MSc. Applied Data Science Thesis

By

Justin Kuiper

(7630220)

How do the sentiments and topics on Fijn Wonen's primary platforms differ from each other, and how can this analysis inform the optimization of their social media strategy to enhance user engagement and drive website traffic?

First supervisor: dr. Florian Kunneman
Second reader: dr. Pierre Albert
Daily supervisor: dhr. Benno Hakvoort

July 7, 2023

Abstract

This study explores the use of sentiment analysis (SA) and topic modeling (TM) to evaluate the performance of Fijn Wonen's social media posts, primarily on LinkedIn, and their impact on the website www.fijn.com. Fijn Wonen, a company dedicated to building advanced high-quality homes using robotic technology, leverages LinkedIn to engage with customers and partners. This study aims to uncover insights into user engagement by comparing sentiments and topics in Fijn Wonen's posts and the reactions they receive. The findings contribute to understanding how social media interactions influence website traffic and customer engagement for a technologically innovative company like Fijn Wonen. This research employs the Valence Aware Dictionary and Sentiment Reasoner (VADER) for SA and Latent Dirichlet Allocation (LDA) for TA to analyze multiple datasets from LinkedIn, www.nu.nl, and www.fijn.com. The datasets, which vary in structure and format, were pre-processed to ensure data quality and consistency. Key pre-processing steps included handling missing values, removing duplicates, tokenization, and lemmatization. Cohen's kappa and Krippendorff's alpha were utilized to measure inter-rater agreement on data labeling. Additionally, Pointwise Mutual Information (PMI) was applied to identify word associations within the data.

Keywords: Topic Modeling – Sentiment Analysis – Social Media Platforms

Acknowledgement

I would like to express my gratitude to my supervisor, Florian Kunneman. I am deeply indebted to him for sharing his expertise and providing sincere and valuable guidance. Florian has been an approachable mentor, who fosters stimulating and pleasant interactions.

Additionally, I extend my thanks to Fijn Wonen that I can do my thesis at their company. Especially Benno Hakvoort for being my daily supervisor and Laura Kuiper for giving me access to extract all the data necessary for this research.

Table of Contents

List of Tables

List of Figures

Acronyms

TM	Topic Modeling
SA	Sentiment Analysis
VADER	Valence Aware Dictionary and Sentiment Reasoner
MNB	Multinomial Naïve Bayes
LR	Logistic Regression
LDA	Latent Dirichlet Allocation
PMI	Pointwise mutual information
NLP	Natural Language Processing
NLTK	Natural Language Toolkit
API	Application Programming Interface

1 Introduction

1.1 Motivation and Context

In the evolution of digital marketing understanding the impact of social media on consumer behavior and business performance has become critical. More and more companies across various industries use social media platforms to engage with customers. Build their brand awareness and drive their sales to a higher level. However, the effectiveness of these efforts varies widely and is influenced by factors such as content quality, platform choice, and audience engagement. This study seeks to address the broader issue of optimizing social media strategies to maximize Fijn Wonen's business benefits.

The problem this research addresses is that Fijn Wonen wants to know what are the best ways to increase clicks, websites traffic and sales. The posts on LinkedIn gets likes but doesn't always trigger the users. Fijn wonen recently added a lobbyist to increase their visibility in the Parliament. They struggled to translate their online presence into business benefits. Therefore, understanding which actions give better response are added value for Fijn Wonen.

By analyzing data from LinkedIn, Nu.nl and Fijn Wonen's website traffic. This research will uncover strategies that lead to higher engagement rates, increased website traffic, and improved conversion rates. The focus is on understanding the relationship between topics and sentiments of Fijn Wonen's post's and the posts they receive.

The importance of this research lies in its ability to inform better practices for Fijn Wonen's digital marketing efforts. Optimizing social media strategies based on this research can increase the customer satisfaction and grow stronger community relations. By offering a detailed analysis specific to Fijn Wonen, this study aims to deliver actionable insights that directly address the company's marketing challenges. It also contribute to more meaningful and effective social media interactions.

1.2 Case Study: Fijn Wonen

To illustrate these broader issues, this study examines the social media strategies of Fijn Wonen, a company that focuses on building future homes using advanced robotic technologies. Fijn Wonen has multiple social media platforms with LinkedIn, Nu.nl and podcasts serving as their primary social media platforms that facilitate their customer engagement and partner communication. However, Fijn Wonen aims to optimize its overall social media strategy on multiple platforms to increase visibility and audience connection. By analyzing Fijn Wonen's social media performance, this study seeks to provide insights into how businesses in the home construction industry can leverage digital platforms more effectively.

Fijn Wonen faces the challenge of identifying which social media strategies best drive website traffic and engagement, ultimately leading to increased sales. They aim to acquire knowledge about the most effective content types, posting frequencies, and platform-specific tactics. This understanding is crucial for fine-tuning their digital marketing efforts to achieve better business outcomes. This case study exemplifies the application of comprehensive social media strategies within a specific business context, providing valuable insights into optimizing digital marketing efforts to improve overall business performance.

1.3 Research question

Building on the central problem, the focus of this research, and the context provided by the Fijn Wonen case, the following research question will be investigated in this thesis:

"How do the sentiments and topics on Fijn Wonen's primary platforms differ from each other, and how can this analysis inform the optimization of their social media strategy to enhance user engagement and drive website traffic?"

By utilizing VADER and the Multinomial Naïve Bayes Model, the study analyzes the sentiment of both Fijn Wonen's posts and the reactions they receive. The sentiment scores provide insights into whether the sentiments are positive, neutral, or negative. The Multinomial Naïve Bayes Model is used to label data on topic categories. To identify the differences in topics, Latent Dirichlet Allocation (LDA) is used, uncovering the main themes and topics in both the posts and the reactions. This analysis shows how the topics discussed by Fijn Wonen align or differ from the topics of interest to their audience. Cohen's Kappa ensures the reliability of manually labeled data by measuring inter-rater agreement, validating that the sentiment and topic labels used in training the models are consistent. Krippendorff's Alpha provides a robust measure of agreement across various types of data and multiple raters, further ensuring the reliability of the labeled data. Additionally, Pointwise Mutual Information was applied to identify word associations between the two datasets. Enhancements in user engagement and website traffic will be validated using data from www.fijn.com.

1.4 Literature overview

This study is carried out using multiple data sets by multiple platforms. Therefore the data structure would not be the same. The methodology employed to address this research question involves using VADER for Sentiment Analysis (SA) and Latent Dirichlet Allocation (LDA) to depict the underlying topics discussed on the platforms.

Digital marketing strategies have increasingly relied on social media platforms to engage consumers and enhance business performance (Smith, 2020; Tuten & Solomon, 2017). As businesses strive to optimize their online presence, understanding the impact of social media on consumer behavior becomes pivotal (Duffett, 2015). This literature overview explores key aspects of social media marketing, sentiment analysis, and topic modeling relevant to the study of Fijn Wonen's digital marketing strategies.

Social media platforms have transformed traditional marketing landscapes by enabling direct engagement with consumers (Mangold & Faulds, 2009). They serve as powerful tools for building brand awareness, fostering customer loyalty, and influencing purchase decisions (Hajli, 2014). The effectiveness of these platforms varies depending on factors such as content quality, platform selection, and engagement strategies (Hanna et al., 2011). Understanding these dynamics is crucial for businesses aiming to maximize their digital marketing investments (Kaplan & Haenlein, 2010).

To achieve meaningful outcomes from social media efforts, businesses must adopt strategies that resonate with their target audiences (Kumar & Mirchandani, 2012). This involves crafting diverse and compelling content, leveraging advanced technologies, and strategically deploying posts across different platforms (Leung et al., 2020). Effective social media strategies not only drive engagement but also translate into tangible business benefits such as increased website traffic and enhanced conversion rates (Kumar et al., 2016).

Sentiment analysis has emerged as a valuable tool for assessing public perceptions and attitudes towards brands on social media (Thelwall et al., 2012). Techniques like VADER provide efficient sentiment scoring by analyzing textual data for polarity (positive, negative, neutral) without the need for extensive training data (Hutto & Gilbert, 2014). This approach enables businesses to gauge audience sentiment in real-time and tailor their marketing strategies accordingly (Gandomi & Haider, 2015).

Topic modeling, particularly using methods like Latent Dirichlet Allocation (LDA), allows businesses to uncover latent themes within large sets of textual data (Blei et al., 2003). By identifying prevalent topics in consumer discussions, businesses can refine their content strategies and address topics of interest to their audience (Chang et al., 2009). This approach supports data-driven decision-making in digital marketing, enhancing content relevance and engagement (Wang et al., 2016).

The literature reviewed underscores the relevance of optimizing social media strategies to achieve business objectives such as enhanced engagement and increased website traffic (Kumar et al., 2016). By applying sentiment analysis and topic modeling techniques, this study on Fijn Wonen aims to uncover insights into audience preferences and content effectiveness (Chang et al., 2009; Hutto & Gilbert, 2014). These insights will inform strategies to improve user engagement, drive traffic to Fijn Wonen's website and support its business growth.

2 Data

To address the research question, multiple datasets were obtained through multiple platforms. Each with a own motivation. The motivation for the first dataset is that Fijn Wonen's posts on a professional social media site will be studied. It is through this analysis that we can tell which kinds of posts generate the most interest and how different metrics such as views, clicks, and likes are related to user interaction.

The second dataset is about Fijn Wonen's video post on Nu.nl. It was reacted upon publicly in this dataset. In this way, one can get a bigger picture regarding how good people feel about Fijn Wonen's innovative housing solutions beyond LinkedIn by just analyzing those comments.

The third dataset is information shows the correlation between social media and results since it looks at website traffic and users' actions within FijnWonen.nl. It helps to check if the efficiency of using social media platforms to enhance customer engagement levels within a company site.

2.1 Data description

The first dataset was collected from LinkedIn, where data was gathered using LinkedIn's API and web scraping techniques to capture all relevant posts from Fijn Wonen. The raw data was then formatted into CSV files for analysis. This dataset consists of a total of 112 posts with an average word count of 98.5, and it includes the following key data fields that will be used for the research:

- Titel bijdrage: Textual content of the post
- Aangemaakt: Date of the post
- Weergaven: Shows the views
- Klikken: Unique clicks on the content, company name, or logo by a registered user
- Doorklikfrequentie (CTR): Click-through Rate are the clicks divided by total views
- Interessant: How many people liked the posts
- Commentaren: How many comments
- Interactiepercentage: Number of interactions plus clicks and followers divided by the number of views
- Interactiepercentage: Shows the number of interactions plus the number of clicks and followers, divided by the number of views

The second dataset covers all comments on a video post on www.nu.nl related to Fijn Wonen's innovative housing solutions. This video message is included in the research to gauge public response and engagement with Fijn Wonen's initiatives in advanced robotic technologies for home construction. Each comment in the dataset includes a count of respect, reflecting the level of engagement from users. With 453 records with an average word count of 18.5, this dataset provides valuable insights into how the public perceives and interacts with Fijn Wonen's content on a prominent news platform like www.nu.nl.

- Bericht: Textual content of the comment
- Respect: Number of users who respected the comment

The third dataset focuses on the traffic on the website www.Fijn.com. Website traffic refers to the number of visitors who access a website and the pages they view. Understanding website traffic helps assess the effectiveness of a website in reaching its audience and achieving its goals. This dataset is included in the research to determine the impact of Fijn Wonen's social media strategies on their website engagement and conversion rates. By analyzing this data, which requires multiple CSV files, the study aims to link social media activity to tangible business outcomes, providing a comprehensive view of how digital marketing efforts translate into website performance and ultimately drive sales.

- Bezoekers: Are the users on the website
- Sessies: Are the actions of a user
- Datum: Is the date

A session is the collective term for the actions a user performs on your site. One user can open multiple sessions. These sessions can take place on the same day or spread over several days, weeks or months. Once a session is ended, a new session can be started.

Value	Date	Visitors
Max	20-06-2024	901
Min	31-01-2024	1
Mean	10-04-2024	235.5

Table 1: Describe figure 1

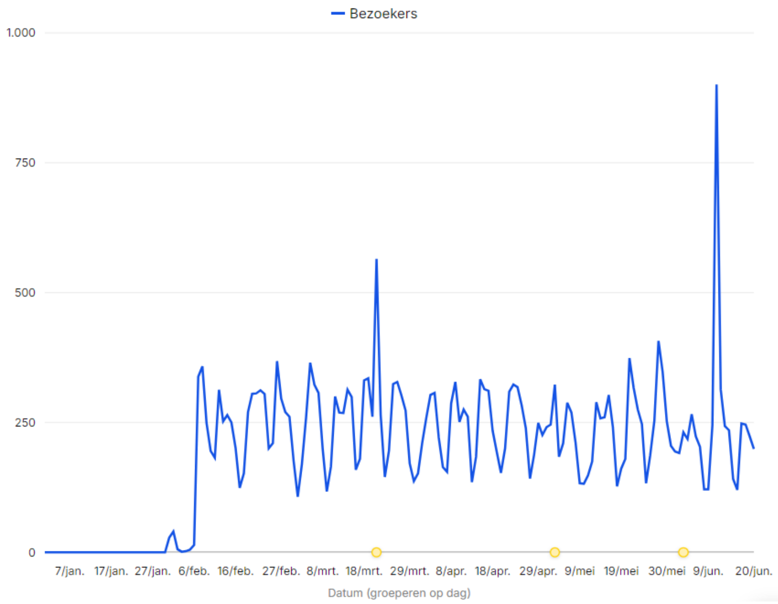

Figure 1: Visitors www.fijn.com

2.2 Data pre-processing

In this study, datasets were sourced from multiple online platforms. Access was gained through the marketing team of Fijn Wonen. To effectively address this noise, an essential first step involved gaining familiarity with the dataset. This was achieved through the execution of a frequency analysis, offering valuable insights into the dataset's composition. The goal is to ensure data quality, consistency, and suitability for the subsequent stages of analysis and modeling. See the table below:

Step	**Discription**	**Result**
Total Posts Collected	Raw posts before cleaning	112
Duplicate Posts Removed	Identified and removed duplicate posts	8
Cleaned Posts	Posts remaining after cleaning	104
Tokenization	Splitting posts into individual words	11302 tokens
Stop Words Removal	Removing common words (e.g., "the", "is")	101 stop words removed
Lemmatization	Converting words to their base form (e.g., "running" to "run")	-
Final Dataset	Prepared dataset for analysis	104

Table 2: LinkedIn pre-processing steps

8 duplicates were found. The reason for these duplicates was the repost on LinkedIn. The tokenization made 104 cleaned posts, which were split into 11,302 individual words (tokens). This prepared the text for further analysis. A list of 101 common words (stop words) were removed to focus on the more meaningful content of the posts. The lemmatization step, while mentioned, did not produce a separate count but would have standardized words to their base forms to improve analysis accuracy.

For the Nu.nl dataset we scraped the data straight from Nu.nl. Documented all the reactions on the video in a excel file. See the table below for more info:

Step	Discription	Result
Total Posts Collected	Raw posts before cleaning	453
Duplicate Posts Removed	Identified and removed duplicate posts	0
Cleaned Posts	Posts remaining after cleaning	453
Tokenization	Splitting posts into individual words	16598 tokens
Stop Words Removal	Removing common words (e.g., "the", "is")	101 stop words removed
Lemmatization	Converting words to their base form (e.g., "running" to "run")	-
Final Dataset	Prepared dataset for analysis	453

Table 3: Nu.nl pre-processing steps

There were no duplicates in this dataset. Therefor the dataset has 453 reactions from users with 16598 tokens after tokenization. The same list of 101 common words (stop words) were removed to focus on the more meaningful content. This conclude that all the 453 nu.nl posts are used.

The data pre-processing for this study involved several meticulous steps to ensure data quality, consistency, and suitability for subsequent analysis and modeling. Initially, raw posts were collected from various social media platforms, including LinkedIn and Nu.nl related to Fijn Wonen. To maintain the uniqueness of each entry, duplicate posts were identified and removed, resulting in a dataset of unique posts. Tokenization was then performed using Python's Natural Language Toolkit (NLTK) to split the text into individual words. Following this, stop words, such as "the" and "is," were removed using a predefined list from NLTK, eliminating 101 common words that do not contribute significant meaning to the analysis. Lemmatization was applied to convert words to their base forms, enhancing the standardization and quality of the dataset. The prepared dataset, consisting of 104 cleaned posts, was then ready for sentiment analysis using the VADER (Valence Aware Dictionary and sEntiment Reasoner) tool and topic modeling through Latent Dirichlet Allocation (LDA). The sentiment analysis process involved calculating polarity scores for the lemmatized tokens, while topic modeling was conducted using the Gensim library to identify underlying themes within the text. This comprehensive pre-processing ensured that the data was of high quality and ready for insightful analysis.

3 Methods

This chapter will provide a detailed description of the used methods. The methods are divided into two sections: section 3.1 will address the the SA techniques and section 3.2 the used TM models. The first section 3.1 will explain SA with the models VADER and Multinomial Naïve Bayes Model. In the second section 3.2 TM are explained with the models LDA and PMI.

3.1 Sentiment Analysis

3.1.1 VADER

Given that the data are unlabeled, the chosen tool for sentiment analysis (SA) is the Valence Aware Dictionary and Sentiment Reasoner (VADER) (Hutto et al., 2014). This lexicon and rule-based SA tool, specifically designed for social media text data, has been demonstrated to be effective across platforms such as Twitter, Facebook, and other social networks (Abiola et al., 2023). VADER can handle a wide range of linguistic elements, such as words, emojis, slang, and abbreviations, that are frequently used in social media (Hutto, 2014). Furthermore, it serves as a rapid solution for SA without needing training data, making it faster than traditional Machine Learning algorithms.

VADER employs a compound score calculation to assess sentiments, categorizing text into negative, neutral, or positive sentiment categories. The compound score serves as a likelihood measure, ranging from -1 (indicative of a strong negative sentiment) to +1 (representing a strong positive sentiment), with 0 denoting neutrality (Abiola et al., 2014). To determine the compound score, VADER conducts a thorough examination of the text to identify known sentimental features. It then proceeds to modify the intensity and polarity of these features based on predefined rules. The scores of the identified features are added together, and the final score undergoes normalization to fall within the range of (-1, 1). Sentiment labels are subsequently assigned based on the calculated compound scores. Scores greater than 0.5 are labeled as positive, those less than -0.5 as negative, and scores between -0.5 and 0.5 as neutral.

3.1.2 Multinomial Naïve Bayes Model

For this model, we had multiple people label our data. We then used this data to train the sentiment/topic model in Python. The Natural Language Toolkit (NLTK) and scikit-learn libraries are used to perform this task. The Multinomial Naive Bayes (MNB) model is a probabilistic learning method commonly used for text classification and natural language processing tasks, such as spam detection and sentiment analysis. It belongs to the Naive Bayes family of algorithms, which operate on Bayes' theorem with a strong (naive) assumption of independence among features.

For all models, accuracy is calculated; this demonstrates how effectively the model performs. Literature specifically citing a target accuracy of 80% for sentiment analysis models may not be prevalent as accuracy thresholds can vary based on specific datasets, tasks, and research objectives (B. Pang et al, 2007). For this research, an accuracy of 70% or higher is required. If the accuracy falls below 70%, more data will need to be labeled to enhance the model's accuracy.

3.2 Topic Modeling

Topic modeling is an unsupervised technique used to identify the predominant themes within a vast collection of unstructured documents (Blei, 2012). This method theorizes that documents comprise mixtures of various topics, and it defines topics as distributions of words found within the documents. This approach provides a more profound understanding of the underlying thematic structure.

3.2.1 Latent Dirichlet Allocation (LDA)

To extract topics from comments and posts, Latent Dirichlet Allocation (LDA) has been chosen as the method for this task. LDA is a commonly used model for uncovering topics from a set of texts (Blei et al., 2003).

Unsupervised Learning Capability: LDA is particularly suitable for unsupervised learning tasks where there is no predefined set of topics. It automatically identifies underlying topics by modeling documents as mixtures of topics, each characterized by a distribution over words. This characteristic makes it versatile for analyzing diverse datasets without the need for labeled data (Blei et al., 2003).

Flexibility and Adaptability: Unlike some topic modeling techniques that require predefined topic structures, LDA adapts well to various types of textual content. It does not assume prior knowledge of the number or nature of topics present in the dataset, making it suitable for exploratory analyses and hypothesis generation (Blei et al., 2003).

Interpretability of Results: LDA outputs are highly interpretable. It assigns a probability distribution of words to each topic and a distribution of topics to each document, allowing researchers to discern meaningful themes or topics within the text corpus (Blei et al., 2003).

Scalability: LDA is scalable to large datasets, making it feasible to analyze extensive collections of documents commonly found in social media platforms, news articles, and academic literature. Its efficiency in handling large volumes of text data ensures robust performance in digital marketing and social media analytics (Blei et al., 2003).

LDA posits that each document is generated from a mixture of topics, each composed of a series of words. The objective of this model is to identify the latent topic present in the text. By assigning each document a set of topics and each topic a set of keywords, LDA can find potential topics in the text. Furthermore, this was also chosen since it is suitable for working with large text datasets and automatically discovers potential topics in text collections without having to define the number or structure of topics in advance (Griffiths et al., 2004). Furthermore, it is also an unsupervised algorithm, which is particularly useful for analyzing large amounts of unlabeled social media data.

In the process of using the LDA model, we performed additional pre-processing steps on the LinkedIn and nu.nl that went pre-processing as explained in section 2.2. These extra steps included using the 'remove_emoji' function to remove emojis. Subsequently, the data has also been vectorized. This was done by first, creating a dictionary from the pre-processed text that contains all the unique words. Next, each document is transformed into a vector represented by a bag-of-words model. Each of these documents is represented by the terms it contains and how often they occur. This vectorization method is designed to convert text data into a format that LDA models can process efficiently so that the model can identify and analyze the topic distribution in the document. Through this transformation, can effectively apply the LDA model for TA.

3.2.2 Pointwise mutual information (PMI)

Pointwise Mutual Information (PMI) is a measure from statistics and information theory, used to quantify the association between two events or variables. It assesses how much more (or less) likely two events are to occur together than if they were statistically independent (Levy & Goldberg, 2014).

Pointwise Mutual Information is a powerful measure for quantifying the association between two events, particularly in the field of natural language processing. It helps identify significant word pairs and assess semantic similarity, providing valuable insights into the structure and meaning of text (Bouma, 2009). It will be used to check for relationships between the words in the LinkedIn and Nu.nl datasets.

4 Experimental Set-up

This section outlines the procedure for applying and evaluating the methods to achieve the research objectives. The manual labeling of the data are described in the first section. Next, the experimentation and evaluation procedures are described for each experiment.

4.1 Experiment 1: Manual labeling of data

The first experiment involved the manual labeling of data to create a benchmark for analyzing the topics and sentiments in Fijn Wonen's social media posts. The LinkedIn and the Nu.nl both are labeled at there own way. Manual labeling of the data ensures high-quality, accurate annotations that will help by the training of our machine learning model.

4.1.1 Data annotations

The sentiment was categorized into 'Negative', 'Neutral', 'Positive'. These three categories provide a clear and straightforward way to get a overall sentiment without introducing excessive complexity. In addition to the three-valued stance classes we included separate classes grouped under 'Topic Categories', 'Intent Categories', 'Info Categories'. With these additional categorizations, we aimed to achieve a precise understanding of all potentially relevant characteristics of the posts in relation to their topics and sentiments. This comprehensive labeling is intended to enhance the performance of machine learning models in analyzing social media content.

The **Topic** Categories were divided into 'Call to action', 'Post' or 'Other'. The Topic Categories were only used in the LinkedIn dataset because it helps streamline content creation. Making it easier to identify the purpose and format of the content.

The **Intent** Categories were divided into 'Answer', 'Question', 'Suggestion', 'Complaint', 'Praise' and 'Other'. The Intent Categories were only used for the Nu.nl dataset. This is a effective way to classify interactions based on the underlying intent behind each communication. This classification helps to understand the user's purpose.

Finally the **Info** Categories were divided into 'Discussion' , Affordable housing' or 'Company'. The Intent Categories were only used for the Nu.nl dataset to classify the underlying intent behind the user interactions or content. We expect that a lot of reaction posts were about a discussion of affordable housing. We expect that because when we scraped the data we went through all the posts.

Category type	Category	Definition	Example Post (in Dutch)
Sentiment	Positive	Expresses approval or satisfaction.	"Hier gaan huizen weg voor 6000 per m2. Als warme broodjes. Ik vond en vind dat nog steeds prima."
	Neutral	Factual or impartial.	"Wat bedoel je Hiermee"
	Negative	Expresses disapproval or dissatisfaction.	"de prijzen staan gewoon niet meer in verhouding wat er geleverd wordt. 4000E/m2 voor een appartement in een opgeknapt kantoor vind ik gewoon ver boven wat redelijk is"
Topic Categories	Call to action	Urges the audience to take action.	"Duik dieper in de toekomst van bouwen en download jouw informatiepakket om meer te weten te komen over onze woningen 👉 https://bit.ly/3UyDMaL"
	Post	General content or updates.	"Vanmorgen sprak Jan Paternotte, Tweede Kamerlid bij D66, in Goedemorgen Nederland van Omroep WNL over de woningfabriek en de woningen die we hier dagelijks produceren. ''We hebben te weinig bouwvakkers om voldoende huizen te kunnen bouwen, maar in zo'n fabriek werken ze met robottechniek. Daardoor kan er wél snel gebouwd worden, op termijn zelfs 4.000 woningen per jaar in één fabriekshal'' aldus Paternotte.

	Other	Content not fitting other categories.	"Door deze nieuwe manier van bouwen, zorgen we voor een snellere realisatie met minder mensen en minder afval."
Intent Categories	Answer	Provides information or solutions.	"Als het filmpje kijkt en goed luistert, dan hoor je dat dat incl installatie, grond etc is."
	Question	Asks for information or clarification.	"heb je het filmpje wel bekeken?"
	Suggestion	Proposes improvements or changes.	"Is zo'n fabriek niet in een megastal op te zetten? Goed voor de werkgelegenheid."
	Complaint	Expresses dissatisfaction or reports issues.	"Totaal onmogelijk en marketing praat. Behalve als die €175.000,00 een totale bruto prijs."
	Praise	Commends or appreciates.	"als je tien procent grondstoffen bespaard, is ieder 11e huis wat dat betreft gratis in vergelijking met reguliere bouw. bij een gewoon bouwproject sneuvelt er nogal eens een steen, cement of kuub beton teveel, tikt nog best aan op een heel huis."
	Other	Content not fitting other categories.	"de prijzen staan gewoon niet meer in verhouding wat er geleverd wordt. 4000E/m2 voor een appartement in een opgeknapt kantoor vind ik gewoon ver boven wat redelijk is"
Info Categories	Discussion	Exchanges of ideas or opinions.	"Kwestie van de video kijken, daarin wordt dat letterlijk gezegd."
	Affordable housing	Information about affordable housing.	"Ten eerste is 175.000 euro voor een huis (alleen het huis) niet heel goedkoop. Ten tweede moet je dan nog grond erbij kopen. En laat dat nou hartstikke duur zijn."
	Company	Details related to the company "Fijn Wonen/Van Wijnen".	"Leuk dat goedkoper bouwen... JA voor Van Wijnen NIET voor de koper!!"

Table 4: Specification of the annotation categories

4.1.2 Intercoder Reliability

We calculated inter-annotator agreement by Cohen's Kappa and Krippendorff's Alpha, which accounts for different annotators and empty values. Cohen's kappa evaluates the agreement between two raters beyond what would be expected by chance. The value of kappa ranges from -1 to 1:

κ = 1: Indicates perfect agreement
κ = 0: Agreement equal to what would be expected by chance alone.
κ < 0: Agreement is worse than chance.

Krippendorff's Alpha to be 0.80 or higher indicates a strong agreement. Values between 0.667 and 0.80 can be acceptable. However, values lower than 0.667 suggest that the agreement among raters is not sufficient.
More information about the Cohen's kappa and the Krippendorff's alpha is in Appendix E. See the results in the table below:

Intercoder Reliability					
	LinkedIn dataset		**Nu.nl dataset**		
Model	Sentiment	Topic Categories	Sentiment	Intent Categories	Info Categories
Cohen's kappa	0.93	0.75	0.87	0.41	0.80
Krippendorff's alpha	0.89	0.73	0.82	0.34	0.77

Table 5: Intercoder Reliability

4.1.3 Evaluation

The intercoder reliability analysis reveals strong agreement in sentiment analysis across both datasets, suggesting that the sentiment classification process is highly reliable. Topic Categories and Info Categories are moderated reliable they both reflect good but suggest some variability in categorization. The Intent Categories in the Nu.nl dataset show low agreement. Both Cohen's kappa and Krippendorff's alpha are relatively low. This suggests that intent categorization needs improvement in either the definition of categories or the training of raters.

Overall the analysis suggests that while most categories are reliable, there is a need for improvement in certain areas. Especially in the intent categorization that will probably come from the more categories there are.

4.2 Experiment 2: Sentiment Analysis

The second experiment will compare the sentiment analysis (SA) performance of a rule-based model and a probabilistic model across three categories: Positive, Neutral, and Negative. All systems will undergo evaluation on their classification abilities to determine the effectiveness of each classifier model. To develop and assess a machine learning sentiment classifier, high-quality annotated data is essential, as detailed in Section 4.1.1.

4.2.1 Set-up VADER, MNB and LR

The categories for annotation are listed in the tables below:

Category			
Sentiment	Positive	Neutral	Negative
Topic Categories	Call to action	Other	Post

Table 6: LinkedIn categories annotations

Category						
Sentiment	Positive	Neutral	Negative			
Info Categories	Affordable housing	Discussion	Company			
Intent Categories	Other	Question	Praise	Answer	Complaint	Suggestion

Table 7:Nu.nl categories annotations

20-25% of the data was labeled by all annotators for both datasets. The annotators got a manual to fulfill this task. There was no option to discuss their answers with other annotators. When the annotator was uncertain, the label was classified as either neutral or other.

4.2.2 Evaluation

To evaluate the performance of the sentiment classifiers, precision, recall, F1 scores, and support are calculated. By implementing these baseline models, you can get a quick sense of how well models perform on your dataset. Each model (VADER, MNB, LR) is evaluated on the test dataset to assess its performance in sentiment classification. Accuracy, precision, recall, and F1 scores for each sentiment class (positive, neutral, negative) are computed. The performance of the rule-based VADER model is compared with that of the probabilistic models (MNB and LR) to identify their strengths and weaknesses.

4.3 Experiment 3: Topic Modeling

The third experiment aims to uncover the main themes in the LinkedIn and Nu.nl datasets using Latent Dirichlet Allocation (LDA) and to investigate the relationship between the words in these datasets using Pointwise Mutual Information (PMI).

4.3.1 Set-up LDA and PMI

Use the preprocessed text for LDA and also calculate the PMI. Compute PMI scores for word pairs within each dataset. Calculate PMI between words in the LinkedIn dataset and words in the Nu.nl dataset to identify cross-dataset relationships. Identify high PMI word pairs to understand significant associations. Analyze how topics in LinkedIn relate to topics in Nu.nl by examining common high PMI word pairs.

4.3.2 Evaluation

To evaluate the results of the topic modeling and PMI analysis, we will take the following steps:

1. Calculate the topic coherence scores for the topics extracted by the LDA to ensure they are interpretable and meaningful.
2. Use coherence scores to compare and validate the quality of the topics generated for both datasets.
3. Evaluate the significance of the PMI scores by comparing them with a baseline, such as randomly shuffled word pairs.
4. Visualize high PMI word pairs to highlight the relationships between the LinkedIn and Nu.nl datasets.
5. Perform a manual inspection of a sample of high PMI word pairs to validate their relevance.

5 Results

In this section, the results of the three experiments are presented successively:
Manual labeling of the data in section 5.1, SA in section 5.2, TM in section 5.3 and 5.4 data from www.fijn.com.

5.1 Manual labeling of the data

5.1.1 Example data with Cohen's Kappa

Index	Bericht	Rater 1	Rater 2	Rater 3
0	Ik heb een keer een poging gedaan om zoiets te...	Neutraal	Negatief	Neutraal
1	Al eeuwen geldt het adagium: de plek, de plek ...	Positief	Positief	Positief
2	Inclusief grond? Ja in Groningen misschien. Ni...	Neutraal	Neutraal	Neutraal
3	Huis 175k, grond een veelvoud daarvan, en zo k...	Neutraal	Neutraal	Neutraal
4	Meer precies: oost Groningen.. Het is me ook n...	Positief	Positief	Positief

Table 8: Annotations from raters

Cohen's kappa between rater1 and rater2: 0.93
Cohen's kappa between rater1 and rater3: 0.93
Cohen's kappa between rater2 and rater3: 0.86

5.2 Sentiment Analysis

Results accuracy on sentiment analysis with VADER, MNB and LR are listed below.

Model	Name	Accuracy
Rule Based	VADER	75%
Probabilistic classifier	MNB	85%
Baseline, statistical	LR	54%

Table 9: Accuracy SA models

These models represent different approaches to sentiment analysis. Rule based model VADER and Probabilistic classifier MNB. Scores higher on the accuracy then LR, therefore we will use these two models on the datasets.

Sentiment	VADER LinkedIn	Multinomial Naïve Bayes LinkedIn	VADER Nu.nl	Multinomial Naïve Bayes Nu.nl
Positive	19	81	32	7
Neutral	88	31	414	442
Negative	5	0	32	4

Table 10: LinkedIn & Nu.nl sentiment by the models VADER & MNB

5.3 Topic Modeling

This section presents the outcomes of TA on posts that are placed by Fijn Wonen marketing. Also the posts that are commented on the www.nu.nl. The tables below show the most frequent words count.

Word	Amount
We	171
Woningen	130
Onze	94
Fijn	93
Wonen	82
Nieuwe	60
Bouwen	51
Samen	51
Project	40
Betaalbaar	39

Table 11: Most frequent words LinkedIn

Word	Amount
Grond	144
Wel	115
Huis	111
Prijs	58
Kosten	57
Bouwen	53
Alleen	53
Waar	52
Woning	48
We	48

Table 12: Most frequent words nu.nl

The two tables have the words: ‘We’ and ‘Bouwen’ in common. These are highlighted in the table with the same color. Fijn Wonen talks a lot about ‘Betaalbaar’ this responds to ‘Prijs’ and ‘Kosten’, affordable housing is a hot topic.

5.3.1 PMI

In Group 1 is all the data of Nu.nl and Group 2 is all the data of LinkedIn. Let's take a look at what these top 10 lists look like if we rank terms by PMI instead:

```
Top 10 Distinctive Terms in Group 1 (based on PMI):
Term: 'huis' - Total PMI Score: 1023.00
Term: 'wel' - Total PMI Score: 1011.52
Term: 'grond' - Total PMI Score: 954.81
Term: 'kosten' - Total PMI Score: 751.19
Term: '000' - Total PMI Score: 721.02
Term: 'huizen' - Total PMI Score: 665.33
Term: 'prijs' - Total PMI Score: 659.81
Term: 'alleen' - Total PMI Score: 630.67
Term: 'bouwen' - Total PMI Score: 627.09
Term: 'we' - Total PMI Score: 625.86

Top 10 Distinctive Terms in Group 2 (based on PMI):
Term: 'we' - Total PMI Score: 1196.78
Term: 'woningen' - Total PMI Score: 915.53
Term: 'onze' - Total PMI Score: 867.39
Term: 'fijnwonen' - Total PMI Score: 836.14
Term: 'wonen' - Total PMI Score: 708.98
Term: 'fijn' - Total PMI Score: 672.82
Term: 'samen' - Total PMI Score: 644.06
Term: 'project' - Total PMI Score: 585.58
Term: 'vanwijnen' - Total PMI Score: 558.38
Term: 'én' - Total PMI Score: 546.87
```

Key observations are labeled in the table below:

Observation	Group	Conclusion
Distinctive Context	Group 1	Focuses heavily on terms related to housing, costs, and building. There's a significant presence of terms related to financial aspects and property features.
	Group 2	Highlights a more communal or positive context with terms related to living quality, community, and projects. Terms suggest a focus on collective experiences and aspirations.
Common Terms	Group 1 & 2	The term 'we' appears in both groups but with different PMI scores, showing that it's relevant but used distinctively in each context.
Term Usage	Group 1	Shows a more technical or financial perspective related to housing.
	Group 2	Emphasizes quality of life, community, and collaborative projects.
Unique Terms	Group 1	Includes specific terms like 'grond', 'kosten', and '000' that point to property and financial discussions.
	Group 2	Features terms like 'fijnwonen' and 'samen', indicating a focus on positive living conditions and community aspects.

Table 13: Observation PMI

5.3.2 LDA

To effectively interpret the LDA topic modeling results we determined the number of topics using the elbow method and the coherence scores. We used them to test various values and select the best one based on the model performance. See figures below:

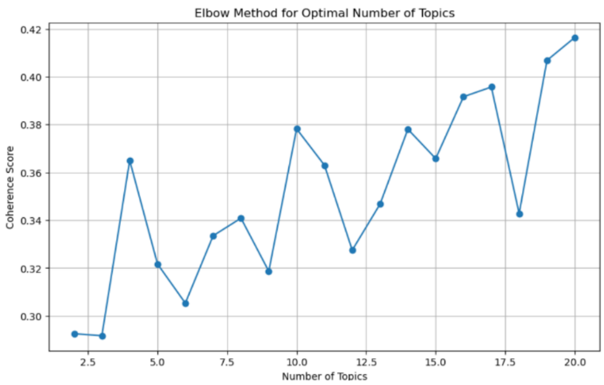

Figure 3: LDA Elbow methode LinkedIn

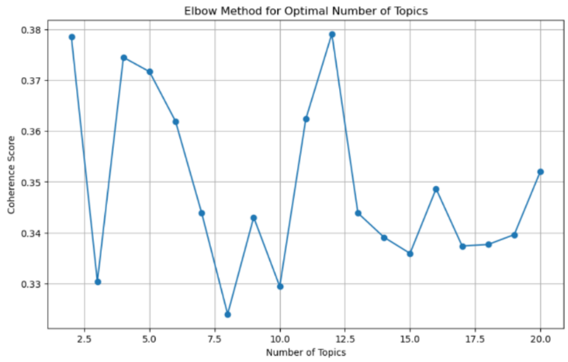

Figure 2: LDA Elbow method nu.nl

Therefore the best number of topics for LinkedIn data is 20 topics. We capped it on 20 topics because it would only rise more as seen in the figure in Appendix F. To compare the 2 datasets we use 10 topics from LinkedIn dataset.

In this figure you can see that the best number of topics for nu.nl are 2 and 12. 2 Topics suggest a very broad and high-level classification. The topics would be very general and capturing only the most overarching themes in the dataset. 12 topics will obtain a more detailed breakdown of the content, which might capture more nuanced topics. The 12 topics from nu.nl and 10 LinkedIn topics both have a coherence score of 0.379.

5.3.2.1 LDA LinkedIn

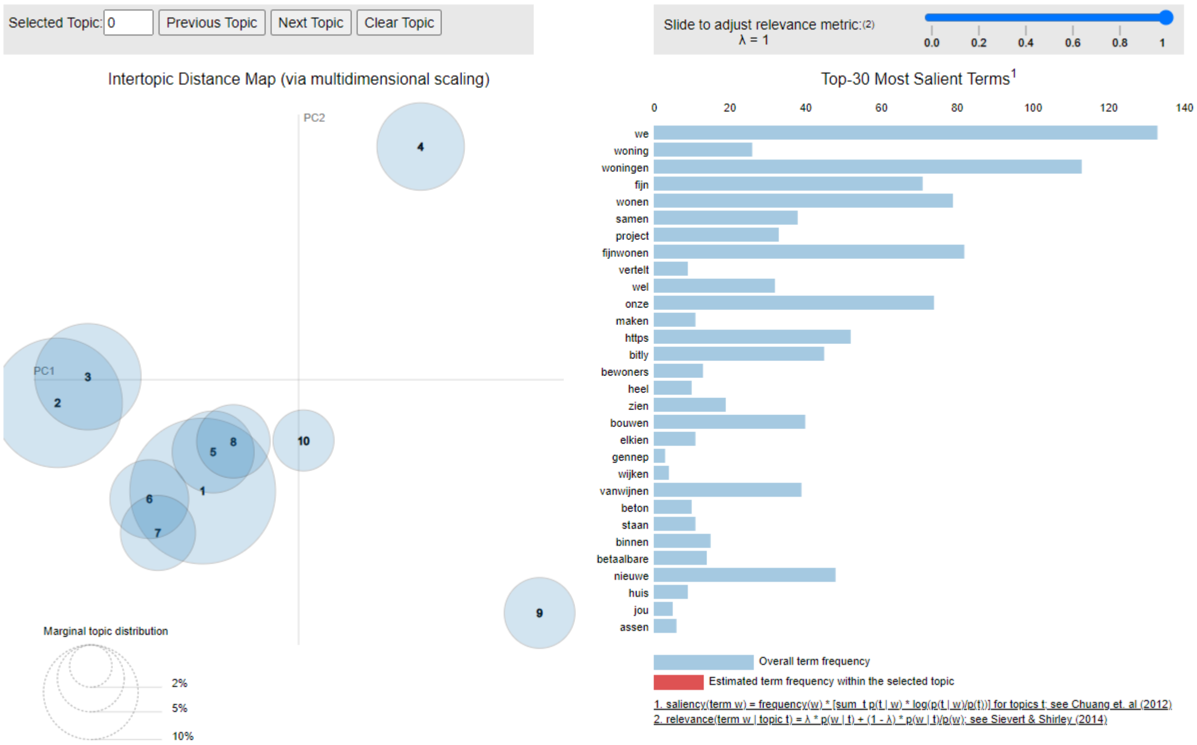

Figure 4: LDA visualization using pyLDAvis LinkedIn

The coherence score you provided for the LinkedIn dataset is 0.379. This coherence score indicates the interpretability and coherence of the topics generated by the LDA model on the LinkedIn dataset. Here's how to interpret this score:

Interpretation: Coherence scores typically range between 0 and 1, where higher scores indicate more coherent and interpretable topics. Score (0.379) suggests moderate coherence. It indicates that the topics extracted from the LinkedIn dataset have some level of interpretability, but there may be room for improvement.

5.3.2.2 LDA Nu.nl

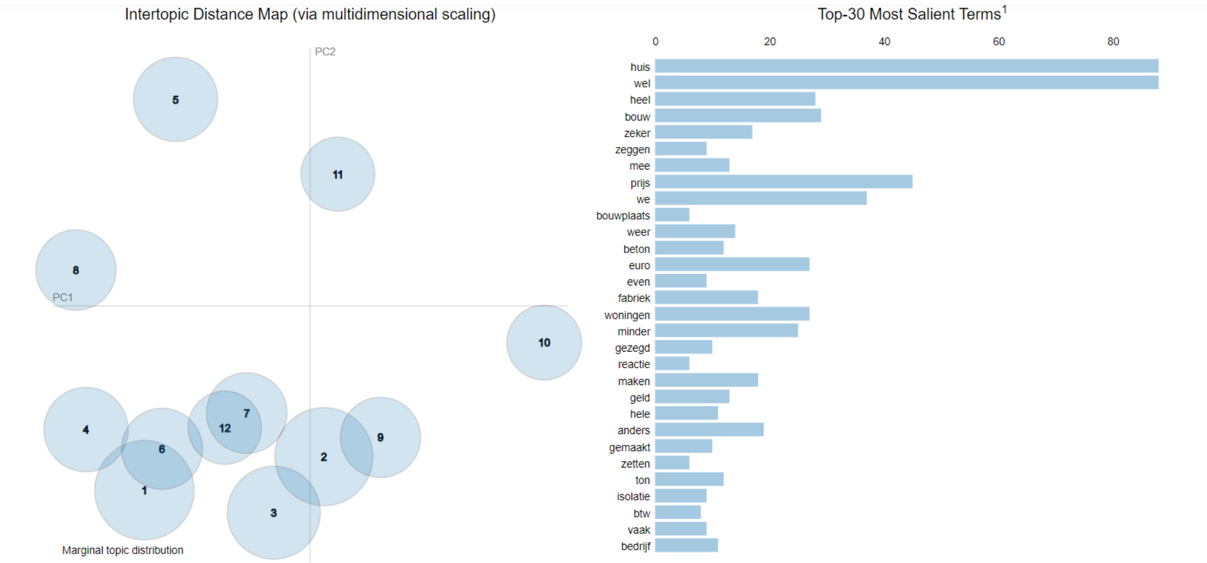

Figure 5: LDA visualization using pyLDAvis nu.nl

The coherence score for the Nu.nl dataset that you provided is 0.379. This score suggests a moderate level of coherence. It indicates that the topics extracted from the Nu.nl dataset have some level of interpretability, but there may still be room for improvement.

Both topics include terms related to housing and construction, indicating that both datasets deal with these subjects. Terms like bouw, beton, woningen, and fabriek appear in both datasets. LinkedIn has terms related to technological and innovative aspects of construction (robots, architectuur, woningfabriek). This implies a focus on modern construction methods and innovations.

5.4 Website Fijn.com

In figure 7 the visitors from channels are plotted, special dates are shown in table 10.

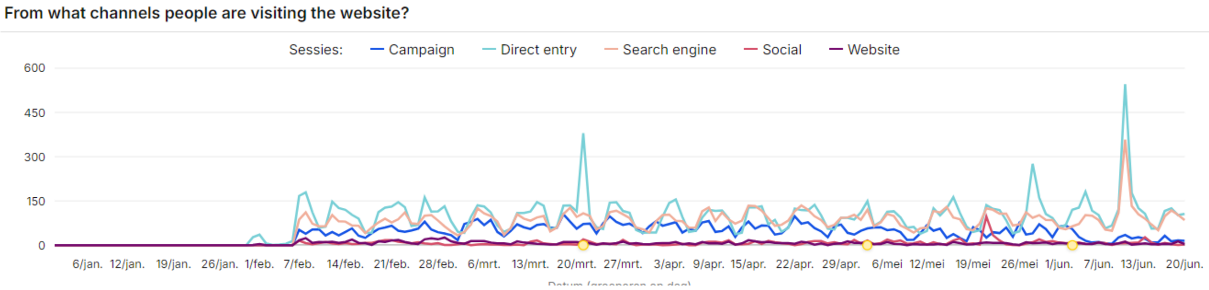

Figure 6: From what channels do visitors come from

Action	Datum	Visitors
Message Paternotte gives high-value	21-03-2024	565
Podcast/radio interview Peter Hutten	03-05-2024	323
Podcast/radio interview Peter Hutten	03-06-2024	231
PROVADA real estate fair, RAI Amsterdam	11-06-2024	901

Table 14: Actions on the website

Post LinkedIn:
“Woningen uit een fabriek, waaien die niet om? - Nee, juist niet! Heb jij beschikbare bouwgrond en wil je daarop een woonwijk plaatsen? Kom langs in onze woningfabriek en ontdek zelf de kwaliteit, variatie én betaalbaarheid van industrieel bouwen. Meld je vandaag nog aan voor een fabriekstour, check de link in de comments. 👇

Vanmorgen sprak Jan Paternotte, Tweede Kamerlid bij D66, in Goedemorgen Nederland van Omroep WNL over de woningfabriek en de woningen die we hier dagelijks produceren. ''We hebben te weinig bouwvakkers om voldoende huizen te kunnen bouwen, maar in zo'n fabriek werken ze met robottechniek. Daardoor kan er wél snel gebouwd worden, op termijn zelfs 4.000 woningen per jaar in één fabriekshal'' aldus Paternotte.

#FijnWonen #VanWijnen #hetkanwel #woningfabriek #industrieelbouwen #conceptueelbouwen?”

You can see that the website traffic was connected to the LinkedIn post from 21-03-2024. LinkedIn showed this data:

Views (total)	Uniek Views	Clicks (total)	Reacties
9615	5430	873	215

Table 15: LinkedIn Table Jan Paternotte

This was labeled as a positive post but also a call to action post. Therefor the Sentiment Positive, Topic call to action are generating more clicks and views then a neutral/negative post or a Other post.

Based on the analysis of the visitor data for Fijn Wonen's website, it is evident that specific actions and events significantly impact the daily traffic. The average number of visitors per day is 235.5. However, certain high-profile activities lead to substantial spikes in visitors.

For instance, on March 21, 2024, when Jan Paternotte highlighted the value of Fijn Wonen, the website attracted 565 visitors, more than double the average daily traffic. Similarly, podcast and radio interviews with Peter Hutten on May 3 and June 3, 2024, resulted in 323 and 231 visitors respectively, showing a notable increase compared to the daily average, particularly for the May interview.

The most significant surge in traffic occurred on June 11, 2024, during the PROVADA real estate fair in RAI Amsterdam, with 901 visitors. This event brought nearly four times the average daily visitors, underscoring the influence of major industry events on website traffic.

These findings indicate that strategic actions, such as media appearances and participation in industry fairs. Play a crucial role in driving engagement and attracting potential customers to Fijn Wonen's website. It is recommended that Fijn Wonen continues to leverage these opportunities to enhance visibility and engagement. Additionally, continuous monitoring of such events' impact on website traffic can provide valuable insights for optimizing future marketing strategies.

For more data about www.fijn.com go to appendix D.

5.5 Additional Analysis

"Vorige week vrijdag was de eer aan Raquel Garcia Hermida-van der Walle, Jan Paternotte en Danny van der Weijde - Hoogstad van D66 om onze woningfabriek, waarin we een nieuwe generatie woningen maken, te bezoeken. Op deze manier beleefden zij industrieel bouwen met eigen ogen en konden zij ervaren hoe kwaliteitswoningen in vele variatie die ook nog eens betaalbaar zijn anno nu wél mogelijk zijn. 👀

Nogmaals bedankt voor jullie komst naar Heerenveen, 'fijn' dat jullie er waren! 🏠

hashtag#FijnWonen hashtag#VanWijnen hashtag#woningfabriek hashtag#industrieelbouwen hashtag#fabriekstour hashtag#kijkjeachterdeschermen"

This post was right after the Podcast from Peter Hutten on 3 june 2024. D66 came to visit Heerenveen. As you can see in the graph the visitors were above average 150 views on the website. This was labeled as a Neutral and normal post.

Date	Views website Fijn.com
3-6-2024	231
4-6-2024	218
5-6-2024	266

Table 16: Website view June

6 Conclusion and Discussion

This study aimed to explore the differences between the sentiments and topics on Fijn Wonen's primary platforms differ from each other, and how can this analysis inform the optimization of their social media strategy to enhance user engagement and drive website traffic.

The sentiment analysis conducted using VADER and the Multinomial Naïve Bayes (MNB) model revealed significant differences in the sentiments expressed by Fijn Wonen and those expressed by users in their reactions. While Fijn Wonen's posts primarily convey positive sentiments, user reactions exhibit a broader range of sentiments, including a substantial number of neutral and negative responses. This disparity indicates a potential gap between the company's messaging and the audience's reception.

Topic modeling using Latent Dirichlet Allocation (LDA) identified distinct themes in Fijn Wonen's posts and user comments. Fijn Wonen frequently emphasizes topics such as affordable housing, new projects, and community initiatives. In contrast, user comments on platforms like nu.nl focus more on issues related to housing costs, construction quality, and personal experiences with housing.

PMI analysis highlighted associations between specific terms across LinkedIn and nu.nl. Revealing that while Fijn Wonen discusses industry-specific and promotional topics, users often bring up concerns and practical issues. The alignment of topics such as "betaalbaar" (affordable) from Fijn Wonen's posts with user concerns about "prijs" (price) and "kosten" (costs) suggests that affordability is a mutual concern, but the perspective and framing of this topic differ between the company and its audience.

The sentiment analysis indicates a need to address the negative and neutral sentiments expressed by users. By directly responding to user concerns and providing more detailed information on contentious issues, Fijn Wonen can build trust and improve engagement (Keyhole) (Emplifi).

While promotional content is important, incorporating more user-centric information that addresses common concerns about costs and construction quality can make the posts more relatable and engaging (Keyhole) (Emplifi). This will have impact on the users from the nu.nl forum. Highlighting positive user experiences can help counterbalance negative sentiments and create a more positive overall perception. Introducing more interactive content such as Q&A sessions, polls, and discussion threads can encourage user participation and provide valuable insights into user preferences and concerns (Emplifi).

Continuous monitoring of user feedback and sentiment analysis can help Fijn Wonen adapt their strategy in real-time, ensuring that their content remains relevant and engaging (Emplifi). Utilize High PMI Word Pairs in Content Creation: Creating content that naturally incorporates high PMI word pairs can ensure that the language used by Fijn Wonen resonates more with users, potentially enhancing relatability and engagement.

The analysis demonstrates that Fijn Wonen's strategic, positive, and innovation-focused posts effectively promote their brand. User engagement is significantly influenced by how well the content addresses practical concerns and broader industry topics. By incorporating more customer-centric information and timely responses to queries and feedback, Fijn Wonen can foster deeper engagement and build stronger relationships with its audience. Moreover, understanding the nuances in user reactions provides valuable insights for refining their social media strategy to better meet the needs and expectations of their potential and current customers.

6.1 Limitations

This study faces some limitations, such as the reliability of the data sources. Some of the data, being self-scraped, may contain inaccuracies that affect the analysis. Combining datasets from multiple sources can introduce inconsistencies and errors, particularly if the sources have different formats, standards, or quality.

The second limitation is the Time Frame. The study appears to analyze sentiments and topics over a longer period. However, the data from www.fijn.com is only here from the beginning of the year 2024. By addressing these limitations, future research can build upon the foundation laid by this study, offering a more comprehensive and nuanced analysis of sentiments on social media platforms.

6.2 Ethical implications and consideration

Privacy and confidentiality are taken into account in this research. The data used is anonymous, ensuring that no personally identifying information is collected. For the ethical use of data, this research only uses specific data for the purposes for which it was collected. Adhering to legal and regulatory standards is not only an ethical obligation but also a legal requirement. Regularly reviewing and updating practices to ensure ongoing compliance as laws and regulations evolve.

6.3 Recommendations for Marketing Fijn Wonen

This study has been conducted in collaboration with Marketing from Fijn Wonen in an attempt to automatically extract substantive topics and their sentiment on monthly base.

Firstly, TM models are capable of capturing topics that are discussed on primary platforms. The findings provide interesting and applicable insights, but also present challenges. Therefor in the results there are words that come back a lot. “huis”, “we” by filtering out these words other words and topics will be shown more.

Secondly, regarding the SA techniques, the sentiment classifier demonstrated a reasonably good performance on sentiment classification. It is of importance to use the outcome of the sentiment classifier with caution. The model is able to give a general impression about the sentiment, however will struggle with the lack of context and sarcasm. Seen in the result below:

Stukje grond van 200k er bij en je kan wonen.	Neutraal	Other	Affordable housing

This is clearly sarcasm but the model labeled it as neutral and other. Also adding “Not clear” to sentiment analysis would drop the amount of neutral tweets.

Furthermore, we can see that certain posts, podcasts, and lobbyists will highly increase the traffic on the website www.fijn.com. As seen in the results, the high views, clicks and traffic. This is something to keep focus on.

7 List of references

Abiola, S., Kumar, A., & Alaba, F. A. (2014). Sentiment analysis on social media platforms: Challenges and prospects. In Proceedings of the World Congress on Engineering and Computer Science (Vol. 1).

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. Journal of Machine Learning Research, 3(Jan), 993-1022.

Bouma, G. (2009). Normalized (Pointwise) Mutual Information in Collocation Extraction. In Proceedings of the Biennial GSCL Conference.

Chang, J., Boyd-Graber, J., Gerrish, S., Wang, C., & Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. In Advances in Neural Information Processing Systems (pp. 288-296).

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20(1), 37-46.

Duffett, R. G. (2015). Facebook advertising's influence on intention-to-purchase and purchase amongst Millennials. Internet Research, 25(4), 498-526.

Emplifi. (n.d.). In Emplifi - Unified CX Platform for Social Media, Marketing, and Customer Care. Retrieved from https://emplifi.io

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. International Journal of Information Management, 35(2), 137-144.

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. Proceedings of the National Academy of Sciences, 101(Suppl 1), 5228-5235.

Hajli, N. (2014). The role of social support on relationship quality and social commerce. Technological Forecasting and Social Change, 87, 17-27.

Hanna, R., Rohm, A., & Crittenden, V. L. (2011). We’re all connected: The power of the social media ecosystem. Business Horizons, 54(3), 265-273.

Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of the 8th International AAAI Conference on Weblogs and Social Media.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. Business Horizons, 53(1), 59-68.

Keyhole. (n.d.). In Keyhole - Social Media Analytics and Influencer Marketing. Retrieved from https://keyhole.co

Kumar, A., & Mirchandani, R. (2012). Increasing engagement in social media: The role of reply tactics and individual-level characteristics. Journal of Interactive Marketing, 26(2), 74-83.

Kumar, A., Bezawada, R., Rishika, R., Janakiraman, R., & Kannan, P. K. (2016). From social to sale: The effects of firm-generated content in social media on customer behavior. Journal of Marketing, 80(1), 7-25.

Krippendorff, K. (1970). Estimating the reliability, systematic error and random error of interval data. Educational and Psychological Measurement, 30(1), 61-70.

Leung, D., Law, R., Van Hoof, H., & Buhalis, D. (2013). Social media in tourism and hospitality: A literature review. Journal of Travel & Tourism Marketing, 30(1-2), 3-22.

Mangold, W. G., & Faulds, D. J. (2009). Social media: The new hybrid element of the promotion mix. Business Horizons, 52(4), 357-365.

Ranganathan, C., Mehta, V., Valkunde, T., & Moustakas, E. (2020). Topics, Trends, and Sentiments of Tweets About the COVID-19 Pandemic: Temporal Infoveillance Study. Journal of Medical Internet Research, 22(10), e22624. https://doi.org/10.2196/22624

Smith, A. N. (2020). Understanding social media marketing: A case study on building brand awareness. Journal of Business Studies Quarterly, 11(3), 1-12.

Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., & Kappas, A. (2012). Sentiment strength detection in short informal text. Journal of the American Society for Information Science and Technology, 63(1), 163-173.

Tuten, T. L., & Solomon, M. R. (2017). Social Media Marketing. Sage Publications.

Wang, C., Blei, D. M., & Heckerman, D. (2016). Continuous time dynamic topic models. In Proceedings of the 23rd International Conference on Neural Information Processing Systems (pp. 358-366).

8 Appendix

8.1 A

Guide to label data

Guide to label data

Introduction

This manual is intended for test persons who need to manually label data in certain categories. The first categories will be positive, neutral or negative. After that there will follow some more categories (shown here below in the table). The goal is to ensure a consistent approach, so that all test takers assess and label data in the same way.

Categories Tabel

Sentiment Analysis	Social Sentiment	Topic Categories	Intent Categories	Emotion Categories	Info Categories
Positive	Real	Product Feedback	Complaint	Happiness	Company
Neutral	Sarcasm	Customer Service	Inquiry	Sadness	Affordable housing
Negative		Company News	Praise	Anger	Discussion
		Marketing/Promotion	Suggestion	Fear	
		Application message	Spam	Surprise	
		Industry Trends	Promotion	Disgust	

Requirements

- Access to the data that needs to be labeled
- A central system or tool to input the labels (spreadsheet, database, etc.)
- This guide for reference

Steps for Labeling

Step 1: Understand the Categories

Positive

- Definition: Data with a positive connotation, where the overall tone is optimistic, happy, satisfied, or supportive.
- Examples:
    - “The product works excellently!”
    - “I am very pleased with the service”

Neutral

- Definition: Data that does not express strong positive or negative emotions, but rather provides objective, factual, or descriptive information.
- Examples:
    - “The product was delivered on Tuesday”
    - “The jacket is blue”

Negative

- Definition: Data with a negative connotation, where the overall tone is pessimistic, angry, disappointed, or disapproving.
- Examples:
    - “The device broke quickly”
    - “I am dissatisfied with the customer service”

Step 2: Individual Labeling

1. Read the Data Carefully: Ensure you fully understand the context and meaning.
2. Determine the Tone: Decide if the tone of the data is positive, neutral, or negative based on the definitions in Step 1.
3. Enter the Label: Record the appropriate label (Positive, Neutral, or Negative) in the central system.

Step 3: Quality Control and Evaluation

1. Double-Check your Labels and compare the labels with the different test persons.
2. Discuss Discrepancies: Analyze any discrepancies and discuss them.
3. Update the Guide if Necessary: Revise this guide if new insights or definitions are needed to improve labeling accuracy.

Conclusion

Consistent and accurate data labeling requires attention, diligence, and collaboration. By following this guide, test persons can ensure that data is labeled uniformly and reliably, enhancing the quality and usability of the analyzed information.

8.2 B

Sentiment Analysis

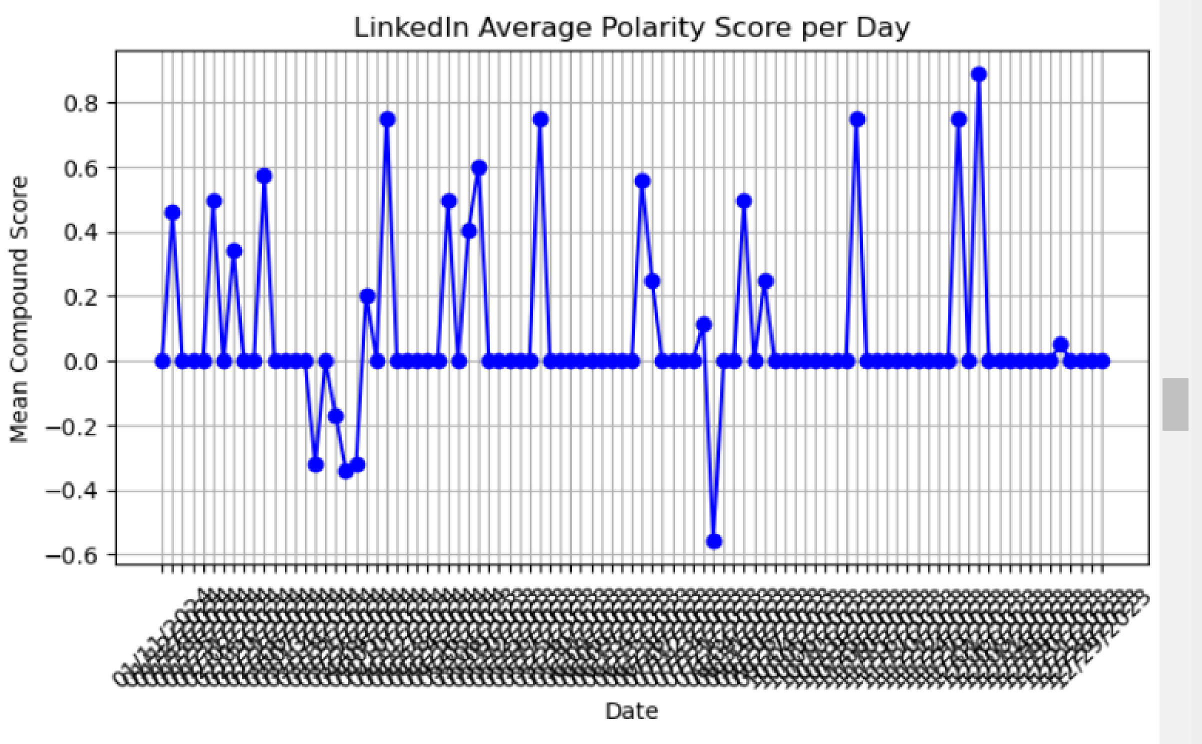

Figure 7: LinkedIn Polarity score a day

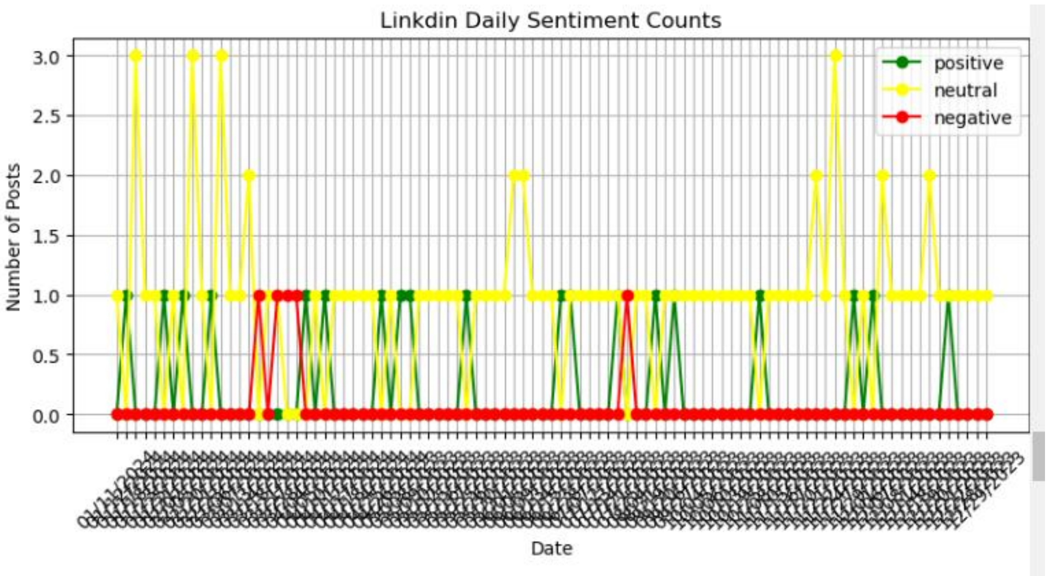

Figure 8: LinkedIn Daily Sentiment counts

8.3 C

Topic Modeling

Figure 9: TM wordcloud LinkedIn

Figure 10:TM wordcloud Nu.nl

8.4 D

www.Fijn.com

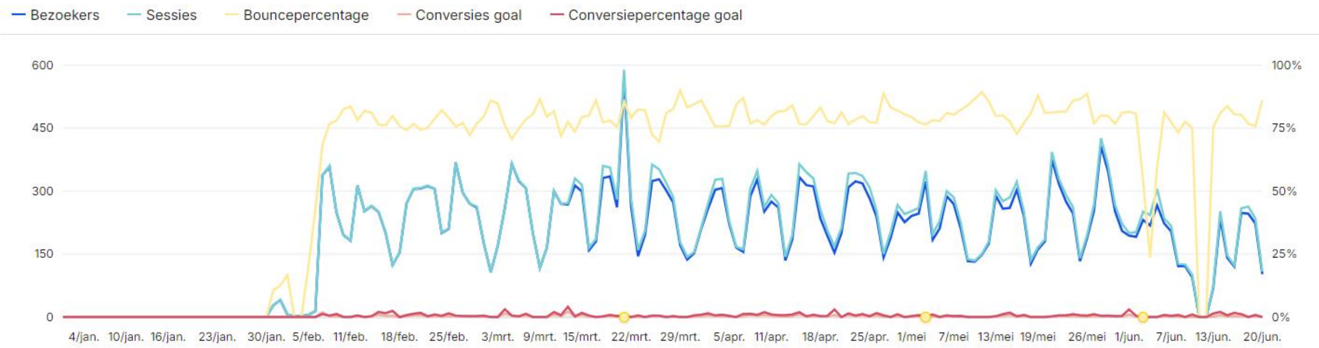

Figure 11: Bouncepercentage, visitors and sessions www.fijn.com

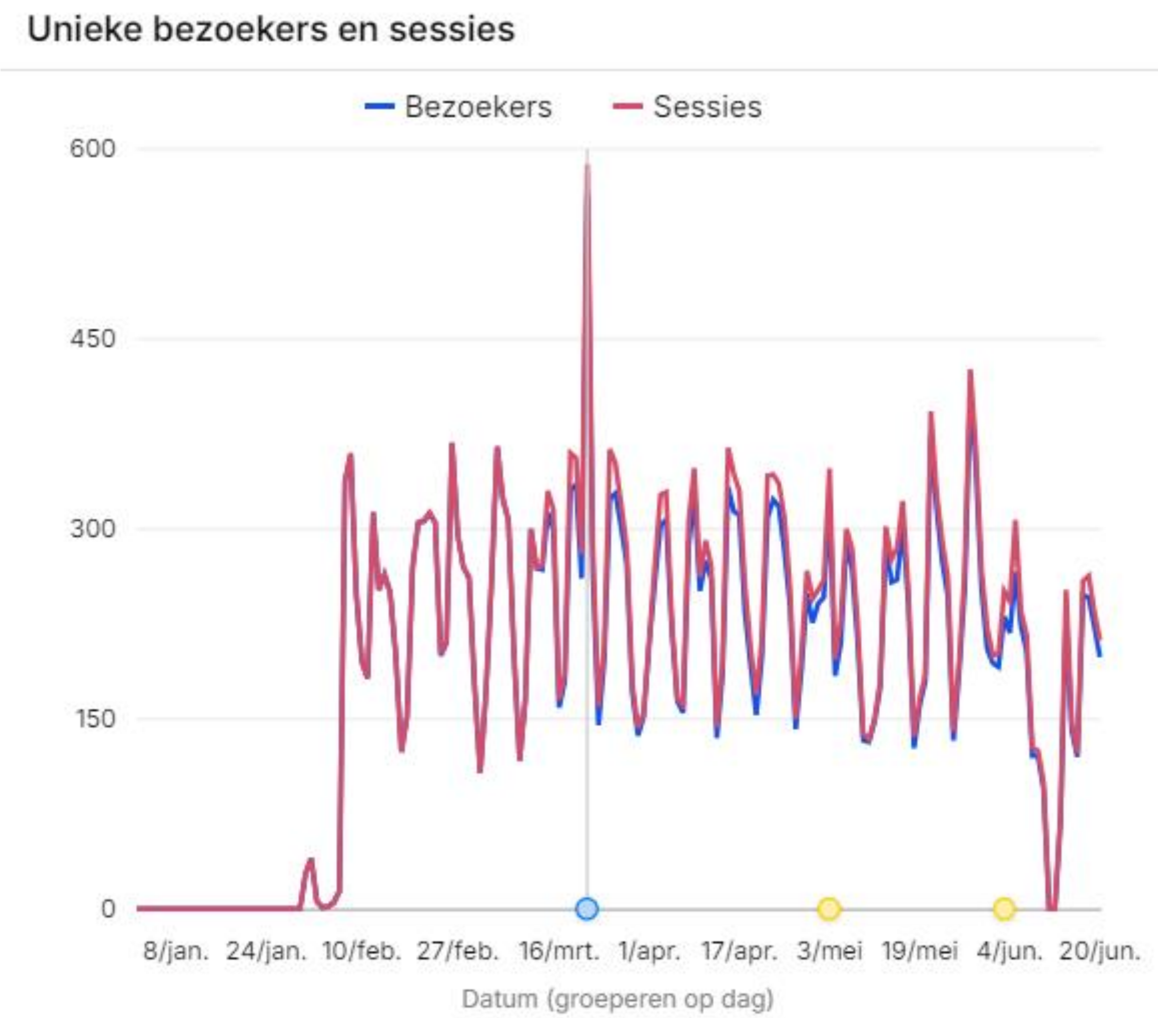

Figure 12:Visitors and sessions www.fijn.com

Download informatiepakket	43 19,82%	Direct entry
Download informatiepakket	33 15,21%	Search engine
Download informatiepakket	27 12,44%	Campaign
Aanvraag rondleiding fabriek	23 10,6%	Direct entry
Aanvraag kennismaking	13 5,99%	Direct entry
Aanvraag rondleiding fabriek	11 5,07%	Campaign
Aanvraag rondleiding fabriek	11 5,07%	Search engine
Aanvraag kennismaking	9 4,15%	Campaign

Figure 13: Actions on www.fijn.com

Figure 14: New vs Returning visitors

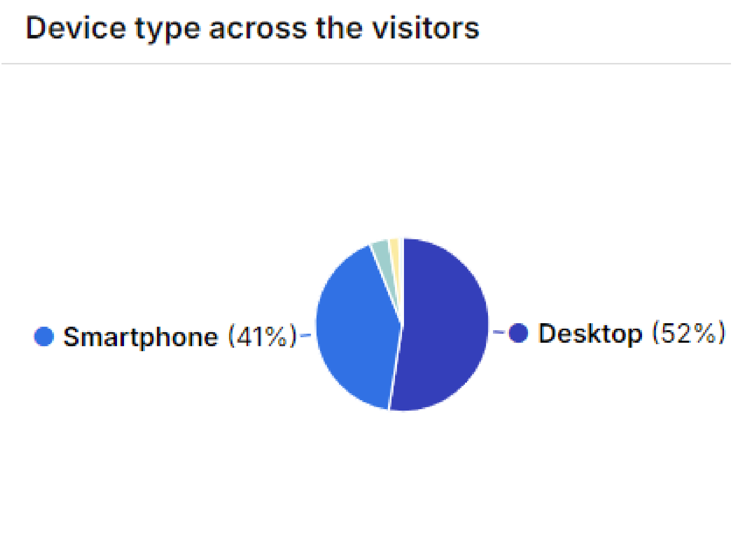

Figure 15: Device type across visitors

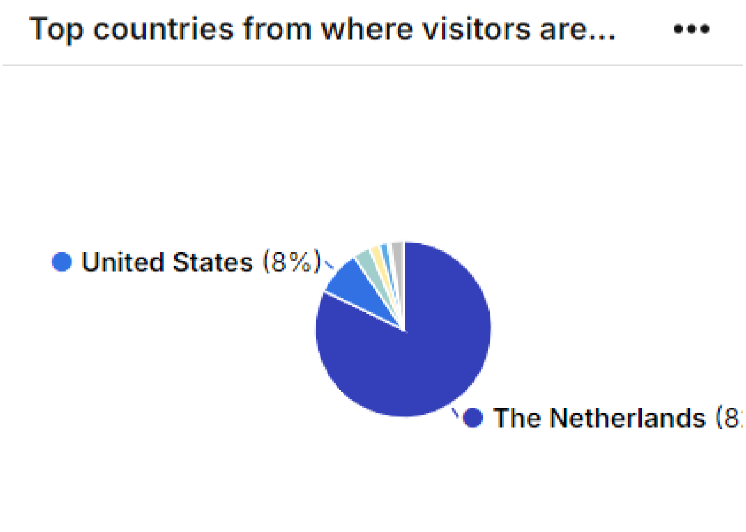

Figure 16: Top countries from visitors

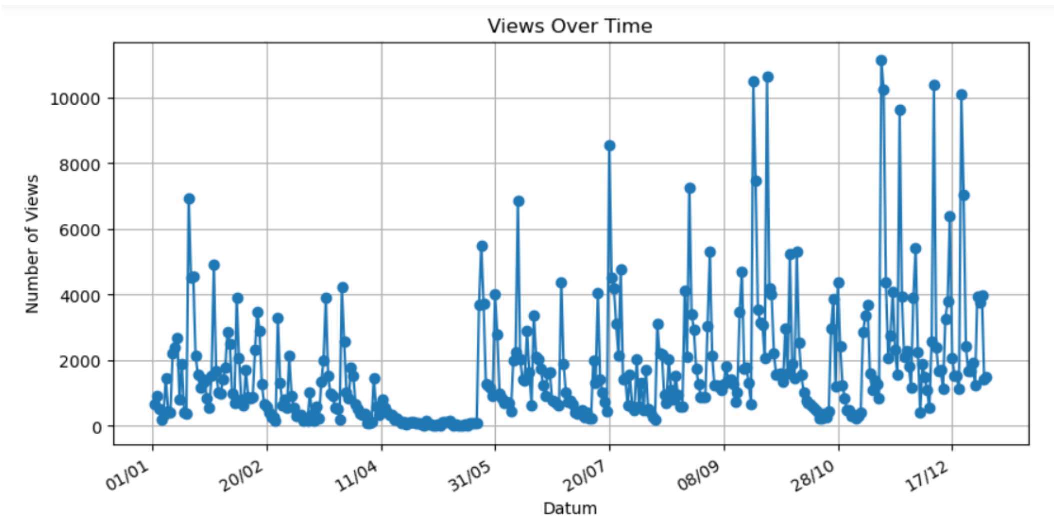

Figure 17: Views over time www.fijn.com

8.5 E

8.5.1 Cohen's kappa

Cohen's kappa (κ) is a statistic used to measure inter-rater agreement for categorical items. It was introduced by Jacob Cohen in 1960 and is widely used in various fields, including psychology, sociology, medicine, and linguistics. Cohen's kappa assesses the agreement between two raters who classify items into mutually exclusive categories. It corrects for the agreement that could be expected by chance alone (Cohen, 1960).

Cohen's kappa evaluates the agreement between two raters beyond what would be expected by chance. The value of kappa ranges from -1 to 1:
κ = 1: Indicates perfect agreement, which means the agreement between raters is completely beyond what would be expected by chance
κ = 0: Agreement equal to what would be expected by chance alone.
κ < 0: Agreement is worse than chance.

8.5.2 Krippendorff's alpha

Krippendorff's alpha is a measure of agreement that is widely used in the field of communication and beyond. Developed by Klaus Krippendorff in the 1970s, this measure quantifies the extent to which raters (or observers) agree on categorical decisions or judgments. Unlike Cohen's kappa, which is specifically used for nominal categories, Krippendorff's alpha is more versatile and can handle various types of data, including nominal, ordinal, interval, and ratio levels (Krippendorff, 1970).

Krippendorff's alpha is calculated based on the ratio of the observed disagreement among raters (Do) to the expected disagreement by chance (De). The formula for Krippendorff's alpha is:

$$\alpha = 1 - \frac{D_o}{D_e}$$

Figure 18: Krippendorff's alpha formula

Do is the observed disagreement among raters.
De is the expected disagreement by chance.

Krippendorff's alpha is an adaptable measure of agreement that has found widespread application in fields such as communication, media studies, psychology, sociology, and beyond. Its ability to handle various types of data and account for chance agreement makes it a valuable tool for assessing the reliability and consistency of judgments or decisions made by multiple raters or observers (Krippendorff, 2018).

8.6 F

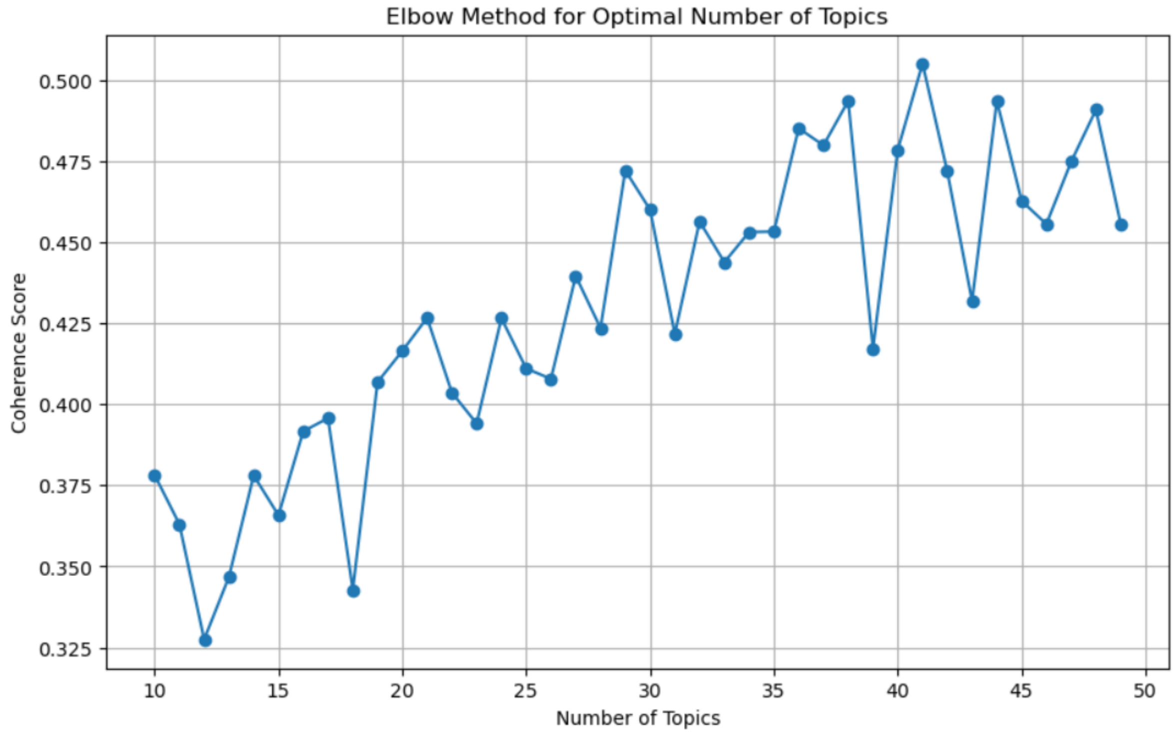

Figure 19: Elbow method LinkedIn v2

8.7 H

Topic 5

Topic 6