Journal of Business and Economics, ISSN 2155-7950, USA
December 2021, Volume 12, No. 12, pp. 1184-1201
Doi: 10.15341/jbe(2155-7950)/12.12.2021/004

http://www.academicstar.us

Twitter Sentiment Analysis During Covid: The Case of Tourist Destination Crete Greece in Summer of 2021

Georgios Kokkinis
(International Hellenic University, Greece)

Abstract: Tourism has been hit hard by Covid-19 for the past two years. Large traditional destinations in the sector have lost a very large percentage of their turnover. The way of choosing a destination has changed, the selection of the Covid free destination has been included in the criteria, as well as the respective legislation for restrictive measures both in the place of vacation and the health protocols of the country of departure.

Compared to traditional advertising, word of mouth (WOM) advertising has impressive advantages, such as significantly lower costs and much faster dissemination, and this is especially true with the popularity of online social networks (eWoM). However, eWom should not be considered as a substitute to other promotional methods, rather than a major advertising medium.

This study has to do with the sentiment analysis on social media and specifically through Twitter, regarding one of the largest summer tourist destinations in Greece, Crete during summer of 2021.

We observe that although Covid 19 has influenced the tourism sector, tourist customers express themselves positively about the tourism destination, even without any organization interacting with the user to manage and meet his needs, showing that even if they did not succeed to choose a tourist package in this place this year, most likely they will do it when the health conditions allow it

Key words: sentiment analysis, electronic word of mouth, Covid-19, vacations, Crete, Greece
JEL codes: M031, Z033

1. Introduction

New technologies have made their presence felt in almost the whole range of life and in combination with the management and extraction of information they have become a very useful tool so that companies can do a short term - and not only - planning.

Sentiment analysis is one such a tool. It is not a tool that will be able to formulate business policy and to develop a strategic roadmap, at least for the time being, but a tool that will show whether the customer is positively or negatively affected and what impact the decisions made by the company takes, almost in real time. Sentiment analysis can be used as a tool to quickly and economically measure the views and feelings of those interested in making at least an initial assessment of decisions.

Social Media today are especially popular in its online format. This is because the Internet is full of people

Georgios Kokkinis, Dr., International Hellenic University; research areas: marketing, strategy, and sentiment analysis. E-mail: kokkinis@mkt.teithe.gr & kokkinis@ihu.gr.

who want to meet other people, develop friendships or business alliances, look for work, get involved and share experiences about their hobbies and interests. In terms of business strategy, Social Media are referred to as a means of creating a corporate image, information, communication and relationships with their customers and has incredible power. Social networks are so popular with such a big audience that the view that we are experiencing the “Social Media revolution” it is not accidental or exaggerated. In addition, social networks are divided into direct networks, which provide access to a user’s profile without the need for his consent, and indirect networks, in which it is necessary for one user to allow the other to connect with his profile.

The former are mainly used for advertising and marketing and the latter produce word-of-mouth communication, which leads to the formation of consumer behavior. The huge amount of content on social media makes it impossible to edit it manually. Therefore, automatic sentiment analysis has become very popular.

To achieve this goal it was chosen to use the technical analysis of emotions from data extracted Tweets in a period of summer of 2021 and analyzing the text itself, here it is worth noting that this research was conducted in a period of time when there are health restrictions and flight protocols that act as a deterrent to travel from other countries.

2. Literature Review

2.1 Sentiment Analysis

Automated emotion analysis is the computational understanding of one’s position, attitude, or view toward an entity, person, or subject 错误!未找到引用源。. Emotion analysis aims to create predictive models for emotions based on annotated data sets where learning is automated 错误!未找到引用源。. This approach creates a set of features of each text entry in which certain aspects or word frequencies are quantified to train standard machine learning tools and interpret them against annotated reference texts.

According to studies 错误!未找到引用源。, extracting information from twitter is quite difficult since tweets are short messages with many slang words and spelling mistakes, so it is suggested to analyze emotions at the sentence level. This is done in 4 phases:

1) Creating a Database
2) Pre-edit Tweets
3) Export features
4) Classification of emotions

2.2 Description & Definitions

Emotion analysis aims to determine the polarity of emotions conveyed through a piece of text in relation to a particular entity 错误!未找到引用源。. Big data of personal emotions is available on the internet, which was very valuable for business intelligence applications such as product search engines 错误!未找到引用源。.

One study 错误!未找到引用源。 shows that most modern studies approach emotion analysis as a classification problem. Depending on the detail of the text areas, emotion analysis can be performed at three levels, namely document level, sentence level and point of view.

2.2.1 Emotion Analysis Levels

According to Hu 错误!未找到引用源。 the levels of emotion analysis are 3 and a brief reference to each is shown below:

a) Document-level emotion analysis: The goal at this level is to determine if an entire case file conveys a positive or negative emotion

b) Sentence level emotion analysis: The goal at this level concerns the sentences and decides regardless of whether each sentence expressed a positive, negative or neutral emotion. Neutral conclusion generally does not imply any conclusion. This level of examination is consistently related to the subjectivity device, which recognizes propositions (called objective propositions) that express data from propositions (called subjective propositions) that express subjective perspectives and emotions.

c) Emotion level analysis: The goal at this level is to draw a conclusion through the tweet, to examine the written text from different angles. The perspective level was previously called the labeling level (component-based emotion extraction and collapse). Instead of looking at what is being written (files, sections, sentences, layouts, or expressions), the angle level specifically takes a look at the case itself. It depends on the probability that an evaluation includes a conclusion (positive or negative) and a goal (emotion).

2.2.2 Emotion Analysis Approaches

In this section, an overview of key approaches to emotion analysis applications will be provided. The approaches will be divided based on the methodologies used. A key criterion for grouping approaches is the way in which they manage features for language modeling. Figure 1 shows the hierarchy of different techniques. Techniques can be divided into two main categories, those that use machine learning (supervised and unsupervised) and those that use semantic orientation techniques (unsupervised).

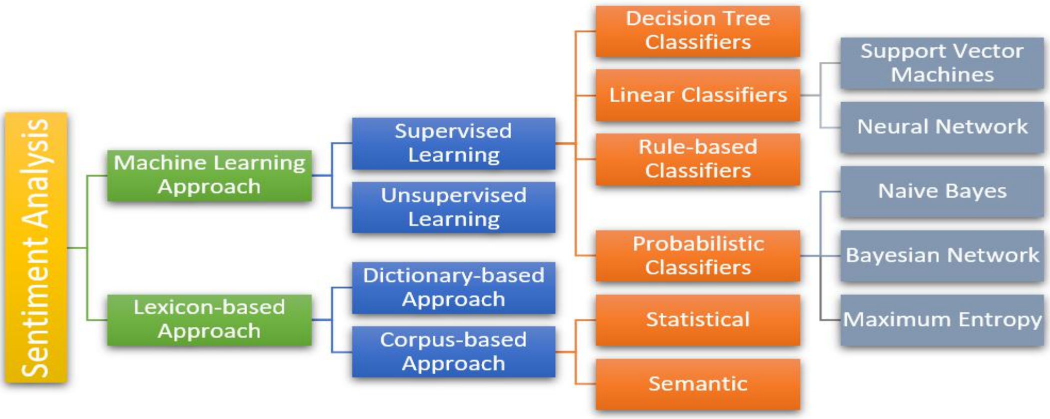


Figure 1 Approaches of Sentiment Analysis

2.2.3 Learning and Machine Learning

Learning: The use of data that one receives from the environment (e.g., audio, visual signals) so that it gradually improves when performing a function. Learning is a key component of intelligence.

Examples of functions:

• Understanding the language

•Speech production

• Identification of persons and objects

• Strategy development in different situations (e.g., games, social behavior)

Machine Learning: The use of data by an algorithm that is executed on a computer machine so that it is gradually improved during the execution of a function. The functions are equivalent to those that fall into human intelligence, such as:

• Mechanical language processing and speech production (natural language processing and understanding)

• Mechanical pattern recognition

• The development of strategy in various situations (e.g., games, space navigation)

Machine learning techniques are divided into three main subcategories, supervised, unsupervised, and learning enhancement techniques.

Supervised Learning Algorithms: Templates are used whose category is known in advance. The algorithm uses the input patterns together with the desired output value (target). As long as the algorithm does not produce the desired output for each template, it adjusts its internal parameters in such a way as to give the desired output for each template.

Unsupervised Learning Algorithms: The number of categories is not known in advance and must be determined during training. The algorithm uses only input templates to train without using targets. It regulates its internal parameters, in such a way as to find similar characteristics in our data (clustering) and to determine the distribution of our data in the input space.

Reinforcement learning: The algorithm uses only the input patterns without knowing if it is going well or not, except at the end. Only then does it know if he has achieved his goal and uses this information to improve next time **错误!未找到引用源。**.

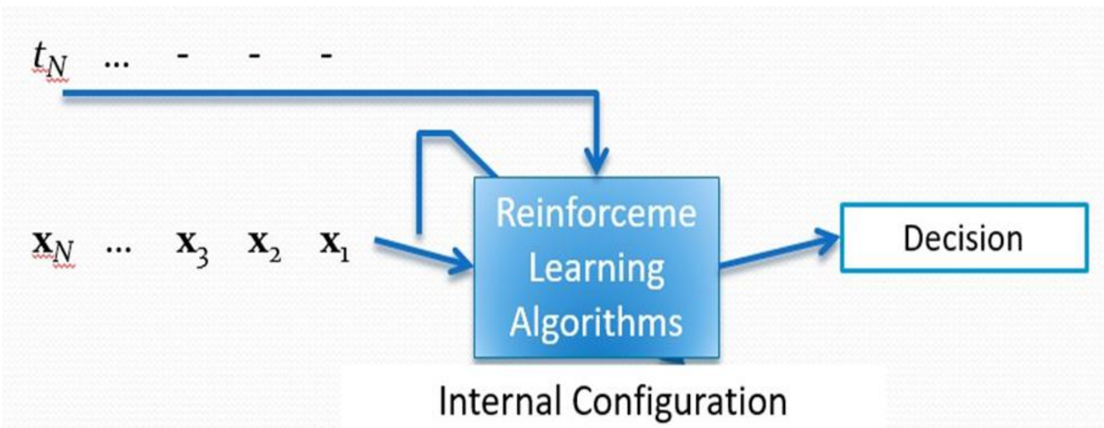


Figure 2 Reinforcement Learning

To check how well the Classifier generalizes:

• We divide all our data (standards) into Training Standards (Training Set) and Recall Standards (Test set) in a ratio of 70%-30%.

• We train it with all the Standards of Education.

• After training, we test the ability to generalize on unknown Recall Standards.

• The Revocation Standards we use for auditing must not have been used during training.

• We calculate the training error and the control error.

Advantages disadvantages of machine learning

Advantages: Traditional Machine Learning gives excellent results in well-defined problems with fixed rules (e.g., chess, diagnosis of defects in industrial components or machines, diagnosis of diseases, etc.) and if it is it is possible to use experts to extract the necessary rules.

Disadvantages of the traditional approach are:

• Every problem requires new rules

• Lack of flexibility.

• The inability to adapt to new data

• The difficulty of extracting all the rules. In many cases, it is not easy to find the right “expert” to provide the rules, or the rules are complex, or they are extremely difficult to formulate.

2.2.4 Semantic Orientation

These techniques are based on the assumption that the semantic orientation (positive, negative or neutral) of a tweet is determined by the orientation of the component of the individual terms (words or phrases). The orientation of each word or the orientation of a package of words (phrase) is extracted and a grade is given in this package with a positive or negative sign, at the end we find the algebraic sum of the grades and a conclusion is drawn. Depending on how this calculation is done, they are divided into two subcategories: semantic orientation based on texts (corpus-based) and lexicon-dictionary based.

1) Text-Based Semantic Orientation (corpus-based)

In this kind of approach we accept that when a word appears too often in texts wanting to give a positive sign, we will enter it in the words with a positive orientation (e.g., Superb) while on the contrary a word that appears in many texts wanting to give a negative sign in the negative (e.g., ugly). In order to accept these words as findings, we must first record several texts and analyse them ourselves (without the use of the computer program) at the same time classifying the words in question in statistical tables, trying to draw a safe conclusion. A major difficulty is that a large collection of texts is required to reliably determine the orientation of each word. A major difficulty is that a large collection of texts is required to reliably determine the orientation of each word. This means that enough texts will have to be collected so that the sample is representative (in one approach) and the statistics are reliable.

Another difficulty we encounter is the slang language that is usually used when composing a tweet, the irony that can characterize the text as well as the “not” ending that users usually write at the end of the sentence that completely changes the result.

2) Lexicon-based Semantic Orientation

The lexicon-based method assumes that the emotional orientation of a text can be inferred from the emotional orientation of its individual words and phrases. In contrast to machine learning-based methods, the dictionary-based approach does not require training of a classifier. Instead, it uses emotion dictionaries to convey the emotion of emotionally charged words in the text, there are many such free dictionaries to use such as WordNet, SenticNet and SentiWordNet. The performance of a dictionary-based Emotion Analysis method is usually determined by the type of emotions dictionary and the sentiment detection algorithms, the algorithm that used to locate the emotionally charged words in the text and calculate the emotion **错误!未找到引用源。**. For Twitter, “the most common practice is to use a built-in dictionary emotion along with a simple keyword matching algorithm” **错误!未找到引用源。**.

3) Advantages and Disadvantages of Semantic Orientation

The advantage of these techniques is that they do not require the existence of marked data, as in non-supervised learning techniques. It is also quite easy to implement models that utilize these techniques. This is one of the main reasons for their popularity, especially in the early years of emotion analysis research.

On the other hand, semantic orientation methodologies have been observed to be biased in favor of “positive” orientation. The reason is that people tend to use “positive” words more often when expressing themselves. This means that in a text, the “positive” words are usually more in number than the “negative” and the orientation of a text is not real.

2.3 Crete as a Tourist Destination

The Region of Crete is located to the southernmost part of the Aegean Sea in a distance of about 160 km from the mainland. Crete borders with the Cretan Sea in the North and with the Libyan in the South, while it is surrounded by a number of small islands. The total area of the island is 8,336 km2, covering 6.3% of the total Greek territory and its capital is the city of Heraklion. Crete is the largest island in Greece and the 2nd largest in the eastern Mediterranean after Cyprus.

According to the Bank of Greece, the characteristics and the tourist infrastructure in Crete during the year 2019-2020 are the following **错误!未找到引用源。**.

2.3.1 Transport Infrastructure of the Region of Crete

• 3 airports, Heraklion Airport “Nikos Kazantzakis”, Chania Airport “Ioannis Daskalogiannis” and Sitia International Airport “Vitsentzos Kornaros” with 4,455,810 flights per year, a percentage of 21% of all flights to Greece

• 2 main ports are the Port of Heraklion and the Port of Souda (Chania). The ports are mainly for passengers, with ferry connection between Piraeus and the Aegean islands. Also, they are important stations for Cruise Ships, in total in 2019 there were 4,930,407 berths of ships and 397 cruise ships.

• 4 smaller ports. The port of Rethymno, the port of Kissamos, Ag. Nikolaou and Sitia.

In 2020, Crete had 1,619 hotel units that included 96,367 rooms for 187,599 visitors per day, a percentage that corresponds to 22% of hotels in Greece. There are also 24,142 rooms for rent in Crete for 55,758 visitors per day, a percentage of 13% of the total territory. Also available for short term rental 17,623 houses and apartments via Airbnb. Finally in the region of Crete there are currently 15 Campings for 760 tents and caravans and 8 marinas with 1,973 mooring places.

Approximately 5,288,000 tourists arrived in Crete from abroad in 2019, a percentage of about 16% of the total number of tourists in Greece.

Given that the region of Crete has a revenue of tourism of about 3.6 billion (21% of revenues in Greece), we understand that it is one of the leaders of tourist destination of the country.

2.4 WoM and Social Media

Word of mouth (WOM) has been recognized for many years as an important source of influence on what people know, feel and do. In his first book, Marketing Management, Kotler acknowledges that “advertising is one of the many sources of influence on a person’s behavior, and perhaps the least important, in terms of sources of influence such as one person’s colleagues and personal observation” **错误!未找到引用源。**.

Electronic word of mouth is a relatively new concept and has two closely related terms. The term electronic word of mouth (eWOM) is often used interchangeably with viral marketing or buzz marketing. Although the degree of relationship between these terms is strong, there are major fundamental differences.

Derived from the word “virus” viral marketing aims to spread (like the virus). According to Kirby and Marsden, their definition is as follows: “The promotion of a company or its products and services through a persuasive message designed to spread, usually online, from person to person” (Justin Kirby & Paul Marsden, 2006).

Buzz marketing is a penetrating promotion process, not only in face-to-face context, but in the media, and is defined according to Kirby and Marsden as follows: “The promotion of a company or its products and services through well-designed to make people and the media speak positively about that company, product or a service”

(Justin Kirby & Paul Marsden, 2006).

3. The Study: Sentiment Analysis Using Twitter Data

3.1 Data Collection

The research in this work, was done with a piece of programming language in Python. Python is an interpreted, object-oriented, high-level programming language with dynamic semantics.

Python is a great tool for knowledge mining and is widely used and intertwined with data science. Researchers have at their disposal a plethora of libraries for several implementations and data processing. It approaches with ease and speed the discovery of patterns and correlations based on performance evaluation, as well as data reduction, two problems that are crucial in the implementation of correlation analysis in textual data, i.e., the issue addressed by the implementation to be carried out.

The data used are csv files. In the created code, the tweets are read and integrated in a Data Frame in the form of a table, through the read_csv command. The first column of the table displays the emotion that categorizes the tweet text. In the second column, there are all the tweets with their text in natural language. The table has as many rows as the tweets, each tweet corresponds to one line. If a data file contains more than the two columns described above, then these columns are removed with the drop_columns command.

3.2 Data Description

With this tool, 24,214 tweets were collected from Twitter, from relevant searches via hashtags that mention the word “Crete” and keywords “Vacation”, “Travel” and “Tourism”. These 24,214 Tweets collected in 11 weeks, from May 28 to August 13 are tweets from various languages of the world that have not been processed or cleaned.

3.3 Data Cleaning

The text data, now in the second column, is a continuous text of words, as written by Twitter users. In this form it is not easy to analyze their information and it is necessary to edit the text. The following are the steps taken in the word processing stage:

1) Tokenization: The fragmentation of a string into words. Terms can be any character, such as word, number, whole sentence, punctuation, or symbols (.tokenize () method).

2) Convert to lowercase: Some characters or even words may be in uppercase, in order to achieve the emotional analysis of the words, all the characters have been converted to lowercase and to uppercase Latin characters. The conversion was completed using the python .lower () string method

3) Subtraction of numbers and symbols: This data was not useful in this research and did not provide any information about the emotions created by the texts and were removed. Symbols were detected and removed using the .isalpha () method

4) Remove Stopwords: A stopword can be an article, a link or an intention. They can reduce the performance of the algorithm if they remain in the index of words, acting as noise-error (stopwords.words (“English”)).

5) Lemmatization: A process that groups the derivatives and endings of a part of speech, reducing the terms to the original root of the words (lemma = WordNetLemmatizer ())

6) Stemming: Aims to identify the roots of words, regardless of the call or derivation, by removing the suffixes from the words (PorterStemmer.stem (word) and SnowballStemmer.stem (word)).

7) Lexicographic Analysis (Pos Tagging): Recognizes and categorizes what part of speech belongs to each

word from the index of data provided, i.e., if it is a noun, verb, link, adjectives, etc. With this categorization of words, any researcher can cut and use only the parts of speech that are most useful in semantic content (pos_tag (word).

3.4 Data Highlighting and Descriptive Results

From the total of Tweets collected, after the liquidation, which was done as described above, 10,562 final and purified data were left which were processed. Most of these data-tweets come from accounts that have declared Greece as a country (14%) and we say they have stated because Twitter allows each user to declare that country of residence is followed by Ukraine while Sweden is very close, United Kingdom and Germany as shown in Figure 3. This fact is very important because it shows that Crete as a tourist destination receives Tweets from all over Europe mainly but also from other parts of the world, while the Greek Tweets, although they are the most, constitute a small percentage.

Most tweets are made in weekends and the fewest on Tuesdays while we observe that the peak hours are from 12 to 4 at noon while the fewest 11 in the evening until 6 in the morning (Figure 4).

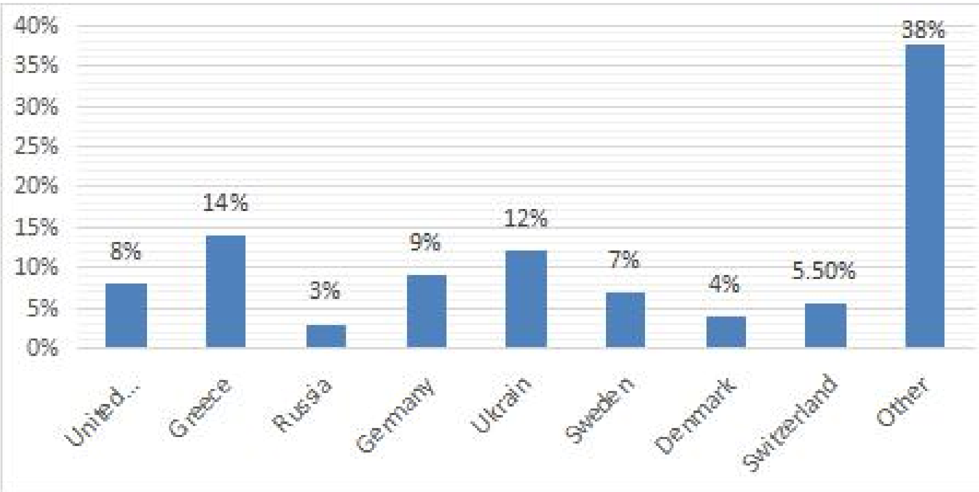
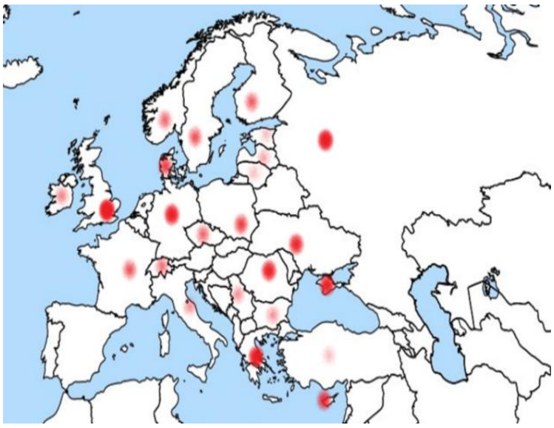


Figure 3 Tweets per Country

Tweets from the world



Tweets from Europe

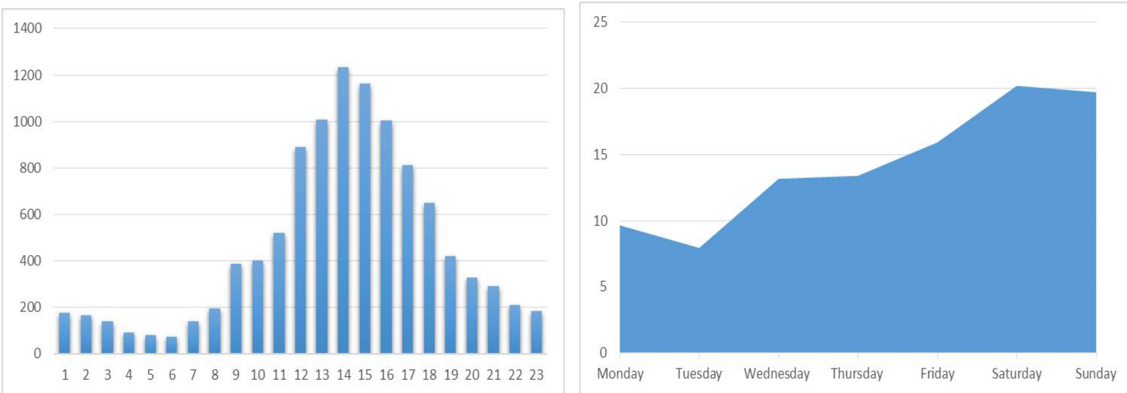


Figure 4 Daily and Hourly Distribution of Tweets

Table 1 shows the usernames, the number of tweets and the number of followers of each user associated with the faster and wider spread of eWoM. Specifically, these users are the ones who have the greatest impact on the reputation of the tourist destination and are considered hubs from which information and news about Crete, as a tourist destination, pass to the general public. User names are collected, but their names remain unpublished, as this information has nothing to do with the study.

Table 1 Top 20 Users With the Highest Number of Tweets

Name	Tweets	Followers
Crete-Tweeter #1	29	57600
Crete-Tweeter #2	27	32000
Crete-Tweeter #3	27	1351
Crete-Tweeter #4	26	4502
Crete-Tweeter #5	23	3899
Crete-Tweeter #6	19	12458
Crete-Tweeter #7	18	41798
Crete-Tweeter #8	16	325
Crete-Tweeter #9	16	1012
Crete-Tweeter #10	15	38
Crete-Tweeter #11	15	11254
Crete-Tweeter #12	14	458
Crete-Tweeter #13	13	321
Crete-Tweeter #14	11	847
Crete-Tweeter #15	10	112
Crete-Tweeter #16	10	332
Crete-Tweeter #17	10	21
Crete-Tweeter #18	10	113
Crete-Tweeter #19	9	214
Crete-Tweeter #20	8	387

The same type of analysis was used to identify the 20 most common hashtags tweeted by media users (Table 2). The hashtags with the highest frequency, are common and refer to expected topics for someone who visits or wants to visit Crete for vacation or just wants to describe Crete and have been highlighted in gray. Such hashtags are #crete, #greece, #summer, #heraklion, #vacations, #mediterranean and #chania.

Hashtags that give a positive sign are highlighted in green, such as #sunvitamin, #hospitality, #friendly, #priceless while contrasting red, that give a negative one such as #canceled that generally come from users' tweets addressed to flights and tour packages canceled from the beginning of August onwards due to the limited number of international flights due to Covid 19 and the hashtag #suzanneeaton.

Hashtags that have not been highlighted are also of great interest for categorization because without the rest of the tweet text they can be ranked positively, negatively or indifferently in any context. These are Hashtags that refer to the Covid 19 pandemic but also to the weather and the Sea of Crete and by themselves cannot give any sign. We still see the Hashtag #fire that can describe the August fires in Greece or wants to emphasize an event in the slang language, most tweets with this Hashtag appeared from 2 to 7 of August which leads us to the conclusion that it concerns the first case.

Table 2 Top 20 Frequent Hashtags

#crete	272
#greece	196
#sunvitamin	181
#summer	144
#masks	144
#hospitality	143
#friendly	141
#heraklion	124
#vacations	124
#covid19	116
#quarantine	108
#weather	104
#priceless	100
#drinks	97
#mediterranean	95
#fire	90
#hottest	78
#sea	65
#canceled	61
#chania	38
#suzanneeaton	21

3.5 Word Clouds

Our analysis included the creation of word clouds that represent the most common words adopted by users. Our main goal is not only to distinguish the most popular words, but also to try to draw an initial conclusion in the evaluation of the polarity of emotions. In figure 5 we present the most popular words found in tweets. The most common words are: “Best”, “Include”, “Travel”, “Friends” and “Free”. If we consider the conditions and precautionary measures taken for the coronavirus as well as the cancellations in tourist packages due to the health protocols and the increase of cases in Crete due to the Delta mutation, it is very encouraging that the passengers use these words, as they represent the majority of their feelings and needs at that time.

3.6 Sentiment Analysis

We have analyzed emotions, at the sentence level, using a dictionary-based approach. In this approach, a ready-made dictionary of emotions was used to calculate the polarity scores of a text, compiling emotion scores throughout the sentence. We use Lexicon TextBlob and open source VADER for this analysis. First, we will present the most important results obtained from the analysis of the emotions of the data, such as the number of emotions of the tweet, the subjectivity versus the polarity and the geographical distribution of the emotion.

After processing, we can finally focus on our main goal in this project. We will calculate the emotional characteristics of tweets such as polarity and subjectivity using TextBlob. Gives us these values using predefined word scores. Polarity is a value that changes between -1 and 1. It shows us how positive or negative the sentence is. The positive sign shows us the positive emotion while the opposite, the negative. When it does not recognize an emotion or an update is made in a tweet, it classifies it as neutral (Figure 5).

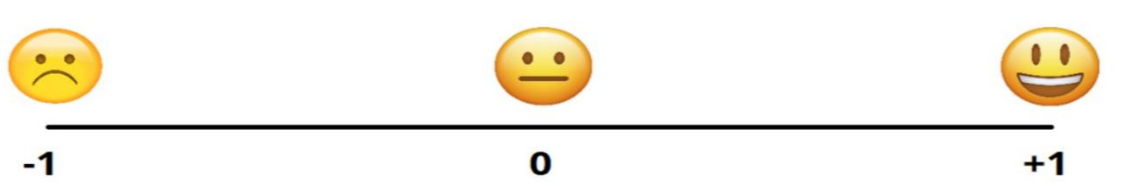


Figure 5 Ranking of Tweets

Out of the total of 10562 tweets that were finally used for the research, 3572 (33.82%) are classified as positive and 1497 (14.17%) as negative. The remaining 5493 were classified as neutral and were mainly about Covid updates as well as weather and hotel updates. The results are shown in the following pie chart (Figure 6).

Moving to the left, the emotion is classified as negative while as we move to the right, it becomes positive. Moving towards the lower objectivity score, the range of polarity score increases. This observation could provide strong evidence that the more subjective a tweet is, the more intense its measured emotion.

In addition, it is noted that in terms of frequency in negative tweets, the graph is denser than the center below, which means that most negatively rated tweets are objective and therefore more reliable. On the right we see that there is an equal distribution between objectivity and subjectivity (Figure 7).

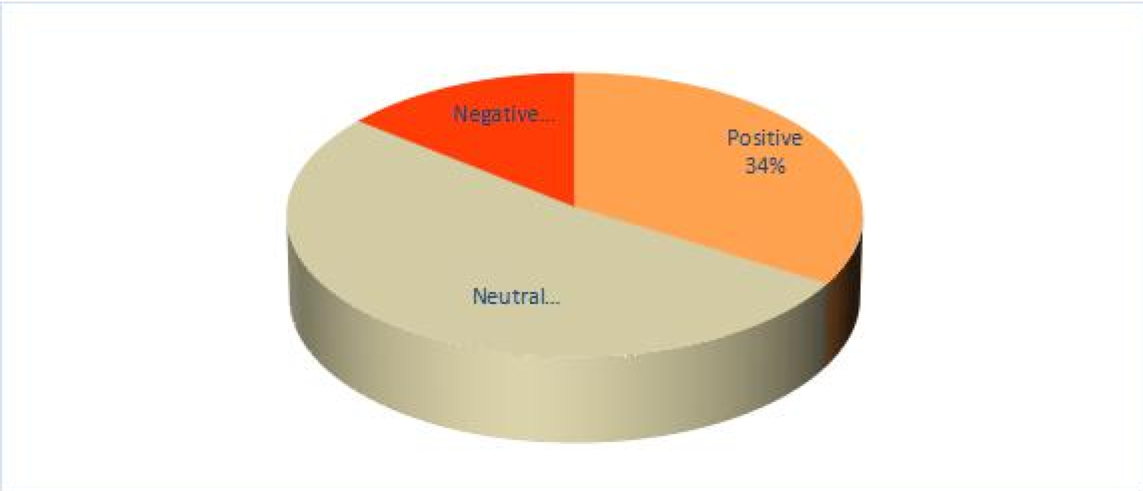


Figure 6 Polarity of Tweets

Figure 7 Subjectivity vs Polarity

An attempt was made to correlate the rating of Twitter users' feelings with the increase of Covid-19 cases on the island of Crete as well as with the announcements made by the Greek government and the ECDC (European Center for Disease Prevention and Control) regarding restricting one's activities to focus and social distancing (Figures 8 and 9).

We notice that during the first weeks that tourism in Greece opened for foreign tourists there is a gradual increase in the positive emotional measurement that is accompanied by the reduction of cases on the island and the lifting of any remaining restrictions. From the 5th to the 7th week we see that the algebraic sum of the emotions remains positive but decreases. We observe a small increase in the algebraic sum in the 8th week and from the 9th to the end of the research we observe a decrease with small ups and downs. Even in the 11th week that the Greek government put Heraklion and Chania in a local lockdown, although the score is the lowest since the beginning of the measurement, the sign is positive with the average of the algebraic sums of the research close to +0.15.

Figure 8 Sentiment Polarity per Week

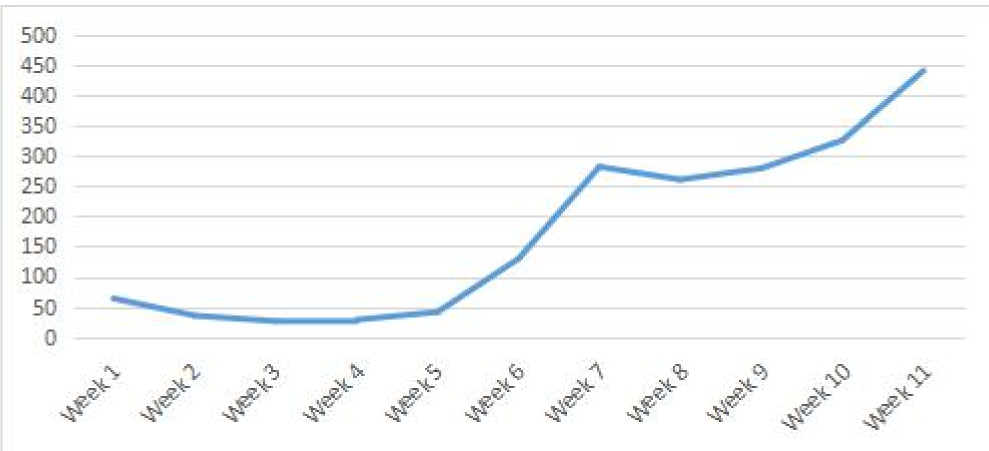


Figure 9 Average Cases of Covid 19 Per Week

The turning points that are worth mentioning and interpret in some way the above data are:

Week 1: Opening of tourism for residents abroad

Week 3: Music in nightclubs and clubs

Week 5: All restrictions removed; masks not obligatory outdoors anymore

Week 7: ECDC ranks Crete in the orange destination (3/5)

Week 9: ECDC ranks Crete in the red destination (4/5)

Week 11: The ECDC ranks Crete in the deep red destination (5/5) and the Greek government puts in a mini lockdown Heraklion and Chania, with a ban on music 24 hours a day and a traffic ban 00: 30-5: 00 early morning.

Figure 10 shows how the emotion component is distributed in European countries, given that most tweets come from there, we also have a clear picture of emotion on this continent. The positive emotions are colored green, while the negative ones are colored red and the neutral emotions are colored gray. The color scale between these two is dynamic, and reflects with great accuracy the overall feeling of each country towards the tourist destination of Crete.

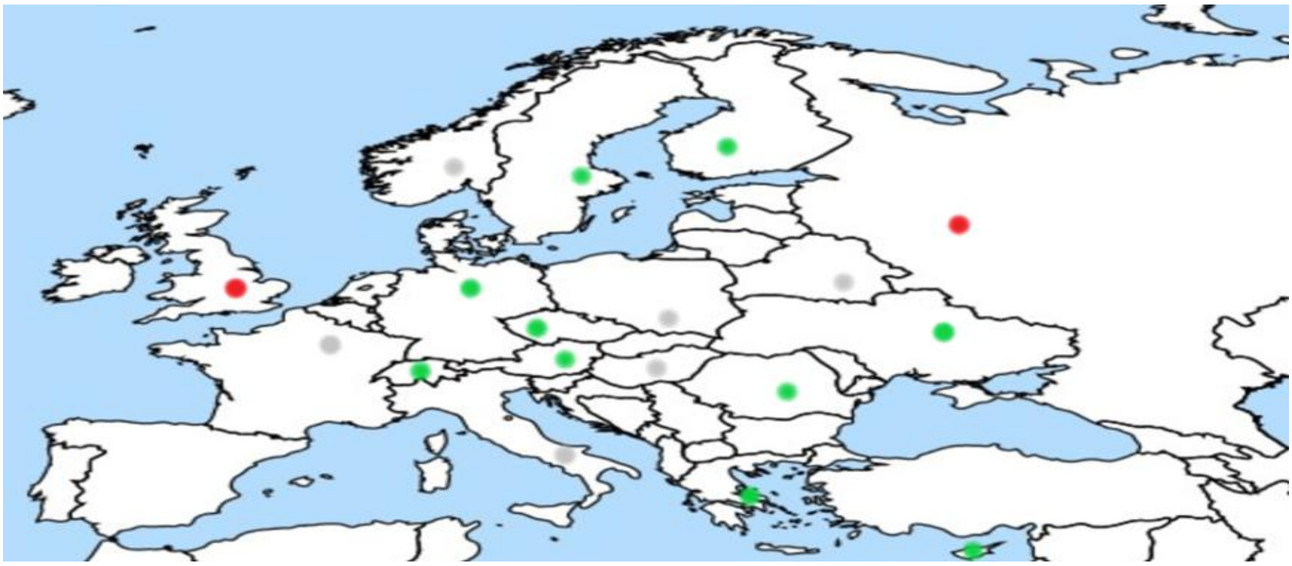


Figure 10 Sentiments in Europe

We notice that the sign is positive in most European countries and in some others it is neutral. We see a negative sign in the United Kingdom and Russia. The negative sign in these two countries may come from the additional restrictions on the first one whose inhabitants had to be quarantined for 15 days on their return from Crete, which created a negative feeling and the refusal of the Greek government to accept the vaccine Sputnik V as a certified vaccine from Russia, forcing tourists to submit a negative 2-day PCR test.

Negative tweets from these two countries confirm the declining trend in the number of tourists from these countries in the pre-Covid era, making the discomfort of these tourists even more intense (Table 3) **错误!未找到引用源。**

Table 3 Thousands of Tourists By Country

	2016	2017	2018	% (2016-2018)
United Kindom	604	592	491	-18,70860927
Russia	150	104	45	-70
Germany	1059	1296	1412	33,33333333
Ukraine	41	49	59	43,90243902
Holland	261	321	323	23,75478927
Denmark	186	194	250	34,40860215
Switzerland	151	171	182	20,52980132
Other	1455	1555	1667	14,57044674

In order to conduct a deeper investigation of emotions, we also present three more word clouds divided into categories according to their category of emotions. In Figure 11 we present the word cloud in terms of the positive emotion, in Figure 12 the negative and Figure 13 depicts the neutral tweets.

From the words that indicate a positive emotion we distinguish: “INCLUDE”, “SUN”, “LIFE”, “BLUE” and “LIGHT”. They are words that can create tourist experiences and show that tourists continue to distinguish the island of Crete and enjoy this tourist destination.

From the words that indicate a negative emotion we distinguish: “MASK”, “PCR”, “TEST”, “QUARANTINE”, “FORESTS”, “DESTROYING”, “CANCEL”, and “COVID”. These words represent certain demands or complaints of passengers, and in addition, show in which areas Crete as a tourist destination through its agencies, should focus on tourist service. Once again we must emphasize that the COVID-19 epidemic left the tourist destination Crete exposed to negative comments.

Figure 11 Positive Sentiments

Figure 12 Negative Sentiments

Figure 13 Neutral Sentiments

From the words that indicate a neutral emotion we distinguish: “ORDER”, “INCLUDE”, “LIFE”, “YEAR” and “ISLAND”. They are words that they use to describe a tourist package or even to describe the island of Crete geographically or as a tourist destination.

In Figure 14, we try to give a network representation of the eWoM spread about Crete as a tourist destination. The purpose of this is to better understand the dynamics of eWoM from the available data. The central node represents Crete and the blue nodes show the 20 top users who posted about Crete as a tourist destination. The green color of the edges indicates the spread of the positive eWoM by the specific user, while the red one indicates a negative one. We can observe that a much larger number of followers are exposed to positive emotions, compared to the negative. In addition, eWoM, passes from user to user and to its own followers respectively, proving that the spread of eWoM starting from a single person, can greatly affect the reputation. In addition, it is important for the tourist destination to have a good relationship with people who are considered “influences”, as they can greatly affect the reputation of the destination to the public.

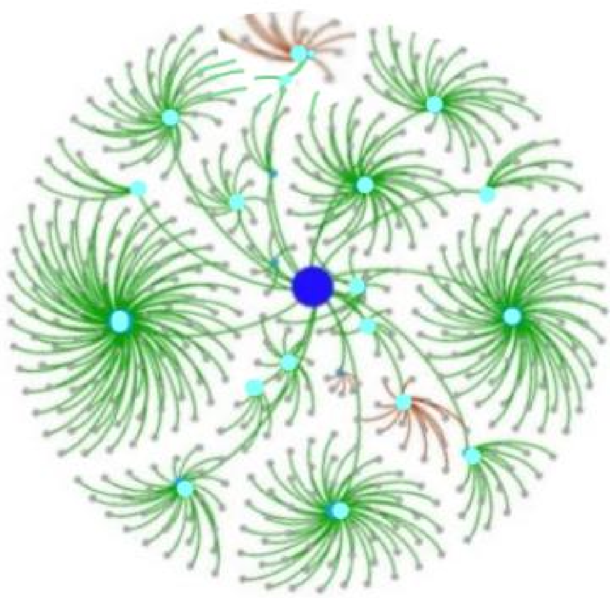


Figure 14 Network of Electronic Word of Mouth

4. Conclusions

In this work, we tried to investigate the feeling of Twitter users about Crete as a tourist destination. The data used covers a short time frame of 11 weeks in the summer of 2021 and includes the period when the protection measures taken by the Greek government for protection against COVID-19 were lifted and ends with the week that the Greek government put again in partial lockdown the two major urban centers of Crete and under increased surveillance the other two. Word clouds played a key role in the research findings because positive and negative word clouds contained different words. This can be translated as the absence of an organization that interacts on Twitter with the users of the social media that are potential customers of the tourist destination of Crete. Perhaps a tourist destination management company — and not the Ministry of Tourism which has an institutional role — could manage and meet the needs of tourists-customers in many cases, users who, for example, do not know if they are entitled to a refund from the cancellation of a travel package due to restrictive measures and wrote about it on Twitter.

During this work, we further monitored some of the important events that affected the emotional imprint for the brand name “Crete” on Twitter, such as the increase in Covid-19 cases, the placement of the geographical destination in a high-risk area by the ECDC and the restrictions in music and the closing of nightclubs for several hours at night, which showed a drop in the algebraic sum of emotions. Despite the pandemic, many people traveled for a summer vacation in Crete with all the necessary precautions.

After analyzing the emotions in the text of the tweets, we noticed that the positive emotion prevails over the negative ones and given the polarity and subjectivity scores we conclude that the positive emotion not only excels in quantity but also in content quality. In other words, more tweets were posted with positive emotion and most importantly, positive emotion is “stronger” than negative. Looking at the rating of emotion, we have noticed many fluctuations. This variation represents eWoM for each post. Although we see a correlation between the increase in cases and the announcements of what strict measures and the magnitude of the positive emotion, we must focus on the fact that it has not turned into a negative sign throughout the research and that the recovery of the curve of positivity of emotions comes much faster than expected and all this without the presence of an organization that could manage and meet the needs of tourists-customers in many cases. Users based on the experiences of senses, emotions, thinking, action and correlation raise the level in the analysis of emotions. In addition, the fact that there is a decent level of fluctuation can be considered as having a positive effect on the Crete brand as a tourist destination. On the other hand, the rapid decline during the announcements and consequently the negative tweets can be considered a lack of coordination and reflexes on the part of the Ministry of Tourism.

Finally, in order to present a clear picture of the eWoM effect, we have created a network that depicts the spread of positive and negative comments about the Crete brand as a tourist destination. The network displays the central nodes, which are the top twenty users, communicating the emotion to their own followers. Of course, the positive eWoM is reproduced by the fans of the nodes and so on. The same goes for negative eWoM. Therefore, given the fact that the positive eWoM outweighs the negative, it is safe to say that eWoM works as a useful and economical tool.

References

Bing L. (2012). *Sentiment Analysis and Opinion Mining.* Chicago: Morgan and Claypool Publishers.

Chien-wen Shen and Chin-Jin Kuo (2015). *Learning in Massive Open Online Courses: Evidence From Social Media Mining Computers in Human Behavior*.

Diamandaras and Mpotsis. (2019). *Mashine Learning*, Thessaloniki: Kleidarithmos.

Forbes L. P. and Vespoli E. M. (2013). "Does social media influence consumer buying behavior? An investigation of recommendations and purchases", *Journal of Business & Economics Research*, Vol. 11, No. 2, pp. 107-111.

Hu (2004). "Mining and summarizing customer reviews", in: *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Seattle, WA, USA: ACM, pp. 168-177.

INSETE (2020). *Actions for Forecasting and Monitoring Changes in the Tourism Sector*, Athens: Bank of Greece.

Justin Kirby and Paul Marsden (2006). *Connected Marketing*, London: Butterworth-Heinemann.

Lerman K. B. G. S. (2009). "Sentiment summarization: Evaluating and learning user preferences", in: *12th Conference of the European Chapter of the Association for Computational Linguistics*, Athens, pp. 514-522.

Neethu M. S. and Rajasree R. (2013). "Sentiment analysis in twitter using machine learning techniques", in: *Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*, Trivandrum, 695016, India: IEEE, pp. 1-5.

Philip Kotler (1967). *Marketing Management: Analysis, Planning, and Control*, Prentice-Hall.

Sherman R. (2015). *Business Intelligence Guidebook From Data Integration to Analytics*, Waltham, USA: Elsevier Inc.

Walaa Medhata, Ahmed Hassanb and Hoda Korashy (2014). "Sentiment analysis algorithms and applications: A survey", *Ain Shams Engineering Journal*, pp. 1093-1113.

Zhang Lei and Bing Liu (2012). *Mining Text Data*, Chicago, USA: Springer Link.