

Article

# AthPPA: A Data Visualization Tool for Identifying Political Popularity over Twitter

Alexandros Britzolakis <sup>1,\*</sup>, Haridimos Kondylakis <sup>2,\*</sup> and Nikolaos Papadakis <sup>1</sup>

<sup>1</sup> Department of Electrical and Computer Engineering, Hellenic Mediterranean University, Estavromenos Campus, 71410 Heraklion, Greece; npapadak@cs.hmu.gr

<sup>2</sup> Institute of Computer Science, Foundation for Research and Technology-Hellas (FORTH), 70013 Heraklion, Greece

\* Correspondence: alexbritz@hmu.gr (A.B.); kondylak@ics.forth.gr (H.K.)

**Abstract:** Sentiment Analysis is an actively growing field with demand in both scientific and industrial sectors. Political sentiment analysis is used when a data analyst wants to determine the opinion of different users on social media platforms regarding a politician or a political event. This paper presents Athena Political Popularity Analysis (AthPPA), a tool for identifying political popularity over Twitter. AthPPA is able to collect in-real-time tweets and for each tweet to extract metadata such as number of likes, retweets per tweet etc. Then it processes their text in order to calculate their overall sentiment. For the calculation of sentiment analysis, we have implemented a sentiment analyzer that is able to identify the grammatical issues of a sentence as well as a lexicon of negative and positive words designed specifically for political sentiment analysis. An analytic engine processes the collected data and provides different visualizations that provide additional insights on the collected data. We show how we applied our framework to the three most prominent Greek political leaders in Greece and present our findings there.

**Keywords:** data visualization; lexicon-based sentiment analysis; political sentiment analysis; natural language Processing

**Citation:** Britzolakis, A.; Kondylakis, H. AthPPA: A Data Visualization Tool for Identifying Political Popularity over Twitter. *Information* **2021**, *12*, 312. <https://doi.org/10.3390/info12080312>

Received: 9 July 2021

Accepted: 29 July 2021

Published: 31 July 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The emergence of Web 2.0 has shaped the way by which Internet users navigate and communicate through it. Easy data sharing as well as collaboration and interoperability are some of the basic aspects of modern websites compared to the ones of the first generation. In more simple terms, Web 2.0 users are empowered to engage and co-operate, creating new online groups, unlike the first generation of websites where users could passively access content. Data transparency, metadata, semantics, responsive content, rich user interface, and scalability tolerance are additional Web 2.0 characteristics [1,2].

Social networking sites (such as Facebook, MySpace), wikis, blogs, multimedia distribution sites (such as Twitter, Flickr), mash-ups, and rich web-based platforms can be found among the most popular Web 2.0 applications. Micro-blogging, which originally received relatively less interest, although it eventually became a widely popular networking platform for a large number of people, is one of these practices. In theory, micro-blogging is based on blogs (i.e., web logs) on which users can post thoughts, views, and questions on any subject picked. The key distinction between micro- and conventional blogs is that the text size is strictly limited [3,4]. Twitter is currently the most popular online micro-blogging site, allowing its users to send and receive text-based posts composed of up to 280 characters, known as “tweets”. Created in 2006, Twitter now records more than 330 million active users who send more than 340 million tweets a day [5]. Micro-blogging sites, such as Twitter, have grown with their increasingly growing success into a functional way of expressing views about nearly all facets of daily life [6]. Tweets' strict

character limitation forces users to be straightforward and consequently more articulate than they are for social networks and blogs. Micro-blogging posts are thus imbued with emotional data and are considered rich sources of opinion mining data [2,7]. Additionally, it is possible to handle tweets more quickly than long blog posts and articles.

Polarity detection is often called sentiment analysis and involves detecting the connotation or opinion expressed in a given text [8]. This involves detection of the expressed emotion over a particular topic. These emotion indicators can usually be positive, negative and/or neutral. Even before Twitter and micro-blogging platforms, sentiment analysis was a common research field to a range of domains, as it can deliver benefits from revenue forecasts, politics as well as investors' choices [9–11]. In addition, automated sentiment analysis on textual corpora has been a research subject for many approaches. Examples include, among others, product and services reviews [12], articles on the Web [13], and news feeds [14]. Specifically, the expressive ability and immediacy of Twitter have prompted researchers to attempt to use it in politics [15], tourism [16], as well as many other disciplines. Especially, financial and economic modelling can be one of the most promising disciplines of Twitter sentiment analysis.

The real-time analysis based on the emotions of users is a challenging task that also requires analysts searching through nearly endless papers and news feeds manually. As such, analysis on Twitter is one of the best options for the automated detection of opinions. To this direction, a prevalent sentiment analysis research area is machine learning-based sentiment classifiers. In the case of tweets, though, there are issues of accuracy [17], since classifiers usually have to examine syntactically inconsistent terms due to the character limitation. An additional limitation is that classifiers usually distinguish sentiment into classes (positive, negative and neutral), assigning a corresponding score to the post as a whole, regardless the fact that many aspects of the same "notion" may be discussed in a single post. We argue that a *single score for each tweet is not enough and more detailed analysis is required*.

Complementing sentiment analysis, sentiment visualization is part of the more general research field of text visualization, understood as a research task in data visualization (InfoVis) and visual analytics to interpret sentiment found in textual content. Sentiment visualization applications and activities include, for example, public opinion tracking in social media, literature analysis for digital humanities, or support for sentiment and position studies in linguistics and NLP [17]. However, tools providing sentiment visualization usually *offer limited visualization options, limiting the exploration potential of the users*.

In this paper, we introduce AthPPA, a data visualization tool that is able to visualize any available data that Twitter provides, applied specifically to the top three Greek politicians. More specifically,

- Our tool is the first tool for analyzing tweets in the Greek Language and visualize them. It communicates with Twitter by using its API, collects tweets based on a set of criteria and visualizes the result using numerous graphs.
- We have also implemented a sentiment analyzer that utilizes a lexicon specifically designed for political sentiment analysis. Instead of calculating a single score for each tweet, our approach distinguishes the individual domain characteristics of each tweet and assigns respective sentiment scores for each individual characteristic, resulting in a more thorough analysis of the sentiments of a statement given. This results overall in a more elaborate analysis of post opinions regarding a specific topic.
- As a proof of concept, we apply our tool to the three most prominent Greek politicians, i.e. Kyriakos Mitsotakis leader of New Democracy (liberal-right wing) and current Prime Minister of Greece; Alexis Tsipras, leader of SYRIZA party (radical-left wing); and Fofi Gennimata, leader of the political party Movement for Change (center-left wing), and we present the high-value insights identified by the various visuals our tool provides.

To the best of our knowledge, our tool is currently the only one available online (Website for AthPPA: <https://athppa.cs.hmu.gr/>, accessed on 30 July 2021) for identifying

political popularity over Twitter in Greek Language. In addition, we have to note that AthPPA is the result of a master thesis in the Hellenic Mediterranean University [18].

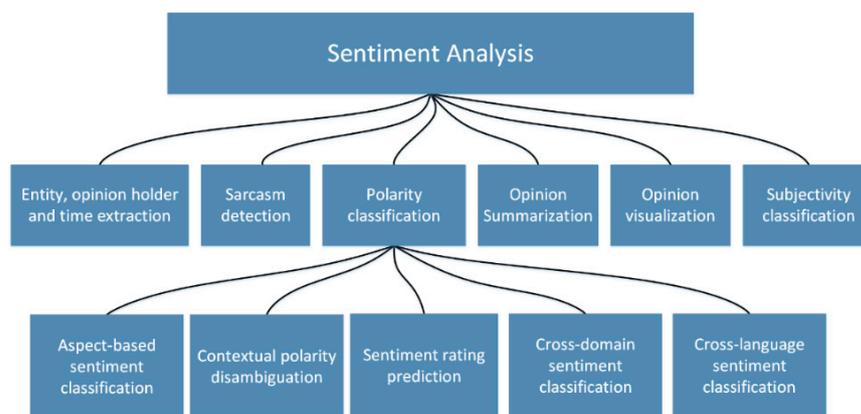
This paper is organized as follows: Section II presents a related work and preliminaries regarding the latest trends, topics, and implementations on sentiment analysis over Twitter; Section III presents a technical overview of the AthPPA architecture, whereas Section IV shows the tool in a proof-of-concept scenario, showing that interesting visualizations can offer interesting insights on the popularity that a political leader has. Section V presents some open topics regarding the implementation; Section VI concludes this paper.

## 2. Preliminaries and Related Work

Sentiment Analysis focuses mostly on the detection of emotional states from a given text and together with the Opinion Mining they can be characterized as fields of Text Mining, which is the process of obtaining crucial information over un-structured textual data. Opinion Mining has greater marketable value than data mining, since it is the most natural way to store data in text format. Since it involves the handling of unstructured and non-defined data, it is a far more complicated task than structured data mining.

While both terms may appear to be similar fields due to the conventional text mining or fact-based analysis, they differ significantly. Although word sentiment can be characterized easily based on sentiment lexicons, the polarity classification is far more difficult as for a given topic there might be many sentiments involved, whereas the topic itself might not be clear.

On the other hand, Opinion Mining utilizes a set of further functionalities than identification of sentiment such as summarization [19]. In other words, sentiment analysis seeks in a particular text to identify words or expressions that indicate an emotion while opinion mining seeks from a particular text to extract and analyze people's thoughts about an entity or an event. A classification of the various fields of work on sentiment analysis is presented in Figure 1 based on Pozzi et al. [20].



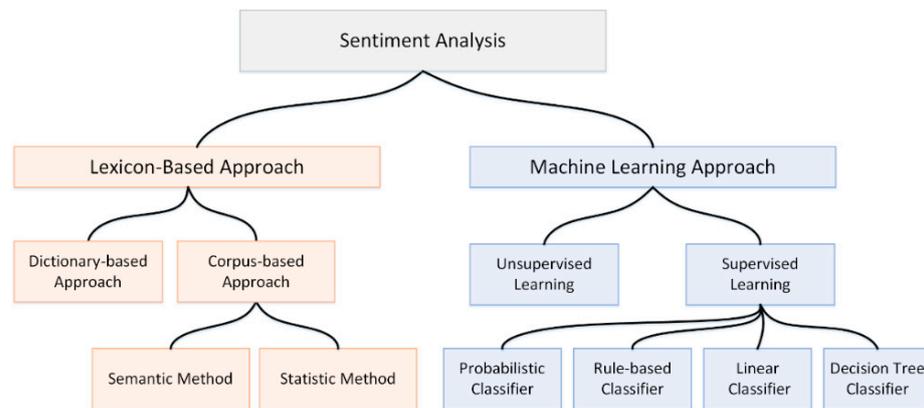
**Figure 1.** Sentiment analysis tasks as listed by Pozzi et al. [20].

Further, Bo and Lee [21] mention that Sentiment Analysis (SA) can be separated into two different methods, presented in Figure 2, which are:

1. The lexicon-based method where a given text is parsed to locate certain phrases or words with its main purpose to identify their respective sentiment value or emotion. The approach is to obtain with their given orientations the initial seed set of words and then search online dictionaries such as WordNet etc. The Corpus-based approach effectively provides corpus analytics to evaluate words of sentiment, although it is not as successful as the dictionary-based method. It is useful for finding the domain and context of specific sentiment words against the corpus data. This

scheme is beneficial when we seek data in order to determine sentiment words [22,23].

- The Machine Learning approach in which a set of marked-up collections known as datasets and feature lists are used as the primary source of knowledge on which a mathematical algorithm relies when classifying other marked-up collections (test sets).



**Figure 2.** Diagram of sentiment analysis methods.

It is worth noting that the Hybrid approach is also a mixture of the two approaches described above, and it is quite popular with sentiment lexicons that play a vital role in most methods.

### 2.1. Levels of Sentiment Analysis and Features

Ankitkumar et al. [24] suggested that Sentiment Analysis can be divided into three stages. The first stage namely as "Message or document" has as its purpose to set the emotional classification of the entire message that indicates an opinion. For example, let's assume that we have a tweet text from an individual. The system will determine if the entire text of the tweet expresses a positive or a negative opinion about something e.g., a politician or an event [18]. In more simple terms, each sentence is an opinion indicator and thus it can be positive, negative, or neutral. If the sentence does not indicate an opinion, then it is neutral. At the second stage, namely as "sentence", each sentence that a text contains is accordingly classified. Thus, each sentence that a text contains is an opinion indicator over an entity or an event. More specifically, it determines whether each sentence expresses an opinion that can be positive, negative, or neutral if no sentiment is identified. The third stage, namely as "Entity or Aspect", is considered as a fine-grained analysis compared to the above-mentioned stages. For instance, the following sentence "Pineapples are tasty but rot very easily" is an opinion indicator over pineapples where the first indication is that they taste fine (which is positive) and the second one is that they rot very easily (which is negative).

There are several ways that we can obtain data in sentiment analysis. Starting with the keyword spotting, where the text is listed in this system based on the presence of unambiguous words present in it [19]. Thus, with regard to sentiment analysis, the words or keywords found in the text are of significance. On the basis of a broad training corpus, the statistical method measures the sentiment or target of affective keywords and word co-occurrence frequencies. Sentiment analysis can be performed in such cases using detected opinion-bearing lexicon objects. Sentiment characteristics (features) are as follows: Presence and frequency of terms where these characteristics are nothing more than actual words or word n-grams and their set of frequencies [25]. Finalizing the opinion, words and phrases are sentences or word indicators that may express an opinion regarding

product or service in the text (e.g., awesome or awful, like or dislike) [19]. In some cases, sentences might convey an opinion even if there are non-opinion words present.

## 2.2. Related Work

In this section, we provide an overview of the related work in the area of sentiment analysis in Twitter, which has been active and constantly rising [26–28], and then we focus on the approaches of sentiment analysis in identifying political popularity.

**Generic Sentiment Analysis.** Wei and Sebastiani [29] proposed a framework that focuses on the allocation or intensity of emotion groups in the collection they study. Bouazizi and Tomoaki [30] suggested a pattern-based approach to sentiment quantification in Twitter. The authors identified two criteria to quantify the accuracy of sentiment classification and demonstrated that sentiment quantitative analysis can be more accurate than regular multi-class classification. Baumgarten et al. [31] discuss a keyword-based classifier for sentiment mining using short messages. It presents a basic classification system that could be expanded in the long term in order to include more sentiment dimensions. This might eventually lead to a better understanding of user preferences, which might then be used to impact future development efforts or marketing campaigns in real time. In addition, approaches like [32] use Bayesian networks along with sentiment analysis in order to perform generic opinion mining.

**Sentiment analysis over Twitter for movies.** Other implementations include the use of lexicon-based techniques in order to identify tweet's overall sentiment from particular movie reviews throughout Twitter such as in the case of Azizan et al. [33]. Bhoir et al. [34] designed a methodology for seeking subjectivity of sentences by using a rule-based framework in order to determine the feature-opinion pair and using another technique, the orientation of the extracted opinion. Mandal et al. [35] implemented a lexicon-based algorithm in order to predict and analyze the emotions available in online movie reviews. This lexicon-based algorithm uses positive, comparative, and superlative levels of comparison on words.

**Distinguishing legitimate from fake accounts.** Carrucio et al. [25] proposed an innovative method for distinguishing legitimate from fake social media profiles. To characterize typical patterns of fake accounts, the methodology uses information automatically collected from big data. They tested their proposed methodology on the Twitter social network and found it to be effective in terms of discriminating capabilities. Carrucio et al. [36,37] proposed a relaxed functional dependencies (RDF) [38] identification technique based on a previously utilized lattice-structured search space, novel pruning strategies, and a novel candidate RDF validation approach. An experimental evaluation reveals the proposed algorithm's discovery performance on real datasets, as well as a comparison to other algorithms.

**Emotions through emoticons & slang.** Boia et al. [39], in their work, depended on the detection of emoticons in order to identify the polarity of the tweets while Manuel et al. [40], in their work, proposed slang identification on tweets in order to get sentiment score from online texts.

**Public sentiment and breakpoints.** Akcora et al. [41] suggested a tool for determining shifts in public sentiment over time and defining the headlines that contributed to breakpoints in public opinion.

**Greek sentiment analysis.** For the Greek language there has also been research for sentiment analysis over Twitter data. For example, [42] examines discussions around COVID-19 in Greece, identifying the fear of transmission of the virus in the community along with a mood, positive and negative emotion analysis. In [43], a Spark-based software-based architecture was proposed for the sentiment analysis of streaming data, whereas [44] exploits machine learning for categorizing a text's sentiment into positive, negative, or neutral.

### 2.2.1. Sentiment Analysis for Political Popularity

Regarding political popularity, the works that are closer to our approach are [15,45–48], whereas a review on the area is also available [49]. Table 1 presents an overview of the main features of each work, enabling also a quick comparison with our proposal.

More specifically, Zhou et al. [45] focused on the Australian federal election 2010, trying to capture the sentiment of the specific political candidates. The authors proposed the specific tool in order to predict trends instead of using polls, and they were able to identify positive, negative, or neutral tweets for a given entity (political person here).

Tumasjan et.al [15], on the other hand, focused on the German federal election and investigated whether Twitter is used as a forum for political deliberation and whether messages on Twitter validly mirror offline political sentiments. According to their findings, the mere number of messages mentioning a party reflects the election result.

Rezapour et al. [46] implemented and tested an improved model that integrates manually-annotated descriptive hashtags into a lexicon to increase the precision of sentiment analysis. The introduction of those descriptive hashtags increases prediction accuracy by about 7%. The model is used to identify and rank the candidates of the Republican and Democratic Party of the 2016 New York primary election by the decreasing ratio of tweets that mentioned these individuals and had positive valence, and compares our results to the election outcome.

**Table 1.** Comparison of the works focusing on sentiment analysis for political popularity.

	Context	Target	Language	Sentiment Analysis	Intelligent Visualization	Available Online
Zhou et al. [45]	Australian federal election 2010	Predict trends, to be used instead of polls.	English	Identify positive, negative or neutral	No	No
Tumasjan et.al [15]	German Federal election, 2009	Evaluate online political sentiment, Predict election result	German auto translated to English	Identify future orientation, past orientation, positive emotions, negative emotions, sadness, anxiety, anger, tentativeness, certainty, work, achievement, and money	No	No
Rezapour et al. [46]	New York primary election 2016	Identify and rank candidate, compare results to the election outcome	English	Identify negative, positive and neutral sentiment	No	No
Ramteke et al. [47]	US elections 2016	Identify the popularity of the candidates	English	Identify negative, positive and neutral sentiment	No	No
Sahu et al. [48]	Approval rating of the President	Analyze the relationship between tweets	English	An floating point number between -1	No	No

	of the US (Donald Trump)	generated by POTUS and his approval rating		(negative) and 1 (positive).		
AthPPA	Greek Election 2019		Greek	Identify Happiness, Surprise, Sadness, Neutral, Fear, Disgust & Anger	Yes	Yes

Ramteke et al. [47] conducted the creation of a dataset using the Twitter streaming API. Then pre-processing was performed on the data of the dataset in order to exclude special characters. At the final stage, data labeling is handled manually, using hashtag labeling with the VADER tool, which is a lexicon and rule-based sentiment analysis tool.

Sahu et al. [48] analyzed the relationship between tweets generated by President of the United States (POTUS) and his approval rating using sentiment-analytics and data visualization tools. The authors used NLP techniques and TextBlob in order to calculate the polarity and subjectivity of each Twitter feed.

Comparing AthPPA with the rest of the approaches in the area, we can see that it is the only tool that is currently available online, providing multiple intelligent analytics and implemented natively for the Greek language—all other works are implemented for the English language with the exception of Tumasjan et.al [15], in which German is auto translated to English.

### 3. System Architecture

AthPPA is a tool for identifying the popularity among different political leaders as well as the latest trends that occur in social media and especially in Twitter. The architecture of our system, used for collecting and analyzing political popularity, is shown in Figure 3.

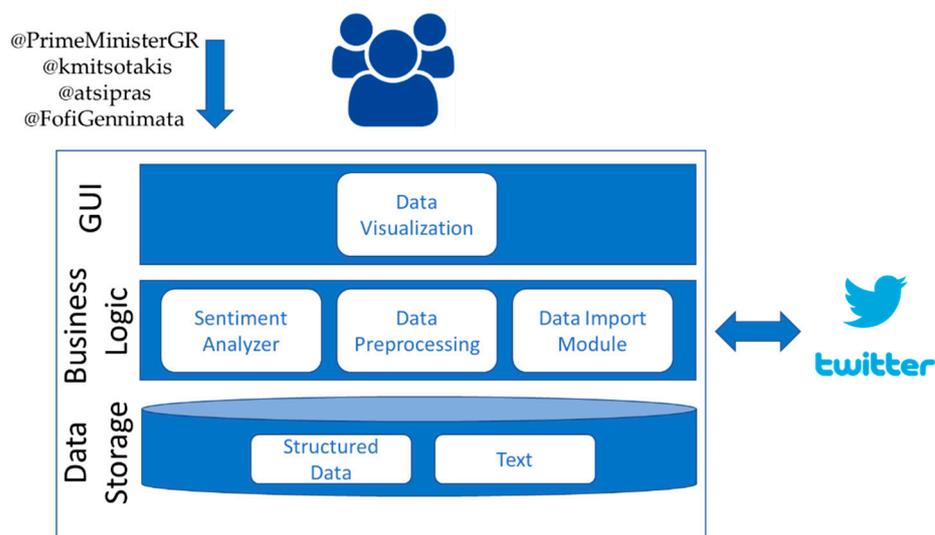


Figure 3. High-level architecture of the AthPAA system.

As shown, it consists of three layers, the graphical user interface (GUI), the business logic, and data storage layer. Once the application is used by the user, the data import module communicates with Twitter using its API, in order to obtain the samples of live tweets. Once these samples are assembled, there are two types of information which the application handles.

The first type is the structured information where it provides exact numerical values such as the number of likes, retweets, character count per tweet, and so on. Based on this structured information, we analyze the number of likes, retweets, and maximum characters per tweet, as well as the total number of subscribers.

The second type of information is the unstructured data, which is a set of raw information that has to be processed in order to extract actual numerical values. The unstructured information is the text available in the tweet. The textual information of a tweet might contain some special characters such as emoticons and exclamation marks, which, in our case, are not useful. For this reason, the data preprocessing module implements a text format parser of each obtained tweet using regex. This technique removes any unwanted special characters. This process is crucial as these special characters are not sentiment value indicators and thus have to be excluded.

Next, the sentiment analyzer module properly processes the tweet's entire text in order to extract its overall sentiment. The analyzer examines the set of tweets stored on each data array, whereas a loop checks all the elements of the data frame and each sentence is processed by the sentiment analyzer in order to extract a proper value by performing grammar checks, tokenization etc. In order to extract sentiment from tweets written in Greek language, we used SpaCy (<https://github.com/eellak/gsoc2018-spacy>, accessed on 30 July 2021), which is a Natural Language Processing tool, in conjunction with a well-known sentiment lexicon (<https://github.com/MKLab-ITI/greek-sentiment-lexicon>, accessed on 30 July 2021) designed specifically for political sentiment analysis [50]. The sentiment analyzer of AthPPA was based on the additional submodule provided in spaCy's Greek implementation repository on Github [22]. This sentiment analyzer utilizes the NLP features of spaCy in order to tokenize each word of a sentence, afterwards it matches them with keywords located in the lexicon of Tsakalidis et al. [50], and in the final stage it merges the tokenized sentence and prints its corresponding emotion, its subjectivity, and its overall score. As the lexicon has been constructed by multiple human annotators, the identified emotion is also annotated as *objective*, or *strongly* or *weakly subjective*. The subjectivity is the average numeric score of the subjectivity based on the individual annotator's score. A sentence example is followed by being written in Greek language and the corresponding output, which this sentiment analyzer provides (Figure 4):

**Sentence:** Έχω μείνει έκπληκτος! Πώς γίνεται αυτό; Η έκπληξη είναι τόσο μεγάλη! Α, τώρα εξηγούνται όλα.  
**Produced Output:**      Subjectivity: 16.6%      Main emotion: Surprise.      Emotion score: 33.3%

**Figure 4.** Initial sentiment analyzer module of spaCy (Greek version).

We have extended this sentiment analyzer in order to obtain each tweet and to analyze it with the same fashion as described above. The following table (Table 2) presents the sentiment values, which our implemented sentiment analyzer assigns.

**Table 2.** Sentiment values used for the labelling process by the sentiment analyzer used in AthPPA

Emotion Type:	Happiness	Surprise	Sadness	Neutral	Fear	Disgust	Anger
Sentiment Value:	3	2	1	0	-1	-2	-3

Once the information is assembled and extracted, then it is visualized by the data visualization module. The main purpose of AthPPA is not to demonstrate a highly accurate sentiment analyzer but to present a web application tool that is capable of presenting any available information that will identify how popular a political leader is throughout Twitter. Although the implemented sentiment analyzer brings fairly good results (since it uses all the NLP capabilities of spaCy tool and that of a solid sentiment lexicon), our main

subject is to visualize any available data efficiently in graphs in order to present which political leader has better presence on Twitter.

We have to note that we have currently disabled the automatic collection of new tweets and we only show online the result of the analysis documented in this paper. However, our system can be used for collecting and analyzing daily the tweets from a predefined set of accounts and hashtags. Besides, users can navigate to the available graphs, and on each graph, they can further filter and refine the presented results as the charts are dynamically updated upon user selections.

### 3.1. Implementation Details

For the creation of this application, Python 3.8 was used in accordance with Tweepy, a Python module that allows the application to communicate with Twitter and fetch data from the platform. Furthermore, Python Dash is used, which is a Python web visualization framework that provides a plethora of features for the creation of dynamic graphs.

The structured data taken from Twitter are the total number of likes, retweets, and text length per posted tweet, as well as the total number of subscribers per account for a set sample of 200 tweets. Furthermore, negative hashtag counter is also a functionality that these classes perform. Figure 5 presents a component diagram of the web application. Note also that the whole implementation is also available online on a Github repository (Github link for AthPPA: <https://github.com/CodeBrakes/AthPPA>, accessed on 30 July 2021).

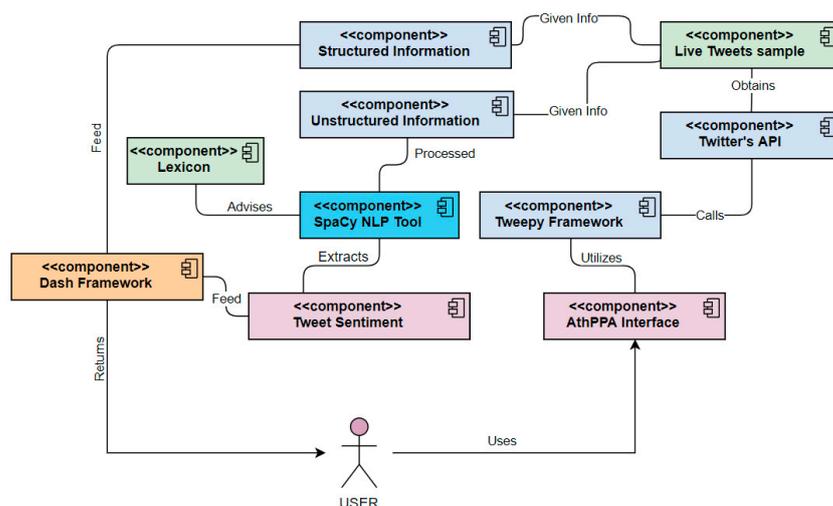
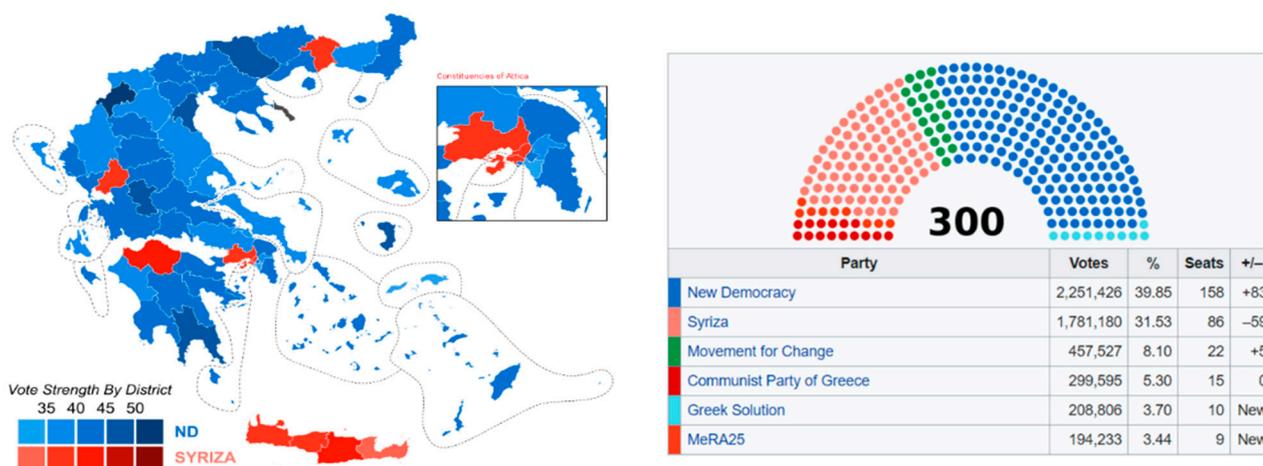


Figure 5. Component diagram of AthPPA.

## 4. Proof of Concept Experimentation

We applied our tool to the Greek elections of 2019. The last legislative elections in Greece were held on 7<sup>th</sup> of July 2019, where the New Democracy centre-right conservative party of Kyriakos Mitsotakis won with 158 from the overall 300 seats of the Greek parliament, leading to an outright majority. Figure 6 depicts the election results.



**Figure 6.** Results of the last Greek legislative election, showing the vote strength of the party winning a plurality in each electoral district.

In addition, it shows the Greek parliament seat ratio among the different political parties. SYRIZA (Coalition of the Radical Left) won 86 seats, making it the official opposition party in Greece; Movement for Change (or KINAL), which is a centre-democratic socialist party won 22 seats; and afterwards follow the rest of the political parties, which have a small share of seats in the Greek parliament (34 seats in total). For the purposes of this research, the three most prominent political parties and their representing leaders are included. These ones are the current Greek Prime Minister Kyriakos Mitsotakis, which is leading the New Democracy party; afterwards is Alexis Tsipras, which is leading the coalition for the Radical Left, making it the official opposition party; and finally, Fofi Gennimata, which is leading the Movement of Change party. Table 3 depicts the identified Twitter accounts and the data obtained. We also examine the Twitter accounts of their respective official political parties as well as the frequency of negative hashtags which Twitter's users include in their tweets.

**Table 3.** Sentiment values used for the labelling process by the sentiment analyzer.

Twitter Account	Type	Person/Entity	Representation
@PrimeministerGR	Politician	Kyriakos Mitsotakis	ND (Majority)
@kmitsotakis	Politician	Kyriakos Mitsotakis	ND (Majority)
@neademokratia	Political Party	New Democracy	ND (Majority)
@atsipras	Politician	Alexis Tsipras	SYRIZA (2 <sup>nd</sup> Opposition)
@syriza_gr	Political Party	SYRIZA	SYRIZA (2 <sup>nd</sup> Opposition)
@FofiGennimata	Politician	Fofi Gennimata	KINAL (3 <sup>rd</sup> Opposition)
@kinimallagis	Political Party	KINAL	KINAL (3 <sup>rd</sup> Opposition)

Person:       Entity:

From those accounts, we obtain for each one of them a dynamic sample of 200 posted tweets using Tweepy Python library, overall collecting 800 tweets. The data visualized are the number of likes, retweets, and text character count for each tweet of the obtained set. We also identified negative hashtags for the two prominent political parties that Twitter's users tend to use in their posted tweets. Table 4 depicts those identified hashtags. We collected 100 tweets per hashtag, resulting in 900 tweets. The two collections were independent.

**Table 4.** Identified negative hashtags per political party and negative/positive hashtags for covid-19 measures.

Identified hashtag	Hashtag Relation	Tweet Sample	Data Visualized
#ΝΔ_θελατε	Negative for ND	100	Date posted frequency
#ΝΔ_ξεφτίλες	Negative for ND	100	Date posted frequency
#ΝΔ_ρομπες	Negative for ND	100	Date posted frequency
#ΣΥΡΙΖΑ_ξεφτίλες	Negative for SYRIZA	100	Date posted frequency
#συριζωα	Negative for SYRIZA	100	Date posted frequency
#Συριζα_απατεώνες	Negative for SYRIZA	100	Date posted frequency
#ΚΙΝΑΛ_ξεφτίλες	Negative for KINAL	100	Date posted frequency
#πανδημία_ηλιθίων	Negative for Covid-19 restriction measures	100	Date posted frequency
#σηκώνουμε_μανίκια	Positive for Covid-19 restriction measures	100	Date posted frequency

At this stage, we can present the graphs and discuss the visualized results. The following figures depict statistics taken from the structured data that tweepy module fetches, such as likes and retweets, character count per tweet from the political accounts described on table 3, as well as negative hashtag frequency included by Twitter's users as described on table 4. Furthermore, sentiment analysis is calculated by our sentiment analyzer, which parses the tweets of each political leader and identifies their overall sentiment by advising the lexicon of Tsakalidis et al. [50]. The total dynamic sample is 800 tweets taken from the Twitter accounts of the three most prominent Greek political leaders, a 600 tweet sample taken by their respective official political parties in Twitter, and a 900 tweet sample for the identified negative hashtags about their respective political parties as well as for the government measures regarding Covid-19 pandemic.

Figure 7 depicts the number of user reactions, likes, and retweets of each tweet from a dynamic sample of 200 tweets taken from Kyriakos Mitsotakis' (@kmitsotakis) account. Axis y depicts the number of likes and retweets that a single tweet has, while the x axis depicts the number of each tweet over this 200-tweet sample.

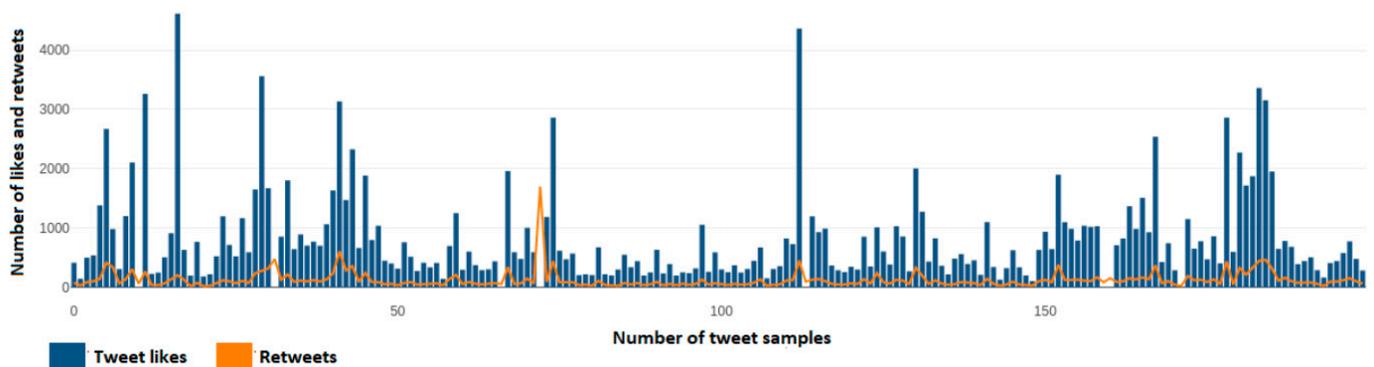
**Figure 7.** Number of likes and retweets per tweet for @kmitsotakis (200 tweet sample).

Figure 8 depicts the number of user reactions, likes, and retweets of each tweet from a dynamic sample of 200 tweets taken from the Hellenic Prime Minister (@primeMinisterGR) account, which currently Kyriakos Mitsotakis administers. Axis y depicts the number of likes and retweets that a single tweet has, while the x axis depicts the number of each tweet over this 200-tweet sample.

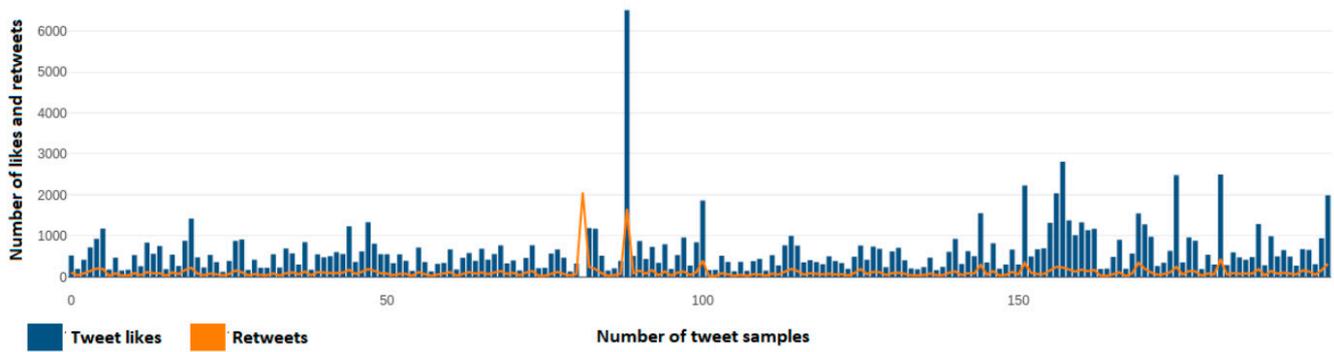


Figure 8. Number of likes and retweets per tweet for @primeMinisterGR (200 tweet sample).

Figure 9 depicts the number of user reactions, likes, and retweets of each tweet from a dynamic sample of 200 tweets taken from Alexis Tsipras' (@atsipras) account. Axis y depicts the number of likes and retweets that a single tweet has, while the x axis depicts the number of each tweet over this 200-tweet sample.

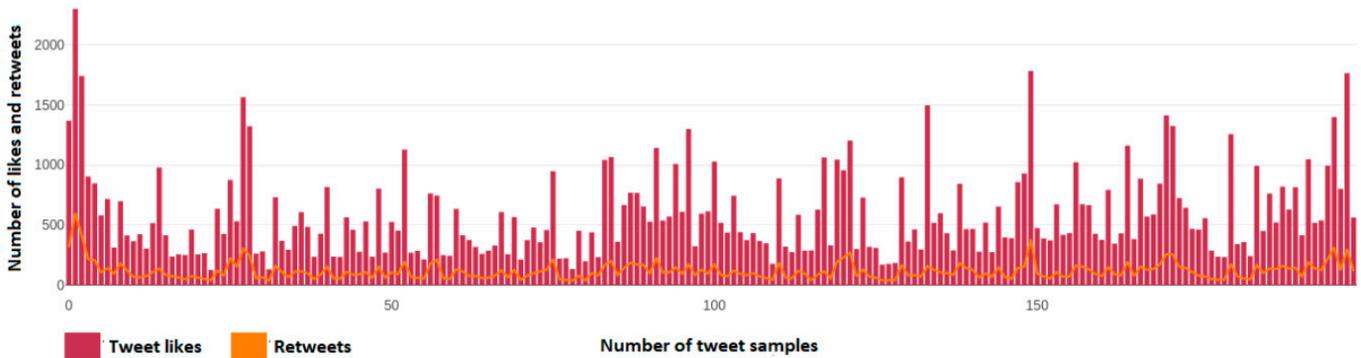


Figure 9. Number of likes and retweets per tweet for @atsipras (200 tweet sample).

Figure 10 depicts the number of user reactions, likes, and retweets of each tweet from a dynamic sample of 200 tweets taken from Fofi Gennimata's (@fofigennimata) account. Axis y depicts the number of likes and retweets that a single tweet has, while the x axis depicts the number of each tweet over this 200-tweet sample.

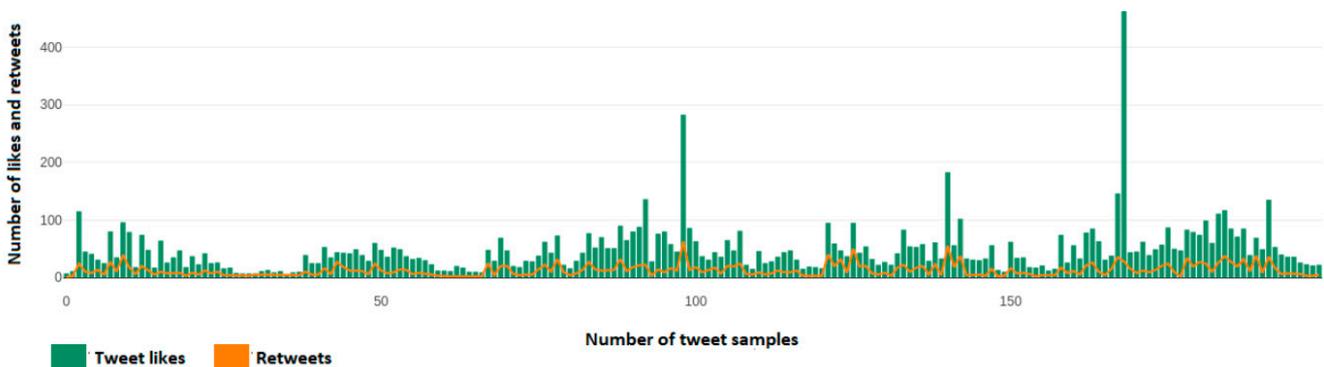


Figure 10. Number of likes and retweets per tweet for @fofigennimata (200 tweet sample).

Figure 11 depicts the number of characters per tweet of each tweet over the dynamic sample of 200 tweets taken from Kyriakos Mitsotakis's (@kmitsotakis) account. Axis y depicts the number of characters that a single tweet has, while the x axis depicts the number of each tweet over this 200-tweet sample. This feature is useful as it is an indication on how a political leader conforms with the character limitation that Twitter has.

Furthermore, a tweet that has not too many characters is likely easier for a user to memorize, thus having a direct affect over the voters.

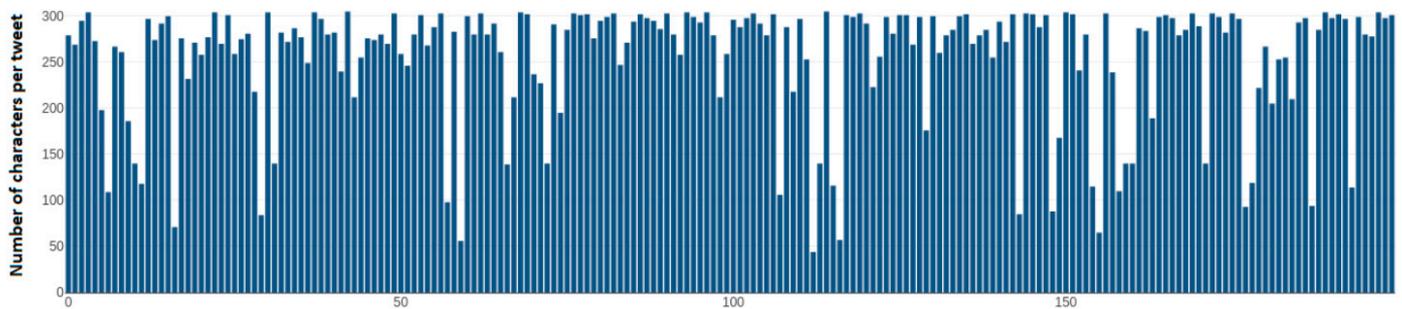


Figure 11. Text length per tweet taken from @kmitsotakis (200 tweet sample).

Figure 12 depicts the number of characters per tweet of each tweet over the dynamic sample of 200 tweets taken from Alexis Tsipras's (@atsipras) account. Axis y depicts the number of characters that a single tweet has, while the x axis depicts the number of each tweet over this 200-tweet sample. This feature is useful as it is an indication of how a political leader conforms with the character limitation that Twitter has. Furthermore, a tweet that has does not have too many characters is likely easier for a user to memorize, thus having a direct affect over the voters.

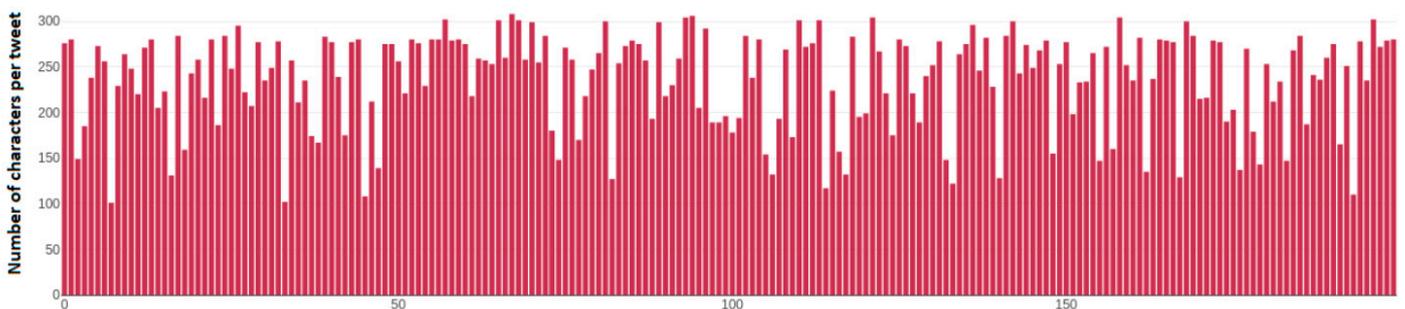


Figure 12. Text length per tweet taken from @atsipras (200 tweet sample).

Figure 13 depicts the number of characters per tweet of each tweet over the dynamic sample of 200 tweets taken from Fofi Gennimata's (@fofigennimata) account. Axis y depicts the number of characters that a single tweet has, while the x axis depicts the number of each tweet over this 200-tweet sample. This feature is useful as it is an indication on how a political leader conforms with the character limitation that Twitter has. Furthermore, a tweet that does not have too many characters is likely easier for a user to memorize, thus having a direct affect over the voters.

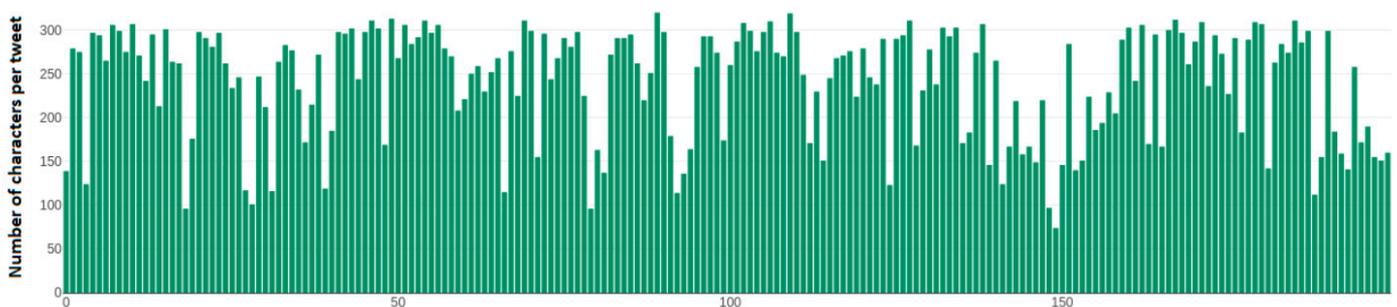
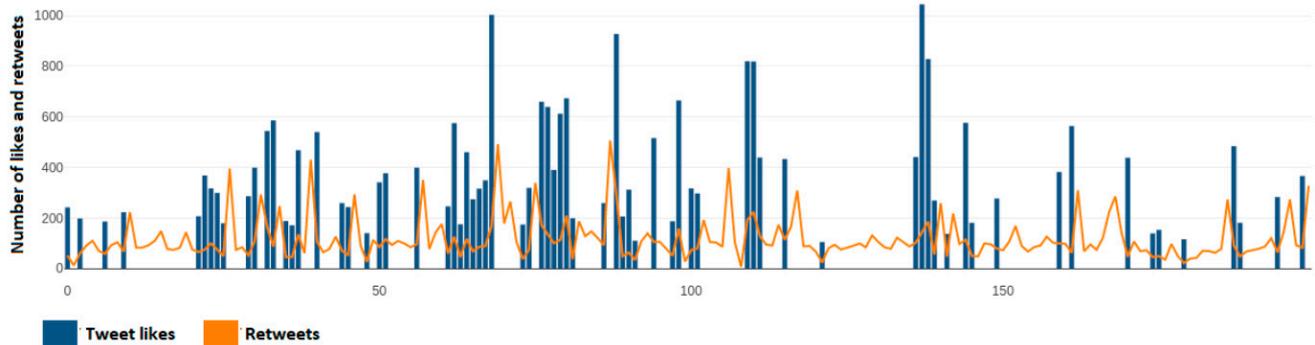


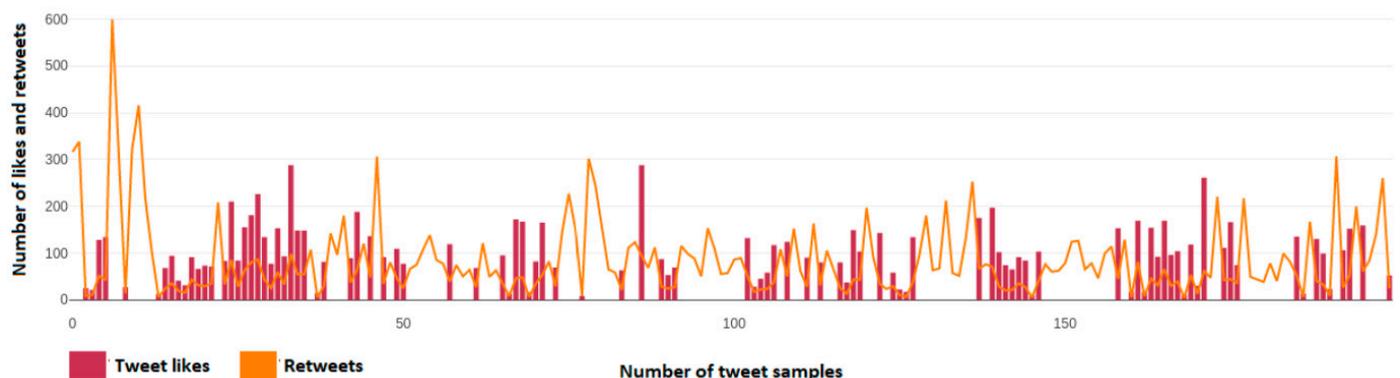
Figure 13. Text length per tweet taken from @ fofigennimata (200 tweet sample).

Figure 14 depicts the number of user reactions, likes, and retweets of each tweet from a dynamic sample of 200 tweets taken from the official New Democracy (@neademokratia) political party account on Twitter. This political party won the Parliamentary majority on the last Greek legislative elections of June 2019, since it received 39.85% of the votes (158 seats) and currently its leader is Kyriakos Mitsotakis, who is also the current Prime Minister of the Hellenic (Greek) Republic. Axis y depicts the number of likes and retweets that a single tweet has, while the x axis depicts the number of each tweet over this 200-tweet sample.



**Figure 14.** Number of likes and retweets per tweet for @neademokratia (200 tweet sample).

Figure 15 depicts the number of user reactions, likes, and retweets of each tweet from a dynamic sample of 200 tweets taken from the official Coalition of the Radical Left – Progressive Alliance (@syriza\_gr) political party account on Twitter. This political party is commonly known as SYRIZA and on the last Greek legislative elections of June 2019 received the majority of the Parliamentary opposition since it received 31.53% of the votes (86 seats), thus making it the official opposition. Currently, its leader is Alexis Tsipras, who was the former Prime Minister of the Hellenic (Greek) Republic from September 2015 until June 2019. Axis y depicts the number of likes and retweets that a single tweet has, while the x axis depicts the number of each tweet over this 200-tweet sample.



**Figure 15.** Number of likes and retweets per tweet for @syriza\_gr (200 tweet sample).

Figure 16 depicts the number of user reactions, likes, and retweets of each tweet from a dynamic sample of 200 tweets taken from the official Movement for Change (@kinimalagis) political party account on Twitter. This political party is commonly known as KINAL and on the last Greek legislative elections of June 2019 received the lesser Parliamentary opposition since it received 8.14% of the votes. Currently, its leader is Alexis Tsipras, who was the former Prime Minister of the Hellenic (Greek) Republic from September 2015 until June 2019. Axis y depicts the number of likes and retweets that a single tweet has, while the x axis depicts the number of each tweet over this 200-tweet sample.

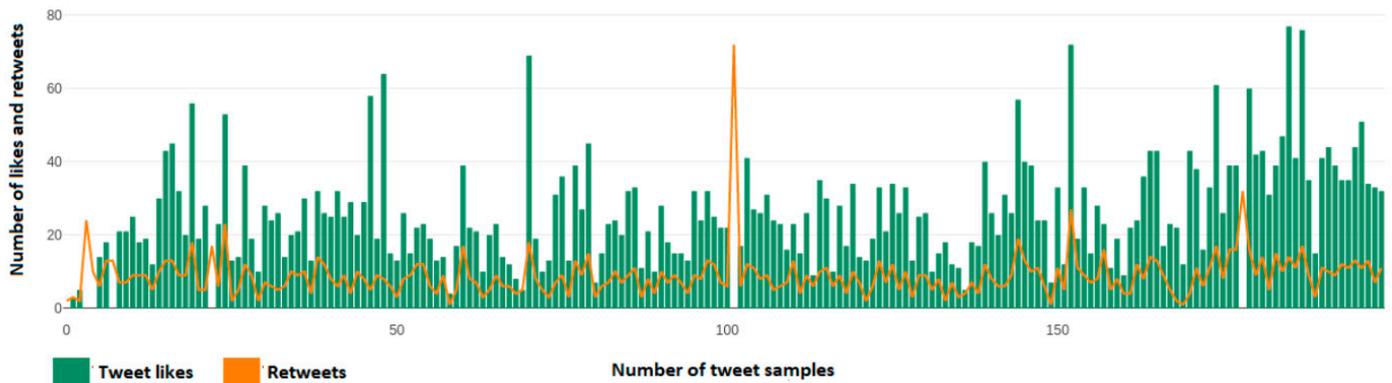


Figure 16. Number of likes and retweets per tweet for @kinimallagis (200 tweet sample).

Figure 17 depicts the number of identified tweets, which include the negative hashtag "#ΝΔ\_θελατε" from a dynamic sample of 100 tweets taken throughout Twitter. This hashtag is negative towards New Democracy party and is an indication of frustration of voters towards the policies of this particular political party. Axis y depicts the number of each tweet over the total 100 tweet sample, while the x axis depicts the date and time frequency of each tweet.

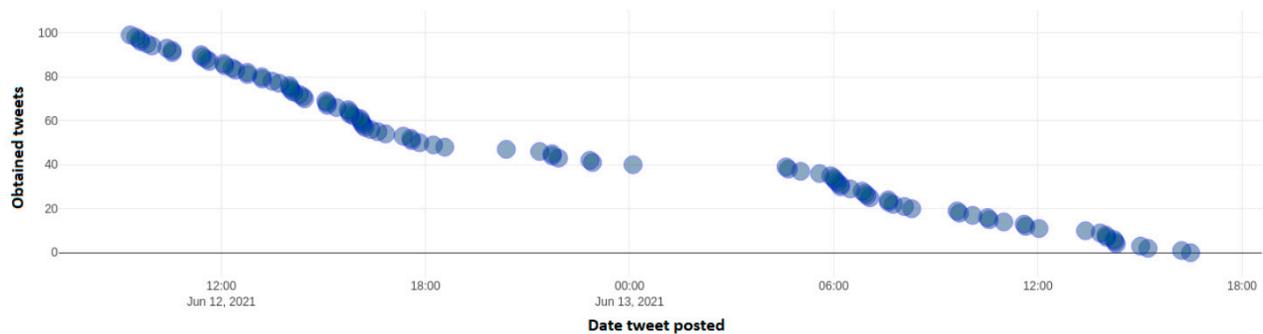


Figure 17. Mined tweets based on negative hashtag (#ΝΔ\_θελατε) (100 tweet sample) for New Democracy party.

Figure 18 depicts the number of identified tweets, which include the negative hashtag "#ΝΔ\_ξεφτίλες" from a dynamic sample of 100 tweets taken throughout Twitter and include this particular hashtag. This hashtag is negative towards New Democracy party and is an indication of frustration of voters towards the policies of this particular political party. Axis y depicts the number of each tweet over the total 100 tweet sample, while the x axis depicts the date and time frequency of each tweet.

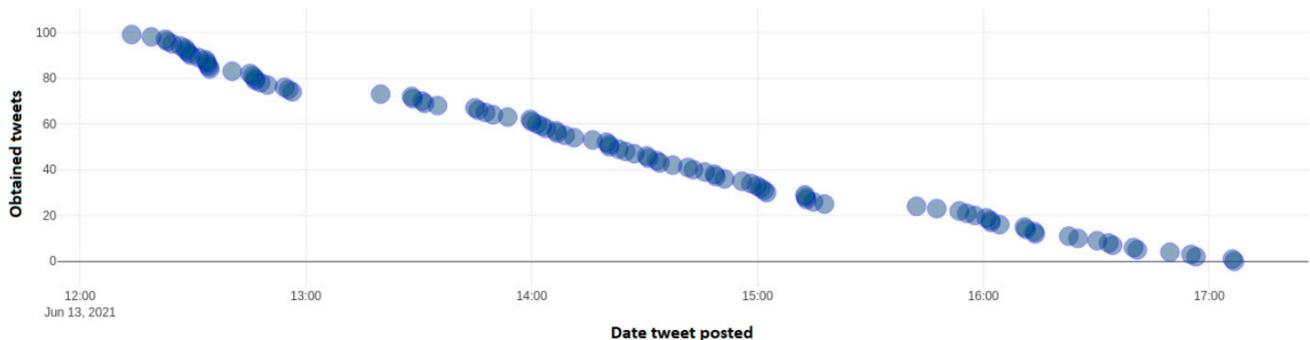
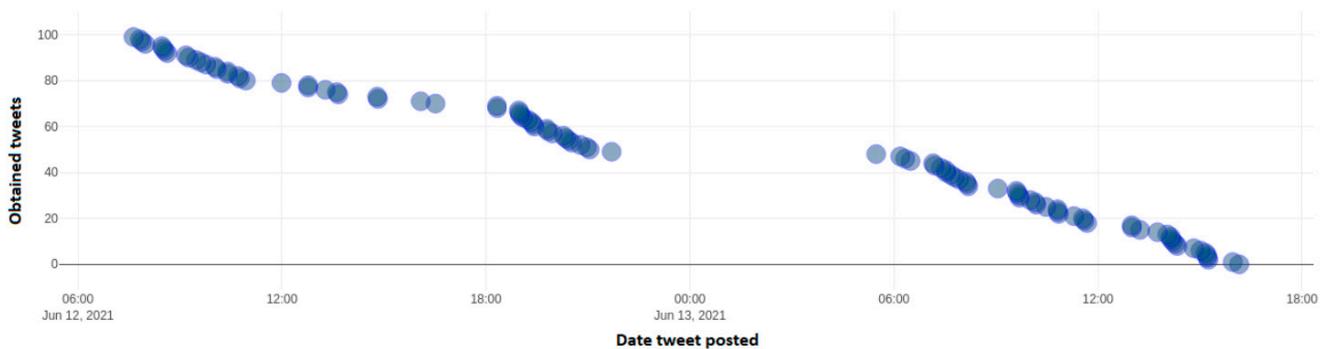


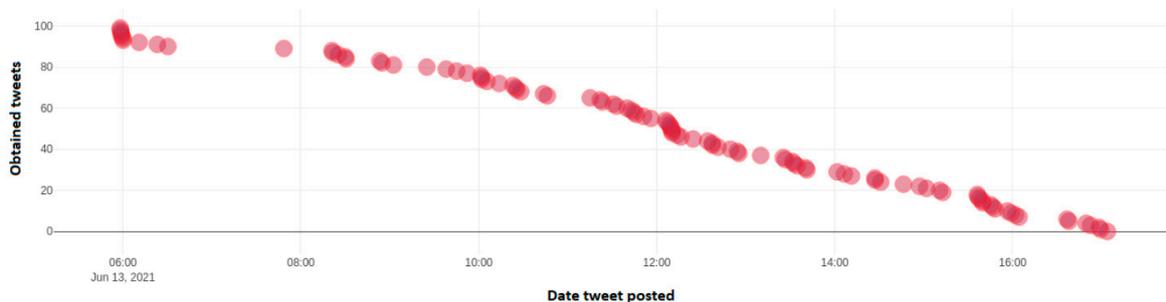
Figure 18. Mined tweets based on negative hashtag (#ΝΔ\_ξεφτίλες) (100 tweet sample) for New Democracy party.

Figure 19 depicts the number of identified tweets, which include the negative hashtag "#ΝΔ\_ρομπες" from a dynamic sample of 100 tweets taken throughout Twitter. This hashtag is negative towards New Democracy party and is an indication of frustration of voters towards the policies of this particular political party. Axis y depicts the number of each tweet over the total 100 tweet sample, while the x axis depicts the date and time frequency of each tweet.



**Figure 19.** Mined tweets based on negative hashtag (#ΝΔ\_ρομπες) (100 tweet sample) for New Democracy party.

Figure 20 depicts the number of identified tweets, which include the negative hashtag "#ΣΥΡΙΖΑ\_ξεφτίλες" from a dynamic sample of 100 tweets taken throughout Twitter. This hashtag is negative towards Coalition of the Radical Left – Progressive Alliance (SYRIZA) party and is an indication of frustration of voters towards the policies of this particular political party. Axis y depicts the number of each tweet over the total 100 tweet sample, while the x axis depicts the date and time frequency of each tweet.



**Figure 20.** Mined tweets based on negative hashtag (#ΣΥΡΙΖΑ\_ξεφτίλες) (100 tweet sample) for SYRIZA party.

Figure 21 depicts the number of identified tweets, which include the negative hashtag "#συριζωα" from a dynamic sample of 100 tweets taken throughout Twitter. This hashtag is negative towards Coalition of the Radical Left – Progressive Alliance (SYRIZA) party and is an indication of frustration of voters towards the policies of this particular political party. Axis y depicts the number of each tweet over the total 100 tweet sample, while the x axis depicts the date and time frequency of each tweet.

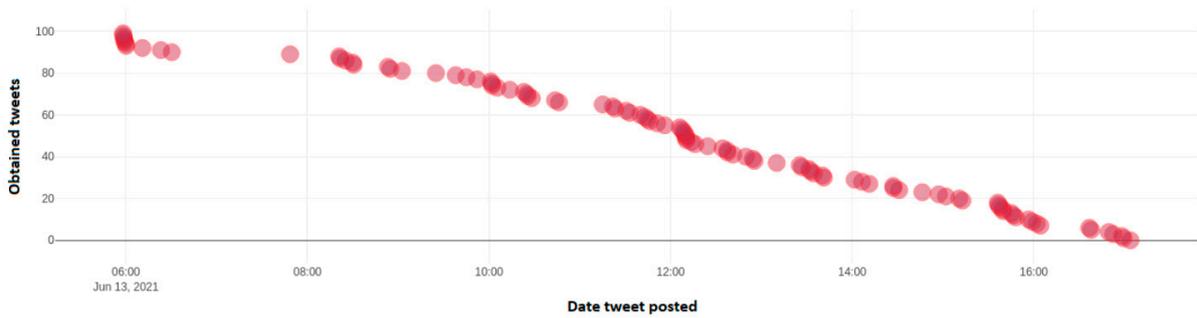


Figure 21. Mined tweets based on negative hashtag (#συριζωα) (100 tweet sample) for SYRIZA party.

Figure 22 depicts the number of identified tweets, which include the negative hashtag “#Συριζα\_απατεώνες” from a dynamic sample of 100 tweets taken throughout Twitter. This hashtag is negative towards Coalition of the Radical Left – Progressive Alliance (SYRIZA) party and is an indication of frustration of voters towards the policies of this particular political party. Axis y depicts the number of each tweet over the total 100 tweet sample, while the x axis depicts the date and time frequency of each tweet.

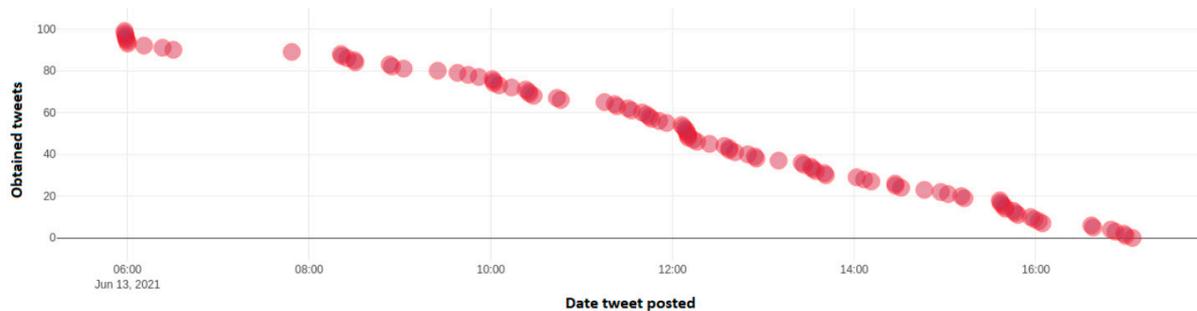


Figure 22. Mined tweets based on negative hashtag (#Συριζα\_απατεώνες) (100 tweet sample) for SYRIZA party.

For the #KINAA\_ξεφτιλες” hashtag, only 6 tweets were available, whereas for the #πανδημία\_ηληθίων there was only one, showing the small impact of the corresponding negative campaign.

Figure 23 depicts the number of identified tweets, which include the positive hashtag “#σηκώνουμε\_μανίκια” from a dynamic sample of 100 tweets taken throughout Twitter. This hashtag is positive towards Government policies of vaccination schedule in order to limit the spread of Covid-19 (SARS-CoV-2) pandemic throughout the Country. Axis y depicts the number of each tweet over the total 100 tweet sample, while the x axis depicts the date and time frequency of each tweet.

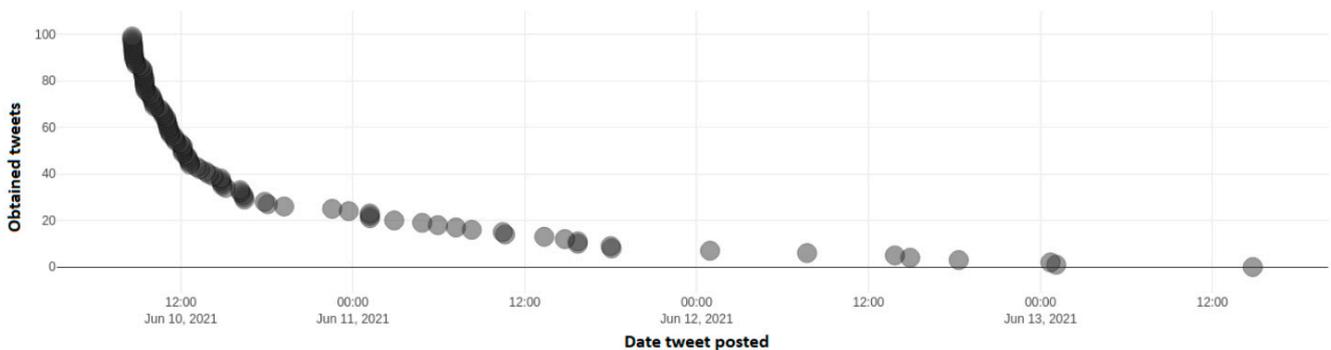


Figure 23. Mined tweets based on positive hashtag (#σηκώνουμε\_μανίκια) (100 tweet sample) for government Covid-19 restriction measures.

Figure 24 depicts the number of total subscribers per political account on Twitter namely for Kyriakos Mitsotakis (@kmitsotakis), Alexis Tsipras (@atsipras), and Fofi Gennimata (@FofiGennimata). The number of subscribers is an indication of how many fans political leaders attract and their strong presence over this particular social media. Axis y depicts the number of total subscribers, while the x axis depicts the number of political leaders.

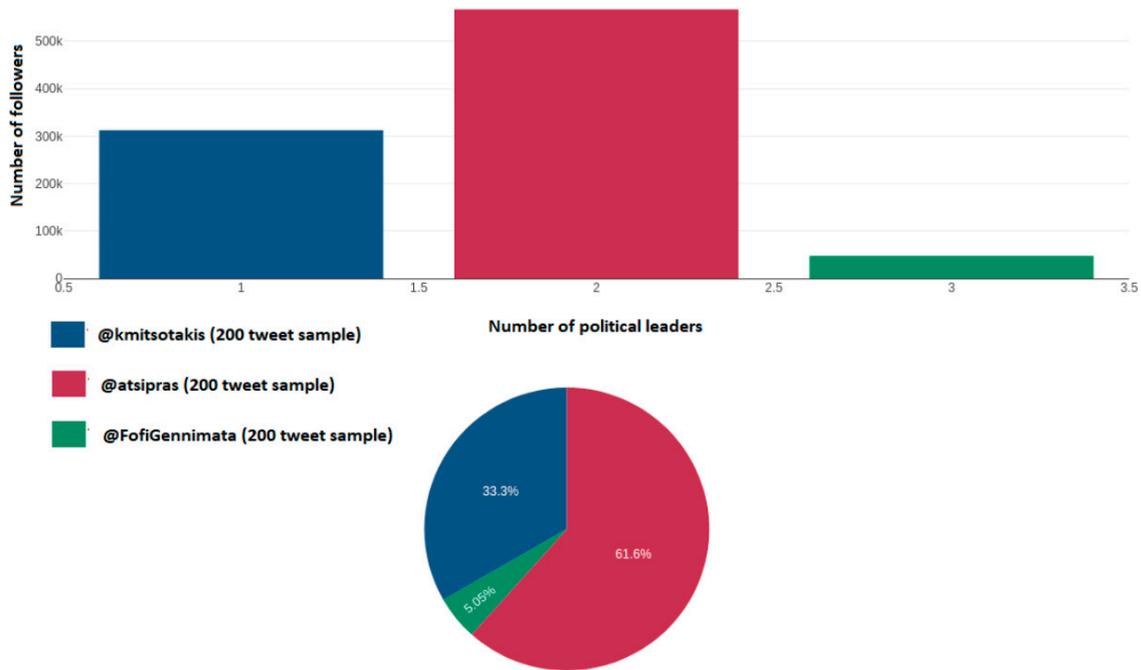


Figure 24. Number of registered subscribers per Twitter account for the top three Greek political leaders.

Figure 25 depicts the sentiment analysis per political account on Twitter namely for Kyriakos Mitsotakis (@kmitsotakis), Alexis Tsipras (@atsipras), and Fofi Gennimata (@FofiGennimata). The indications are based on emotion scores, which are Happiness (value: 3), Surprise (value: 2), Sadness (value: 1), Neutral (value: 0), Fear (value: -1), Disgust (value: -2), and Anger (value: -3). We can think of this as a fine-grained result from very positive to very negative.

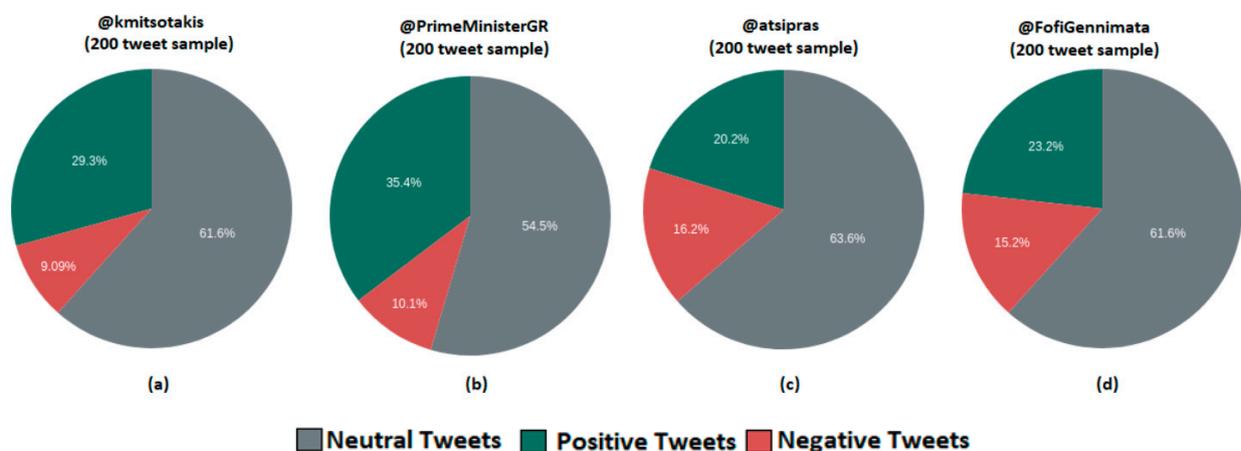
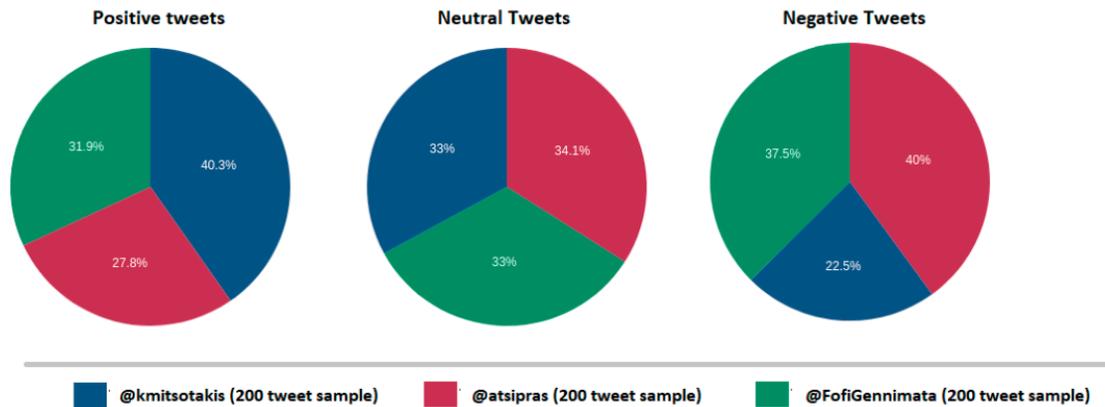


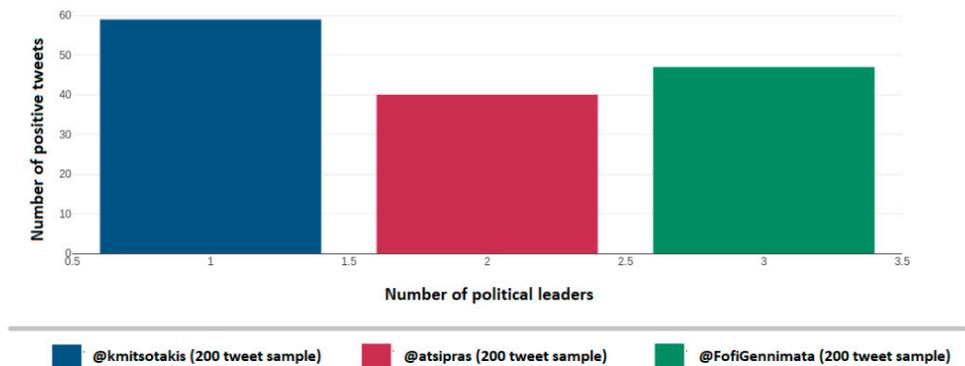
Figure 25. Sentiment analysis of tweets mined from @kmitsotakis, @primeministerGR, @atsipras, and @FofiGennimata accounts (800 tweet sample; 200 tweets per account).

Figure 26 depicts the identified sentiment analysis in a comparison form per political account on Twitter namely for Kyriakos Mitsotakis (@kmitsotakis), Alexis Tsipras (@atsipras), and Fofi Gennimata (@FofiGennimata). Neutral tweets are excluded from the graph as they do not indicate any form of emotion.



**Figure 26.** Sentiment Analysis comparison for all political leaders—600 tweet sample; 200 for each one of them (@primeministerGR is excluded).

Figure 27 depicts the identified tweets, which indicate positive mood per political account on Twitter namely for Kyriakos Mitsotakis (@kmitsotakis), Alexis Tsipras (@atsipras), and Fofi Gennimata (@FofiGennimata). Axis y depicts the number of identified positive mood tweets over the 200-tweet sample, while the x axis depicts the number of political leaders.



**Figure 27.** Positive identified tweets for the top three Greek political leaders (600 tweet sample).

Figure 28 depicts the identified tweets that indicate neutral mood per political account on Twitter namely for Kyriakos Mitsotakis (@kmitsotakis), Alexis Tsipras (@atsipras), and Fofi Gennimata (@FofiGennimata). Axis y depicts the number of identified neutral mood tweets over the 200-tweet sample, while the x axis depicts the number of political leaders.

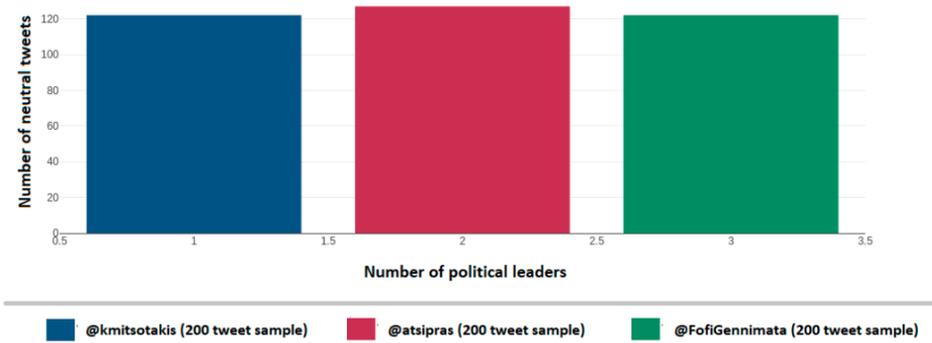


Figure 28. Neutral identified tweets for the top three Greek pollical leaders (600 tweet sample).

Figure 29 depicts the identified tweets that indicate neutral mood per political account on Twitter namely for Kyriakos Mitsotakis (@kmitsotakis), Alexis Tsipras (@atsipras), and Fofi Gennimata (@FofiGennimata). Axis y depicts the number of identified neutral mood tweets over the 200-tweet sample, while the x axis depicts the number of political leaders.

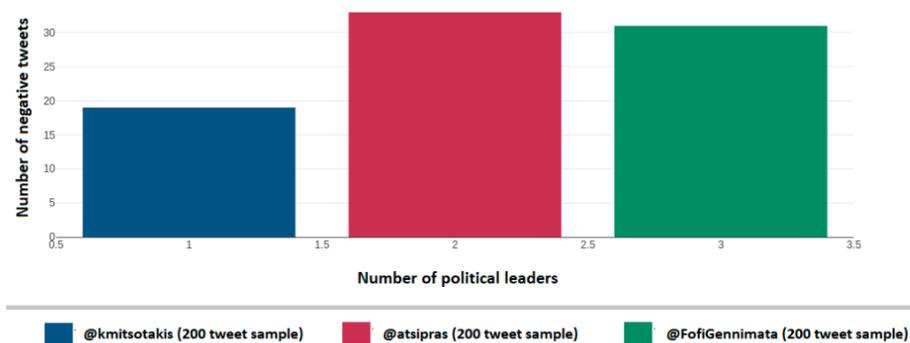


Figure 29. Negative identified tweets for the top three Greek pollical leaders (600 tweet sample).

Figure 30 presents the number of likes per tweet over the sample of 600 tweets (200 per political leader) taken from the accounts of Kyriakos Mitsotakis (@kmitsotakis), Alexis Tsipras (@atsipras), and Fofi Gennimata (@FofiGennimata). Axis y depicts the number of likes per tweet over the 200-tweet sample per political leader, while the x axis depicts the number of tweet samples taken for each political leader and not the total sample.

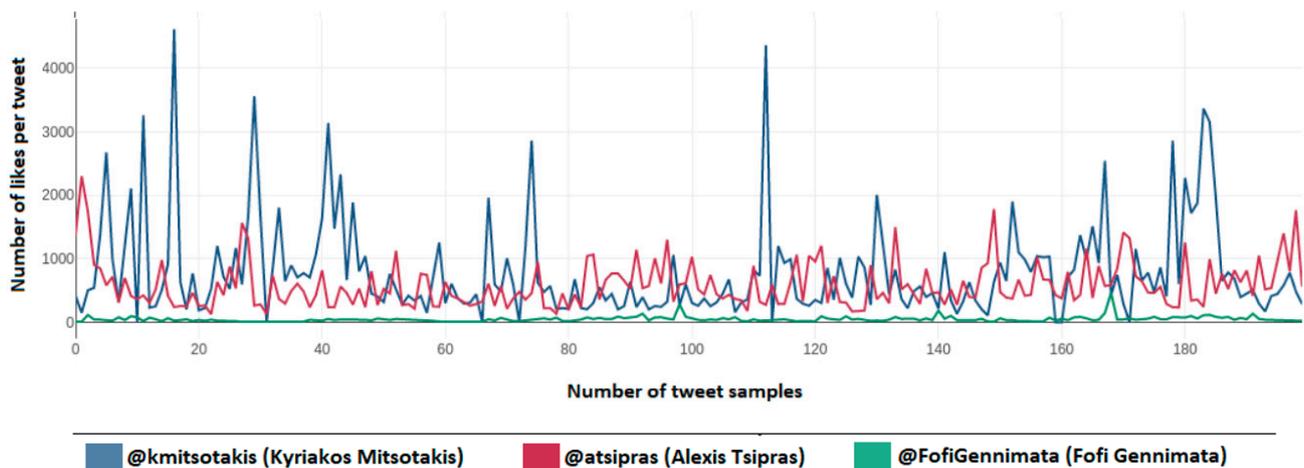
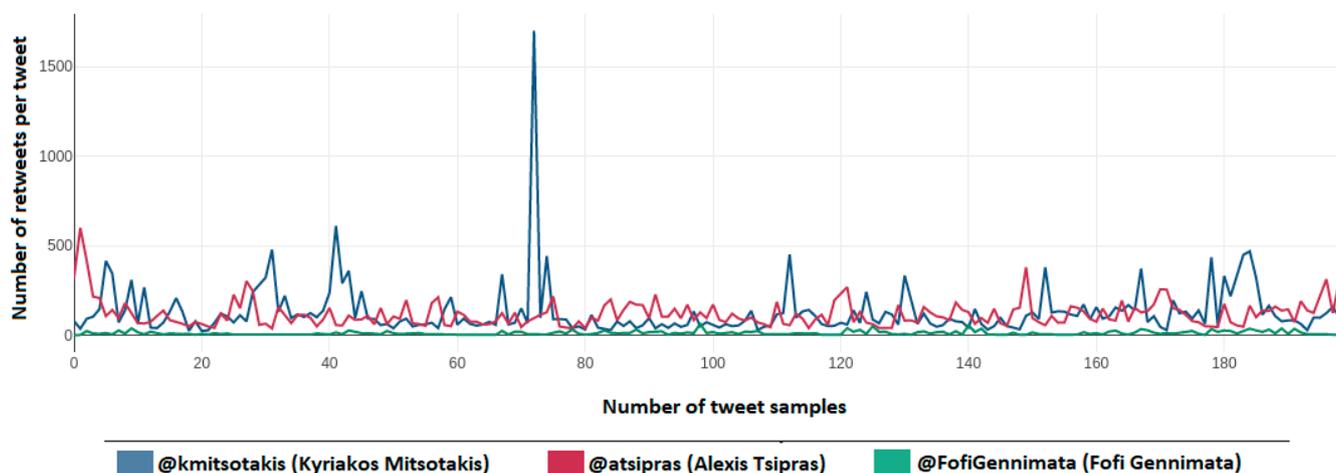


Figure 30. Comparison of user likes per posted tweet for the top three Greek political leaders—600 tweet sample; 200 for each one of them (@primeministerGR is excluded).

Figure 31 depicts the number of retweets per tweet over the sample of 600 tweets (200 per political leader) taken from the accounts of Kyriakos Mitsotakis (@kmitsotakis), Alexis Tsipras (@atsipras), and Fofi Gennimata (@FofiGennimata). Axis y depicts the number of retweets per tweet over the 200-tweet sample per political leader, while the x axis depicts the number of tweet samples taken for each political leader and not the total sample.



**Figure 31.** Comparison of user retweets per posted tweet for the top three Greek political leaders—600 tweet sample; 200 for each one of them (@primeministerGR is excluded).

From all above charts, we can observe that sentiment analysis data were pretty much close to the last Greek legislative election results, which took place on July 2019. The calculation of sentiment accuracy is produced using the following formula:

$$\frac{\text{emotion score} \times 100}{\text{Number of all matched tokens}}$$

The number of all matched tokens is the total number of emotions identified by our analyzer; emotion score is the score produced by the analyzer using sentiment lexicon. The average of identified positive or negative words indicates the emotion of the entire sentence/tweet.

Furthermore, the presentation of structured data provides us further conclusions on the user engagement that each political leader has. For example, the number of followers that each political leader has on their Twitter accounts indicates the voters that he/she can attract or that he/she has attracted more young voters rather than the elderly, given the fact that elder people do not have experience with social media, let alone if they are aware of its meaning or its existence. The number of likes and retweets indicate the engagement of voters as there is a probability that the user will press the like button when he/she likes the content of this tweet. The same thing applies with the retweets although the chances here are fairly lower (in comparison with the user likes), as the user might retweet a tweet in order to mock its content. The frequency of negative hashtags that people include in their tweets regarding political party is also a strong indicator of voter's disappointment, and this reflects as well to the image of the leader, which represents this political party. The frequency of negative and positive hashtags, which users include in their tweets regarding the government lockdown measures due to the Covid-19 pandemic, are also an indicator of voter's frustration or approval regarding the government measures and actions in response to this crisis.

## 5. Open Topics and Limitations

There are some topics that need further revision regarding our implementation. The first issue has to do with the graphs and their need for an update; it is necessary to convert these graphs into instant live graphs.

Currently, our graphs use Tweepy module in order to communicate with Twitter, and once a tweet is uploaded, it will be fetched by our application after 1 or 2 hours approximately. Regarding the validity of our sentiment analyzer, it relies entirely on the lexicon of Tsakalides et al. [50] and on the NLP capabilities of spaCy to produce results. We could update it using LSTM classification, thus creating a hybrid approach. This choice of using LSTM classification occurs due to the reason that it can learn from the sentences each time the model is trained. Although this approach would definitely be something different from the current approach and shifts towards the implementation of a highly accurate sentiment analyzer, it is not applicable to our current topic, which is the implementation of a data visualization tool for political popularity identification.

Other improvements could be made on the overall performance of the executed code, for example, using less resources and thus making it a light-weight application. AthPPA could also be updated in order to include other political parties and political entities, thus making it a poll chart application where pollsters can use as advice in order to estimate political results. Further, we have to note that our system does not detect or give any special treatment to fake accounts, and as such, might be prone to false reports.

## 6. Conclusions

In this paper, we report the first Greek language, web-based data visualization application for political tendency identification of Twitter's users.

Twitter is a useful tool to extract the sentiment of users and to predict a political result. We have identified crucial structured data such as the number of likes, re-tweets, text length, number of subscribers per account, as well as the frequency of negative hashtags that users include in their posted tweets. The number of likes and retweets allows us to observe how popular a political leader is and the number of active followers they have on Twitter. This can help us identify the fan base group of a politician, which is a strong indicator of political popularity. The same occurs for political parties that have their official accounts in Twitter, where likes and re-tweets per posted tweet are an indicator of a fan base pool, voters, or supporters. In addition to that, political parties that have official accounts in Twitter are the first source of advertising a political leader that represents it. For example, according to the Greek political standards, if a political leader, let us say, leader A, posts something on his/her account, it is likely that this post will be re-tweeted by the official Twitter account of the political party that leader A represents. Also, the number of subscribers that political leaders have on their accounts is another crucial indicator of political popularity, as it indicates the interest of users on staying tuned to the news that a political leader tweets. This includes also the negative hashtags, as users might use them as trend rather than a true source of opinion indicator. For this reason, we are focusing mostly on how frequently a tweet with a negative hashtag about a political party is posted on the Twitter rather than the emotion that these tweets imply.

Finalizing the structured data that we obtained from Twitter, the text length of a posted tweet is an indicator to what extent a politician uses Twitter, for example, small texts are an indicator of expressing an announcement or something very crucial that wants to reach many people and that will ensure that many people will read it; while long texts indicate an effort to express an opinion or to affect the thoughts of a certain group of people e.g. invocation of emotion etc. although there is high risk of being unnoticed by users as they might be bored of reading it. That is why the character limitation that Twitter provides is a crucial restriction in order to make sure that the message of the tweet will reach everyone (or at least everyone will read it). As we mentioned previously, structured data are a fairly good source of obtaining the popularity of a politician or of a political party, but there are also fake accounts that might lead us to false results. That is why we are also analyzing the text of the tweet that a politician posts on Twitter in order to identify its expressed emotion, and thus, its impact on voters. SpaCy is an efficient Natural Language Processing tool that helps us to deploy Natural Language Processing techniques on a text, and it is also an efficient commercial tool with strong community and with a good

documentation. Also, the lexicon made by Tsakalidis et al. [50] is designed specifically for political sentiment analysis. Emotion-based lexicons are efficient when it comes to political sentiment analysis, as the emotions are the main indicators that affect the opinion of a voter (e.g., anger about a tweet that a politician tweeted etc.). In the field of data visualization, the Python Dash framework provides many capabilities and features for creating graphs and charts in a web-based environment.

To conclude, this paper presented a web-based data visualization tool for political sentiment analysis; similar applications have been created with machine learning or Natural Language processing techniques, although not many of them are web-based tools that people can observe and usually exclude structured data. The research question of this paper was designed to determine how our daily activities through social media have an impact over the political landscape, especially in countries with a political crisis, like Greece. Furthermore, we can conclude that social media analytics can be proven useful for analyzing the political landscape of countries. AthPPA can also be extended to other domains beyond those of political science, such as the marketing industry, by analyzing the popularity of a commercial product. That would also need a major modification in the lexicon used in order to characterize negative and positive words based on the interests of that domain. To further our approach for sentiment analysis, polarity detection could also be used for creating document-based recommendations, as detecting the polarity of persons on specific topics would enhance the quality of the proposed recommendations [51,52].

**Author Contributions:** Methodology, N.P.; software, A.B.; supervision, H.K. and N.P.; validation, A.B.; writing—original draft, A.B.; writing—review & editing, H.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** Authors can confirm that all relevant data are included in the article.

**Acknowledgments:** The authors would like to thank Stelios Sfakiannakis for the interesting discussions during the development of this tool.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Skiba, D.J. Web 2.0: Next great thing or just marketing hype? *Nurs. Educ. Perspect.* **2006**, *27*, 212–214.
- Kondylakis, H.; Plexousakis, D. Ontology evolution in data integration: Query rewriting to the rescue. In *International Conference on Conceptual Modeling*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 393–401.
- Kontopoulos, E.; Berberidis, C.; Dergiades, T.; Bassiliades, N. Ontology-based sentiment analysis of twitter posts. *Expert Syst. Appl.* **2013**, *40*, 4065–4074, doi:10.1016/j.eswa.2013.01.001.
- Kaplan, A.M.; Haenlein, M. The early bird catches the news: Nine things you should know about micro-blogging. *Bus. Horizons* **2011**, *54*, 105–113, doi:10.1016/j.bushor.2010.09.004.
- Twitter—Statistics & Facts. Available online: <https://www.statista.com/topics/737/Twitter/> (accessed on 30 July 2021).
- Chatziadam, P.; Dimitriadis, A.; Gikas, S.; Logothetis, I.; Michalodimitrakis, M.; Neratzoulakis, M.; Papadakis, A.; Kontoulis, V.; Siganos, N.; Theodoropoulos, D.; et al. TwiFly: A Data Analysis Framework for Twitter. *Information* **2020**, *11*, 247, doi:10.3390/info11050247.
- Pak, A.; Paroubek, P. Twitter as a corpus for sentiment analysis and opinion mining. *LREc* **2010**, *10*, 1320–1326.
- Liu, B. Sentiment analysis and opinion mining. In *Synthesis Lectures on Human Language Technologies*; Morgan & Claypool Publishers: San Rafael, CA, USA, 2012.
- Liu, Y.; Huang, X.; An, A.; Yu, X. ARSA: A sentiment-aware model for predicting sales performance using blogs. In Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Amsterdam, The Netherlands, 23–27 July 2007.
- Park, S.; Ko, M.; Kim, J.; Liu, Y. The politics of comments: Predicting political orientation of news stories with commenters' sentiment patterns. In Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work, Hangzhou, China, 19–23 March 2011.
- Dergiades, T. Do investors' sentiment dynamics affect stock returns? Evidence from the US economy. *Econ. Lett.* **2021**, *116*, 404–407, doi:10.1016/j.econlet.2012.04.018.
- Kang, H.; Seong, J.Y.; Dongil, H. Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews. *Expert Syst. Appl.* **2012**, *39*, 6000–6010.

13. Prem, M.; Gryc, W.; Lawrence, R.D. Sentiment analysis of blogs by combining lexical knowledge with text classification. In Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, 28 June–1 July 2009.
14. Moreo, A.; Romero, M.; Castro, J.; Zurita, J. Lexicon-based Comments-oriented News Sentiment Analyzer system. *Expert Syst. Appl.* **2012**, *39*, 9166–9180, doi:10.1016/j.eswa.2012.02.057.
15. Tumasjan, A.; Sprenger, T.O.; Sandner, P.G.; Welpe, I.M. Predicting elections with Twitter: What 140 characters reveal about political sentiment. In Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, Washington, DC, USA, 23–26 May 2010.
16. Claster, W.B.; Cooper, M.; Sallis, P. Thailand—Tourism and conflict: Modeling sentiment from Twitter tweets using naïve Bayes and unsupervised artificial neural nets. In Proceedings of the Second International Conference on Computational Intelligence, Modelling and Simulation, Washington, DC, USA, 28–30 September 2010.
17. Kostiantyn, K.; Paradis, C.; Kerren, A. The State of the Art in Sentiment Visualization. *Comput. Graph. Forum* **2017**, *37*, 71–96, doi:10.1111/cgf.13217.
18. Britzolakis, A. Design and Development of a Web-Based Data Visualization Software for Political Tendency Identification of Twitter's Users Using Python Dash Framework. Master's Thesis, Hellenic Mediterranean University Institutional Repository (Apothesis), Iraklion, Greece, 2020.
19. Pawar, K.K.; Pukhraj, P.S.; Deshmukh, R.R. Twitter sentiment analysis: A review. *Int. J. Sci. Eng. Res.* **2015**, *6*, 957–964.
20. Pozzi, F.A.; Fersini, E.; Messina, E.; Liu, B. *Sentiment Analysis in Social Networks*; Morgan Kaufmann: Burlington, MA, USA, 2016.
21. Bo, P.; Lee, L. Opinion Mining and Sentiment Analysis. *Found. Trends Inf. Retr.* **2008**, *2*, 1–135.
22. Nausheen, F.; Sayyada, H.B. Sentiment analysis to predict election results using Python. In Proceedings of the 2nd International Conference on Inventive Systems and Control (ICISC), Coimbatore, India, 19–20 January 2018.
23. Koltsova, O.Y.; Alexeeva, S.; Kolcov, S. An opinion word lexicon and a training dataset for Russian sentiment analysis of social media. In Proceedings of the International Conference "Dialogue 2016", Moscow, 1–4 June Russia 2016.
24. Ankitkumar, D.; Badre, R.; Kinikar, M. A Survey on Sentiment Analysis and Opinion Mining. *Int. J. Innov. Res. Comput. Commun. Eng.* **2014**, *2*, 6633–6639.
25. Caruccio, L.; Desiato, D.; Polese, G. Fake account identification in social networks. In Proceedings of the IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 10–13 December 2018.
26. Calderon, N.A.; Fisher, B.; Hemsley, J.; Ceskavich, B.; Jansen, G.; Marciano, R.; Lemieux, V. Mixed-initiative social media analytics at the World Bank: Observations of citizen sentiment in Twitter data to explore "trust" of political actors and state institutions and its relationship to social protest. In Proceedings of the IEEE International Conference on Big Data (Big Data), Santa Clara, CA, USA, 29 October–1 November 2015.
27. Ferrara, E.; Zeyao, Y. Quantifying the effect of sentiment on information diffusion in social media. *PeerJ Comput. Sci.* **2015**, *1*, e26.
28. Yaqub, U.; Chun, S.A.; Atluri, V.; Vaidya, J. Sentiment based Analysis of Tweets during the US Presidential Elections. In Proceedings of the 18th Annual International Conference on Digital Government Research, Staten Island, NY, USA, 7–9 June 2017.
29. Wei, G.; Sebastiani, F. Tweet sentiment: From classification to quantification. In Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Paris, France, 25–28 August 2015.
30. Bouazizi, M.; Tomoaki O. Sentiment analysis in Twitter: From classification to quantification of sentiments within tweets. In Proceedings of the IEEE Global Communications Conference (GLOBECOM), Washington DC, USA, 4–8 December 2016.
31. Baumgarten, M.; Mulvenna, M.; Rooney, N.; Reid, J. Keyword-Based Sentiment Mining using Twitter. *Int. J. Ambient. Comput. Intell.* **2013**, *5*, 56–69, doi:10.4018/jaci.2013040104.
32. Zervoudakis, S.; Marakakis, E.; Kondylakis, H.; Goumas, S. OpinionMine: A Bayesian-based framework for opinion mining using Twitter Data. *Mach. Learn. Appl.* **2021**, *2*, 100018.
33. Azizan, A.; Jamal, N.N.S.A.; Abdullah, M.N.; Mohamad, M.; Khairudin, N. Lexicon-Based Sentiment Analysis for Movie Review Tweets. In Proceedings of the 1st International Conference on Artificial Intelligence and Data Sciences (AiDAS), Ipoh, Malaysia, 19 September 2019.
34. Bhoir, P.; Shilpa, K. Sentiment analysis of movie reviews using lexicon approach. In Proceedings of the IEEE International Conference on Computational Intelligence and Computing Research (ICIC), Madurai, India, 10–12 December 2015.
35. Mandal, S.; Sumit, G. A Lexicon-based text classification model to analyse and predict sentiments from online reviews. In Proceedings of the International Conference on Computer, Electrical & Communication Engineering (ICCECE), Kolkata, India, 16–17 December 2016.
36. Caruccio, L.; Deufemia, V.; Polese, G. Mining relaxed functional dependencies from data. *Data Min. Knowl. Discov.* **2020**, *34*, 443–477, doi:10.1007/s10618-019-00667-7.
37. Caruccio, L.; Deufemia, V.; Naumann, F.; Polese, G. Discovering relaxed functional dependencies based on multi-attribute dominance. *IEEE Trans. Knowl. Data Eng.* **2020**, *1*. Doi: 10.1109/TKDE.2020.2967722.
38. Breve, B.; Caruccio, L.; Cirillo, S.; Deufemia, V.; Polese, G. Dependency Visualization in Data Stream Profiling. *Big Data Res.* **2021**, *25*, 100240, doi:10.1016/j.bdr.2021.100240.
39. Boia, M.; Faltings, B.; Musat, C.; Pu, P. A ) Is Worth a Thousand Words: How People Attach Sentiment to Emoticons and Words in Tweets. In Proceedings of the International Conference on Social Computing, Alexandria, VA, USA, 8–14 September 2013.

40. Manuel, K.; Indukuri, K.V.; Krishna, P.R. Analyzing Internet Slang for Sentiment Mining. In Proceedings of the International Conference on Information Technology for Real World Problems, Warangal, India, 9–11 December 2010, pp. 9–11.
41. Akcora, C.G.; Bayir, M.A.; Demirbas, M.; Ferhatosmanoglu, H. Identifying breakpoints in public opinion. In Proceedings of the First Workshop on Social Media Analytics, Washington, DC, USA, 25–28 July 2010.
42. Kydros, D.; Argyropoulou, M.; Vrana, V. A Content and Sentiment Analysis of Greek Tweets during the Pandemic. *Sustainability* **2021**, *13*, 6150, doi:10.3390/su13116150.
43. Karageorgou, I.; Liakos, P.; Delis, A. Just-in-Time Sentiment Analysis for Streamed Data in Greek. In *Next-Gen Digital Services*; Springer International Publishing: Cham, Switzerland, 2021; pp. 249–263.
44. Giannakis, S.; Valavani, C.; Alexandris, C. A Sentiment Analysis Web Platform for Multiple Social Media Types and Language-Specific Customizations. In Proceedings of the International Conference on Human-Computer Interaction Washington, DC, USA, 24–29 July 2021; pp. 318–328.
45. Zhou, X.; Tao, X.; Yong, J.; Yang, Z. Sentiment analysis on tweets for social events. In Proceedings of the 2013 IEEE 17th International Conference on Computer Supported Cooperative Work in Design (CSCWD), Whistler, BC, Canada, 27–29 June 2013; pp. 557–562.
46. Rezapour, R.; Wang, L.; Abdar, O. Identifying the overlap between election result and candidates' ranking based on hashtag-enhanced, lexi-con-based sentiment analysis. In Proceedings of the IEEE 11th International Conference on Semantic Computing (ICSC), San Diego, CA, USA, 30 January–1 February 2017.
47. Ramteke, J.; Shah, S.; Godhia, D.; Shaikh, A. Election result prediction using Twitter sentiment analysis. In Proceedings of the International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 26–27 August 2016.
48. Sahu, K.; Bai, Y.; Choi, Y. Supervised Sentiment Analysis of Twitter Handle of President Trump with Data Visualization Technique. In Proceedings of the 10th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 6–8 January 2020.
49. Britzolakis, A.; Kondylakis, H.; Papadakis, N. A Review on Lexicon-Based and Machine Learning Political Sentiment Analysis Using Tweets. *Int. J. Semant. Comput.* **2020**, *14*, 517–563, doi:10.1142/s1793351x20300010.
50. Tsakalidis, A.; Papadopoulos, S.; Voskaki, R.; Ioannidou, K.; Boididou, C.; Cristea, A.I.; Liakata, M.; Kompatsiaris, Y. Building and evaluating resources for sentiment analysis in the Greek language. *Lang. Resour. Evaluation* **2018**, *52*, 1021–1044, doi:10.1007/s10579-018-9420-4.
51. Stratigi, M.; Kondylakis, H.; Stefanidis, K. Fairness in group recommendations in the health domain. In Proceedings of the IEEE 33rd International Conference on Data Engineering (ICDE), San Diego, CA, USA, 19–23 April 2017; pp. 1481–1488.
52. Stratigi, M.; Kondylakis, H.; Stefanidis, K. FairgreCs: Fair group recommendations by exploiting personal health information. In Proceedings of the International Conference on Database and Expert Systems Applications, Regensburg, Germany, 3–6 September 2018; pp. 147–155.