Analyzing Public Sentiments on Urban Transportation in Montreal Using GPT-4o

Alireza Lorestani

A Thesis

in

The Department

of

Concordia Institute for Information Systems Engineering(CIISE)

Presented in Partial Fulfillment of the Requirements

for the Degree of

Master of Applied Science (Quality Systems Engineering) at

Concordia University

Montréal, Québec, Canada

June 2025

© Alireza Lorestani, 2025

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: **Alireza Lorestani**

Entitled: **Analyzing Public Sentiments on Urban Transportation in Montreal Using GPT-4o**

and submitted in partial fulfillment of the requirements for the degree of

Master of Applied Science (Quality Systems Engineering)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the Final Examining Committee:

______________________________ Chair
Dr. Fuzhan Nasiri

______________________________ Examiner
Dr. Fuzhan Nasiri

______________________________ Examiner
Dr. Chun Wang

______________________________ Supervisor
Dr. Ursula Eicker

Approved by ______________________________
Dr. Chun Wang, Chair
Department of Concordia Institute for Information Systems Engineering(CIISE)

____________ 2025

____
Mourad Debbabi, Dean
Faculty of Engineering and Computer Science

Abstract

Analyzing Public Sentiments on Urban Transportation in Montreal Using GPT-4o

Alireza Lorestani

This thesis investigates how people in Montreal feel about transportation by analyzing posts on X (formerly Twitter) using a Large Language Model (LLM) called GPT-4o. Montreal is a unique city with French and English speakers, making public opinion mining challenging. However, GPT-4o can directly process both languages, making the analysis more accurate and efficient.

Unlike traditional methods that often struggle to capture the nuances of language, GPT-4o generates precise sentiment analysis, helping us understand the emotions behind people's opinions. this tool was used to categorize tweets into transportation modes (e.g., bus, metro, and train), specify aspects (e.g., safety, cost, and punctuality), and overall sentiment (positive, negative, or neutral). Local terms like ”REM” and ”STM” were included to ensure the AI understood the context. AI-generated aspects were then grouped into standardized categories like reliability, cost, safety, and environmental impact to enhance clarity and consistency.

This approach showcases the flexibility and scalability of LLMs for multilingual public opinion mining. The study revealed significant differences in public sentiment across transportation modes and aspects, such as safety concerns in cycling and punctuality issues for public transit. These insights are valuable for transportation planners and policymakers seeking to improve urban mobility.

Future research could explore other public opinion sources or use this technology for real-time sentiment tracking to aid urban infrastructure planning.

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Ursula Eicker, for her persistent support, insightful guidance, and encouragement throughout this journey. Her expertise and patience have been instrumental in shaping the direction of this thesis. I sincerely thank all the CERC Lab members. Their knowledge, feedback, and company created an inspiring and supportive environment for this research. Your insights and assistance have greatly contributed to the progress of this work.

I want to give special thanks to my family, especially my Parents, for their love, encouragement, and sacrifices, which have been a constant source of motivation. To my partner Nasim, thank you for your support, understanding, and love, which have given me the strength to persevere through the challenges of this journey.

Finally, I would also like to acknowledge the developers of GPT-4o and related tools, whose groundbreaking technology made this research possible. The insights gained from their work have been invaluable.

To everyone who contributed to this thesis in ways big and small, I am sincerely thankful.

Contents

List of Figures x

List of Tables xi

1 Introduction 1

1.1 Background and Context . . . 1

1.2 Overview of the Framework . . . 1

1.3 Problem Statement . . . 3

1.4 Research Objectives . . . 4

1.5 Research Contributions . . . 5

1.6 Thesis Structure . . . 5

2 Literature Review 7

2.1 Introduction . . . 7

2.2 Public Sentiment Analysis . . . 7

2.3 Definition of Sentiment Analysis . . . 7

2.4 Use of Social Media for Data Collection . . . 8

2.5 Sentiment Analysis Techniques . . . 8

2.6 Aspect-Based Sentiment Analysis . . . 9

2.7 Multilingual Challenges and LLM Capabilities . . . 10

2.8 Summary . . . 10

3 Methodology **11**

3.1 Introduction . . . 11

3.2 Tools and Technologies . . . 11

3.2.1 GPT-4o . . . 11

3.2.2 Apify: Tweet Scraper V2 by API Dojo . . . 12

3.3 Data Collection . . . 13

3.3.1 Keywords Used for Data Collection . . . 14

3.3.2 Data Overview . . . 14

3.4 Data Cleaning and Preprocessing . . . 15

3.5 Categorization and Sentiment Analysis with GPT-4o . . . 17

3.5.1 Prompt Structure . . . 17

3.5.2 API Call for GPT-4o . . . 18

3.5.3 Advantages and Challenges . . . 19

3.5.4 Example Classifications . . . 19

3.6 Validation of GPT-4o Outputs . . . 19

3.6.1 Overview . . . 19

3.6.2 Method . . . 20

3.6.3 Results . . . 20

3.6.4 Interpretation . . . 21

3.7 Aspect Standardization . . . 21

3.8 Analysis and Insights Extraction . . . 23

3.8.1 Net Sentiment Score (NSS) . . . 23

3.8.2 Visualization Techniques . . . 23

3.8.3 Mode- and Aspect-Level Insights . . . 24

3.8.4 Temporal Analysis . . . 24

3.8.5 Location-Based Analysis . . . 24

3.9 Summary . . . 25

4 Results **26**

4.1 Introduction 26

4.2 Sentiment Analysis 26

4.2.1 Overall Sentiment 26

4.2.2 Sentiment by Transportation Modes 27

4.2.3 Sentiment by Aspect 28

4.2.4 Monthly Sentiment Trends 30

4.2.5 Sentiment by Aspect for Each Mode 31

4.2.6 Sentiment by Aspect for Each Mode Over Time 32

4.3 Mode Trends Over Time 34

4.4 Aspect Distribution within Modes 35

4.5 Location-based analysis 36

4.5.1 Aspect and Mode Distribution within Location 37

4.5.2 Sentiment Distribution within Location 39

4.5.3 Mode and Mode-Aspect Sentiment Analysis by Location 40

4.6 Summary 41

5 Discussion **43**

5.1 Introduction 43

5.2 Key Insights and Implications 43

5.2.1 Overall Sentiment Trends 43

5.2.2 Mode-Specific Sentiment Analysis 43

5.2.3 Aspect-Based Sentiment Analysis 44

5.2.4 Temporal Sentiment Trends 44

5.2.5 Location-Based Sentiment Analysis 44

5.2.6 Methodological Contributions 45

5.3 Policy Implications 45

5.4 Challenges and Limitations 46

5.5 Conclusion 47

6 Conclusion **48**

6.1 Summary of Findings 48

6.2 Contributions 49

6.2.1 Methodological Contributions 49

6.2.2 Theoretical Contributions 49

6.2.3 Practical Contributions 49

6.3 Limitations and Future Work 50

6.3.1 Limitations 50

6.3.2 Future Work 50

6.4 Final Overview: What This Research Found and Why It Matters 51

Appendix A Detailed Sentiment Distribution by Aspect for Each Mode **52**

A.1 Bus 52

A.1.1 Table: Sentiment Distribution by Aspect for Bus 52

A.1.2 Figure: Sentiment Distribution by Aspect for Bus 54

A.2 Walking 54

A.2.1 Table: Sentiment Distribution by Aspect for Walking 54

A.2.2 Figure: Sentiment Distribution by Aspect for Walking 54

A.3 Bicycle 54

A.3.1 Table: Sentiment Distribution by Aspect for Bicycle 54

A.3.2 Figure: Sentiment Distribution by Aspect for Bicycle 54

A.4 Car 54

A.4.1 Table: Sentiment Distribution by Aspect for Car 54

A.4.2 Figure: Sentiment Distribution by Aspect for Car 54

A.5 Train 54

A.5.1 Table: Sentiment Distribution by Aspect for Train 54

A.5.2 Figure: Sentiment Distribution by Aspect for Train 54

A.6 Ride-share 54

A.6.1 Table: Sentiment Distribution by Aspect for Ride-share 54

A.6.2 Figure: Sentiment Distribution by Aspect for Ride-share 54

Bibliography **57**

List of Figures

Figure 1.1 Workflow of the research methodology. 2

Figure 3.1 Workflow of the data collection process. 14

Figure 3.2 Tweet Frequency Over Time. 16

Figure 3.3 Tweets Word Cloud. 16

Figure 4.1 Overall Sentiment Distribution. 27

Figure 4.2 Sentiment Distribution by Transportation Modes. 28

Figure 4.3 Sentiment Distribution by Aspects. 30

Figure 4.4 Monthly Net Sentiment Trends Across the Entire Dataset. 31

Figure 4.5 Mode Trends Over Time. 34

Figure 4.6 Aspect Distribution within Each Mode. 36

Figure 4.7 Tweet Volume by Location. 37

Figure 4.8 Mode Distribution within Each Location. 38

Figure 4.9 Aspect Distribution within Each Location. 38

Figure 4.10 Sentiment Distribution across Standardized Locations. 40

Figure A.1 Sentiment Distribution by Aspect for Bus. 53

Figure A.2 Sentiment Distribution by Aspect for Walking. 54

Figure A.3 Sentiment Distribution by Aspect for Bicycle. 54

Figure A.4 Sentiment Distribution by Aspect for Car. 55

Figure A.5 Sentiment Distribution by Aspect for Train. 56

Figure A.6 Sentiment Distribution by Aspect for Ride-share. 56

List of Tables

Table 1.1 Comparison of challenges in sentiment analysis and how GPT-4o addresses them. 4

Table 3.1 Initial and Additional Extracted Keywords Used in Data Collection 14

Table 3.2 Location Distribution 17

Table 3.3 Sample Tweet Classifications with Modes, Aspects, Sentiments, and Standardized Aspects 20

Table 3.4 Agreement between GPT-4o and baseline models across 100 tweets. 20

Table 3.5 Grouped Mappings of Suggested Aspects to Final Categories 22

Table 4.1 Overall Sentiment Distribution. 27

Table 4.2 Sentiment by Transportation Modes. 28

Table 4.3 Sentiment by Aspect. 29

Table 4.4 Aspect Distribution within Each Mode (Percentage). 35

Table 4.5 Mode of Transportation vs. Location Distribution 39

Table 4.6 Aspect Distribution by Location 39

Table 4.7 Sentiment Distribution Across Standardized Locations 40

Table 4.8 Net Sentiment Score (NSS) Across Standardized Locations 40

Table 4.9 Mode and Mode-Aspect Sentiment Analysis by Location 42

Table A.1 Sentiment Distribution by Aspect for Bus. 52

Table A.2 Sentiment Distribution by Aspect for Walking. 53

Table A.3 Sentiment Distribution by Aspect for Bicycle. 54

Table A.4 Sentiment Distribution by Aspect for Car. 55

Table A.5 Sentiment Distribution by Aspect for Train. 55
Table A.6 Sentiment Distribution by Aspect for Ride-share. 56

Chapter 1

Introduction

Chapter 1 introduces the research context, objectives, and the importance of public opinion mining in transportation.

1.1 Background and Context

Transportation is a vital part of urban life, affecting millions of people's daily lives. In the city of Montreal, which is known for its bilingual population and different urban challenges, public transportation is essential for residents. Studying public sentiment concerning transportation systems is crucial for policymakers and planners to understand their weaknesses and improve their infrastructure and services. Social media platforms, especially X (formerly known as Twitter), provide a rich source of real-time public opinions on various aspects of urban transportation Pak and Paroubek (2010). Analyzing this data offers valuable insights into user experiences, concerns, and expectations. However, such analysis is complex in multilingual settings like Montreal Bouazizi and Ohtsuki (2017), where tweets are in English and French, and sentiments often depend on cultural and contextual factors.

1.2 Overview of the Framework

This section provides a high-level overview of the methodology used in this study. The workflow diagram in Figure 1.1 illustrates the sequential steps used to analyze public sentiment about transportation systems in Montreal.

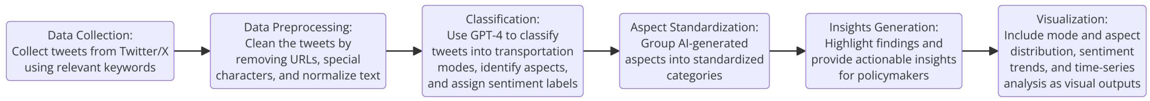

Figure 1.1: Workflow of the research methodology.

The methodology consists of the following steps:

(1) **Data Collection:** Tweets were collected from X (formerly Twitter) using an initial standard list of keywords related to transportation. Then, an iterative process was employed to refine the list by identifying additional keywords from the collected data. This assured complete coverage of transportation-related discussions in Montreal.

(2) **Text Cleaning and Preprocessing:** The collected tweets were cleaned to remove irrelevant information, such as URLs and special characters while maintaining linguistic diversity in both languages.

(3) **AI-Powered Classification:** Each tweet was categorized into:

- **Predefined Transportation Modes:** For example, bus, metro, train, etc.
- **AI-Generated Aspects:** These aspects were identified by GPT-4o and reflect features such as cost, safety, and reliability.
- **Sentiment:** Tweets were labeled with three sentiment categories: Positive, Negative, or Neutral.

(4) **Aspect Grouping:** The AI-generated aspects were further standardized into broader categories, such as reliability, cost, and safety, to ensure consistency and ease of interpretation.

(5) **Sentiment and Aspect Analysis:** The data was analyzed to identify patterns, trends, and public concerns across transportation modes and aspects. Key findings were derived by aggregating results over time and by mode.

This framework demonstrates the effectiveness of state-of-the-art language models over the traditional analytical methods to extract valuable insights from public opinion data in multilingual and urban contexts.

1.3 Problem Statement

Sentiment analysis plays a critical role in understanding public opinion on various urban systems, including transportation. However, traditional sentiment analysis methods face significant challenges when applied to real-world, multilingual, and domain-specific datasets. One of the major issues is the inability of conventional models to effectively analyze French-language text, as most sentiment analysis tools are trained primarily on English datasets Shahriar, Lund, Mannuru, et al. (2024). Attempting to bridge this gap through translation introduces information loss, inaccuracies, and a reduction in context-specific insights, ultimately compromising analytical accuracy.

Beyond language-specific challenges, context-awareness remains a major limitation of traditional sentiment analysis techniques. Many models struggle with complex sentence structures, sarcasm, and implicit sentiment that require a deeper understanding of linguistic nuances. In social media data, where sentiment is often expressed through informal language, abbreviations, and emojis, traditional machine learning models tend to misclassify opinions due to their reliance on predefined lexicons and rule-based classifiers.

Another fundamental challenge is domain-specific terminology in urban transportation. Many transportation-related terms such as ”REM,” ”STM,” and ”BIXI” in Montreal have distinct meanings that general-purpose sentiment models may fail to recognize. Without proper domain adaptation, these models may misinterpret or misclassify transportation-related tweets, leading to unreliable insights.

Lastly, Code-Mixing and language switching between English and French are common in Montreal's social media discourse. Many tweets contain a mix of both languages within a single sentence or post, making it difficult for monolingual sentiment models to accurately process and classify them.

This research aims to overcome these challenges by leveraging GPT-4o, a large language model, to perform multilingual, context-aware, and aspect-based sentiment analysis of public transportation tweets. By directly analyzing tweets in English and French without requiring translation, GPT-4o enhances accuracy and preserves linguistic nuances. Moreover, it enables fine-grained categorization of tweets by transportation mode and aspect, ensuring a more structured and interpretable analysis of public opinion. This approach establishes a robust framework for real-time sentiment tracking, providing valuable insights that can inform policy decisions and urban transportation planning.

Challenge	**Traditional Methods**	**GPT-4o Approach**
Multilingual Data	Struggles with French-language text due to limited training data, requiring translation that introduces inaccuracies.	Processes English and French natively, preserving linguistic nuances without translation.
Context Awareness	misinterpreting sarcasm, abbreviations, emojis, and complex sentence structures, leading to inaccurate sentiment classification.	Captures nuanced sentiment using deep contextual understanding, recognizing sarcasm and informal language.
Domain-Specific Terminology	Fails to recognize transportation-related terms like "REM," "STM," and "BIXI," leading to misclassification.	Identifies and categorizes domain-specific transportation terms accurately, improving sentiment analysis precision.
Code-Mixing	Struggles with tweets that contain a mix of English and French within the same post, reducing classification accuracy.	Seamlessly handles mixed-language content, ensuring accurate sentiment classification in multilingual discourse.

Table 1.1: Comparison of challenges in sentiment analysis and how GPT-4o addresses them.

1.4 Research Objectives

The objectives of this research are:

- To categorize tweets into transportation modes (e.g., bus, metro, train) and aspects (e.g., cost, safety, punctuality).
- To perform sentiment analysis, classifying tweets as positive, neutral, or negative.
- To group AI-generated aspects into standardized categories for consistency in analysis.
- To extract insights from public sentiment data related to urban transportation systems.

1.5 Research Contributions

This research makes the following contributions:

- **Bilingual Sentiment Analysis**: Developed a methodology to analyze tweets in both English and French without translation, leveraging GPT-4o's multilingual capabilities to improve accuracy and preserve context.
- **Context-Aware Sentiment Analysis**: Demonstrated the use of a Large Language Model (LLM) for sentiment analysis, which captures nuanced, context-aware sentiments compared to traditional methods often trained primarily on English datasets.
- **Standardized Aspect Categorization**: Introduced a two-step classification approach where aspects were generated by AI and subsequently grouped into predefined categories (e.g., safety, cost, convenience) for consistent analysis.
- **Scalable Methodology for Public Opinion Mining**: Proposed a flexible and replicable framework for analyzing public opinion in multilingual settings, adaptable to other domains beyond transportation.
- **Insights into Montreal's Transportation**: Provided actionable insights into public sentiment across transportation modes and aspects in Montreal, highlighting key challenges such as safety concerns with cycling and punctuality issues in public transit.

1.6 Thesis Structure

This thesis is organized into the following chapters:

- **Chapter 1: Introduction** — Provides the background, motivation, objectives, and structure of the thesis.
- **Chapter 2: Literature Review** — Reviews existing research on sentiment analysis, public transportation systems, and multilingual NLP models.

- **Chapter 3: Methodology** — Details the data collection, preprocessing, and classification methods used in the study.
- **Chapter 4: Results and Analysis** — Presents the findings, including sentiment trends, mode-specific insights, and key patterns in the data.
- **Chapter 5: Discussion** — Interprets the results, discusses limitations, and suggests potential applications.
- **Chapter 6: Conclusion and Future Work** — Summarizes the research contributions and outlines opportunities for further studies.

Chapter 2

Literature Review

2.1 Introduction

This chapter reviews existing literature on public sentiment analysis, the use of social media for data collection, advancements in sentiment analysis techniques, aspect-based sentiment analysis, and addressing multilingual challenges in urban contexts. The review provides a foundation for understanding the methodologies and tools used in this research.

2.2 Public Sentiment Analysis

Analyzing public sentiment is crucial for understanding perceptions of transportation systems Zheng, Capra, Wolfson, and Yang (2014). Studies have shown that sentiment analysis can reveal valuable insights into user satisfaction, service quality, and policy impacts Cambria and White (2017); Ortigosa, Martín, and Carro (2014). Specifically, in urban contexts, sentiment analysis has been used to assess public transport systems, revealing key areas for improvement Zheng et al. (2014).

2.3 Definition of Sentiment Analysis

Sentiment analysis, also known as opinion mining, is a natural language processing (NLP) technique used to determine the emotional tone expressed in a piece of text. The goal is to classify the sentiment as positive, negative, or neutral. It is widely used in various domains, including business, politics, and public service management, to assess public perception and trends Pang and Lee (2008). Sentiment analysis can be conducted at different levels:

- **Document-Level Sentiment Analysis**: Determines the overall sentiment of a document or text block.
- **Sentence-Level Sentiment Analysis**: Analyzes individual sentences to classify their sentiment.
- **Aspect-Based Sentiment Analysis (ABSA)**: Extracts sentiment toward specific aspects within a text Pontiki, Galanis, Pavlopoulos, et al. (2016).

2.4 Use of Social Media for Data Collection

Social media platforms, particularly Twitter (now X), are widely used for public sentiment analysis due to their real-time and large-scale data availability. Researchers have successfully leveraged Twitter data to evaluate urban mobility and public transportation services Bollen, Mao, and Zeng (2011). Social media data provides dynamic insights that traditional survey-based approaches may lack, allowing policymakers to assess public perception in real-time Burnap and Williams (2015).

However, challenges persist in using social media for sentiment analysis. These include the presence of spam, bot-generated content, evolving language trends, and informal expressions, which necessitate robust preprocessing techniques Mahmud et al. (2012). Advanced NLP techniques, such as noise filtering and neural embeddings, have been introduced to address these challenges, improving the accuracy of sentiment classification Conneau et al. (2020).

2.5 Sentiment Analysis Techniques

Several techniques are used in sentiment analysis, ranging from traditional approaches to deep learning-based methods:

- **Lexicon-Based Approaches**: These methods rely on predefined sentiment dictionaries, such as SentiWordNet or VADER, to assign sentiment scores to words and aggregate them for classification Taboada, Brooke, Tofiloski, et al. (2011).

- **Machine Learning Approaches**: Supervised learning techniques such as Support Vector Machines (SVM), Naïve Bayes, and Random Forest have been used to train sentiment classifiers Pang and Lee (2008).

- **Deep Learning Approaches**: Neural networks, particularly transformer-based architectures like BERT and GPT-4o, have significantly improved accuracy by capturing contextual meanings and sentiment nuances Brown et al. (2020); Devlin et al. (2019).

Recent advancements in LLMs, like GPT-4o, offer enhanced capabilities in handling nuanced and contextual sentiment analysis. Unlike earlier models, LLMs can infer sentiments based on surrounding context, effectively handling complex expressions such as sarcasm and negation Devlin et al. (2019).

2.6 Aspect-Based Sentiment Analysis

Aspect-based sentiment analysis (ABSA) goes beyond generic sentiment classification by identifying specific aspects or features being discussed. This approach has been applied in diverse domains, including transportation, to understand user opinions on punctuality, safety, and cost Pontiki et al. (2016). ABSA provides granular insights, making it a critical component of public sentiment analysis.

Recent research has focused on improving ABSA performance through the use of attention mechanisms and domain-adaptive language models. Transformer-based models, such as BERT and GPT-4o, have shown improved accuracy in extracting aspect-level sentiments Devlin et al. (2019). Additionally, unsupervised and semi-supervised approaches are being explored to reduce dependency on labeled datasets Conneau et al. (2020).

2.7 Multilingual Challenges and LLM Capabilities

Montreal's bilingual context introduces complexities in public sentiment analysis, as tweets are often in both English and French. Studies have highlighted the limitations of traditional tools in handling multilingual data **?**. Conventional machine learning models struggle with mixed-language text, often requiring separate models for each language, which increases computational complexity and potential inconsistencies Conneau et al. (2020).

LLMs, such as GPT-4o, have demonstrated superior performance in processing multilingual content without requiring translation, preserving data integrity, and improving accuracy Devlin et al. (2019); Shahriar et al. (2024). These models are trained on large multilingual datasets, allowing them to understand and process language variations more effectively. Furthermore, transfer learning techniques have enabled fine-tuning of these models for specific applications, such as sentiment analysis in urban transportation Shahriar et al. (2024).

2.8 Summary

The reviewed literature highlights the evolution of sentiment analysis techniques, the importance of aspect-based analysis, and the unique challenges caused by multilingual contexts. Traditional sentiment analysis methods have limitations in handling informal language, sarcasm, and mixed-language text, whereas modern LLMs like GPT-4o provide robust solutions by incorporating deep contextual understanding and multilingual capabilities. These insights inform the methodology employed in this research, leveraging the capabilities of LLMs for a comprehensive analysis of public sentiment in Montreal's transportation system.

Chapter 3

Methodology

3.1 Introduction

This chapter outlines the methodology used in this research, including the data collection process, data cleaning and preprocessing, classification, sentiment analysis, and final analysis. Additional sections describe the tools employed, the data characteristics, and the Net Sentiment Score (NSS) calculation, ensuring reproducibility and clarity.

3.2 Tools and Technologies

The research leveraged various tools and technologies for data collection, preprocessing, and analysis. These tools enabled efficient data handling, sentiment analysis, and visualization.

3.2.1 GPT-4o

GPT-4o, an advanced Large Language Model (LLM) by OpenAI, was employed for sentiment analysis and aspect categorization. This tool was chosen for its multilingual capabilities and deep contextual understanding, essential for accurately processing transportation-related tweets in both English and French.

How GPT-4o Works:

- **Language Understanding:** GPT-4o is trained on a vast corpus of multilingual data, enabling it to understand syntax, semantics, and context in diverse languages. This allows it to process English and French tweets natively without translation, preserving nuances and reducing errors.

- **Contextual Comprehension:** The model uses transformers—a type of neural network architecture—to understand the relationships between words in a sentence. This is particularly useful for identifying sentiments (positive, neutral, negative) and categorizing aspects (e.g., cost, safety) within complex tweet structures.

- **Fine-Grained Analysis:** For this study, GPT-4o was prompted with specific instructions to:

  - Categorize tweets by transportation mode and aspect.

  - Analyze sentiment polarity while considering the broader context of the tweet (e.g., sarcasm, idioms).

Advantages of GPT-4o:

- Handles mixed-language content seamlessly.

- Adapts to predefined categories (e.g., reliability, cost) with precision.

- Reduces the need for manual intervention during categorization and sentiment tagging.

Challenges:

- Dependence on API-based processing, requiring robust computational resources.

- Potential biases in sentiment tagging due to the model's training data.

3.2.2 Apify: Tweet Scraper V2 by API Dojo

For data collection, the Tweet Scraper V2 actor by API Dojo, available on the Apify platform, was used instead of Tweepy. This tool allowed efficient scraping of tweets using predefined keywords.

Key Features of Tweet Scraper V2:

- **Keyword-Based Collection:** Supports targeted scraping of tweets using user-defined search terms (e.g., “bus,” “metro,” “train”).
- **Geographical Filtering:** Allows location-based filtering to ensure tweets are relevant to Montreal and surrounding areas.
- **Multilingual Support:** Captures tweets in various languages, including English and French, without language-specific restrictions.

Advantages:

- No need for Twitter Developer API access, simplifying setup.
- Handles large-scale data scraping efficiently.

Limitations:

- Reliance on public data may exclude private tweets or accounts with limited visibility.
- Requires additional preprocessing to filter noise and irrelevant content.

3.3 Data Collection

The dataset comprises tweets related to transportation in Montreal, collected from Twitter using predefined keywords. The tweets reflect public sentiment over three years (January 2022 to December 2024), providing a comprehensive view of trends across transportation modes and aspects. The data collection process involved:

(1) Using an initial list of keywords (e.g., “bus,” “train,” “bicycle”).

(2) Extracting additional keywords from the collected tweets (e.g., “STM,” “gare”).

(3) Conducting a second round of collection with the expanded keyword list to refine coverage.

The workflow is illustrated in Figure 3.1.

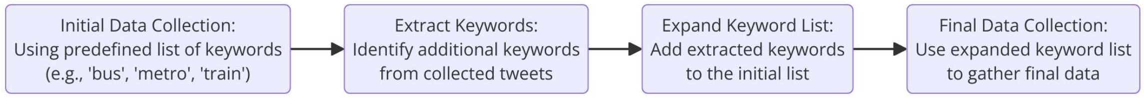

Figure 3.1: Workflow of the data collection process.

3.3.1 Keywords Used for Data Collection

Table 3.1 lists the keywords used in the collection process.

Table 3.1: Initial and Additional Extracted Keywords Used in Data Collection

Initial Keywords	**Additional Extracted Keywords**
bus, bicycle, Car, Ride Share, train, Walk, autobus, vélo, voiture, covoiturage, marche, à pied, Uber	station, gare, STM, métro, transport, cyclist, driver, ride

3.3.2 Data Overview

This section describes the data's scope and characteristics.

Dataset Characteristics

- **Volume:** Approximately 26,000 tweets were collected, with 17,500 retained after cleaning.
- **Languages:** The dataset includes both English and French tweets to reflect Montreal's bilingual context.
- **Key Fields:**
  - **Tweet Content:** Cleaned text of the tweet, free from unnecessary symbols and links.
  - **Mode and Aspect:** Categorization of tweets by transportation mode and associated aspect (e.g., bus: safety).
  - **Sentiment:** Labeled as positive, neutral, or negative.
  - **Metadata:**
    - *Timestamp:* The exact date and time when the tweet was posted.

* *Geographical Area:* Categorized into different cities or areas within the greater Montreal area.

- **Engagement Metrics:**
  * *Retweet Count:* Number of times a tweet was re-shared.
  * *Reply Count:* Number of replies received.
  * *Like Count:* Number of likes received.

The diversity of this dataset ensures a rich basis for analyzing public sentiment trends and transportation challenges.

Dataset Statistics

- **Tweet Frequency Over Time:** To analyze trends in public discussions on Montreal transportation, we visualized tweet frequency over time (3.2). The dataset was grouped by month to observe fluctuations in activity. The bar chart illustrates how tweet volume changed from January 2022 to December 2024, highlighting peaks that may correspond to major transit events, disruptions, or policy changes. This analysis helps identify periods of high public engagement and potential causes, such as severe weather, strikes, or new transportation initiatives.

- **Most Common Words in Montreal Transportation Tweets:** The common words visualization 3.3 can reveal recurring topics in Montreal transportation tweets.

3.4 Data Cleaning and Preprocessing

The raw data underwent preprocessing to ensure quality and consistency:

- Removed duplicates, irrelevant tweets, and special characters.
- Performed language detection to retain English and French tweets.
- Tokenized and lemmatized text for better classification and sentiment tagging.

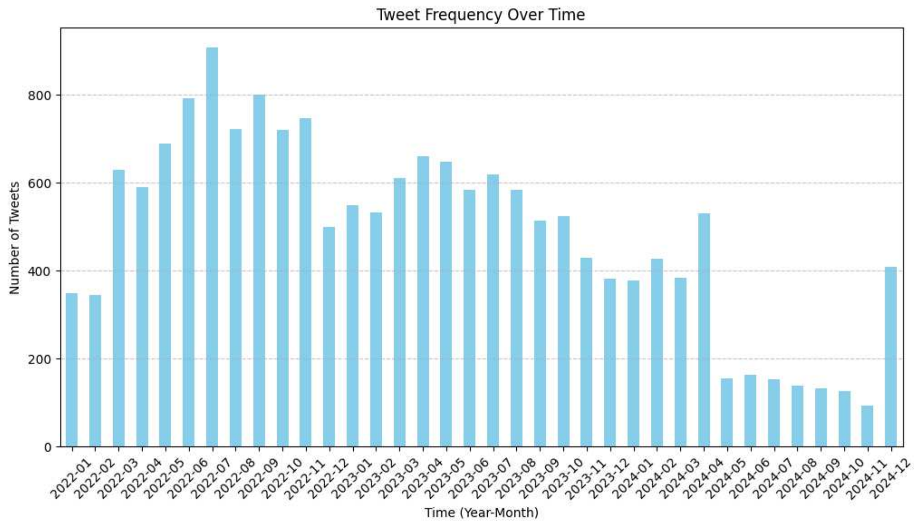

Figure 3.2: Tweet Frequency Over Time.

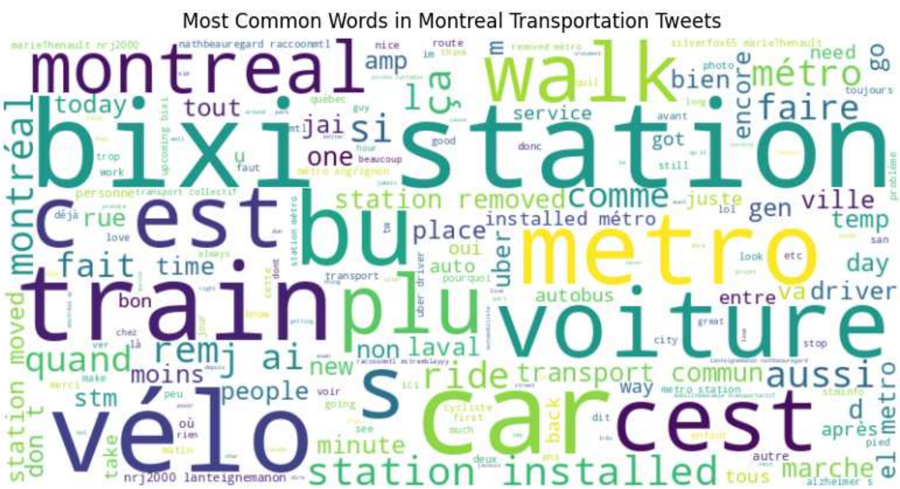

Figure 3.3: Tweets Word Cloud.

- Standardizing the Location into ”Montreal”, ”South Shore”,”North Shore”, ”West Island”, ”Off-Island Areas”, and ”Others”. The distribution is shown in table 3.2

As part of the preprocessing pipeline, irrelevant tweets were also filtered out. This step was handled directly during the classification process with GPT-4o. The prompt explicitly instructed the model to ignore tweets that contained transportation-related keywords but were not actually about transportation. For example, a tweet saying “I took my dog to get trained today” might include the word “train” but is unrelated to public transit. GPT-4o was directed to assign such tweets the label `unrelated`, and these were later excluded from the analysis.

Location	Count
Montreal	12604
North Shore	1994
Other	1182
South Shore	1044
West Island	587
Off-Island	208

Table 3.2: Location Distribution

3.5 Categorization and Sentiment Analysis with GPT-4o

Tweets were categorized into three dimensions using GPT-4o:

- **Transportation Mode:** Predefined categories such as bus, metro, train, car, walking, and bicycle.
- **Aspect:** AI-generated categories reflecting key features such as cost, punctuality, safety, and convenience.
- **Sentiment:** Positive, neutral, or negative sentiments.

GPT-4o was chosen for its ability to handle multilingual data (English and French) and its contextual understanding, ensuring accurate sentiment tagging and aspect classification.

3.5.1 Prompt Structure

The classification process used a structured prompt designed to provide GPT-4o with precise instructions. The prompt ensured consistency and accuracy by including:

- A brief task description instructing the model to classify tweets by transportation mode, aspect, and sentiment.
- Definitions of terms specific to Montreal's transportation system, such as REM," STM," BIXI," and OPUS Card," to guide contextual understanding.

- Examples of tweet classifications to establish a clear format and expectation.

Below is the structure of the prompt used:

> *Classify the following tweet by the mentioned transportation mode, aspect, and sentiment in the format* `mode:aspect:sentiment`*. 'Aspect' refers to a feature or quality of the transportation mode, such as cost, punctuality, safety, or convenience. Sentiments: positive, negative, or neutral.*

To improve the model's understanding and reduce ambiguity, example classifications were included:

- *"The bus was late again today, so frustrating!"* → `Bus:Reliability:Negative`
- *"Love how convenient the metro is during rush hour."* → `Metro:Convenience:Positive`
- *"BIXI bikes are affordable and easy to use!"* → `Bicycle:Cost:Positive`

Incorporating examples has been shown to improve model performance, particularly in few-shot learning scenarios Brown et al. (2020). By providing clear instructions and sample outputs, the prompt effectively reduced misclassifications and improved consistency in the dataset.

3.5.2 API Call for GPT-4o

The GPT-4o API was used to classify tweets in batches, ensuring efficient processing while effectively handling potential errors. The classification process for each tweet involved the following steps:

- Generating a dynamic prompt using the tweet's text.
- Sending the prompt to GPT-4o via an API call.
- Parsing the response to extract the classified label in the format `mode:aspect:sentiment`.

A retry mechanism was implemented to handle transient API errors to enhance reliability. Each failed request was retried up to three times. If all attempts failed, a default label of `unrelated:unrelated:unre` was applied to avoid missing data.

3.5.3 Advantages and Challenges

Advantages:

- **Contextual Understanding:** The model's ability to process multilingual content natively (English and French) ensured accurate and nuanced classifications.
- **Precision:** Explicit instructions and examples minimized ambiguity in classifications.
- **Efficiency:** The use of API calls allowed for automated classification of thousands of tweets with minimal manual intervention.

Challenges:

- **Error Handling:** Temporary API errors required implementing a retry mechanism.
- **Bias Risks:** Potential biases in GPT-4o's training data could affect sentiment and aspect classifications.
- **Resource Dependence:** The classification process relied heavily on computational resources and API access.

3.5.4 Example Classifications

Table 3.3 provides examples of tweets and their classifications by GPT-4o, illustrating the model's ability to handle various contexts and modes.

3.6 Validation of GPT-4o Outputs

3.6.1 Overview

To assess the reliability of GPT-4o's predictions, a validation step was conducted by comparing its outputs with those of two other advanced language models: Claude 3 Opus and Gemini 1.5 Pro. The comparison focused on agreement across three dimensions: transportation mode, aspect, and sentiment.

Table 3.3: Sample Tweet Classifications with Modes, Aspects, Sentiments, and Standardized Aspects

Tweet (Cleaned)	Mode	Aspect	Sentiment	Standardized Aspect
où sont les autobus	Bus	Availability	Negative	Convenience
is there one day that the mtl metro is not having problems	Train	Reliability	Negative	Reliability
its no longer safe for cyclists to share the road with cars drivers today are much worse because most of them never rode	Bicycle	Safety	Negative	Safety
oh goodness i hope that there wont be any scary happenings on this midnight walk im taking the dreaded vampyr	Walking	Safety	Negative	Safety

3.6.2 Method

A set of 100 randomly selected and cleaned tweets was used for this comparison. All tweets were classified by GPT-4o, Claude, and Gemini using the same prompt and configuration.

Agreement was calculated as the percentage of tweets where the model's output matched GPT-4o's output for each classification dimension.

3.6.3 Results

The agreement percentages between GPT-4o and the two baseline models are shown in Table 3.4. Results are reported separately for mode, aspect, and sentiment, along with the overall agreement across all labels.

Table 3.4: Agreement between GPT-4o and baseline models across 100 tweets.

Dimension	Gemini Agreement (%)	Claude Agreement (%)
Mode	87	84
Aspect	72	79
Sentiment	80	78
Overall	79	80

3.6.4 Interpretation

The results show that GPT-4o's outputs are largely consistent with those of Claude and Gemini. Agreement on transportation mode was highest, followed by sentiment. Aspect classification showed lower agreement, which is expected due to the greater variability and ambiguity in aspect-related language. These findings suggest that GPT-4o performs reliably across all three dimensions and is suitable for large-scale analysis of public sentiment in this context.

3.7 Aspect Standardization

AI-generated aspects were grouped into 8 predefined categories, including reliability, cost, and safety, to ensure consistency in analysis. This two-step approach—raw aspect generation followed by grouping—allowed for deeper exploration while maintaining interpretability. Table 3.5 details the aspect mappings.

Table 3.5: Grouped Mappings of Suggested Aspects to Final Categories

Final Category	**Mapped Suggested Aspects**
Reliability	Punctuality, Efficiency, Delivery, Service Disruption, Frequency, Speed, Duration, Timing, Travel Time, Waiting, Route Change, Operation, Pace
Cost	Cost, Economic Impact, Purchase, Ticketing
Safety	Safety, Personal Security, Accident Risk, Health, Legality, Regulation, Driving Skill, Weather, Privacy, Policy, Enforcement, Mental Health, Security, Road Conditions
Convenience	Availability, Usage, Route Planning, Accessibility, Utility, Location, Destination, Purpose, Seasonal Use, Necessity, Seasonal Availability, Seasonal Preparation, Mobility, Distance, Recreation, Effort, Choice, Freedom, Usability, Exploration
Comfort	Comfort, Cleanliness, Seating, Quality, Space, Noise, Appearance, Design, View, Feature, Sound, Scenery, Conditions
Environmental Impact	Environmental Impact, Emissions, Air Quality, Social Impact, Environment
Customer Service	Service, Trust, Public Perception, Lost Item, Perception, Popularity, Communication, Organization, Announcement, Branding, Information, Engagement, Behavior, Competence, Management, Comparison, Procedure, Etiquette, Inclusivity, Interaction, Community Engagement
Infrastructure	Infrastructure, Innovation, Traffic, Capacity, Urban Planning, Connectivity, Development, Stations, Maintenance, Planning, Coverage, Parking, Modernization, Installation, Future, Progress, Construction, Project, Facility, Bike Paths

3.8 Analysis and Insights Extraction

This section describes the methods used to analyze categorized data, extract key trends, and derive actionable insights.

3.8.1 Net Sentiment Score (NSS)

The Net Sentiment Score (NSS) was used as a primary metric for analyzing sentiment trends. NSS simplifies sentiment representation by combining positive and negative opinions into a single value, normalized by the total number of sentiment-labeled tweets. It is calculated as:

$$\text{NSS} = \frac{\text{Positive Sentiment Count} - \text{Negative Sentiment Count}}{\text{Total Sentiment Count}}$$

Role of NSS in Analysis:

- Simplifies the representation of sentiment trends over time and across transportation modes and aspects Boiy and Moens (2009).
- Highlights shifts in public sentiment, allowing for easy identification of peaks and troughs.
- Avoids the need for separate positive, neutral, and negative sentiment lines in visualizations, enhancing clarity.

3.8.2 Visualization Techniques

Data visualizations were crucial for uncovering trends and patterns. The following techniques were employed:

- **Line Charts:** Used to visualize temporal sentiment trends (e.g., monthly NSS trends).
- **Stacked Bar Charts:** Represented the distribution of aspects within each mode, showing the focus of public discussion.

Tool Support:

- Python libraries such as Matplotlib and Seaborn were used for creating detailed and publication-ready charts.

3.8.3 Mode- and Aspect-Level Insights

The analysis compared sentiment and aspect distributions across transportation modes to identify key differences:

- **Mode-Level Analysis:** Highlighted the modes receiving the most positive (e.g., walking) or negative (e.g., buses) sentiment.
- **Aspect-Level Analysis:** Identified the aspects driving public perception, such as safety concerns for cars and comfort for trains.

3.8.4 Temporal Analysis

Sentiment trends were analyzed over time using NSS to detect seasonal or event-driven shifts. Key methods included:

- Aggregating NSS values monthly to track sentiment peaks and valleys.
- Comparing trends across modes to identify shared patterns or mode-specific deviations.

3.8.5 Location-Based Analysis

To capture geographic variations in public sentiment, a location-based analysis was incorporated:

- **Grouping by Location:** Tweets were grouped by standardized locations (e.g., Montreal, North Shore, South Shore, West Island, Off-Island) to capture localized trends.
- **Visualization:** Location-specific visualizations, such as bar charts and heatmaps, were created to facilitate comparisons of sentiment and aspect distributions across regions.
- **Analytical Focus:** This approach allows for the identification of regional differences in public discussion and sentiment regarding transportation.

3.9 Summary

This chapter described the methodology, including data collection, tools, preprocessing, and analysis steps. The next chapter presents the findings derived from this approach.

Chapter 4

Results

4.1 Introduction

This chapter presents the results of the sentiment analysis on tweets related to transportation in Montreal. The findings are organized to show overall sentiment trends, sentiment by transportation modes and aspects, geographical trends, and temporal patterns. Each section provides data visualizations and tables where helpful to summarize the trends and distributions clearly.

4.2 Sentiment Analysis

4.2.1 Overall Sentiment

The overall sentiment distribution of tweets is as follows:

- **Positive Sentiment:** 28.12% of tweets showed satisfaction with transportation.
- **Neutral Sentiment:** 33.22% of tweets were factual or had no strong opinion.
- **Negative Sentiment:** 38.66% of tweets expressed dissatisfaction, mostly about cost and safety.

The overall sentiment distribution is presented in Table 4.1 and Figure 4.1.

Negative sentiment dominates the overall distribution, primarily due to concerns about cost and safety. Neutral sentiment reflects factual reporting or mixed opinions, while positive sentiment

Table 4.1: Overall Sentiment Distribution.

Sentiment Category	Percentage (%)
Positive	28.12
Neutral	33.22
Negative	38.66

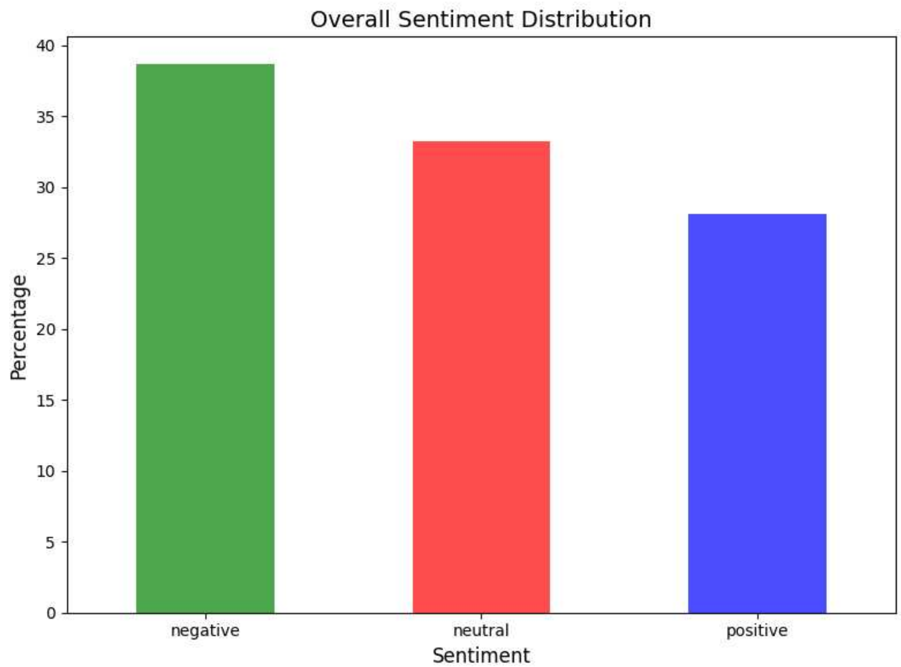

Figure 4.1: Overall Sentiment Distribution.

highlights areas where transportation systems meet public expectations.

4.2.2 Sentiment by Transportation Modes

Sentiment trends vary across different transportation modes. Below are key observations:

- **Walking:** The most positive mode, with 50.11% positive sentiment.
- **Car:** 53.42% negative sentiment, mostly due to congestion and cost.
- **Bus:** 52.23% negative sentiment, highlighting reliability issues.
- **Bicycle:** Balanced sentiment, with 36.69% positive, 33.86% neutral, and 29.45% negative sentiment.

- **Ride-share:** Moderate sentiment, with 48.38% negative, 33.65% neutral, and 17.97% positive sentiment.

- **Train:** Mixed sentiment, with 41.95% negative, 34.26% neutral, and 23.79% positive sentiment.

The sentiment distribution by transportation modes is presented in Table 4.2 and Figure 4.2.

Table 4.2: Sentiment by Transportation Modes.

Mode	Positive (%)	Neutral (%)	Negative (%)
Walking	50.11	37.44	12.45
Bicycle	36.69	33.86	29.45
Train	23.79	34.26	41.95
Bus	18.96	28.81	52.23
Ride-share	17.97	33.65	48.38
Car	15.22	31.36	53.42

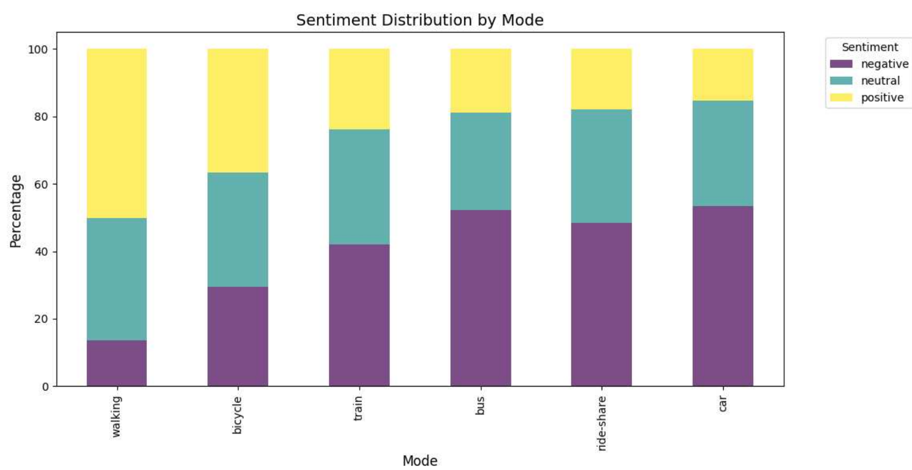

Figure 4.2: Sentiment Distribution by Transportation Modes.

Walking emerged as the most positively rated mode, likely due to its affordability and convenience. Cars and buses received significant negative sentiment due to congestion, high costs, and reliability issues.

4.2.3 Sentiment by Aspect

The sentiment distribution for specific aspects, such as cost, safety, and comfort, in all modes, is as follows:

- **Cost:** A major issue, with 62.70% negative sentiment, 21.37% neutral, and 15.92% positive sentiment.

- **Safety:** The most negatively perceived aspect, with 71.58% dissatisfaction, 17.05% neutral, and only 11.36% positive sentiment.

- **Comfort:** The most positively perceived aspect, with 44.38% positive sentiment, 39.12% neutral, and 16.49% negative sentiment.

- **Convenience:** Mixed sentiment, with 39.71% negative, 34.19% neutral, and 26.11% positive sentiment.

- **Customer Service:** Predominantly neutral (43.93%), with 32.62% negative and 23.45% positive sentiment.

- **Environmental Impact:** Relatively positive, with 34.14% positive sentiment, 48.64% negative, and 17.22% neutral.

- **Infrastructure:** Mixed sentiment, with 36.51% negative, 38.17% neutral, and 25.32% positive sentiment.

- **Reliability:** Criticized heavily, with 61.03% negative sentiment, 21.09% neutral, and 17.88% positive sentiment.

Table 4.3 and Figure 4.3 show the sentiment distribution by aspects.

Table 4.3: Sentiment by Aspect.

Aspect	**Positive (%)**	**Neutral (%)**	**Negative (%)**
Comfort	44.38	39.12	16.49
Environmental Impact	34.14	17.22	48.64
Convenience	26.11	34.19	39.71
Infrastructure	25.32	38.17	36.51
Customer Service	23.45	43.93	32.62
Reliability	17.88	21.09	61.03
Cost	15.92	21.37	62.70
Safety	11.36	17.05	71.58

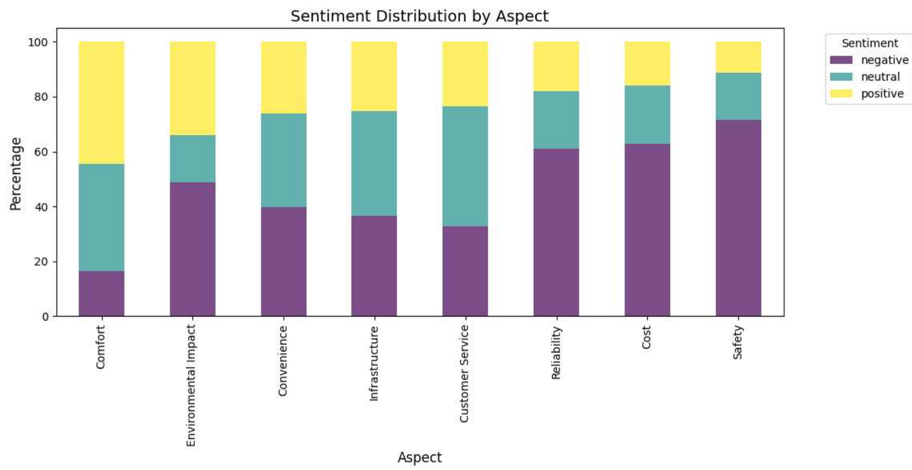

Figure 4.3: Sentiment Distribution by Aspects.

Safety and cost were the most negatively perceived aspects across all modes, indicating key public concerns. In contrast, comfort stood out as the most positively rated aspect, especially for walking and cycling.

4.2.4 Monthly Sentiment Trends

This subsection presents the overall sentiment trends across the entire dataset over time, using the Net Sentiment Score (NSS) to simplify the visualization. NSS combines positive, neutral, and negative sentiments into a single score, ranging from -1 (entirely negative) to +1 (entirely positive).

Key observations include:

- **Sentiment Peaks and Valleys:** Significant fluctuations in NSS are observed, with notable negative sentiment drops during late 2022 and late 2023, reflecting periods of public dissatisfaction.

- **Improved Sentiments:** Periods of increased NSS indicate improvements in public perception, especially in early 2023 and mid-2024.

Figure 4.4 shows the NSS trends month by month, providing a concise view of sentiment dynamics.

The NSS revealed notable negative sentiment dips during late 2022 and late 2023, possibly linked to service disruptions or fare increases. Periods of improved NSS, such as early 2023 and

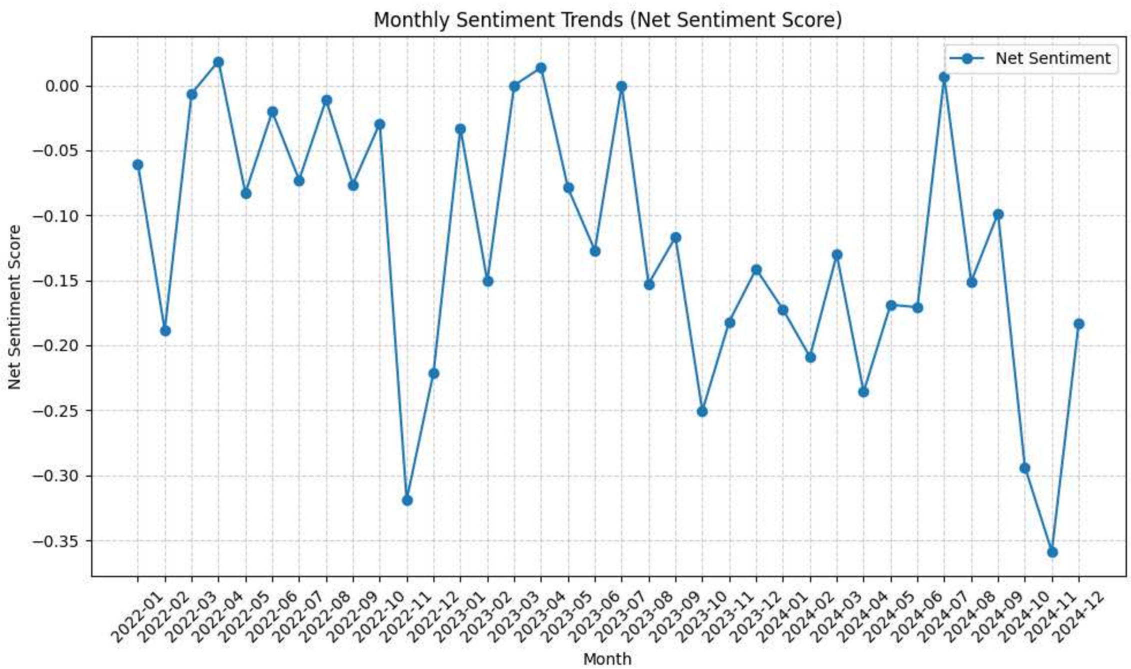

Figure 4.4: Monthly Net Sentiment Trends Across the Entire Dataset.

mid-2024, indicate positive public responses to seasonal changes or potential service improvements.

4.2.5 Sentiment by Aspect for Each Mode

The sentiment distribution by aspects for each transportation mode highlights the following key observations:

- **Bus:** Safety received the most negative sentiment (70.00%), followed by reliability (68.92% negative sentiment).
- **Walking:** The most positive aspect is cost (60.00% positive), and comfort is also rated highly (55.40% positive).
- **Bicycle:** Comfort (66.82% positive) and environmental impact (61.70% positive) are the most appreciated aspects.
- **Car:** Safety received the most negative sentiment (76.98%), followed by environmental impact (66.15% negative sentiment).

- **Train:** Safety had the highest negative sentiment (81.22%), while environmental impact is the most positive aspect (40.74% positive).

- **Ride-share:** Cost received the most negative sentiment (76.79%), followed by safety (90.74% negative sentiment).

Figures for sentiment distribution by aspects for each mode are available for reference in the appendix (Figures A.1 to A.6).

Detailed tables summarizing sentiment distribution by aspect for each mode are included in the appendix (Tables A.1 to A.6).

Safety consistently received the most negative sentiment across all modes, with the highest dissatisfaction observed in buses and trains. Cost also emerged as a significant concern, particularly for ride-shares and cars. Walking stood out with the most positive sentiment for cost and comfort, reflecting its accessibility and affordability.

4.2.6 Sentiment by Aspect for Each Mode Over Time

This section examines the Net Sentiment Score (NSS) trends for various aspects across different transportation modes over time. The NSS provides a single metric reflecting the balance of positive and negative sentiment while accounting for the total sentiment occurrences. Key observations for each mode and aspect are outlined below.

Mode: Bicycle

- **Comfort:** NSS is consistently positive, with significant peaks in 2024-09 and 2024-10 (1.00). A sharp drop is observed in 2024-08 (-0.50).

- **Convenience:** Mixed sentiment trends, fluctuating between positive and negative. NSS dropped sharply in late 2024 to -1.00 for November and December.

- **Cost:** Predominantly negative, reflecting dissatisfaction, though 2024-03 shows a notable positive spike (1.00).

- **Safety:** NSS trends are mostly negative, with occasional spikes such as 2024-07 (1.00).

Mode: Bus

- **Comfort:** Generally negative NSS, with minor positive deviations, such as 2024-11 (0.50).
- **Convenience:** Persistent negative sentiment dominates, with extreme lows in 2024-01 and 2024-11 (-1.00).
- **Cost:** Strong dissatisfaction is reflected in consistently negative NSS values. Positive spikes appear in 2024-02 and 2024-03 (1.00).
- **Safety:** Overwhelmingly negative NSS, with values often at -1.00.

Mode: Car

- **Comfort:** NSS trends are balanced, alternating between positive (2024-04 at 0.73) and negative.
- **Cost:** Consistently negative sentiment indicates dissatisfaction.
- **Safety:** NSS remains persistently negative, showing strong dissatisfaction.

Mode: Ride-Share

- **Comfort:** Mixed NSS trends, with positive peaks such as 2022-02 (0.75) and sharp dips like 2023-08 (-0.67).
- **Cost:** Strong dissatisfaction is evident, with NSS frequently at -1.00 across multiple months.
- **Reliability:** NSS shows consistently negative trends, reflecting dissatisfaction.
- **Safety:** Overwhelmingly negative sentiment dominates.

Mode: Train

- **Comfort:** Mixed NSS trends, with minor positive peaks such as 2024-10 (0.38).
- **Cost:** Predominantly negative sentiment with rare positive spikes, such as 2024-05 (1.00).
- **Safety:** NSS trends show consistently negative sentiment.

Mode: Walking

- **Comfort:** NSS is consistently positive, reflecting high satisfaction.
- **Convenience:** Alternating sentiment, with sharp dips (2023-02 at -1.00) and peaks (2023-05 at 1.00).
- **Cost:** Limited data shows predominantly positive NSS, indicating satisfaction.

4.3 Mode Trends Over Time

The number of tweets about each mode over time is presented in Figure 4.5.

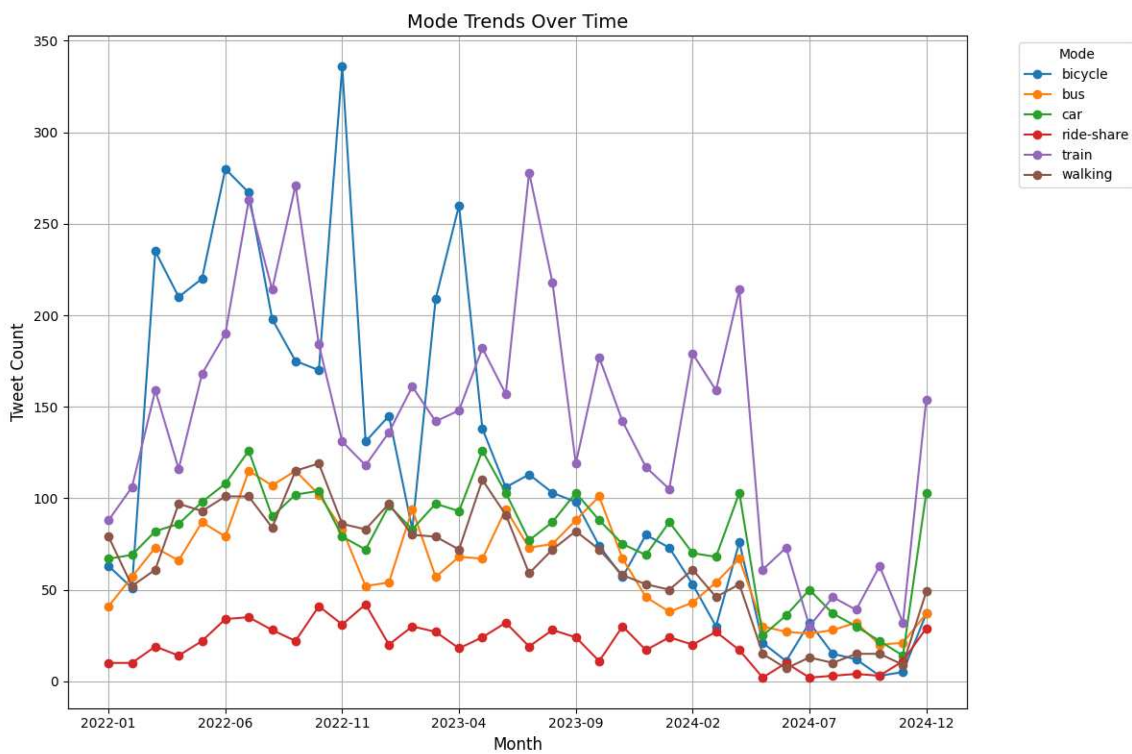

Figure 4.5: Mode Trends Over Time.

4.4 Aspect Distribution within Modes

The distribution of aspects within different transportation modes highlights the varying focus areas for each mode without considering the sentiment. Key observations include:

- **Bicycle:** The dominant aspect is *Infrastructure*, accounting for 47.41% of the focus, reflecting concerns about cycling infrastructure. Other significant aspects include *Convenience* (16.31%) and *Comfort* (15.47%).

- **Bus:** *Reliability* is the most prominent aspect, comprising 24.65% of the total focus, followed by *Comfort* (21.02%). *Infrastructure* (14.32%) is also an important aspect for buses.

- **Car:** The aspect distribution for cars shows a significant focus on *Safety* (20.14%) and *Infrastructure* (24.74%). *Cost* also plays a critical role, making up 11.40% of the focus.

- **Ride-share:** The highest focus is on *Comfort* (28.51%), followed by *Convenience* (17.16%) and *Cost* (15.14%). *Reliability* also holds substantial weight at 19.86%.

- **Train:** For trains, *Comfort* dominates the aspect distribution with 27.92%, followed by *Infrastructure* (22.32%) and *Reliability* (12.33%). Safety is a smaller concern compared to other modes at 9.22%.

The detailed distribution of aspects across modes is shown in Table 4.4, and Figure 4.6 provides a visual representation.

Table 4.4: Aspect Distribution within Each Mode (Percentage).

Aspect	Bicycle	Bus	Car	Ride-share	Train
Comfort	15.47	21.02	14.09	28.51	27.92
Convenience	16.31	14.54	11.82	17.16	15.88
Cost	1.37	7.84	11.40	15.14	4.61
Customer Service	7.05	7.27	7.89	7.70	6.67
Environmental Impact	2.25	1.18	4.60	0.68	1.05
Infrastructure	47.41	14.32	24.74	3.65	22.32
Reliability	1.13	24.65	5.31	19.86	12.33
Safety	9.02	9.19	20.14	7.30	9.22

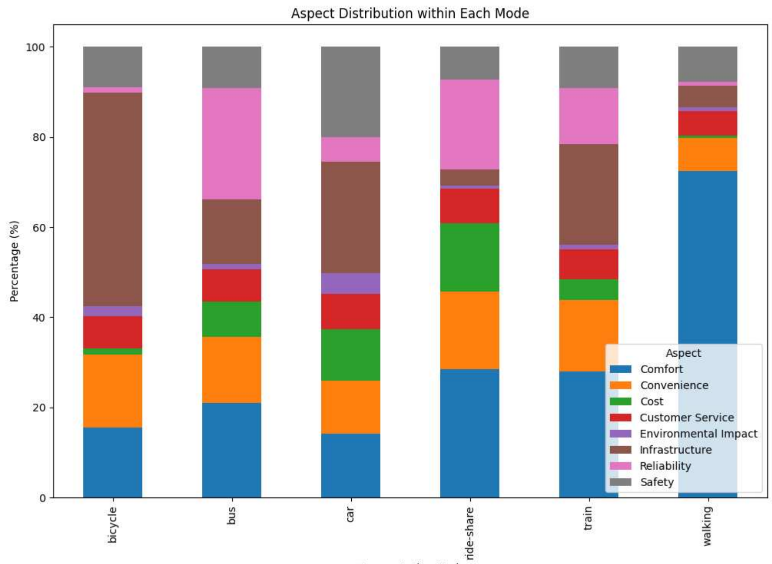

Figure 4.6: Aspect Distribution within Each Mode.

The focus on infrastructure for bicycles and reliability for buses underscores public concerns in these areas. Safety emerged as a significant focus for cars, while comfort dominated discussions for ride-shares and trains, suggesting opportunities for targeted improvements.

4.5 Location-based analysis

As mentioned in Section 3.4, a mapping was performed on the locations to convert them into five standard categories: ”Montreal”, ”South Shore”, ”North Shore”, ”West Island”, ”Off-Island Areas”, and ”Others”. In this section, we will analyze the tweets based on these standardized locations. As shown in figure 4.7, the greatest number of tweets belong to ”Montreal” with 71.6%, followed by the ”North Shore” in second position with 11.3%.

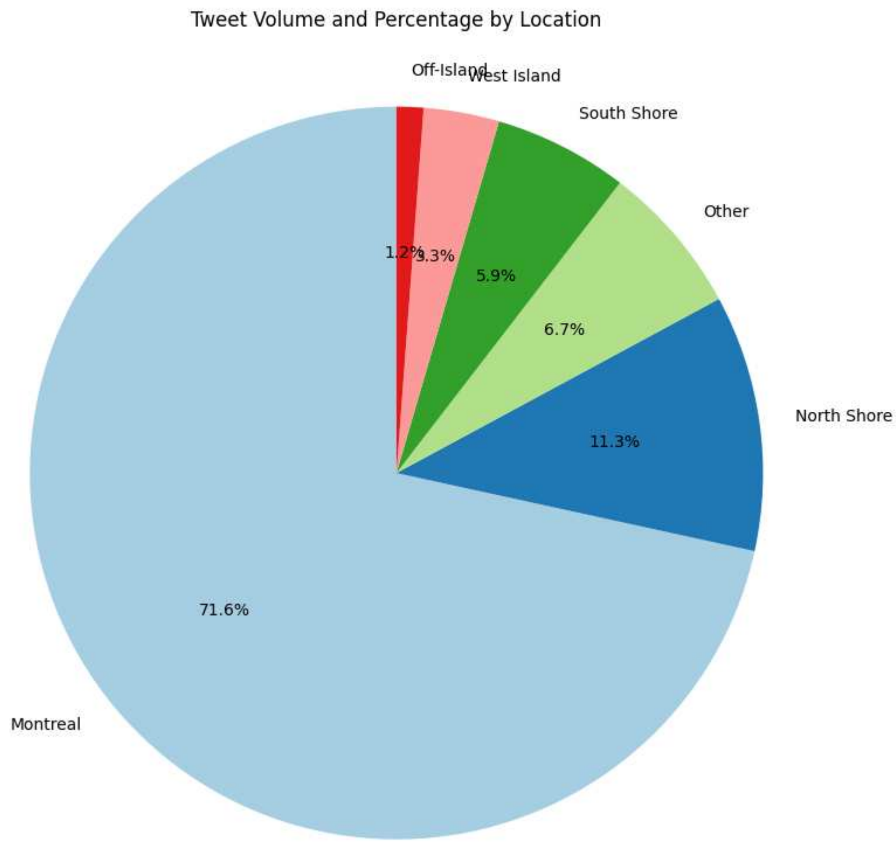

Figure 4.7: Tweet Volume by Location.

4.5.1 Aspect and Mode Distribution within Location

As shown in Figures 4.9 and 4.8, most of the tweets are related to Montreal. The modes of transport most frequently discussed in Montreal are Train and Bicycle, while Comfort and Infrastructure are the most discussed aspects. In the North Shore, Bicycle and Car dominate as the primary modes of transport, with Infrastructure and Comfort being the main aspects. The tables 4.6 and 4.5 show the numbers for the rest of the locations.

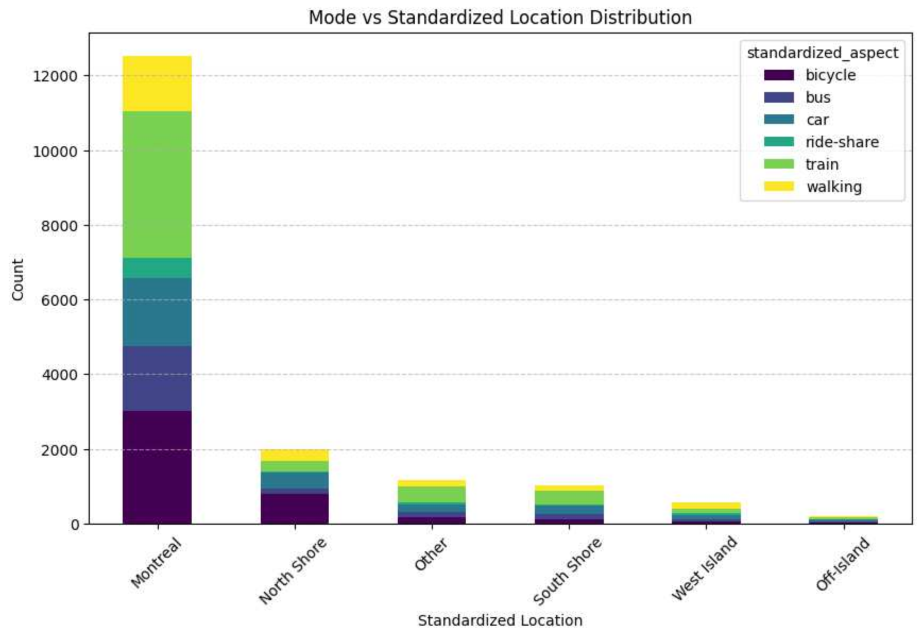

Figure 4.8: Mode Distribution within Each Location.

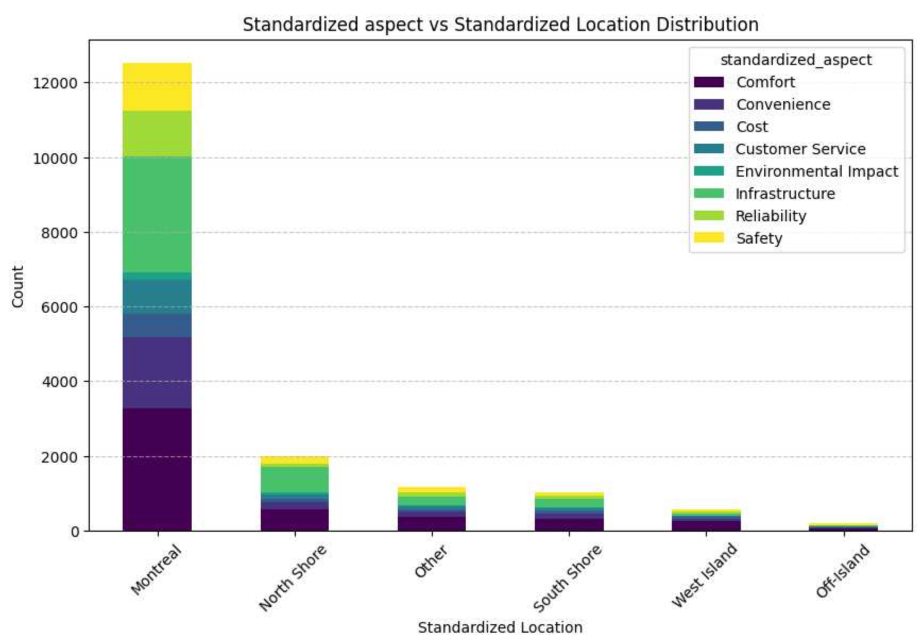

Figure 4.9: Aspect Distribution within Each Location.

Standardized Location	Bicycle	Bus	Car	Ride-share	Train	Walking
Montreal	3023	1731	1811	555	3930	1475
North Shore	787	166	406	37	278	304
Other	167	154	190	55	417	185
South Shore	122	136	233	35	354	157
West Island	55	64	116	56	118	174
Off-Island	16	33	69	2	43	44

Table 4.5: Mode of Transportation vs. Location Distribution

Standardized Location	Comfort	Convenience	Cost	Customer Service	Environmental Impact	Infrastructure	Reliability	Safety
Montreal	3279	1908	610	920	198	3107	1209	1294
North Shore	578	193	86	106	72	661	95	187
Other	373	129	80	84	14	224	115	149
South Shore	310	149	86	64	27	213	77	111
West Island	262	63	39	27	15	61	57	59
Off-Island	60	21	16	10	5	23	7	65

Table 4.6: Aspect Distribution by Location

4.5.2 Sentiment Distribution within Location

As shown in figure 4.10, The sentiment analysis reveals notable differences across standardized locations. Montreal shows a higher proportion of negative sentiment (39.86%) compared to positive (26.41%), with a neutral sentiment of 33.73%. This trend is mirrored in the ”Off-Island” and ”South Shore” regions, both displaying similarly high negative sentiment percentages.

In contrast, the West Island stands out with a relatively higher positive sentiment (36.88%) compared to its negative sentiment (37.56%), making it the most balanced region. The ”North Shore” also shows a more evenly distributed sentiment, with positive sentiment (35.49%) slightly exceeding negative (32.66%). Table 4.7 presents the remaining data.

The Net Sentiment Score (NSS) values, as shown in table 4.8, further highlight these trends. Montreal, Off-Island, and South Shore all have an NSS of -0.13, indicating a stronger lean toward negativity. Conversely, the North Shore has a positive NSS of 0.03, suggesting a slight positive sentiment balance. The West Island, with an NSS of -0.01, appears to have a near-neutral sentiment distribution.

Overall, the data suggests that Montreal and surrounding regions experience more negativity in transportation-related tweets, while the North Shore and West Island demonstrate a more balanced or slightly positive outlook.

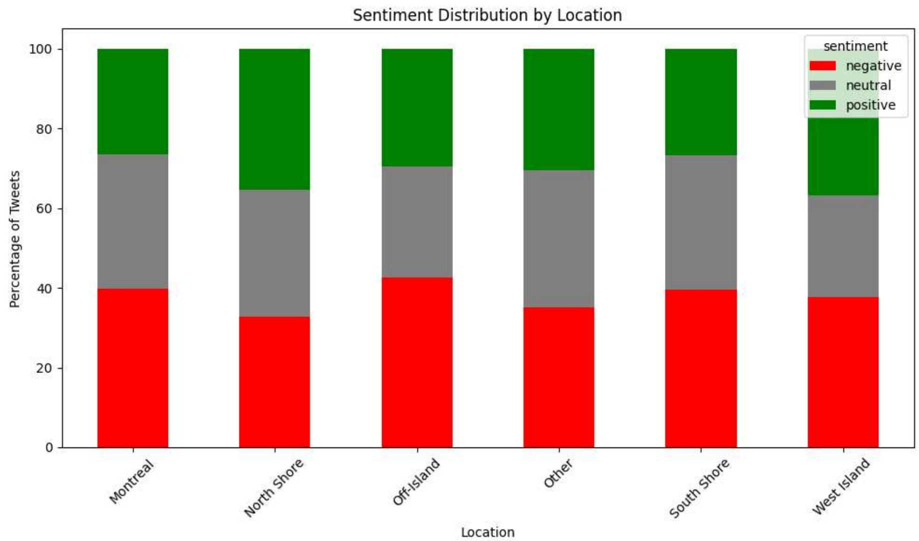

Figure 4.10: Sentiment Distribution across Standardized Locations.

Standardized Location	Negative (%)	Neutral (%)	Positive (%)
Montreal	39.86	33.73	26.41
North Shore	32.66	31.85	35.49
Off-Island	42.51	28.02	29.47
Other	35.10	34.33	30.57
South Shore	39.44	33.75	26.81
West Island	37.56	25.56	36.88

Table 4.7: Sentiment Distribution Across Standardized Locations

Standardized Location	NSS
Montreal	-0.13
North Shore	0.03
Off-Island	-0.13
Other	-0.05
South Shore	-0.13
West Island	-0.01

Table 4.8: Net Sentiment Score (NSS) Across Standardized Locations

4.5.3 Mode and Mode-Aspect Sentiment Analysis by Location

Table 4.9 summarizes the refined results of the sentiment analysis by location. Notably, the minimum-group-size filter reduces the prevalence of single-tweet categories. Key findings by area are as follows:

- **Montreal:** Walking yields the highest positive sentiment (47.73%), while the car is the most negatively perceived (53.12%). The mode-aspect combination of *bicycle:Comfort* attains 69.71% positive sentiment, whereas *ride-share:Safety* reaches 90.48% negative.
- **North Shore:** Bicycles register the strongest positive sentiment (51.21%), whereas cars lead in negative sentiment (60.10%). The *train:Environmental Impact* aspect garners 70.00% positive sentiment, while *bicycle:Cost* stands at 87.50% negative.
- **South Shore:** Walking is most positively viewed (47.13%), whereas ride-share shows the highest negative sentiment (62.86%). *bicycle:Comfort* ranks as the most positive mode-aspect (63.16%), and *ride-share:Cost* hits 100.00% negative.
- **West Island:** Walking again prevails in positive sentiment (71.84%), whereas buses hold the highest negative share (64.06%). The combination *walking:Comfort* achieves 78.57% positive, while *car:Environmental Impact* is entirely negative (100.00%).
- **Off-Island:** Walking dominates positively at 61.36%, in contrast to cars at 73.91% negative. *bicycle:Comfort* receives 85.71% positive sentiment, whereas *train:Safety* shows 100.00% negative sentiment.
- **Other:** Walking secures the strongest positive sentiment (55.14%), while buses lead in negative sentiment (49.35%). *bicycle:Comfort* earns 64.15% positive sentiment, but *car:Safety* is notably negative (76.47%).

4.6 Summary

This chapter provided results from the sentiment analysis of transportation-related tweets. The findings cover overall trends, mode-specific insights, and temporal changes, supported by data visualizations and tables for better understanding.

Location	Category	Mode/Aspect	Sentiment (%)
Montreal	Most Positive Mode	Walking	47.73
	Most Negative Mode	Car	53.12
	Most Positive Mode-Aspect	Bicycle:Comfort	69.71
	Most Negative Mode-Aspect	Ride-share:Safety	90.48
North Shore	Most Positive Mode	Bicycle	51.21
	Most Negative Mode	Car	60.10
	Most Positive Mode-Aspect	Train:Environmental Impact	70.00
	Most Negative Mode-Aspect	Bicycle:Cost	87.50
South Shore	Most Positive Mode	Walking	47.13
	Most Negative Mode	Ride-share	62.86
	Most Positive Mode-Aspect	Bicycle:Comfort	63.16
	Most Negative Mode-Aspect	Ride-share:Cost	100.00
West Island	Most Positive Mode	Walking	71.84
	Most Negative Mode	Bus	64.06
	Most Positive Mode-Aspect	Walking:Comfort	78.57
	Most Negative Mode-Aspect	Car:Environmental Impact	100.00
Off-Island	Most Positive Mode	Walking	61.36
	Most Negative Mode	Car	73.91
	Most Positive Mode-Aspect	Bicycle:Comfort	85.71
	Most Negative Mode-Aspect	Train:Safety	100.00
Other	Most Positive Mode	Walking	55.14
	Most Negative Mode	Bus	49.35
	Most Positive Mode-Aspect	Bicycle:Comfort	64.15
	Most Negative Mode-Aspect	Car:Safety	76.47

Table 4.9: Mode and Mode-Aspect Sentiment Analysis by Location

Chapter 5

Discussion

5.1 Introduction

This chapter interprets the results presented in Chapter 4 within the context of existing literature and the research objectives outlined in Chapter 1. The discussion emphasizes the implications of the findings for urban transportation planning and explores how the methodology addressed challenges in sentiment analysis, especially in multilingual contexts.

5.2 Key Insights and Implications

5.2.1 Overall Sentiment Trends

The dominance of negative sentiment (38.66%) aligns with prior studies highlighting public dissatisfaction with urban transportation services. Issues such as cost and safety reflect systemic challenges that have been previously documented. The high proportion of neutral tweets suggests an opportunity for urban planners to engage more deeply with public discourse to better understand nuanced opinions.

5.2.2 Mode-Specific Sentiment Analysis

Walking emerged as the most positively rated mode, reflecting its accessibility and affordability. This finding corroborates existing research that promotes walking as a sustainable and cost-effective mode of transportation. Conversely, the high levels of negative sentiment for buses and cars, driven by concerns over congestion, cost, and reliability, underscore the need for targeted interventions in these areas.

5.2.3 Aspect-Based Sentiment Analysis

Safety and cost emerged as the most negatively perceived aspects, consistent with global trends in urban transportation dissatisfaction. The consistently positive sentiment for comfort, particularly in walking and cycling, suggests that improvements in these areas may offer significant gains in public satisfaction.

5.2.4 Temporal Sentiment Trends

Temporal analysis revealed key periods of public dissatisfaction, such as late 2022 and late 2023, which may correlate with service disruptions or fare increases. Conversely, positive sentiment peaks in early 2023 and mid-2024 suggest successful interventions or seasonal improvements. These trends emphasize the importance of timely and transparent communication during periods of disruption.

5.2.5 Location-Based Sentiment Analysis

Figure 4.7 shows that most tweets originate from Montreal (71.6%), reflecting its large population and extensive transportation network. As seen in Figure 4.10, Montreal also has one of the highest proportions of negative sentiment (39.86%), often linked to concerns about congestion, cost, and safety. By contrast, the West Island stands out for having comparatively higher positive sentiment (36.88%), suggesting that lower congestion or better infrastructure may contribute to more favorable views there.

The North Shore, which accounts for 11.3% of all tweets, exhibits nearly balanced sentiment: 35.49% positive and 32.66% negative. This balance could reflect a moderate acceptance of alternatives like bicycles and trains, offset by worries about reliability or cost. Meanwhile, the South Shore's sentiment pattern largely mirrors Montreal's, emphasizing challenges such as commute distances and limited public transit coverage that lead to higher negative perceptions.

Table 4.9 provides additional insights on how transportation modes and mode-aspect combinations vary by location. For example, in Montreal, *walking* is viewed most positively (47.73%), whereas *cars* draw the highest negative sentiment (53.12%). In the West Island, *walking* again ranks highest in positive sentiment (71.84%), compared with the *bus* at 64.06% negative. These findings highlight how geographic variations, such as infrastructure quality or population density, shape different attitudes toward the same modes of transportation. Understanding these local differences can help planners and policymakers focus on location-specific improvements—for instance, addressing high-cost or congestion issues in Montreal's core, while preserving the factors that contribute to the West Island's more positive outlook.

5.2.6 Methodological Contributions

The use of GPT-4o demonstrated the viability of LLMs for nuanced sentiment analysis in multilingual contexts, addressing limitations in traditional methods. This study introduces a structured bilingual sentiment analysis pipeline leveraging GPT-4o's multilingual capabilities, integrating automated aspect classification and sentiment tagging to ensure high accuracy in English and French tweets without requiring translation.

Additionally, a standardized aspect categorization system was developed to enhance sentiment consistency across datasets. The Net Sentiment Score (NSS) provided an effective mechanism for summarizing complex sentiment trends into a single, interpretable metric, simplifying temporal and comparative analysis.

To enhance computational efficiency, batch processing of LLM API calls, error handling strategies, and optimized data storage were implemented. The integration of retry mechanisms ensured uninterrupted data processing, making large-scale sentiment analysis feasible for real-time applications.

5.3 Policy Implications

The findings offer actionable insights for policymakers:

- Addressing public concerns about safety, cost, and reliability should be a priority for improving satisfaction.
- Increasing engagement with public feedback, particularly during periods of negative sentiment spikes, can help build trust.
- Leveraging positive aspects, such as comfort and accessibility, can enhance the appeal of walking and cycling.
- The location-based analysis reveals that Montreal, South Shore, and Off-Island areas exhibit comparatively higher negative sentiment, whereas North Shore and West Island present a more balanced or slightly positive outlook. Policymakers can:
    - Focus on mitigating the high cost of travel and congestion in Montreal's core, where negative sentiment is prevalent.
    - Reinforce the successful modes (e.g., walking, cycling) in West Island and North Shore, while improving bus reliability or addressing car-related grievances in regions where they rank poorly.
    - Implement targeted outreach or public engagement in negatively skewed areas, aiming to identify local pain points more precisely (e.g., fare policies, limited routes) and tailor solutions accordingly.

5.4 Challenges and Limitations

While this study provided valuable insights, applying LLMs to sentiment analysis presents several computational and methodological challenges:

- **Computational Costs and API Constraints:** Conducting large-scale sentiment analysis using an API-based LLM incurs latency and financial costs.
- **Lack of Model Explainability:** Unlike traditional sentiment classification models, LLMs function as black-box systems, making it difficult to interpret the rationale behind sentiment classifications.

- **Social Media Bias:** The dataset may not represent the broader population due to demographic biases in social media usage.
- **Model Bias:** GPT-4o's sentiment and aspect classification may reflect biases inherent in its training data.
- **Text-Only Analysis:** Excluding multimedia data, such as images and videos, limits the scope of the analysis.
- **Location Data Quality:** The accuracy of location-based findings can be influenced by incomplete geotagging or ambiguous user-reported locations, potentially skewing regional sentiment estimates.

5.5 Conclusion

The discussion highlights the study's contributions to understanding public sentiment on urban transportation in Montreal. These findings pave the way for more inclusive and data-driven policy interventions.

Chapter 6

Conclusion

6.1 Summary of Findings

This study analyzed public sentiment on urban transportation in Montreal using multilingual social media data. Key findings include:

- Safety and cost were identified as the most significant public concerns.
- Walking and cycling received the most positive feedback, emphasizing the importance of sustainable modes of transport.
- Temporal sentiment analysis revealed key periods of public dissatisfaction and improvement.
- Location-based analysis indicated that Montreal, South Shore, and Off-Island areas exhibited more negative sentiment overall, whereas the North Shore and West Island displayed a more balanced or slightly positive outlook.

6.2 Contributions

6.2.1 Methodological Contributions

- This study introduces a structured bilingual sentiment analysis pipeline leveraging GPT-4o's multilingual capabilities. The pipeline integrates automated aspect classification and sentiment tagging, ensuring high accuracy in English and French tweets without requiring translation.

- Additionally, a standardized aspect categorization system was developed to enhance sentiment consistency across datasets.

- This research introduces a Net Sentiment Score (NSS)-based sentiment trend analysis, enabling real-time tracking of public perception in urban transportation. The integration of temporal sentiment visualization and mode/aspect-specific NSS calculations provides an efficient mechanism for detecting shifts in public sentiment.

- To enhance computational efficiency, this study implemented batch processing of LLM API calls, error handling strategies, and optimized data storage. The integration of retry mechanisms ensured uninterrupted data processing, making large-scale sentiment analysis feasible for real-time applications.

6.2.2 Theoretical Contributions

The study provided novel insights into the public perception of urban transportation, contributing to the literature on aspect-based sentiment analysis in multilingual settings.

6.2.3 Practical Contributions

Actionable insights were identified for policymakers, highlighting areas for improvement in Montreal's transportation system, such as addressing safety and cost concerns.

6.3 Limitations and Future Work

6.3.1 Limitations

- **Computational Costs and API Constraints:** Conducting large-scale sentiment analysis using an API-based LLM incurs latency and financial costs.
- **Lack of Model Explainability:** LLMs function as black-box systems, making it difficult to interpret the rationale behind the classifications.
- **Model Bias:** GPT-4o's sentiment and aspect classification may reflect biases inherent in its training data.
- **Text-Only Analysis:** Excluding multimedia data, such as images and videos, limits the scope of the analysis.
- **Data Bias:** Social media data may not fully represent all demographics.

6.3.2 Future Work

- **Fine-tuning LLMs for domain-specific sentiment analysis:** Addressing domain adaptation challenges could improve accuracy in transportation discourse, particularly in sarcasm detection.
- **Incorporating multimodal data:** Future work could combine text, images, and audio sentiment analysis, allowing for a more holistic understanding of public transportation experiences.
- **Expanding Data Sources:** Incorporating surveys, real-time public feedback, and other social platforms could provide a more comprehensive view.
- **Evaluating alternative and local models:** Future studies could explore other LLMs, including open-source or locally deployed models, to reduce API dependency and improve adaptability to local context.

6.4 Final Overview: What This Research Found and Why It Matters

In an era where public trust in urban transportation is shaped not just by performance but by perception, this study set out to understand how people in Montreal feel about their daily commutes—through the lens of what they share on social media. Using over 17,000 tweets in both English and French, posted over a three-year period, the research used GPT-4o, one of today's most advanced language models, to analyze public sentiment toward buses, metros, ride-shares, walking, cycling, and more. The system automatically detected what each tweet was about, categorized it into a mode (e.g., bus), an aspect (e.g., safety or cost), and its sentiment (positive, neutral, or negative), all without manual translation or labeling.

The findings highlight a clear mismatch between public expectations and certain transportation services. Safety and cost emerged as the most concerning issues, driving negative sentiment across most modes, especially for ride-shares, cars, and buses. For example, over 90% of tweets about ride-share safety were negative, while train safety also received extremely poor ratings. In contrast, walking and biking received more positive sentiment, with walking especially praised for affordability and comfort. Over time, the data showed recurring drops in sentiment in late 2022 and 2023, likely related to winter service disruptions or fare increases, and occasional spikes in positivity, potentially due to seasonal changes or improvements in services.

This thesis demonstrates how AI can help cities tap into real-time, organic feedback from the public—without waiting for surveys or formal consultations. The approach is scalable, bilingual, and adaptable to other urban domains beyond transportation. For Montreal, the insights offered here point to clear opportunities for improvement: from targeting safety issues and reducing cost burdens to reinforcing what's already working well, like promoting walkability and cycling. In a time where policy must be agile and citizen-centered, this work shows how public opinion mining using tools like GPT-4o can offer a faster, smarter way forward.

Appendix A

Detailed Sentiment Distribution by Aspect for Each Mode

A.1 Bus

A.1.1 Table: Sentiment Distribution by Aspect for Bus

Table A.1: Sentiment Distribution by Aspect for Bus.

Aspect	**Negative (%)**	**Neutral (%)**	**Positive (%)**
Comfort	33.96	43.13	22.92
Convenience	46.99	21.99	31.02
Cost	65.36	15.08	19.55
Customer Service	46.99	34.94	18.07
Environmental Impact	62.96	14.81	22.22
Infrastructure	38.84	41.90	19.27
Reliability	68.92	19.18	11.90
Safety	70.00	20.95	9.05

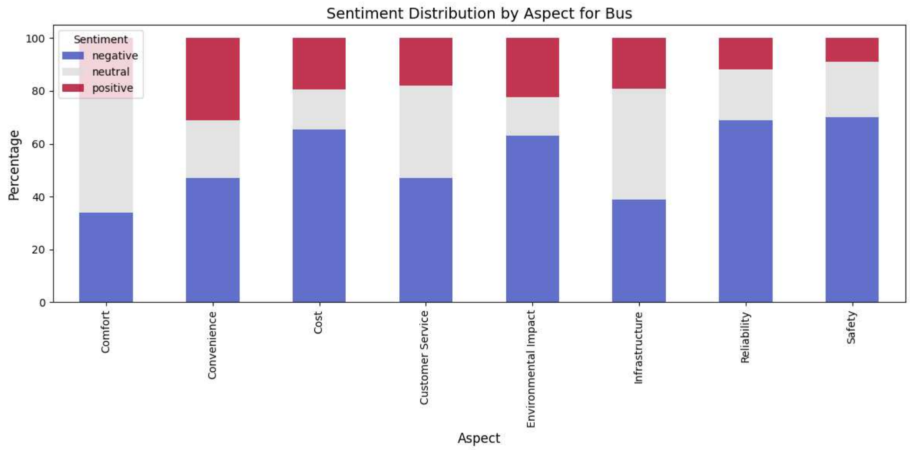

Figure A.1: Sentiment Distribution by Aspect for Bus.

Table A.2: Sentiment Distribution by Aspect for Walking.

Aspect	**Negative (%)**	**Neutral (%)**	**Positive (%)**
Comfort	6.44	38.16	55.40
Convenience	32.18	25.86	41.95
Cost	10.00	30.00	60.00
Customer Service	14.84	52.34	32.81
Environmental Impact	52.38	28.57	19.05
Infrastructure	40.18	28.57	31.25
Reliability	26.32	36.84	36.84
Safety	40.66	22.53	36.81

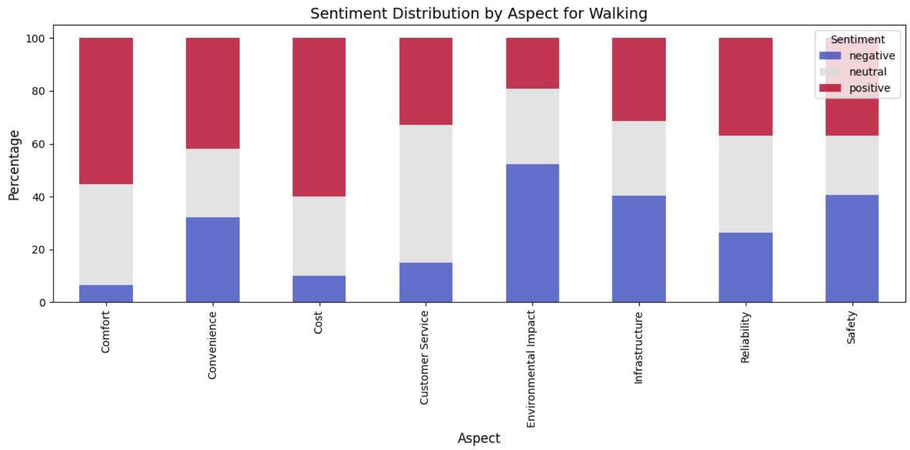

Figure A.2: Sentiment Distribution by Aspect for Walking.

Table A.3: Sentiment Distribution by Aspect for Bicycle.

Aspect	Negative (%)	Neutral (%)	Positive (%)
Comfort	3.72	29.46	66.82
Convenience	51.91	22.65	25.44
Cost	43.86	33.33	22.81
Customer Service	18.71	44.56	36.73
Environmental Impact	22.34	15.96	61.70
Infrastructure	25.14	41.63	33.23
Reliability	23.40	14.89	61.70
Safety	64.36	19.41	16.22

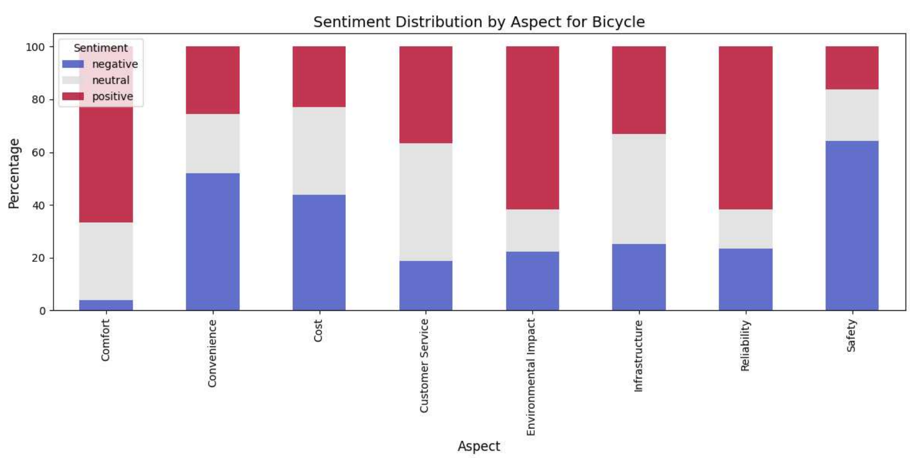

Figure A.3: Sentiment Distribution by Aspect for Bicycle.

A.1.2 Figure: Sentiment Distribution by Aspect for Bus

A.2 Walking

A.2.1 Table: Sentiment Distribution by Aspect for Walking

A.2.2 Figure: Sentiment Distribution by Aspect for Walking

Table A.4: Sentiment Distribution by Aspect for Car.

Aspect	Negative (%)	Neutral (%)	Positive (%)
Comfort	20.35	46.48	33.17
Convenience	38.62	42.22	19.16
Cost	60.56	27.02	12.42
Customer Service	42.15	45.29	12.56
Environmental Impact	66.15	17.69	16.15
Infrastructure	60.23	29.18	10.59
Reliability	43.33	27.33	29.33
Safety	76.98	18.28	4.75

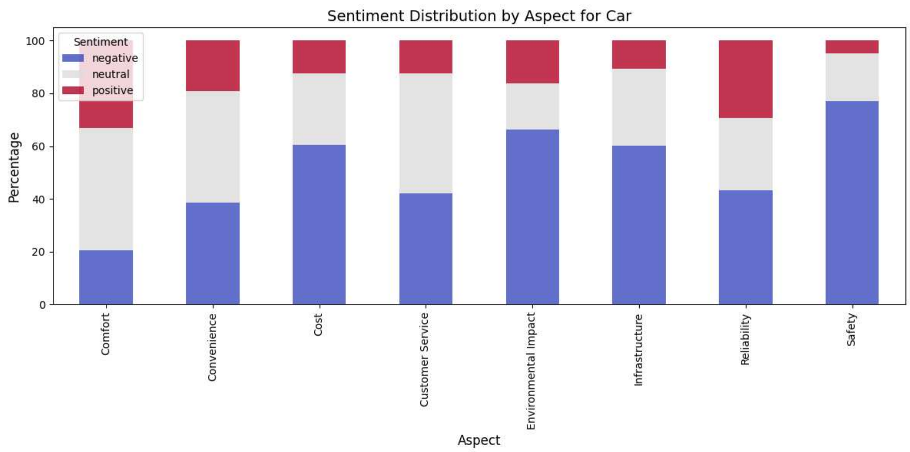

Figure A.4: Sentiment Distribution by Aspect for Car.

Table A.5: Sentiment Distribution by Aspect for Train.

Aspect	Negative (%)	Neutral (%)	Positive (%)
Comfort	26.13	40.21	33.66
Convenience	29.17	45.83	25.00
Cost	63.71	18.99	17.30
Customer Service	36.15	43.73	20.12
Environmental Impact	44.44	14.81	40.74
Infrastructure	40.71	37.31	21.97
Reliability	61.83	19.87	18.30
Safety	81.22	11.18	7.59

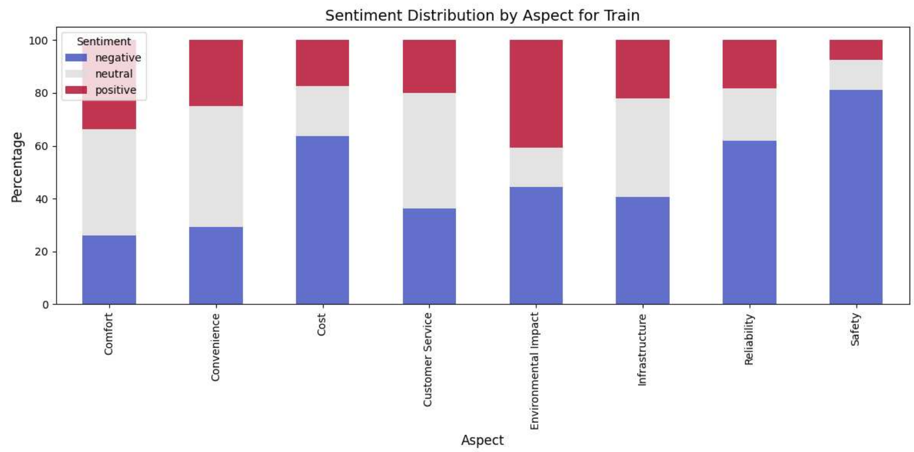

Figure A.5: Sentiment Distribution by Aspect for Train.

Table A.6: Sentiment Distribution by Aspect for Ride-share.

Aspect	Negative (%)	Neutral (%)	Positive (%)
Comfort	23.70	45.97	30.33
Convenience	36.22	43.31	20.47
Cost	76.79	13.39	9.82
Customer Service	43.86	43.86	12.28
Environmental Impact	40.00	20.00	40.00
Infrastructure	33.33	48.15	18.52
Reliability	61.90	27.21	10.88
Safety	90.74	5.56	3.70

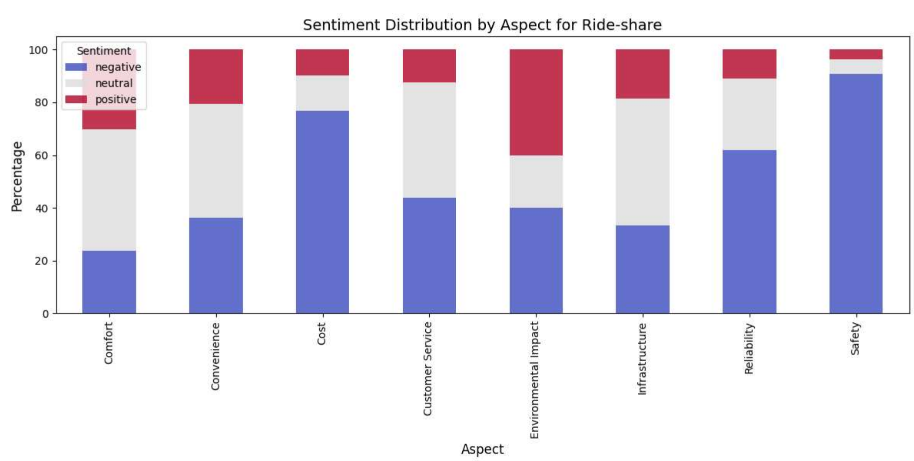

Figure A.6: Sentiment Distribution by Aspect for Ride-share.

References

Boiy, E., & Moens, M.-F. (2009). Machine learning for social media sentiment analysis. *Journal of the American Society for Information Science and Technology*, *60*(7), 1358–1370.

Bollen, J., Mao, H., & Zeng, X.-J. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, *2*(1), 1–8.

Bouazizi, M., & Ohtsuki, T. (2017). A pattern-based approach for multi-class sentiment analysis in twitter. *IEEE Access*, *5*, 20617-20639. doi: 10.1109/ACCESS.2017.2740982

Brown, T., et al. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, *33*, 1877–1901.

Burnap, P., & Williams, M. L. (2015). Detecting tension in online communities with computational twitter analysis. *Technological Forecasting and Social Change*, *95*, 96–108. Retrieved from https://www.sciencedirect.com/science/article/pii/S0040162513000899 doi: 10.1016/j.techfore.2013.04.013

Cambria, E., & White, B. (2017). Affective computing and sentiment analysis. *IEEE Intelligent Systems*, *32*(2), 102–107.

Conneau, A., et al. (2020). Unsupervised cross-lingual representation learning at scale. *arXiv preprint arXiv:1911.02116*.

Devlin, J., et al. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. *NAACL-HLT*, *1*, 4171–4186.

Mahmud, J., et al. (2012). Tweet analysis for real-time event detection and earthquake reporting system. *PloS one*, *7*(4), e36127.

Ortigosa, A., Martín, J. M., & Carro, R. M. (2014). Sentiment analysis in facebook and its application to e-learning. *Computers in Human Behavior*, *31*, 527–541.

Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. , 1320–1326.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, *2*(1-2), 1–135.

Pontiki, M., Galanis, D., Pavlopoulos, J., et al. (2016). Semeval-2016 task 5: Aspect-based sentiment analysis. , 19–30.

Shahriar, S., Lund, B., Mannuru, N. R., et al. (2024). Putting gpt-4o to the sword: A comprehensive evaluation of language, vision, speech, and multimodal proficiency. *arXiv preprint arXiv:2407.09519*.

Taboada, M., Brooke, J., Tofiloski, M., et al. (2011). Lexicon-based methods for sentiment analysis. *Computational linguistics*, *37*(2), 267–307.

Zheng, Y., Capra, L., Wolfson, O., & Yang, H. (2014). Urban computing: Concepts, methodologies, and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *5*(3), 38. doi: 10.1145/2629592