	100	Negative
470	100	Negative
480	100	Negative
490	100	Negative
500	100	Negative
510	100	Negative
520	100	Negative
530	100	Negative
540	100	Negative
550	100	Negative
560	100	Negative
570	100	Negative
580	100	

The scatter plot (Figure 4) for the r/Canada subreddit reveals a compelling narrative. Despite an increase in the number of posts from 100 to 1000 for this subreddit, the volume of posts is lower, yet the engagement in terms of comments and scores is significantly higher, which is understandable since this subreddit has more members. This suggests that each post garners more attention and interaction. The sentiment of these posts, however, is predominantly negative, with only two exceptions that are positive. These positive outliers are found amidst the sea of negative sentiments, one at lower comment numbers and scores, and another standing alone at higher values. This pattern indicates a more negative connotation

associated with the housing discussions in the r/Canada subreddit compared to the r/CanadaHousing subreddit, suggesting that most of these Reddit members are deeply concerned about the Canadian housing issue.

Method 2: Topic Modeling

Produced Graphs For r/CanadaHousing Subreddit



Figure 5. Top Words In Each Topic From LDA Analysis For r/CanadaHousing

Figure 5 illustrates the top words associated with each of the ten topics identified in our LDA analysis of the r/CanadaHousing subreddit. The x-axis of each bar graph represents the frequency of words, while the y-axis lists the words themselves. The topics cover a diverse range of areas related to housing. For instance, Topic 2 seems to focus on pricing, with words like “price” and “market” appearing frequently. Topic 5 appears to be about locations, with place names and terms like “city” and “area” being common. Topic 8 seems to discuss types of housing, with words like “apartment”, “condo”, and “rent” being prevalent. These results provide a comprehensive overview of the discussions happening in the r/CanadaHousing subreddit, highlighting the key themes and concerns of the community members.

This bar graph, Figure 6, illustrates the average proportions of the ten topics identified in our LDA analysis of the r/CanadaHousing subreddit. The x-axis represents the topics, labeled from Topic 1 to Topic 10, while the y-axis represents the average proportions, ranging from 0.0000 to 0.0200. The graph shows that Topics 4 and 5 have the highest average proportions, indicating that they are the most prevalent topics in the subreddit. On the other hand, Topic 8 has the lowest average proportion, suggesting that it is less frequently discussed.

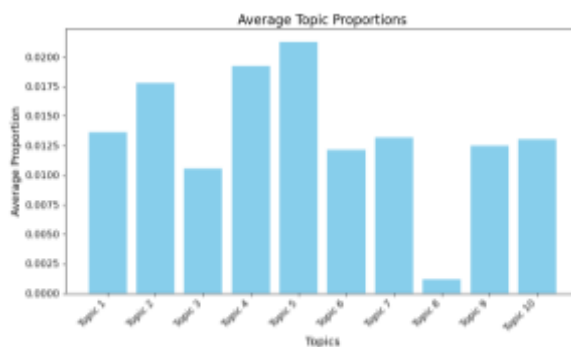


Figure 6. Average Proportions of Topics From LDA Analysis For r/CanadaHousing

Document 15 [score: 0.99189115761750]: He said he wants to make sure we can "access all the opportunities the generations before us had".
....that's never going to happen with respect to housing. Things may have gotten better on the social front (minority rights, women's rights, etc). Children benefit, dental benefit, 99 student loan interest are also all nice. But housing is totally fucked now and there is no fixing it.
It's going to become increasingly common for kids to grow up in 2000-2000 sq ft houses with like 7 family members, 3 or 4 of them having incomes.
If that kind of family support network isn't available, then kids are going to be growing up in really tight quarters because the average 3-parent household won't be able to afford a lot of space.
This is very clearly a step back from what millennials grew up with. I'm not saying it's all bad. Living with grandma/grandpa can even be better for the kids, but in terms of material prosperity it's clearly a downgrade.
Instead of making an unfulfillable promise or trying to put lipstick on a pig, I wish the government would just acknowledge this new reality. But I guess they can't and we n't because it would be too big of a failure to admit to.
Document 14 [score: 0.99108134048571]: I'm an American but both sides of my family are originally Canadian and moved to the states. My grandparents always said "America is the best for making money, Canada is the best for living" so I figured I look into seeing if I could get a Canadian passport. I haven't been to Canada since I was a kid in the 90s several days back then and it's 100% in Texas so I want to escape the heat. I got on this Reddit and I'm shocked by the amount of despair. I always thought Canadians were average but it's better than Americans. Are the housing crisis and cost of living really getting as bad as Reddit says? Does that sound all these problems?
Edit: well look get back from the video lol, there actually was a hell rider from Alberta there lol. This blew up! thank you all for taking so much time to write. The charts are crazy. I will never complain about the price of housing in Texas again! It seems that unless you are very wealthy or already own property Canada is a very hard place to live. I'm really sorry that this happened to y'all, I hope it gets fixed or it's easy for you to come here.

Figure 7. Top 2 Posts Associated With Topic 5 From LDA Analysis For r/CanadaHousing

Taking a look at topic 5 since it had the highest average proportion, the two posts shown in Figure 7 are the most representative of Topic 5 in our LDA analysis of the r/CanadaHousing subreddit. These posts, which scored 0.99189 and 0.99108 respectively, provide

valuable insights into the discussions that dominate this topic. The first post (Document 15) discusses the future of housing, expressing concerns about the increasing commonality of larger families living in confined spaces due to unaffordable housing prices. The user criticizes the government for making unfulfillable promises and not acknowledging the severity of the housing crisis. This post provides a

personal perspective on the housing situation, highlighting the challenges many individuals face due to the high cost of housing. The second post (Document 14) shares an anecdote about an American with family ties to Canada. The user contrasts the living conditions between the two countries, implying that he now sees Canada's housing crisis in a more negative light than America, which also has a housing problem. These posts provide a snapshot of the discussions happening in the r/CanadaHousing subreddit, particularly around Topic 5. They highlight the concerns and experiences of community members, offering a deeper understanding of the context and content of this topic.

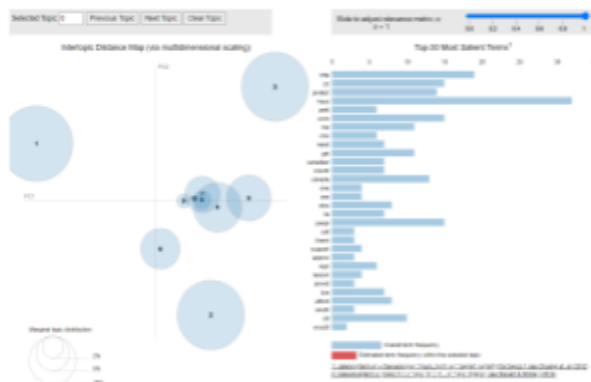


Figure 8. Interactive Visualization of Topic Model From LDA Analysis For r/CanadaHousing

“call”, “thank”, “support”, “approval”, “resign”, “provide”, “live”, “afford”, “another”, “exactly”, are the most relevant words across all topics. Most of these words suggest a need for housing and the the urge to “protest” and call this scenario a “crisis”. We can also assume that the presence of “http” is because people on Reddit like sharing links that better convey the real-world situation of Canada’s housing. As such, this interactive visualization provides a comprehensive overview of the topics and their relationship with each other, as well as the most relevant terms within these topics.

Produced Graphs For r/Canada Subreddit

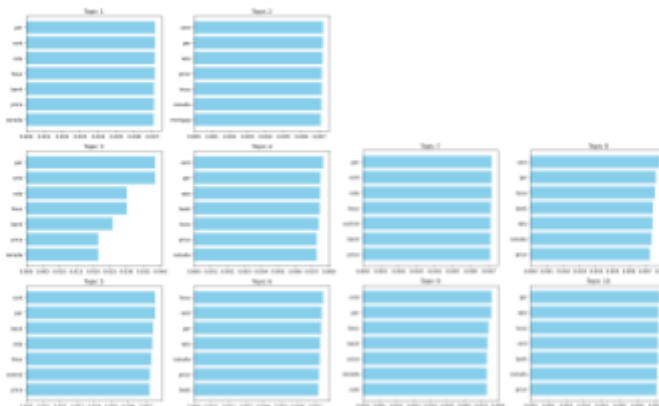


Figure 9. Top Words In Each Topic From LDA Analysis For r/Canada

subreddit, highlighting the key themes and concerns of the community members.

This figure, Figure 8, presents an interactive visualization of the topic model derived from our LDA analysis of the r/CanadaHousing subreddit. The map shows that most of the smaller topics overlap around the center. While the bigger topics are more spread out and distinct, suggesting that they cover different aspects of the housing discussion. On the right, the bar graph displays the Top-30 Most Salient Terms across all topics. These terms, which include “http”, “protest”, “house”, “work”, “mp”, “crisis”, “need”, “get”, “canadian”, “month”, “canada”, “one”, “see”, “also”, “people”,

This figure, Figure 9, presents the top words in each of the 10 topics identified in our LDA analysis of the r/Canada subreddit. Each bar graph represents a different topic, with the words that are most representative of that topic displayed along the y-axis and their respective weights displayed along the x-axis. For instance, Topic 1 includes words like “hill”, “ottawa”, and “canada”. Topic 2 includes province names and “canada”. Topic 3 includes terms like “price” and “market”. These results provide a snapshot of the discussions happening in the r/Canada

This bar graph, Figure 10, illustrates the average proportions of the ten topics identified in our LDA analysis of the r/Canada subreddit. The x-axis represents the topics, labeled from Topic 1 to Topic 10, while the y-axis represents the average proportions, ranging from 0.000 to 0.030. The graph shows that Topic 3 has a significantly higher average proportion, indicating that it is discussed more frequently or garners more attention in the subreddit. On the other hand, the other topics have similar, much lower proportions, suggesting that they are less frequently discussed. These results offer valuable insights into the prevalence of various topics in the r/Canada subreddit, highlighting the dominant themes and less common ones, thereby providing a useful understanding of the main concerns and interests of the community members.

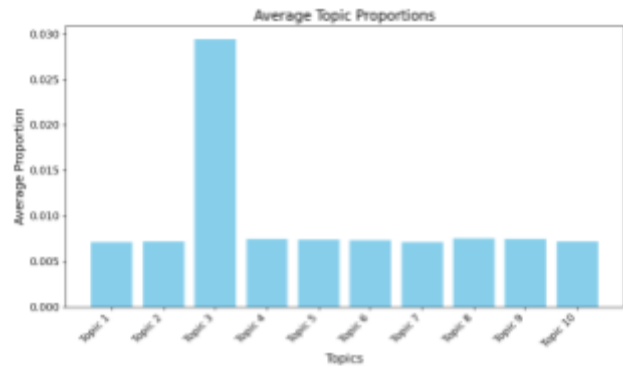


Figure 10. Average Proportions of Topics From LDA Analysis For r/Canada

The two posts shown in Figure 11 are the most representative of Topic 3 in our LDA analysis of the r/Canada subreddit. These posts provide valuable insights into the discussions that dominate this topic. The first post, titled “First Canadian to orbit moon in attempt to find affordable housing” and tagged as satire, humorously suggests a Canadian is going to the moon to find affordable housing. It has received 17.9k upvotes and has 222 comments. The comments joke about finding prime property on the moon’s Sea of Tranquility. The second post, titled “Grade 8 Student Calls for Ban on Housing Minister’s Personal Investments In Housing Market”, discusses a call for a ban on the Housing Minister’s personal investments in the housing market. It has received 7.8k upvotes. A comment mentions the rising rent prices in Toronto and criticizes political apathy towards it. These posts provide a snapshot of the discussions happening in the r/Canada subreddit, particularly around Topic 3. They highlight the concerns and experiences of community members, offering a deeper understanding of the context and content of this topic that most users talk about when mentioning housing.



Figure 11. Top 2 Documents Associated With Topic 3 From LDA Analysis For r/Canada

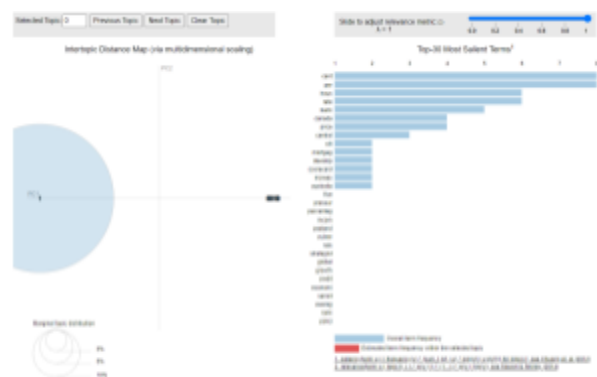


Figure 12. Interactive Visualization of Topic Model From LDA Analysis For r/Canada

The interactive visualization (Figure 12) of the topic model from the Latent Dirichlet Allocation (LDA) analysis for the subreddit r/Canada presents 10 topics. The most prevalent topic, Topic 1, is represented by the largest circle on the Intertopic Distance Map. Its size suggests a high prevalence in the discussions within the subreddit. The nine remaining smaller topics are clustered closely together, indicating their similarity or relatedness. Additionally, the bar graph titled “Top-30 Most

Salient Terms” lists the most informative terms across all topics, with ‘cent’, ‘per’, ‘hous’, ‘state’, ‘bank’, ‘canada’, ‘price’, ‘city’, and ‘mortgage’ being among the most salient. These terms, related to finance and economy, have the highest frequencies, as indicated by the length of the bars. This visualization provides a comprehensive understanding of the underlying topics in the Reddit posts, their relevance, and relationships, offering valuable insights into the discussions happening in the subreddit r/Canada.

Discussion and Conclusions

Connecting the results of the research with the theories discussed in the introduction and literature review provides a clearer picture of how people feel about housing costs in Canada on social media.

The sentiment analysis of two important subreddits, r/CanadaHousing and r/Canada, revealed interesting trends in discussions about housing costs. The word clouds showed that terms like "housing," "affordable," "crisis," and location indicators like "Canada," "Ontario," and "Vancouver" were frequently mentioned. These terms confirm that discussions are focused on housing affordability and market conditions in specific regions of Canada, especially those with high housing costs (aligning with Smith, 1983). Furthermore, most of the sentiment expressed in the posts was negative, suggesting that many online and Reddit users are dissatisfied and worried about the housing situation in Canada, or are at least aware of the current situation. This aligns with the findings of Farhan (2023) who highlighted the widespread concerns about affordability. The scatter plots also supported this, showing that the majority of posts were classified as negative or neutral, indicating a prevailing pessimistic view among Reddit users regarding housing costs in Canada.

Supporting this, the topic modelling analysis provided deeper insights into the specific themes and discussions within these subreddits. Topics such as pricing (correlating with Revington & August, 2019's discussion of financialization), locations, types of housing, and government policies showed up as prevalent themes across both r/CanadaHousing and r/Canada. These findings align with the literature review, which highlighted various challenges related to housing affordability in Canada, including economic, social, and political factors (Farhan, 2023; Kipfer & Sotomayor, 2024).

Moving on to the research question, "What do social media users think about housing costs in Canada?", it guided the analysis of sentiments expressed on platforms like Reddit. Our hypothesis suggested that social media discussions would include mostly negative and possibly neutral opinions, but rarely any positive ones about housing costs in Canada. This idea was based on existing research indicating widespread concerns about affordability and the varied nature of social media conversations, which can sometimes be more critical.

The analysis confirmed the hypothesis, showing that negative sentiments were prevalent in discussions on both the r/CanadaHousing and r/Canada subreddits. These sentiments highlighted deep-rooted concerns and frustrations among social media users about the challenges of finding affordable housing in Canada, particularly regarding high pricing and lack of government intervention. Common themes such as pricing, government policies, and regional differences further enriched our understanding of public discourse on this issue, aligning with the arguments presented by Kipfer and Sotomayor (2024) on systemic capitalism issues worsening housing problems.

However, it is important to acknowledge that social media data may not be entirely representative of the entire Canadian population. Additionally, sentiment analysis techniques have limitations in capturing the nuances of human language. Future research could explore public opinion on housing affordability using a wider range of social media platforms and potentially complement social media analysis with traditional survey methods.

By connecting theory with real-world findings from social media, this study provides a detailed perspective on the complexities of housing affordability in Canada. The information gained from social media discussions can inform political parties and researchers working to address this important issue. Additionally, they can help the public better understand the factors influencing housing costs, leading to more informed discussions and awareness efforts.

References

- Armstrong, P. (2024, February 7). Analysis | Bank of Canada governor says interest rate relief won't deliver New Housing Supply | CBC News. CBCnews. <https://www.cbc.ca/news/business/armstrong-housing-bank-of-canada-macklem-1.7106867>
- Farhan, B. Y. (2023). Canada's leadership and housing affordability: Evidence from the Canadian real estate market. *Journal of Urban Management*. Advance online publication. <https://doi.org/10.1016/j.jum.2023.11.001>
- Garcia, J., & Chen, L. (2022). The Influence of Social Media on Public Perceptions of Housing Affordability: A Case Study of Canadian Millennials. *Journal of Housing Studies*, 15(3), 321-339.
- Housing challenges remain for vulnerable populations in 2021. Statistics Canada. (2022, July 21). <https://www150.statcan.gc.ca/n1/daily-quotidien/220721/dq220721b-eng.htm>
- Kipfer, S., & Sotomayor, L. (2024). Housing beyond land rent? A critique of market housing solutionism. *Radical Housing Journal*, 6(1). <https://doi.org/10.54825/GQAU9087>
- Nguyen, H., & Patel, R. (2024). Social Media as a Tool for Advocacy: Mobilizing Public Opinion on Housing Affordability in Canada. *Journal of Public Policy & Marketing*, 43(1), 56-72.
- Ramage, K., Bell, M., Zaretsky, L., Lee, L., & Milaney, K. (2021). Is the right to housing being realized in Canada? Learning from the experiences of tenants in affordable housing units in a large Canadian city. *Societies*, 11(2), 53. <https://doi.org/10.3390/soc11020053>

Revington, N., & August, M. (2019). Making a market for itself: The emergent financialization of student housing in Canada. *Sage Journals*, 52(5). <https://doi.org/10.1177/0308518X19884577>

Smith, L. B. (1983). The Crisis in Rental Housing: A Canadian Perspective. *American Behavioral Scientist*, 26(4), 58-75. <https://doi-org.myaccess.library.utoronto.ca/10.1177/0002716283465001006>

Wang, Y., & Li, X. (2023). Exploring Public Discourse on Housing Costs in Canada: A Social Media Analysis Approach. *Canadian Journal of Sociology*, 46(2), 187-205.