INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Social media insights about Nutri-Score topic: a text mining approach

Inês Gomes de Carvalho

Master in **Integrated Business Intelligence Systems**

Supervisor
PhD. Professor Vítor Manuel Basto Fernandes, Assistant Professor with habilitation,
ISCTE – University Institute of Lisbon

Co-Supervisor
PhD. Candidate Ulpan Tokkozhina, Invited Assistant,
ISCTE – University Institute of Lisbon

November, 2022

TECNOLOGIAS
E ARQUITETURA

Department of Information Science and Technology

Social media insights about Nutri-Score topic: a text mining approach

Inês Gomes de Carvalho

Master in **Integrated Business Intelligence Systems**

Supervisor
PhD. Professor Vítor Manuel Basto Fernandes, Assistant Professor with habilitation,
ISCTE – University Institute of Lisbon

Co-Supervisor
PhD. Candidate Ulpan Tokkozhina, Invited Assistant,
ISCTE – University Institute of Lisbon

November, 2022

“Ambition is the path to success. Persistence is the vehicle you arrive in.” – Bill Bradley

Aknowledgements

From my perspective, the completion of the master's degree has a much greater meaning than the delivery of this thesis. It means, besides it, a personal overcoming, the closing of a long-desired chapter, and a feeling of duty done.

Despite the personal achievement, I am fully aware that none of this could have been done without the support of my supervisor and co-supervisor, Professor Vítor Fernandes and Ulpan Tokkozhina, who were always available to answer any and all questions I had and who guided me to achieve this final result.

A thank you to those who are present not only in this but in all other moments of my life:

To my parents, for always encouraging me to run after my goals and overcome adversity. For believing in me and demonstrate me how proud they are of the path I am building.

To my brother, Daniel, for knowing me and knowing the moments when he needs to push for me and give me the motivation and strength I need.

To my grandparents - to grandmother Bia, for being an example and accompanying me throughout my school journey, for always encouraging me to achieve my goals. To grandmother Lena and grandfather Mário, who, even if physically far away, were always present in important moments and to whom I cannot express how grateful I am for giving me the strength I need and the motivation when it seems to be lacking.

To the other family that life made me meet and which I chose - to Mónica, for having been present throughout the writing of this dissertation, for being by my side when I needed company to write it (and for making me discover places where I could find that inspiration needed), for never letting me go after the not so good thoughts and relativizing everything that I thought could be a problem, for reminding me (every time I forgot) mainly of the "why", a thank you is not enough. To my friends that the Management course in ISCTE made me meet (Ferreira, Margarida, Hugo, Franjão, ...) and to whom I thank for the moments of fun and also encouragement.

To all the other people, that despite not naming them individually, friends, master degree colleagues and co-workers, my sincere thanks.

Resumo

Os sistemas de informação nutricional estão cada vez mais presentes no quotidiano dos consumidores, o que traduz também a necessidade da Comissão Europeia de introduzir um sistema único de partilha de informação nutricional de um produto.

O Nutri-Score é um destes sistemas, e, apesar de já estar a ser implementado de forma voluntária em vários países europeus, alguns consumidores, responsáveis políticos e mesmo organizações da indústria alimentar mostram-se reticentes a tal sistema.

O objetivo desta dissertação é apoiar na tomada de decisão dos responsáveis relativamente à aplicação do Nutri-Score como o sistema de informação nutricional nos produtos na Europa partilhando a opinião dos consumidores relativamente a este tópico. Para isso, e tendo em conta os estudos já realizados diretamente aos consumidores (nomeadamente, questionários), alavancar os mesmos numa ótica de análise de sentimento. Nesse sentido, foram extraídos dados da rede social Twitter que contivessem alguma ligação a este sistema de informação nutricional.

Com a informação obtida foram aplicadas técnicas de pré-processamento dos dados, aplicados dois métodos de análise de sentimento e, com os resultados obtidos, foi feita uma comparação (ao nível do país) com os resultados dos questionários já realizados noutras pesquisas. A demonstração dos resultados obtidos e subsequente comparação foi suportada por visualizações as quais podem então servir de ferramenta para posterior análise e tomada de decisão.

Palavras-Chave: Redes sociais, Nutri-Score, Front-of-Package Labeling Systems, Sistemas de Informação Nutricional, Análise de sentimento, Opinião pública

Abstract

Front-of-package labeling systems are increasingly present in consumers' daily lives, which also reflects the European Commission's need to introduce a single system for sharing a product's nutritional information.

Nutri-Score is one of these systems, and although it is already being implemented on a voluntary basis in several European countries, some consumers, policy makers, and even food industry organizations are reluctant to such a system.

The goal of this dissertation is to support the decision making of policy makers regarding the implementation of Nutri-Score as the nutritional information system in Europe by sharing the consumers' opinion on this topic. For this, and considering the studies already conducted directly to consumers (namely, questionnaires), leverage them in a sentiment analysis perspective. In this sense, data was extracted from the social network Twitter that contained some connection to this nutritional information system.

With the information obtained, data pre-processing techniques and two sentiment analysis methods were applied and, with the results obtained, an analysis of it and a comparison was made (at the country level) with the results of questionnaires already conducted in other research. The demonstration of the results obtained and subsequent comparison was supported by visualizations which can then serve as a tool for further analysis and decision making.

Keywords: Social media, Nutri-Score, Front-of-Package Labeling Systems, Nutrition Information Systems, Sentiment Analysis, Public Opinion

Contents

Aknowledgements iii
Resumo v
Abstract vii
Chapter 1 Introduction 1
1.1 Motivation and Topic Relevance 2
1.2 Question and Research Goals 3
1.3 Methodological Approach 4
1.4 Structure and Organization of Dissertation 5
Chapter 2 Introductory Concepts 7
2.1 Front-of-Package Systems 7
2.2 Nutri-Score Labelling System 9
2.3. Potential unintended negative consequences of front-of-package food labels 10
2.4 Text Mining 12
2.5 Sentiment Analysis 14
Chapter 3 Literature Review 17
3.1 Review Methodology 17
3.2 Data Analysis applied to Front-Of-Package Label Systems 18
3.2.1 Text Mining on Front-of-Package Labelling Systems topic 20
3.2.2 Nutri-Score Countries Evaluation 22
3.3 Text Mining applied to public opinion on healthcare topics 24
3.4 Major Findings of the Systematic Literature Review 26
Chapter 4 Methodology 29
4.1 Business Understanding 30
4.2 Data Understanding 31
4.2.1 Information Source Understanding and Search Terms 31
4.2.2 Data Extraction method 33
4.3 Data Preparation 36
4.3.1 Data Pre-Processing tasks 39
4.4 Modeling for Knowledge Extraction 44
4.4.1 Sentiment Analysis of Nutri-Score tweets 44
4.4.2 Sentiment Analysis Methods Comparison 52
4.4.3. Sentiment Analysis by Country 54

Chapter 5 Knowledge Extraction and Analysis 57
5.1 Knowledge Extraction of Sentiment Analysis 62
Chapter 6 Conclusions and Future Work 69
6.1 Limitations and Future Work 70
References 73

List of Tables

Table 1 - Literature review search queries 18
Table 2 - Method and Conclusions of "Sentiment Analysis" article 26
Table 3 - Method, Objective and Conclusion of Nutri-Score Countries articles 27
Table 4 - Method, Techniques and Conclusions on Text Mining Applied to Healthcare articles 28
Table 5 – SNScrape columns retrieved 35
Table 6 - SNScrape columns extracted 37
Table 7 - Example of a French-English translation 38
Table 8 -Type of data extracted 40
Table 9 - Example of stop words removal 42
Table 10 - Columns created based on Pre-Processing transformations 43
Table 11 - Examples of VADER results 47
Table 12 - Overall dataset VADER polarity result (%) 47
Table 13 - Clean Tweet column TextBlob result (examples) 49
Table 14 - Lowercased Clean Tweet without Stop Words column TextBlob result (examples) 50
Table 15 - Lowercased Clean Tweet without Stop Words, with POS Tagging and Lemma column TextBlob result (examples) 51
Table 16 - Comparison of TextBlob result with VADER result 53
Table 17 - Summary of the overall result of questionnaires 65

List of Figures

Figure 1 - Nutri-Score nutrition labelling 1
Figure 2 - CRISP-DM Model 5
Figure 3 – Countries with either mandatory or voluntary FOPL systems on packages foods 8
Figure 4 - Nutri-Score nutrition labelling 9
Figure 5 - Inter-relationship among text mining techniques and their core functionalities 14
Figure 6 - Sentiment analysis methods 16
Figure 7 - PRISMA 2020 flow diagram for Nutri-Score systematic review 20
Figure 8 - Number of "Text mining applied to healthcare" related articles 25
Figure 9 – (left) CRISP-DM Model and (right) Adapted CRISP-DM Model 29
Figure 10 - Information Source System understanding (Twitter) 32
Figure 11 - Exporting process data flow 34
Figure 12 - Example of "Nutri-Score" tweet 34
Figure 13 - Example of "Nutriscore" tweet 34
Figure 14 - Number of Nutri-Score tweets over time (2017 to 2022) 36
Figure 15 - Number of tweets by keyword 36
Figure 16 - Number of English tweets extracted 38
Figure 17 - Translation and duplicates removal data flow 39
Figure 18 - Example of tweet with emojis 41
Figure 19 – (left) Text Blob vs (right) VADER results 52
Figure 20 - Total Number of tweets and year analysis 57
Figure 21 - Number of tweets per month 58
Figure 22 - Google search results of Nutri-Score news 58
Figure 23 - Number of Nutri-Score news over time (years) 59
Figure 24 - Number of Nutri-Score papers published in Scopus and PubMed 60
Figure 25 - Number of tweets per country in 2017 61

Figure 26 - Number of tweets per country and year 62
Figure 27 - (left) Text Blob vs (right) VADER results 62
Figure 28 - Analysing Polarity and Subjectivity using TextBlob 63
Figure 29 - Sentiment polarity score trend using TextBlob (above) and VADER (below) 64
Figure 30 - TextBlob results per country group 66
Figure 31 - VADER results per country group 66
Figure 32 – (above) TextBlob results per country group (bottom level) and (below) VADER results per country 67
Figure 33 - Subjectivity and Polarity analysis by country group using TextBlob 68

Abbreviations

CEAP	**C**ircular **E**conomy **A**ction **P**lan
CNL	**C**olour **N**utrition **L**abel
CRISP-DM	**CR**oss-**I**ndustry **S**tandard **P**rocess for **D**ata **M**inning
CSV	**C**omma **S**eparated **V**alues
EU	**E**uropean **U**nion
FoP	**F**ront-**o**f-**P**ackage
FoPL	**F**ront-**o**f-**P**ackage **L**abeling
FoPLS	**F**ront-**o**f-**P**ackage **L**abeling **S**ystems
NB	**N**aïve-**B**ayes
NLP	**N**atural **L**anguage **P**rocessing
NLTK	**N**atural **L**anguage **T**ool**K**it
POS	**P**art-**O**f-**S**peech
PRISMA	**P**referred **R**eporting **I**tems for **S**ystematic reviews and **M**eta-**A**nalysis
RQ	**R**esearch **Q**uestions
SOM	**S**ocial **O**pinion **M**ining
SVM	**S**upport **V**ector **M**achine
VADER	**V**alence **A**ware **D**ictionary and s**E**ntiment **R**easoner
WHO	**W**orld **H**ealth **O**rganization

CHAPTER 1

Introduction

The Front-of-Package Labeling Systems (FoPLS) makes part of the public health nutrition strategy, being recommended by World Health Organization (WHO) that guides countries to develop, implement, and monitor such systems. The principal goal of these systems is to make it easier to educate the population about nutrition and to facilitate the access to nutritionally healthy food, allowing the consumer, in an autonomous and voluntary way, see and compare the nutritional value of pre-packaged foods (Egnell, Talati, et al., 2020).

Analysing on a European perspective, there is still no agreement on which FoPLS should be used by all countries, allowing it to have a harmonized way of showing a simplified nutrition information of products. However, the aim is to by the end of 2022 formally agree on the best way to show this kind of information (Egnell, Talati, et al., 2020).

Nutri-Score makes part of the Front-of-package (FoP) labels that have the main goal of helping consumers to make healthier food choices. This statement was validated by (Temple & Fraser, 2014)where it was demonstrated that food labels have a lot of value in what regards the assistance of millions of people in making healthier food choices.

The Nutri-Score is also known as 5-Colour Nutrition Label (5-CNL), as it consists of an image from the letter A to E (represented in figure 1), each in a different colour, to consider the products as healthier and least healthy, respectively (Hagmann & Siegrist, 2020).

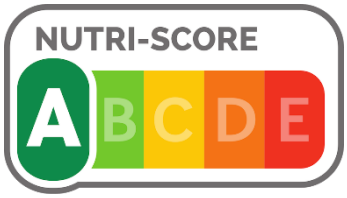


Figure 1 - Nutri-Score nutrition labelling[1]

[1] (*File:Nutri-Score-A Light Background Logo.Svg - Wikimedia Commons*, 2022)

Although in most of the studies conducted it is uniformly agreed that an important strategy for helping consumers identify the various types of foods and their degree of healthiness resides in a piece of explicit and easy-to-understand nutrition information, there is still an absence of agreement about which format best communicates it to the consumers (Hagmann & Siegrist, 2020).

Adding to this, when speaking about Nutri-Score, some questions arise on the opponent's side, whether about the way that the algorithm is calculated or about the way that it leads with ultra-processed foods.

More than the government's opinion and subsequent decision, it is also important to understand what is the consumer's opinion on this topic. There were principally two kinds of approaches already taken to answer this problem: online/real-life questionnaires and social media sentiment analysis.

Septia Irawan et al., 2022, did similar work to this current research, where it extracted Twitter data and analysed the text sentiment in what regards to FoPLS.

The current approach extracted also Twitter data but focused on the Nutri-Score topic, to understand what the sentiment of users is when speaking about such labelling system. There were principally two sentiment analysis techniques applied and, in the end, more than analyzing the sentiment polarity it was also done a country study, where it was compared the results of this Twitter data with the online/real-life questionnaires made by other researchers.

1.1 Motivation and Topic Relevance

The food system has been included in the main action lines for the Circular Economy Action Plan (CEAP) of Europe. More than the food system itself, an important point that is highlighted is the fact that, to guarantee that new measures are well succeeded, there must be a shift in people's diets (Directorate-General for Health and Food Safety (European Commission) & European Commission, 2020).

Population diet is also part of the action plan for the prevention of noncommunicable diseases developed by WHO. More specifically, it is recommended that food nutrition labelling acts as a policy tool, assisting consumers in defining which are the foods with less sugar, for example (Health Organization Regional Office for Europe, 2016).

In the last few years, Front-of-Package nutrition labels, i.e., supplementary nutrition information printed on the front of food products packages, have been implemented as a way to contribute to the improvement of the population's knowledge of proper lifestyle, directly related to the prevalence of obesity and its health consequences on the population (Carruba et al., 2022).

The debate about the harmonization of FoPLS in Europe gains notoriety. As part of the "Farm to Fork" strategy from the Green Deal, the European commission announced that the limit to select a single front-of-pack nutrition labelling is the end of 2022 (Egnell, Talati, et al., 2020).

With the greater access to the Internet and the subsequent increase of social media popularity, many people started to express their opinions online, especially on Twitter. With 450 million monthly active users known in 2022, Twitter is an undoubted source and spread of information (Pandey et al., 2022).

In spite of being present on French product packages since the middle of 2017, the Nutri-Score front of package label is still not widely accepted. While some experts sign a petition to make Nutri-Score a label mandatory for all European countries and even the population tested defend the idea that it is the best front of package label (Mazzù et al., 2021), others tend to disagree with such premise, arguing that there are gaps in the way that the algorithm is calculated (Julia et al., 2022) .

In a study done by Septia Irawan et al., 2022 on the evaluation of tweets sentiment regarding Front-Of-Packages labelling systems it was understood that the discussions on Twitter are heavily influenced by the customer journey.

1.2 Question and Research Goals

The main goal of this research, as previously mentioned, was to identify on Twitter posts the user's sentiment when speaking about the Nutri-Score topic, and in addition try to leverage the studies already done via direct questionnaires and add to them a text mining approach to analyse the comments present on social media, that can represent the opinion of some users.

After identifying the sentiment behind the Twitter posts, another goal was to analyse the tendency of the countries, to identify whether the topic analysed is more accepted or not, compared to the questionnaire results. Moreover, the analysis of the sentiment throughout the years is also something that was taken into consideration.

To achieve this objective, it was formulated the following set of research questions (RQ):

RQ1. What is the trend when speaking about Nutri Score on the Twitter social network – positive or negative?

RQ2. From which countries are the users with positive/negative thoughts related to Nutri-Score?

RQ3. Is there a difference between 2017 and 2022 in terms of sentiment polarity, when speaking about Nutri-Score?

1.3 Methodological Approach

The industry-proven way for carrying out data mining projects is the framework developed in 1996 and published in 1999 – CRoss Industry Standard Process for Data Mining (CRISP-DM). It was designed by a consortium of leading data mining users and suppliers and it was also partly sponsored by the European Commission (Wirth & Hipp, 2000).

The goal of this methodology is to assist in the development of a project, making sure that all questions – from a business perspective and a technical one – are answered. The communication part of a project is not mistreated, as planning and communicating the results (and also documenting the whole project process) is guaranteed when using this methodology (Wirth & Hipp, 2000).

The workflow of CRISP-DM is divided into six stages (Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation and Deployment), presented in figure 2. The sequence of the phases is not strict because, even if the outcome of each phase determines the input of the next phase, there is the possibility of moving back and forth, depending on the evolution and analysis during the project workflow, adapting it as necessary.

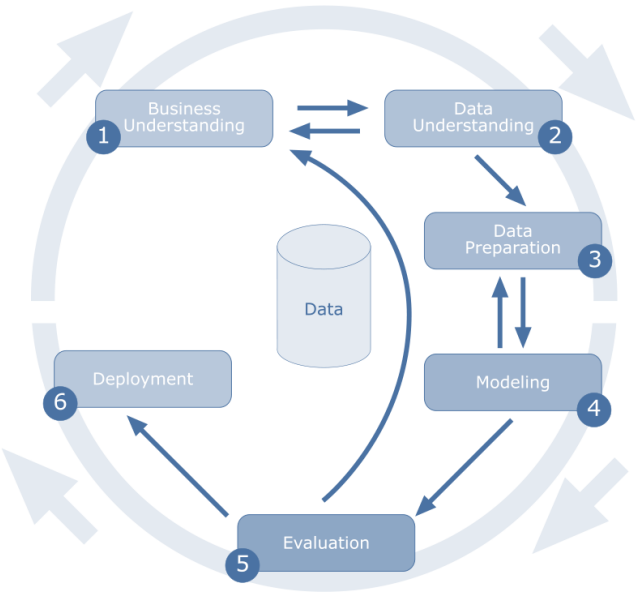


Figure 2 - CRISP-DM Model [2]

1.4 Structure and Organization of Dissertation

Having in mind the motivation and topic relevance and to which research questions this dissertation aims to respond, defined in **Chapter 1**, the following structure of this research is to, in a first instance, in **Chapter 2**, present the introductory concepts of the Front-of-Package Labelling Systems, and, in more detail, on the Nutri-score label. It also presents a review of text mining papers, namely in the sentiment analysis area, to try to identify what has been the evolution in terms of techniques used and the different methods of doing this kind of analysis.

Chapter 3 presents a systematic literature review of the state-of-the-art of data analysis applied to FoPLS, gathering papers that study this topic and the methodologies used in each of them. As the sentiment analysis on FoPLS is still scarce it is also presented a review of papers related to text mining applied to healthcare topics in general.

Next, the main practical part of the current dissertation is presented in **Chapter 4**, where the methodology adopted is used as chapter sections, describing all the steps taken to reach the main goal of the dissertation – sentiment analysis of the Nutri-Score topic.

[2] (Roy, 2018)

In **Chapter 5** the results of the work done are presented, and also the comparison of the analysis performed with other works on this topic previously investigated in chapter 2.

Finally, in **Chapter 6** the conclusions and limitations of the research are outlined, as also as suggestions for future work.

CHAPTER 2

Introductory Concepts

In this chapter, it is presented the introductory concepts in order to contextualize the work done.

Based on the Literature Review, concepts like “Front-of-Package Labelling Systems” and “Nutri-Score” were deeply analysed. With the intention of reviewing the evolution in terms of techniques used in text mining and, more specifically, in the sentiment analysis area, it was also made a review on these topics.

2.1 Front-of-Package Systems

In what regards the nutrition supplementary labelling, printed images on the front of the packages (FOP: Front of Package) have been designed, confronted and tested by consumers all over the world, including, but not limited to: Reference Intakes, Warnings, Multiple Traffic Light and Health Star Rating (Carruba et al., 2022).

In figure 3, it is presented the implementation of this package labelling system in some countries all around the world. It is very explicit the variety of systems that exist, where in the major part of the countries these systems are implemented voluntarily (represented by the orange colour) being the decision of determining which score label to choose by each country’s government (Carruba et al., 2022).

The FoPLS makes part of the public health nutrition strategy, as it complements other measures – its principal goal is to make it easier to educate the population about nutrition and to facilitate access to nutritionally healthy food, allowing the consumer, on an autonomous and voluntary way, see and compare the nutritional value of pre-packaged foods (Hagmann & Siegrist, 2020; Hercberg et al., 2022; Julia et al., 2018).

WHO and the Organization for Economic Co-operation and Development believe that this way of presenting the nutritional value of foods helps consumers by making better and healthier decisions – which will then have an impact on nutrition-related diseases (helping to reduce them) (Julia et al., 2018).

In the "Impact of color-coded and warning nutrition labelling schemes" study, it was concluded that exists clear evidence to support policymakers to push forward mandatory Front-of-Package Labeling (FoPL) policies. This would help not only to direct consumers to a greater food choice, as previously mentioned, but also to encourage the reformulation of some products in the food industry in order to improve the nutritional quality of the goods produced (Song et al., 2021).

In Europe, Front of Package Labeling was first introduced in the late 1980s in Sweden through the adoption of the Keyhole Logo. Since then, a lot of different FoPLS were implemented in other European countries, but there is one that stands out from the others for being not only the most diffused summary label at the European Union (EU) level but also the one that adheres the most to the processes, concepts and validations studies that were published by WHO: Nutri-Score (Mazzù et al., 2021).

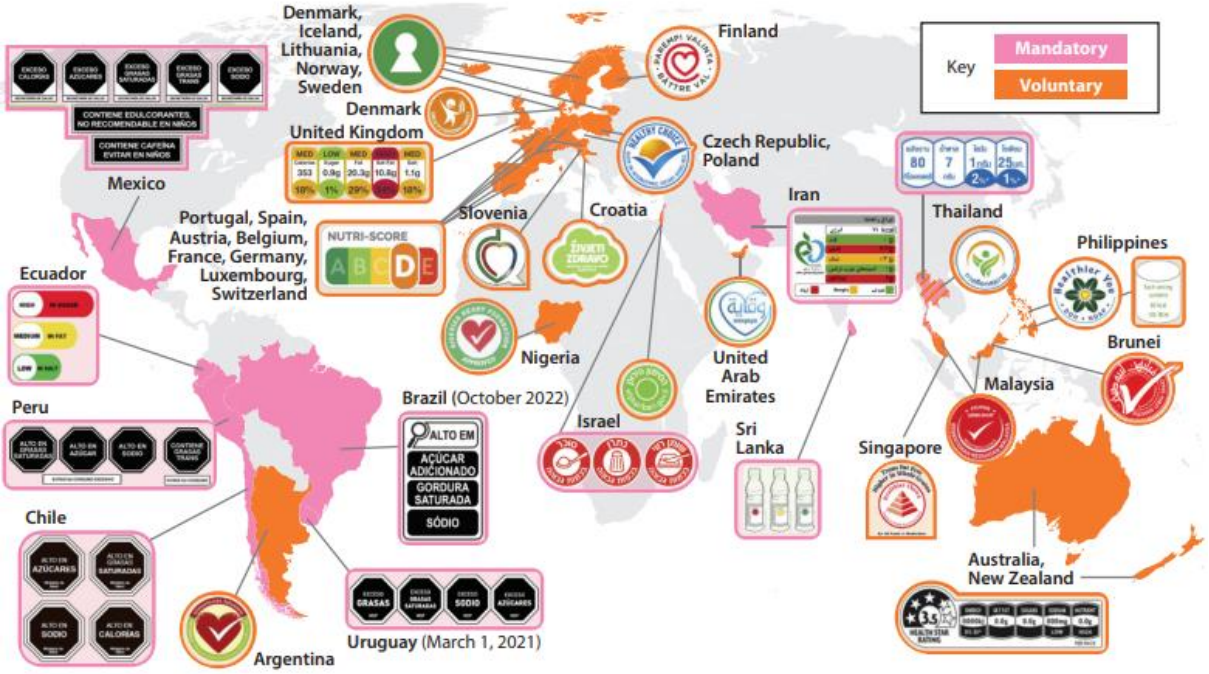


Figure 3 – Countries with either mandatory or voluntary FOPL systems on packages foods [3]

[3] (GAIN Convening Paper Series, 2020)

2.2 Nutri-Score Labelling System

Nutri-Score is a front-of-pack labelling system that has its origin in France. In 2016, a randomised control trial was taken and conducted in real-world conditions, where the Nutri-Score label was selected as the official front of the package system on the 31st October 2017 by the French ministry. Nowadays Nutri-Score label is still voluntary, but in a study made in 2019, it was concluded that at that time it was present in 250 brands in France, representing 25% of the French market (Nutri-Score: The Story so Far - EuroHealthNet, 2019).

This system consists of a colour-coded label based on a five-colour scale (from dark green to red) associated with five letters, from A to E, respectively, and represented once again in figure 4. Products that have a higher nutritional quality are rated as A, and, on the other side, products with lower nutritional quality are rated as E – as so, it is not a dichotomy nutritional label that characterizes food as "healthy" or "unhealthy" - it was essentially built to compare products of the same category, allocating them on one of the 5 categories and let consumers know which one is more or less favourable to health (Hercberg et al., 2022).

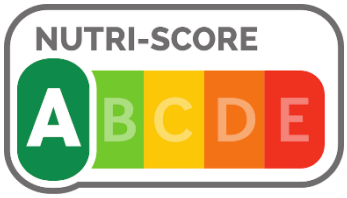


Figure 4 - Nutri-Score nutrition labelling [4]

To guarantee its improvement, the algorithm behind this FOP labelling system is planned to be updated every three years. It was first adapted from the 2005 UK Food Standards and it has as its base the content of the food per 100 grams, using all the components of the back-package nutrition table – i.e., the algorithm awards bad N points to the values of energy, sugars, saturated fat and sodium, and good C points based on the quantity of fiber, protein, fruit and vegetables (Drewnowski et al., 2021; Hercberg et al., 2022).

[4] (File:Nutri-Score-A Light Background Logo.Svg - Wikimedia Commons, 2022)

The good C score points will be subtracted from the bad N points, so, the Nutri-Score can be seen as a partly compensatory score. In what regards to the points for protein, if the food does not contain at least 80% of fruit or vegetables, protein content does not contribute to the final score. The algorithm point results range from 0 to 5 – as an example, foods with a total protein content of ≤1.6 g/100 g are given 0 points, whereas those with >8 g/100 g are given full 5 "C" (good) points (Drewnowski et al., 2021).

The score that results after applying this algorithm is represented by the letters coloured, and it was validated in various epidemiological studies – in volunteer and in representative random samples (Hercberg et al., 2022).

Although in most of the studies conducted it is uniformly agreed that an important strategy for helping consumers identify the various types of foods and their degree of healthiness resides in a piece of explicit and easy-to-understand nutrition information, there is still an absence of agreement about which format best communicates it to the consumers (Hagmann & Siegrist, 2020).

There have been a lot of studies already done [(Mazzù et al., 2021) , (Ares et al., 2018), (Pettigrew et al., 2022)], comparing the different types of nutrition labels, and that let the consumers in real-life situations choose which is the best way to select healthier foods. This topic will be deeply analysed in Chapter 3 – Related Work.

The debate around an EU-wide FOPL model continues to gather momentum as the harmonization of the front-of-pack nutritional labelling is even more needed, as the plurality of FOPL existing on the market may potentially lead to confusion for consumers. Due to the arising of this topic, the European Commission has launched a discussion among member states, to stimulate a reflection on the harmonization of FoPLs (Dréano-Trécant et al., 2020).

2.3. Potential unintended negative consequences of front-of-package food labels

As it could be verified, some studies take the Nutri-Score as the ideal FOPL system to implement on pre-packaged food products. On the other hand, there are other papers

where this algorithm is considered not as much accurate as others to identify correctly the nutritional value of food, as it will be analyzed next.

This topic reveals to be controversial when, in other studies, different arguments against the mandatory usage of FOPL systems arise. In these papers, it is defended that this FOP scheme may risk giving a "healthy aura" to ultra-processed foods, as there are not many labels that have into consideration the processed level of the product (Julia & Hercberg, 2018).

Adding to this, in spite of the main purpose of FOPL systems being to calculate the nutritional value of a given product, one of the problems encountered is the fact that they do not consider the food after it is cooked. Taking Nutri-Score as an example, and as already noted, the way of calculating its algorithm benefits certain nutrients while assigning a negative value to others, but it does not consider how the products are metabolized by the human body itself. The example given in (Roberto et al., 2021) is about granola cereal bars, to which dehydrated fruit is often added and because of this the granola bar can get an improvement in its nutrition score value, even if this fruit has sugars and sodium, that may not be taking into consideration.

When Nutri-Score was considered in France as the principal front-of-pack nutrition label, six big manufacturers, named "Big6" (Nestlé, Coca-Cola, PepsiCo, Unilever, Mars and Modelez) have taken a stand against such label, implementing an alternative label in France that was the result of an adapted version of the British label – Multiple Traffic Light system. After some time, and to prevent confusion among consumers, Big6 removed this adapted label (Julia & Hercberg, 2018).

Another argument of the Nutri-Score opposition is related to the impact that adding this label to traditional foods can have. Foods like French cheese and olive oil, have a high density in terms of calories, explained by the amount of salt and saturated fat, which are considered as having a low nutritional value – due to this, the Nutri-Score value of these products is not good, leading some agro-industry stakeholders and even some political parties stand against such label (Julia et al., 2022).

2.4 Text Mining

The principal purpose of text mining is to find patterns, information or knowledge on a large amount of unstructured data – that can include documents, quoting text, reviews, and others (Wahyuni et al., 2021).

The text mining process can be divided into seven techniques:

- **Document clustering and document classification:** The first one is an unsupervised technique related to finding intrinsic structures in information and grouping similar documents/files into understandable subclasses. While the second is a supervised learning method because the categories of the documents are known in advance. It includes alert detection (e.g., detection of spam on email folders – commonly known as spam filtering) and document categorization (finding the accurate topics for each document) (Dang & Ahmad, 2014).

- **Information Retrieval:** This technique is present in most text mining projects, as it aims to handle textual data by extracting a group of relevant documents from a large collection. It also tries to match those documents, do optimization on search engines and websites based on keywords, and an inverted index that returns a ranked list of documents that may be relevant to a specific user search. Some examples of information retrieval applications are search engines like Google and Yahoo. Academic literature search tools – like Google Scholar – also make use of this technique, as the results retrieved are mainly based on the scope to which documents are connected by web links and citations (Krallinger & Valencia, 2005).

- **Web Mining:** The principal goal of this method is to use web data to extract and find information patterns. It is divided into three categories, to be used according to different analysis targets and resources: web content mining, which analyzes web content (mostly textual data, but also images, videos, sounds, etc.) using mainly Natural Language Processing (NLP) technologies having as an example the identification of web pages relevance; web usage mining, also called log mining, analyses the usage of web by users, determining the navigation behaviour of them (e.g. visited pages, data used); and web structure analysis that focuses on

explaining the structures and links of websites, after extracting the structure of websites (Jawad & Mughal, 2018; Ting, 2008).

- **Concept Extraction:** The concept extraction subdiscipline tries to make the understanding of textual data for automated systems tasks easier, aligning the automated work with the human understanding of data. Defining the polarity of a text analyzed (sentiment analysis), and word association make part of the techniques used in this group (Miner et al., 2012).

- **Information Extraction:** This field aims to, by using a huge amount of text, change that information into a readable one and then extract meaningful knowledge and patterns from it. In the transformation part, the huge amount of text is, in most of the cases, split into tokens, and the extraction of entities/tokens and their relationships is a result of applying this information extraction technique (Tandel et al., 2019).

- **Natural Language Processing:** NLP tasks are used in text mining mostly for statistical processes, on data understanding and pre-processing tasks such as tokenization, lemmatization and Part-of-Speech (POS) (Miner et al., 2012).

In figure 5, it is presented those seven techniques and also some of the core functionalities mentioned above.

Text mining has been implemented in a variety of sectors – from the health sector, using text mining to screen electronic healthcare records in cardiovascular trials (van Dijk et al., 2021), the education sector, to improve text predictions (Ferreira-Mello et al., 2019), until the political one, using text mining to deep in the e-government concept (Abu-Shanab & Harb, 2019).

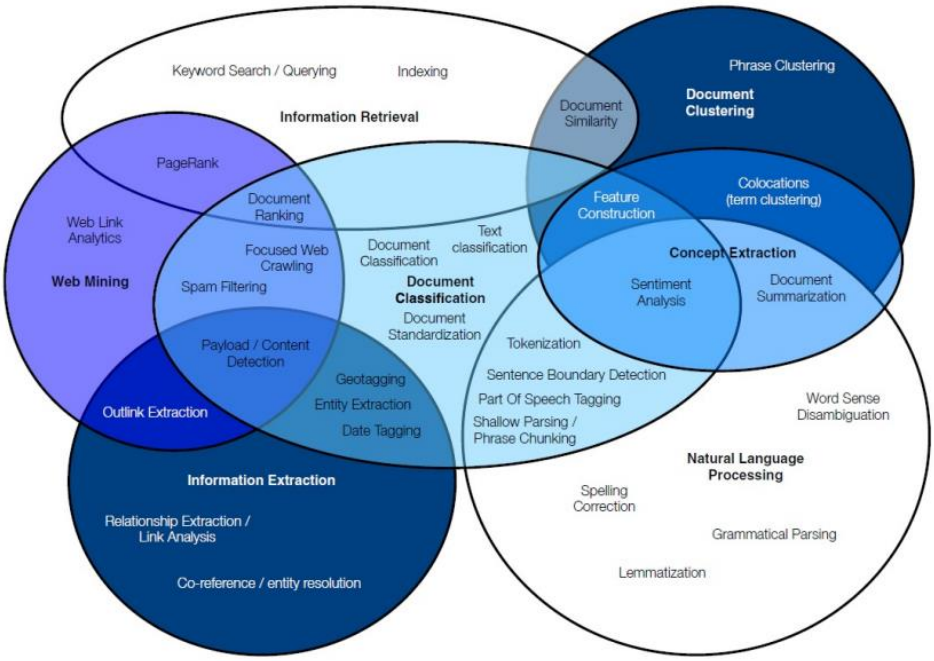


Figure 5 - Inter-relationship among text mining techniques and their core functionalities [5]

2.5 Sentiment Analysis

The branch of text mining that has as a main goal to extract subjectivity and polarity from text is denominated sentiment analysis, determining whether a text leaves a positive, negative, or neutral impression (Aljedaani et al., 2022).

Sentiment analysis can be divided into two main ways of performing it: lexicon-based and machine learning.

Regarding the first approach, it involves calculating the sentiment of a document based on the semantic orientation of words or phrases in that document. This is obtained with the support of a lexicon, also called a "valence dictionary" or "library" - that can be manually adapted by each researcher - that comprises a large number of words (mostly adjectives but also adverbs) that are ranked based on their polarity score. Once a sentence is analyzed, depending on the number of positive/negative words counted it is classified as having a positive/negative sentiment score, respectively (Aljedaani et al., 2022).

This first approach is called "dictionary-based approach" and can also be expanded by using other tools. WorldNet for example, allows the researchers to, by using the libraries already created, incorporate synonyms and antonyms of those words (present

[5] (Talib et al., 2016)

in WorldNet), expanding the dictionary until no new words can be added (N. Gupta & Agrawal, 2020).

In addition, the lexicon-based approach can also be addressed by using a corpus-based approach that considers the sentence context. This approach can also be split into two methods: statistical and semantic – the first one analyses the appearance of the words in positive/negative texts, assigning the polarity to where it appears more frequently (on positive or negative sentences), and the second one has into consideration the synonyms and antonyms of the word analyzed, assigning sentiment values to those words that are close to the one that it is being evaluated (N. Gupta & Agrawal, 2020).

This type of technique is predominantly used in unsupervised machine learning problems (dataset without previous annotation) as the words used in the analysed text will be compared with the ones present in the library (Aljedaani et al., 2022). Gudivada et al., 2015 also claims that a way of increasing the match rate between entities is by using pre-processing steps such as rewriting, stemming, and replacing a word with its synonyms.

In what regards the second approach – machine learning – it requires the split of the data into test and training data, in order to automatically classify the text. The training data consists of text previously labelled, i.e., sentences that are classified as positive, negative or even neutral which represents the lexicon of machine learning approaches; it is with this data that the algorithm will learn from and will try to classify correctly the unlabelled test data. The most well-known methods for sentiment classification are using Naïve Bayes, Random Forest and Support Vector Machine models (Kumaresh et al., 2019).

On sentiment analysis while using machine learning techniques a common method used is the bag-of-words and also n-gram features. The first one assumes the independence of words, not considering the semantic information in the text, ignoring also the subjective information on it (Agarwal & Mittal, 2016). The second method specified, related to the n-gram features, splits the text into words of equal length, and then organizes those words in groups of N-words (Psomakelis et al., 2015).

A disadvantage of machine learning methods is that it involves a manual annotation of the train data, and in the lexicon-based approach this is taken for granted. In recent years, what started to be done was to use a hybrid approach – taking advantage of the positive points of each method (Psomakelis et al., 2015).

In figure 6 it is presented a graph that illustrates what was described above, defining sentiment analysis as two main approaches: lexicon-based and machine Learning. Both of them are split then into more branches, that have specific techniques associated with each.

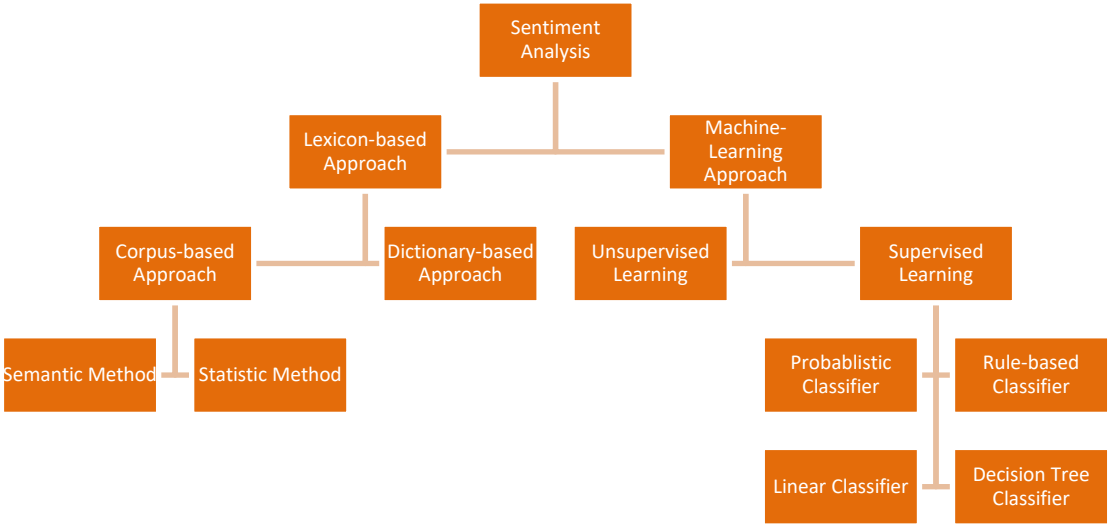


Figure 6 - Sentiment analysis methods (adapted from [6])

Having into consideration the research taken in this chapter, it can be verified that:

1. There is still no harmonization of the FOPL systems used
2. Various studies defend the Nutri-Score as the primary label to be used on products, as its evaluation of readiness, comprehensibility and liking is the best one.
3. There are also other studies that defend that Nutri-Score, when compared to other FOPL systems, is not the best choice, having in mind the accuracy to choose the healthiest product and the way that the algorithm is calculated, and should be taken out from the public health policy makers consideration.
4. There are different ways of addressing a text mining problem, and, more specifically a sentiment analysis one. The one that adheres most to the current research will be deeply analysed in the next chapter, where it will be understood what are the methods normally used in this kind of topic.

[6] (Sharma et al., 2020)

CHAPTER 3

Literature Review

3.1 Review Methodology

The existing literature on Nutri-Score was already presented in the previous chapter, on this one, the intent is to clarify the existing literature gap, defining to what the current work can contribute for. Therefore, it is necessary to search for relevant articles on the topic under study, finding works related to the subject of this research.

The Preferred Reporting Items for Systematic reviews and Meta-Analysis (PRISMA) statement updated in 2020 was used to assist in the systematic review analysis improving the reporting of it, and it consists of a 27-item checklist and a three-phase flow diagram (Page et al., 2021).

It was used two databases for the systematic review: Scopus, that it is an Elsevier database that covers only peer-reviewed literature ranging from articles, conference papers, journals, and books from 1996 forward (Ballew, 2009); and PubMed, a database more related with biomedical topics, being the most widely accessible database of resources of this kind (Williamson & Minter, 2019).

The decision of considering these two databases was principally because Scopus is considered the largest abstract and citation database, increasing the probability of finding articles related to Nutri-Score. Regarding PubMed, as already specified, it is a database related to biomedical topics, and as the current research is on a nutrition/healthcare topic, it was useful to look into a database specific to the subject.

The database search is done based on the search terms that should appear on the documents. These search terms are combined on a query, and the ones used in this systematic review are presented in table 1.

Table 1 - Literature review search queries

Query	Literature Review Queries
1	((nutriscore) OR (nutri-score) OR (front-of-pack) OR (front-of-package)) AND ((data mining) OR (data analysis) OR (text mining))
2	(sentiment analysis) AND (public health OR nutrition OR healthcare)

In the first query, the goal was to find works related to Nutri-Score (or the front-of-package label systems in general) that have some data analysis associated.

To understand what has been the usage of public opinion sentiment analysis on public health topics, and more specifically on nutrition ones, there was also made a research that included these terms, and the second query was used for this purpose.

3.2 Data Analysis applied to Front-Of-Package Label Systems

The PRISMA methodology, represented by a flowchart in figure 7, is explaining, in a visual format, the number of articles analyzed in a first instance, and those that were excluded throughout the analysis by using various techniques.

Applying the first query, the number of articles retrieved on the 24th of August 2022, after applying a filter criteria, to only include documents that have those search terms on the title, abstract or keywords, there were 112 results on Scopus and 95 on PubMed.

The first iteration made was regarding the duplicate records. When combining both articles (the ones from Scopus and PubMed) 63 documents were duplicated – these were removed, remaining 144 for further analysis.

After this, those documents were individually analysed – each document that had a title that did not match the current search criteria was removed and after this, after going through each abstract, the ones that were out of the scope of the current research were also excluded. For the title exclusion criteria, 84 were removed, and for the abstracts reason, 30 did not fit with the topic in analysis.

The final screening included a careful reading of the full 30 document's text, where 9 reports were excluded due to the following reason:

- seven documents related to product evaluation and comparison regarding which products had the best Nutri-Score, regardless of the country where the analysis was done, were excluded. The reason why was to try to focus the systematic review as much as possible on articles that analyze people's opinions on the front-of-package labelling implementation, and also on Nutri-score, more specifically.
- Not up-to-date information - although it was not applied any time frame condition, after reading two of the documents, from 2012, the information was not accurate due to the risen of different front-of-packages labels.

In the end, 21 documents were included in the systematic review. These papers were grouped into two categories: text mining on front-of-package labelling systems topic and Nutri-Score countries evaluation.

The first category is more related to the current research, and, as will be seen in the next section, it was only found one article directly related to it – which highlights that there is a lack of research in this area. In regards to the second category, while each of the 21 articles were being read, it was realized that they had a strong cross-country component. Whether it was an analysis of the Nutri-Score in a specific country, using questionnaires, or, by the same methodology, an analysis in several countries at the same time.

Thus, the main goal was to find works that fit into the first category, since there was no possibility of finding more papers than the one found, it was given priority to all articles that in some way contributed to the perception of the opinion of the population (in this case, using questionnaires and surveys) on the Nutri-Score topic.

Besides this represented methodology, it was also important to guarantee that all the analyzed papers that were interesting for the present study could be properly referenced. For this, it was chosen the Mendeley tool (version 2.76.0), which allows, in a single platform, to store all citations/references, ensure that there are no duplicate articles to be analyzed and with a direct link to the writing tool, which allows a better organization of the references chapter.

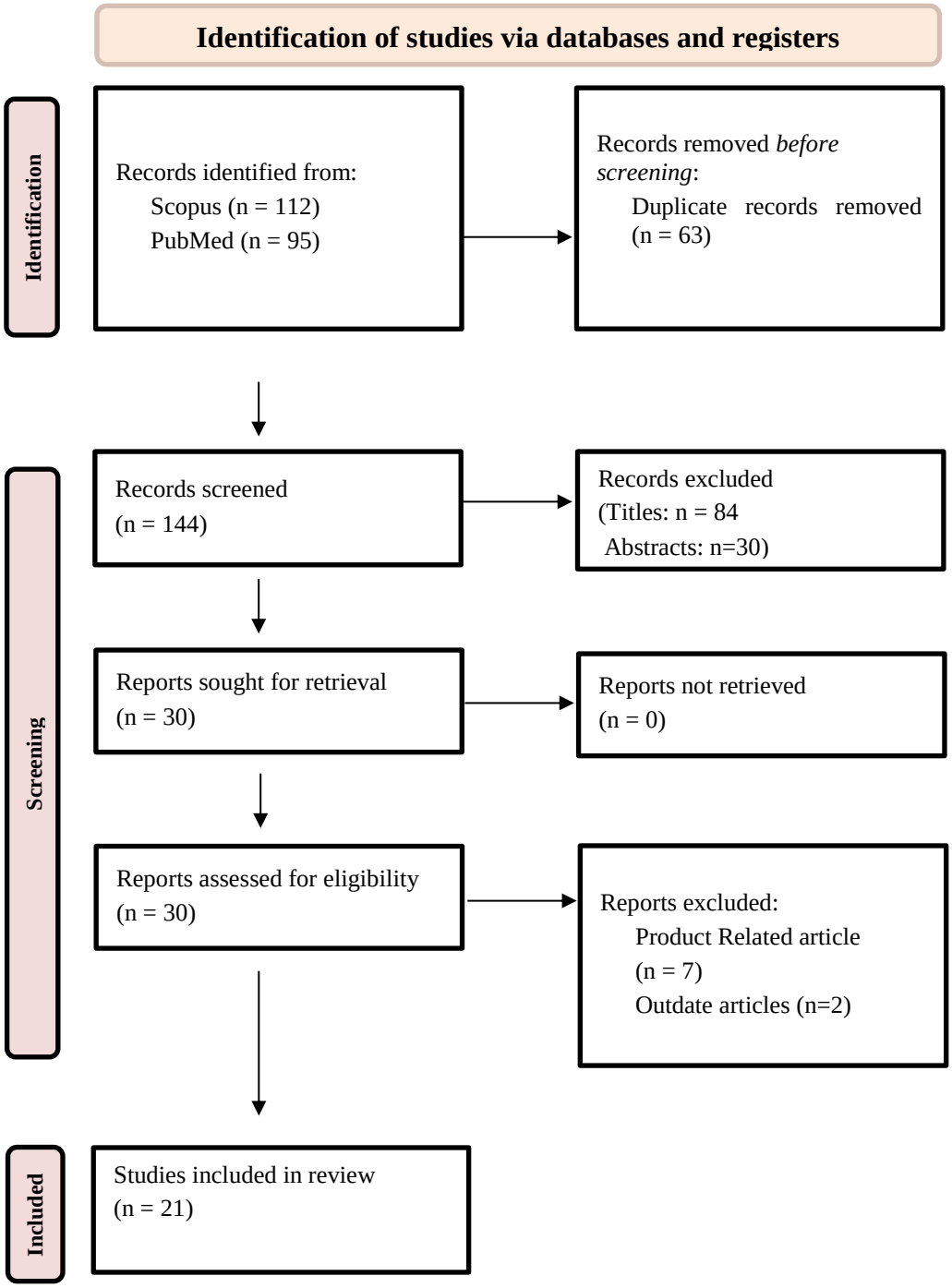


Figure 7 - PRISMA 2020 flow diagram for Nutri-Score systematic review

3.2.1 Text Mining on Front-of-Package Labelling Systems topic

To the best of the author's knowledge, the article published in April 2022 named "Analysis of Content, Social Networks, and Sentiment of Front-of-Pack Nutrition Labeling in the European Union on Twitter" is the first manuscript to evaluate the public discourse on

Twitter about FOPL systems in the EU and the only one that completely matches the current research topic.

In this article it was analysed the tweet content, sentiment, and a mapping of the network characteristics in order to understand not only how much people speak about such topics but also to evaluate the sentiment polarity of tweets.

The Twitter Application Programming Interface (API) V2.0 was used to scrap all tweet data independently of the time when the tweet was published or the language of it. The filter applied was related to the origin of those tweets – users should be from the EU (the UK included).

The search terms used included not only the Nutri-Score FOPL system but also other nutrition labels such as “keyhole”, “heart symbol”, “traffic light”, “nutrition labelling battery”, “healthy choices tick”, “healthy leaving guarantee mark” and the Croatian one “zivjeti zdravo”. In addition to the FOPL systems, it was also applied to the search criteria words related to “Food labelling”, “FOP labelling”, “front/of/package labelling” and “nutrition labelling”.

The data transformations done by the authors were translation into the English language, using Google Translator software, as the tweets retrieved were composed of eight different languages.

In what regards the sentiment analysis results, after applying the automatic sentiment analysis by Nvivo software, the overall sentiment on general topics and discussions towards the political conflicts was negative, while on the other side, the users perceived positive sentiments toward the food industry.

One of the topics concluded in the discussion part was related to the fact that Nutri-Score is at the heart of the debate on Twitter, existing many more tweets related to this FOPL when compared with tweets of the FOPL system in general and other nutrition labelling systems in particular. The tweets encountered ranged from an acceptance of the appliance of a mandatory FOP system and the debate of the right nutrition score value of a specific product (Septia Irawan et al., 2022).

3.2.2 Nutri-Score Countries Evaluation

Cross-Country experimental studies have been made to understand the effect on consumers' accuracy in identifying healthier snack options and liking of the front-of-package nutrition labels, comparing different kinds of it.

While comparing the NutrInform Battery (an Italian nutrition label) with Nutri-Score, in a cross-country study taken in various European countries, the first one showed a greater consistency and effectiveness in comprehensibility, help-to-shop and complexity (Mazzù et al., 2021).

In Portugal, a recent study showed that 40% of all people surveyed do not understand nutritional information on food labels. Despite it, Nutri-Score has revealed as the easiest to understand, showing a clear superiority when compared to all other FOPL systems analysed in the study (Goiana-Da-silva et al., 2021).

Bottari & Mark-Herbert, 2022 did a study where it was collected data from semi-structured interviews (focusing on Italy and Sweden) and strategic documents (such as policy documents) to identify the topics of the debate regarding the development of a single FOPL in the EU. The focus was on answering three research questions, and one of those questions has a special interest in this current research as it tries to identify the key stakeholders (that includes the association of consumers, food industries and of retailers; state agencies, EU Commission, and EU Council) opinions regarding the harmonization of the FOPL.

In this paper, the conclusion was that, although the key stakeholders have the same opinion on the benefit of a unified label on a European level, they disagree on the features that the new labelling system should have, having very contradictory opinions on this topic (Bottari & Mark-Herbert, 2022).

Van den Akker et al., 2022 evidenced that Nutri-Score has a positive impact on the food choice of Dutch consumers. In an experiment done in a Dutch supermarket, 300 participants were asked to make a choice when comparing 3 FOPL schemas, and Nutri-Score was the one that promoted the choice of the healthiest products.

The research done by Egnell et al., 2020 was the first one to investigate the performance of Nutri-Score among Swiss consumers. In this paper, consumers were compared in terms of perception and understanding of five front-of-package label

systems, and the results were that all five FoPLS were favourably perceived, with Nutri-Score standing out from the others.

On a questionnaire done on 21,702 participants in France, four formats of a FOP nutrition label were compared, having the Nutri-Score received more positive responses, having identified also that this system is appealing to subjects that normally have lower adherence to nutritional recommendations (Julia et al., 2017). Likewise, Ducrot et al., 2016 concluded that Nutri-Score is effective in promoting healthier food choices among French consumers. Similarly, when comparing with the British FSA/Ofcom nutrient profiling score, the Nutri-Score had the greatest accuracy in identifying the healthier of two snacks (Hagmann & Siegrist, 2020).

To finalize this systematic review in a European context, it is also important to highlight a cross country made in 12 European countries, where the objective understanding of Nutri-Score was evaluated. This study was done on more than 12,000 subjects and it was concluded that this FOPL system is the most effective label in terms of eliciting consumers to properly classify food accordingly to their nutritional value, not only on the countries as a whole but also in each of the 12 countries [1] (Egnell, Talati, et al., 2020).

Finally, on an international level, Nutri-Score was also studied among consumers. One of them was a study conducted in Uruguay. In this research, it was concluded that the Health Star Rating nutrition label, when compared to Nutri-Score, performed worse based on the opinion of the consumers (Ares et al., 2018).

The other study taken outside Europe was in Morocco, where 814 participants were asked to choose among three food classes with different nutrition profiles and labels and after it, a series of questions were made regarding their preferences and the attractiveness of the labels. Nutri-Score received the highest number of positive answers being ranked as the preferred label by more than half of the subjects (Aguenaou et al., 2021).

In an 18-country analysis, that contained more than 18,000 participants, five FoPL were analysed in terms of effectiveness. Nutri-Score performed best across the total sample on an attraction and aversion level (Pettigrew et al., 2023).

3.3 Text Mining applied to public opinion on healthcare topics

In recent years, there has been an increase in sentiment analysis in healthcare topics, principally explained by the start of the COVID-19 pandemic, as can be seen in figure 8.

While doing a research in Scopus and PubMed, applying the second query defined earlier in the chapter, there were a lot of documents retrieved, and a great part was related to the coronavirus.

Of the 380 results retrieved in PubMed, 157 were related to COVID, while in Scopus there were 325 document results, and 176 were on the coronavirus topic.

The goal of this section is to analyse what are the techniques more used when speaking about healthcare domains, and which ones have a better performance on the sentiment analysis topic. To reach such an objective, it was decided to limit the number of documents to "review" or "systematic review" type.

In the Scopus database, after the type of documents filter was applied, the analysis ended up in 10 articles. However, from these 10 articles, only three were relevant to include in this dissertation, based on the paper topic.

I. Gupta & Joshi, 2021 studied the negation role appliance in Twitter sentiment analysis, doing a review of what has been done until 2021. In healthcare, they have concluded that it is particularly important the consideration of negation while doing sentiment analysis. One of the popular scopus used for negation scope is BioScope (among rule-based approach), NegEx and NegFinder.

Cortis & Davis, 2021 have done a systematic review of the user-generated content related to some opinion expression, also denominated social opinion mining (SOM), analysing documents related to this topic, in a twelve-year period (between 2007 and 2018). A total of 485 published studies were reviewed, focusing principally on the social media platforms used as a data source, tools and techniques used, etc. It was concluded that Twitter is the social media most used in this kind of study. In terms of techniques, deep learning has the highest increase in terms of usage, and should be further explored, namely transformed-based models, e.g., RoBERTa. Sentiment analysis and subjectivity detection are the dimensions of SOM where most of the studies focus. Sarcasm detection

and irony detection still need further research. The sector where opinion mining is more applied is politics, followed by, marketing, finance, technology, and healthcare.

Doing the systematic review, the majority of the lexicon-based studies (using for example VADER) found used one or two lexicons, being SentiWordNet the most commonly used. Approximately 1/3 of the studies found used machine learning-based approaches, whereas only four of these studies used unsupervised machine learning algorithms (namely Continuous Bag-Of-Word and K-Means (KM)). Regarding supervised algorithms, NB and SVM were the most popular ones (Cortis & Davis, 2021).

Gohil et al., 2018 objective was to review the methods used in the sentiment analysis of healthcare tweets and determine which method better fits such topic. A total of twelve papers were reviewed, where half of them used tools specifically produced by the researchers for their analysis – 5 of these studies annotated manually a sample of their data, and used algorithms and machine learning techniques to pursue their analysis (such as NB, K-Nearest Neighbours , SVM). Three papers used open source software – one of them used a tool called package twitteR R; another tool used in another paper was SentiStrenght, which measures the sentiment of short informal messages, and it was compared with other lexical classifiers, and it outperformed; the last open source software used was TextBlob. One of the conclusions was that it exists a large disparity in terms of types of methods used in sentiment analysis applied to healthcare topics on Twitter.

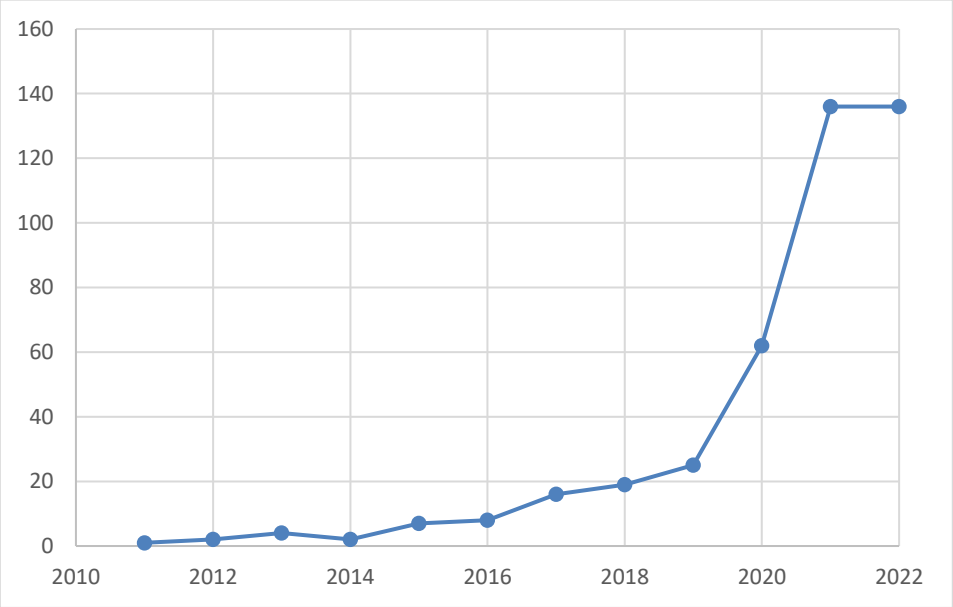


Figure 8 - Number of "Text mining applied to healthcare" related articles

3.4 Major Findings of the Systematic Literature Review

The main purpose of the systematic review was to search for papers directly related to the topic under study - sentiment analysis on the Nutri-Score topic.

As it was verified, it was only found one study that fully corresponds to the topic being studied in this dissertation, and this paper was considered as being part of one of the two categories defined: text mining on front-of-package labelling systems topic.

The other category taken into consideration was the Nutri-score countries evaluation as it was identified a tendency of papers that analyse and compare FoPLS between a group of countries.

As previously stated, in the first category defined present in section 3.2.1, it was only found one article that completely fulfilled the criteria. In table 2 is a summary of the method used and conclusions taken on the Nutri-Score sentiment analysis article.

Table 2 - Method and Conclusions of "Sentiment Analysis" article

Reference	Country	Method	Conclusion
(Septia Irawan et al., 2022)	European Union	Analyse data from Twitter using Twitter API and extract sentiments, social network and content analysis	Nutri-Score is the nutrition label more spoken about on Twitter when compared to other FoPL. The sentiment when speaking about Nutri-Score varies when it is implied from a political perspective (negative) to when talking about the food industry as a whole (positive).

The next category described in section 3.2.2 was Nutri-Score countries evaluation. This category was defined because even though it does not fit completely the purpose of this dissertation it is a fundamental resource to do the comparison conclusions represented by the RQ2.

The main study method used in the papers of this category was questionnaires. These questionnaires aimed to identify what was the population's opinion on the Nutri-Score topic as described in table 3.

Table 3 - Method, Objective and Conclusion of Nutri-Score Countries articles

Reference	Country	Method	Objective	Conclusion
(Mazzù et al., 2021)	France, Germany, Greece, Italy, Portugal, Romania and Spain	Online questionnaire	Compare NutrInform Battery label with Nutri-Score	NutrInform Battery performed better in terms of subjective comprehension and liking.
(Goiana-Da-silva et al., 2021)	Portugal	Online questionnaire	Understand the effectiveness of the usage of FoPL systems in identifying healthy products	Nutri-Score was the easiest to understand among the other labels analysed.
(Bottari & Mark-Herbert, 2022)	Itlay and Sweden	Semi-structured interviews and strategic documents	Find debate topics regarding the development of a single FoPL in the EU	Key stakeholders have a positive sentiment in what regards a unified label on a European level. They have a contradictory opinion on what the new labelling system should have.
(van den Akker et al., 2022)	Netherlands	Supermarket experiment	Compare different FoPLS	Nutri-Score promoted the choice of the healthiest products.
(Egnell, Galan, et al., 2020)	Switzerland	Supermarket experiment	Compare Nutri-score with other FoPL on products	Nutri-Score was the most efficient in informing Swiss consumers.
(Hagmann & Siegrist, 2020)	Switzerland	Online experiment	Compare two snacks using British FSA with Nutri-Score	Nutri-Score had the greatest accuracy.
(Julia et al., 2017)	France	Questionnaire	Compare four formats of FoPL	More positive responses on the Nutri-Score label.
[1] (Egnell, Talati, et al., 2020)	Bulgaria, Denmark, France, Germany, Spain, UK, Belgium, Italy, Netherlands, Poland, Portugal, Switzerland	Online survey	Compare the effectiveness of five FoPLS (including nutri-score) on consumer understanding	Nutri-Score was the most effective
(Ares et al., 2018)	Uruguay	Experimental task	Compare three FoPL systems	Consumers had a better final opinion of Nutri-Score

(Aguenaou et al., 2021)	Morocco	Online survey	Compare three food classes with different FoPL	Nutri-Score was ranked as the preferred nutrition label
(Pettigrew et al., 2023)	18 countries	Online survey	Compare five FoPL	Nutri-score had a better performance

On the overall, Nutri-Score was considered among consumers as a good FoPL system, even when compared with others nutrition labels.

Finally, in section 3.3, where it was highlighted text mining techniques applied to healthcare, multiple approaches were used, and some of them, for example deep learning, have been increasing the usability by researchers.

In table 4 it is presented a summary of the techniques used in this section, where it was quite clear the raise of text mining applied to healthcare particularly in the last two years due to the covid-19 pandemic. Although, it was decided not to focus specifically on each of the articles, instead of it, use systematic reviews.

Table 4 - Method, Techniques and Conclusions on Text Mining Applied to Healthcare articles

Reference	**Method**	**Techniques**	**Conclusion**
I. Gupta & Joshi, 2021	Systematic Review	Negation handling: BioScope, NegEx and NegFinder	In healthcare, it is particularly important the consideration of negation while doing sentiment analysis.
(Cortis & Davis, 2021)	Systematic Review	NB, KNN SentiWordNet Continuous Bag-Of-Word, VADER	Twitter is the social media most used for social opinion mining; Deep learning is being increasingly used.
(Gohil et al., 2018)	Review	NB, KNN, SVM Open Source Software SentiStrenght, TextBlob	It exists a large disparity in terms of the types of methods used

CHAPTER 4

Methodology

The following work, as already mentioned, was developed having as base a six stages methodology called CRISP-DM. The six phases are divided into Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation and Deployment presented in figure 9.

CRISP-DM has the ability to combine the business questions of a data science problem and also the technical development of it. In the end, to validate if the project was successfully implemented, it should be evaluated if those business questions were answered.

For this specific project, it was made an adoption of the CRISP-DM methodology, renaming/removing some of the steps, to make it, as much as possible, adapted to the current analysis. This adoption is presented in figure 9.

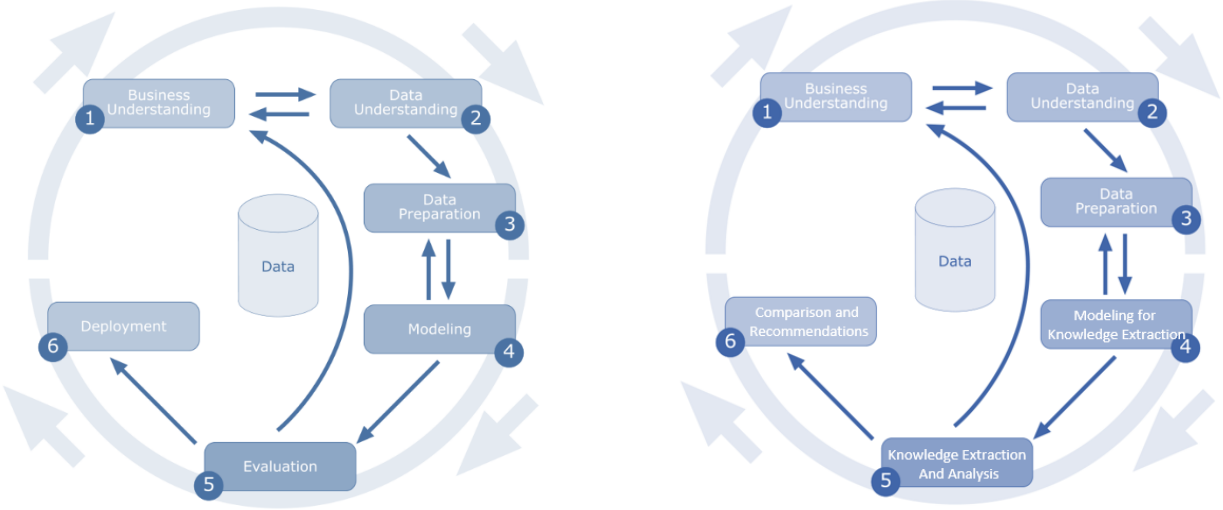


Figure 9 – (left) CRISP-DM Model[6] and (right) Adapted CRISP-DM Model

4.1 Business Understanding

CRISP-DM stands out from the variety of methodologies used for directing data mining projects because it focuses first on the definition of a project's business objectives. The developers of this methodology defend that the first stage of a data science project should be focused on understanding the project objectives and requirements – identifying exactly to which questions that project will answer, what will be the impact of it, and then, after this, convert the knowledge gained into a data mining problem definition (Wirth & Hipp, 2000).

In this stage, besides defining the objectives and business questions the goal is also to describe which kind of data was needed to address those questions and what were the technologies and tools used.

What is the desired output of the project?

For this work, the definition of the problem and its objectives were already defined in the Literature Review and Related Work chapter, where the main keywords related to Nutri-Score were presented, and in what this specific work will contribute – the major problem identified was that FoPLS are being implemented all around the world, but first, there is a need to decide which nutrition label better fit the current needs, and second, after deciding which one will be chosen, there is a need to standardize it on all countries.

The main goal of the project can be defined as the need to have more ways of identifying the opinion of the population in what regards the Nutri-Score topic. To recap, the research questions defined in chapter 1 are presented below. The answer to these questions can be used as a way to measure if the objectives were accomplished or not:

*(**RQ1**) What is the trend when speaking about Nutri Score on the Twitter social network –positive or negative?*

*(**RQ2**) From which countries are the users with positive/negative thoughts related to Nutri-Score?*

*(**RQ3**) Is there a difference between 2017 and 2022 in terms of sentiment polarity, when speaking about Nutri-Score?*

To try to answer all of the research questions above, as this is a data science project the main resource of it should have a minimum level of quality – the data. Even though it was verified in the Chapter 3 that the text on Twitter (and on social media as a whole) can have many symbols to be removed and the kind of language used can be difficult to analyse, it must be guaranteed on a first stage of the project development that the quality of data is good enough to continue the analysis.

The technologies that were used during the project development was the Jupyter Notebook (version 6.3.0) for all the data preparation and exploration part, using Python as a programming language. For the visualizations, it was used not only some Python (version 3.8.8) packages (namely seaborn and matplot libraries) but also Power BI (version 2.11) visuals (to turn the visualization of the results more dynamic and easier to analyse as needed).

4.2 Data Understanding

The second stage of the process is understanding which type of data is needed compared to what exists – and where it is going to be collected. This stage is directly connected to business understanding because it is at this stage that it is going to be clarified if the business questions can be answered, and if the objectives defined in the first stage can be completed (Dongo et al., 2020a).

4.2.1 Information Source Understanding and Search Terms

As it was decided, the input method to obtain data was through one of the most popular information sources available on the Web - the social network Twitter. This social network also called “microblogging service”, attracts several researchers towards it because - but not limited to – the characteristics of tweets: the users tweet about a variety of topics, from events and personalities to day-to-day activity and health (Alam & Yao, 2019; Dongo et al., 2020a).

The main source of information is the tweet – a message published on Twitter – that can contain not only the message itself but also another type of information marked in figure 10. To sum up, a tweet, besides its content has a profile name (1); hashtag (2) it is

not mandatory, but in case the user wants to sign a message to be part of a specific topic is using the # sign before the specific word/combination of words; in case there is a need to mention another user to capture its attention, the @*mention* mark (3) can be used.

In terms of ways to interact with a tweet there is the possibility of answering/commenting (4), share a specific tweet on the user's own page - referred to as "retweet" (5) -, like a specific tweet (6) and lastly, share a tweet outside Twitter (7).

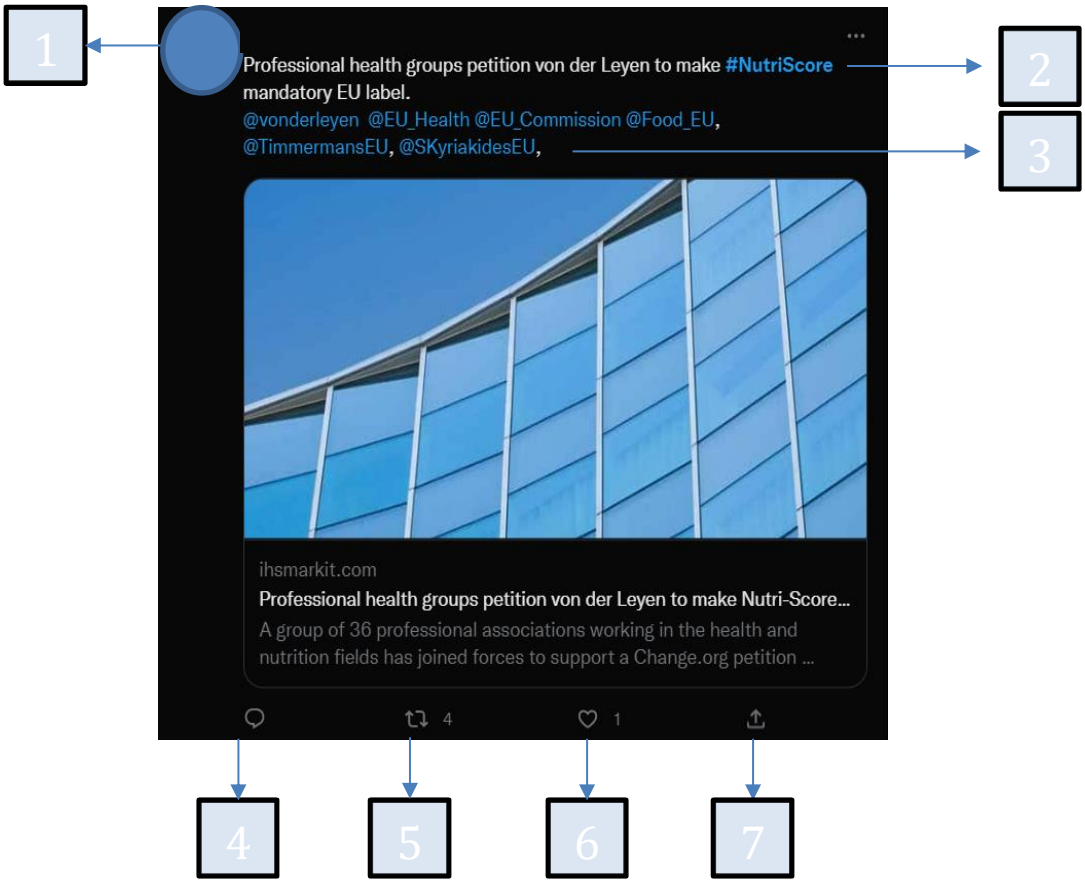


Figure 10 - Information Source System understanding (Twitter)

The very first step was to define what was going to be the data collection criteria, namely in what regards the period to analyse.

Nutri-Score was released at the end of 2017, so to conduct this study the tweets exported had this timeframe in consideration – from January 1, 2017, until July 1, 2022.

Moreover, it is important to outline what were the searched keywords used – there are other ways of mentioning this FOPL system, as it was understood during the literature

review, but to minimize the data noise, it was decided to use only two keywords: “Nutri-Score” and “Nutriscore”. Due to this, these two keywords were used for the search terms on the web scraping.

4.2.2 Data Extraction method

What was the web scrapping method used to collect data?

Dewi et al., 2019 state that web scrapping is a process that can change unstructured data collected from the internet into structured data that can be stored and validated in a database.

In this specific case, the data that was collected was in a structured format (the tweets are only textual data, and any type of audio/video is out of the scope of this project).

There are several ways to extract data from Twitter:

Twitter API – It is Twitter’s official way to extract Twitter data, allowing not only to read but also write data. Although using this API guarantees compliance with Twitter’s terms of service, it has the disadvantage of not being as flexible or efficient as using scraping services as the authors in the “Web Scraping versus Twitter API: A Comparison for a Credibility Analysis” study published in 2021 concluded. In addition to that, the number of tweets that can be extracted is also limited in terms of time – only the tweets from the last two weeks can be extracted – which leads to a tremendous limitation (Dongo et al., 2020b).

Tweepy – This Python library accesses the Twitter API and offers a range of functionalities, e.g., post, retrieve, and engage with Tweets, timelines, post and receive direct messages, retrieve trends, etc. In terms of limitations, this standard API only allows to retrieve tweets up to seven days ago, and in terms of number of tweets it can only extract up to 18,000 tweets. Another important point to have in mind while using this API is that it requires Twitter developer credentials. The process to achieve it is simple but it is always important to have in mind that it can take more than a couple of days to receive the necessary permissions (*Tweepy Documentation — Tweepy 4.12.1 Documentation*, 2022)

SNScrape – After the GetOldTweets3 library became obsolete, SNScrape was one of the Python libraries that stood out in order to scrap tweets without the restrictions of

Tweepy. From the most well-known and most used scrapping methods, SNScrape is the one that allows the extraction of more attributes from each tweet as can be seen in table 5 and it has the advantage of allowing to get historical data (older than two weeks ago) (*GitHub - JustAnotherArchivist/Snscrape: A Social Networking Service Scraper in Python*, 2022)

SNScrape was the web scrapping method that was selected for this research, allowing to have an important resource for the analysis: history, location information and scrapping time.

Using the search terms and period already defined, it was extracted 136,808 tweets. While starting the analysis of the data extracted some difficulties were encountered regarding the velocity of running the code. To try to surpass it, the analysis was divided into data frames of ~30,000 tweets and then exported to a CSV file, to make it easier to do the further analysis. This transformation allowed an increase in terms of processing time because the web scrapping task was limited to a one-time task. The data flow of the exporting process is defined in figure 11 and an example of the tweets extracted are presented in figure 12 and 13.

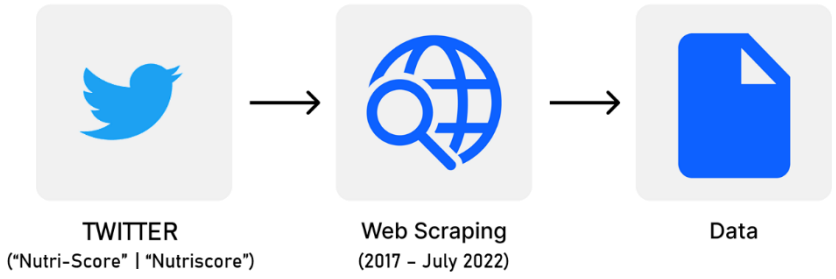


Figure 11 - Exporting process data flow

Example of "Nutri-Score" tweet:

"Our fight at the @Europarl_EN is to try to leave out of the Nutri-Score labelling the ingredients like olive oil" #Nutriscore

@AdrianVL1982 via @euronews 👉 https://t.co/sBw4vS2RI1 https://t.co/icRochE46z

Figure 12 - Example of "Nutri-Score" tweet

Example of "Nutriscore" tweet:

"In Europe, reformulation efforts will see all Kellogg children's cereals achieve a #NutriScore A or B"
On top of providing at-a-👀 info to consumers, evaluative nutritional labels incentivise food makers to improve recipes to achieve better scores 👇

Figure 13 - Example of "Nutriscore" tweet

In regards to the type of data extracted, the columns received from the web scraping are presented in table 5 and it is also important to denote that each line of the dataset corresponds to one tweet written by one user.

Table 5 – SNScrape columns retrieved

Field Name	Description
url	A link that points to the location of the tweet
date	Date when the tweet was created
content	Main content (text) of the tweet
Id	Unique identifier of the tweet
user	The column containing a range of user information (e.g. username, description, etc.)
replyCount	Number of replies that specific tweet had
retweetCount	Number of retweets that specific tweet had
likeCount	Number of likes that specific tweet had
quoteCount	Number of users that quoted the tweet and replied
conversationId	Sames as tweet id
lang	Generated by an algorithm that retrieved the language of the tweet
source	Where the tweet was posted from
sourceUrl	Where the tweet was posted from URL
sourceLabel	Where the tweet was posted from Label (e.g. Web, iPhone, etc.)
outlinks	Full URL
tcooutlinks	Short URL
media	Any kind of media that was published alongside the tweet
retweetedTweet	ID of the original tweet (in case it is a retweeted tweet)
quotedTweet	ID of the original tweet (in case it is a quotes tweet)
inReplyToTweetId	ID of the original tweet (in case of a reply tweet)
inReplyToUser	ID of the original tweet (in case of a reply tweet)
mentionedUsers	Users that were mentioned in the tweet
coordinates	From where the tweet was published
place	From where the tweet was published
hashtags	Hasthats present in the tweet
cashtags	Cashtags (stock symbols) present in the tweet

After the data extraction and the examination of the data format, the number of rows and columns, field identities, and available features it is important to do the data exploration. It is in this phase that the work related to answering the business questions starts, validating also if the data quality is appropriate enough to answer them.

The first statistical information extracted was related to the number of tweets during the indicated period. The number of total tweets extracted as mentioned was 136,808, and after removing the duplicates, the total to pursue the analysis was 80,496 tweets. It is possible to visualize the data extracted over time on Twitter in figure 14.

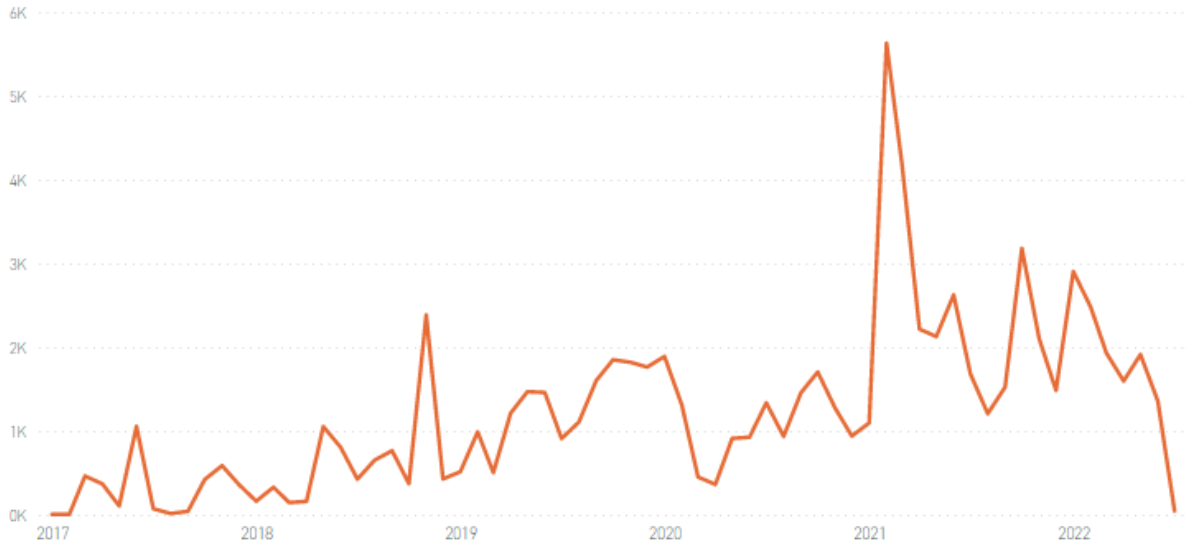


Figure 14 - Number of Nutri-Score tweets over time (2017 to 2022)

Besides the number of tweets, it was also visualized the number of tweets per keyword “Nutriscore”, and “Nutri-score” as presented in figure 15:

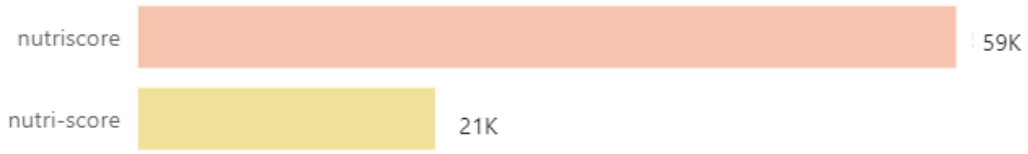


Figure 15 - Number of tweets by keyword

4.3 Data Preparation

The main goal of it is to clean and prepare the data making it in the most uniform way possible for any type of modelling/major analysis work of the project that will be done next.

In this specific case, the step after the data preparation was the sentiment analysis and the answers to the business questions defined earlier.

Having in mind the scope of the current project it was only selected four main columns presented in table 6 (from the 27 attributes that SNScrape retrieves) – alongside the columns that are used it is a valid explanation of why it was important to have those columns:

- **Date:** to understand when the tweet was posted and do some further analysis in order to get some insights on the sentiment variation over the years (RQ3);
- **Content:** where the main analysis was focused, in order to retrieve the sentiment polarity of each tweet (RQ1, RQ2, RQ3);
- **Coordinates:** to answer the question related to the sentiment level grouped by country (RQ2) and to understand if the results from the questionnaires described in the literature review have the same behaviour compared to Twitter data; However, it should be noted that the location information is only shared if the user permits to do so; In this case, this column was mostly empty, and only 3% of the dataset had values filled in this column.
- **Language:** the tweet language might also be important to do the translations, or even to analyse the tweets by country (RQ2);

Table 6 - SNScrape columns extracted

Name	Data type	Example
Date	date	2018-12-29 19:19:10+00:00
Content	String	@7detable Espérons que lorsque de plus en plus de pays l'auront adopté, le Nutri-Score deviendra obligatoire en Europe (et pourquoi pas ailleurs !).
Coordinates	String	*NULL*
Lan	String	fr

Upon the data extraction, it was identified several languages in the tweets. Only 9% of tweets were in English (as it can be seen in figure 16 identified as “en” language code), so there was here a long work to be done in terms of data preparation, specifically in translation. One option that could have been chosen was to directly apply sentiment analysis to the dataset based on the different languages, however, this option was not chosen because not many languages have so many resources available for text mining and specifically to sentiment analysis as the English language has (Cortis & Davis, 2021).

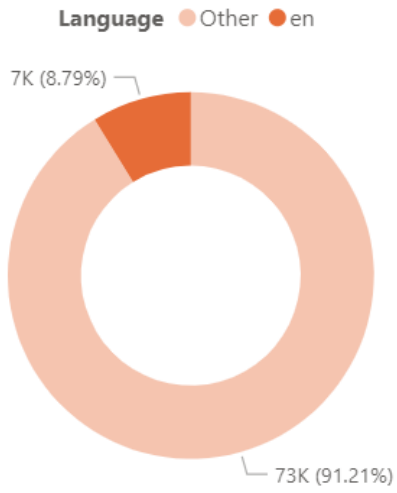


Figure 16 - Number of English tweets extracted

Although it can be considered a difficult computer problem, translating text code gets easier to solve as technological solutions progress. Google, for example, is one of the big companies that share its API to make it possible to use it in coding projects. In order to overcome the translation challenge, it was used the TextBlob library from Python – this library uses Google Translator API to detect a text's language and translate textblobs, sentences and words into other languages.

The goal of using TextBlob is first to identify what is the text language and then, if it is different from English, translate it to this language. An example of this translation is presented in table 7.

Table 7 - Example of a French-English translation

Language	Original Tweet	Translated Tweet
French	Avec 1 adulte sur 2 obèse ou en surpoids, difficile de croire que les européens ne sont pas libres de manger (et même trop)... Le #Nutriscore est un outil utile pour le consommateur qui lui permet de faire un choix plus éclairé. Est-ce pour cela que certains s'y opposent ?	With 1 adult out of 2 obese or overweight, hard to believe that Europeans are not free to eat (and even too much) ... Nutriscore is a useful tool for the consumer who allows him to make a more enlightened choice. Is that why some people oppose it?

One important keynote lies in the fact that it was not possible to do the translation on all of the ~80,000 tweets that exist with the keyword "nutri-score" or "nutriscore" at the

same time, due to the fact that the computer capacity was not sufficient enough for such analysis.

To overcome such a challenge, the translation had to be done segment by segment (having each dataset around 30,000 tweets). This had a huge impact on the time spent during the analysis, but it was the only way found by the researchers to overcome the computer capacity problem.

To present visually the work done in the previous section that resulted in a dataset used for the sentiment analysis described in this part, the data flow present on figure 17.

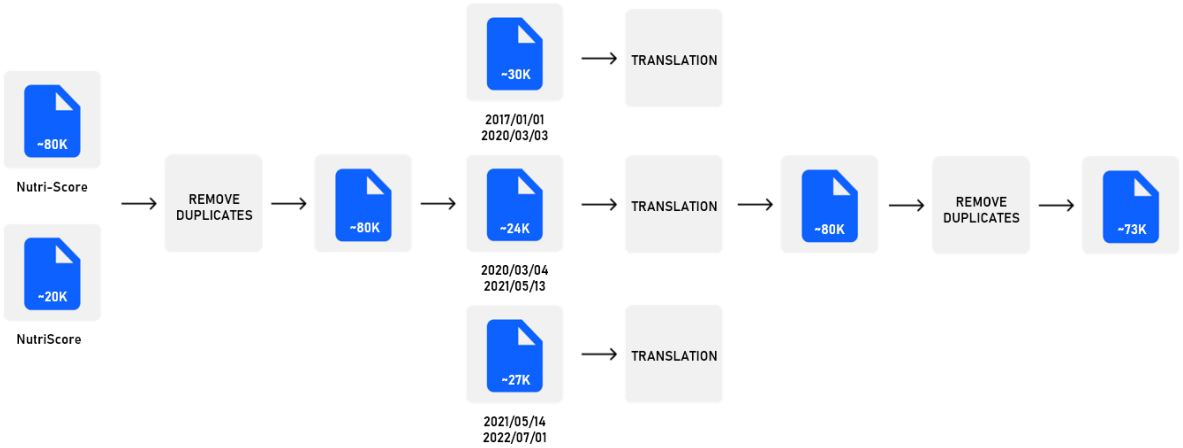


Figure 17 - Translation and duplicates removal data flow

In figure 17, it can be seen how the translation part was carried out (splitting the datasets into smaller subsets). It was also important to guarantee the data didn't have any duplicates doing the removal of it after the merge of the two main datasets and also after the translation part. In the end, there were approximately 73,000 tweets to be analysed.

4.3.1 Data Pre-Processing tasks

The pre-processing part of a text mining project consists of converting data to something a computer can understand and work with. In this specific case, it was not applied any machine learning algorithm, so the data pre-processing part focused more on data cleaning. The data cleaning phase can be described as the task that involves raising the data quality to the level required by the analysis techniques selected (Dongo et al., 2020a).

Different types of transformations can be used in text pre-processing. Some are considered a "must do" independently of the type of problem analysed, such as: special characters removal and lowercasing; some are a "should do" like standardizing identical words e.g. b4 meaning "before". The third type is "task-dependent" and it consists of advanced normalization (e.g. addressing out-of-vocabulary words), stop-word removal, stemming / lemmatization and text enrichment/ augmentation (Dongo et al., 2020a).

Twitter data is considered sparse and noisy. In table 8, it is clarified how the pre-processing should be handled for the different types of texts.

Table 8 -Type of data extracted (adapted from [7])

	Domain-Specific / Noisy Texts	**General / Well-Written Texts**
Lots of data	- Moderate pre-processing - Text enrichment could be helpful	- Light pre-processing - Text enrichment could be helpful, but not critical
Sparse data	- Heavy pre-processing - Text enrichment is important	- Moderate pre-processing - Text enrichment could be helpful

In this sense, it is presented next what were the tasks carried out as part of the data transformation, having in mind that the data cleaning and all of the following steps in the project development were taking into consideration the translated dataset:

1. **Remove special characters and numerical:** A tweet, besides the normal text data, may contain special characters such as emojis, hashtags, URLs, punctuation, numbers, etc.

   Emoticons can be described as a way of expressing a user's feeling, however, in some cases, users use emojis as sarcasm on social media which is a complicated linguistic approach while doing sentiment analysis of that text – the model may consider a text with a positive sentiment if it has a smiley face, but the user might only be using that emoji to represent anger. It exists many articles regarding this

[7] Text Preprocessing for Machine Learning & NLP - Kavita Ganesan, PhD (kavita-ganesan.com)

sarcasm detection on emojis (Chauhan et al., 2022; Subramanian et al., 2019), but that identification is out of the scope of this project.

Regarding hyperlinks (URLs) they are frequently present in tweets, nevertheless, they don't add any additional information to it, at least, for sentiment analysis (Agarwal & Mittal, 2016)

In terms of punctuation, some researchers stratify it makes a significant difference in sentence meaning (Agarwal & Mittal, 2016; Alenzi et al., 2022). An advantage of one of the sentiment analysis models used in this work (Valence Aware Dictionary and sEntiment Reasoner - VADER) is that it takes into consideration the punctuation to understand the emotion of a sentence. In section 4.4.1it will be analysed more in deep this topic.

However, in other papers it is considered as not providing further information about the type of language and can be considered redundant in terms of sentimental analysis, making sense to remove them for the analysis.

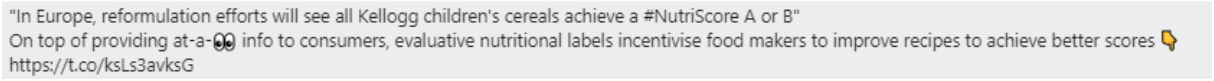


Figure 18 - Example of tweet with emojis

2. **Lowercasing:** It consists of lowercasing all letters in all words. Such transformation is done because whether a word starts with an uppercase most often does not affect its meaning. Another reason is that if on the tweets it exists for example the word "Nutri-Score" and "nutri-score", even though it is the same word, they will be treated as two different words; so, to prevent this duplication all words were converted to their lowercase format.

3. **Tokenization:** The tweet was converted into smaller parts – each sequence of characters was converted into a sequence of tokens using the Natural Language Toolkit (NLTK) library. This process is also beneficial for the following text processing techniques: Part-Of-Speech (POS) Tagging and Lemmatization.

4. **POS Tagging:** It is one of the fundamental techniques when speaking about natural language processes. Using also the NLTK library the POS tagger dictionary

was applied to each token – this consists of assigning morphosyntactic subtypes of fundamental syntactic categories, such as nouns, verbs and adjectives, to each token. If they are correctly mapped to each token it can facilitate the process of identifying their meaning and letting the model know which words (tokens) make part of the same "group". To support the following technique explained next (Lemmatization) the wordnet dictionary was applied. Wordnet can be described as a freely and publicly available lexical database for the English language, and splits the words categorization on these forms: "n" for nouns ['NN', 'NNS', 'NNP', 'NNPS'], "v" for verbs ['VB', 'VBD', 'VBG', 'VBN', 'VBP', 'VBZ'], "a" for adjectives ['JJ', 'JJR'], "r" for adverbs ['RB', 'RBR', 'RBS'] and "s" for satellite adjectives ['JJS'] (Talib et al., 2016).

5. **Stop words removal:** It consists on the removal of any auxiliary verbs, articles, pronouns, prepositions and other common terms that do not add any value to the analysis and so are only taking up space in the dataset. Words like "and" and "the" do not have much meaning, do not influence the polarity of the sentence. To remove stop words, it can be created a list of words that should be considered as words to exclude for the analysis for a specific project. Another way is to use a pre-built library – in this research, the words were discarded using the NLTK pre-built stop words library to facilitate its removal and because it provides good performance as some software authors state (Agarwal & Mittal, 2016).

   In table 9 it is presented an example of using the stop words list on the Nutri-Score tweets dataset.

Table 9 - Example of stop words removal

Original Tweet	Tweet with Stop Words applied
' "25% of the veggie segment in our plant-based portfolio has been renovated with Nutri-Score 🦃. More than 80% of the portfolio now scores **A** or B" EUNutriscore ConsumerDebates	25 veggie segment plant based portfolio renovated nutri-score 80 portfolio scores or b eunutriscore consumerdebates

6. **Lemmatization:** It can be described as the process of converting a word to its base form, grouping together the different inflected forms of a word so they can be analysed as a single item, linking words with similar meanings to one word

which brings context to them. Lemmatization works better when applied after using the POS Tagger pre-processing method because it takes into consideration the part of speech parameter – if this parameter is not supplied, lemmatization would consider that the default was "noun". The WordNetLemmatizer module from the NLTK library was the model chosen for this transformation, the WordNet provides a lexical relation understanding for the connections of the words, making the process of lemmatize more straightforward (Talib et al., 2016).

After each data transformation it was created an additional column with the transformation applied to the translated column, in order to better organize and decide which transformation to use in the models (further analysed in section 4.4.1).

To specify and summarize the columns created with the transformations above explained, it is presented in table 10 the name of those columns and a brief explanation of the transformations applied to each of them.

Table 10 - Columns created based on Pre-Processing transformations

Column Name	Transformation(s) Applied
Clean_Tweet	Remove special characters (URLs, hashtags, mentions)
Clean_Punc_Tweet	Clean_Tweet transformations + Remove punctuation
Clean_Lower_Tweet	Clean_Punc_Tweet transformations + lowercase
Clean_Lower_SW_Tweet	Clean_Lower_Tweet + remove stop words
Clean_SW_Tweet	Clean_Tweet + remove stop words
Clean_Token_Tweet	Clean_Tweet + tokenization
Clean_Lower_SW_Token_Tweet	Clean_Lower_SW_Tweet + tokenization
Clean_SW_Token_Tweet	Clean_SW_Tweet + tokenization
Clean_Token_POS_Tweet	Clean_Token_Tweet + POS Tagging
Clean_Lower_SW_Token_POS_Tweet	Clean_Lower_SW_Tweet + POS Tagging
Clean_SW_Token_Lemm_Tweet	Clean_SW_Token_Tweet + Lemmatize
Clean_Token_POS_Lemm_Tweet	Clean_Token_POS_Tweet + Lemmatize
Clean_Lower_SW_Token_POS_Lemm_Tweet	Clean_Lower_SW_Token_POS_Tweet + Lemmatize

4.4 Modeling for Knowledge Extraction

The final segment of this chapter consists in trying to answer the research questions defined in chapter 1. To answer those questions, a sentiment analysis was done on the tweets previously extracted.

4.4.1 Sentiment Analysis of Nutri-Score tweets

Sentiment analysis, as previously stated, can be described as a computational process which categorizes a piece of text (that in most of the sentiment analysis projects is an opinion) that expresses a positive, negative, or neutral attitude of a writer towards a particular product, event or personality (Alam & Yao, 2019).

In this project it was decided also to analyse this category because as Nutri-Score is not a very well-known topic, it was important also to find what percentage of the tweets were considered neutral, to not lose any kind of data. Furthermore, (Koppel & Schler, 2006) in the "The importance of neutral examples for learning sentiment" article stated that the use of neutral examples in sentiment analysis facilitates better distinction between positive and negative examples

The technique used in this research was a lexicon sentiment analysis (also called "Rule-based sentiment analyser"), however, as seen in the literature review, there are two other techniques that sentiment analysis can be categorized into: machine learning-based sentiment analysis, and hybrid techniques (combining both the lexicon and the machine learning one).

Most of the machine learning approaches while doing sentiment analysis are based on datasets previously divided into test and training – meaning this that, in most of the cases, there was a previous manual annotation done by (presumably) human experts, which identified such opinions as having a positive, negative or neutral polarity. There are pros and cons to these approaches, while doing a machine learning analysis the accuracy can be compared, but at the same time, the manual annotation of the polarities can be subjective to the mental state of the expert and similarly, the bias element cannot be fully ignored (Aljedaani et al., 2022).

In this project, the lexicon-based sentiment analysers used were VADER and TextBlob:

VADER

This lexicon-based sentiment analysis tool was developed by C. J. Hutto & Gilbert, 2014 to solve the problem of analyzing language, symbols, and style of texts in the field of social media.

This rule-based sentiment analyser uses a list of lexical features, in this case 7,500 emotion-related words, to calculate a text sentiment. This list is labelled as positive or negative according to the semantic orientation of the word in order to calculate the text sentiment (C. J. Hutto & Gilbert, 2014)

VADER is specifically attuned to sentiments expressed in social media and, additionally, can be applied to the text without pre-processing techniques applied, i.e., VADER has in consideration punctuation (having the exclamation point in its lexicon), capitalization, conjunctions, abbreviations (such as LOL) and even emoticons (Liew & Lee, 2021).

When compared to other machine learning techniques, VADER can execute the same amount of data in much less time, being very computationally efficient. More than this, unlike other machine learning models, VADER doesn't require any form of training data to produce accurate results, because it is based on a human-validated lexicon (Liew & Lee, 2021).

It retrieves three polarity categories: positive, negative and neutral, as exemplified below, having the total of the three categories give a total of 100% (1):

"The food was great!"
Positive: 99%
Negative:1%
Neutral: 0%

In addition to the polarity, it also calculates the intensity of the emotion, having for this a compound metric as an output that gives the overall sentiment ranging from -1 to 1 (C. J. Hutto & Gilbert, 2014)

- if the sentiment is positive the compound score is greater than or equal to 0.05;
- if neutral the score is in the interval [-0.05, 0.05];
- And lastly, if the sentiment is negative, the compound score is less than or equal to -0.05.

Applying the VADER technique on the Nutri-Score tweets extracted, had into consideration the column with only the hashtags and URL removed (Clean_Tweet column defined on table 10), meaning this that there were no more pre-processing transformations done on the column used for VADER analysis. The reason why was, as specified before, VADER is specifically attuned to sentiment analysis in social media and has in consideration different characters. If those characters were removed (and used another column with more transformations applied) the VADER results might not be so accurate.

It is presented next, in table 11 some examples of the sentiment categorization along with the Negative, Neutral, Positive and Compound columns. There is an example for each polarity and, additionally, an example of a wrong categorization done by VADER (false positive).

Table 11 - Examples of VADER results

Translated Content	Pos	Neu	Neg	Compound	VADER result
We all agree to have common labeling system, but not system that damages our regional products so seriously 🤒 lucky, we have taken forward in our proposal to postpone the nutriscore food labeling	0.14	0.64	0.22	-0.5	Negative
The nutri-score is made to compare food of the same category. Thus not to compare cheese with cereals. 5/8	0.00	1.00	0.00	0.00	Neutral
We all agree on NOVA approach. Nonetheless, within each food category, Nutriscore can help consumers to choose more balanced products (i.e. less saturated fats, sugar, salt)	0.18	0.82	0.00	0.64	Positive
More than 50% of the products marked in the nutriscore are ultra-processed and as such tend to be harmful to health and only useful for generating huge profits for the multinationals that support Nutriscore	0.27	0.74	0.00	0.87	Positive (incorrectly categorized)

As it can be seen in the fourth tweet, the user is speaking about the ultra-processed products, and that Nutri-Score doesn't have such level of processing in consideration when categorizing the score label. VADER classified the text as having a positive sentiment, however, it is wrongly classified.

To analyse the overall dataset, an analysis was done, and the results are presented in table 12, where the percentages of VADER results are shown.

Table 12 - Overall dataset VADER polarity result (%)

Sentiment Polarity	VADER Result
Positive	42.76%
Neutral	33.54%
Negative	23.70%

TextBlob

The second technique used for the sentiment analysis was a lexicon-based model library available in Python (2 and 3) that uses NLTK to perform the tasks, called TextBlob.

TextBlob uses as input a single sentence and has polarity, subjectivity and intensity as output. It is built using essentially two sentiment analysis implementations: NaiveBayesAnalyzer (an NLTK classifier) and PatternAnalyzer (based on a library named pattern) and because it is a lexicon-based approach the output is based on the semantic orientation and the intensity of each word in the sentence (Loria, 2020).

TextBlob retrieves essentially two columns:

Polarity – It is the output score most examined in sentiment analysis. It categorizes texts as positive, neutral or negative, depending on the polarity score that lies between 1 and -1. This polarity score is based on the words that make part of the sentence analysed.

Subjectivity – quantifies the amount of personal opinion and factual information contained in the text. A score closer to 1 means that the text contains personal opinions rather than factual information. The way that this subjectivity score is calculated is based on the intensity, which determines if a word modifies the next word (Shah, 2020).

If the sentence that it is being analysed does not contain any words that had a polarity in the NLTK set or if TextBlob uses a weighted average sentiment score over all the words in the sample, the library returns exactly 0 for polarity (neutral) and 0 for subjectivity (neither personal opinion nor factual information) (Shah, 2020).

To understand the impact of each transformation on the Nutri-Score data set it was done 3 experiments: one using the column without the mentions, hashtags URL and punctuation (Clean Tweet); another using the cleaned tweets on the first iteration described plus the lowercased tweets and without stop-words (Lowercased Clean Tweets without Stop Words); the last experiment used the tweets with the transformations already described on the last two tests and, additionally, the POS tagger on each token and the tokens lemmatized (Lowercased Clean Tweets without Stop Words, with POS Tagging and Lemma).

In table 13, 14 and 15 it is presented some examples of the tests explained above, where it is used the same tweet to compare the three experiments. The examples shown represent, by order, the comparison of a: positive tweet, neutral, positive, negative, negative, wrong negative classification (on the three experiments), and wrong positive classification (on the three experiments).

Table 13 - Clean Tweet column TextBlob result (examples)

Clean Tweet	Subjectivity	Polarity	TextBlob Result
According tosciensano study Belgians think Nutri Score finds the best system to assess the nutritional value of packaged foods But how should you read and understand Nutri Score and other front of Pack labels	0.34	0.44	Positive (v)
"The interest of nutritional logo and the effectiveness of the Nutri Score logo 5c are demonstrated for the first time**"**	0.33	0.25	Positive (x)
"Hello France issued recent press release December 17 2021 reinforced support for the Nutri Score system who is already counting over 870 French healthcare professionals"	0.00	0.00	Neutral (x)
Nesquik is 74 sugar but it has NutriScore This is caused by the prepare formula suggest by Nestle 13 5g of product 200ml semi skimmed milk although real consumers add more product usually	0.35	0.15	Positive (x)
"Look the Nutri Score system is fundamentally flawed It is confusing misleading and often leads consumers to make poor buying decisions Its also gift for the processed food industries who can profitably tweak their product ranges to fit outdated criteria**"**	0.51	-0.24	Negative (v)
"Nutriscore does not attack any product but gives information on the nutritional quality of food It indicates that classified as it cold cuts and Italian or Swiss cheeses French Spanish can be consumed but in reasonable quantities frequencies**"**	0.32	-0.08	Negative (x)
There is an error in olive oil that Nutriscore classifies isand notas the image shows widely used by lobbies Olive oil is classified C the best score of all added oils and fats	0.35	0.45	Positive (x)

Using this first experiment presented in table 13, the overall dataset had a 43.19% of positive tweets, 41.31% of neutral ones and 15.50% of negative.

Table 14 - Lowercased Clean Tweet without Stop Words column TextBlob result (examples)

Lowercased Clean Tweets without Stop Words	Subjectivity	Polarity	TextBlob Result
according tosciensano study belgians think nutri score finds best system assess nutritional value packaged foods read understand nutri score front pack labels	0.30	1.00	Positive (v)
interest nutritional logos effectiveness nutri score logo 5c demonstrated	0.00	0.00	Neutral (v)
hello france issued recent press release december 17 2021 reinforced support nutri score system already counting 870 french healthcare professionals	0.00	0.00	Neutral (x)
nesquik 74 sugar has nutriscore caused prepare formula suggest nestle 13 5g product 200ml semi skimmed milk although real consumers add product usually	0.28	-0.03	Negative (v)
look nutri score system fundamentally flawed confusing misleading often leads consumers make poor buying decisions also gift processed food industries profitably tweak product ranges fit outdated criteria	0.51	-0.24	Negative (v)
“Nutriscore does not attack any product but gives information on the nutritional quality of food It indicates that classified as it cold cuts and Italian or Swiss cheeses French Spanish can be consumed but in reasonable quantities frequencies”	0.32	-0.08	Negative (x)
error olive oil nutriscore classifies isand notas image shows widely used lobbies olive oil classified c best score added oils fats	0.35	0.45	Positive (x)

After applying the TextBlob sentiment analysis on the Lowercased Clean Tweets without Stop Words column, the overall dataset results were: 40.24% of positive tweets, 44.58% of neutral ones and 15.17% of negative.

Table 15 - Lowercased Clean Tweet without Stop Words, with POS Tagging and Lemma column TextBlob result (examples)

Lowercased Clean Tweets without Stop Words, with POS Tagging and Lemma	Subjectivity	Polarity	TextBlob Result
accord tosciensano study belgian think nutri score find best system assess nutritional value package food read understand nutri score front pack labels	0.30	1.00	Positive (v)
interest nutritional logos effectiveness nutri score logo 5c demonstrated	0.00	0.00	Neutral (v)
hello france issued recent press release december 17 2021 reinforce support nutri score system already count 870 french healthcare professional	0.05	0.05	Positive (v)
nesquik 74 sugar has nutriscore cause prepare formula suggest nestle 13 5g product 200ml semi skim milk although real consumer add product usually	0.28	-0.03	Negative (v)
look nutri score system fundamentally flawed confusing misleading often leads consumers make poor buying decisions also gift processed food industries profitably tweak product ranges fit outdated criteria	0.51	-0.24	Negative (v)
“Nutriscore does not attack any product but gives information on the nutritional quality of food It indicates that classified as it cold cuts and Italian or Swiss cheeses French Spanish can be consumed but in reasonable quantities frequencies”	0.32	-0.08	Negative (x)
error olive oil nutriscore classifies isand notas image show widely use lobby olive oil classify c best score add oil fat	0.35	0.45	Positive (x)

Finally, in the last iteration tested, using Lowercased Clean Tweets without Stop Words, with POS Tagging and Lemma the results were: 38.74% of positive tweets, 44.65% of neutral ones and 16.60% of negative.

4.4.2 Sentiment Analysis Methods Comparison

In terms of results on the Nutri-Score dataset, when comparing both methods (TextBlob with the 3rd experiment and VADER without pre-processing) the number of neutral sentiment sentences identified was almost a half of the total dataset in TextBlob whereas in VADER although there is more positive tweets the number of neutral ones is also considerable, as presented in figure 19.



Figure 19 – (left) Text Blob vs (right) VADER results

The difference on the two lexical methods can be explained by the lexicon dictionary used behind each one. In table 16 it is being shown a comparison of the two methods in one of the cases where the result was different (two examples of a wrongly classification in VADER, and two examples of wrongly classification in TextBlob) along with the words present in each of the lexicons, that explain the polarity result.

Table 16 - Comparison of TextBlob result with VADER result

Original Tweet	TextBlob Dictionary[8]	TextBlob Result	VADER Dictionary[9]	VADER Result
!! Defend the Nutriscore logo against lobbies who try to distort it and try to prevent it from becoming compulsory in Europe!! 🖉 to make it compulsory in 2023 and for real nutritional transparency, sign the petition:	Real (0.2)	Positive	Defend (0.4) Lobby (0.1) Distort (-1.3) Prevent (0.1)	Negative (wrong)
🕵France might stand back from the Nutriscore label. According to Italy's agriculture minister Patuanelli, Draghi and Macron discussed the topic "but it seems quite clear that even France is backing away from that unhealthy idea of colouring food and labelling it good or bad."	Clear (0.1) Unhealthy (-0.4) Good (0.7) Bad (-0.7)	Negative	Backs (-0.2) Clear (1.6) Unhealthy (-2.4) Ideal (2.4) Good (2) Bad (-2.5)	Positive (wrong)
🔎The Nutri-Score helps us obtain information on the global nutritional quality of packaged products. Solve all your doubts with the didactic chips offered to us 👇👇 Eroskiconsumer Hispacoop Nutriscore	Didactic (-0.5)	Negative (wrong)	Help (1.7) solve (0.8) doubt (-1.5)	Positive
🕵nutri score destroys the healthy image of oil from Oliva, matching it to Colza oil, given the passivity of Spain ES The collision of consumer associations	Destroy (-0.2) Healthy (0.5)	Positive (wrong)	Destroys (-2.6) Healthy (1.7) Passive (0.8) Collision (-1.5)	Negative

[8] (*TextBlob: Simplified Text Processing — TextBlob 0.16.0 Documentation*, 2022)

[9] (C. Hutto, 2022)

4.4.3. Sentiment Analysis by Country

With the intention of pursuing an analysis of the sentiment demonstrated on tweets regarding Nutri-Score in each of the users' countries, two options could be taken.

A primary solution found was to use the "Coordinates" column (that has information from where the user is writing) to do the further analysis. This solution would enrich the analysis since a map-like graph could be developed showing where each of the tweets had been posted, and furthermore identify in each of the zones the sentiment polarity of the tweets. However, as verified in the Data Understanding section, this column is 97% empty, and using it could lead to a solution that should not be extrapolated.

The second solution was based on the "Lang" column of the dataset – this column, as it could be seen in the Data Pre-Processing section, was used for the translation's transformations, and it was also one of the options found to solve the missing values in the "Coordinates" column, as it did not exist any empty value on this column.

Using this second middle-stage solution it was possible to get the country from where the user might be physically (or might be from) – although this solution also has its inconsistencies: there are, for example, 29 countries that use French as their official language, and when the column "Lang" identifies a French tweet (filling the column with the letters FR) that tweet may be assigned to France, Belgium, Canada, etc. , i.e., all the French speaking countries.

This inconsistency of the second solution can also be minimized because, as already noted in the Related Work chapter, the Nutri-Score label is at the centre of the debate in Europe mostly, which reduces a lot the number of countries where the tweets might be coming from.

In order to prevent the assumption of a French tweet (for example) to just only one country, the analysis was taken calling the category in analysis not "country" but instead of it "region group" that may be explained as being a region where some specific language is spoken. For example, for French tweets it was assigned the group of countries: France, Belgium, Luxembourg and Switzerland; for English tweets it was assigned United Kingdom and Ireland. The rest of region groups may be better analysed when looking into the visualizations built in the next chapter.

Neither of the solutions are 100% consistent, although it was decided that the second one could retrieve more data for further analysis.

The visuals that were generated are presented in the next chapter, but the approach taken was, based on the result of the polarity for each tweet, to group them by country (doing a transformation to switch the “Lang” column to a country group name one) and then, based on this additional column created, display the sentiment analysis by country.

CHAPTER 5

Knowledge Extraction and Analysis

In Chapter 3 (Related Work) it was presented some papers that analysed, via surveys and questionnaires, the population's opinion regarding Nutri-Score.

This chapter aims to bring together some visualizations done using the Power BI platform, analysing not also the results retrieved on this current research (Nutri-Score tweets sentiment analysis) but also comparing the results on a country level with the outcomes of the questionnaires.

There were two main data sources of these visualizations - one was the CSV file previously transformed, being used the final version of it, with some of the data pre-processing transformations described in Chapter 4 applied; the other one was the one that resulted from the transformations applied in the sentiment analysis part.

The first analysis presented is related to the total number of tweets (before the translation), in figure 20. With this graphic, it is possible not only to see the number of tweets but also the tendency of it by year, being very clear a spike in the number of tweets.

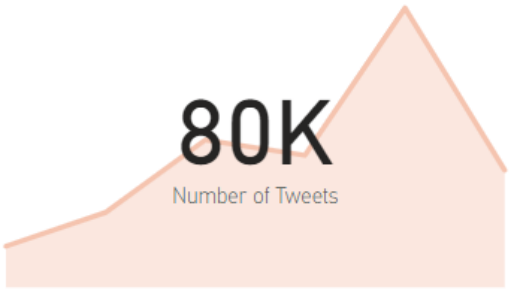


Figure 20 - Total Number of tweets and year analysis

Studying more in dept the number of tweets over time, another analysis that was done was regarding the number of tweets per month as presented in figure 21.

As it can be seen, it was at the beginning of 2021 (February) that a spike in tweets occurred. At the end of 2019 and the beginning of 2020, there was a marked decrease. In

general, there is not a consistent decrease/increase in the number of Nutri-Score-related tweets over the years.

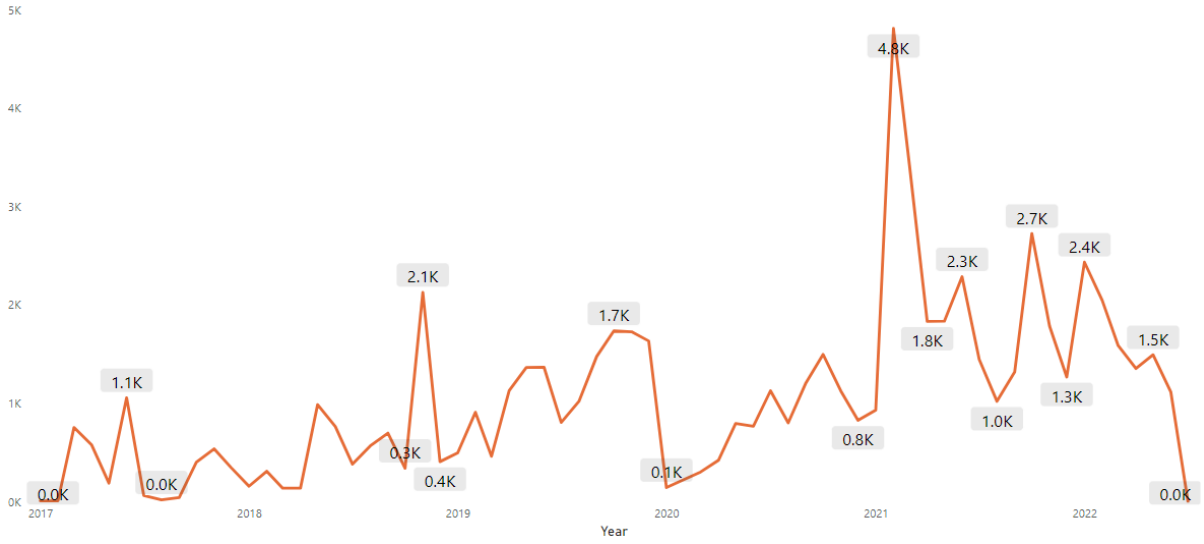


Figure 21 - Number of tweets per month

To try to find an explanation of the number of tweets, and why there was a spike in 2021, it was essentially made two analyses: the first one was to try to get the number of news published on Nutri-Score topic over the years; the second one is directly related with the Literature Review and the Related Work chapter, analysing the number of papers published by Year/Month.

Applying the first experiment, it was done a search in Google News, using the search query “Nutri-Score news” to get the news published on this topic. The number of results retrieved when doing this search was 332,000 (on 16th November 2022) as it can be seen in figure 22.

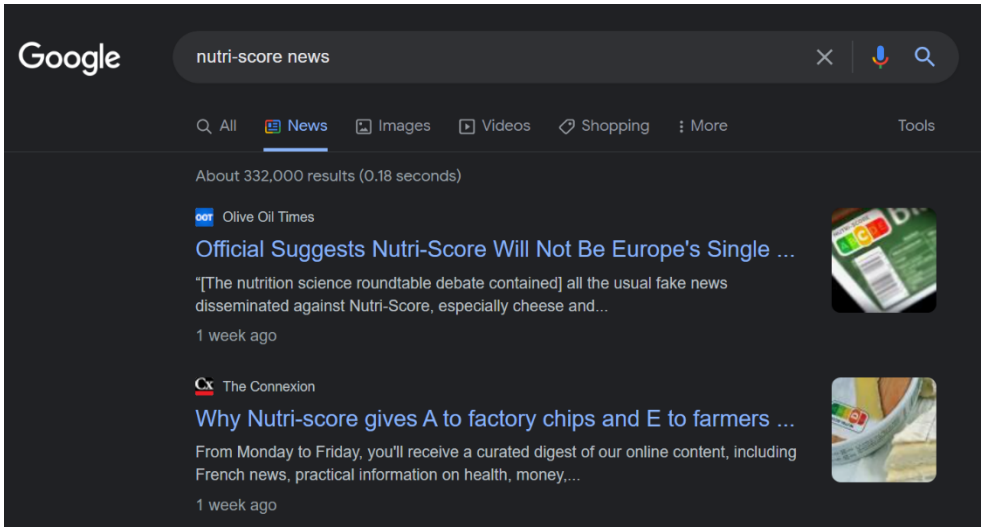


Figure 22 - Google search results of Nutri-Score news

To do the scrapping, it was used Power Automate (an automation tool from Microsoft), and in the end it was extracted 271 news. However, this is only a small sample of the real number of news related to Nutri-Score as it could be seen in figure 22. The reason why it was only possible to extract 271 news was because Google has an internal algorithm that ensures that only the most relevant results are retrieved, in this case, only 30 pages of results were considered relevant (Sedigh & Roudaki, 2003).

The main goal of this experiment was to try to understand if there was a correlation between the number of news of Nutri-Score and the number of tweets, to try to explain the spike of tweets in 2021.

The number of news analysed over time is shown in figure 23, whereas it is also important to denote that not only the number of news retrieved are a sample but they are also not sorted by date. As the number of news retrieved is small in comparison with the real number of news, the analysis becomes impractical.

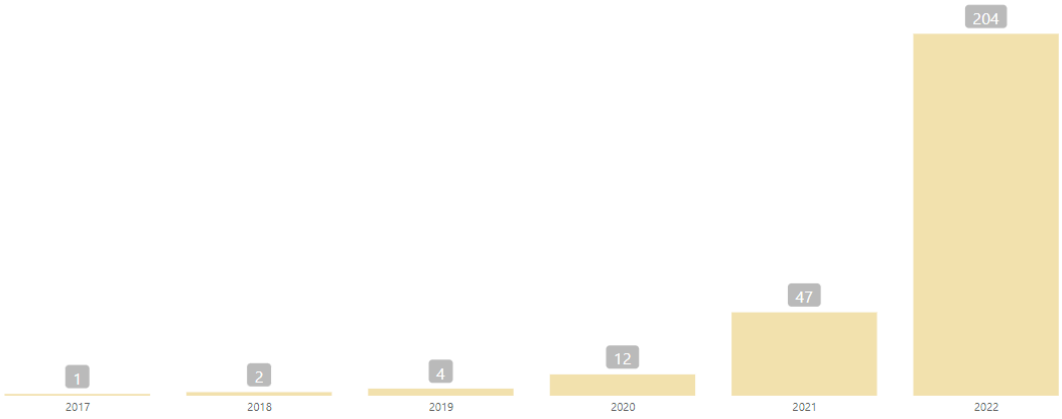


Figure 23 - Number of Nutri-Score news over time (years)

The second analysis done, as already mentioned, was using the number of articles published. It was used both databases already described in chapter 3, although, in opposition to what was done on that chapter, the search query did not have any filters and only included the words “Nutri-score” OR “Nutriscore”. In figure 24, it is presented both results from Scopus and PubMed.

By analysing the number of papers published in figure 24, it is understood an increase all over the years, on both databases. The Nutri-Score has been more and more analysed, and the number of papers reflects that. Nevertheless, it cannot be stated that exists a direct correlation between the number of papers versus the number of tweets.

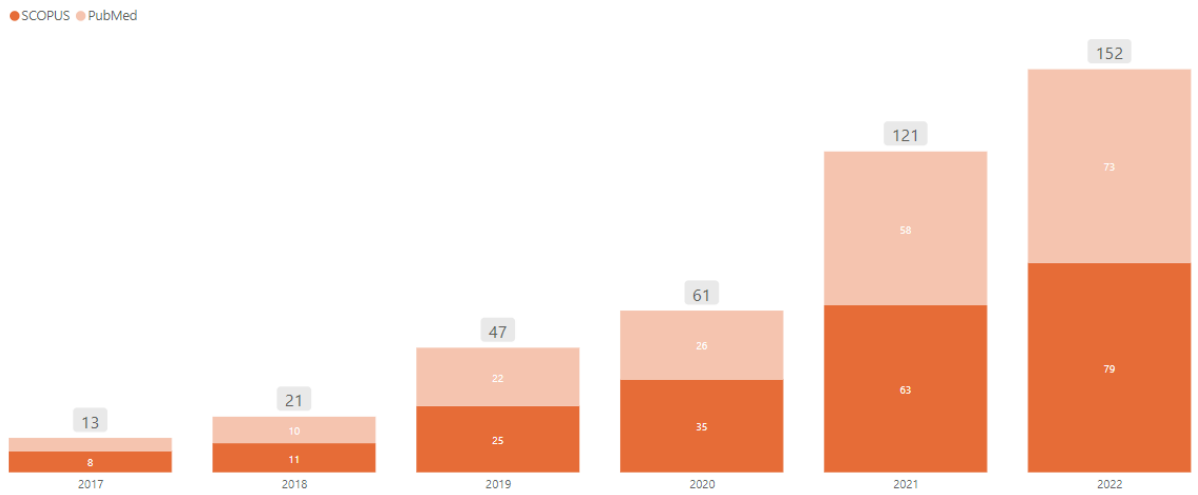


Figure 24 - Number of Nutri-Score papers published in Scopus and PubMed

By using these two experiments it was not possible to find an explanation for why there was a spike in the number of tweets in 2021.

More than this analysis, it was also important to get to know from where the tweets were being published (in order to find any correlation between the time when Nutri-Score was implemented in that specific country and the number of tweets published). To do this, the Lang column transformed to return the countries names was used, having in mind all the limitations of using this column (already described in section 4.4.3).

In terms of temporal analysis, and at a European level, as described in the Literature Review chapter, Nutri-Score was implemented in France in 2017. In November 2018, Spain adopted the Nutri-Score label on their food products (on a voluntary basis) although it was in 2021 that this country officially reinforced the suggestion of adopting this label by organizing its first directive committee meeting. Still in 2018, Belgium started to introduce this labelling format. Switzerland in 2019 chose Nutri-Score as the preferable system to present the FoPL and in 2021 it revised its Nutri-Score website. In 2019 Netherlands introduced it on its products in November, while at the beginning of 2022 implemented it. In 2020 February, Luxembourg introduced it on the front of packages but it was only in 2021 that it adopted a regulation to allow the country's food operators to distribute their products with Nutri-Score on it. Still in the year 2020, was Germany that implemented this nutrition labelling system. Finally, regarding Italy, this country has ruled against the use of Nutri-Score since its implementation in France and, in 2021, the

Italian Competition Authority launched its own review of the system (Peonides et al., 2022).

In figure 25, it is presented the number of tweets in the year when Nutri-Score was implemented for the first time (2017). The major number of tweets are coming from a French-speaking country – which in this case was considered Luxembourg, Belgium and Switzerland as already mentioned), which makes sense since it was in France that Nutri-Score was implemented for the first time.



Figure 25 - Number of tweets per country in 2017

Analysing from 2017 until 2022 in figure 26, French-speaking countries stand out from the other countries in terms of the number of tweets in all years, except in the year 2021, where it was Spain that had the most number of tweets. As it could be seen previously, it was in 2021 that Spain reinforced the suggestion of Nutri-Score as a FoPL system to be applied to products, which can be a reason why in this year there was an increase in the number of tweets from this country.

In 2019, after a French-speaking country and Spain, it was the Netherlands that had the greatest number of tweets, which as it was understood matches the year when Nutri-Score was implemented in this country.

Although with fewer tweets, Italy is also one of the countries that contribute to the number of tweets related to Nutri-Score, especially in 2021.

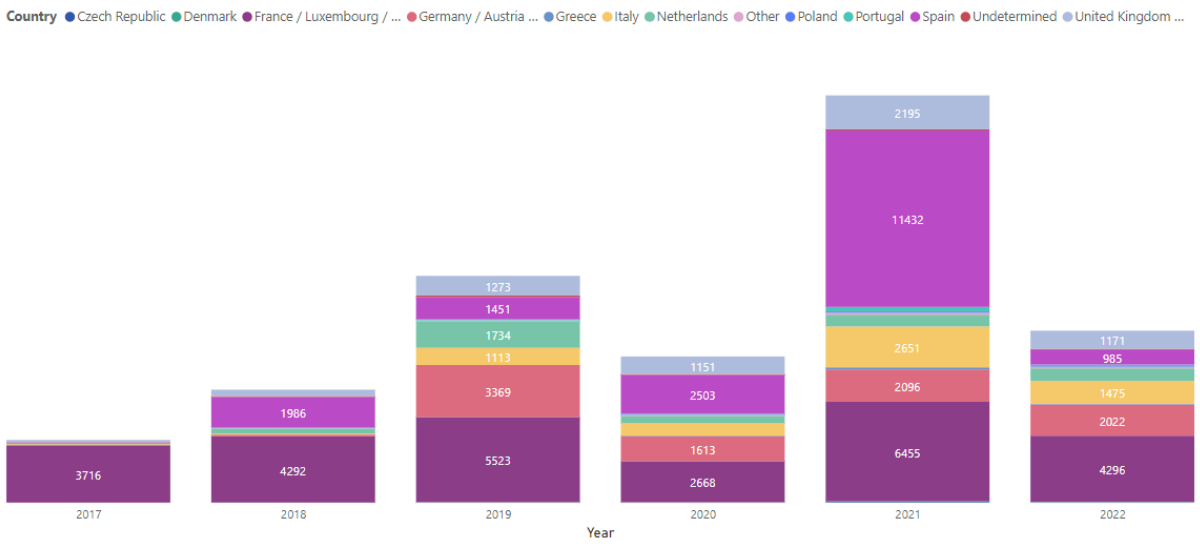


Figure 26 - Number of tweets per country and year

Summing up, it can be concluded that exists a correlation between the number of tweets shared and the country where Nutri-Score was implemented/more discussed.

5.1 Knowledge Extraction of Sentiment Analysis

Besides the sentiment analysis already done by (Septia Irawan et al., 2022) on the Nutri-score topic, the only way until now to get to know the population's opinion on the FoPL systems and, more specifically, on the Nutri-Score system, it was through surveys (both online and in real-life scenarios) as already described in chapter 3.

Regarding the current work, in figure 27 it is presented the overall sentiment retrieved after applying TextBlob and VADER. Although it was already presented in the previous chapter it is also important to add it to this section as it will be relevant to do the comparison.



Figure 27 - (left) Text Blob vs (right) VADER results

Both methods retrieve a lot of neutral sentiment, meaning this that it was not found either a positive or negative sentiment when speaking about Nutri-score in those tweets. TextBlob makes the analysis of sentiments richer in the sense that it calculates not only the polarity but also the subjectivity.

As it was described previously, subjectivity allows one to distinguish whether a text is informative or opinionative. To represent the results of subjectivity when applying TextBlob to Nutri-Score tweets a graph was built and it is presented in figure 28. On the X-axis it is measured the polarity (ranging from -1 to 1) representing the negative, neutral and positive sentiments, while on the Y-axis it is presented the subjectivity (ranging from 0 to 1).

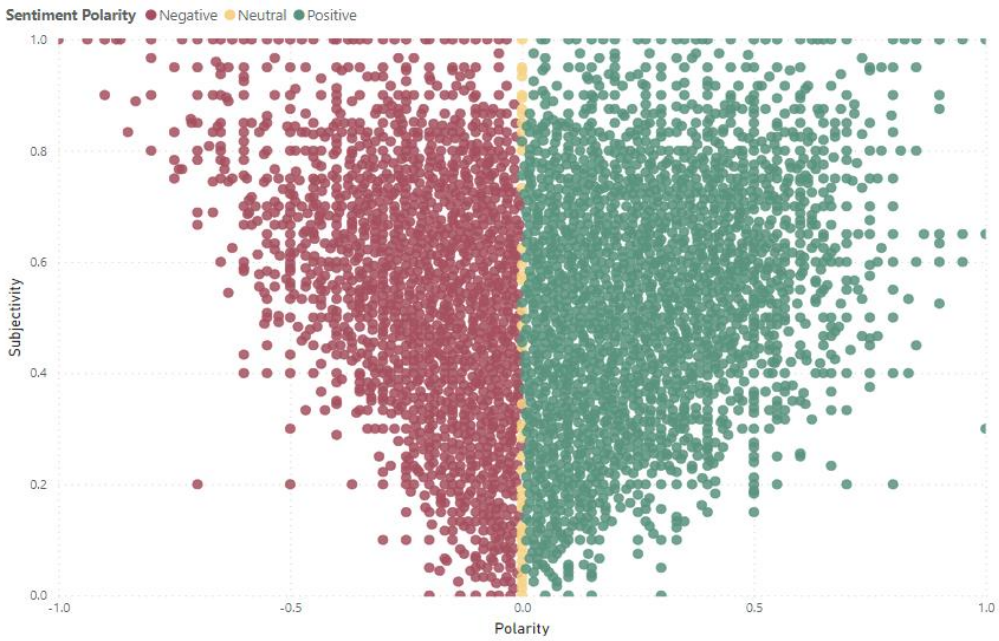


Figure 28 - Analysing Polarity and Subjectivity using TextBlob

By analysing this graph, it can be concluded that when comparing by polarity, there are more objective tweets on the positive side, meaning that the tweets that are considered with having positive sentiments are more informative. Nevertheless, on both polarities, there is a majority of subjective tweets.

Sentiment analysis over time is quite different when comparing both methods. By analysing the graph in figure 29 it is understood that the TextBlob (blue line) has a trend that does not change that much over the months from 2017 until 2022, whereas VADER

represented with the orange line has a negative trend. Although it has a negative trend, it does not have any point (when calculating the average of compound score) less than 0.

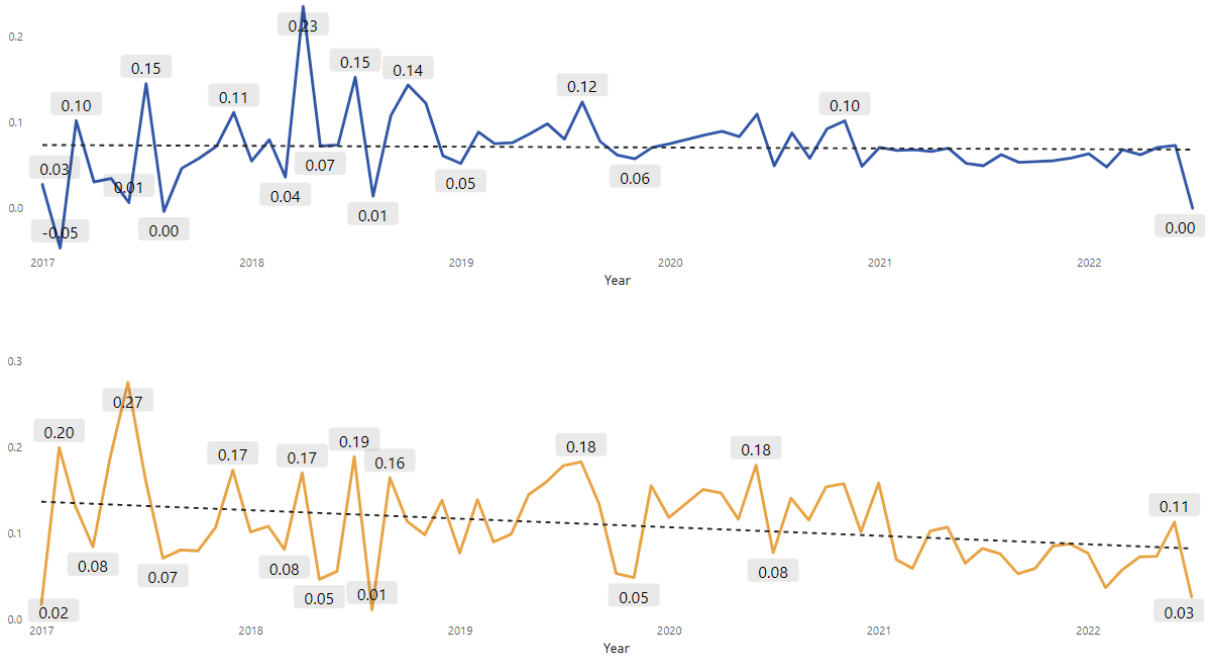


Figure 29 - Sentiment polarity score trend using TextBlob (above) and VADER (below)

5.1.1 Sentiment Analysis by Country

Although (Septia Irawan et al., 2022)has applied a text mining approach and retrieved the overall sentiment of FoPL systems-related tweets, it was not done a cross-country analysis.

As it was understood in the previous section when doing an analysis of the number of tweets by country the major part were tweets from European countries. Having this in mind, the countries used in this current part are from Europe.

The main goal of this section is to compare the questionnaire results (doing an average by country) with the results retrieved from this current research. In table 17 it is presented a summary of the overall result of questionnaires made in different countries when comparing Nutri-Score with other nutrition labels or only to realize the understanding of Nutri-Score itself by consumers. The values in the “Overall Result”

column represent the positive and/or negative sentiment that the consumers from each country have regarding Nutri-Score.

Table 17 - Summary of the overall result of questionnaires

Country	Overall Result	Reference
Belgium	Positive	(Egnell, Talati, et al., 2020)
Bulgaria	Positive	(Egnell, Talati, et al., 2020)
Denmark	Positive	(Egnell, Talati, et al., 2020)
France	Positive / Negative	(Ducrot et al., 2016; Egnell et al., 2021; Egnell, Talati, et al., 2020; Julia et al., 2017; Mazzù et al., 2021)
Germany	Positive / Negative	(Egnell, Talati, et al., 2020; Mazzù et al., 2021)
Greece	Positive	(Mazzù et al., 2021)
Italy	Positive / Negative	(Egnell, Talati, et al., 2020; Mazzù et al., 2021)
Netherlands	Positive	(Egnell, Talati, et al., 2020; van den Akker et al., 2022)
Poland	Positive	(Egnell, Talati, et al., 2020)
Portugal	Positive / Negative	(Egnell, Talati, et al., 2020; Goiana-Da-silva et al., 2021; Mazzù et al., 2021)
Romania	Positive / Negative	(Mazzù et al., 2021)
Spain	Positive / Negative	(Egnell, Talati, et al., 2020; Mazzù et al., 2021)
Sweden	Positive / Negative	(Bottari & Mark-Herbert, 2022)
Switzerland	Positive	(Egnell, Galan, et al., 2020; Egnell, Talati, et al., 2020; Hagmann & Siegrist, 2020)
UK	Positive	(Egnell, Talati, et al., 2020)

When doing an analysis of the sentiment by countries in the current research dataset, and using for this the same column that was used when analysing the number of tweets per country, the following two graphs presented were built in order to present both TextBlob (in figure 30) and Vader results (in figure 31).

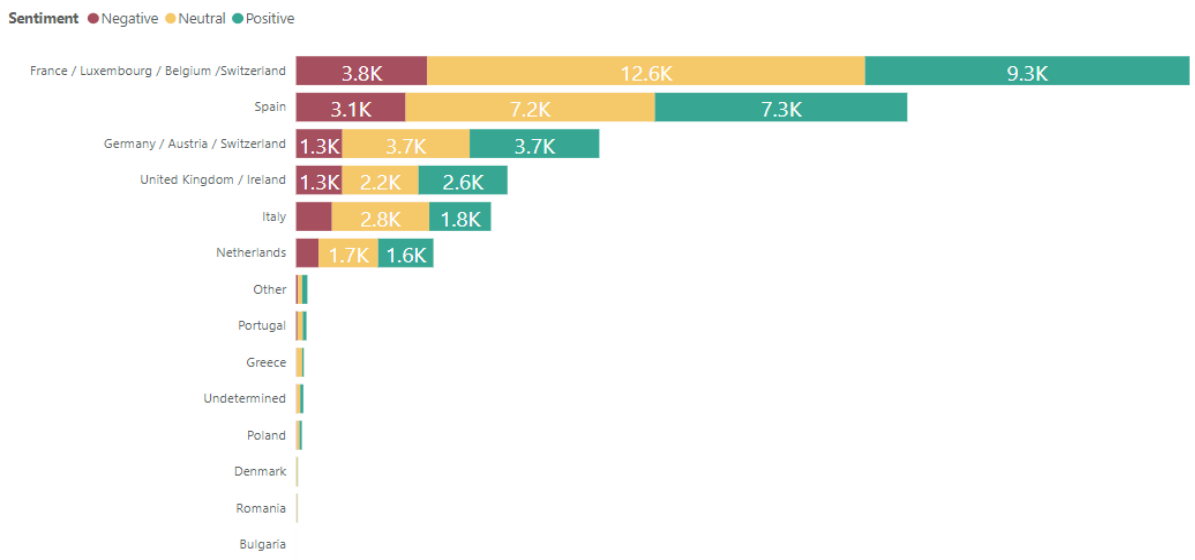


Figure 30 - TextBlob results per country group

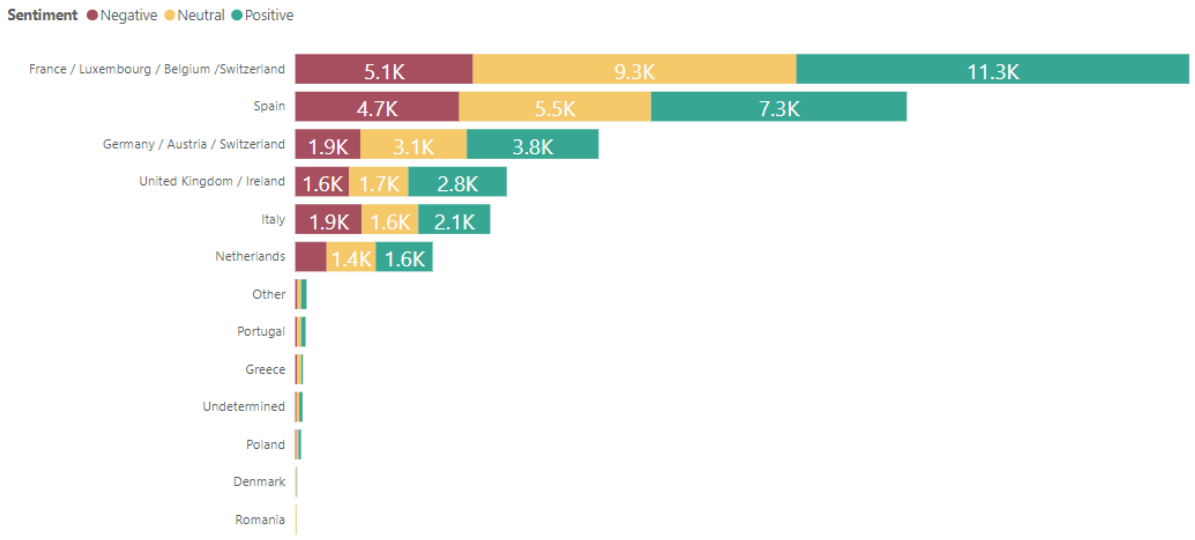


Figure 31 - VADER results per country group

The number of neutral sentiments stands out from the other two polarities, however in order to compare only the positive and negative sentiment, the following graph presented in figure 32 was built to understand clearly the value of each sentiment on the countries at the bottom of the list in figure 30 and figure 31.

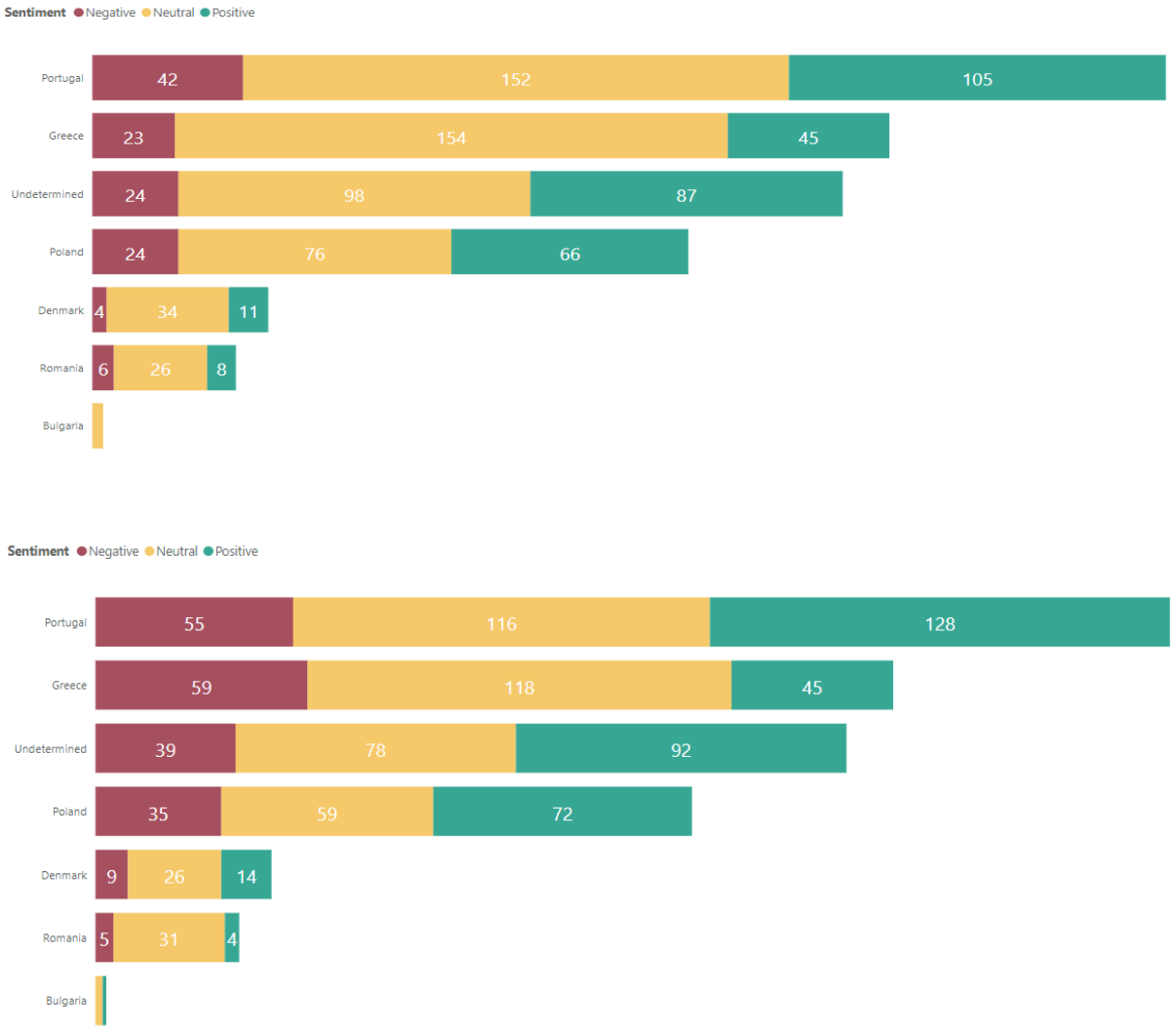


Figure 32 – (above) TextBlob results per country group (bottom level) and (below) VADER results per country

The number of positive sentiments is higher than the negative ones in all countries, except three: Greece, Romania and Bulgaria (when using VADER). Italy is the country that has the most "balanced" dataset in terms of positive/negative values.

When comparing the questionnaire results, it can be stated that the results are similar as most of the countries have a positive sentiment regarding Nutri-Score on both questionnaires and tweets. As an exception, Bulgaria and Greece, as already mentioned, have more negative tweets than positive ones.

The last analysis made was using the subjectivity/polarity calculation of TextBlob. Here it was used the average value of polarity/subjectivity of each country, build a graph with it, comparing all countries. In figure 33, it is presented this, where it can be determined what was seen in figures 30 and 32 that were presenting the number of

positive, neutral and negative sentiments of each country. With this graph, it is possible to quickly conclude that Romania has the lowest polarity of all countries (although in terms of the number of tweets it is much less in comparison with, for example, a French-speaking country). More than this it is also clear that almost all countries have a high level of subjectivity in their tweets.

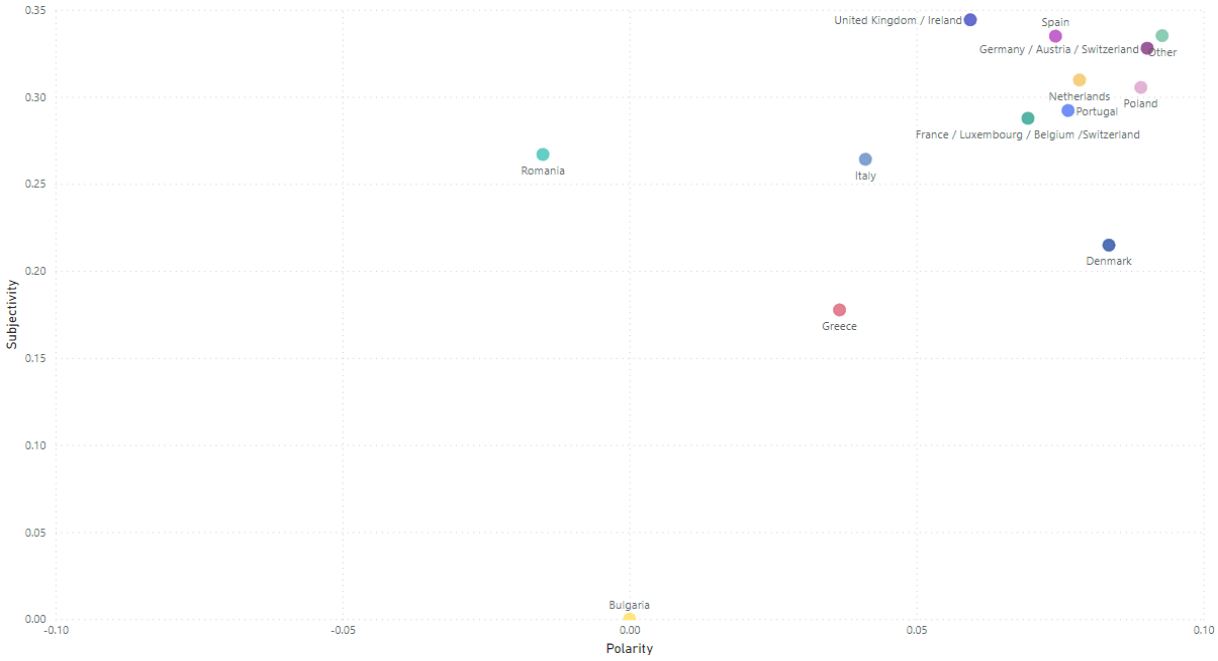


Figure 33 - Subjectivity and Polarity analysis by country group using TextBlob

Having into consideration the analysis taken in this chapter, it can be verified that:

1. There is not an increase in the number of tweets that correlates with the number of news and/or papers published under the Nutri-Score topic

2. With regards to the countries comparison, it can be stated that the number of tweets published by users of each country can be related to the year when Nutri-Score started to gain more notoriety in that specific country.

3. In general, there are more positive sentiments than negative sentiments although the majority are considered neutral. From a country perspective, all countries except three, demonstrate a positive sentiment on the Nutri-Score topic.

CHAPTER 6

Conclusions and Future Work

The purpose of this research was to have more ways of identifying the opinion of the consumers in what regards the Nutri-Score topic applying for this a text mining approach. The aim was to give the decision-makers some additional input to try to understand if the population finds useful this kind of information on the product packages.

To the best of the author's knowledge, the only work that completely matches the current research topic is the one published by (Septia Irawan et al., 2022). This work gets data from Twitter that includes search terms related to FoPL systems and that have origin in Europe. More than defining the sentiment of the tweets, a categorization of the topics discussed in each of them was also built. The major differences from this current research is the number of tweets collected (approximately 5000 versus approximately 80,000), the methods undertaken to analyse the tweet's sentiment and the analysis done after it.

When compared to this research described, an additional point that this current research did was to analyse/compare each country in terms of sentiment on the Nutri-Score topic.

Besides this work other types of papers that analysed the opinion of consumers in what regards to the Nutri-Score labelling system were questionnaires and online surveys. While not including a text mining analysis, these articles try to understand what consumers find more important in a FoPL system, by comparing some of them or by analysing only one in particular (Nutri-Score).

To summarize the findings of the questionnaires undertaken on the papers found, it was built a table that categorizes if a country classifies Nutri-Score as a positive thing to implement or not (on average). The conclusion was that consumers find Nutri-Score as a tool to improve population knowledge of the nutritional value of products, and, more than this, assist in the decision of which product contributes to the consumers' health.

This current work aimed to answer four main research questions and it is safe to affirm that after the analysis was done, those questions were answered:

RQ1: What is the trend when speaking about Nutri Score on the Twitter social network – positive or negative? In overall, the most correct answer to this question is "neutral" as the sentiment that highlighted from the others was the neutral one (especially when using TextBlob). However, it can be said that between positive and negative sentiments, the tweets analysed had more positive polarity than negative.

RQ2: From which countries are the users with positive/negative thoughts related to Nutri-Score? With exception of three countries, all countries had a positive sentiment toward Nutri-Score. In all of them, the neutral sentiments were more than the two other categories, but, if that polarity is "ignored" it can be said that almost all countries demonstrated a positive sentiment when speaking about Nutri-Score. The countries with more negative sentiment than positive were countries where there were not many tweets being published (especially in Romania and Bulgaria) and it's also important to denote that this result was only when using VADER sentiment analysis.

RQ3: Is there a difference between 2017 and 2022 in terms of sentiment polarity, when speaking about Nutri-Score? It was analysed the trend of sentiment over these years - while in the TextBlob method there is not a significant positive/negative trend, when using VADER it can be verified a negative trend.

More than answering the research questions, just like as Nutri-Score and the Front-of-Package Labelling systems as a whole aim to contribute to a society with better nutrition knowledge and free from diet-related diseases (Julia et al., 2018), also this work aims to contribute to that goal. By providing information on the consumer's opinion on the Nutri-Score labelling system, the decision-making of those in charge can be assisted.

6.1 Limitations and Future Work

Some limitations should be considered in this study. First, as (Liew & Lee, 2021) also highlighted in its work, although the social media platform is among the most widely social networks used, it must be identified that it only exemplifies social media data and

may not fully represent users of other social media platforms or even a global perspective (outside social media).

In what regards the data extracted, and the transformations (pre-processing) done, as only 9% of the tweets were in English language, it was not possible to correct abbreviations on the text and/or wrong-spelling lexicons (that are better tuned for the English language). Still in the language topic, the data translation may lead to distortion in the meaning of tweets, and that should also be denoted as a work limitation.

In this work it was used a lexicon-based approach technique, that it has the limitation of not being as in-depth as a manually conducted qualitative analyses, i.e., an analysis that it is based on a human annotating a subset of data with having a positive, neutral, or negative sentiment.

As a future work, the dataset generated by this research may lead to further analysis on this topic, more than the decision support on the topic, it may also be useful to be used on a different kind of analysis. An example of it may be, for example, the usage of supervised machine learning techniques, by using a test/train dataset to extend the usability to another type of models that can convey better results in this type of data.

A lexicon specific to this problem may also be created, based on the nature of this target domain, in order to improve the categorization of positive/negative sentiments.

More than a sentiment analysis it can also be extended to an emotion analysis to identify if a user is showing happiness, anger, sadness, etc. on this specific topic.

References

Abu-Shanab, E., & Harb, Y. (2019). E-government research insights: Text mining analysis. *Electronic Commerce Research and Applications*, *38*, 100892. https://doi.org/10.1016/J.ELERAP.2019.100892

Agarwal, B., & Mittal, N. (2016). *Machine Learning Approach for Sentiment Analysis*. 21–45. https://doi.org/10.1007/978-3-319-25343-5_3

Aguenaou, H., Ammari, L. el, Bigdeli, M., Hajjab, A. el, Lahmam, H., Labzizi, S., Gamih, H., Talouizte, A., Serbouti, C., Kari, K. el, Benkirane, H., Berri, H. el, Al-Jawaldeh, A., & Yahyane, A. (2022). *Comparison of appropriateness of Nutri-Score and other front-of-pack nutrition labels across a group of Moroccan consumers: awareness, understanding and food choices*. https://doi.org/10.1186/s13690-021-00595-3

Alam, S., & Yao, N. (2019). The impact of preprocessing steps on the accuracy of machine learning algorithms in sentiment analysis. *Computational and Mathematical Organization Theory*, *25*(3), 319–335. https://doi.org/10.1007/S10588-018-9266-8/FIGURES/5

Alenzi, B. M., Khan, M. B., Hasanat, M. H. A., Saudagar, A. K. J., Alkhathami, M., & Altameem, A. (2022). Automatic Annotation Performance of TextBlob and VADER on Covid Vaccination Dataset. *Intelligent Automation & Soft Computing*, *34*(2), 1311–1331. https://doi.org/10.32604/IASC.2022.025861

Aljedaani, W., Rustam, F., Mkaouer, M. W., Ghallab, A., Rupapara, V., Washington, P. B., Lee, E., & Ashraf, I. (2022). Sentiment analysis on Twitter data integrating TextBlob and deep learning models: The case of US airline industry. *Knowledge-Based Systems*, *255*. https://doi.org/10.1016/J.KNOSYS.2022.109780

Ares, G., Varela, F., Machin, L., Antúnez, L., Giménez, A., Curutchet, M. R., & Aschemann-Witzel, J. (2018). Comparative performance of three interpretative front-of-pack nutrition labelling schemes: Insights for policy making. *Food Quality and Preference*, *68*, 215–225. https://doi.org/10.1016/J.FOODQUAL.2018.03.007

Ballew, B. S. (2009). Elsevier's Scopus® Database. *Http://Dx.Doi.Org/10.1080/15424060903167252*, *6*(3), 245–252. https://doi.org/10.1080/15424060903167252

Bottari, F., & Mark-Herbert, C. (2022). Development of uniform food information –the case of front of package nutrition labels in the EU. *Archives of Public Health*, *80*(1), 1–15. https://doi.org/10.1186/S13690-022-00915-1/TABLES/6

Carruba, M. O., Caretto, A., de Lorenzo, A., Fatati, G., Ghiselli, A., Lucchin, L., Maffeis, C., Malavazos, A., Malfi, G., Riva, E., Ruocco, C., Santini, F., Silano, M., Valerio, A., Vania, A., & Nisoli, E. (2022). Front-of-pack (FOP) labelling systems to improve the quality of nutrition information to prevent obesity: NutrInform Battery vs Nutri-Score. *Eating and Weight Disorders*, *27*(5), 1575–1584. https://doi.org/10.1007/S40519-021-01316-Z

Chauhan, D. S., Singh, G. V., Arora, A., Ekbal, A., & Bhattacharyya, P. (2022). An emoji-aware multitask framework for multimodal sarcasm detection. *Knowledge-Based Systems*, *257*, 109924. https://doi.org/10.1016/J.KNOSYS.2022.109924

Cortis, K., & Davis, B. (2021). Over a decade of social opinion mining: a systematic review. *Artificial Intelligence Review*, *54*(7), 4873. https://doi.org/10.1007/S10462-021-10030-2

Dang, S., & Ahmad, P. H. (2014). Text Mining : Techniques and its Application Text Mining View project Text Mining: Techniques and its Application. *IJETI International Journal of Engineering & Technology Innovations*, *1*. www.ijeti.com

Dewi, L. C., Meiliana, & Chandra, A. (2019). Social Media Web Scraping using Social Media Developers API and Regex. *Procedia Computer Science*, *157*, 444–449. https://doi.org/10.1016/J.PROCS.2019.08.237

Directorate-General for Health and Food Safety (European Commission), & European Commision. (2020). *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS A Farm to Fork Strategy for a fair, healthy and environmentally-friendly food system - Publications Office of the EU*. https://op.europa.eu/en/publication-detail/-/publication/ea0f9f73-9ab2-11ea-9d2d-01aa75ed71a1/language-en

Dongo, I., Cadinale, Y., Aguilera, A., Martínez, F., Quintero, Y., & Barrios, S. (2020a). Web Scraping versus Twitter API: A Comparison for a Credibility Analysis. *ACM International Conference Proceeding Series*, 263–273. https://doi.org/10.1145/3428757.3429104

Dongo, I., Cadinale, Y., Aguilera, A., Martínez, F., Quintero, Y., & Barrios, S. (2020b). Web Scraping versus Twitter API: A Comparison for a Credibility Analysis. *ACM International Conference Proceeding Series*, 263–273. https://doi.org/10.1145/3428757.3429104

Dréano-Trécant, L., Egnell, M., Hercberg, S., Galan, P., Soudon, J., Fialon, M., Touvier, M., Kesse-Guyot, E., & Julia, C. (2020). Performance of the Front-of-Pack Nutrition Label Nutri-Score to Discriminate the Nutritional Quality of Foods Products: A Comparative Study across 8 European Countries. *Nutrients*, *12*(5). https://doi.org/10.3390/NU12051303

Drewnowski, A., Amanquah, D., & Gavin-Smith, B. (2021). Perspective: How to Develop Nutrient Profiling Models Intended for Global Use: A Manual. *Advances in Nutrition (Bethesda, Md.)*, *12*(3), 609–620. https://doi.org/10.1093/ADVANCES/NMAB018

Ducrot, P., Julia, C., Méjean, C., Kesse-Guyot, E., Touvier, M., Fezeu, L. K., Hercberg, S., & Péneau, S. (2016). Impact of different front-of-pack nutrition labels on consumer purchasing intentions: A randomized controlled trial. *American Journal of Preventive Medicine*, *50*(5), 627–636. https://doi.org/10.1016/J.AMEPRE.2015.10.020

Egnell, M., Galan, P., Farpour-Lambert, N. J., Talati, Z., Pettigrew, S., Hercberg, S., & Julia, C. (2020). Compared to other front-of-pack nutrition labels, the Nutri-Score emerged as the most efficient to inform Swiss consumers on the nutritional quality of food products. *PLOS ONE*, *15*(2), e0228179. https://doi.org/10.1371/JOURNAL.PONE.0228179

Egnell, M., Galan, P., Fialon, M., Touvier, M., Péneau, S., Kesse-Guyot, E., Hercberg, S., & Julia, C. (2021). The impact of the Nutri-Score front-of-pack nutrition label on purchasing intentions of unprocessed and processed foods: post-hoc analyses from three randomized controlled trials. *International Journal of Behavioral Nutrition and Physical Activity*, *18*(1). https://doi.org/10.1186/S12966-021-01108-9

Egnell, M., Talati, Z., Galan, P., Andreeva, V. A., Vandevijvere, S., Gombaud, M., Dréano-Trécant, L., Hercberg, S., Pettigrew, S., & Julia, C. (2020). Objective understanding of the Nutri-score front-of-pack label by European consumers and its effect on food choices: an online experimental study. *International Journal of Behavioral Nutrition and Physical Activity*, *17*(1), 1–13. https://doi.org/10.1186/S12966-020-01053-Z/FIGURES/3

Ferreira-Mello, R., André, M., Pinheiro, A., Costa, E., & Romero, C. (2019). Text mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *9*(6). https://doi.org/10.1002/WIDM.1332

File:Nutri-score-A light background logo.svg - Wikimedia Commons. (2022). Retrieved July 15, 2022, from https://commons.wikimedia.org/wiki/File:Nutri-score-A_light_background_logo.svg

GAIN Convening Paper Series. (2020). *Front-of-Pack labelling and visual cues as tools to influence consumer food choices in low-and-middle-income countries*. https://doi.org/10.36072/cp.6

GitHub - JustAnotherArchivist/snscrape: A social networking service scraper in Python. (2022). Retrieved August 30, 2022, from https://github.com/JustAnotherArchivist/snscrape

Gohil, S., Vuik, S., & Darzi, A. (2018). Sentiment Analysis of Health Care Tweets: Review of the Methods Used. *JMIR Public Health Surveill 2018;4(2):E43 Https://Publichealth.Jmir.Org/2018/2/E43*, *4*(2), e5789. https://doi.org/10.2196/PUBLICHEALTH.5789

Goiana-Da-silva, F., Cruz-E-silva, D., Nobre-Da-costa, C., Nunes, A. M., Fialon, M., Egnell, M., Galan, P., Julia, C., Talati, Z., Pettigrew, S., Darzi, A., Araújo, F., & Hercberg, S. (2021). Nutri-Score: The Most Efficient Front-of-Pack Nutrition Label to Inform Portuguese Consumers on the Nutritional Quality of Foods and Help Them Identify Healthier Options in Purchasing Situations. *Nutrients 2021, Vol. 13, Page 4335*, *13*(12), 4335. https://doi.org/10.3390/NU13124335

Gudivada, V. N., Rao, D., & Raghavan, V. v. (2015). Big Data Driven Natural Language Processing Research and Applications. *Handbook of Statistics*, *33*, 203–238. https://doi.org/10.1016/B978-0-444-63492-4.00009-5

Gupta, I., & Joshi, N. (2021). A Review on Negation Role in Twitter Sentiment Analysis. *Https://Services.Igi-Global.Com/Resolvedoi/Resolve.Aspx?Doi=10.4018/IJHISI.20211001.Oa14*, *16*(4), 1–19. https://doi.org/10.4018/IJHISI.20211001.OA14

Gupta, N., & Agrawal, R. (2020). Application and techniques of opinion mining. *Hybrid Computational Intelligence*, 1–23. https://doi.org/10.1016/B978-0-12-818699-2.00001-9

Hagmann, D., & Siegrist, M. (2020). Nutri-Score, multiple traffic light and incomplete nutrition labelling on food packages: Effects on consumers' accuracy in identifying healthier snack options. *Food Quality and Preference*, *83*, 103894. https://doi.org/10.1016/J.FOODQUAL.2020.103894

Health Organization Regional Office for Europe, W. (2016). *Action Plan for the Prevention and Control of Noncommunicable Diseases in the WHO European Region*. http://www.euro.who.int/pubrequest

Hercberg, S., Touvier, M., & Salas-Salvado, J. (2022). The Nutri-Score nutrition label. *International Journal for Vitamin and Nutrition Research. Internationale Zeitschrift Fur Vitamin- Und Ernahrungsforschung. Journal International de Vitaminologie et de Nutrition*, *92*(3–4), 147–157. https://doi.org/10.1024/0300-9831/A000722

Hutto, C. (2022). *GitHub - VADER Sentiment Analysis*. Retrieved August 28, 2022, from https://github.com/cjhutto/vaderSentiment

Hutto, C. J., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), 216–225. https://doi.org/10.1609/ICWSM.V8I1.14550

Jawad, M., & Mughal, H. (2018). Data Mining: Web Data Mining Techniques, Tools and Algorithms: An Overview. *IJACSA) International Journal of Advanced Computer Science and Applications*, *9*(6). www.ijacsa.thesai.org

Julia, C., Etilé, F., & Hercberg, S. (2018). Front-of-pack Nutri-Score labelling in France: an evidence-based policy. *The Lancet. Public Health*, *3*(4), e164. https://doi.org/10.1016/S2468-2667(18)30009-4

Julia, C., & Hercberg, S. (2018). Big Food's Opposition to the French Nutri-Score Front-of-Pack Labeling Warrants a Global Reaction. *American Journal of Public Health*, *108*(3), 318. https://doi.org/10.2105/AJPH.2017.304284

Julia, C., Péneau, S., Buscail, C., Gonzalez, R., Touvier, M., Hercberg, S., & Kesse-Guyot, E. (2017). Perception of different formats of front-of-pack nutrition labels according to sociodemographic, lifestyle and dietary factors in a French population: cross-sectional study among the NutriNet-Santé cohort participants. *BMJ Open*, *7*(6). https://doi.org/10.1136/BMJOPEN-2017-016108

Julia, C., Touvier, M., Kesse-Guyot, E., Galan, P., & Hercberg, S. (2022). Nutri-Score in tug-of-war between public health and economic interests in the European Union. *Nature Food*, *3*(3), 181. https://doi.org/10.1038/S43016-022-00476-0

Koppel, M., & Schler, J. (2006). THE IMPORTANCE OF NEUTRAL EXAMPLES FOR LEARNING SENTIMENT. *Computational Intelligence*, *22*(2), 100–109. https://doi.org/10.1111/J.1467-8640.2006.00276.X

Krallinger, M., & Valencia, A. (2005). Text-mining and information-retrieval services for molecular biology. *Genome Biology*, *6*(7). https://doi.org/10.1186/GB-2005-6-7-224

Kumaresh, N., Janardhan, N., Bonta, V., & Janardhan, N. (2019). A Comprehensive Study on Lexicon Based Approaches for Sentiment Analysis. *Asian Journal of Computer Science and Technology*, *8*(S2), 1–6. https://doi.org/10.51983/ajcst-2019.8.S2.2037

Liew, T. M., & Lee, C. S. (2021). Examining the Utility of Social Media in COVID-19 Vaccination: Unsupervised Learning of 672,133 Twitter Posts. *JMIR Public Health Surveill 2021;7(11):E29789 Https://Publichealth.Jmir.Org/2021/11/E29789*, *7*(11), e29789. https://doi.org/10.2196/29789

Loria, S. (2020). *Advanced Usage: Overriding Models and the Blobber Class — TextBlob 0.16.0 documentation*. https://textblob.readthedocs.io/en/latest/advanced_usage.html#sentiment-analyzers

Mazzù, M. F., Romani, S., Baccelloni, A., & Gambicorti, A. (2021). A cross-country experimental study on consumers' subjective understanding and liking on front-of-pack nutrition labels. *International Journal of Food Sciences and Nutrition*, *72*(6), 833–847. https://doi.org/10.1080/09637486.2021.1873918

Miner, G. D., Elder, J., & Nisbet, R. A. (2012). Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications. *Practical Text Mining and Statistical Analysis for Non-Structured Text Data Applications*. https://doi.org/10.1016/C2010-0-66188-8

Nutri-Score: The story so far - EuroHealthNet. (2022). Retrieved July 27, 2022, from https://eurohealthnet-magazine.eu/nutri-score-the-story-so-far/

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S.,

… Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ (Clinical Research Ed.)*, *372*. https://doi.org/10.1136/BMJ.N71

Pandey, D., Pradhan, B., Wangmo, & Wairya, S. (2022). Understanding COVID-19 response by twitter users: A text analysis approach. *Heliyon*, *8*(8), e09994. https://doi.org/10.1016/J.HELIYON.2022.E09994

Peonides, M., Knoll, V., Gerstner, N., Heiss, R., Frischhut, M., & Gokani, N. (2022). Food labeling in the European Union: a review of existing approaches. *International Journal of Health Governance*, *27*(4), 460–468. https://doi.org/10.1108/IJHG-07-2022-0072/FULL/PDF

Pettigrew, S., Jongenelis, M. I., Jones, A., Hercberg, S., & Julia, C. (2023). An 18-country analysis of the effectiveness of five front-of-pack nutrition labels. *Food Quality and Preference*, *104*. https://doi.org/10.1016/J.FOODQUAL.2022.104691

Psomakelis, E., Tserpes, K., Anagnostopoulos, D., & Varvarigou, T. (2015). Comparing methods for Twitter Sentiment Analysis. *KDIR 2014 - Proceedings of the International Conference on Knowledge Discovery and Information Retrieval*, 225–232. https://doi.org/10.48550/arxiv.1505.02973

Roberto, C. A., Ng, S. W., Ganderats-Fuentes, M., Hammond, D., Barquera, S., Jauregui, A., & Taillie, L. S. (2021). The Influence of Front-of-Package Nutrition Labeling on Consumer Behavior and Product Reformulation. *Annual Review of Nutrition*, *41*, 529–550. https://doi.org/10.1146/ANNUREV-NUTR-111120-094932

Roy, A. (2018). *Chapter 1 - Introduction to CRISP DM Framework for Data Science and Machine Learning*. https://www.linkedin.com/pulse/chapter-1-introduction-crisp-dm-framework-data-science-anshul-roy/

Sedigh, A. K., & Roudaki, M. (2003). *IDENTIFICATION OF THE DYNAMICS OF THE GOOGLE'S RANKING ALGORITHM*. www.pandia.com/searchworld/

Septia Irawan, A., Shahin, B., Wangeshi Njuguna, D., Nellamkuzhi, N. J., Thiện, B. Q., Mahrouseh, N., & Varga, O. (2022). Analysis of Content, Social Networks, and Sentiment of Front-of-Pack Nutrition Labeling in the European Union on Twitter. *Frontiers in Nutrition*, *9*. https://doi.org/10.3389/FNUT.2022.846730

Shah, P. (2020, June 27). *Sentiment Analysis using TextBlob | by Parthvi Shah | Towards Data Science*. https://towardsdatascience.com/my-absolute-go-to-for-sentiment-analysis-textblob-3ac3a11d524

Sharma, D., Sabharwal, M., Goyal, V., & Vij, M. (2020). Sentiment analysis techniques for social media data: A review. *Advances in Intelligent Systems and Computing*, *1045*, 75–90. https://doi.org/10.1007/978-981-15-0029-9_7

Song, J., Brown, M. K., Tan, M., MacGregor, G. A., Webster, J., Campbell, N. R. C., Trieu, K., Ni Mhurchu, C., Cobb, L. K., & He, F. J. (2021). Impact of color-coded and warning nutrition labelling schemes: A systematic review and network meta-analysis. *PLoS Medicine*, *18*(10). https://doi.org/10.1371/JOURNAL.PMED.1003765

Subramanian, J., Sridharan, V., Shu, K., & Liu, H. (2019). *Exploiting Emojis for Sarcasm Detection*. https://www.facebook.com/

Talib, R., Hanif, M. K., Ayesha, S., & Fatima, F. (2016). Text Mining: Techniques, Applications and Issues. *IJACSA) International Journal of Advanced Computer Science and Applications*, *7*(11). www.ijacsa.thesai.org

Tandel, S. S., Jamadar, A., & Dudugu, S. (2019). A Survey on Text Mining Techniques. *2019 5th International Conference on Advanced Computing and Communication Systems, ICACCS 2019*, 1022–1026. https://doi.org/10.1109/ICACCS.2019.8728547

Temple, N. J., & Fraser, J. (2014). Food labels: A critical assessment. *Nutrition*, *30*(3), 257–260. https://doi.org/10.1016/J.NUT.2013.06.012

TextBlob: Simplified Text Processing — TextBlob 0.16.0 documentation. (2022). Retrieved September 19, 2022, from https://textblob.readthedocs.io/en/dev/

Ting, I.-H. (2008). *Web mining techniques for on-line social networks analysis; Web mining techniques for on-line social networks analysis*. https://doi.org/10.1109/ICSSSM.2008.4598506

Tweepy Documentation — tweepy 4.12.1 documentation. (2022). Retrieved August 30, 2022, from https://docs.tweepy.org/en/stable/

van den Akker, K., Bartelet, D., Brouwer, L., Luijpers, S., Nap, T., & Havermans, R. (2022). The impact of the nutri-score on food choice: A choice experiment in a Dutch supermarket. *Appetite*, *168*. https://doi.org/10.1016/J.APPET.2021.105664

van Dijk, W. B., Fiolet, A. T. L., Schuit, E., Sammani, A., Groenhof, T. K. J., van der Graaf, R., de Vries, M. C., Alings, M., Schaap, J., Asselbergs, F. W., Grobbee, D. E., Groenwold, R.

H. H., & Mosterd, A. (2021). Text-mining in electronic healthcare records can be used as efficient tool for screening and data collection in cardiovascular trials: a multicenter validation study. *Journal of Clinical Epidemiology*, *132*, 97–105. https://doi.org/10.1016/J.JCLINEPI.2020.11.014

Wahyuni, S., Sitompul, O. S., Nababan, E. B., & Sihombing, P. (2021). Social Network Analysis Text Mining on Networks Publication Citation. *2021 International Conference on Data Science, Artificial Intelligence, and Business Analytics, DATABIA 2021 - Proceedings*, 35–39. https://doi.org/10.1109/DATABIA53375.2021.9650327

Williamson, P. O., & Minter, C. I. J. (2019). Exploring PubMed as a reliable resource for scholarly communications services. *Journal of the Medical Library Association : JMLA*, *107*(1), 16. https://doi.org/10.5195/JMLA.2019.433

Wirth, R., & Hipp, J. (2022). *CRISP-DM: Towards a Standard Process Model for Data Mining*.