

Louvain School of Management

How has the global chip shortage crisis from 2020-2023 influenced brand popularity across affected sectors and what was the public opinion toward it?

Author(s): Marco Giuliano
Supervisor(s): Vincenzo Verardi
Academic year 2024-2025
Thesis for the Master [120] in Management
Sciences: Consumer insights
Daytime schedule

During the preparation of this master's thesis, the author utilized ChatGPT and DeepL for the following purposes:

1.Translation: DeepL has been used to translate some thoughts more quickly and efficiently.

2.Reformulation: DeepL (Write version) was also often used to reformulate the text in sustained, academic English. The text was subsequently re-read by myself and adapted to my satisfaction.

3.Data interpretation and classification: In the context of this study, ChatGPT was utilized to facilitate the interpretation and coding of qualitative variables, including “comm_strat” (communication strategy), “SCV_level” (supply chain visibility), and “brand”, based on documentary sources. For the “buzz_score” variable, ChatGPT helped contextualize Reddit engagement metrics. All classifications and interpretations were made within a predefined analytical framework and were critically reviewed and validated by the author.

After using ChatGPT, the author diligently reviewed and edited the content produced by the tool. I take full responsibility for the final content presented in this thesis.

By signing this declaration, I affirm that the content of this master's thesis reflects my original work, augmented by the responsible use of AI.

01/05/2025

Marco Giuliano

Acknowledgements:

I would like to thank all the people who helped me write this thesis, whether directly or indirectly.

I would like to express my gratitude to my supervisor, Mr. Vincenzo Verardi, for his invaluable guidance and support during the composition of this thesis. His insightful comments and judicious advice have considerably enhanced the quality of my work.

In addition, I would like to express my gratitude to my family and especially my father for their unwavering support and encouragement throughout the process of composing this master's thesis and throughout my academic journey.

I would also like to thank my friend, Thomas, who wishes to remain anonymous, for proofreading my thesis.

Table of Contents

1. Introduction	1
2. Literature Review	3
2.1. Topics within a specific area in literature	3
2.1.1. <i>Supply Chain Disruptions and the Global Chip Shortage</i>	3
2.1.2. <i>Impact on Brand Popularity and Consumer Perception</i>	3
2.1.3. <i>Social Media as a Tool for Sentiment Analysis</i>	3
2.2. Theoretical model or framework	4
2.2.1. <i>Supply Chain Visibility Framework</i>	4
2.2.2. <i>Rationale for Choosing the SCV Framework</i>	4
2.2.3. <i>Situational Crisis Communication Theory</i>	4
2.2.4. <i>Three-Stage Model of Crisis Management</i>	5
2.3. Hypotheses	5
3. Methodology and Data	7
3.1. Model	8
3.2. Hypotheses and Tests	9
3.3. Variables	10
3.3.1. <i>Dependent variable</i>	10
3.3.2. <i>Explanatory variables</i>	10
3.4. Sampling approach	11
3.4.1. <i>Archival data</i>	11
3.4.2. <i>How Orange Datamining works</i>	13
3.4.3. <i>VADER sentiment analysis</i>	14
3.4.4. <i>SentiART sentiment analysis</i>	15
4. Results	16
4.1. Hypothesis 1: The global semiconductor shortage negatively impacted brand popularity across affected sectors	17
4.1.1. <i>Results by source</i>	18
4.1.1.1. Wikipedia	18
4.1.1.2. The Guardian	18
4.1.1.3. NY Times	19
4.1.1.4. Reddit	20
4.2. Hypothesis 2: Public sentiment on social media platforms during the shortage was predominantly negative	21
4.2.1. <i>Results by source</i>	21
4.2.1.1. Reddit	21

4.3.	Hypothesis 3: Brands that effectively communicated their mitigation strategies were better able to maintain consumer trust	22
4.3.1.	<i>Results by source</i>	23
4.3.1.1.	Wikipedia.....	23
4.3.1.2.	The Guardian.....	24
4.3.1.3.	NY Times.....	25
4.3.1.4.	Reddit.....	26
4.4.	Hypothesis 4: Industries with greater supply chain visibility experienced less negative impact on brand popularity	27
4.4.1.	<i>Results by source</i>	28
4.4.1.1.	Wikipedia.....	28
4.4.1.2.	The Guardian.....	28
4.4.1.3.	NY Times.....	29
4.4.1.4.	Reddit.....	29
4.5.	Hypothesis 5: Sentiment analysis of social media and media sources can serve as an early indicator of brand popularity fluctuations during supply chain crises	30
4.5.1.	<i>Results by source</i>	30
4.5.1.1.	Wikipedia.....	30
4.5.1.2.	The Guardian.....	31
4.5.1.3.	NY Times.....	32
4.5.1.4.	Reddit.....	32
5.	Conclusion	34
5.1.	Key findings.....	34
5.2.	Recommendations for companies	36
5.3.	Limits of paper	36
5.4.	Suggestions for further research	37
6.	Bibliography	39
7.	Appendix	42
	Appendix A: Research Model	42
	Appendix B: Orange workflows	43
	<i>Data extraction and initial analysis</i>	43
	<i>Extensive data analysis and concatenation between 2020 and 2023</i>	44
	<i>Detailed description of widgets used</i>	45
	Appendix C: Results for the first hypothesis	46
	<i>Sentiment Analysis of Wikipedia Articles</i>	46
	<i>Sentiment Analysis of The Guardian Articles</i>	47
	<i>Sentiment Analysis of The New York Times Articles</i>	48
	<i>Sentiment Trends on Reddit</i>	49

Appendix D: Results for the second hypothesis	50
<i>Emotional Analysis on Reddit</i>	50
Appendix E: Results for the third hypothesis	51
<i>Communication Strategies and Sentiment on Wikipedia</i>	51
<i>Communication Strategies and Sentiment in The Guardian</i>	52
<i>Communication Strategies and Sentiment in The New York Times</i>	53
<i>Communication Strategies and Sentiment on Reddit</i>	54
Appendix F: Results for the fourth hypothesis	55
<i>Impact of Supply Chain Visibility on Wikipedia</i>	55
<i>Impact of Supply Chain Visibility in The Guardian</i>	56
<i>Impact of Supply Chain Visibility in The New York Times</i>	57
<i>Impact of Supply Chain Visibility on Reddit</i>	58
Appendix G: Results for the fifth hypothesis	59
<i>Sentiment-Popularity Correlations on Wikipedia</i>	59
<i>Sentiment-Popularity Correlations in The Guardian</i>	60
<i>Sentiment-Popularity Correlations in The New York Times</i>	61
<i>Sentiment-Popularity Correlations on Reddit</i>	62

1. Introduction

Crises and supply chain disruptions have always played a significant role in shaping global markets, influencing industries, businesses, and consumer behavior. Among recent crises, the global chip shortage from 2020 to 2023 stands out due to its widespread impact across multiple sectors, ranging from automotive and consumer electronics to healthcare and telecommunications. This shortage, precipitated by a confluence of pandemic-related disruptions, geopolitical tensions, and heightened demand for semiconductors, engendered production delays, financial losses, and shifts in consumer preferences. Given the scale and intricacy of this crisis, it is imperative to examine not only its economic ramifications but also its impact on brand popularity and public perception.

This study aims to understand how the global chip shortage crisis influenced brand popularity across the affected sectors and how the public reacted to this crisis through social media discourse. Specifically, the research will investigate shifts in consumer sentiment, brand loyalty, and the ways in which different companies managed their communication strategies in response to supply chain challenges. By analyzing public opinion through Reddit, a prominent online platform, this study seeks to uncover patterns in consumer reactions, whether they manifested as frustration, brand switching behavior, or increased awareness of supply chain vulnerabilities.

The research question of this thesis therefore comes naturally as:

“How has the global chip shortage crisis from 2020-2023 influenced brand popularity across affected sectors and what was the public opinion toward it?”

Addressing this research question is of both academic and practical significance. From an academic perspective, this study contributes to a multidisciplinary understanding of crisis management by integrating insights from marketing, consumer psychology, supply chain management, and media studies. The chip shortage represents a unique case where economic disruptions intersect with digital

consumer reactions, making it an ideal subject for exploring how crises influence brand perception in the digital age. This research will contribute to extant literature on crisis communication, consumer trust, and market resilience, offering novel perspectives on how companies navigate supply chain disruptions while maintaining their brand reputation.

This study adopts a data mining approach, emphasizing the extraction of insights from large-scale textual corpora rather than traditional hypothesis testing based on inferential statistics. Rather than aiming to generalize results to a broader population, the analysis is oriented toward the construction of evidence through multiple converging indicators across diverse sources, including Reddit, Wikipedia, and international news media.

The study employs sentiment analysis techniques to construct bundles of evidence that support or challenge the proposed hypotheses. This approach aligns with contemporary practices in consumer insights and crisis communication research, particularly in contexts where the complexity and unstructured nature of data impede the applicability of formal statistical inference.

On a practical level, the findings of this study will provide valuable insights for businesses, policymakers, and consumers. For companies operating in industries dependent on semiconductor supplies, understanding the effects of the chip shortage on brand perception can help refine future crisis management strategies.

The study will shed light on which brands effectively allayed consumer concerns and retained brand loyalty during the crisis, as well as those that encountered difficulties in adapting. In addition, a thorough examination of public sentiment will assist companies in aligning their communication and marketing strategies with consumer expectations, ensuring more effective responses during times of uncertainty.

2. Literature Review

2.1. Topics within a specific area in literature

2.1.1. Supply Chain Disruptions and the Global Chip Shortage

The global semiconductor chip shortage that transpired between 2020 and 2023 has been the subject of extensive analysis. Studies have identified the causes of this shortage to include the impact of the pandemic of 2019 novel coronavirus (SARS-CoV-2) on manufacturing hubs in East Asia, increased demand for consumer electronics, and geopolitical tensions affecting supply chains (Wassen et al., 2022).

This shortage has had a disruptive effect on various industries, including the automotive, consumer electronics, and healthcare sectors, resulting in production delays and financial losses (Xiong et al., 2024).

2.1.2. Impact on Brand Popularity and Consumer Perception

A body of research suggests a significant relationship between supply chain disruptions and brand resilience, as well as consumer perception. The effective management of such disruptions is therefore crucial for maintaining consumer trust and brand reputation (Holloway, 2024). However, the extant literature does not thoroughly address the specific impact of the chip shortage on brand popularity and public sentiment, thus highlighting the need for further investigation.

2.1.3. Social Media as a Tool for Sentiment Analysis

In the contemporary age, social media platforms have emerged as valuable resources for the analysis of public opinion during times of crisis. The employment of sophisticated text-mining techniques facilitates the extraction of consumer sentiments from prominent platforms such as Twitter and Reddit, thereby providing real-time insights into the prevailing public perception (Sheikhattar et al., 2022).

2.2. Theoretical model or framework

2.2.1. *Supply Chain Visibility Framework*

This study employs the Supply Chain Visibility (SCV) framework, which underscores the significance of transparency and the real-time exchange of information within supply networks. SCV is imperative for the expeditious identification of disruptions and the implementation of efficacious responses (Özkanlısoy, 2024).

2.2.2. *Rationale for Choosing the SCV Framework*

The SCV framework is particularly pertinent to this study, as it pertains to the necessity of transparency in supply chains, which is imperative for preserving consumer confidence during periods of disruption. In contrast to other models that may prioritize operational efficiency exclusively, the SCV framework incorporates consumer perception, rendering it apt for the analysis of the repercussions of the chip shortage on brand popularity (Cserhat, 2024) (Novotny, 2024).

2.2.3. *Situational Crisis Communication Theory*

The Situational Crisis Communication Theory (SCCT), developed by Coombs (2007), provides a robust framework for understanding how organizations should communicate during crises. SCCT proposes that the efficacy of a crisis response strategy is contingent upon the perceived level of responsibility attributed to the organization and the nature of the crisis itself. The model delineates between victim, accidental, and preventable crises and recommends customized communication strategies (e.g., denial, diminishment, rebuilding) for each scenario.

In the context of the global semiconductor shortage, SCCT offers a useful perspective for evaluating how brands adjusted their messaging based on their perceived culpability or control over the situation. It is evident that brands that proactively communicated their mitigation efforts aligned with SCCT's recommendations for rebuilding trust during high-responsibility crises. The theory is consistent with the SCV framework insofar as it emphasizes the communicative dimension of crisis

management, particularly with respect to shaping public perception and maintaining reputational resilience (Coombs, 2007).

2.2.4. Three-Stage Model of Crisis Management

The Three-Stage Model of Crisis Management conceptualizes crises as dynamic processes unfolding in three sequential phases: pre-crisis, crisis response, and post-crisis. Each phase of the crisis management process entails specific managerial and communicative actions, ranging from early detection and preparedness to immediate response strategies and, ultimately, learning, adaptation, and reputation repair.

This model provides a time-sensitive lens through which organizational behavior can be understood and evaluated during disruptive events. In the context of the semiconductor shortage, it enables an analysis of how brands reacted during the height of the crisis, as well as their preparedness beforehand and their strategic recovery efforts afterward. The model's temporal structuring of crisis management serves to complement existing frameworks, such as SCV and SCCT, which prioritize transparency and communication. The model's integration of these elements across all stages of the disruption is a notable strength (Bundy et al., 2017).

2.3. Hypotheses

Hypothesis 1: The global chip shortage that occurred from 2020 to 2023 had a negative impact on brand popularity across the affected sectors.

Supply chain disruptions have been associated with a decline in brand equity and consumer trust. The chip shortage resulted in product delays and reduced availability, which led to a decrease in consumer satisfaction and brand loyalty (Holloway, 2024).

Hypothesis 2: During the chip shortage, public sentiment on social media platforms was largely negative.

In the context of supply chain crises, consumers frequently articulate feelings of

frustration and dissatisfaction through social media channels. The pervasive repercussions of the chip shortage are likely the catalyst for the dissemination of such negative sentiments in online forums (Sheikhattar et al., 2022).

Hypothesis 3: Brands that effectively communicated mitigation strategies during the chip shortage were able to maintain higher levels of consumer trust.

The efficacy of crisis communication in preserving or enhancing consumer trust during periods of disruption is a subject of considerable interest. Brands that have demonstrated a commitment to transparency, particularly in the context of addressing significant challenges such as the chip shortage, have often been observed to establish more robust consumer relationships. The development of clear mitigation strategies has also been identified as a key factor in fostering these relationships (Jézéquel & Gérard, 2017).

Hypothesis 4: Industries with superior supply chain visibility experienced a reduced negative impact on brand popularity during the chip shortage.

Enhanced supply chain visibility enables organizations to respond more rapidly to disruptions, thereby minimizing adverse outcomes (Özkanlısoy, 2024). Industries that implement robust SCV practices tend to demonstrate greater resilience, which in turn preserves brand popularity.

Hypothesis 5: Social media sentiment analysis can serve as an early indicator of shifts in brand popularity during supply chain disruptions.

Real-time sentiment analysis of social media data provides immediate insights into consumer perceptions, enabling brands to respond proactively (Sheikhattar et al., 2022). During the chip shortage, for instance, monitoring social media could have alerted brands to emerging public concerns.

3. Methodology and Data

This section delineates the model employed to analyze the impact of the global semiconductor shortage (2020-2023) on brand popularity and public opinion, meticulously detailing the methodological approach and data sources. The methodology is structured to provide a clear framework for assessing how different factors may influence brand perception during the crisis.

The initial phase of this study involves the introduction of a conceptual model, which delineates the relationship between brand popularity and pivotal variables such as crisis communication strategies and media coverage trends & consumer behavior shifts. This model provides a foundation for understanding how these elements interact to shape public sentiment.

Subsequently, research hypotheses are presented, focusing on the expected negative impact of the semiconductor shortage on brand popularity and the role of effective communication in mitigating consumer dissatisfaction. The analytical approach employed to test these hypotheses relies on sentiment analysis techniques, which are grounded in computational linguistics and data mining rather than traditional statistical inference.

The subsequent section delineates the variables employed in the study, distinguishing between dependent and explanatory variables, and elucidating any data transformations applied to ensure reliability. The subsequent section details the data sources, encompassing archival data from social media and news articles. The study employs sentiment analysis using Orange Data Mining to extract insights from Reddit, Wikipedia, and media reports, thereby capturing public opinion dynamics over time.

The integration of quantitative analysis is a distinctive feature of the methodology, offering a comprehensive approach to understanding the effects of the semiconductor shortage on brand perception.

3.1. Model

The model (see Appendix A)¹ employed in this thesis was developed to best answer the research question.

The explanation of the variables included in this model can be described as follows:

- **Y** represents brand popularity affected by the semiconductor shortage and/or public sentiment expressed towards these brands.
- **X1** represents brand crisis communication strategies.
- **X2** includes media coverage trends and shifts in consumer behavior influenced by the crisis.

The proposed model directly addresses the research question by establishing a structured relationship between brand popularity and key influencing factors during the global chip shortage crisis (2020-2023). Specifically, it examines how brand crisis communication strategies and media coverage trends, along with shifts in consumer behavior influenced public perception and brand resilience.

Brand crisis communication strategies are defined as the way companies responded to the semiconductor shortage through their messaging, transparency, and crisis management tactics. The model's analysis of this factor enables the assessment of whether brands that implemented proactive communication strategies maintained stronger consumer trust and loyalty compared to those that did not. Furthermore, the analysis encompasses the intensity and sentiment of news reports, social media discussions, and consumer reactions, as influenced by shifts in consumer behavior. The media played a crucial role in shaping public perception of the crisis, and by incorporating this factor, the model evaluates whether negative or positive media framing influenced brand popularity.

By examining the interplay between these variables and brand popularity, the model offers a comprehensive understanding of how diverse factors influenced consumer

¹ See Appendix A for full details.

sentiment and brand resilience. This methodological approach enables an empirical assessment of the extent to which public opinion, influenced by media narratives and brand communication efforts, has driven shifts in brand perception across affected sectors.

3.2. Hypotheses and Tests

The hypotheses tested in this study are as follows:

- **H1:** The global semiconductor shortage negatively impacted brand popularity across affected sectors.
- **H2:** Public sentiment on social media platforms during the shortage was predominantly negative.
- **H3:** Brands that effectively communicated their mitigation strategies were better able to maintain consumer trust.
- **H4:** Industries with greater supply chain visibility experienced less negative impact on brand popularity.
- **H5:** Sentiment analysis of social media and media sources can serve as an early indicator of brand popularity fluctuations during supply chain crises.

The proposed model directly addresses the research question by establishing a structured relationship between brand popularity and key influencing factors during the global chip shortage crisis (2020-2023). Specifically, it examines how brand crisis communication strategies and media coverage trends, along with shifts in consumer behavior influenced public perception and brand resilience.

Brand crisis communication strategies are defined as the way companies responded to the semiconductor shortage through their messaging, transparency, and crisis management tactics. The model's analysis of this factor enables the assessment of whether brands that implemented proactive communication strategies maintained stronger consumer trust and loyalty compared to those that did not. Furthermore, the analysis encompasses the intensity and sentiment of news reports, social media discussions, and consumer reactions, as influenced by shifts in consumer behavior. The identification of patterns in public sentiment during the crisis is enabled by the previously mentioned approach.

While some statistical comparisons, including ANOVA and t-tests, were utilized to illustrate patterns, the analysis remains primarily exploratory and descriptive in nature. The objective of this study was not to estimate parameters or construct predictive models, but rather to identify converging trends and patterns across multiple data sources. This evidence-based approach aims to support the hypotheses through triangulated signals rather than formal statistical inference.

3.3. Variables

3.3.1. *Dependent variable*

In this study, the dependent variable is brand popularity, which is measured using two main indicators: the number of mentions on Reddit and public opinion indicators derived from media sources. Reddit mentions offer a direct measurement of the frequency and context in which a brand is discussed by users on a prominent social media platform.

This metric is advantageous in that it reflects real-time discussions and sentiments, providing insights into the brand's visibility and consumer engagement. Conversely, public opinion indicators from media sources encompass factors such as the tone and frequency of brand mentions in news outlets, blogs, and other forms of media. Collectively, these indicators offer a more comprehensive understanding of a brand's reputation and recognition in the public sphere.

The integration of these two data sources enables a multifaceted assessment of brand popularity, encompassing both user-generated content and professional media coverage.

3.3.2. *Explanatory variables*

The explanatory variables in this study are the crisis communication strategies employed by brands and the intensity and tone of media coverage. Crisis communication strategies refer to the ways in which a brand responds to crises or negative events that might impact its reputation. These strategies may encompass public apologies, transparency, or proactive engagement with the public to rebuild

trust and manage perceptions.

The investigation focuses on the effectiveness of these strategies in influencing brand popularity. The intensity and tone of media coverage are also critical explanatory variables. Intensity, defined as the extent of brand coverage in media during and following a crisis, and tone, denoting the nature of media coverage as positive, neutral, or negative, are pivotal in this regard. The amalgamation of these variables will facilitate the determination of the influence that both the brand's own communication efforts and the external media environment exert on the public's perception of the brand's popularity.

3.4. Sampling approach

3.4.1. Archival data

The sampling approach for this study involves the use of archival data from a range of secondary sources, which are crucial for analyzing the crisis and its impact on brand popularity. The initial secondary source is Reddit, which offers significant insights into consumer discourse concerning the crisis. Reddit, a popular platform where users openly share their thoughts and experiences, serves as a rich source for understanding public sentiment and reactions during a crisis.

The second secondary source is Wikipedia, which offers a historical overview and explanation of the shortage, providing context and background information that helps to understand the broader narrative surrounding the event. The third secondary source is news articles, which offer expert and media perspectives on the crisis. These articles are significant because they provide an external viewpoint on how the crisis is perceived and reported by the media, helping to identify the tone and intensity of media coverage.

The analysis period for this study extends from January 2020 to December 2023, enabling a comprehensive review of the crisis over time and its effects on the brand. The data collection process was facilitated by employing Orange Data Mining software, which was utilized to extract, process, and analyze sentiment from these

diverse sources. This software assists in the automation of the extraction of relevant data from large data sets, ensuring that the analysis is both efficient and accurate. The data obtained from Reddit, Wikipedia, and news articles were meticulously processed to assess sentiment and identify trends in public opinion and media coverage.

To ensure the relevance and quality of the dataset, specific inclusion and exclusion criteria were applied during the data collection phase.

For Reddit, only English-language posts that explicitly mentioned keywords such as “chip shortage” “semiconductors”. Posts that were identified as spam, off-topic, promotional, or unrelated to the context of crisis were systematically excluded, as they do not contribute to the analysis of consumer sentiment or media coverage of the crisis. The filtering process was executed through the implementation of keyword filters.

For media articles, specifically those published by Wikipedia, The Guardian and The New York Times, the selection included articles published between January 2020 and December 2023 that mentioned keywords such as “chip shortage” “semiconductors”. The articles were required to be classified into pertinent sections such as “Business”, “Technology” or “Economy”. Editorials, opinion pieces, and non-factual columns were excluded to maintain consistency and focus on reported events.

For this study, the sample was limited to company-specific pages within industries affected by the shortage, including but not limited to the automotive, consumer electronics, and information technology industries. To ensure the relevance and currency of the articles, those that were not related to the mentioned sectors or had not been updated during the crisis period were excluded from consideration.

After the collection of the datasets, they were subjected to preprocessing using Orange Data Mining’s “Preprocess Text” widget. This widget ensures the cleanliness and uniformity of the text prior to the application of any sentiment analysis. Specifically, the system performs the following functions:

- The text is converted to lowercase, a process referred to as case normalization.
- The function removes stop words, which are defined as common but uninformative words such as “the”. “and”. and “of”.
- It has been demonstrated that this process effectively removes both punctuation and special characters.
- Stemming and lemmatization are optional processes that can be applied; however, these linguistic processes were not utilized in this study in order to maintain lexical nuances.

This preprocessing stage is crucial to minimize noise in the data and improve the accuracy and reliability of sentiment analysis algorithms like VADER and SentiART.

3.4.2. How Orange Datamining works

Orange Data Mining is an open-source data visualization and machine learning tool that has gained renown for its intuitive widget-based graphical interface. Its “Text Mining” extension empowers users to conduct sentiment analysis visually, obviating the necessity for programming expertise, thereby catering to non-technical users seeking to explore textual data from diverse sources, including social media, press articles, and online forums.

As illustrated by the provided screenshots (see Appendix B)², the sentiment analysis process is applied to diverse textual sources, including Wikipedia, The Guardian, The New York Times, and Reddit. The workflow is structured in a modular fashion, with textual data being imported into the system via the Corpus widget, a feature tailored to each individual source. Subsequently, the data undergoes a preprocessing phase within a module designated as “Preprocess Text”. This module is tasked with the removal of special characters and “stop words”, a process that can encompass operation such as lower-casing, with the objective of enhancing the efficacy of subsequent analyses.

Following the preparation of the texts, the Sentiment Analysis widget is employed to

² See Appendix B for full details.

evaluate the affective orientation of each document. Orange offers a variety of approaches to sentiment analysis, the most prevalent of which are based on lexicons, such as VADER (adapted to English texts), Liu & Hu, or multilingual modules, which can also integrate customized dictionaries. These methods assign sentiment scores to each text, often in the form of positive, negative, or neutral values, or even an overall composite score. The two sentiment analysis methods employed in this study are VADER and SentiART. A more thorough exposition of these methods can be found in the subsequent sections (Biolab, 2018).

The results of this analysis are then explored using visualization widgets such as Bar Plot, Distributions, Word Cloud, or Heat Map. This facilitates the identification of prevalent trends in the sentiments expressed, the comparison of different sources with each other, and the identification of dominant terms or emotions. Other widgets, such as Data Table or Corpus Viewer, facilitate the consultation of individual results or analyzed texts. Finally, the Save Data widget is employed to export enriched data for further analysis or archiving.

Orange Data Mining offers a comprehensive visual framework for sentiment analysis, encompassing data cleansing, automatic language interpretation, and the visualization of results. This tool is particularly effective in transforming massive volumes of raw text into actionable insights, making it relevant in fields as varied as media monitoring, academic research, and opinion polling.

3.4.3. VADER sentiment analysis

VADER (Valence Aware Dictionary and Sentiment Reasoner) is a lexical sentiment analysis method designed to interpret the emotional tone of texts, particularly those from social networks. It is predicted on a pre-established lexicon that assigns valence value ranging from -1 (very negative) to +1 (very positive) to each word. This lexicon is augmented by heuristic rules that consider linguistic context, such as negation, intensification (e.g., “very good”), and emphatic punctuation (e.g., capital letters or exclamation marks). This approach enables VADER to produce detailed sentiment scores, categorizing them as positive, neutral, negative, and an overall composite score (Kurniawan & Anggraini, 2023).

However, it is important to note that VADER is not without its limitations. Its efficacy may be diminished when applied to more complex or literary texts, where emotions are expressed in more subtle ways. Additionally, while VADER is primarily designed for English, its effectiveness may be diminished when used for other languages without appropriate adaptations.

3.4.4. SentiART sentiment analysis

SentiART is a sentiment analysis method developed to evaluate emotions in literary texts, particularly in the field of computational poetics. Distinct from methodologies that rely on emotional lexicons, SentiART employs vector space models (VSMs) to represent words in a multidimensional semantic space. This representation facilitates the calculation of metrics such as the emotional valence and the affective-aesthetic potential (AAP) of words and texts (Jacobs, 2019).

A notable advantage of SentiART is its capacity to function without a predefined emotional lexicon, thereby ensuring its applicability across over 150 languages for which VSMs are available. This flexibility is particularly advantageous for the analysis of literary texts, where emotions are often expressed implicitly.

However, it should be noted that SentiART necessitates substantial computational resources to generate and manipulate VSMs, and its effectiveness is contingent on the quality of the vector models available for the language and domain of interest. Additionally, while SentiART presents an innovative approach to sentiment analysis in literary texts, its adaptation to non-literary texts or specific domains may necessitate further refinement.

4. Results

This section presents the results of an empirical analysis based on four separate sources: The following sources were utilized: Reddit, The Guardian, The New York Times, and Wikipedia. Each of these sources provides a complementary perspective on brand perception, ranging from institutional media discourse to spontaneous expression by Internet users. The objective of this study is to examine the evolution of brand sentiment during the global semiconductor shortage and the factors that are likely to influence this perception.

The initial corpus was derived from Reddit, encompassing a total of **1,729** observations spanning the years 2020 to 2023. The primary variables encompass a sentiment score (sentiment, compound), the brand evoked (brand), the date, as well as categorical data associated with communication strategy (comm_strat) and logistical visibility (SCV_level).

The second corpus is based on The Guardian, with **1,713** articles containing explicit references to brands affected by the crisis. The emotional tone of each extract was meticulously calculated and then associated with metadata, including brand, publication date, and explanatory variables utilized in the hypotheses.

The third data set is derived from the New York Times, with a total of **245** articles undergoing analysis. Despite the reduced sample size, this corpus provides a valuable qualitative perspective on American media coverage, facilitating an examination of the relationship between editorial discourse and brand perception during periods of tension.

Wikipedia, an encyclopedic source, is represented by **18** articles dealing with the companies concerned. While its style is, by definition, more neutral, the sentiment scores enable us to identify notable variations between pages. The analysis of this source offers a valuable addition to the existing literature on traditional and social media, providing a comprehensive overview of brands' institutional image.

The construction of certain variables necessitated the use of inference, assisted by artificial intelligence, particularly ChatGPT, within a controlled methodological framework. Consequently, the brand variable was assigned from the analyzed text content in instances where it was not explicitly specified.

For the variable “comm_strat”, which indicates whether a brand has implemented a communication strategy considered strong or weak during the crisis, a classification was carried out based on documentary research (via ChatGPT) concerning official speeches, press releases, or public visibility actions of the brands studied. In a similar manner, the SCV_level (Supply Chain Visibility) variable was allocated according to criteria of logistical transparency, determined from information available in open sources and interpreted via IA according to predefined rules (references to production lines, suppliers, or inventory management).

The “buzz_score”, employed as a metric of popularity, was derived from the relative frequency of brand mentions in each corpus. This score has undergone adjustment to ensure its comparability across brands and sources, with adjustments made through either logarithmic transformation or normalization.

The integration of all the data into the Orange Data Mining platform was instrumental in facilitating the sentiment analyses, statistical comparisons, and visualizations by subgroups that were necessary to evaluate the five hypotheses formulated in this research.

4.1. Hypothesis 1: The global semiconductor shortage negatively impacted brand popularity across affected sectors

The initial hypothesis posits that the global semiconductor shortage experienced from 2020 to 2023 exerted a deleterious influence on the popularity of the brands in question. To assess this, the present study expanded its analysis to include several online information sources, including Wikipedia, The Guardian, Reddit and the New York Times. Two automatic language processing methods were applied: VADER, which is oriented towards lexical sentiment, and SentiART, which is more focused on psycholinguistic dimensions such as dominance, valence, and arousal.

4.1.1. *Results by source*

4.1.1.1. *Wikipedia*

Wikipedia's overall tone between 2020 and 2023 (see Appendix C)³ regarding the microprocessor shortage is mostly neutral, but there are occasional negative shifts. The VADER method will be utilized to demonstrate Wikipedia's neutrality during the crisis, while the SentiART method will be employed to ascertain the prevalence of negative sentiments within the total sample of 18 articles on Wikipedia.

A comparative analysis of the graphs corresponding to the two distinct methods reveals that the SentiART analysis identifies three articles with a predominantly negative sentiment, while the VADER method categorizes all articles as neutral.

4.1.1.2. *The Guardian*

In the case of the British newspaper The Guardian, the impact of the semiconductor shortage on brand perception is particularly evident. A review of articles (see Appendix C)⁴ published between 2020 and 2023 in the "Business", "Technology", "Global Development", "Business", "Money", "Science", and "World News" sections reveal a persistent use of terms with negative connotations, particularly in 2022, a period characterized by the most significant tensions in global supply chains.

The VADER method has been instrumental in identifying "compound" sentiment scores that frequently fall below zero. This phenomenon is particularly evident in articles addressing financial losses incurred by automakers, price increases for consumers, and challenges in accessing certain electronic products.

This phenomenon is further accentuated by the implementation of the SentiART model (see Appendix C)⁵, which underscores a heightened level of arousal (emotional excitement). This suggests that the management of the crisis is not merely negative but is also accompanied by a certain degree of media stress. This type of discourse is likely to amplify the effect on public perception of brands, particularly when they are

³ See Appendix C (Sentiment Analysis of Wikipedia Articles) for further details.

⁴ See Appendix C (Sentiment Analysis of The Guardian Articles) for further details.

⁵ See Appendix C (Sentiment Analysis of The Guardian Articles) for further details.

mentioned in contexts of crisis management or strategic failure. The findings obtained from The Guardian substantiate the notion that the written press, particularly when it adopts an engaged or critical tone, has contributed to shaping a degraded image of companies affected by the shortage.

4.1.1.3. NY Times

The New York Times, as an international generalist media outlet, offers a more diverse but equally revealing perspective on the effects of the chip shortage on brand perception. A content analysis of articles from various sections including “Business”, “Technology”, “Automobiles”, “Business Day”, and “World” reveals that the tone adopted varies according to the target audience and the nature of the content. However, a constant theme emerges: the sectors impacted are often presented as vulnerable, ill-prepared, or dependent on a globalized system in crisis.

The findings (see Appendix C)⁶ of the VADER analysis demonstrate analogous trends to those documented in The Guardian, exhibiting consistently unfavorable “compound” sentiment scores in the economic sections, particularly during periods of peak disruption (2021-2022).

The SentiART method (see Appendix C)⁷ lends further support to this reading, demonstrating a significant decline in valence among articles pertaining to semiconductor companies. Additionally, there is notable variability in dominance, which exhibits fluctuations in accordance with the stance of the sectors referenced.

This narrative dynamic suggests that brand representation extends beyond mere mention of a logistical issue but rather constitutes a component of a crisis framework that is perceived as unveiling inherent systemic deficiencies. These results, derived from a medium with a robust international presence, thus substantiate the hypothesis that the shortage has exerted a deleterious effect on brands’ popularity and credibility, by positioning them within a protracted crisis narrative.

⁶ See Appendix C (Sentiment Analysis of The New York Times Articles) for further details.

⁷ See Appendix C (Sentiment Analysis of The New York Times Articles) for further details.

4.1.1.4. *Reddit*

The initial hypothesis posits that the global semiconductor shortage experienced from 2020 to 2023 has exerted a deleterious influence on the popularity of the brands in question. To evaluate this hypothesis, an analysis was conducted of the sentiments expressed by users on Reddit using the VADER algorithm. The “compound” sentiment score, ranging from -1 (very negative) to +1 (very positive), was utilized as a pivotal perception indicator.

Aggregated data (see Appendix C)⁸ disaggregated by month demonstrate a marked decline in sentiment commencing in mid-2022, with a nadir observed in October 2022, when the mean score declined to -0.74. This period corresponds to a time of particularly high tension in the global supply chain.

Beginning in 2023, a gradual recovery in sentiment was observed (see Appendix C)⁹, with levels reaching neutral or positive territory by December 2023 (compound = 0.61). This recovery could be indicative of an improvement in the situation or the effectiveness of communication efforts by brands.

The observed correlation between periods of crisis and declines in sentiment on Reddit indicates a potential loss of brand popularity or satisfaction during the shortage. The emotions expressed frequently indicate frustration or impatience, as evidenced by the negative scores.

A post stated: *“Still can’t get a PS5 because of the chip shortage. This is beyond ridiculous.”* This statement was assigned a compound score of -0.65 by VADER, indicating a predominant negative sentiment. A manual inspection of the elements substantiates this assertion. The language employed in this text conveys a clear sense of frustration and blame directed towards supply chain issues, thereby confirming the model's capacity to accurately capture the emotion expressed. This example lends support to the hypothesis that such outbursts are indicative of a decline in brand

⁸ See Appendix C (Sentiment Trends on Reddit) for further details.

⁹ See Appendix C (Sentiment Trends on Reddit) for further details.

popularity, particularly when frustrations are attributed to perceived incompetence or lack of availability.

4.2. Hypothesis 2: Public sentiment on social media platforms during the shortage was predominantly negative

It is hypothesized that the predominant sentiment expressed on Reddit between 2020 and 2023 regarding the semiconductor shortage was predominantly negative. To verify the proposition, an in-depth analysis of the Reddit corpus was conducted. The corpus was gathered around terms related to the crisis, and the analysis employed SentiART, which enables a more nuanced assessment of the underlying emotional charge through dimensions such as anger, fear, or sadness.

4.2.1. Results by source

4.2.1.1. Reddit

A more extensive examination of trends on Reddit (see Appendix D)¹⁰ from 2020 to 2023 reveals a persistent and pervasive presence of negative emotions, irrespective of variations in the overall sentiment score. This observation, beyond the identification of specific negative peaks, underscores a continuous and widespread prevalence of negative sentiments. During the entire period under consideration, emotions such as anger persist, with average intensity level estimated at approximately 0.4 for anger (on a scale of 0 to 1). This emotional consistency suggests that users experienced sustained discomfort, which cannot be attributed to isolated events.

The months with the strongest combinations of these negative emotions corresponded to periods when the effects of the shortage had accumulated: queues for technological products, price rises for graphics cards, or the unavailability of new vehicles.

A representative remark from the mid-2022 period is provided herewith: *“My GPU preorder just got pushed back again. What’s the point anymore?”* The post received a

¹⁰ See Appendix D (Emotional Analysis on Reddit) for further details.

VADER score of -0.52. The communication conveys a combination of resignation and disappointment, rather than overt anger. Despite its negative tone, it contrasts with more assertive posts. This example demonstrates that negative sentiment on Reddit was not monolithic but rather exhibited a range of emotional states, all of which were associated with consumer dissatisfaction. This finding supports the hypothesis and underscores the necessity for nuanced emotional classification.

These findings imply that Reddit functioned not only as a platform for objective criticism but also as an emotional outlet, where Internet users articulated their feelings of anxiety, fatigue, and even resignation in the face of protracted shortages. The predominant emotions identified serve to reinforce the notion of a crisis that is perceived as profound, not only from an economic perspective, but also from a psychological standpoint, particularly for individuals who are dependent on electronic or automotive products that have been impacted.

4.3. Hypothesis 3: Brands that effectively communicated their mitigation strategies were better able to maintain consumer trust

To test hypothesis 3, the analysis centered on the influence of brand communication strategy during the crisis on maintaining public trust. This assumption is predicated on the notion that companies that have implemented a clear, regular, and proactive communication strategy in response to the semiconductor shortage have been better able to maintain a positive perception among consumers and stakeholders.

To translate this concept into data, two variables were utilized. The first, “comm_strat”, is a categorical variable indicating the perceived level of communication strategy adopted by each brand. The coding of these systems typically occurs in two modalities: The communication style can be characterized as either “strong” or “weak”. A “strong” communication style is defined by its active nature, transparency, and reassurance. In contrast, a “weak” communication style is characterized by its limited nature, absence, or perception of inefficiency. This classification was established based on documentary research on official statements, press releases, public campaigns, or crisis management initiatives. It was supplemented by assisted interpretation via ChatGPT, allowing for the

homogenization of the judgments in the entire corpus.

The second variable, “brand”, serves to identify the brand to which each observation pertains. The text was automatically extracted, where possible, or derived from context using a guided semantic analysis. The utilization of these two variables enabled the execution of cross-brand comparisons and the aggregation of observations according to the adopted form of communication.

4.3.1. Results by source

4.3.1.1. Wikipedia

An analysis of Wikipedia data reveals a significant correlation between brands’ communication strategy and the sentiment expressed in their encyclopedic treatment. A binary classification of strategies was employed, with strategies characterized as “strong” or “weak”. The overall sentiment indicator was also studied, which was comprised of both the compound and sentiment. The results demonstrated that brands with a communication that was qualified as “strong” experienced a significantly more positive representation.

The initial graph (see Appendix E)¹¹ (bars colored by comm_strat) offers a clear illustration of this trend: brands such as Apple, Intel, and Ford, identified as having conducted proactive communication during the crisis, stand out with compound scores regularly above zero, or even very positive (up to +1). Conversely, brands with “weak” or absent communication (notably grouped under the “Other” category) show much lower scores, often negative or even close to -1. This phenomenon of brand perception can be understood through the lens of brand communication, which can be categorized as informative, reassuring, or anticipatory.

The second graph (see Appendix E)¹² (box plot of sentiment by brand) further refines this analysis: the average sentiment scores for Apple (0.2168) and Intel (0.1581) are significantly higher than those for “Other” brands (0.0448). The statistical significance of the observed results was confirmed by the analysis of variance

¹¹ See Appendix E (Communication Strategies and Sentiment on Wikipedia) for further details.

¹² See Appendix E (Communication Strategies and Sentiment on Wikipedia) for further details.

(ANOVA) test, with a p-value of 0.010. This finding allows us to reject the null hypothesis of equality of means and validates the effect of the communication strategy on emotional perception. We observe that variance is also more controlled for well-perceived brands, suggesting narrative consistency in their editorial treatment.

In summary, even on an impartial platform such as Wikipedia, the communication strategies employed by brands have a discernible impact on the prevailing sentiment and the image they project. The hypothesis that effective communication is conducive to maintaining trust is substantiated in this instance.

4.3.1.2. The Guardian

A review of articles published by The Guardian between 2020 and 2023 (see Appendix E)¹³ reveals that the quality of brands' crisis communication is explicitly revealed. The analysis of sentiment reveals a marked disparity in compound scores between brands that have been identified as having communicated proactively and those whose efforts have been deemed inadequate. The Student's t test conducted in Orange substantiates that this discrepancy is statistically significant, with a p-value < 0.05, thereby reinforcing the hypothesis's validity.

Brands that demonstrate robust communication (e.g., Volkswagen, Ford, Samsung) are frequently characterized as transparent, with concrete plans, and favorably mentioned in the "Business" and "Technology" sections. Their average sentiment scores approach +0.2. Conversely, respondents whose statements were absent or ambivalent (e.g., Dell, Lenovo) exhibited negative scores, frequently accompanied by emotions of anger or disappointment. This observation indicates that media coverage does not merely convey facts but also entails an implicit evaluation of communication behavior. Therefore, Hypothesis 3 has been largely confirmed in the British trade press.

¹³ See Appendix E (Communication Strategies and Sentiment in The Guardian) for further details

4.3.1.3. *NY Times*

A thorough analysis of the New York Times articles, employing sentiment scores (sentiment and compound) and the “comm_strat” variable, unequivocally substantiates the hypothesis that brands that effectively communicated during the shortage were able to sustain a more favorable public image. The two graphs from the analysis in Orange provide a clear illustration of this differential effect.

Regarding the overall sentiment (see Appendix E)¹⁴, brands demonstrating robust communication exhibited an average score of 0.127, in contrast to the 0.1027 average score of those with deficient communication. The discrepancy between the means is found to be statistically significant, as evidenced by the results of the Student’s t-test, which yielded a p-value of 0.000, falling well below the 5% level of significance ($\alpha = 0.05$). This outcome permits the rejection of the null hypothesis, which proposes equality of the means. This finding indicates that the perception of brands that have communicated is not only more positive, but consistent across the data set.

The second graph (see Appendix E)¹⁵, which is based on compound scores from the VADER analysis, further corroborates this trend. It was observed that brands demonstrating effective communication received an average score of 0.135, while those exhibiting weaker communication received scores lower, reaching -0.160. Once more, statistical analysis corroborates the significance of the observed discrepancy ($p = 0.017$). It has been demonstrated that brands that have clearly articulated their strategic approaches, such as Apple, Intel, or GM, are associated with a more neutral or slightly favorable tone in the articles. Conversely, those that have remained ambiguous or silent are often characterized by a more critical and even anxiety-inducing tone.

These findings demonstrate that, even in a milieu traditionally characterized by its objectivity, crisis communication exerted a pivotal influence in shaping public perception of brands. Hypothesis 3 is fully validated in this case, and the effect of

¹⁴ See Appendix E (Communication Strategies and Sentiment in The New York Times) for further details

¹⁵ See Appendix E (Communication Strategies and Sentiment in The New York Times) for further details

communication is evident not only in the average feelings, but also in the dispersion and coherence of readership reactions relayed through media discourse.

4.3.1.4. *Reddit*

The findings from Reddit, a spontaneous discussion platform, unequivocally substantiate the impact of communication strategy on the sentiment expressed by users. A comparison between brands with “strong” and “weak” communication, as determined by compound scores calculated using the VADER algorithm, reveals a significant difference in the overall tone of comments.

The initial graph (see Appendix E)¹⁶ displays the mean values for each brand. The findings indicate that brands recognized for their proactive communication, including Apple, GM, Ford, and Intel, consistently achieve notably high scores, with GM and Apple attaining particularly elevated values above 0.4. Conversely, brands classified as “weak” or “other” struggle to achieve neutrality, with scores ranging from 0.01 to 0.02. These scores demonstrate a high degree of variability, reflecting more disordered, even critical reactions.

As illustrated in the second graph (see Appendix E)¹⁷, which presents a box plot by communication group (comm_strat), the overall effect is confirmed. Brands with robust communication exhibit an average compound of 0.216, while those with ineffective communication display an average of 0.016.

The Student’s t test, which is associated with the data, yielded a result that is highly significant ($p = 0.001$). This finding allows us to confidently conclude that crisis communication directly influenced the emotional perception expressed on Reddit. The presence of a brand in the minds of consumers was associated with more measured and benevolent reactions, as opposed to the hostility and frustration that was observed in the absence of brand presence.

Conversely, brands that proactively communicated during the crisis received more favorable reactions. A representative post reads: *“Intel seems to be doing a decent job keeping people informed. Props to them.”* The comment is interpreted as moderately

¹⁶ See Appendix E (Communication Strategies and Sentiment on Reddit) for further details

¹⁷ See Appendix E (Communication Strategies and Sentiment on Reddit) for further details

positive, as indicated by a VADER compound score of 0.36. Preliminary findings from a manual review suggest that users have acknowledged and expressed appreciation for Intel's transparency. This finding serves to substantiate the algorithm's output and lends support to the hypothesis that effective communication has a positive impact on consumer sentiment on Reddit.

This finding is particularly salient in the context of Reddit, which functions as an uncontrolled channel, wherein brands do not dictate the message. This finding lends further credence to the observed effect, suggesting that communication not only fosters reassurance within an official setting but also permeates the freest spaces of expression. Hypothesis 3 is thus fully validated in this context, with robust statistical support and discernible emotional differences between the groups.

4.4. Hypothesis 4: Industries with greater supply chain visibility experienced less negative impact on brand popularity

To test hypothesis 4, the analysis centered on the impact of Supply Chain Visibility (SCV) on brands' public image during the global semiconductor shortage. This assumption is predicated on the notion that enterprises demonstrating enhanced transparency in their logistics and procurement practices have experienced diminished exposure to reputational risks during periods of disruption.

To operationalize this concept, a categorical variable has been introduced. This variable is referred to as "SCV_level". This variable employs a classification system to categorize brands according to their level of visibility, which is typically categorized as "high" or "low" visibility. This categorization is determined based on publicly available information related to the brands' logistics and communication processes. The levels were awarded based on a supervised qualitative evaluation, based on the analysis of official sites, press releases, CSR reports, and sectoral summaries. The categorization in question was validated using contextual interpretations enhanced by artificial intelligence-assisted research, specifically the ChatGPT model.

Despite the "SCV_level" status as approximation, it establishes a functional framework for comparing the degree of transparency of brands. This variable was employed as a subgroup criterion in Orange Data Mining, enabling the visualization

and statistical comparison of feelings distribution according to levels of visibility. The results were examined separately for each source to assess their robustness.

4.4.1. Results by source

4.4.1.1. Wikipedia

An analysis of the “SCV_level” variable in the Wikipedia data (see Appendix F)¹⁸ reveals a marked over-representation of brands with high logistics visibility. The distribution graph indicates that most documented cases pertain to companies that have clearly communicated their supply chain management, such as Apple or Intel.

Brands with low or indeterminate visibility remain marginal, limiting direct comparisons in terms of sentiment. This imbalance can be explained by the encyclopedic nature of Wikipedia, which favors verifiable sources and the most visible companies. Therefore, even in the absence of extensive statistical testing, the concentration around the most transparent brands appears to substantiate the notion that visibility in the supply chain contributes to an enhanced public image.

4.4.1.2. The Guardian

The box plot of sentiment scores (see Appendix F)¹⁹ by logistic visibility level (SCV_level) demonstrates clear disparities. It has been demonstrated that brands with high visibility exhibit a more positive average sentiment than those classified as low or indeterminate.

The results of the one-way analysis of variance (ANOVA) test confirm that this difference is statistically significant ($p = 0.000$), thereby validating the hypothesis that greater supply chain transparency correlates with a more favorable image in the media. Consequently, The Guardian’s reporting on brands demonstrating superior control over their supply chains is more favorable.

¹⁸ See Appendix F (Impact of Supply Chain Visibility on Wikipedia) for further details

¹⁹ See Appendix F (Impact of Supply Chain Visibility in The Guardian) for further details

4.4.1.3. *NY Times*

The graph (see Appendix F)²⁰ illustrates that brands with high supply chain visibility attain significantly more positive sentiment scores (mean = 0.135) in comparison to those with low (mean = -0.112) or indeterminate (mean = -0.084) visibility.

The results of the one-way analysis of variance (ANOVA) test confirm the significance of the observed difference ($p = 0.000$), thereby validating the hypothesis that logistical transparency helps preserve brand image in mainstream media. The New York Times has demonstrated a discernible preference for companies that have exhibited more effective communication in addressing the crisis.

4.4.1.4. *Reddit*

On the online platform Reddit, sentiment scores (see Appendix F)²¹ exhibit a discernible variation according to the level of supply chain visibility (SCV_level). The mean score for high-visibility brands is -0.037, while the mean score for low-visibility brands is -0.235, and for indeterminate brands, the mean score is -0.236.

Despite the persistent negative overall scores, the discrepancy is substantial (ANOVA, $p = 0.000$), indicating that even in an environment of free expression, brands that are transparent about their logistics have a competitive advantage in terms of image. This finding aligns with observations from other sources, thereby substantiating the hypothesis that supply chain visibility functions as a reputational protection factor.

While not overtly concerned with supply chains, posts frequently imply assessments of brand transparency. To illustrate this point, consider the following example: “*At least Apple acknowledged the delays. They’re doing better than most*”. VADER assigned a compound score of 0.42, and qualitatively, this sentiment reflects appreciation for Apple's openness—implicitly tied to supply chain visibility. This finding lends further support to the prevailing notion that brands that are both visible and proactive are perceived more favorably, even in informal social discourse.

²⁰ See Appendix F (Impact of Supply Chain Visibility in The New York Times) for further details

²¹ See Appendix F (Impact of Supply Chain Visibility on Reddit) for further details

4.5. Hypothesis 5: Sentiment analysis of social media and media sources can serve as an early indicator of brand popularity fluctuations during supply chain crises

To test Hypothesis 5, an investigation was conducted into the relationships between sentiment scores expressed in different types of sources and various popularity indicators associated with brands. The objective of this hypothesis is to ascertain the extent to which emotions perceived in digital discourse can serve as precursors of variations in brand awareness or media attention. To this end, an indicator called “buzz_score” was mobilized. This is a proxy for popularity, constructed from the volume of mentions associated with each brand, then normalized to enable comparisons between sources.

In certain instances, this score has been enriched or corrected by the general tone (positive or negative) to reflect not only the quantity, but also the quality of media attention. While this measure remains an approximation, it provides an operational representation of visibility dynamics. To assess the consistency or specificity of the observed trends, analyses were conducted separately for each of the sources utilized in this research.

4.5.1. *Results by source*

4.5.1.1. *Wikipedia*

The objective of this study was to test the hypothesis that online sentiment analysis can predict variations in brand popularity during a crisis. To this end, a scatter plot enriched with brand-specific linear regressions was produced from data extracted from Wikipedia.

The plot (see Appendix G)²² demonstrates the correlation between the overall sentiment, measured as the aggregate score derived from lexical analysis, and the level of popularity, approximated via a “buzz_score”. While the correlation coefficients vary across brands, several noteworthy trends emerge.

²² See Appendix G (Sentiment-Popularity Correlations on Wikipedia) for further details

The Apple brand, for instance, demonstrates a significant positive correlation ($r = 0.70$), indicating that periods of enhanced perception coincide with an increase in its visibility on Wikipedia. Conversely, Ford exhibits a pronounced negative correlation ($r = -0.93$), indicating that its visibility increases precisely when it encounters more critical scrutiny, a phenomenon consistent with its notoriety stemming from issues associated with the semiconductor crisis.

It has been demonstrated that other brands, including Intel ($r = -0.34$), as well as entries classified as “Other” ($r = 0.79$), offer less extreme but nevertheless significant correlations.

These results offer partial support for the hypothesis that sentiment can precede or reflect fluctuations in popularity.

4.5.1.2. *The Guardian*

A cross-analysis of sentiment and media popularity, based on articles published in The Guardian, reveals fewer clear-cut results overall than on other platforms. To this end, a scatter plot (see Appendix G)²³ with brand-specific linear regressions was utilized to measure the correlation between average sentiment and “buzz_score”.

Most brands demonstrate an absence of correlation between the two variables. Apple ($r = 0.02$), GM ($r = -0.07$), and Intel ($r = -0.05$) demonstrate no significant trend, indicating that the volume of coverage is not substantially influenced by the perceived tone of the news. This stability is indicative of a consistent editorial approach, potentially less responsive to emotional fluctuations than social platforms such as Reddit.

However, it should be noted that there are exceptions to this general rule. A moderate positive correlation is demonstrated by Dell ($r = 0.20$), suggesting a modest concurrence between periods of heightened sentiment and augmented buzz. In contrast, Samsung ($r = -0.92$) exhibits a robust negative correlation, which may be indicative of heightened media attention during periods of crisis or criticism.

²³ See Appendix G (Sentiment-Popularity Correlations in The Guardian) for further details

4.5.1.3. *NY Times*

An analysis of data from the New York Times reveals a more nuanced relationship between sentiment expressed in content and brand popularity, measured by `buzz_score`. To this end, a scatter plot (see Appendix G)²⁴ augmented with linear regressions was utilized to assess the correlation on a brand-by-brand basis.

The overall correlation coefficients ranged from weak to moderate, thereby underscoring the intricate nature of the relationship between perceived tonality and media attention. This phenomenon is exemplified by prominent brands such as Apple ($r = -0.21$) and Ford ($r = -0.32$), which demonstrate an inverse trend, indicating that as the tone becomes more critical, the brands' coverage increases. This phenomenon may be indicative of heightened visibility during periods of tension or controversy, a common occurrence in media environments that are reactive to negative news.

In contrast, brands such as Toyota ($r = 0.45$) or Intel ($r = 0.15$) demonstrate a positive correlation, suggesting that media valorization is more closely associated with favorable tones, potentially attributable to positively perceived innovations, announcements, or campaigns.

The Volkswagen brand, though exhibiting a modest presence in this context, displays a pronounced incline, albeit on a reduced number of points, which constraints interpretation.

This visualization lends support to the notion that sentiment occasionally, though not invariably, serves as an indicator of media popularity. The nature of this relationship is found to be contingent upon the context of coverage, the news specific to each brand, and the editorial treatment of the newspaper.

4.5.1.4. *Reddit*

On Reddit, a platform characterized by spontaneous and often emotional user interactions, the analysis of hypothesis 5 assumes relevance. For this source, a scatter plot (see Appendix G)²⁵ was generated from the mean sentiment scores and a

²⁴ See Appendix G (Sentiment-Popularity Correlations in The New York Times) for further details

²⁵ See Appendix G (Sentiment-Popularity Correlations on Reddit) for further details

popularity indicator “buzz_score”. The latter is defined here as a proxy for the relative volume of mentions.

It is imperative to note that the data categorized as “Other” has been systematically excluded from the final graph. Indeed, the high volume of the data set would have resulted in a visual saturation of the representation space, thereby obscuring the trends of the primary brands and rendering the graph difficult to interpret.

The remaining results indicate moderate yet predominantly positive correlations for most brands. The correlation coefficients (r) for GM, Intel, and Apple are 0.16, 0.18, and 0.10, respectively. These results indicate that there is a positive relationship between the popularity of these brands on the platform and the sentiment expressed by users. A very weak correlation is exhibited by Ford ($r = 0.03$), while AMD appears to demonstrate a similar trend.

The highest correlation is found in the Other brand (not shown), with an $r = 0.75$. This phenomenon likely reflects the aggregate weight of news and brands that are not systematically categorized, thereby substantiating the notion that topics that are perceived as highly positive tend to generate more extensive discourse.

The post, titled *“Of course GM delays again. Not even surprised anymore.”* This piece received a score of -0.24 from VADER. While not overtly hostile, the tone reflects sarcasm and fatigue, which can still damage brand perception. It is noteworthy that this post garnered a significant number of interactions, as indicated by the high volume of likes and replies. This observation suggests a potential correlation between emotional tone and visibility. While the algorithm detects it as mildly negative, human analysis reveals a more profound emotional undercurrent. This finding lends further credence to the notion that emotional responses, even those that are subtle, can occasionally occur in advance of surges in brand-related online discourse.

While the data on Reddit demonstrates a general correlation between sentiment and popularity, it is important to note that this should not be interpreted as indicative of robust predictive power. While sentiment may accompany visibility trends, causal or predictive claims require further validation. The platform’s conversational environment appears to perpetuate this dynamic, as users respond promptly to positive or negative perceptions associated with brands.

5. Conclusion

5.1. Key findings

This research examined the repercussions of the global semiconductor shortage on brand perception, employing a multifaceted approach that encompassed traditional media, encyclopedic sources, and social platforms. By evaluating five successive hypotheses, it elucidates the mechanisms by which a supply crisis can impact not only the image of the companies involved, but also the dynamics of attention they receive in the digital public sphere.

The findings demonstrate, firstly, that the crisis exerted a substantial influence on brand perception, particularly in unmediated discussion spaces such as Reddit, where the prevailing tone exhibited a notable decline during periods of scarcity. Conversely, brands that employed active and transparent communication methods experienced heightened positive sentiment levels, thereby underscoring the pivotal role of a communication strategy in a crisis context.

Furthermore, supply chain visibility defined as a brand's capacity to elucidate its actions, decisions, and trade-offs emerges as a pivotal factor in shaping public perception. In general, brands that have maintained clear and explicit logistics have demonstrated a more robust reputation. Finally, analyses carried out on emotional signals and attention levels suggest that sentiment expressed online can, in some cases, anticipate or reflect popularity dynamics, paving the way for forward-looking uses of sentiment analysis in strategic intelligence.

The subsequent table offers a synoptic overview of the findings in relation to the five research hypotheses. The study provides a comprehensive overview of the extent to which each hypothesis was supported by the data collected from various sources, including social media, news articles, and encyclopedic content. This consolidated perspective enables a comprehensive understanding of the evolution of sentiment and brand perception during the global chip shortage.

It is noteworthy that all five hypotheses were confirmed, albeit with varying degrees of strength, contingent on the source and method of analysis. No hypothesis was rejected, indicating a consistent pattern across different types of data. This reinforces

the robustness of the evidence-based approach used in this study, which prioritized convergence across multiple indicators over isolated statistical testing.

Hypothesis	Result	Key Insight
H1 – The shortage negatively affected brand perception	✓ Confirmed	Sentiment scores dropped during peak crisis periods, especially on Reddit
H2 – Sentiment on social media was predominantly negative during the shortage	✓ Strongly confirmed	Social media sentiment remained predominantly negative throughout the entire 2020–2023 period
H3 – Brands with strong communication better maintained trust	✓ Strongly confirmed	Clear correlation between strong communication and more positive sentiment
H4 – Supply chain visibility reduced reputational damage	✓ Confirmed	Brands with high visibility in logistics were perceived more positively
H5 – Sentiment can predict brand popularity fluctuations	✓ Partially confirmed	Correlations vary by brand and media; stronger and clearer patterns observed on Reddit

5.2. Recommendations for companies

Considering the findings from this study, a series of pragmatic recommendations can be put forward for companies grappling with supply chain crises.

It is imperative for brands to adopt a proactive and transparent communication strategy. Brands that regularly inform the public and explain their mitigation measures are perceived more positively and are more resilient to criticism.

It is imperative to enhance supply chain visibility. Public disclosure of supply chain operations, partners, and contingency plans has been demonstrated to build consumer trust and media credibility.

The utilization of sentiment monitoring as a strategic instrument is imperative. Platforms such as Reddit function as precursors to consumer dissatisfaction or approval. Real-time sentiment tracking has the potential to provide public relations teams with actionable signals, enabling them to adapt their communication in advance.

These recommendations can assist companies in safeguarding their reputation, fostering consumer trust, and adeptly managing crisis communication during periods of disruption.

5.3. Limits of paper

This work is not without its limitations, which are particularly associated with the nature of the data (e.g., noise, selection bias, and variable representativeness according to source) and the tools employed (e.g., the relative precision of lexical models). Nevertheless, it proposes a robust, transdisciplinary, and reproducible approach, shedding light on the reputational effects of a major industrial crisis on a global scale.

It is imperative to acknowledge the limitations inherent to the sentiment analysis tools utilized in this study.

VADER, a widely utilized tool for the analysis of social media data, operates by employing a static emotional lexicon and a set of heuristic rules. This methodological

framework renders VADER less adept at deciphering complex linguistic expressions, including but not limited to sarcasm, irony, and implicit meaning. Furthermore, VADER is optimized for English and may not yield reliable results with multilingual content unless adapted.

Conversely, SentiART employs semantic vector models to assess emotional content, a methodology that is particularly well-suited for literary texts. However, this method can be computationally intensive and requires well-trained word embeddings, which may not always capture nuances in informal or technical discussions. The performance of the system may also be subject to variation depending on the structure and domain of the input texts.

These limitations indicate that while both tools offer valuable insights, the results must be interpreted with caution, especially in contexts involving implicit sentiment or domain-specific jargon.

It is imperative to acknowledge the inherent demographic bias in Reddit as a data source. The platform exhibits a tendency to overrepresent younger users, those who are technologically proficient, and those who primarily speak English. This demographic distribution may impede the generalizability of the findings and should be considered during the interpretation of public sentiment trends (Duggan & Smith, 2013).

Furthermore, while some correlations between sentiment and popularity were observed, these findings should not be overinterpreted. The capacity of sentiment analysis to predict future events is constrained by temporal delays, platform-specific dynamics, and exogenous events that influence public discourse. To validate sentiment as a reliable forecasting tool, further longitudinal and controlled studies are necessary.

5.4. Suggestions for further research

This study establishes a foundation for future research, particularly concerning the integration of advanced natural language processing models and the expansion of the study to encompass other types of crises that affect global supply chains.

Integrating sentiment data with operational indicators, such as product availability or sales trends, would also be relevant. The establishment of such a connection between discourse and performance could facilitate the assessment of the practical implications of public perception.

Another promising direction of research lies in the longitudinal study of sentiment dynamics across multiple crises. For instance, a comparative analysis of consumer reactions during the semiconductor shortage with those observed during other disruptions (e.g., the pandemic, energy shortages, or geopolitical conflicts) could yield insight into the consistency or variability of brand perception under pressure. A comparative analysis of these practices could identify the best practices in crisis communication that are effective regardless of sector or crisis type.

Subsequent studies could investigate how consumer sentiment varies across regions or cultures by incorporating data from non-English media and social platforms. This would facilitate the identification of communication approaches that are more effective in specific cultural contexts.

Ultimately, Subsequent studies could investigate the role of visual content, including images, videos, and memes, in shaping public sentiment during crises. On social media platforms such as Twitter and Instagram, such content disseminates expeditiously and elicits profound emotional responses, thereby rendering visual sentiment analysis a valuable extension to text-based approaches.

6. Bibliography

Biolab. (2018, December 11). *Orange3 text mining documentation*. Biolab.

<https://dac.lip6.fr/wp-content/uploads/2019/01/orange3-text.pdf>

Black, B. (2021, June 23). How the Internet of Things Could Help Solve the Chip Shortage.

SupplyChainBrain. <https://www.supplychainbrain.com/blogs/1-think-tank/post/33243-how-iot-can-help-solve-the-computer-chip-shortage>

Bundy, J., Pfarrer, M. D., Short, C. E., & Coombs, W. T. (2017). Crises and crisis management: Integration, interpretation, and research development. *Journal of Management*, 43(6), 1661–1692. <https://doi.org/10.1177/0149206316680030>

Coombs, W. T. (2007). Protecting organization reputations during a crisis: The development and application of situational crisis communication theory. *Springer Nature Link*, 10(3), 163–176. <https://doi.org/10.1057/palgrave.crr.1550049>

Cserhat, N. (2024, February 13). Supply chain transparency: How technology enhances visibility and builds consumer trust. *Morpheus Network*.

<https://news.morpheus.network/supply-chain-transparency-how-technology-enhances-visibility-and-builds-consumer-trust-55bc87b91d14>

Deloitte. (2021). *Anchor of global semiconductor: Asia Pacific takes off*. Deloitte.

<https://www2.deloitte.com/cn/en/pages/technology-media-and-telecommunications/articles/tmt-anchor-of-global-semiconductor-industry-2021.html>

Duggan, M., & Smith, A. (2013). *6% of online adults are Reddit users*. Pew Research Center.

<https://www.pewresearch.org/internet/2013/07/03/6-of-online-adults-are-reddit-users/>

Greaves, F., Ramirez-Cano, D., Millett, C., Darzi, A., & Donaldson, L. (2013). Use of sentiment analysis for capturing patient experience from free-text comments posted online. *Journal of Medical Internet Research*, 15(11). <https://doi.org/10.2196/jmir.2721>

Holloway, S. (2024). Exploring the influence of supply chain disruption management on brand resilience and consumer perception. *Preprints*.

https://www.researchgate.net/publication/381674665_Exploring_the_Influence_of_Supply_Chain_Disruption_Management_on_Brand_Resilience_and_Consumer_Perception

Jacobs, A. M. (2019). Computing the affective-aesthetic potential of literary texts. *Scientific Study of Literature*, 9(1), 4–45. <https://doi.org/10.3390/ai1010002>

Jézéquel, B., Gérard, P. (2017, February 28). Communiquer en période de crise. *Emarketing*. <https://www.e-marketing.fr/Thematique/academie-1078/fiche-outils-10154/COMMUnIQUER-En-PeRIODE-DE-CRISE-314762.htm>

Kurniawan, R., & Anggraini, W. (2023). Sentiment analysis using the VADER model for assessing company services based on posts on social media. *ResearchGate*. <https://doi.org/10.2478/seeur-2023-0043>

Novotny, T. (2024, August 6). *What is supply chain visibility and why is it important?* Sedex Blog. <https://www.sedex.com/blog/what-is-supply-chain-visibility-and-why-is-it-important/>

Özkanlısoy, Ö. (2024). Supply chain visibility: A theoretical framework. *International Journal for Multidisciplinary Research*, 6(5), 1–15. https://www.researchgate.net/publication/383819149_Supply_Chain_Visibility_A_Theoretical_Framework

Semiconductor Industry Association. (2016). *The global semiconductor value chain: How an interconnected industry promotes innovation and growth*. Semiconductor Industry Association. <https://www.semiconductors.org/wp-content/uploads/2018/06/SIA-Beyond-Borders-Report-FINAL-May-6-1.pdf>

Semiconductor Industry Association. (2020). *From microchips to medical devices*. Semiconductor Industry Association. <https://www.semiconductors.org/wp-content/uploads/2020/10/From-Microchips-to-Medical-Devices-SIA-White-Paper.pdf>

Semiconductor Industry Association. (2021a). *Chipmakers are ramping up production to address semiconductor shortage. Here's why that takes time*. Semiconductor Industry Association. <https://www.semiconductors.org/chipmakers-are-ramping-up-production-to-address-semiconductor-shortage-heres-why-that-takes-time/>

Semiconductor Industry Association. (2021b). *State of the U.S. Semiconductor Industry*. Semiconductor Industry Association. <https://www.semiconductors.org/wp-content/uploads/2021/09/2021-SIA-State-of-the-Industry-Report.pdf>

Semiconductor Industry Association. (2021c). *Strengthening the global semiconductor supply chain in an uncertain era*. Semiconductor Industry Association.

https://www.semiconductors.org/wp-content/uploads/2021/05/BCG-x-SIA-Strengthening-the-Global-Semiconductor-Value-Chain-April-2021_1.pdf

Sheikhattar, M. R., Nezafati, N., & Shokouhyar, S. (2022). A thematic analysis-based model for identifying the impacts of natural crises on a supply chain for service integrity: A text analysis approach. *Environmental Science and Pollution Research*, 29(52), 79413–79433.

<https://doi.org/10.1007/s11356-022-21380-x>

Sun, C., & Rose, T. (2015). Supply chain complexity in the semiconductor industry: Assessment from system view and the impact of changes. *Procedia Manufacturing*, 3, 1210-1215. <https://doi.org/10.1016/j.ifacol.2015.06.249>

University of Ljubljana. (2020, April). *Data mining: Working notes for the hands-on course at the Faculty of Economics at University of Ljubljana*. University of Ljubljana.

<https://file.birolab.si/notes/20-ul-ef-data-science-marketing.pdf>

Wakade, S., Shekar, C., Lyszka, K-J., & Chan, C-C. (n.d.). *Text mining for sentiment analysis of Twitter data*. The University of Akron. <https://worldcomp-proceedings.com/proc/p2012/IKE3997.pdf>

Wassen, M., Adel, E., & Laoucine, K. (2022). The global semiconductor chip shortage: Causes, implications, and potential remedies. *Procedia Computer Science*, 207, 476-483.

<https://doi.org/10.1016/j.ifacol.2022.09.439>

Xiong, W., Wu, D. D., & Yeung, J. H. Y. (2024). Semiconductor supply chain resilience and disruption: Insights, mitigation, and future directions. *International Journal of Production Research*, 1–24. <https://doi.org/10.1080/00207543.2024.2387074>

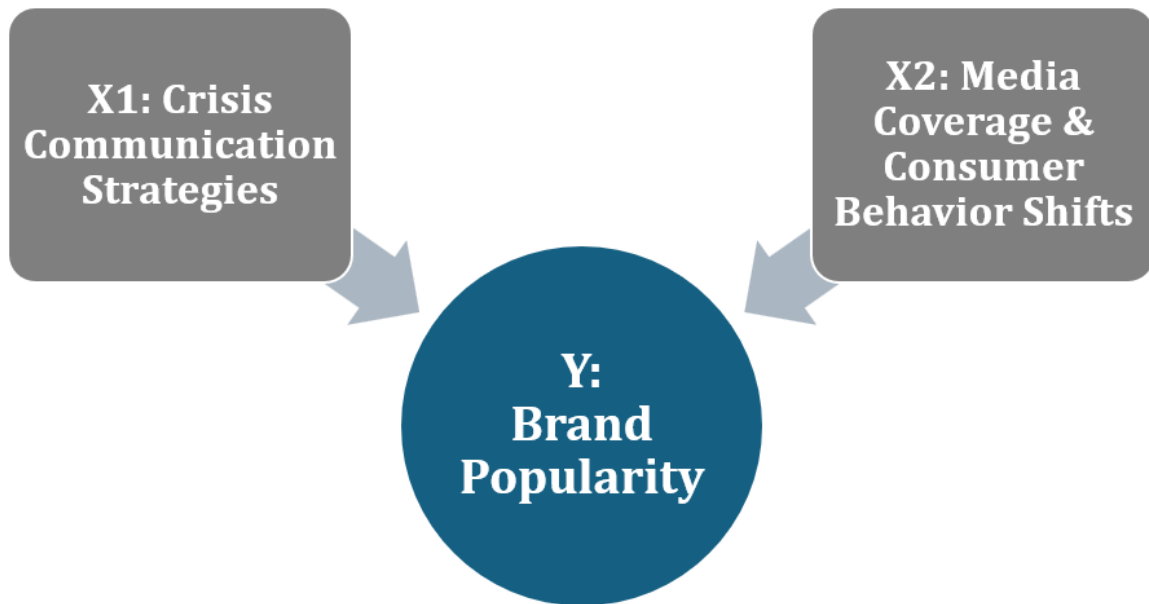
Zolghadri, M., Besbes, M., Bourgeois, V., & Saad, E. (2023). Micro-electronic chips shortages and obsolescence: An empirical study. *Procedia CIRP*, 127, 1570-1575.

<https://doi.org/10.1016/j.procir.2023.09.216>

7. Appendix

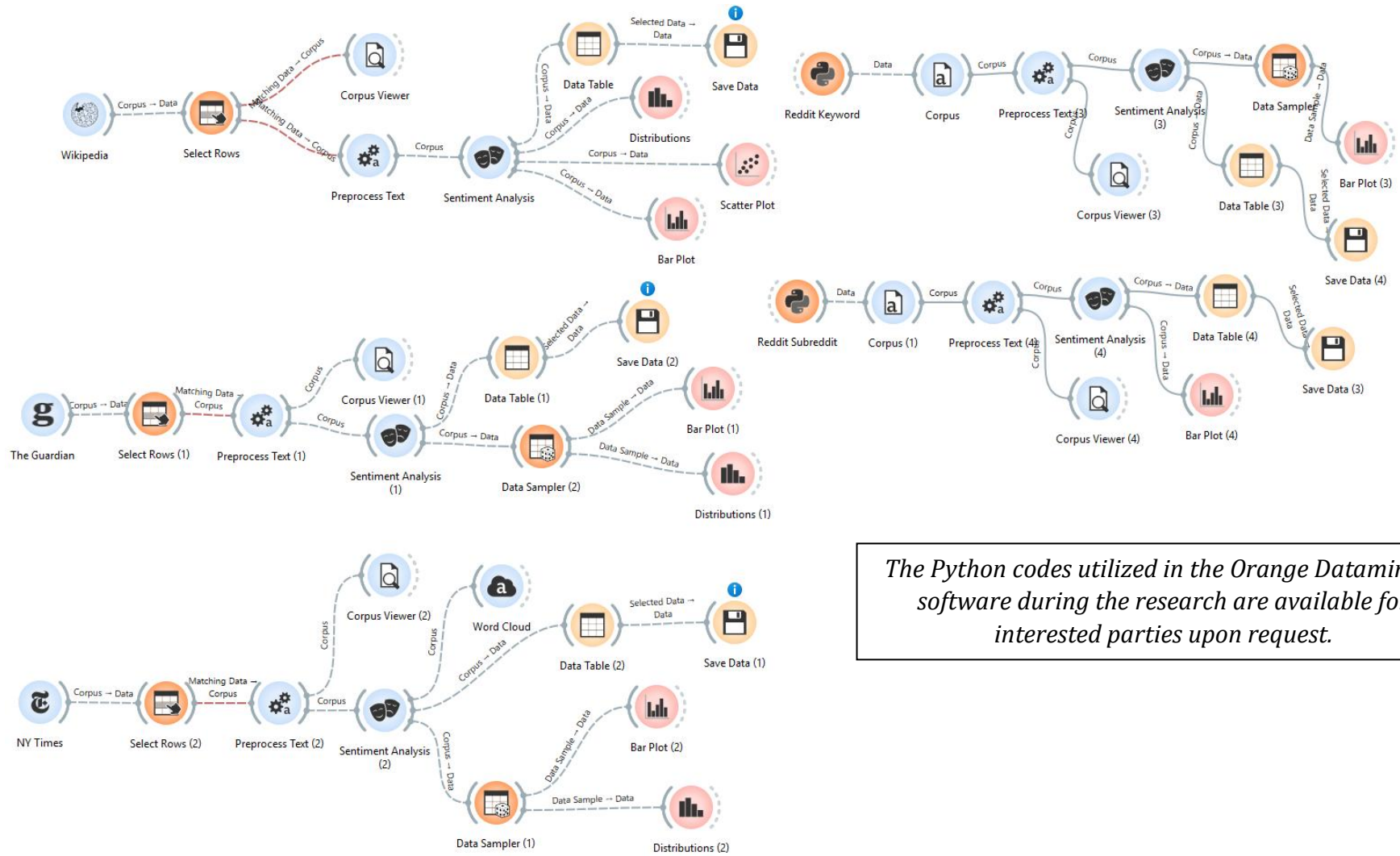
Appendix A: Research Model

Variables influencing brand popularity



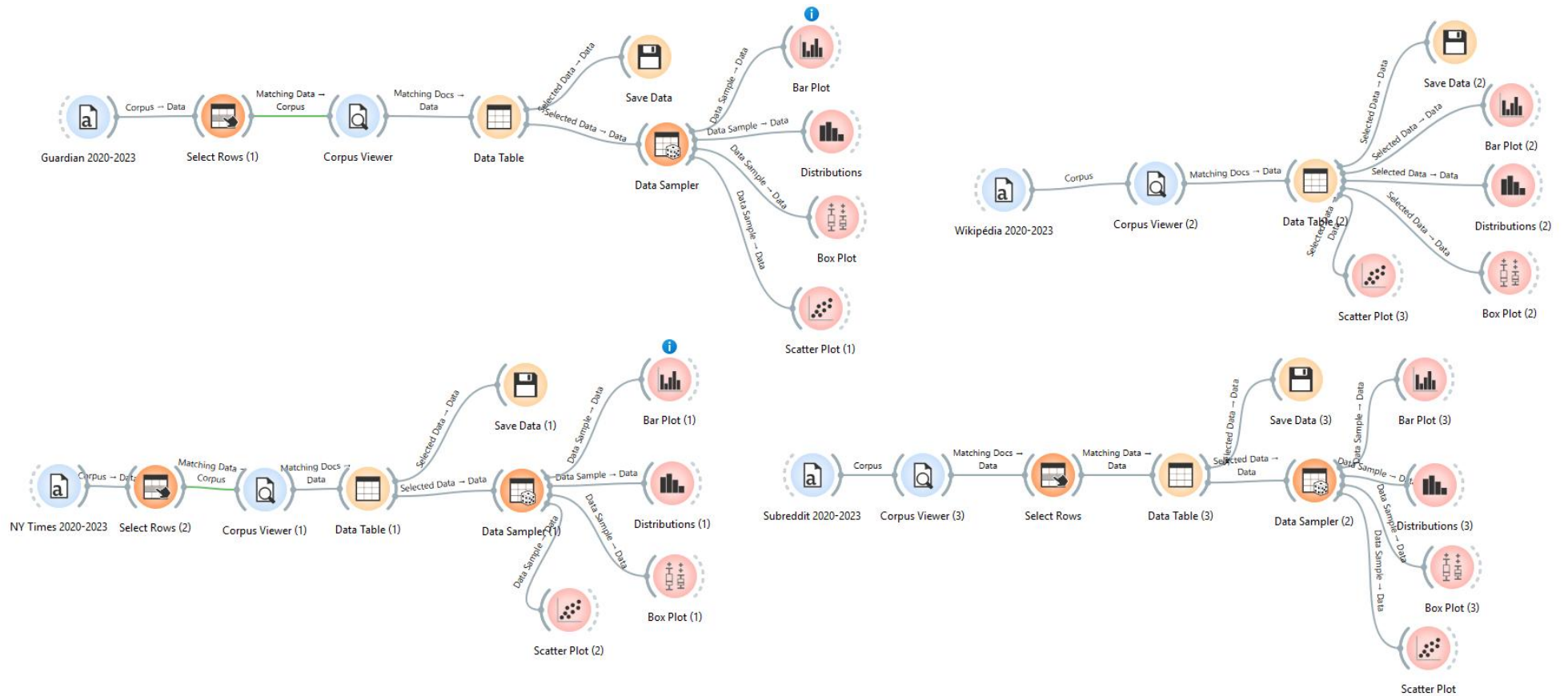
Appendix B: Orange workflows

Data extraction and initial analysis



The Python codes utilized in the Orange Datamining software during the research are available for interested parties upon request.

Extensive data analysis and concatenation between 2020 and 2023

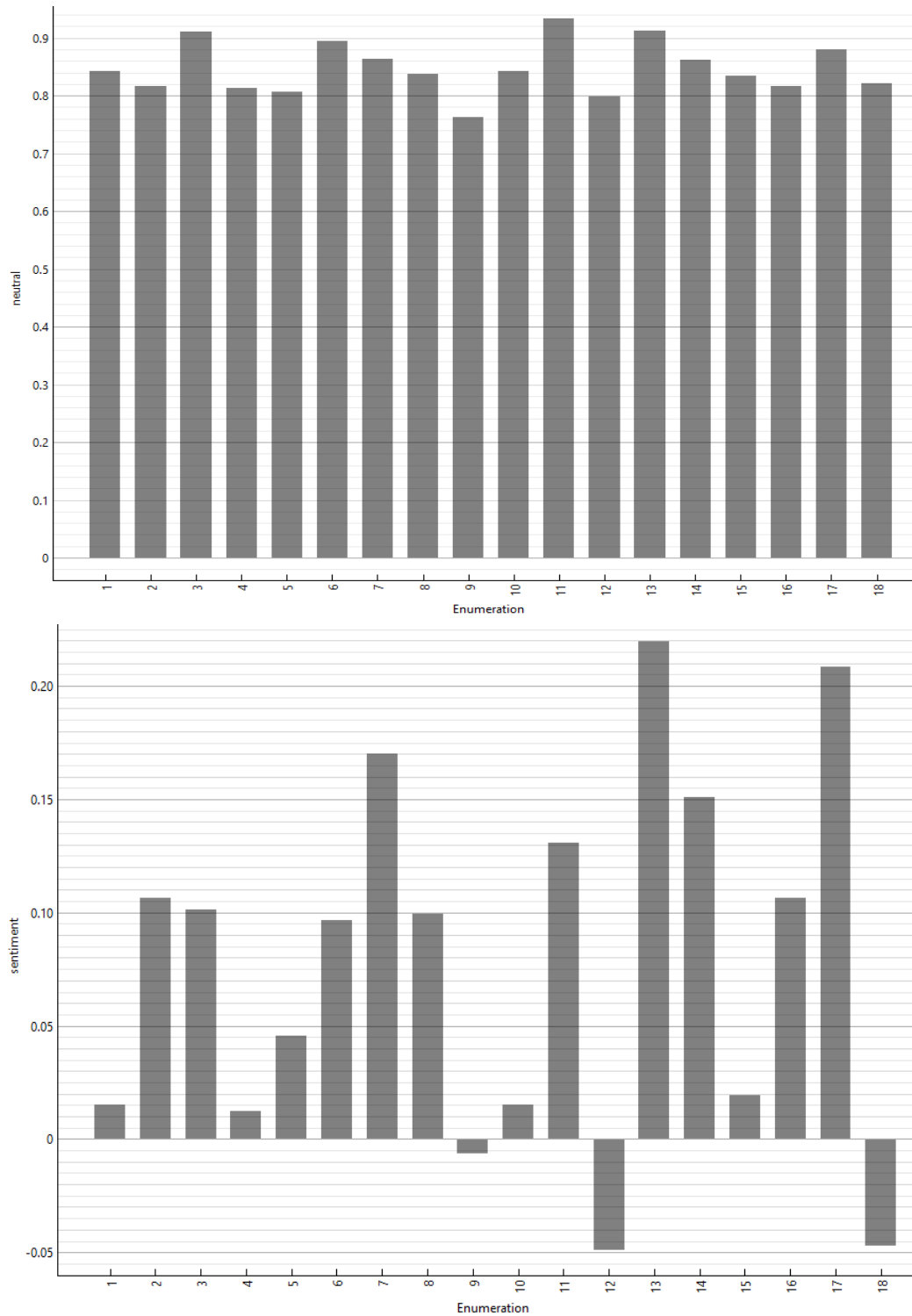


Detailed description of widgets used

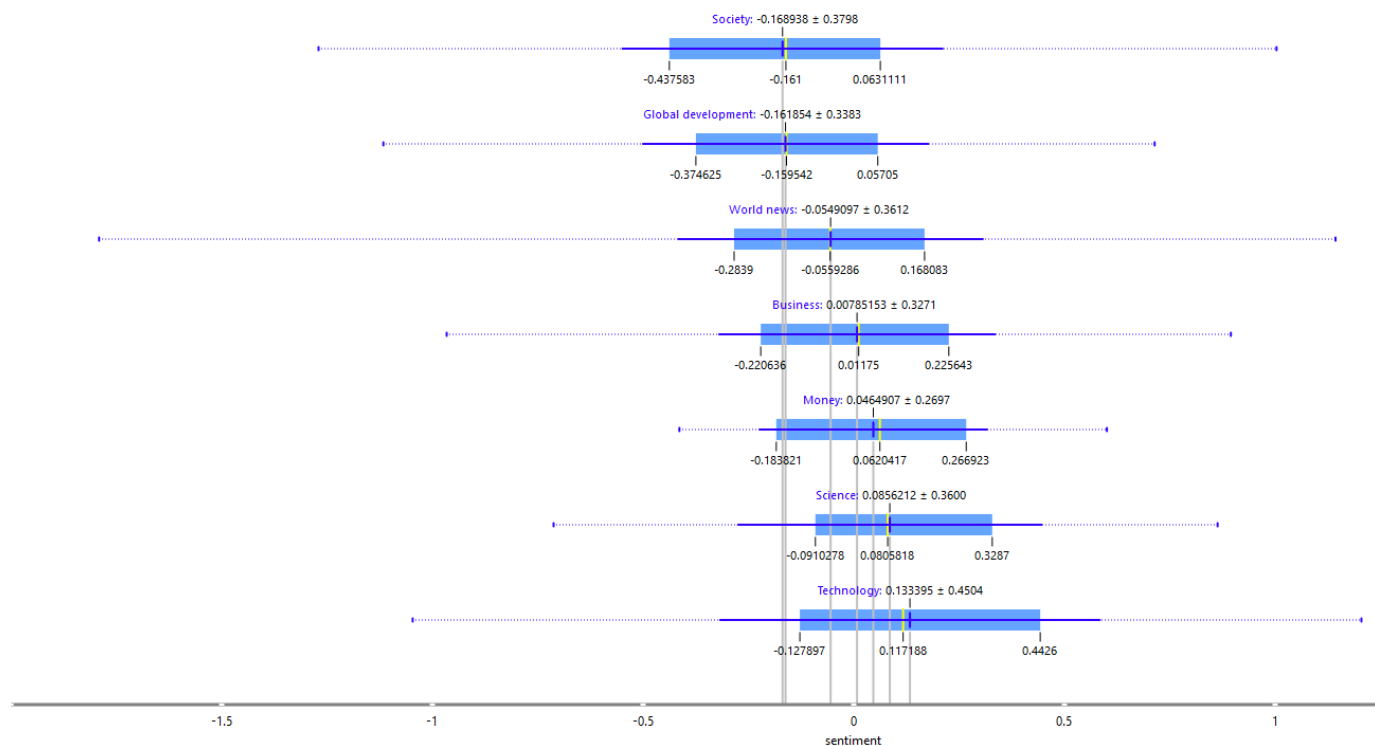
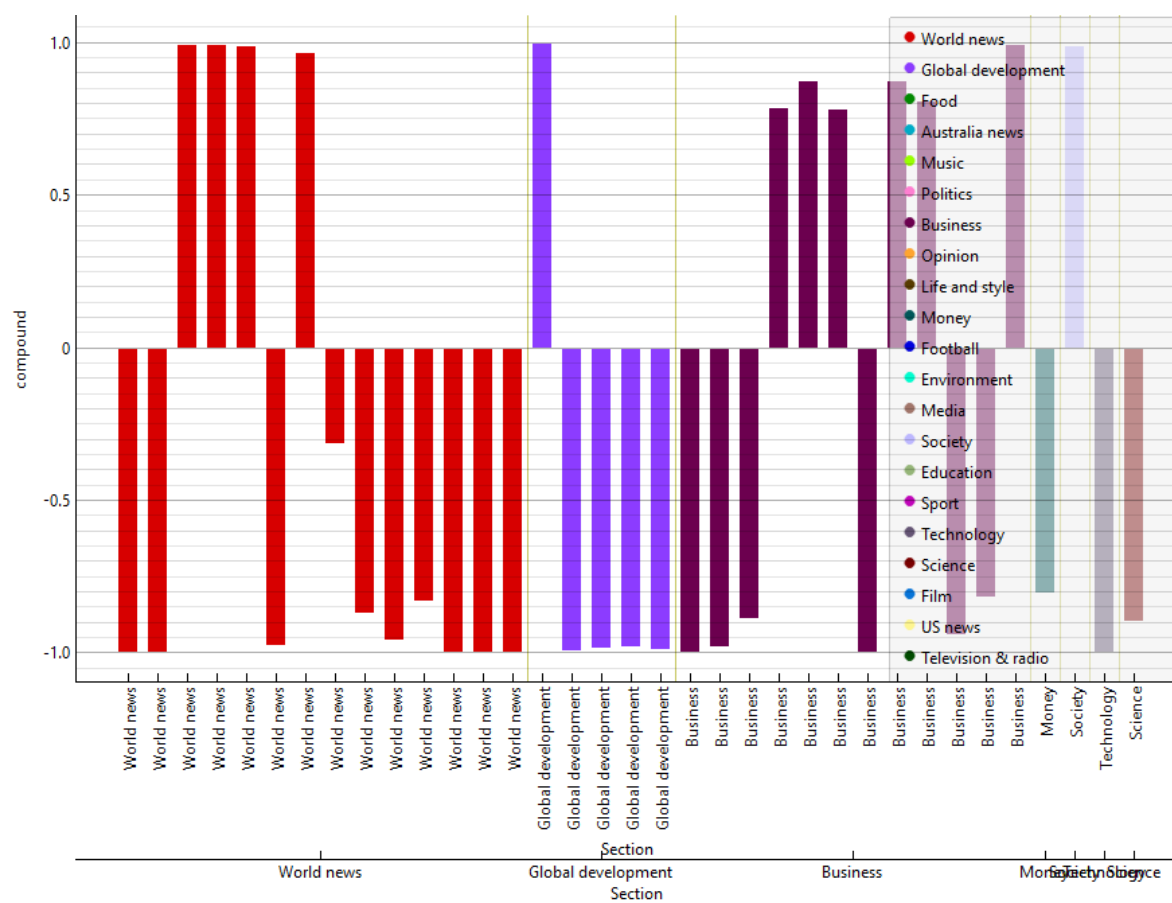
Step	Function	Used for	Specific Notes
Corpus / Input	Import raw textual data	All sources	for Reddit, use of two python scripts
Select Rows	Filter by publication year, brand, or topic keywords	Guardian, NY Times, Subreddit	Ensures temporal consistency (2020–2023) and thematic focus
Corpus Viewer	Inspect and validate the selection of documents	All except Reddit Keyword	Used both before and after filtering for quality control
Preprocess Text	Clean and standardize text (lowercasing, removing stop words and punctuation)	All sources	Crucial for ensuring input quality prior to sentiment scoring
Sentiment Analysis	Compute sentiment scores (VADER or SentiART)	All sources	VADER preferred for press and SentiART for Reddit where more semantic depth is useful
Data Table	View data, filter results, and prepare export	All sources	Used as the central node for export and branching into visualizations
Data Sampler	Reduce dataset size for performance and clarity in visualization	Reddit, NY Times, Guardian	Prevents overload in visual plots due to large volumes
Bar Plot	Compare sentiment scores between brands, timeframes, or platforms	All sources	Highlights variations in emotional perception across dimensions
Distributions	Show sentiment category proportions (positive/neutral/negative)	All sources	Useful to assess general tone in the dataset
Box Plot	Display spread and outliers in sentiment scores	All sources	Adds depth to understanding of score variability

Appendix C: Results for the first hypothesis

Sentiment Analysis of Wikipedia Articles

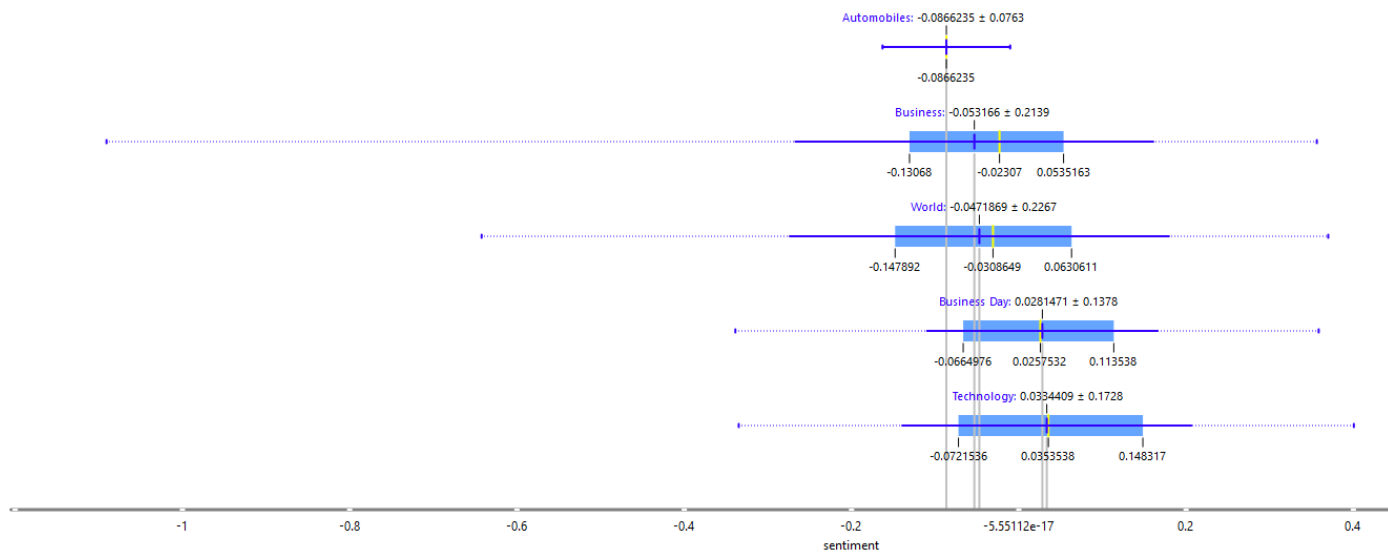
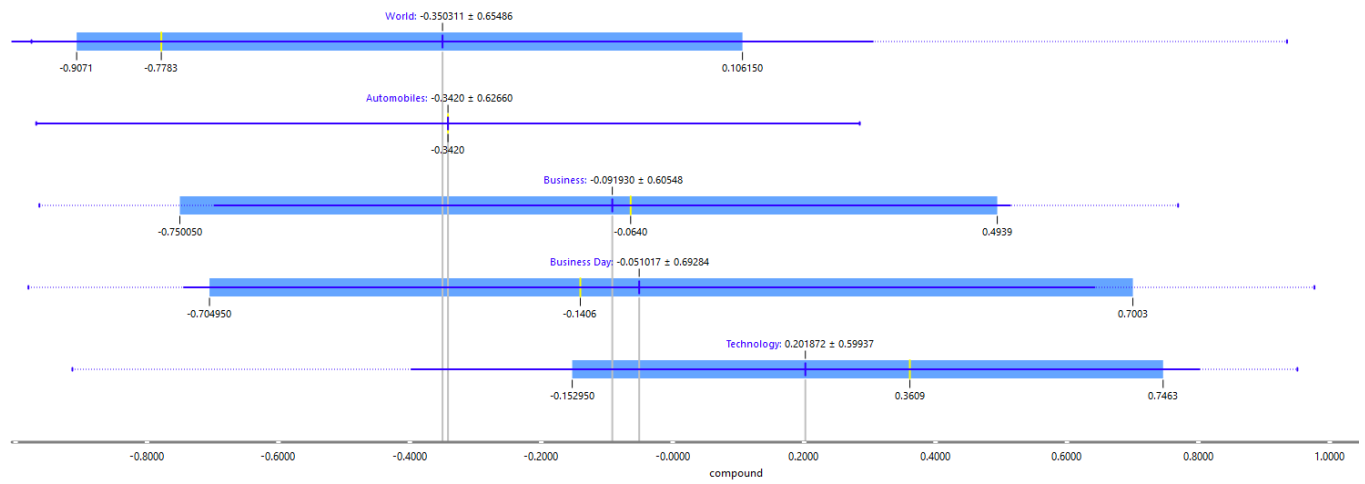


Sentiment Analysis of The Guardian Articles



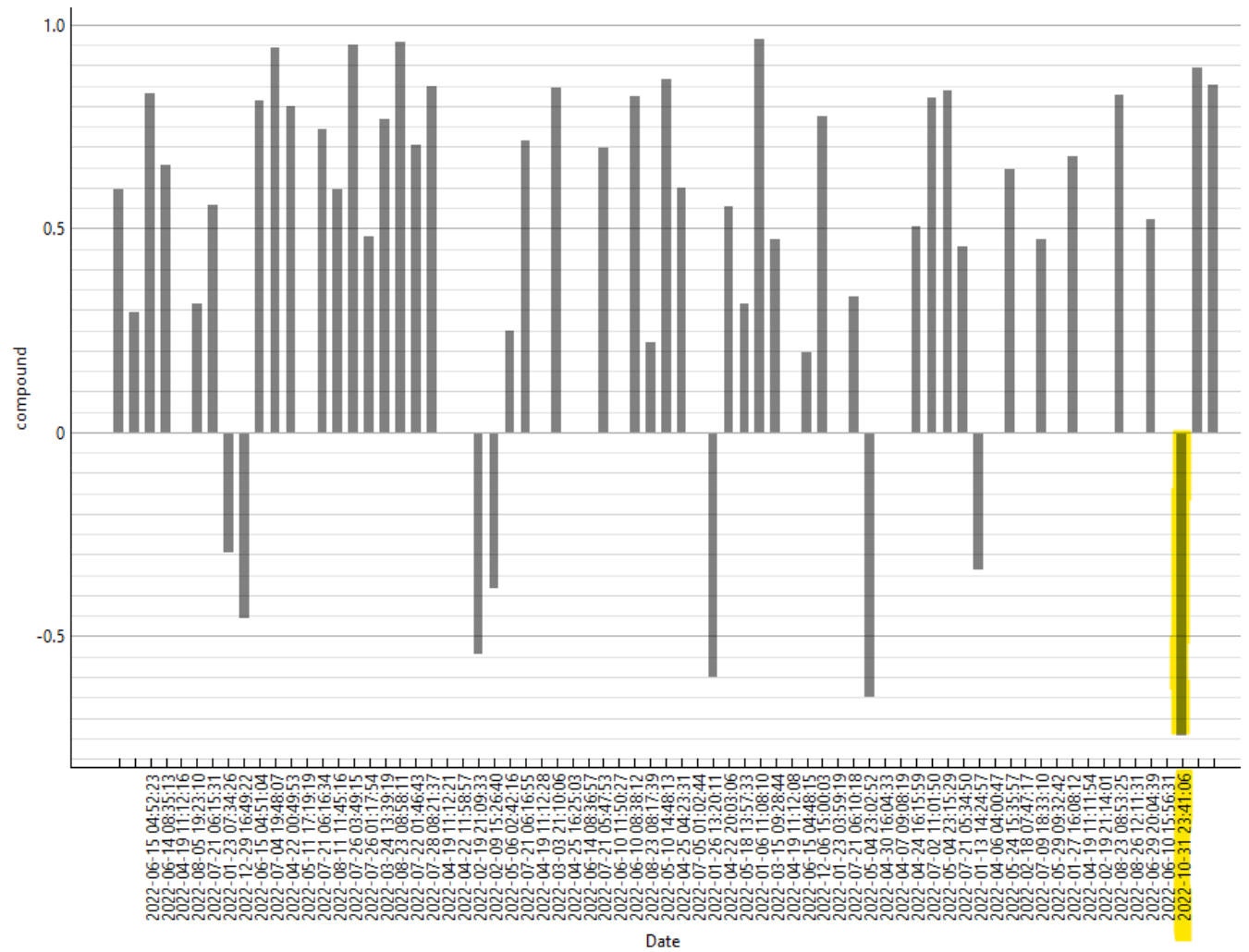
ANOVA: 11.608 ($p=0.000$, $N=1713$)

Sentiment Analysis of The New York Times Articles



Sentiment Trends on Reddit

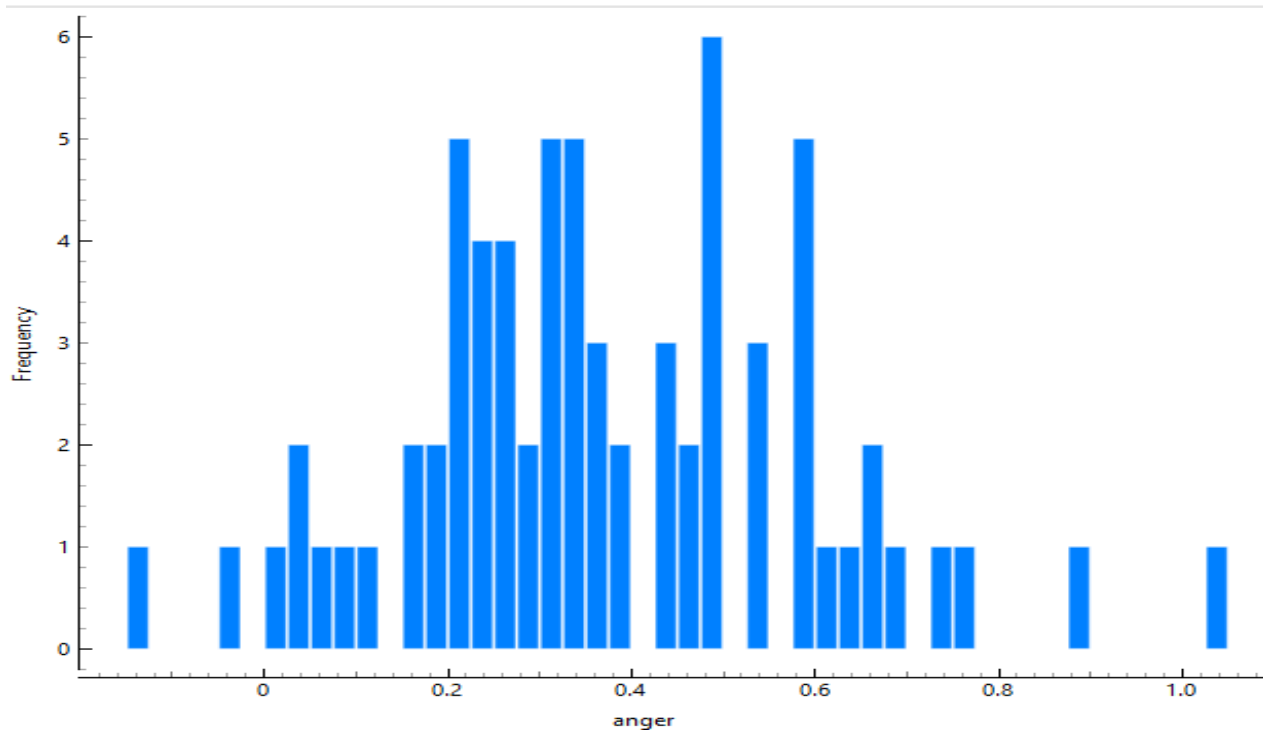
title	Subreddit	Author	Body	Date	Score	positive	negative	neutral	compound
1	HPB_Global	Iconic_Gamech...	Hpb is a ghost ...	2022-10-31 23:4...	2	0.000	0.328	0.672	-0.7435



title	Subreddit	Author	Body	Date	Score	positive	negative	neutral	compound
1	BrainChip_inc	[deleted]	Are BRN shares ...	2023-12-30 00:5...	1	0.500	0.000	0.500	0.6124

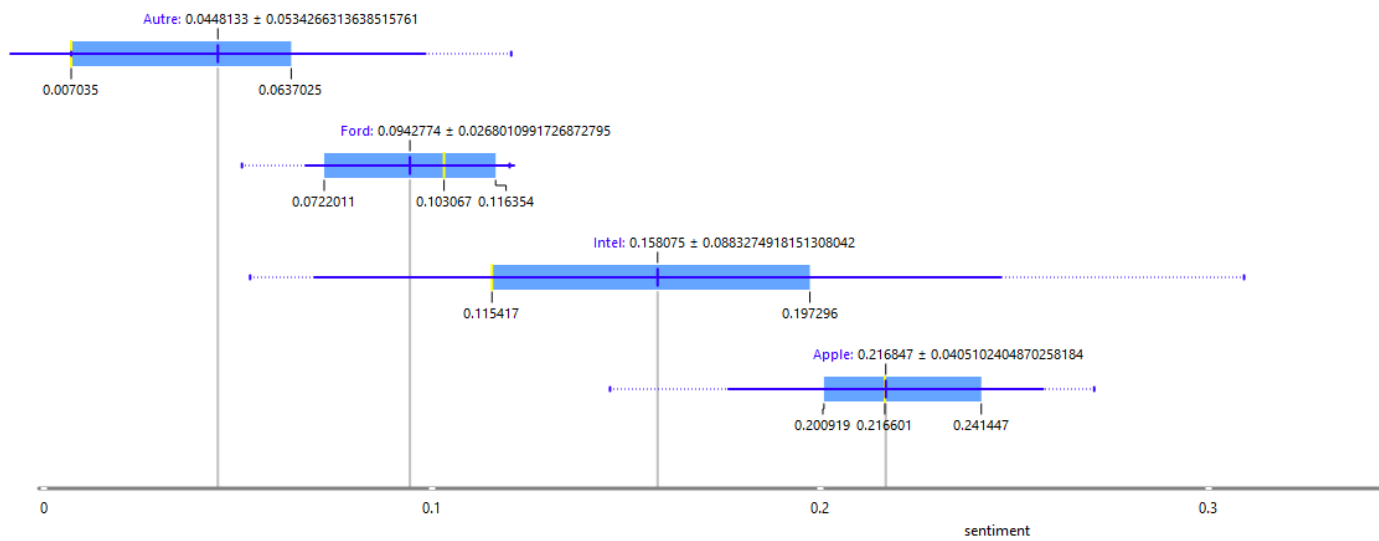
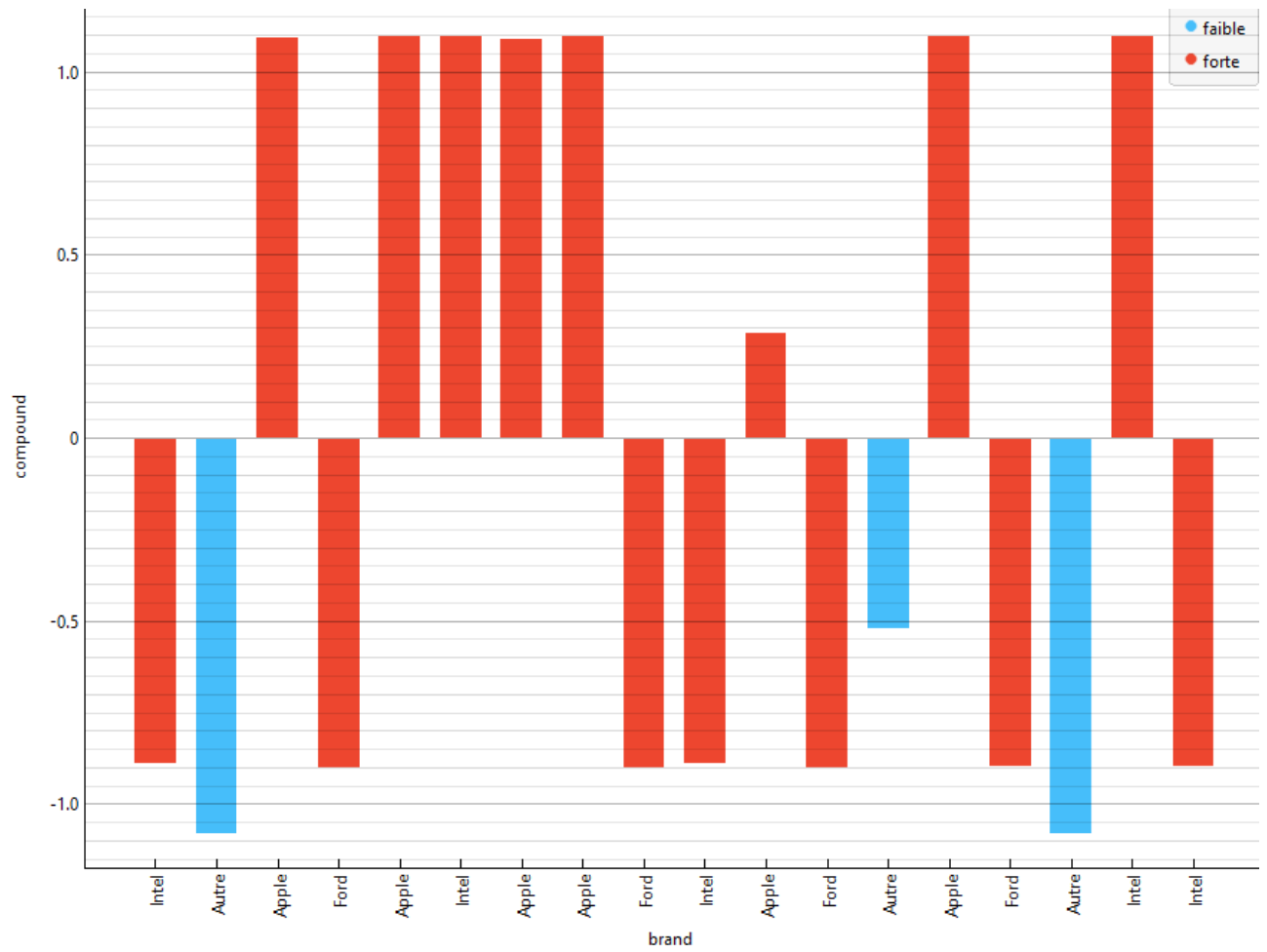
Appendix D: Results for the second hypothesis

Emotional Analysis on Reddit



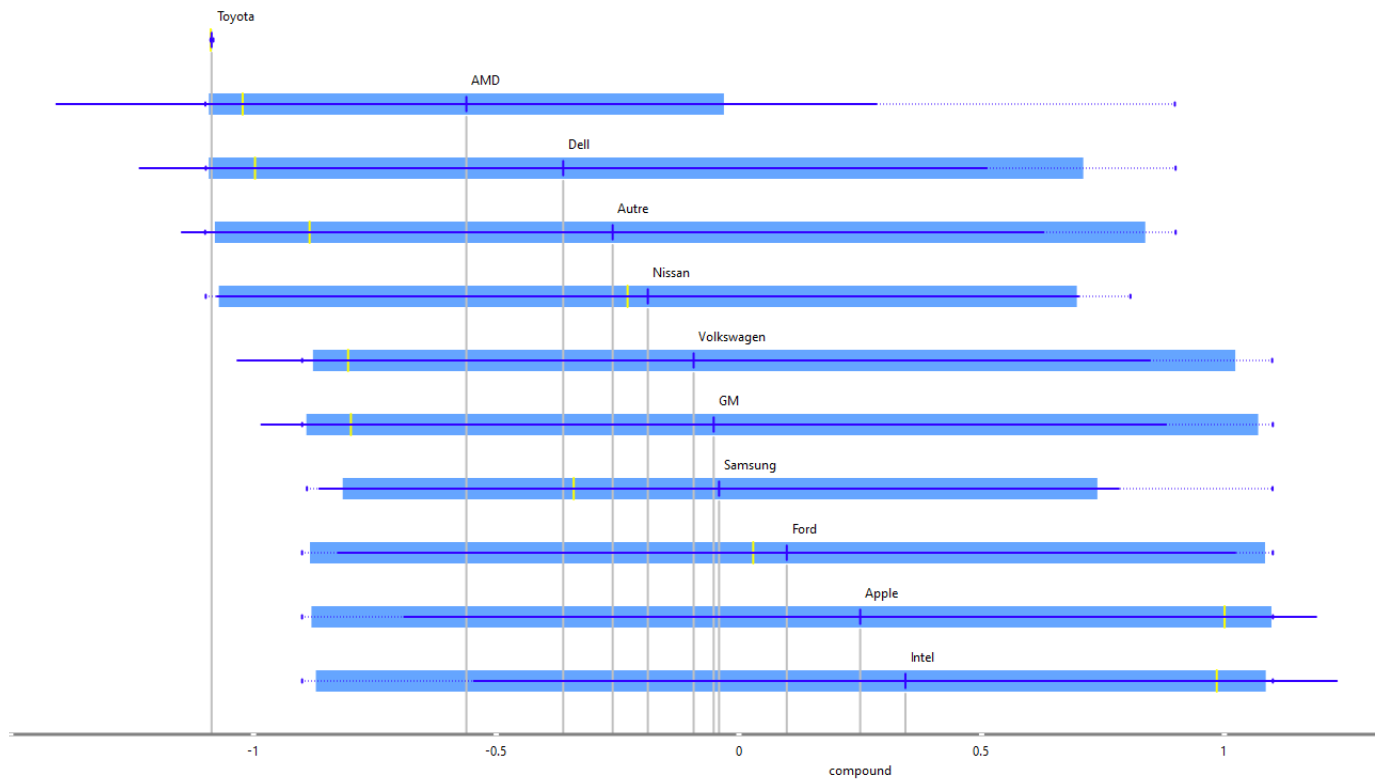
Appendix E: Results for the third hypothesis

Communication Strategies and Sentiment on Wikipedia

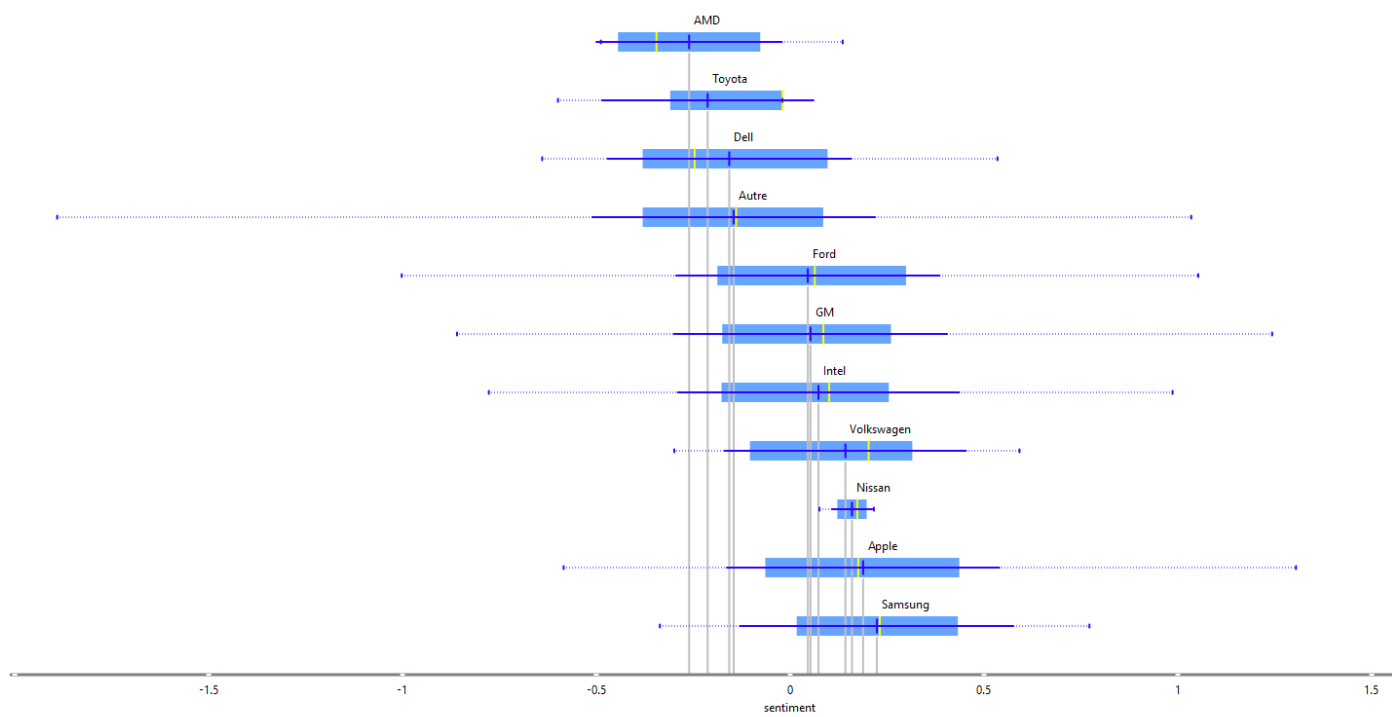


ANOVA: 5.591 (p=0.010, N=18)

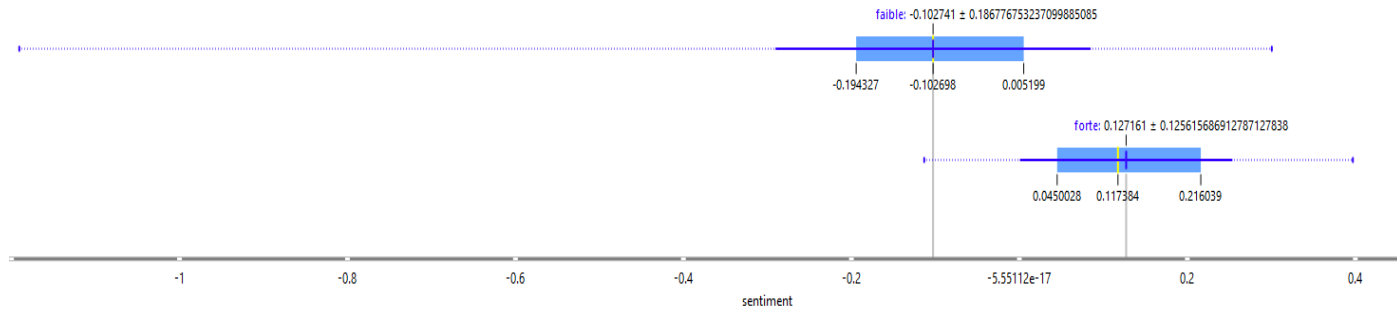
Communication Strategies and Sentiment in The Guardian



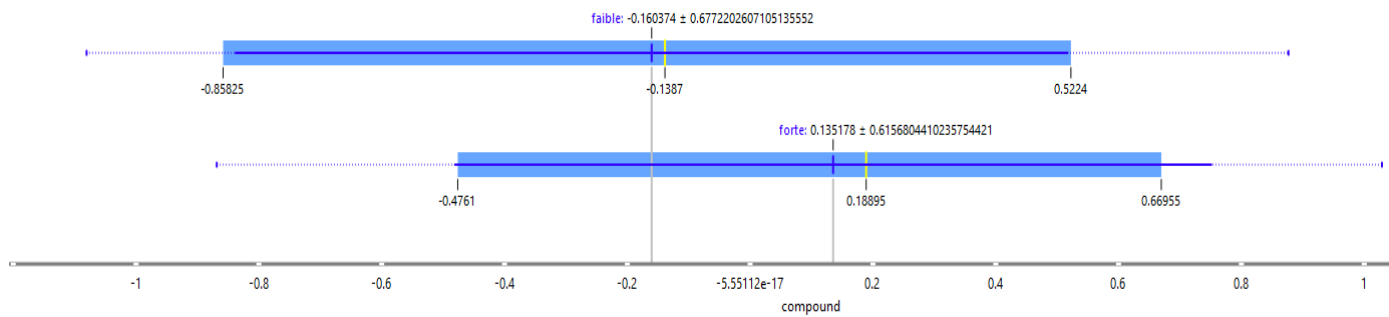
ANOVA: 10.126 ($p=0.000$, $N=1713$)



ANOVA: 19.517 ($p=0.000$, $N=1713$)

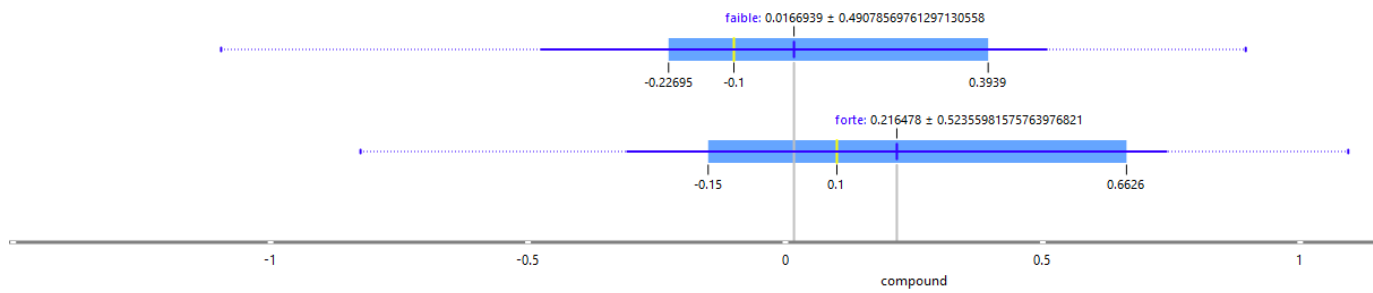
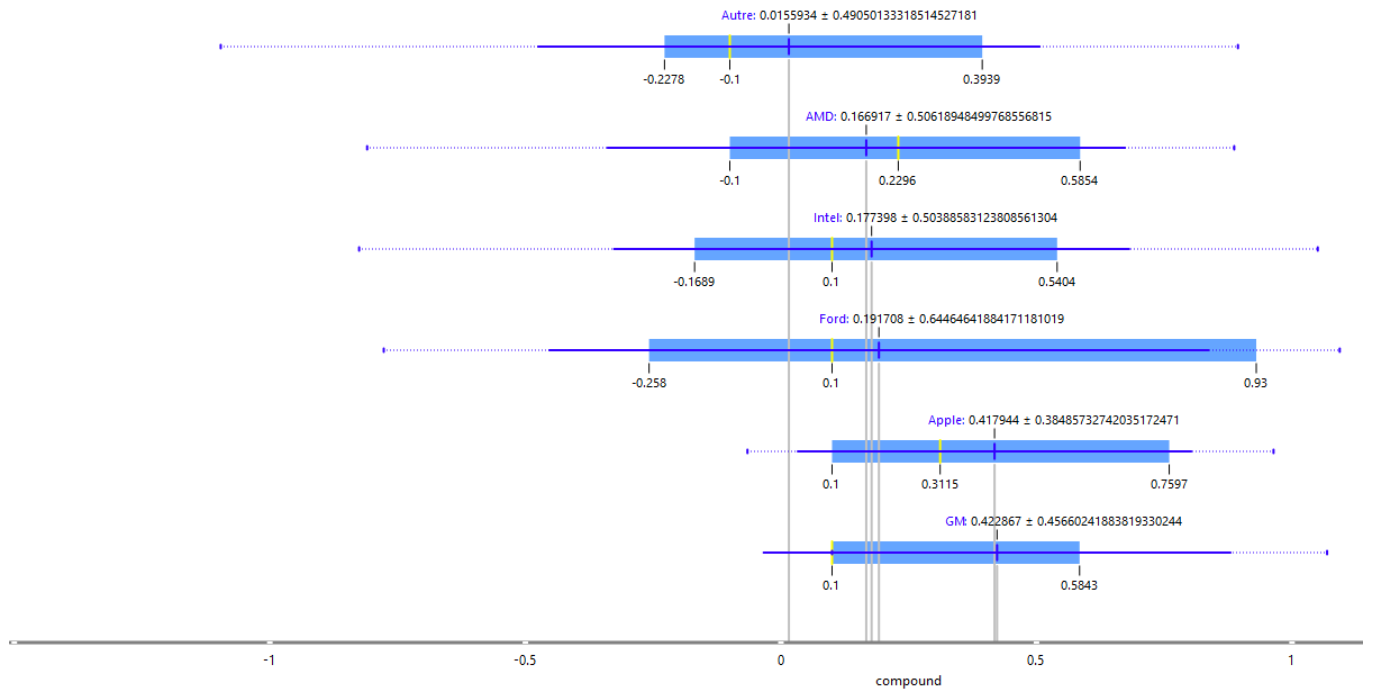
Communication Strategies and Sentiment in The New York Times

Student's t: 8.875 (p=0.000, N=228)



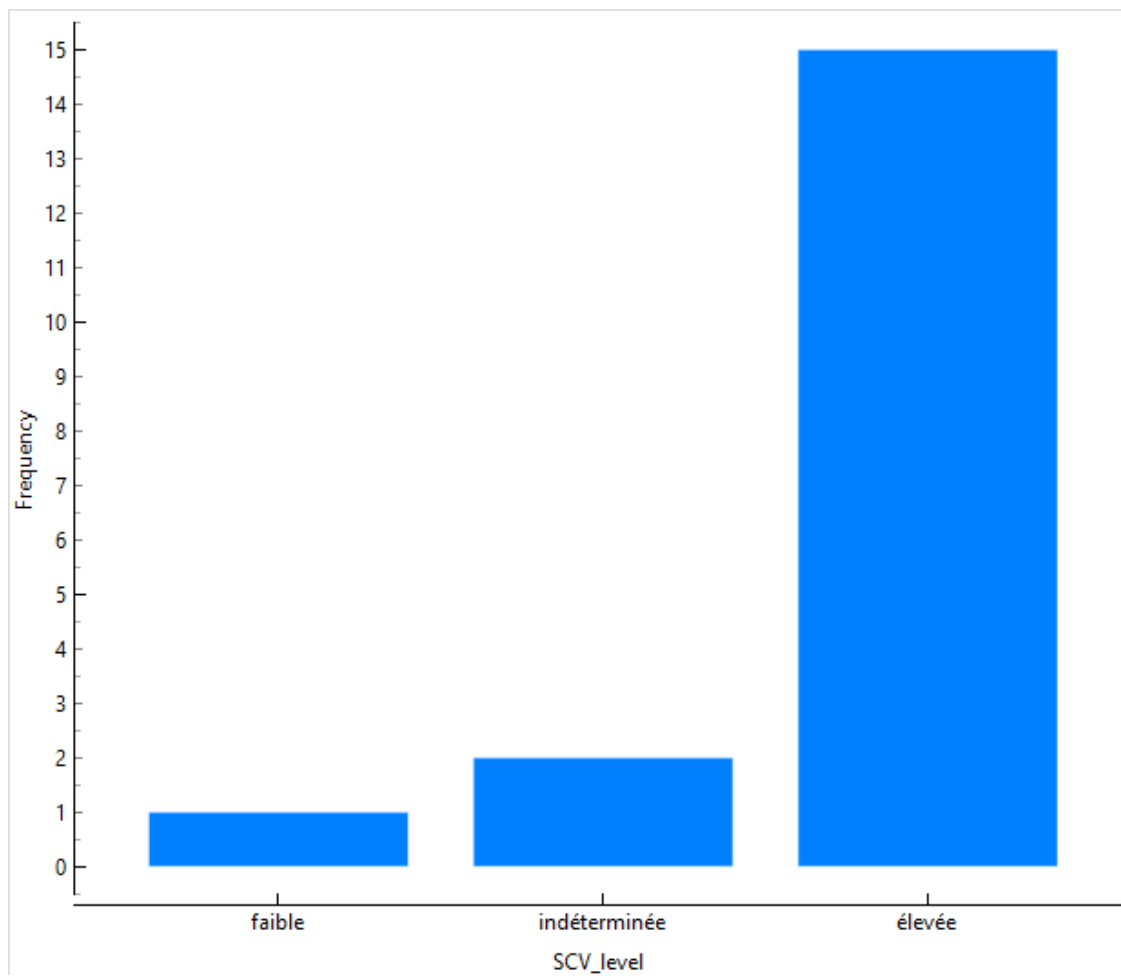
Student's t: 2.481 (p=0.017, N=228)

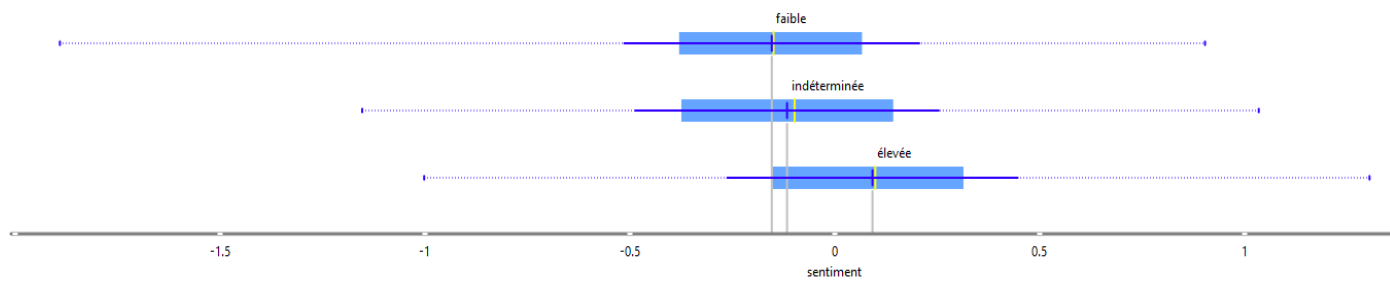
Communication Strategies and Sentiment on Reddit



Appendix F: Results for the fourth hypothesis

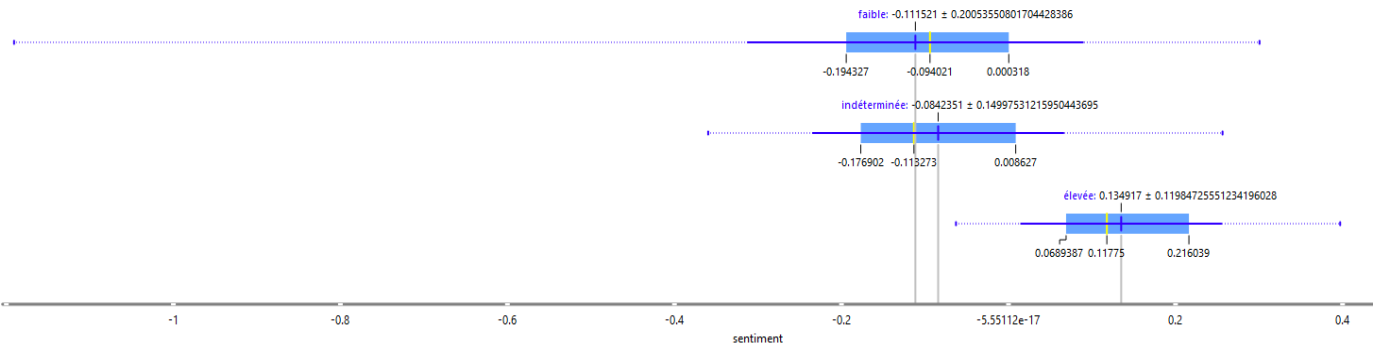
Impact of Supply Chain Visibility on Wikipedia



Impact of Supply Chain Visibility in The Guardian

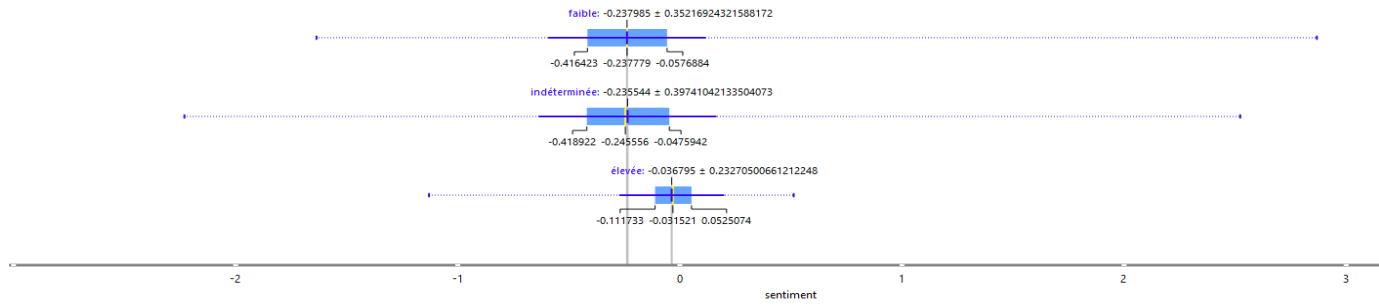
ANOVA: 83.998 (p=0.000, N=1713)

Impact of Supply Chain Visibility in The New York Times



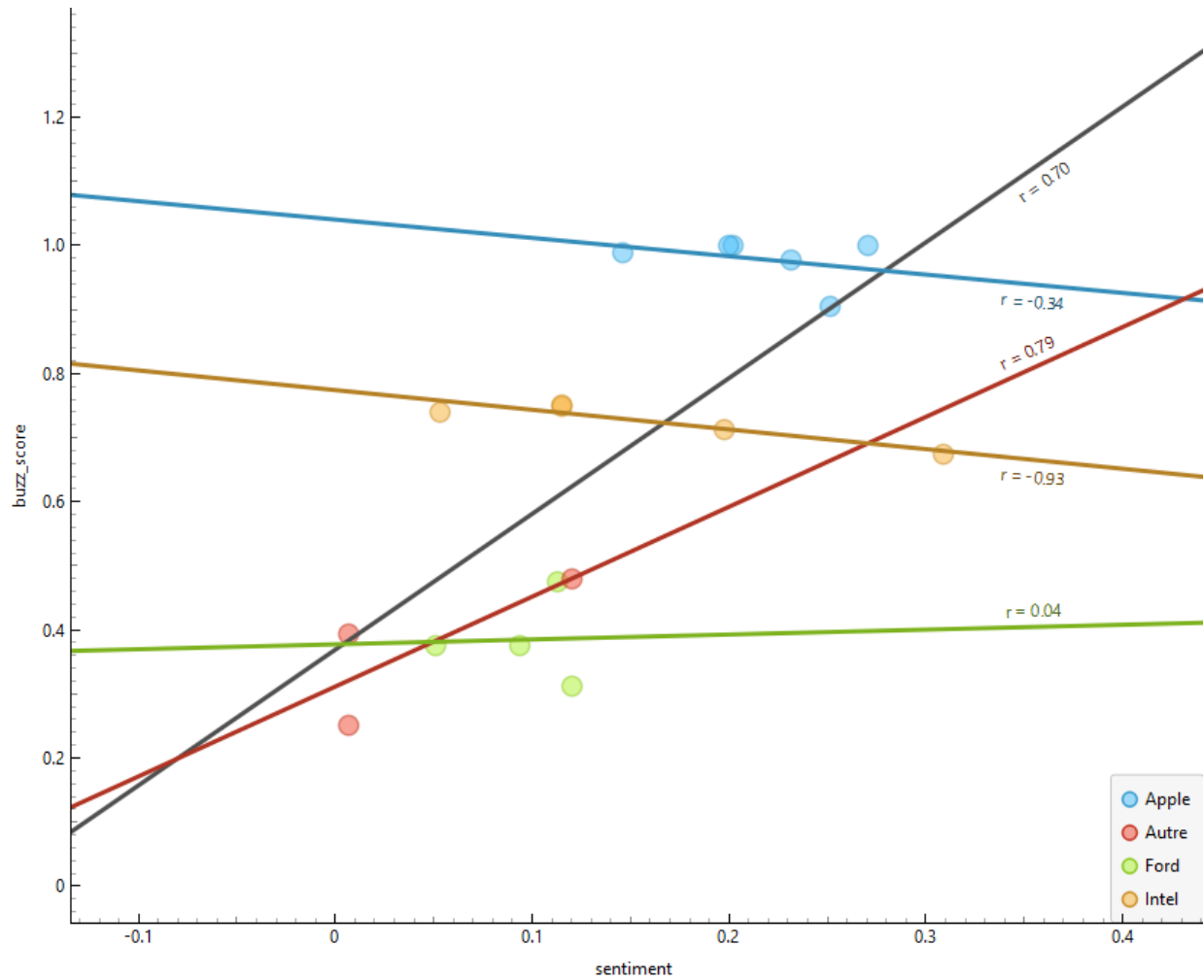
ANOVA: 23.978 (p=0.000, N=228)

Impact of Supply Chain Visibility on Reddit

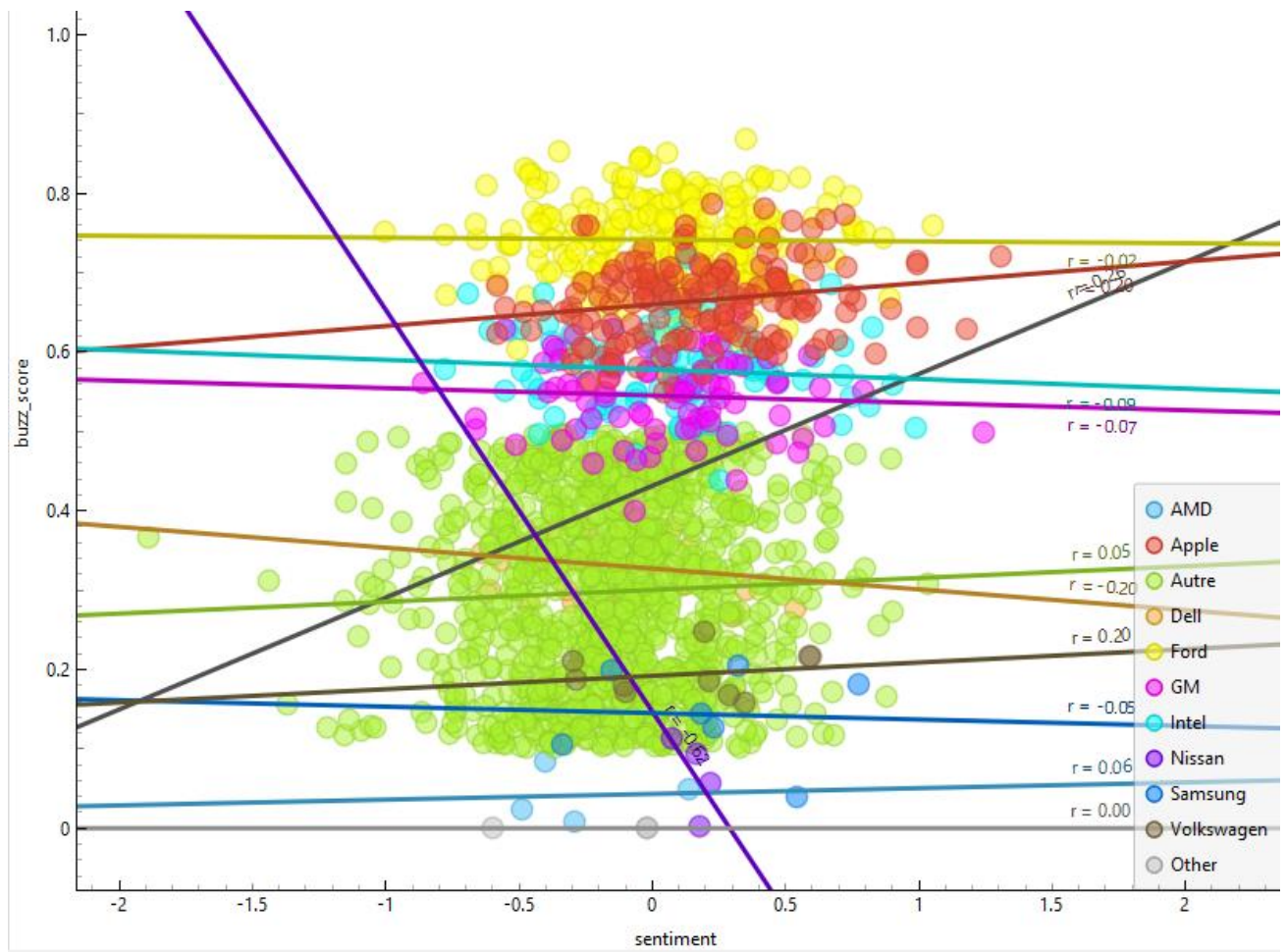


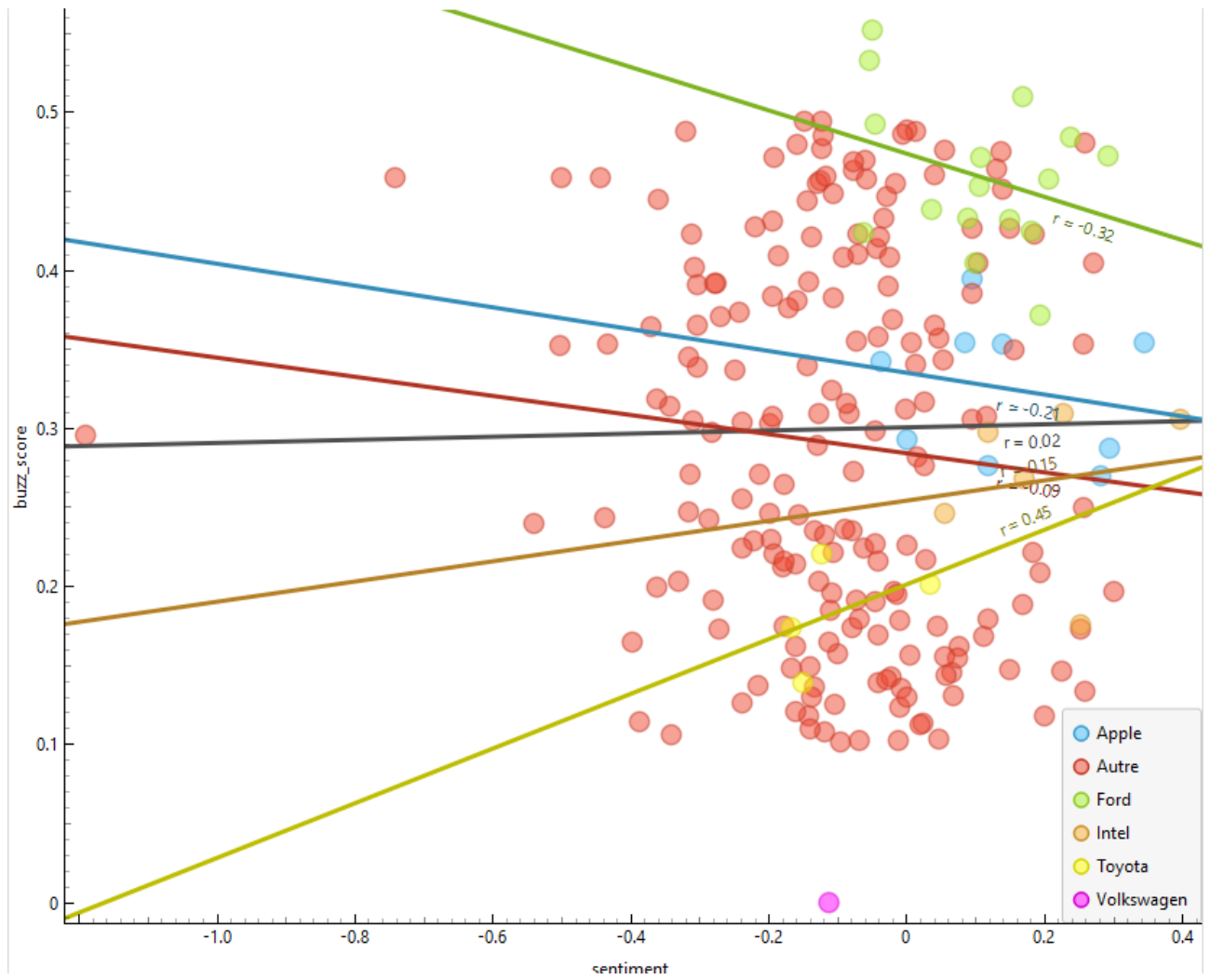
Appendix G: Results for the fifth hypothesis

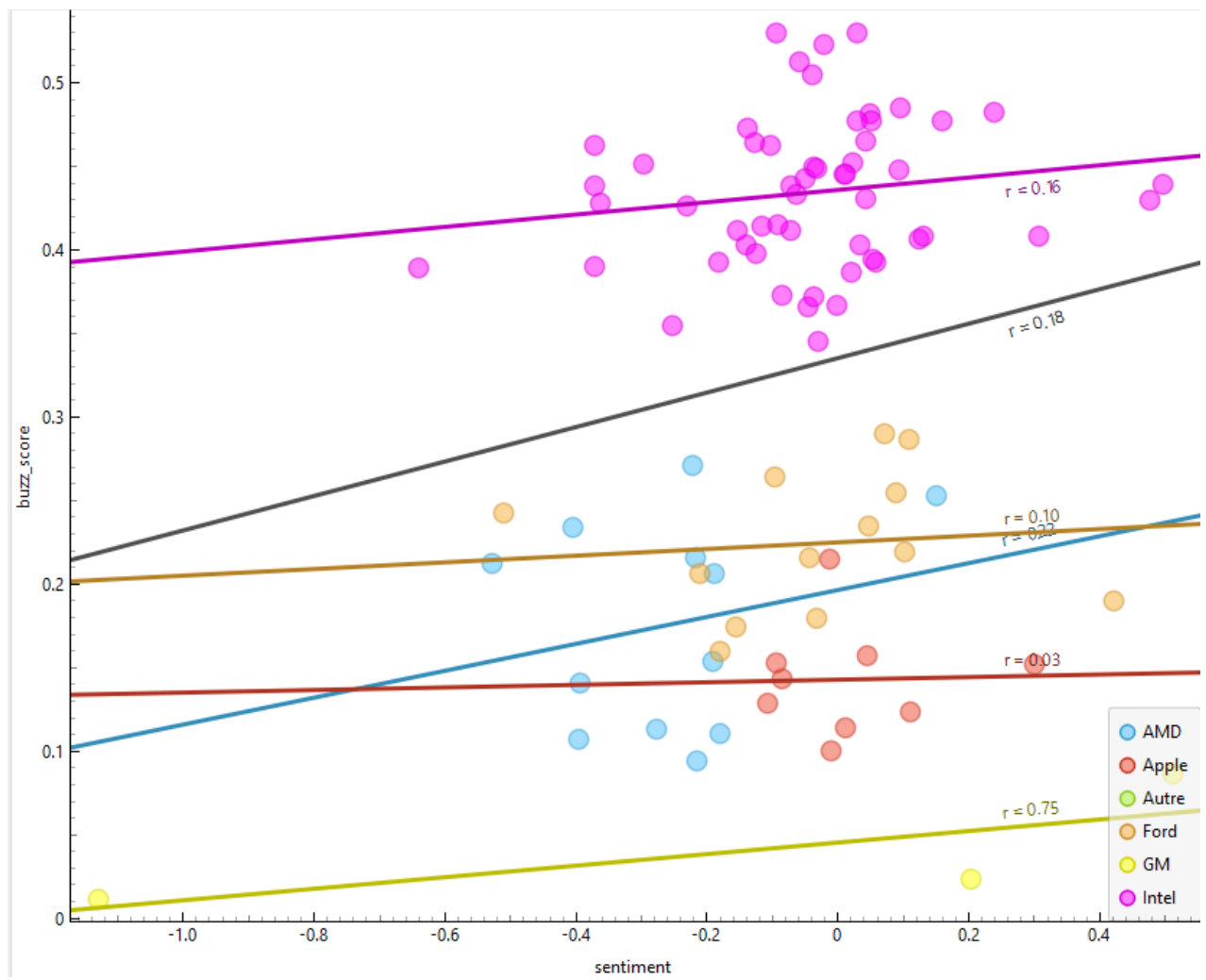
Sentiment-Popularity Correlations on Wikipedia



Sentiment-Popularity Correlations in The Guardian



Sentiment-Popularity Correlations in The New York Times

Sentiment-Popularity Correlations on Reddit

Abstract:

The present research examines the impact of the global semiconductor shortage (2020-2023) on brand perception, mobilizing sentiment analysis methods applied to a variety of digital sources. Five hypotheses were formulated to explore the effects of this crisis on brand popularity, communication strategy, sector perception, logistical transparency, and the predictive capacity of sentiment on the evolution of brand visibility. The study is grounded in quantitative analysis, employing data sourced from prominent digital platforms such as Reddit, Wikipedia, The Guardian, and The New York Times.

The VADER and SentiART algorithms were utilized to assess the emotional sentiment and the prevailing tone of the content. The Orange Data Mining platform was utilized to process all the mentioned elements, thereby enabling the development of evidence through the application of pattern recognition, comparative visualizations, and data mining techniques across a range of sources.

The findings indicate that the semiconductor crisis has exerted a quantifiable influence on brand perception, particularly within the context of social media platforms. The quality of communication and the visibility of the supply chain appear to be major factors in maintaining a favorable image. Finally, although fluctuations in sentiment may coincide with shifts in popularity, this relationship should be interpreted with caution. The predictive potential of emotional signals is context-dependent and inconsistent across different platforms and brands.