

DEPARTMENT OF MANAGEMENT ENEGINEERING

Master degree in Management Engineering in Digital Business and Market Innovation

Brand-generated and Usergenerated content videos on YouTube: characteristics, behavior and user perception

Supervisor: prof. Susan Chen

Nicholas Beccari 899301 Nicola Valerio 905960

Academic Year 2018/2019

List of Contents

ACKNOWLEDGMENTS	7
EXECUTIVE SUMMARY	10
1. INTRODUCTION	16
2. LITERATURE REVIEW	20
2.1 INTRODUCTION	20
2.2 SOCIAL MEDIA MARKETING	20
2.3 YOUTUBE	23
2.4 SOCIAL MEDIA ENGAGEMENT	26 27
2.5 BRAND-GENERATED CONTENT VS. USER GENERATED CONTENT 2.5.1 Authenticity and credibility – Brand	30
2.6 BRAND GENERATED CONTENT	32
2.7 USER GENERATED CONTENT	34
2.8 AUTHENTICITY	
2.9 CREDIBILITY 2.9.1 People credibility 2.9.2 Source credibility	42
3. DATA COLLECTION AND METHODOLOGY	46
3.1 DATA COLLECTION	46
3.2 DATA ON VIDEO LEVEL 3.2.1 Video & Channel characteristics	47 47
3.3 DATA ON BRAND LEVEL	49
3.4 SURVEY	50
3.5 METHODOLOGY FOR THE ANALYSIS 3.5.1 Descriptive analysis 3.5.2 Discriminant analysis 3.5.3 Correlation analysis	52 52
4. RESULTS AND DISCUSSION	54
4.1 VIDEO LEVEL	54 54

4.1.4 Numerical measure vs. Language measure	59
4.1.5 Discussion	
4.2 BRAND LEVEL	63
4.2.1 Network & semantic brand score	
4.2.2 Network & sentiment brand score vs. Numerical measure	
4.2.3 Network & sentiment brand score vs. Language measure	67
4.2.4 Discussion	67
4.3 SURVEY	69
4.3.1 Discussion	
5. CONCLUSION	75
5.1 MANAGERIAL IMPLICATIONS	_
5.2 LIMITATIONS	78
5.3 FUTURE DEVELOPMENT	79
6. BIBLIOGRAPHY	81
7. APPENDIX	90
7.1 SURVEY	90
7.2 CORRELATION MATRIXES	92
7.2.1 Numerical engagement p-value matrixes	92
7.2.2 Language engagement p-value matrixes	
7.2.3 Numerical vs. language colleration and p-value matrixes	
7.2.4 Network & SBS p-value matrixes	
7.2.5 Network vs. Numerical correlation and p-value matrixes	
7.2.6 Network vs. Language correlation and p-value matrixes	
7.2.7 Ferrari correlation and p-value matrixes with numerical engagement n	
7.2.8 Tiffany correlation and p-value matrixes with numerical engagement m	
7.2.9 Ferrari correlation and p-value matrixes with language engagement matrixes. The Tiffany correlation and p-value matrixes with language engagement matrixes.	
7.2.10 Lijjany correlation and p-value matrixes with language engagement n	ieusures90
7.3 T-TEST	97

List of Tables

Table 1 Numerical engagement measures correlation matrix UGC	56
Table 2 Numerical engagement measures correlation matrix BGC	56
Table 3 Language engagement measures correlation matrix UGC	
Table 4 Language engagement measures correlation matrix BGC	58
Table 5 views and language measure correlation matrix UGC	59
Table 6 views and language measure correlation matrix BGC	59
Table 7 likes and language measure correlation matrix UGC	60
Table 8 likes and language measure correlation matrix BGC	60
Table 9 view and language measure correlation matrix UGC	60
Table 10 view and language measure correlation matrix BGC	
Table 11 Network & semantic brand score correlation matrix BGC	65
Table 12 Network & semantic brand score correlation matrix UGC	65
Table 13 views and network & SBS measure correlation matrix BGC	65
Table 14 views and network & SBS measure correlation matrix UGC	65
Table 15 likes and network & SBS measure correlation matrix UGC	66
Table 16 likes and network & SBS measure correlation matrix BGC	66
Table 17 comments and network & SBS measure correlation matrix BGC	66
Table 18 comments and network & SBS measure correlation matrix UGC	66
Table 19 complexity and network & SBS measure correlation matrix UGC	67
Table 20 analytic and network & SBS measure correlation matrix UGC	67
Table 21 sentiment and network & SBS measure correlation matrix UGC	67
Table 22 sentiment and network & SBS measure correlation matrix BGC	67
Table 23 Tiffany correlation matrix BGC	72
Table 24 Ferrari correlation matrix UGC	72

List of Figures

Figure 1 Description of brand importance	28
Figure 2 Average number of subscribers	54
Figure 3 Average number of views	55
Figure 4 Average number of comments	56
Figure 5 Sentiment histogram	57
Figure 6 Emotionality histogram	
Figure 7 Complexity histogram	
Figure 8 Average number of actors	
Figure 9 Average number of links among actors	
Figure 10 Average Semantic Brand Score	
Figure 11 Average answers value for Ferrari	
Figure 12 Average answers values for Tiffany	

ACKNOWLEDGMENTS

We would like to acknowledge all the participants to this study which made possible this thesis project.

Initially, we would like to express our gratitude to our supervisor Prof. Shan Chen, who supported us in every step through her knowledge and patience.

We would like to thank all the participants to the questionnaire, who fulfilled it spending some of their time to contribute to this study.

Finally, a thank to Sara Afshari, who provides us her thesis that is the base of our work.

ABSTRACT

Nowadays, social media are an important factor in the engagement process. In particular, the popularity in YouTube is becoming an important mean to attract customers, promoting one's self, products and services through the videos uploaded. This research is an exploratory study that aims to analyze the differences between the two main categories of videos: Brand-Generated Content (BGC) and User-Generated Content (UGC). BGC are clips posted on brands official pages mainly used for advertising and marketing; meanwhile, UGC ones are uploaded by people on their channels. The analysis starts from a variety of variables: video & channel characteristics, numerical measure of engagement, language measure of engagement and network & semantic brand score (SBS). From this analysis results that BGC performs better in views, while UGC have a higher level of interaction (likes, dislikes and comments). Moreover, a questionnaire was implemented to evaluate authenticity and credibility perceptions of Ferrari and Tiffany & Co. videos. Examining the answers, Ferrari shows in both variables' impressions an equilibrium between official and unofficial channels, while BGC is perceived more authentic and credible for Tiffany & Co.

ABSTRACT

Al giorno d'oggi, i social media sono un fattore importante nel processo di engagement. In particolare, la popolarità su YouTube sta diventando un mezzo importante per attrarre clienti, promuovendo sé stessi, i prodotti e i servizi offerti attraverso i video caricati. Questa ricerca è uno studio esplorativo che ha lo scopo di analizzare le differenze tra le due principali categorie di video: Brand-Generated Content (BGC) e User-Generated Content (UGC). BGC sono i video postati dai brand sulle loro pagine ufficiali e sono utilizzati principalmente per pubblicità e campagne di marketing; UGC, invece, sono quei video caricati dagli utenti sui loro canali. L'analisi parte da una serie di variabili: caratteristiche del video e del canale, misure numeriche dell'engagement, misure del linguaggio dell'engagement e network & semantic brand score (SBS). Da questo lavoro risulta che BGC ha un numero maggiore di visualizzazioni, mentre UGC ha un livello più alto di interazione (like, dislike e commenti). Inoltre, è stato implementato un questionario per valutare l'autenticità e la credibilità dei video di Ferrari e Tiffany & Co. Esaminando le risposte, Ferrari mostra un equilibrio tra i due tipi di canale in entrambe le variabili, mentre BGC viene percepito più autentico e credibile nei video di Tiffany & Co.

EXECUTIVE SUMMARY

The purpose of the thesis is to understand the different behavior of users on YouTube, where they can find videos generated by other users (User-Generated Content) or contents posted directly by the companies (Brand-Generated Content).

To reach the objective, the work is structured in this way: 1) introduction; 2) literature review; 3) how the data were collected, and the methodology used to analyze them; 4) an exposure of the results and the discussion about them; 5) a conclusive part.

1) INTRODUCTION

Social media are by now an emerging trend. Everyone can use them in a variety of ways: people can simply spend time watching videos on YouTube or searching information on Facebook, but they can also create accounts in order to generate contents or to give their opinions about what they found (Goh, Heng, & Lin, 2013). Moreover, customers are changing their behavior; people are more and more active players of the marketing process, co-creating and sharing contents of the brands (Kietzmann, Hermkens, McCarthy, & Silvestre, 2011).

The major five social media are Facebook, YouTube, Instagram, Twitter and Snapchat, especially diffused between the under 30 years old people (Smith & Anderson, 2018). This can explain the reason why companies are making high efforts on social media marketing: indeed, they use these platforms as a mean to advertise products and services and to raise awareness about the brand's reputation.

This research considers YouTube, since there are very little researches that consider it, even if it is one of the most used platforms. Actually, the aim is to define some possible measures of the engagement through YouTube and to explore the users' different behavior, depending if the channel they look for is a brand-generated content type or a user-generated content one.

2) <u>LITERATURE REVIEW</u>

To understand the actual state of art, the main notions are analyzed using the already existing literature. Particularly, the starting point is to delineate the concepts of Social Media Marketing, YouTube and Social Media Engagement; subsequently, the central topics are treated: Brand-Generated Contents, User-Generated Contents, Authenticity and Credibility. Social media marketing is the use of platforms and websites to promote a product or a service; the main issues for the companies is that, in social media environment, people can also talk to each other, where the brands can only influence these conversations (Mangold & Faulds, 2009). Not only the companies make marketing efforts on social media, but also "influencers" do something similar; indeed, these new figures strongly affect the customers' purchasing behavior through reviews and comments that make products attractive (Smith & Gallicano, 2015). One of the most used social media is YouTube, i.e. a platform in which everyone can upload, manage, share and watch videos; consequently, brands usually exploit this social media to upload clips in order to make products promotion.

Social media engagement measures the involvment of users in the platforms. There is not a single method to calculate it, but it assumes different meanings accordingly to different visions. In general, it could be seen as a multidimensional concept (Hollebeck, 2011), composed by behavioral, emotional and cognitive aspects; this allows to find three different measures of engagement:

- Numerical measures: they include all the numbers that describe the popularity of a content, such as views, likes, dislikes and comments (Burgess & Green, 2013);
- ➤ Language measures: they are the dimensions analyzing the language used by who comments, they are related to the feelings expressed by customers in giving their opinions, and to the way of writing;
- Network measures: they include calculations of the conversations created, together with the centrality of the brand in them.

BGC and UGC are two different channels that it is possible to encounter in the social media. Brand-generated content includes messages, images and videos posted on the official page of the brands (Goh, Heng, & Lin, 2013); they create contents in order to attract people and communicate directly with the users through advertising and the share of information about

products. Thanks to the growing importance of digital world, companies are gradually shifting the efforts from traditional advertising to the online media, creating accounts on the social networks to interact with customers (De Vries, Gensler & Leeflang, 2012).

All images, videos, audios and messages posted by the user in online platforms are classifyed as UGC; there are different reasons why people post, such as entertainment, passing the time, social interaction or self-expression (Papacharissi, 2003). The most used forms of UGC are word-of-mouth, reviews, online communities and logically social media, but all of these lead to the generation of conversations and to the sharing of opinions and ideas.

The main differences between BGC and UGC could be described as the diversity in authenticity and credibility perception. The term authenticity indicates the trustworhty and genuineness of the contents; different studies describe it as a combination of different factors, as continuity, reliability, integrity, symbolism, originality and naturalness (Morhart, Malar, Guevremont, Girardin, & Grohmann, 2015; Burhn, Schoenmuller, Schafer, & Heinrich, 2012). On the other hand, credibility is the ability of someone/something to be trusted and believed. It can be simple described as a mix of trustworthiness, expertise, dynamism, composure and sociability (Perloff, 1993).

3) DATA COLLECTION AND METHODOLOGY

The third paragraph is dedicated to the explanation of how the data were collected and the methodology used to analyze them.

The data collection includes not only the gathering of the videos and all the variables linked to them, but also the creation of a survey that allow to assess the perception of the clips of two brands. The brands, from which the videos are collected, were selected between the 100 Best Global Brand; they had to satisfy different conditions, such as the activity on their social media pages and the existence of contents generated by the users on YouTube.

Thanks to the data gathered about these clips, it was possible to identify different kinds of variables: video & channel characteristics (number of subscribers, age and length), numerical measures of engagement (views, likes, dislikes, comments), language measures of engagement (sentiment, emotionality, number of words, analytic, complexity), network engagement (actors and links among them) and semantic brand score.

Subsequently, the focus moved on two brands, a typical-male as Ferrari and a typical-female as Tiffany & Co.; about these, a survey was created in order to understand the users' perception about the two main aspects described in the literature: authenticity and credibility. Each questionnaire included six videos, everyone followed by 8 sentences that led to understand how much authentic and credible Ferrari and Tiffany are perceived.

The methodology part describes the tools used to analyze the data. Firstly, to have a panoramic view of all the data collected, a discriminant analysis was made, with means, standard deviations, minimum, maximum and overall; secondly, a discriminant analysis was introduced, using the t-test to determine if the means of two different samples (in this case BGC and UGC) are considerably different from each other; lastly, a correlation analysis was performed to statistical evaluate the strength of the relationship between two variables.

4) RESULTS AND DISCUSSION

The presentation of the results aims to show the main differences between the two kind of contents generated.

On YouTube, people search what they are interested in, in this way the meaning of subscribers change a little from the other social media; indeed, UGC channels present a very high level of subscribers, where the main reason is to be updated with their contents. People use social media to get information, for entertainment activities and for social connection, leaving the decision of buying something as a minor part.

One interest thing is that more subscribers do not lead to higher views. In fact, the number of views is very high in BGC channels; this is due to the fact that videos are posted in other platforms, gaining views also from them, but also thanks to the world popularity of a brand, that bring a clips the possibility to be seen worldwide, differently to influencers that usually are more related to a location (Chatzopoulou, Sheng, & Faloutsos, 2010).

Even if the higher views in BGC, the level of interaction is greater in UGC: people prefer to ask information or to give feedbacks under the influencers' videos. This leads to a higher level of the network engagement in UGC: indeed, actors are more inclined to create conversation under UGC videos, even if they usually leave only few comments, participating to the conversation to express their opinion about determined topics.

Lastly, the part related the brand importance on the comment sees an equilibrium between BGC and UGC; since comments, in average, are much lower in official pages, the semantic brand score becomes a significative value for the brands. However, this result should be detailed, because without a sentiment analysis it is not possible to say if the brand is mentioned in the comments in a positive or negative way.

All the findings are confirmed by the correlation analysis: different variables result correlated to each other, verifying statistically all the numbers found.

Moving on the survey, there is not a channel that prevails. Indeed, in Ferrari there is an equilibrium, while Tiffany BGC is better perceived than UGC. This difference is probably due to the fact that most of Ferrari UGC channels are influencers with competences and knowledges, leading to consider them at the same level of the official pages people; meanwhile, the UGC videos or Tiffany are made by people that buy product of this brand and expose their collection and their new pieces.

The literature explains that if a video is perceived more authentic and credible, it should increase its popularity, i.e. the number of likes and comments (Lu et al., 2015; Metzger et al., 2003). The absence of the correlation between numerical measures and the two perceived aspects of the questionnaire is in contrast with it, leading to think that the popularity on YouTube is described in different ways, such as ratings and views instead of likes and comments.

5) CONCLUSIONS

It is demonstrated that BGC and UGC have their strengths in different fields; for example, BGC is better regarding views and the brand importance in the comments, while UGC has a higher interaction. The findings from the survey have not a strong statistical evaluation, because of the low number of brands analyzed and of small sample considered. Managers should focus on the popularity, proposing new contents to translate the high views in a bigger interaction with the customers. Moreover, they must not consider YouTube as a stand-alone platform, because it is better to consider an ecosystem in which social media work together to take advantages by a complete experience given to the consumers.

However, this study shows different limitations, the initial sample is selected from 100 brands, but also other companies that are not included in that list could be taken into consideration if they satisfy the conditions imposed. Moving on the questionnaires, different limits should be considered: certainly, the number of brands to take into account could increase, in order to have a higher number of data that could describe better the perception of authenticity and credibility; moreover, the answers received present a high level of subjectivity, consequently, the results may change in accordance to the sample. Possible future researches should be focused on the resolution of the precedent limitations.

1. INTRODUCTION

Nowadays, social media are becoming a fundamental tool for everybody. Indeed, they are more and more used to do a variety of things: for example, companies can attract customers through marketing campaigns or people can share opinions about everything (Goh, Heng, & Lin, 2013). For these reasons, the brand pages on social networks are increasing, and each company should be able to manage all the principal platforms, with the aim to gain new customers and to maintain the loyalty of the old ones.

Moreover, also the consumers' behavior is changing. Traditionally, they went on Internet to read information about everything and maybe to use them to buy products or services; now, more and more consumers use many platforms (as Facebook, YouTube, Twitter and so on) to create, modify and share the contents. In fact, consumers are adopting more and more active roles in co-creation of marketing contents with the brands and consequently companies should be able to be active in social media, with efforts to reach costumers that live online (Hanna, Rohm, & Crittenden, 2011). In addition, to develop social media strategies, the different platforms are too often treated as stand-alone elements, while they should be considered as parts of an integrated system useful to attract customers.

The changes in customers' behavior are confirmed by a study about this trend (Kietzman, Hermkens, McCarthy, & Silvestre, 2011). Furthermore, the emergence of these Internet-based social media allow people to communicate each others about products and the companies that provide them, leading to an expansion of consumer-to-consumer communications that must be considered in the promotion mix of the brands (Mangold & Faulds, 2009). As a consequence, the original sense of "promotion mix" can be divided into the traditional sense, i.e. the companies that directly talk with their customers, and the untraditional one, where consumers talk each others.

Some interesting results are showed by the data collected by Smith (2019), that sintetized them in his article "126 amazing social media statistics and facts" published on Brandwatch. Brandwatch is a web sites in which it is possible to find statistical evidence about very different thing; in particular, the article abovementioned goes in details talking about the social networks and their numbers. The focus is put on the most known social medias, that are also the ones in which the majority of the digital marketing efforts of the brands are concentrated: Facebook, Instagram, Snapchat, Twitter and YouTube.

The previous article lists the number of users for each social media; Facebook stands out, with its more than 2 billion members, and it is followed by YouTube and Instagram (1.9 billion and 1 billion respectively). In comparison, Twitter and Snapchat count very low number of users (330 million and 160 million).

Considering these 5 social media, from a surveys done by Pew Research Center (Smith & Anderson, 2018) emerges that the 88% of people between 18 and 29 years old uses every types of social; a high percentage of members is also that of persons between 30 and 49 years old (78%) and between 50 and 64 (64%), while it falls until the 37% for the 65 and older persons. Particularly, YouTube and Facebook are the most used ones, with a peak of 94% active users of YouTube between 18 and 24 years old.

These data allow to understand why the companies are more and more exploiting these platforms to do marketing campaign with different targets, using the social media as means to advertise products and services through posts or videos and to try controlling the exchange of information and opinions between customers that comment them (Blackshaw & Nazzaro, 2004).

The growing importance of these platforms for marketing is an opportunity but also a complication for managers, because they must be updated with the issues about them (Hofacker & Belanche, 2016). Particularly, digital information are considered as products to be purchased and sold (Lambrecht, et al., 2014), especially because customers may participate in different parts of the process, frequently contributing to create all of the products' value (Mustak, Jaakkola, & Halinen, 2013). Moreover, the competition level increases, leading to higher efforts and investments in this field, since it gives better results (Hong, 2013).

In this case study, the social media considered is YouTube, i.e. a platform where everyone can upload videos and watch the ones posted from the others (Cheng, Dale, & Liu, 2007). This social media was born with a focus on user-generated content, but now it has become a platform for online video worldwide and an important location for some of the most significant trends and controversies in the new-media environment (Burgess J. E., 2011). Indeed, YouTube offers to everybody the possibility to create an account; the two main categories of video maker are persons (the ones for which the social was founded) and brands, that more and more use social networks to attract customers. People that post clips created the so-called User-Generated Content (UGC); they upload videos for a variety of

reasons, as entertainment, reviews, unboxing and so on, influencing the behavior of the watchers (Ding, Phang, Lu, Tan, & Sutanto, 2014). On the other hand, the same authors explain that the brands that upload videos on its official page create the Brand-Generated Content (BGC); these are made especially for advertising, exposing the features of the new product or service to capture the users' attention and to convince them to buy it.

Ding, Phang, Lu, Tan, & Sutanto (2014) add also the importance of these two kind of channels for product-related (for example, the products promotion) and social-related (to build relationships) content. Indeed, the aim of both brand and user generated contents are to promote some products and trying to build and maintain strong relationships with the customers. The second part is well done from both the kinds of channels, while the user-generated are more credible in terms of product promotion; for this reason, maybe the brands should consider to stimulate the UGC in order to increase the credibility perception of their advertising.

Especially on some social networks as YouTube and Instagram, the figures of "Influencers" are emerging: they are an indipendent third party who shape audience attitudes thorugh blogs and the use of social media (Freberg, Graham, McGaughey, & Freberg, 2011). Indeed, they are acquiring more and more importance, modifying the target audience's behavior in compliance with that influence (Scheer & Stern, 1992). This leads to different outcomes depending on the favorable or unfavorable impact: the favorable one improves the negative attitudes, while the unfavorable outcomes principally undermine positive attitudes.

Even if YouTube is one of the most used social media, there are very little researches that consider it, unlike others such as Facebook and Twitter. Consequently, the aim of this thesis is to explore the data related to videos posted on YouTube by these two categories: the brands official pages, i.e. Brand-Generated Content, and the unofficial ones, that are the User-Generated Content.

Hence, the thesis pursues different objectives:

- identify the possible measures and variables that can be find about engagement;
- explore the users' different behavior, comparing the brand-generated content and the user-generated content to underlines the differences in their performances.

The research is divided into five chapters.

The second section is the literature review, that provides some relevant concepts in this field. The aim of this part was to increase the knowledge about the topics included in the thesis, and in particular YouTube, social media marketing, social media engagement, BGC, UGC, authenticity and credibility.

The third chapter regards the description of how the data were collected and the methodology used to analyze them. The data collection is divided in two parts: the gathering of the data about the videos and the answers to the survey related to authenticity and credibility perceptions. They are subsequently analyzed through statistic methods such as t-test and correlation analysis.

In the fourth section, the results of the research are presented and discussed. For each variable, the performances of the two categories are compared, especially highlighting when the differences are remarkable.

The last part is the conclusion, where the main results are summarized; moreover, in the same chapter, the limitations that influenced the study made are listed, together with some possible future development in order to overcome these limits.

2. LITERATURE REVIEW

2.1 INTRODUCTION

The first step in this research is to define the actual state of art, and so the starting point, to develop the new model. It is important to define the criteria that allowed to find the right information given the big amount of data present on the internet.

Before going to the core of the research, the comparison of UGC and BGC videos, it is necessary to study the context. Doing so, the work is structured in this way: a first part regarding a more generic approach, while a second one more focused.

The starting point of the analysis was selecting the right keywords that brought to understand the world in which the research was taking place. In fact, to know better YouTube and the surrounding environment, were selected these three keywords: YouTube, social media marketing and social media engagement. The first one serves to comprehend the potential of this social media which people believe they know how to use it. The other two words are essential to learn how to market on social media, engaging users in an active behavior, providing an experience.

After the context was analyzed, it was possible to focus on the main part of the research: UGC and BCG. These were the two keywords used, together with two more: authenticity and credibility, that are two of the most important factors found from the literature that influence the perception of the two different channels. The first two keywords allowed to figure out the difference between content created by a user and the one made by a brand. While the other two will represents the variables to classify YouTube videos and to compare them, in order to evaluate the content, user or brand generated, in each video.

All the papers, found on Google Scholar, were firstly classified according to the H index of the journal that published them and then were read.

2.2 SOCIAL MEDIA MARKETING

Nowadays, the communication with the customers is very different from the past, especially thanks to the emergency of Social Media. They are new sources of online information that are created and used by consumers to educate each other about products, services and brands

(Blackshaw & Nazzaro, 2004). Consequently, social media marketing is the use of these platforms and website to promote a product or a service.

Social media plays an important role for companies, enabling them to talk with their customers or allowing people to talk to each other, where companies can only influence their conversation but cannot control them (Mangold & Faulds, 2009). Indeed, they become an important factor that influences people behavior, such as information acquisition, purchase behavior and post-purchase communication. For this reason, social media marketing is more and more used by managers, even if they do not receive full appreciation for their role in the company's promotional effort.

Social media marketing provides benefits for the companies. First of all, it's possible to listen customers' suggestions and complains. Second it allows to recognize influencers and third everything can be done without paying. These are the main benefits but there are many more such as: generating exposure, increasing traffic subscribers, building new partnership, etc. (Sajid, 2016).

Kim & Ko (2010) studied luxury brands social media marketing; they found that these platforms have five properties: entertainment, customization, interaction, word of mouth and trend. Every feature result to positively influence customer relationship and purchase intention. These two factors are the core of social media marketing. The two authors, in a later work (Kim & Ko, 2012), confirm the importance of the five properties of social media. They added two attribute, brand and value equity (brand equity is the subjective evaluation of the brand over and above its value; value equity is the objective evaluation of the utility of the brand by the customer), because they both influence purchase intention.

De Vries, Gensler, & Leeflang (2012) focused on the effects of social media marketing; especially they defined some drivers for brand post popularity. To enhance popularity, brands have to post on top of the brand fan page. Likes and comment are two of the main drivers for popularity; there are different technique to increase the two. Vivid and interactive brand post and the share of positive comments enhance the first, while interactive brand post such as question and the share of both positive and negative comment increment the second one. Another factor that can increase popularity of a brand is the use of influencers. Kumar & Mirchandani (2012) in their article demonstrate the effectiveness of a seven-step framework able to find the right influencer for their campaign. The results show positive

effect for the brand the author investigated, a small business chain, but it could be extended to larger global enterprises.

Ashley & Tuten (2014) offer a connection between social media marketing and social media engagement. The authors describe different type of creative strategies such as: functional appeal, emotional appeal, experiential appeal and so on. Then they calculate the correlation between channel, creative strategies and social media engagement metrics. The results show that brands with the greatest number of tweets in a week is the one with highest followers and the highest *Klout* score on twitter. Brands that use most social media channels are the ones with more followers and highest engagement scores. Engagement is strictly related to how and how many social media a brand uses, together with a coherent strategy.

The core of this research is to understand the differences between two kinds of contents generated on social media: Brand-generated and user-generated content (BGC and UGC). Brand-generated content are videos, images, messages, etc., directly posted by the brand on its official page, while the user-generated are the same contents uploaded by people.

Nowadays, people use more and more Internet and technology to do everything, especially the young generation; for this reason, the marketing through social media is a key concept for the brands. Moreover, the so-called "influencers" are figures with which the companies must co-exist; indeed, they are persons that influence the purchasing behavior of costumers, through reviews or comments about products that make them attractive (Smith & Gallicano, 2015).

The main differences between these two types of channel are found in their perception by the users; generally, all the content must be perceived authentic and credible, in order to attract customers and increase the reputation.

However, even though the social media marketing is an emerging trend, there are not so many studies of the phenomenon. Searching the keyword "social media marketing", the majority of the researches takes into consideration Facebook and Twitter; it can be a starting point, because they clearly are social media, but the behavior of these two social networks is different from the one that is central in this research: YouTube.

2.3 YOUTUBE

YouTube is an American video-sharing website. Online videos existed long before YouTube entered the scene. However, uploading videos, managing, sharing and watching them was very cumbersome due to a lack of an easy-to-use integrated platform (Cheng, Dale, & Liu, 2007). Before YouTube, the videos were stand alone, they weren't related with any other video, plus the possibility of reviewing and rating did not exist. The video-sharing website overcame these problems, giving the opportunity to relate videos thanks to keywords linked to each one and allowing also to rate and comment each of them.

YT can be considered also a social media: in fact, there are communities and groups on it, there are statistics and awards for clips and personal channels. There is a connection between videos and users. Users can interact with the videos they watch; in fact, they can give feedbacks in different ways. they have two characteristics: the number of friends and the number of uploaded videos.

An important factor for YouTube is being popular since it helps in marketing services and products. To assess the popularity there are various metrics such as: *viewcount, number of comments, number of favorites, number of ratings* and *average rating*. These metrics capture the reaction of users to a video. Chatzopoulou, Sheng, & Faloutsos (2010) with their work discovered that a tape is commented, rated or added to someone's favorite list once every 400 times is viewed. They also found, talking about uploading, that the daily peak is at 1 PM and the weekly peak is on Sunday for most video categories. Everything is based on the Pacific Time Zone so it's impossible to take 1 PM for grant since YouTube is globally accessible.

Zhou, Khemmarat, & Gao (2010) studied YouTube recommendation system; they found out that this technique is really powerful, in fact, 30% of overall clip views comes from it, resulting in the most important view source for the majority of tapes. In this way video that are placed in a related list of popular ones have a higher chance to become popular. This system results to be a factor driving users in discovering more clips of their interest rather than popular ones only.

Brodersen, Scellato, & Wattenhofer (2012) studied global popularity of videos; they discovered that half of all them have about 70% of their total views in a single country. Indeed, video views grow firstly in the focus location and then they expand across other

regions. It's possible to see that, even if YouTube has global audience, the popularity of clips follows a pattern limited to the geographic locality of interest.

In his research Shifman (2011) wanted to study a particular category of YouTube tapes: memetic ones. It is defined by the author as: "a popular clip that lures extensive creative user engagement in the form of parody, pastiche, mash-ups or other derivative work." The fundamental characteristics of this typology of clips is the participation through mimesis. The researcher found that all the videos have the same six features: focus on ordinary people, flawed masculinity, humor, simplicity, repetitiveness and whimsical content. These attributes mark the videos as incomplete or flawed; this generates engagement since the incompleteness makes the users start further dialogue contributing to the successful spread of the meme. All of this brought to the conclusion that UGC and especially memetic video are important for their ability to engage users and to generate WOM.

2.4 SOCIAL MEDIA ENGAGEMENT

The word engagement means involvement, participation. Social media engagement measures the participation of users in social media. There is not a given formula to calculate the engagement, but every brand can decide how to do it. Usually the main factors used are the numbers of likes, comments and repost. Social media engagement can be considered the primary object of a social media marketing campaign. Harrigan, Evers, Miles, & Daly (2017) delineated customer engagement made of five dimensions: enthusiasm, attention, absorption, interaction and identification. Enthusiasm represents individual excitement or zeal and interest in a brand. Attention refers to a customer's level of focus, consciously or subconsciously, on the brand. Absorption goes further than attention; in fact, consist in a customer's high level of concentration and engrossment in a brand. Interaction involves sharing and exchanging ideas, thoughts and feelings about experiences with the brand and with other customers of the brand. Customer will identify themselves more with brands that match their self-image.

Consequently, the concept of individuals as a mere consumer is outdated; nowadays, the consumers should be seen as active producer of value. Social media strategies should focus on involve customers, making them active, participating in company business. So, the attention should shift in engaging the users on social media (Heinonen K., 2011). To do this

in the tourism field, Cabiddu, De Carlo, & Piccoli (2014) highlighted three type of customer engagement: persistent engagement, customized engagement and triggered engagement. The first represents the possibility of maintaining a dialogue with customers even when they are not at home (hotels posting video on YT or photo on Pinterest). The second consists in the possibility of proposing something customized for each client thanks to information previous collected. The last is the possibility to instigate customers encounter based on another customer-initiated event.

To engage people and consumers, the brand communities are a relevant aspect to be taken into account to build brand trust (Habibi, Laroche, & Richard, 2014). A brand community is a community of social relations formed by people that admire a brand; community engagement is the intrinsic motivation of the consumer to interact and cooperate with the other community members, in order to create strong relationships between consumers.

Hollebeek, Glynn, & Brodie (2014) consider the consumer brand engagement made of three factors: cognitive, emotional and behavioral engagement. The first is the level of elaboration and though processing of a consumer in an interaction with the brand. The second represent the degree of positive affect in a contact with the brand. The last is the effort, energy and time spent by a consumer interacting with the brand. A high level of these three factors bring the consumer to self-connect with the brand together with the intent of brand usage. The same description of the engagement as cognitive and emotional is done by Smith & Gallicano (2015), they added that it may not characterize all the social media. Indeed, they delineated engagement driven by four factors: information consumption, sense of presence, interest immersion and social connectivity. The first factor deals with being informed and staying up to date. The second one is described as cognitive attachment to their online experience. The third element is the personal immersion into content important for people own interest. Last item considers the interaction with friends. Moreover, the investigators focused their attention toward who are among the most active in social media: Millennials. According to their perspective, the authors redefined the four factors before listed. The first factor becomes being up to date with organizational online activities. The second one, since millennials consider themselves engaged with organizations online based on invested time and attention by the brand itself, they expect the organizations to invest consistent interest in them. The third element is the personal research of company, then assessed on its

engagement level based on personal interest. Last item considers offline consumer activities and additional relational interaction online.

The distinctions between the different vision of engagement previously mentioned were done also by Hollebeck (2011): he described the engagement as a multidimensional concept, that includes behavioral, emotional and cognitive aspects. In details, the behavioral includes all the actions did by the potential consumers, as viewing, liking and commenting; the emotional part is formed by all the feelings shared by the language used by the customers; the cognitive includes all the thoughts of the actors and the relations that are created between them. As a consequence, three different measures of engagement could be found: the numerical measures (all the numerical parts as views, likes, dislikes), the language measures (all that regards the feelings and the sentiments transmitted by the users) and the network measures (all the relationships between customers).

2.4.1 Numerical measures of engagement

The popularity of social media encouraged the interaction between brands and consumers; in particular, YouTube allows the share and the discussion about different contents (Burgess & Green, 2013), offering the possibility to comment a video or to judicate it with a like or a dislike. Consequently, the popularity can be seen as the combination of all these characteristics, that lead to understand how many consumers could be attracted by the content posted.

Lim, Hwang, Kim, & Biocca (2015) in their research studied social media engagement during sport tv events, they discovered three dimensions of it: functional, emotional and communal engagement. The functional engagement is the one that interests these measures, while the other two will be better described as Language engagement. Indeed, the functional engagement is the participation of consumers in the social media sites; it can be measured with many indicators, depending on the social media: for example, views, likes, dislikes and comments are good measures of functional engagement for YouTube, while the retweet, hashtags and mentions are optimal ones for Twitter (Bik & Goldstein, 2013, et al.).

2.4.2 Language engagement

As previously mentioned, Lim et al. (2015) explained other two types of engagement: the emotional and the communal ones.

The emotional engagement is related to the feelings and the sensations expressed by the consumers in their opinions; the basic aspect of the sentiment part is the possibility to share it with other viewers, that can agree or not with it (underlining also the importance of the network part). This leads to the third part, the sense of community that is essential to create feelings and links between users.

Talking about emotions, there are many studies that focused on their analysis in several fields: from the sentiment analysis in social media (Desmet & Hoste, 2013) to online news and blogs (Nassirtoussi, Aghabozorgi, Wah, & Ngo, 2014; Li & Xu, 2014), from the Human Computer Interaction (Ali, Hariharan, Yaacob, & Adom, 2015) to the Speech Emotions Recognition. There is not a given number of emotions to be analyzed, but there are models that take into account a categorization of the emotions: the primitives and the secondaries, that resulted from a combination of the first ones (Cowie & Cornelius, 2003). Someone considers four basic emotions, i.e. happy, anger, fear and sadness (Chavhan, Yelure, & Tayade, 2015, et al.); others found six emotions, adding to the previous four surprise and disgust (Balti & Elmaghraby, 2014, et al.); there is also the neutral sentiment, that is used as intermediate when switching between two different emotions.

Alonso et al. (2015) explained that there are two most commonly used dimensions to represent the emotions: the activation (the perception of the intensity of the emotions) and the valence (if they are positive or negative).

2.4.3 Network engagement

The Semantic Brand Score (SBS) is a new measure of brand importance that overcomes some of the limitations of the old measures: it is automatable and fast to compute even on big text data, without surveys or without informing the ones who generate the content; SBS can be calculated for each kind of text documents and it can be applied to different contexts, as newspaper, social media platforms, etc. (Colladon, 2018).

This method is introduced because it seems important, in the online environment, to find a

measure that does not impact on the spontaneous behavior of social media users and that does not use a survey. The author proposed this new method to analyze the brand importance, studying the occurrences of a brand name in the discourse, its embeddedness in text data and the heterogeneity of its text associations. According to this classification, a brand that is used marginally is considered "unimportant", while an important brand is at the core of a conversation, but it can be

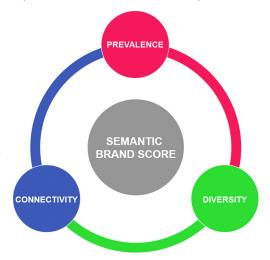


Figure 1 Description of brand importance

associated to both negative and positive comments: for this reason, a more comprehensive measure is obtained combining SBS with a sentiment analysis.

In particular, the brand importance is described as the combination of prevalence, diversity and connectivity, as shown in Figure 1:

- ➤ Prevalence is the frequency with which the brand name appears in a set of documents; higher prevalence implies higher brand mentions, with consequently more visibility and familiarity for the authors of the documents. Aaker (1996) and Keller (1993) suggested a link between prevalence and brand awareness, even if it is only partial, because there can be brands with high awareness but not frequently mentioned in documents.
- ➤ Diversity is the heterogeneity of the words co-occurring with a brand: it correlates the lexical diversity (McCarthy and Jarvis, 2010) and the study of word co-occurrence (Evert, 2005). Grohs et al (2016) added that the uniqueness and the type of the association are very important. For example, a brand with high prevalence but always related to the same words is limited of a very specific context, while diverse associations usually lead to a richer discourse and a higher versatility of the brand. As with the prevalence, the diversity could increase the brand awareness (Keller, 2016; Aaker, 1996).
- ➤ Connectivity describes how often a word (in our case the brand) serves as an indirect link between other words. Higher connectivity indicates that the brand is used as a "bridge" between different words.

Nevertheless, the SBS does not consider if the feelings are positive or negative, but only if the brand is frequently mentioned and central in the conversations. Its results should be interpreted together with the context where it is measured: for example, if a brand is in crisis, it is better to have a lower SBS because all the mentions are probably related to negative feelings. The solution is to make a sentiment analysis; it can be done in different ways: Gloor et al (2017) exposed a method in which the analyst classifies the comments that include the brand name, then they should be divided into two subset (positive and negative), in order to calculate the SBS for each subset and to determine where the brand is stronger. It is also possible to study the polarity of the words surrounding the brand in the co-occurrence network (Basile & Nissim, 2013) or to isolate a set of positive and negative words to evaluate their SBS and their influencing power.

2.5 BRAND-GENERATED CONTENT VS. USER GENERATED CONTENT

People rely in different way on mass media or interpersonal communications and in particular this changed over time; in fact, Mahajan et al (1995) explained that the credibility of external communication as the TV advertisements tends to be higher at the starting point and to decrease over time, while the one of internal influences as the word-of-mouth is lower initially but it increases over time, reaching a peak and then it has a decline.

Also the relative effectiveness of the two different methods is studied: Bruce et al (2012) estimated a dynamic linear model through a Kalman filter and demonstrated that the traditional advertising is more effective in the early stages of the product life cycle, while the word-of-mouth is more effective when consumers gain experience with the product. This could be explained through the idea of advertising repetition wear-in and wear-out: at the beginning, the advertising spending is higher, wear-in occurs and audience attention is very high, but when the consumers gain experience with the product, the effectiveness of the repetition wears-out and advertising becomes less effective than the word-of-mouth. MacInnis et al (2002) added differences in wear-out based on the type of advertisement: they explain that emotional ads have a slower wear-out compared to the rational ones, because emotional ads create a positive relationship between advertising and sales.

These ideas is confirmed by Villanueva et al (2008), who investigated marketing (broadcast media, direct mail) and WOM, using a vector autoregressive model to examine how

customers attracted by these two forms of communication affect the growth of the firm's customer equity; the results showed that marketing is more valuable in the short terms, while the WOM in the long ones.

The study of Lee & Choi (2017) highlighted the difference in terms of content trustworthiness: it is demonstrated that UGC has a stronger effect on customers' cognitive trust, while the BGC on customers' emotional trust. The same authors explained that trustworthiness is only a factor that influences another concept: the credibility. Perloff (1993) treated the same topic, adding other items that affect the credibility: expertise, dynamism, composure and sociability. Moreover, a brand can influence the user emotion through brand-related contents and consequently, a specific brand behind the content could be a key indicator for users to establish their way of thinking and make affective reactions (Li, Oh & Wang, 2017).

Different researches went deeper into other items that describe the differences between BGC and UGC. One of them developed a scale to measure the consumers' perception in terms of four dimensions: reliability, integrity, symbolism and continuity (Morhart, Malar, Guevremont, Girardin, & Grohmann, 2015); particularly, these items are all factors that influence a significant variable: the authenticity of what a brand (or a user) posts.

From the previous articles, it is evident that there are two key concepts that every content generated should observe: the authenticity perception of what is said or written together with its credibility. These are two keywords that assume different meanings if they are associated to a brand or to a user.

2.5.1 Authenticity and credibility – Brand

To have an impression to be authentic, a sincere story is required (Beverland, 2005); the sincerity could be achieved through different characteristics, as uniqueness, passion for the products of the company and the use of modern marketing techniques. However, the sincerity is not the only factor that represent the brand authenticity from a consumer's point of views; indeed, Napoli, Dickinson, Beverland, & Farrelly (2014) explained other two characteristics that correspond to a higher brand authenticity construct: the quality commitment and the heritage. The first indicates both the quality of the materials used, if the owner pays attention to details and if he/she is involved in the production process, while the

heritage can be considered synonimus of culture values and traditions that establish a sense of legitimacy and authenticity among some target groups (Kates, 2004).

The credibility is another factor that is as important as the authenticity. in particular, Erdem & Swait (2004) underlined the importance of the brand credibility on brand choice and consideration; indeed, a higher credibility implies a higher probability of inclusion of a brand in the consideration set. Moreover, the same authors suggested a subdivision of the credibility that is one of the most used in the different analysis: trustworthiness and expertise. The trustworthiness is the reliability of the brand, while the expertise is the accuracy of information together with the presence or not of certifications, credentials and similar.

Brand credibility is studied to understand also different aspects that it could influence; for example, it is demonstrated that the credibility affects also the price sensitivity, and this influence varies across product categories, depending on the uncertainty of them (Erdem, Swait, & Louvriere, 2002)

2.5.2 Authenticity and credibility – User

In reality, there is not an evident difference between user and brand authenticity; also in this case, the authenticity could be described with the sincerity and genuineness of people who write or say something (Beverland, 2005).

The majority of studies about the authenticity of a person were done in tourism, and in particular taking into account the reviews of the holidays spent in a determined place. Pearce & Moscardo (1985) explained that the authenticity of users' description is a key concept to understand the tourists' behaviors and experiences. Neverthless, it is also demonstrated that the concept of authenticity is influenced by subjective and collective views on consensus and creativity: for this reason, the simple binary distinction between authentic and inauthentic can be a little small, suggesting that it could be overcome especially in the approaches to authenticity's tourist role (Olsen, 2002).

Even in the case of credibility, it is possible to describe it as the importance of people trustworthiness and expertise (Erdem & Swait, 2004).

For example, the trstuwrothiness of a person could be influenced by the accent: Lev-Ari & Keysar (2010) demonstrated that the native language is more credible than the non-native

one, making less credible people who routinely communicate in a language different from their mother tongue.

Also the level of expertise is an important measure of credibility: indeed, the higher are the competences and the certifications of the person, the higher is his/her credibility. This is confirmed by Anderegg, Prall, Harold, & Schneider (2010), who explained that the researchers are more reliable if they have experience in a determined field, i.e. if they public many articles about the same topic and if they are mentioned lots of times in other studies.

2.6 BRAND GENERATED CONTENT

Brand-generated content can be defined as any type of content (messages, videos, images, etc.) posted online directly by the brand.

If a brand wants the consumers to engage with itself, the information and content shared must be relevant for the customers (Schmitt, 2012). For this reason, the branded social media activities should be used to increase the brand awareness, promote customers' engagement and loyalty and inspire the consumers' word-of-mouth in order to expand the knowledge of the brand's qualities (Ashley & Tuten, 2014); moreover, Stelzner (2013) explained that a brand can increase their opportunities to increment the target audiences through rising the messages on web sites and social media.

Every brand should define its own customer engagement behavior: in social media, it can be translated in encouraging consumers to participate and share their opinions or directly in the brand-generated content; this is typically measured by observing when the users link each others, blog, connect or submit a demand or an idea (Falls, 2010).

To understand how much users are involved with these contents, Frederick et al (2012), Horton & Wohl (1956) and Rubin, Perse & Powell (1985) introduced the Para-Social Interaction (PSI): it is the relationship between media personalities and media users and it is conceptualized as the interpersonal involvement of the media users with what they consume. The same authors in their studies looked for some "antecedents", i.e. some variables that influence or are correlated with PSI; the results showed that there are two main variables of this kind:

- Attractiveness: the attraction to a media, both social and physical, can be considered as a predictor of PSI; it increases the interactions or the repeat viewing. The more users perceive the media personality similar to themselves, the more likely PSI occurs.
- Attitude homophily: it is a tendency for friendships to form between those who are alike in some designated respect (Turner, 1993); it means that the more users perceive similarity between their beliefs and the media personality, the more likely they will continue the relationship.

Goh et al (2013) introduced the concept that marketers generate contents to advertise their products and to provide information through the web; in particular, the brands update product-related images, videos and posts on Facebook or YouTube to attract people in these channels, or they use social network services to communicate directly with the users. The same authors wrote about the social broadcasting communities: in this environment, marketers can release the latest news about the products to the current and potential customers; in this way, the brand underlines that the interactivity should be critically considered to increase the influence in the communities. So, it is important to encourage communication among users to drive purchases on the platforms.

The most common brand-generated content is the traditional advertising, even if some empirical evidences suggest that firms are gradually shifting their traditional advertising investments to online media, as social medias and blogs. For this reason, many firms created their pages on the social networks as Facebook and post messages to interact with costumers: this could be called firm-to-consumer social messages (De Vries, Gensler & Leeflang, 2012). Kumaret et al (2016) thought that F2C have positive effect on existing customer's expenditure and that F2C and tradition advertising has complementary effects.

In the actual world, advertising in an online context does not mean only the message itself, but also personal channels and link to other people, contents and sites. In particular, on social networking sites the success of a brand can be described as the combination of social connectivity and user interactivity:

➤ Social connectivity: the success of a brand in a social networking context is related to the social connectivity of consumer-networked relations that the brand provides; so, the content of marketing communication must be customized allowing the engagement and connectivity between individuals, groups of people and brand. (Shih, 2009).

➤ User interactivity: it is a user-centered interaction with machines, messages or other users, focusing on the experiential aspect of networking process (Liu & Shrum, 2002); this interactivity allows consumers to participate in social networking by selecting content, time and other main aspects.

From these two aspects it is possible to observe two types of interactive marketing: the interactive digital advertising and the virtual brand community.

- Interactive digital advertising: Taylor (2009) explained five principles based on the digital advertising, most of them related to consumer understanding and insights: considerations of privacy, trust, relevance, experiential values of interactivity and entertaining. Different studies (Hoy & Milne, 2010, et al) demonstrate that users tend to avoid advertising in a context of social networking to protect their privacy, because they are skeptical about the persuasion of advertising in social media.
- ➤ Virtual brand community: virtual brand community is a set of brand-consumer relationships articulated online (Muniz & O'Guinn, 2001, et al). These kinds of community are tools for marketing communication, because they reveal brand loyalty and consumer needs (Casalò et al, 2008).

2.7 USER GENERATED CONTENT

User generated content stand for any type of content, such as image, video, text, and audio, that have been posted by users on online platforms. It can be defined as encompassing of the alteration/production or the sharing features, or both. Moreover, it is possible to identify two key features: UGC involves amateur or petty production of original contents or the alteration of the existing ones; then, it involves sharing it with others, posting on a website or a personal blog (Eynon and Malmberg, 2011, et al). It could be interesting to understand which ones are the gratifications or motives associated with user-generated content online. Papacharissi (2003) analyzed some bloggers and revealed six major motivations for blogging: self-expression, social interaction, entertainment, passing the time, information and professional advancement. Also, Nardi et al (2004) identified five motivations for blogging i.e. documenting one's life, providing commentary and opinions, expressing deeply felt emotions, articulating ideas through writing as well as forming and maintaining community

forums, but both the studies converge in the idea that the more users are gratified with the content generation experiences, the more they will participate in UGC activities.

This kind of content generation is being increasingly used, so the quality of the information, shared by blogs, user-generated websites or similar, becomes a concern. Indeed, everyone can post something on the internet or comment products or services even if he does not know the topic. For these reasons, Chen et al (2011) suggested the utility of a moderation system, already used in some websites as Slashdot: every comment or post is checked by "moderator", i.e. other users, who can assign a score as informative or redundant; this is a method to ensure the content quality.

A method to understand the influence of brand-related UGC on consumers' attitude is the S-O-R model: Mehrabian & Russell (1974) introduced this method, composed by Stimulus, Organism and Response. This model can be utilized to demonstrate the external influences on consumers (S), the internal processes to respond to that influences (O) and the resulting behavior (R). Bagozzi (1983) went in details about that:

- > External influences: they include managerially controllable factors as advertising, price, product design and others non-controllable ones as competition and economic conditions.
- ➤ Internal processes: emotional responses as fear and arousal or cognitive responses as perceived risk and expectations.
- ➤ Behavioral responses: they include activities that lead to choices (actual, past and future), outcomes and the reactions to the choices.

In the tourism, UGC are considered as a new form of word-of-mouth for products, services and their providers; especially since internet revolutionized the distribution and the influence of word-of-mouth, individuals make their ideas easily accessible to other Internet users (Dellarocas, 2003). Online user generated reviews became an important source of information for both consumers and online retailers (Zhu & Zhang, 2006): in fact, hundreds of millions of potential visitors consult online reviews each year and are affected by them. UCG encompasses a variety of different kind of content: from the encyclopedic entries as Wikipedia to blog and post of music videos on YouTube; it includes some forms in which audience-producers are sold as a commodity to advertisers and other activities like the file-sharing (Fuchs, 2009).

In the context of the online news media, Thurman (2008) exposed seven formats for participation online, found thanks to a survey done by some news websites; there are polls, i.e. topical questions to which readers are asked to make a multiple choice or binary response; have your say, that is the features where the journalists post topical questions to which readers send written replies; chat rooms; Q&A, i.e. the interviews with journalists and invited guests where the questions are posed by readers; blogs with comment enabled; post-moderated message boards, where the comments of the users are posted without initial moderation, and the pre-moderated ones, with the vet before publication.

Staying on the same field, Harrison (2010) underlines four categories of UGC that delineate a new journalists' way to relate them with the UGC audience:

- ➤ Unsolicited news story: they can break or create a news story, starting from emails of readers and expanded thanks to traditional journalist techniques of investigation.
- ➤ Solicited content for specific extant news stories: contacts and sources are sought by journalists from the UGC contacts database, that is continuously cultivated and expanded.
- > Expeditious content for specific items and features: it is a part of the forward-planning routine to enhance future stories.
- Audience watchdog content: it is the reaction of the audience that try to influence the reports or covers of a particular news story.

The presence of UGC has changed also the online video market; in parallel with the traditional video-on-demand, new platforms were born in which the users are not only consumers but also publishers. The length of UGC videos is shortened and so is the production time and thanks to their scale, dynamics and decentralization their popularity is ephemeral and unpredictable. Moreover, the unlimited choice of content and the convenience of the web allow to personalize the viewer experience (Cha et al, 2009).

In the following paragraphs, different kind of user-generated content are analyzed.

A particular typology of user-generated content is the word of mouth (WOM): it is the diffusion of information and suggestion between different people, through social networks, blogs or face to face speeches. It is a new frontier of the exchange of information and it can be used in different areas. For example, WOM can be used as a source of travel information: travelers tend to trust more in word-of-mouth than in others commercial sources as travel

agents or accommodation operators, because there is lack of commercial self-interest in WOM recommendations (Litvin et al, 2007).

Another way to use the word-of-mouth is in an online context, where it is called WOM marketing; it occurs when consumers share their experiences and views about products or services they had purchased (Hyung-Park et al, 2007); moreover, Senecal & Nantel (2004) noted that product recommendations from other customers are more important to the prospective purchasers when an experiential product is concerned instead of a tangible product.

Dicther (1966) explained four inter-related motivators of consumer involvement in WOM communication:

- ➤ Product involvement: how WOM acts as a tension-releasing mechanism that drives consumers to share what they know.
- Message involvement: how the discussion is stimulated by messages created by the brand or other consumers.
- > Self-involvement: consumers participate in content generation to enhance their image.
- ➤ Other involvement: the users' genuine desire to help others for example understanding the brand to make the wright purchase decision.

Since the WOM is very diffused, all the brands should be able to mitigate negative online interactions between consumers; for this reason, they should find the negative electronic word-of-mouth (NWOM) to avoid some detrimental effects in all the decision-making process, from the brand evaluation and choice to the purchase behavior and loyalty (Vermeulen and Seegers, 2009, et al).

NWOM has a very strong impact especially in brand evaluations, even more than the positive word-of-mouth (Chiou and Cheng, 2003); in particular, Ahluwalia (2002) explains that negative attributes are believed to be more symptomatic to poor quality product than the positive are for high quality product; consumers pay more attention to NWOM than the positive ones, so the negative evaluations have stronger effects than the positive reviews in terms of reach and impact (Sen and Lerman, 2007).

Reviews are a particular type of word-of-mouth: people can give a rating of whatever they want, for example restaurants in which they eat or hotels and places where they spend the holiday. Goldenberg et al (2001) thought that consumers' decision-making processes are strongly influenced by what other customers think, while Gretzel and Yoo (2008) found that

reviews from other consumers are more up-to-date, enjoyable and reliable than the information provided by travel service providers: these two studies confirm that reviews is one of the most important UGC today.

Another method to generate UGC is the online communities: they are specialized, non-geographically bound community based on a structured set of social relations among admirers of a brand (Muniz and O'Guinn, (2001). McAlexander et al (2002) explained that there are two main characteristics of the online communities: the creation and sharing of meaning. Even if several studies demonstrated that mainly younger people participate in the communities (Hargittai, 2007, et al), it could be shown that there is a trend of older users, especially for the social networks. The reasons why older people use less the communities and social networks are the inability to see how UGC can meet their needs and desires, the fear that the system is too hard to learn, the lack of social influence or some negative attitudes towards digital technologies (Karahasanovic et al, 2009).

Since the knowledge and efforts of users are the primary input to develop a community of user-generated content, it is important to understand how these factors can increase their perceived value; three aspects are taken into account to explain what is previously written:

- Number of contributors: higher the number of contributors is, higher should be the efforts and the energy dedicated to create content and to provide knowledge and abilities for content creation; at the same time, an excessive number of contributors negatively influences the value of UGC; thus it is important to understand when it should be better to stop the number of contributors, to avoid negative impacts (Foutz and Jank, 2010, et al)
- Network embeddedness: higher the degree of a person/project connection with the others, better is for the value of UGC (Grewal et al, 2006, et al).
- Content age: the value generated is different if the content is new or old.

There are many reasons for which the users join a community: to feel socially connected (Sarason, 1974) one of the essential psychological need, but also the desire for social interaction is a motivation for consumers to engage in content generation activities (Henning-Thurau et al, 2004); furthermore, brand communities support their members in terms of sharing necessary information from various sources and emphasizing different values (Szmigin & Reppel, 2001, et al).

Finally, the most interesting UGC is social media: they are a group of internet-based applications that allow the creation and exchange of user-generated contents (Kaplan & Haenlein, 2010). The same authors explained that the two most important aspects of social media are creation and sharing of content, that together with the most important aspects of brand communities previously described form an ideal environment for customers.

A particular speech is reserved for some social networks that are typically considered as user generated contents, for example Facebook, Instagram and MySpace; in fact, Stepchenkova et al (2013) thought that photographs are the method to capture the reality and they are used in these social networks as a way to promote, advertise and distribute goods and to provide precise and accurate information about the services.

2.8 AUTHENTICITY

The term *Authenticity* is a key concept in this work: it comes from Latin and Greek words and it means worthy of acceptance, trustworthy, not imaginary and conforming to an original (Capannelli & Capannelli, 2004). Over time, and depending to the author, authenticity has taken different meanings: the most common is related to genuineness, reality or truth of something (Peterson, 1997 et al). Someone considered authenticity as sincerity, innocence and originality (Fine, 2003) and others a social construction that can be attributed to an object or an event that conforms to an idealized representation of reality (Beverland, 2005 et al). Finally, Boyle (2003) referred to some concepts such as being natural, honest, simple and unspun.

The authenticity is treated especially in the tourism studies. Reisenger & Steiner et. al. (2006) described two different kind of authenticity: the object authenticity and the subject authenticity, which includes constructivism and existentialism.

Object authenticity is defined as the original that confers legitimate authority and power (Wang, 1999); it is the desire to visit some original sites or artifacts and the experience and knowledge resulting.

The constructivism is a mutual meaning-making process where tourist actively construct their meanings in negotiation with different environmental factors: there is not a single objective reality but multiple realities that depend on situation and context. Here the authenticity is socially constructed and evolves over time (Kim & Jamal, 2007).

The existentialism is an authenticity where differences between real and unreal are no longer perceptible (Kim & Jamal, 2007) there is the interest to escape everyday life and an experience of true self in a foreign environment.

Grayson & Martinec et al (2004) underlined another division: the authenticity can be considered as a consequence/output of the tourist experience but also as an antecedent/input for tourist behavior, as it is considered an important driver and value.

2.8.1 Brand authenticity

Now the focus pass on the brand authenticity, i.e. the perceived authenticity of the brand from the point of view of the customers.

Felicitas Morhart et al (2015) reintroduced the three perspectives already mentioned in precedence: objectivism (authenticity as a quality inherent in an object and evaluated by experts), constructivism (authenticity is a socially or personally constructed phenomenon and the reality is the result of different interpretations of what the "real world" looks like) and existentialism (authenticity as being true to one's self).

Moreover, from the point of view of the same authors, these perspectives led to find four dimensions of the brand authenticity: continuity, reliability, integrity and symbolism.

- Continuity: it is the brand ability to never change over time and to survive trends.
- ➤ Reliability: it is the brands' willingness and ability to deliver on their promises; it is important because consumers associate authentic brands with high level of reliability.
- > Integrity: it is the virtue reflected in the brand's intentions and in the values it communicates.
- > Symbolism: it is the symbolic quality of the brand that consumers can use to define who they are or who they are not.

Bruhn et al (2012) also discussed some dimensions of the authenticity; in addition to what it has just wrote, other two dimensions were introduced: originality and naturalness.

- > Originality: it is the brands' ability to be creative, original and innovative.
- Naturalness: it is the brands' capability to be genuine and natural.

Focusing on small-medium enterprises (SME), Fisher-Buttinger & Vallaster et al (2008) conceptualized the brand authenticity in other three dimensions: brand consistency, brand customer orientation and brand congruency.

- ➤ Consistency: it occurs when the promise made to stakeholders aligns with the corporate values, strategies and vision.
- ➤ Customer orientation: it is the organizational process that generates and sustains a shared sense of brand meaning that provides a superior value to customers and stakeholders (Ewing & Napoli, 2005).
- Congruency: a brand should have congruency between individual value and brand value. In order to better understand how much the brand authenticity perception can be a key concept in the brand strategies' decision, some antecedents and some factors, that are influenced by authenticity, will be described.

The antecedents, i.e. the factors that influence the authenticity, can regard behavior unique to the brand and consistent brand behaviors; respectively, the authenticity is influenced by rare brand behavior and stable brand behavior. In particular, there are four manifestations that positively affect the perception of brand authenticity: uniqueness, scarcity, longevity and longitudinal consistency.

Uniqueness is the extent to which consumers perceive a brand as unusual or atypical compared to the competitors; scarcity is the extent to which consumers perceive that the goods or service outlets of the brand are not widely available or accessible; longevity refers to the customers perception of a long period of time existence of the brand; longitudinal consistency is the extent to which consumers perceive that the brand has not changed over time (Moulard et al, 2016).

Authenticity perception affects different key factors through which a brand could construct a competitive advantage on the competitors.

Moulard et al (2016) thought that brand authenticity has positive effects on two perceptual consequences: the expected quality and the brand trust. Indeed, brand's high-quality products may be an indication of its authenticity and additionally, brand managers who are perceived motivated should elicit consumers' trust: for these reasons, authenticity can be considered "antecedents" of both expected quality and brand trust.

The same authors tried to describe authenticity as a "mediator", that carries the effects of the antecedents previously reported on the two outcomes.

Gursoy et al (2015) analyzed the relationships between authenticity and brand equity, in particular brand awareness, brand image, perception of quality and consequently the brand loyalty. Brand awareness is the customers' ability to identify a brand in their memory; brand

image meant to associate a brand with a particular attribute, situation or logo; perception of quality is an intangible feeling about a brand, very important to generate firms' values. These three dimensions influence the brand loyalty, so it is possible to say that the authenticity indirectly affect also the loyalty.

Lastly, Choi et al (2014) introduced brand attachment, i.e. the strong emotion that consumers can obtain from the long-term relationship with brands, brand commitment, that is the continuous passion for maintaining a valuable relationship, and the brand loyalty in the fashion sport brands. It is affirmed that brand authenticity positively affects all these three dimensions and these effects will differ between different brands.

2.9 CREDIBILITY

Credibility is the quality of being trusted and believed in. Previous research defined credibility as a set of different dimensions: the primary two are trustworthiness and expertise, while the secondary ones are dynamism, composure and sociability (Berlo et. al., 1969; Gass et. al., 1999; Hovland et. al., 1953; McCroskey, 1966; Perloff, 1993; Whitehead, 1968). Both people and source credibility are defined by these five factors.

- > Trustworthiness: reliability of a person / explicit policy statement or lack of commercial content
- Expertise: there are two factor that explain this concept
 - o Informativeness: information quality and availability
 - o Reputation: credentials or certification
- > Dynamism: can be communicated through various dimension of appearance (layout, graphics, colors, font)
- ➤ Composure: order and clarity
- Sociability: the quality of being social / ability of being close to the user

2.9.1 People credibility

People credibility is the credibility of a person. To asses it in a determined field it's necessary to evaluate his/her expertise level in that specific field. An example of this was done by Anderegg et. al. (2010), where they examined expert credibility in climate change. In this

study, the authors used two dimensions to define which are the credibile researchers. The two factors are climate-specific expertise and overall scientific prominence; the first is the number of climate-relevant publications authored or coauthored by each researcher, while the second is the number of citations for each of the researcher's four highest-cited papers. Then they used an a priori criterion that a researcher must have authored a minimum of 20 climate publications to be considered a climate research.

Another example regarding the expertise level of a person is the study made by Lev-Ari & Keysar (2010), here they defined the beavibility of a person according to his language. They studied why people do not trust non-native speaker. People are influenced and believe more native speakers. The iusse is raised by the fact that non-native speaker have an accent, and this could cause them to seem less credibile for two main reasons: the accent serves as a signal and the accent makes the speech harder to process. The first motive makes people understand that a person is an outsider, generating prejudice impatting on the credibility of the person. Everything is iniflunctial when the speaker has a mild accent, damaging only the people with a heavy accent. This, due to the increase in mobility, can the reduce the possibility of finding a job for the non-native speakers, especially in job where talking is at the base.

2.9.2 Source credibility

Source credibility is the credibility of a source; it has been defined as the judgments made by perceivers concerning the believability of a communicator (O'Keefe, 1990). For consumers, a credible brand assures the purchasing of quality that they can count on; for companies, a credible brand implies more cost-effective marketing efforts, because of the heightened likelihood of message acceptance.

The first to produce empirical evidences that the believability of a message is strongly influenced by the source were Hovland & Weiss (1951), who demonstrated that the same message proposed by two different sources (a known expert and one less trustworthy) was perceived in different ways, in terms of credibility. After that, the study was ampliated by Berlo et al (1969), that found some factors related to the source credibility, as safety, qualification, dynamism and sociability.

With the increasing amount of information available, the gatekeeping is an activity extremely important, because the gatekeepers should check the information for veracity to ensure the credibility of these information (Salcito, 2009). With the new online media environment, the gates are located not only at the information providers but also at the information consumers: this change has created the shift from the traditional "gatekeeping" to the newer "gatewatching" (Bruns, 2008). The gatewatchers promote and diffuse information by making sources or stories known to other in the media environment.

There are several frameworks to understand credibility assessment. An example is Fogg's (2003b) Prominence-Interpretation Theory, in which it is explained that two things need to happen to make a credibility assessment: the user notices something (prominence) and the user makes a judgment about what he/she notices (interpretation); moreover, the same author listed some factors that influence prominence (involvement, topic of website, task of the user, experience of the user and individual differences of users) and some drivers that affect interpretation (user assumptions, user skill and knowledge, context).

The credibility of online news seems to be positive in general; people trust Internet as much as the other media, except done for the newspaper (Flanagin & Metzger, 2000). Furthermore, Internet is the most common source among under 30 people and second only to the television in the case of general audience. Flanagin & Metzger (2007) explained that the perception of credibility online is strongly influenced by style-related attributes as visual design, which are not related with the content itself; also the supposed gender of the author of the post may influence the perception of the users (Armstrong & McAdams, 2009).

Johnson & Kaye (1998) interviewed some internet users on how they found newspapers, magazines and websites, to understand if the new method of online information are perceived as credible as the traditional ways; the results showed that more than two-thirds of respondent thought that all these kind of information online are credible, and also that the online version of newspaper and literature are judged significantly more credible than their traditional counterparts.

It is necessary to differentiate between various online sources and source levels; in particular, Flanagin & Metzger (2007) measured the perception of three types of credibility, that capture the relevant notions of believability in the web environment and allows some tests of the perceived online credibility.

- Message credibility: it depends on the aspects of the message itself as information quality, accuracy, currency and language intensity; all these characteristics have an impact on the perception of competence and trustworthiness of messages in the online environment (Metzger et al, 2003).
- > Site credibility: it may vary by site features that engender greater or lesser credibility, such as visuals or amount of information.
- > Sponsor credibility: it relies on the evaluations of the website's sponsor.

Some typology of sources credibility can be advertising, media and web sites.

Advertising can be used as a signal of a firm's commitment to its products or services quality (Nelson, 1974); the expenditure on advertising perceived by customers can be considered to infer quality when there is no other information about the products (Kirmani, 1990). Moreover, high-quality firms that can recover their advertising expenditure from future sales use advertising as a signal to ensure that their products or services claims are credible (Rao et al, 1999).

The digital networked media have changed the information landscape: in this new environment, everyone has access to an unprecedent amount of information available for public consumption. Some consequences are that there is a need to research the most credible information and that the filter of information cannot control everything and, as a result, there can be information out of data, incomplete or inaccurate. (Rieh & Danielson, 2007). Moreover, there are few standards for quality control and evaluation, and, at the same time, there are many potential targets of credibility evaluation that work simultaneously, and all these factors affect users' credibility judgments (Flanagin & Metzger, 2011).

Finally, web sites are crucial to ensure sales or to procure services. The first impression of users is essential, because in few seconds they will decide if they stay or move to another site (Robins & Holmes, 2008): many factors influence this choice, two of which are page aesthetics and user's judgement about the site's credibility. The research of the authors is based on two different kind of aesthetic treatments: low and high. The results demonstrated that a low aesthetic treatment creates a "low-budget" impression and consequently the feeling that the content of the site is not credible, while the high aesthetic treatment creates a lucid impression about the site's intentions and so the feeling that the content is credible. Since there were no previous studies about the topic considered, it is performed an exploratory analysis that aims to highlight the differences between BGC and UGC.

3. DATA COLLECTION AND METHODOLOGY

3.1 DATA COLLECTION

To collect the initial database, an external brand ranking was used to create a list of potential brands, from which collect the videos; indeed, Tafesse (2016) explained that an external brand ranking as a sampling frame is a common approach because it provides a list of brands from which a useful sample can be derived. By using the "100 Best Global Brand" list from "Interbrand.com" as a population of brands, it is selected the sample on the basis of the brands' activities in themselves official channel on YouTube with an active audience (so the possibility for users to like or comment), the uniform and global social media strategy, and the existence of UGC regarding the brand on YouTube. Following these characteristics, only eleven out of hundred brands were considered appropriate: Louis Vuitton, Zara, Hermes, Gucci, Porsche, Cartier, Ferrari, Tiffany & Co., Jack Daniel's, Burberry and Prada.

After the choice of the sample, the data were collected during a year. In particular, 25 videos from each brand were taken from the official channel (Brand-Generated Content); on the other hand, 25 videos were randomly collected through Google researching the "Brand Name" as a keyword (User-Generated Content), taking into account only videos from YouTube, created by an influencer and in the time framework previously agreed.

Moreover, for each video the following information were collected:

- The reference number for video:
- > The name of the brand;
- The type of channel: brand-generated content or user-generated content;
- > The name of the sector:
- > The date of data collecting
- > Other brands mentioned: if the brand and product is central or not in the video;
- The link of the video;
- The name of the channel;
- > The number of subscribers to the channel;
- The date of publication (mm/dd/yy);
- > The number of views;
- > The number of likes and dislikes;

- > The number of comments;
- The video duration (s);
- The subject of the video;
- > The main message of the video;
- > The secondary message of the video;
- ➤ The type of influencer: individual or specialized.

3.2 DATA ON VIDEO LEVEL

In this level, the database is formed by all the data collected for each video; consequently, each row includes data related to a single clip. The variables used are different: there are video & channel characteristics, numerical measures of engagement and language measures of engagement.

3.2.1 Video & Channel characteristics

The first element included in this category is the number of subscribers; it indicates how many people follow the page in which videos are posted. The mean of this items for BGC is 236,838, while for UGC is 557,719; however, it is noticed that the clips of each brand's official page count the same amount of subscribers, meanwhile this number is different for each influencer's clip.

Another characteristic of the videos is the age, i.e. the time passed from the moment when they are uploaded on YouTube (calculated in days). On average, BGC clips are online from 313.31 days and UGC ones from 290.81 days.

Lastly, the length of each video is collected; it shows the size in minutes of the clips and so the time a viewer must spend to watch the entire tape. The means of this variable are 1.86 minutes for BGC and 12.31 minutes for UGC.

3.2.2 Numerical measures of engagement

As already mentioned in the literature paragraph, the numerical measures of engagement are views, likes, dislikes and comments. To have an idea of their behavior, some ratios between

them are added: Views/Subscribers, Likes/Views, Dislikes/Views, Comments/Subscribers and Comments/Views.

A view is defined as a person who watch the video. It is counted when the watcher sees the clip for at least 40 seconds or, if the video is shorter, when it is ended. The number of views is more and more synonym of the popularity of a video, since it indicates that many people watch it. In the case studied, the average of views for BGC is 732,146, while for UGC is 254,874.

Differently, likes, dislikes and comments can be left even if the video is not watched. In theory, a like indicates that the clip is enjoyed by the viewers, while the opposite is showed by the dislike. Respectively, the means are 1,102 and 45 for BGC, and 5302 and 352 for UGC.

The comments are all the sentences written by the watchers to express their opinion or to ask questions to the protagonists of the videos. There is also the possibility "to comment a comment", i.e. to write an answer under the opinion of another viewer, with the opportunity to create a network between users (it will be deepened subsequently).

Some further details could be done by ratios between the precedent items. For example, it could be interesting to see how many viewers there are compared to the number of subscribers, that is the only ratio higher than 1 (particularly, 6.5 for BGC and 3.4 for UGC in average). Others could be likes and dislikes compared to the views (respectively 1.44% and 0.07% for BGC, and 2.5% and 0,17% for UGC in average) and comments divided for subscribers and views (respectively 0.02% and 0.10% for BGC, and 1.23% and 0.84% for UGC in average), but they are clearly very small.

3.2.3 Language measures of engagement

Here, there will be included the variables related to the characteristics of the comments through which people give their opinions under the clips. The main fields are the feelings expressed (sentiment and emotionality) and the writing characteristics (number of words, complexity and analytics).

The sentiment measures the positivity (or negativity) of the language used; it has a range between 0 and 1, where 0 is extremely negative and 1 is very positive. On the other hand, emotionality is the deviation from the neutral sentiment, indicating the intensity of the

feelings expressed in the comment. The averages of the sentiment are similar between BGC and UGC (respectively 0.54 vs 0.52 in average) and the same happens for emotionality (0.29 for the official pages and 0.28 for the influencers' pages).

The number of words simply counts the words included in the comments (in average 7.30 for BGC and 10.43 for UGC), while complexity and analytics describe in different ways how much the language used is complicate. In details, the complexity measures how much the comments are complex to understand (5.86 for BGC and 5.76 for UGC), while the analytics captures the degree to which people use words to suggest formal, logical and hierarchical thinking patterns (in average 63.77 for BGC and 58.86 for UGC). Particularly, people low in analytical thinking tend to write in a more narrative ways, focusing on hereand-now and personal experiences; meanwhile, the higher analytical persons perform better in college, with higher college board scores.

3.3 DATA ON BRAND LEVEL

Here, the data collected for each video are aggregated, forming a database in which every row contains data regarding the BGC or UGC part of a brand. In this context, the variables previously described for the video level are the same, considering that the data were aggregated. Moreover, other kinds of variables are taken into account: the network engagement ones (actors and links among actors) and the semantic brand score.

3.3.1 Network engagement

These are variables related to the people that create links with the owner of the video and also with other viewers. Actors and links among actors are the variables that fall in this category.

Actors are the persons who comment under the video; it simply counts the number of users that leaves a comment and so that interacts with the channel. In average, there are 1,141 actors for BGC and 8,641 for UGC.

Links among actors count the interactions between users; this index takes into account the conversations generated through comments and replies between viewers. Following the behavior of the number of actors, its means are 478 for BGC and 3,264 for UGC.

3.3.2 Semantic Brand Score

It represents the strength of the brand in the comments. As already described in the literature paragraph, it counts the number of times that the brand is mentioned in the comments, without considering the sentiment for which it is written. The average is 75 for BGC and 87 for UGC, even if a bigger value does not imply necessarily a better situation.

3.4 SURVEY

After the study of the already existing literature, the focus was on two fundamental characteristics that a viewer must perceive from the video: authenticity and credibility.

To understand the perception of these features, two brands, on which concentrate the efforts, were selected; in particular, it was thought to choose a typical-male brand and a typical-female one, in order to guarantee that the persons who will see the clips are potentially interesting in them; going through the list of brand, the most appropriates were Ferrari as a male brand and Tiffany & Co. as a female brand.

With the aim to have a hit, the comments under each video were analyzed to see if effectively the clips of Ferrari have a majority of men's comments and the ones of Tiffany & Co. have mostly women's comments. This analysis confirmed the previous idea: Ferrari tapes are commented especially by men and the ones of Tiffany & Co. by women.

Then, the structure of the survey was designed. It was found in the literature that the two key aspects to describe authenticity and credibility could be influenced by different variables; in details, the first was divided into reliability (brand's willingness and ability to deliver on their promises) and integrity (the virtues reflected in the brand's intentions and the value it communicates), while the second was split into trustworthiness (reliability of a person or of a source) and expertise (quality and availability of information and the presence of certifications or credentials). It was also added a question to understand the familiarity with the brand at the beginning of the survey and some about the age, the gender and the occupation at the end; for all the statements, excluded the last three (age, gender and occupation), it was used a Likert scale from 1 to 7, where 1 is completely in disagreement with the sentence and 7 is completely in agreement with it. Another problem was that the

videos were too long to be included in the survey: for this reason, after the decision of the right length (between 30 seconds and 1 minute), the clips that exceeded were cut.

The final structure of the survey (appendix 7.1) is the following:

- A question to understand the familiarity with the brand.
- ➤ 6 videos, equally distributed between BGC and UGC; after each clip there are 8 statements:
 - O Two about reliability, in order to understand if the expectations about the brand (or the speakers) are met and if the things said in the clips are believable.
 - Two regarding integrity, to evaluate the genuineness and spontaneity of the clips, together with the impression that who watch the tapes are important for the brand (or the influencer).
 - Two for trustworthiness, to figure out if the video is well made (in terms of sound,
 editing and images) and if the information shared are unbiased and fair.
 - Last two related to expertness, to assess the skills and the competences of the speaker, together with the quality and accuracy of the information shared.
- The last 3 questions regarding the age, the gender and the occupation.

The database has 48 videos for Ferrari (24 BGC and 24 UGC) and 45 for Tiffany & Co. (21 BGC and 24 UGC). There were created 8 different surveys for Ferrari and 8 ones for Tiffany & Co., in order to have different videos per version; since the same number of questionnaire for each brand was desired, it was decide to use two times three Tiffany & Co. BGC tapes with the aim to have 8 survey also for this brand. In order to reach a good number of replies, the threshold was placed at the minimum of 20 answers for each version.

3.5 METHODOLOGY FOR THE ANALYSIS

The procedure used is the same for each level. In particular, it is divided into descriptive analysis, discriminant analysis and exploration of correlation.

3.5.1 Descriptive analysis

It is a first step that gives an idea of the distribution of the data; here, it is calculated the mean, the standard deviation, the maximum value, the minimum value and the overall of each variable, following these formulas:

$$Mean = \frac{\sum_{i=1}^{n} x_i}{n}$$

Standard Deviation =
$$\sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}}$$

$$Minimum = min_{i \in n}(x_i)$$

$$Maximum = max_{i \in n}(x_i)$$

Overall =
$$\sum_{i=1}^{n} x_i$$

3.5.2 Discriminant analysis

It compares the data from two groups (BGC and UGC) by using the t-tests; indeed, they determine if the means of two samples of data are considerably different from each other. All the t-tests were performed on Excel.

Hypotheses t-test
$$\rightarrow$$
 H₀: $\mu_1 = \mu_2$; H₁: $\mu_1 \neq \mu_2$

3.5.3 Correlation analysis

It is a statistical evaluation used to examine the strength of the relationship between two continuous variables (numerically measured); this analysis is made to understand possible

connections between two variables. It can be positive or negative: positive when the two variables increase simultaneously, while negative if when a variable increases the other one decreases; according to this, the coefficient of correlation ranges between -1 and +1. Values close to the extremes are considered as strong correlation, while the ones between -0.1 and 0.1 indicate an absence of correlation. For all the other values, it was necessary to do a significance test: if the p-value of this test is lower than 0.05, the correlation is significant and so it is taken into account for the analysis. All the calculations of correlation and its significance were performed using the software R.

Correlation =
$$\frac{cov\left(A,B\right)}{\sigma_{A}\sigma_{B}}$$
, where Covariance (A,B) = $\frac{\sum_{i=1}^{n}(x_{i}-\bar{x})(y_{i}-\bar{y})}{n}$

4. RESULTS AND DISCUSSION

4.1 VIDEO LEVEL

4.1.1 Video & channel characteristics

In this paragraph the video and channel characteristics (number of subscribers, age and length) are analyzed. T-test is useful to highlight the differences between each channel variable. The result shows the clear divergence between the number of subscribers of the two channels; in fact, the unofficial one, in average, have more subscribers than the official one (410k vs 245k, Figure 2). So, users prefer to follow more influencers than brands; this

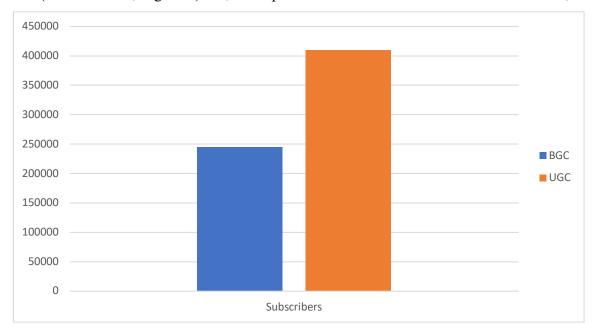


Figure 2 Average number of subscribers

can perhaps be explained by the facts that users feel closer to someone like them; another possible reason could be related by the fact that most brand content is advertisement, in which users are not interested.

A big difference between BGC and UGC is shown by the length. Indeed, the average duration of a brand video is around 1.8 minutes, while it is 12.2 minutes for an influencer video. In accordance with the previous affirmation regarding the advertisement nature of a BGC clip the length reflects fully the shortage of such video. UGC tapes instead last longer

since influencer gives an accurate description and experiences of their reviews and unboxing (these are mainly users' video content).

The age of the videos was not taken into consideration because they are all collected in the same timeframe.

4.1.2 Numerical measure of engagement

This paragraph exposes the differences between the numerical measure of engagement variable, statistically based on t-test. In discordance with subscriber numbers, the views for BGC videos are higher than the ones for UGC. BGC clips nature leads the views to increase not only by YouTube but also thanks to the advertisement before other videos and to the banners that appear in other websites. UGC clips longer duration affects the views (the length of the tape itself bring to a view or not, long videos are usually less watched). To get one view, a clip should be watched for a determined amount of time proportional to its length; for this reason, BGC tapes have a higher probability than UGC ones to receive one (Figure 3).

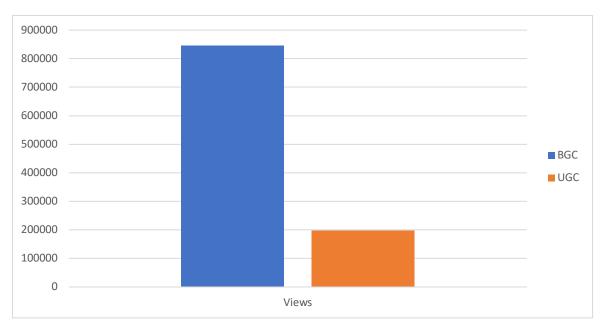
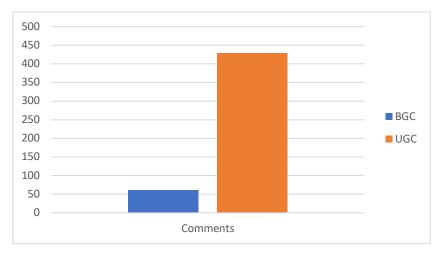


Figure 3 Average number of views

The amount of view would bring to say that the interaction is higher in BGC videos, but for the reasons above explained, in reality, UGC clips have more interaction than the others.



and comments are way more in the influencer channel. The variable that most impact on interaction are comments (Figure 4). UGC videos reach, in average, hundreds of comments, meanwhile

Indeed, likes, dislikes

Figure 4 Average number of comments

the official ones barely reach around twenty-five comments per video.

BGC	Views Likes Dislikes		Comments	
Views	1	0,18**	0,39***	0,12
Likes	0,18**	1	0,62***	0,97***
Dislikes	0,39***	0,62***	1	0,55***
Comments	0,12	0,97***	0,55***	1

Table 2 Numerical engagement measures correlation matrix BGC

UGC	Views	Likes	Dislikes	Comments
Views	1	0,77***	0,87***	0,67***
Likes	0,77***	1	0,47***	0,86***
Dislikes	0,87***	0,47***	1	0,5***
Comments	0,67***	0,86***	0,5***	1

Table 1 Numerical engagement measures correlation matrix UGC

The analysis of correlation between the numerical engagement variables highlights more the difference between BGC and UGC. Indeed, view in the influencer channels results related with likes, dislikes and comments (Table 2), while in the official one the link is low or absent (Table 1). However, the study of the significance level shows that, even if the correlation of view variable with the others is not that strong, they can be considered in relationship. The only link with p-value not acceptable (7%, a bit above the threshold 5%) is the one between view and comment.

4.1.3 Language measure of engagement

Here, the focus shift to the language engagement measures to evaluate the two channels in term of language used and emotions expressed. Sentiment follows approximately a Gaussian distribution for both BGC and UGC, showing that the majority of the comments does not

express what users feel; it is possible to see from Figure 5 that the two peaks are in the center, where the sentiment is neutral.

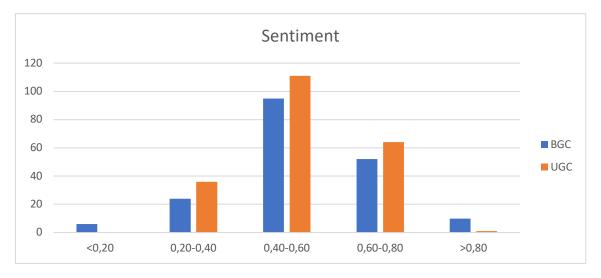


Figure 5 Sentiment histogram

Probably, the reason behind this could be the worry of users to open themselves. When they express their feelings, often these are more positive than negative; there are some isolated cases for BGC where the sentiment is strongly manifested. Also, emotionality can be approximated with a Gaussian distribution in both cases. The central value is included between 0.25 and 0.30, indicating that the distance from the neutral sentiment is really close to zero; this validates the previous idea (Figure 6).

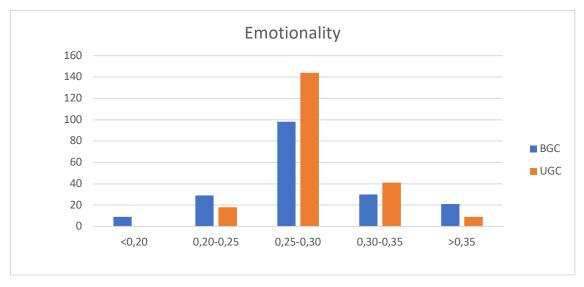


Figure 6 Emotionality histogram

Regarding the complexity, as can be seen from the Figure 7, the comments of UGC are on

average simpler than the BGC ones. The difference is so small that the level of complexity can be considered the same, with some exceptions.

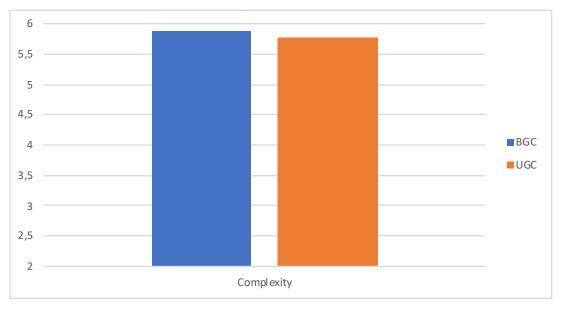


Figure 7 Complexity histogram

Differently from sentiment emotionality and complexity, where BGC and UGC have some similarity, number of words and analytic measurements divide the two channels. In fact, from t-test results that users use more words under an influencer video than below an official one. Moving to analytic, even if the average scores are close to each other (63.9 BGC and 59.1 UGC), comments of BGC videos present a higher variance between the single values, meaning that there are people commenting with really different type of analytical thinking. Instead, for UGC the users are similar, having approximately the same type of reasoning.

BGC	Words	Complexity	Analytic	Sentiment	Emotionality
Words	1	0,01	0,25***	-0,08	-0,08
Complexity	0,01	1	0,34***	-0,07	0,23**
Analytic	0,25***	0,34***	1	0,15*	0,09
Sentiment	-0,08	-0,07	0,15*	1	0,25***
Emotionality	-0,08	0,23**	0,09	0,25***	1

Table 3 Language engagement measures correlation matrix BGC

UGC	Words	Complexity	Analytic	Sentiment	Emotionality
Words	1	0,05	0,15*	-0,14*	-0,08
Complexity	0,05	1	0,69***	-0,27***	-0,01
Analytic	0,15*	0,69***	1	-0,29***	-0,12
Sentiment	-0,14*	-0,27***	-0,29***	1	0,6***
Emotionality	-0.08	-0.01	-0.12	0.6***	1

Table 4 Language engagement measures correlation matrix UGC

The study of correlation between the language measurement starts from the sentiment and emotionality. The two variables as they are defined are related, but for BGC a low coefficient was found between the two (Table 3), this is however denied by the significance test where the p-value show that this correlation is significative. Moving to number of words people would expect a strong connection between it and complexity, but in both cases the link is

really weak with a high p-value too. Analyzing the data both for BGC and UGC, to a higher complexity does not correspond a higher number of words, sometimes it is even below the average. Differently from complexity the relation between analytic and number of words, even with a low coefficient, is justified by the p-value. So, a high number of words lead to a high analytic meaning that the more a person write the more he/she has an analytic way of thinking. Considering the relation between complexity and analytic, from both channels is possible to see that the more analytic a person is the more the writing is complex. In fact, a person with analytical thinking writes in a more specific and accurate way, using also specific terms.

Analyzing the relation between the feeling variables and the others, in the UGC channel it is possible to find a weak negative link between sentiment and number of words (Table 4). It is considered because the significance test shows a low p-value. This connection can be explained by the fact that users who use more words usually are criticizing the video just viewed. As for the number of words also complexity has a weak negative correlation with sentiment, strengthening the fact that users in influencer channel comment negatively, in this case when they write more complex. Lastly both UGC and BGC sentiment have a weak correlation with analytic variable, negative and positive respectively, which are taken into account for their positive significance test. In the official channel the more good sentiment people put the more they result with analytical thinking, while for the unofficial one is the opposite: less is the positive sentiment more is the analytical thinking.

4.1.4 Numerical measure vs. Language measure

In this paragraph the correlation between numerical and language measures has been studied. Starting from views, it is possible to see two different behavior between BGC and UGC part. In the first (Table 5), views presents no relation with all the language variables, while in

BGC	Views
Words	-0,03
Complexity	-0,02
Analytic	0,01
Sentiment	-0,03
Emotionality	-0,06

Table 6 vi	ews a	ınd
language i	measi	ure
7		DO

UGC	Views
Words	-0,15*
Complexity	0,08
Analytic	0,08
Sentiment	-0,29***
Emotionality	-0,17*

Table 5 views and language measure correlation matrix BGC correlation matrix UGC

the second (Table 6), even if there are some weak significative connections (with number of words, sentiment and emotionality), it is hard to find an explanation for the connection between a view and what people write. They can be considered cases.

BGC	Likes
Words	0,1
Complexity	0,06
Analytic	0,04
Sentiment	-0,18*
Emotionality	-0,09

Table 8 likes and language measure correlation matrix BGC

UGC	Likes
Words	-0,16*
Complexity	0,14*
Analytic	0,15*
Sentiment	-0,31***
Emotionality	-0,29***

Table 7 likes and language measure correlation matrix UGC

Analyzing the likes, both channels have a weak negative correlation with sentiment, considered because it is significative, meaning that to a high number of like correspond less sentiment or negative sentiment (Tables 7 & 8). The reason could be

that the people who actually do not like the

videos comment in a negative way. Following the previous reasoning, a smaller number of likes implies more emotionality and so more distance from the neutral sentiment (maybe more negative comments). Looking at the relationships between likes number of words, complexity and analytical, even if they are significative, the low correlation index and the definition of the variables make these links just cases.

Dislike variable, in UGC channel, results to be correlated negatively with the sentiment, so

people who do not like the videos put a dislike plus they leave a negative comment.

Lastly, studying comments correlation with the feeling's variables, in both channels, there is a weak negative correlation, but significative, between sentiment and comment, meaning that

BGC	Comments
Words	0,11
Complexity	0,06
Analytic	0,05
Sentiment	-0,18*
Emotionality	-0,08

Table 10 view and language measure correlation matrix BGC

UGC	Comments
Words	-0,02
Complexity	0,09
Analytic	0,11
Sentiment	-0,33***
Emotionality	-0,28***

Table 9 view and language measure correlation matrix UGC

there are more negative comments. Since in UGC channel the previous link is stronger, it affects the emotionality deviating less from neutral sentiment, meaning that the number of negative comments is reaching the positive ones (Tables 9 & 10).

4.1.5 Discussion

The results above described highlighted the first differences between the two channels at video level. Talking about subscribers, which are more on UGC, in this social media it has a completely different meaning compared to Facebook, Instagram or Twitter. Indeed, people on YouTube look for something particular typed in the search bar. YouTube is the second largest search engine after Google, showing that one of its main usage is search engine one. People search what they look for on YouTube to find a video that: help them in doing things they haven't done, passing the time, decide to buy a product or not (Smith K., 2019). While on the others, persons go for different reasons. Results from existing research on uses and gratification of social media suggest that people use media for three main motives: information processing, entertainment activities and social connection (connect with other people online). This is a big problem for BGC channels because now the users are not anymore just mere consumers, but they should be active part of business value production; the consumer should be the center of the marketing strategies starting from the activities and practices on social media (Heinonen k., 2011; Stafford et al. 2004).

So, people choose to follow what they are interested on Facebook, Twitter and Instagram while on YouTube they don't need to follow a channel to watch the video related to their research. When a person follows a channel is because he/she is interested in being updated with its new content. Users generally follow influencers more than brands because they find their content more engaging than advertisements. In fact, influencers give a story trough their videos; followers are the ones who liked the story and want to keep up with it, having the possibility to find new videos on their homepages. Indeed, from the statistical research about the most followed channel on YouTube appear clearly that YouTubers and not brand are at the top (Boyd J., The Most-Subscribed YouTubers and Channels, 2019). Twitter follows the same path of the previous platform, while Facebook is more homogenous, in the list of pages most-followed there are both person and brand (Boyd, 2019 and Wikipedia, 2019).

It would be expected that more subscribers would bring more views but for this social media is not so. In fact, even if UGC channel has more subscribers, the number of views is higher for BGC one. There are different reasons behind: brand videos are not only posted on YouTube, but also on the other social so they have more visibility and they get more views. Facebook contributions in views is really huge and it brings more views to video that are originally posted on YouTube (Smith K., 2019).

Influencers or users probably do not have all the social media, or they work mostly with one so is not that easy to get visibility. Another reason could be the research results, where the list of the videos can give more visibility to brands than users; indeed, people usually watch one of the first videos. Another factor influencing views derive by the world popularity of the brand, while users' content is usually more restricted to a location and sometimes is difficult that it expands beyond the boundaries of the country, so the viewers for UGC content mostly are local ones (Chatzopoulou, Sheng, & Faloutsos 2010).

Considering the interaction level, given the typology of UGC tapes, users are more interested to ask further information or to give feedbacks to the influencer. There are no reasons why a user would comment below an advertisement, it is rarely done and sometimes do not have anything to do with the content exposed. Users are fundamental for the interaction, and all of them have a role that change the way in which he/she contributes: core member (those who contributed by retrieving, supplying and discussing information), conversationalists (they just focus on discuss information) (Heinonen k., 2011).

All of this is strengthened by the correlation analysis. The big number of views for BGC do not bring equal amount of likes, dislikes and comments. Especially the last one, that resulted no correlated, shows the huge problem for brands: the low engagement.

Differently from Twitter, where there are heated discussions (tweets with more sentiment tend to be retweeted more often and more quickly than the neutral ones) in which people express their feelings asserting their positions against other users, YouTube show a completely different behavior. Users do not express what they feel while writing, and most of comments have no sentiment at all (Stielglitz & Dang-Xuan, 2014).

An interesting result derives from the analytic: it is similar between BGC and UGC, but it has different behavior. In fact, actors' comments on official pages clips are very different, varying from the high analytical thinking to the high narrative type, while they are more similar under the UGC videos. This can be related to the global audience of the official channel, such as the views, because different cultures potentially can imply different way of thinking. Moreover, people commenting under BGC videos are sometimes expertise which bring to a more technical comment compared to the user one, increasing, consequently, the high difference in the analytic (Pennebaker et al. 2014).

A weird founding was the absence of correlation between number of words and complexity; the possible cause is the absence of regularity between the data, but most likely the comments are written with a simple vocabulary even if the level of analytical thinking is high. This just means that people know what they saw and are specific when they express themselves. Furthermore, since everybody can watch the videos, also who leaves a comment try to be understandable by the possible audience.

It's a trend between the social media platforms for the users to express their negative opinions with a long and complex comment; indeed, when people want to criticize a post, they tend to write a lot to be clear as much as possible, trying to be more precise and accurate

in their language. This happens especially for UGC clips, but it could be seen as a stimulus for the influencers to improve and do even better next time. From previous studies results that brand do not use negative comments to improve but they prefer to censor or ignore them. The negative comments are most likely to be posted in response to explicit marketing messages initiated by sponsors. Brands that mix marketing messages and "fun" postings attract significantly fewer negative comments (Dekay, 2012).

These sometimes are texted from the so called "haters", i.e. persons who does not like the influencer/brand and leave negative opinion without considering the content of the post (Dawson, 2018). These ones hide themselves behind the digital world but in reality, they have not the courage to criticize who they target on the social media.

The presence of a negative correlation between comments and dislike with sentiment confirms the precedent statement.

4.2 BRAND LEVEL

4.2.1 Network & semantic brand score

To analyze the network and sentiment brand score the tests were done at brand level, since these variables are available only at this level. Actors is related to the comments, so as the number of comments is higher for UGC videos also the number

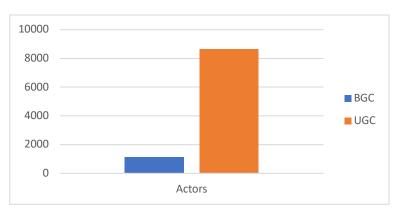


Figure 8 Average number of actors

of actors is greater in unofficial clips than in BGC ones (Figure 8). Indeed, an actor is a person who expresses an opinion under the clip and logically they increase with the increment of comments.

Links among actors count the times when a conversation is generated; this implies different persons that reply each other, increasing the comments. Actually, as it is possible to see from

Figure 9, the values for influencer channel are higher than the BGC ones also for this variable.

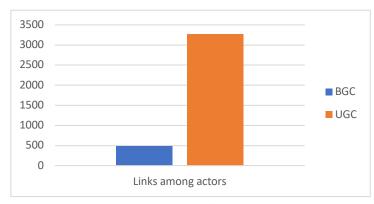


Figure 9 Average number of links among actors

The semantic brand score for UGC is higher than BGC one (Figure 10). Since it counts the number of times the brand is mentioned in the comments, it was expected a bigger value in the influencers' channels, because they have on average

more comments. However, it does not mean that every reference is a positive one, because this index does not take into account the feelings: not always a higher number of mentions is a good thing.

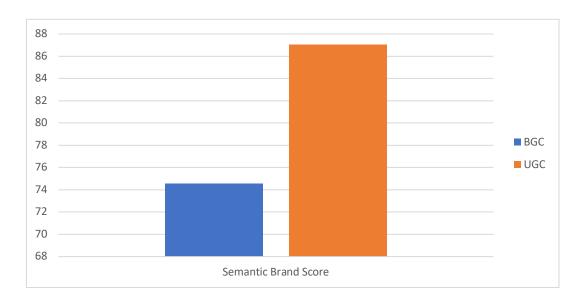


Figure 10 Average Semantic Brand Score

BGC	Actors	Links among Actors	SBS
Actors	1	0,99***	0,85**
Links among Actors	0,99***	1	0,8**
SBS	0,85**	0,8**	1

UGC	Actors	Links among Actors	SBS
Actors	1	0,96***	0,6
Links among Actors	0,96***	1	0,71*
SBS	0,6	0,71*	1

Table 11 Network & semantic brand score correlation matrix BGC

Table 12 Network & semantic brand score correlation matrix UGC

Considering the correlation between the network and semantic brand score variables it possible to see a high Pearson coefficient between all the variables, in both channels. Actors and links among actors by their definition were expected to be related, and the study of correlation confirmed it, for BGC and UGC (Tables 11 & 12). The relationship between the semantic brand score and the other two variables is actually similar between official and unofficial channel, but the study of significance level shows a p-value higher than the threshold for UGC (8.48%) which makes that correlation not significative. The reason behind this can be found in the values of the user's variables where to a high number of actors does not correspond a high semantic brand score.

4.2.2 Network & sentiment brand score vs. Numerical measure

This paragraph explores the correlation, at brand level, between network and SBS variables and numerical engagement ones. Starting from the views, BGC has no correlation while UGC has really strong one. Indeed, as was said before the nature of official channel videos generate low interaction, while the influencers ones generate more interaction (Tables 13 & 14).

BGC	Views
Actors	0,19
Links among Actors	0,19
SBS	-0,09

Table 13 views and network & SBS measure correlation matrix BGC

UGC	Views
Actors	0,88**
Links among Actors	0,86**
SBS	0,82**

Table 14 views and network & SBS measure correlation matrix UGC

So, more view means more actors, more links among actor and so more possibility that the brand is mentioned.

Considering like variable, both the channels have really strong correlation. More likes imply more actors, more links among actors and more brand mentions. Regarding the link between SBS and like, even if SBS does not measure the sentiment, in this case a brand mention can be considered positive because it brings a like (Tables 15 & 16).

BGC	Likes
Actors	1***
Links among Actors	0,99***
SBS	0,86**

Table 16 likes and network & SBS measure correlation matrix BGC

UGC	Likes
Actors	0,96***
Links among Actors	0,94***
SBS	0,68*

Table 15 likes and network & SBS measure correlation matrix UGC

Also, dislike has the same behavior of like variable, but BGC channel present a non-significative correlation with SBS. For UGC the reasoning here is the opposite as before, in fact, in this case a brand mention can be considered negative because it brings a dislike. Comment relation with actors, in both channels, is the only one presenting the maximum value of correlation, generally speaking more people are involved more comments are generated. Comments correlation with links among actors, for BGC and UGC, has a really high value, too; in fact, more users reply to a comment more comments are written (Tables 17 & 18).

BGC	Comments
Actors	1***
Links among Actors	0,99***
SBS	0,85**

Table 17 comments and network & SBS measure correlation matrix BGC

UGC	Comments
Actors	1***
Links among Actors	0,96***
SBS	0,61

Table 18 comments and network & SBS measure correlation matrix UGC

While the last correlation has different behavior; in fact, influencer channel videos present a medium correlation but with low significance (7.8% is really close to the threshold), meaning that people who comments in UGC channel are less incline to mention brand. At the same time, in BGC one, people who comment tend to mention more the brand.

4.2.3 Network & sentiment brand score vs. Language measure

In this section the correlation between network and SBS variables and language ones is analyzed. Starting from number of words, it has no significative correlation in both BGC and UGC videos. So, more words don not bring to more actors, maybe the same actors are writing more. More words do not correspond to more links among actors, perhaps always the same people are generating the conversations. With more words the SBS should increase but persons are writing more without mentioning the brand.

UGC	Complexity
Actors	0,66
Links among Actors	0,74*
SBS	0,9***

Table 19 complexity and network & SBS measure correlation matrix BGC

Considering complexity, as can be seen form Table 19, only in the influencer channel it has

UGC Analytic

a relationship with links among actors and SBS. When users start a conversation, the

UGC	Analytic
Actors	0,71*
Links among Actors	0,85**
SBS	0,88**

Table 20 analytic and network & SBS measure correlation matrix

complexity increase meaning that they go deeper writing more

articulated, increasing also the number of specific terms and so the references to the brand. Also, for analytic in the unofficial channel it is possible to see correlation between it and the other variables (Table 20). Increasing the number of actors, also the analytical thinking increase, due to the participation of more analytic actors, creating a conversation oriented to this way of thinking. As for the complexity, also here the comments present more brand mention.

BGC	Sentiment
Actors	-0,9**
Links among Actors	-0,9**
SBS	-0,87**

Table 22 sentiment and network & SBS measure correlation matrix BGC

UGC	Sentiment
Actors	-0,73*
Links among Actors	-0,85**
SBS	-0,71*

Table 21 sentiment and network & SBS measure correlation matrix UGC

Regarding sentiment, it is possible to find, both for BGC and UGC videos in Tables 21 & 22, strong negative

correlation for all the variables. In both channels when the conversation is more negative the actors, the links between them and the SBS increase, meaning that people participate more in the disccusion where the brand is criticized, mentionig it often.

4.2.4 Discussion

The analysis at brand level helped in discovering other divergences between the two channels. The huge number of actors and links among actors in UGC channels are, as for the comments, due to the content mostly; in fact, people comment more influencers videos. Users do not just type below the video for the action itself, as for the few comments under BGC clips, but they want a conversation, a comparison with other users or with who made the video. By tagging friend or by answering to other users' comments the links increase and that's why the variable is higher for influencer channel (Heinonen, 2011).

The semantic brand score reached a greater value for UGC channel but taking into account the average number of comments for the two channels, this value is more significative for BGC. Indeed, users refer directly to the brand when the video is official, while in the UGC case they still use brand name, but the focus is more on the content and the protagonist of the clip itself. However, this indicator is difficult to interpret since, as already said before, the sentiment is not considered. So, it cannot be possible to say one channel is better than the other because the comment related to the mention is not known. Sbs plays an important role for the brand itself, indeed, the name represents the company and so everything written near the brand name matters for the firm. First thing to do is to recognize who mentioned you: journalist, fans, lost & confused, constant complainers, and enemies. According to this the brand can act in different ways and evaluate also the comment in a different way (Berger, 2015). This helped in understand that BGC channels are more concerned in this metric than UGC one, where it does not have that importance. It can be considered more when the mentions are below a video made by someone collaborating with the brand.

Following this, sbs correlation absence with actors, for UGC, confirms that the comments made by viewers do not present brand name. So, a bigger number of actors does not imply more mentions as for the BGC where the references to the brand increase with the number of actors. In fact, the main argument of the official channel is the brand itself and its products. The correlation analysis confirms the previous thoughts both for BGC and UGC, but it also helped highlighting other divergences. In fact, the views are related with the network variables only in the unofficial channel, this is due to the interaction level, too. The few actors present in BGC channels do not increase with the views, confirming that the users' involvement is too low for brands. Likes outline the similarity between the two channel, showing that the interaction is important for both, denotating even more that the engagement is central in social media (Hollebeek, Glynn, & Brodie, 2014).

The link between comments and actors are the highest, indicating that who write their opinions leave more or less one comment. These lead to the conclusion that there are single

opinions without replies and conversation where different people are replying each other, not always the same. These is evident in all the social media, especially for brand post, where there are lists of single different comments or conversation where the different opinions are exposed just to contrast the other users.

Considering complexity, which is related with links among actors and sbs, only for UGC, it delineates the level of interest users put in influencer channels, commenting more detailed especially when they are having a discussion with other users.

In conclusion, the negative relationship between sentiment and network variables strength the fact that the actors comment mostly negatively mentioning also the brand. As for the video level, also these variables show that the comments are mostly to criticize or express their dissention.

4.3 SURVEY

Previous paragraphs were on the analysis of the different variables; now the focus shift to the study of the two brands used in the survey: Ferrari and Tiffany & Co. Their choice, as already explained in the data collection section, was related to the idea to select a typical-male brand and a typical-female one, in order to potentially capture the attention of all the respondents.

In general, both Ferrari and Tiffany & Co. have higher values for UGC variables. The only exceptions are the number of subscribers for Tiffany & Co. (a very bigger value in BGC), emotionality and complexity for the same brand and sentiment for Ferrari. This implies that influencers are a very important part, with which viewers prefer to interact and give their opinions.

Regarding these two brands, it was implemented a survey to better understand the impression that the viewers have about the videos, both BGC and UGC, in terms of authenticity and credibility. In particular, there were 8 statements for each clip to evaluate them; the analysis of the score obtained helped in describing how much each video is considered authentic and credible from the point of view of who watch it.

The threshold of 20 answer per survey allowed to reach more than 160 respondents both for Ferrari and Tiffany & Co. The data collected were then analyzed trough some t-tests.

The average for the majority of the questions is located between 4 and 5 for both the brands; some values are higher than 5 for Ferrari, while only Q5 is out of this range for Tiffany & Co. (particularly, it is higher than 5 for BGC and lower than 4 for UGC).

T-test helped in comparing the two channels for each brand and through the use of some graphs it was possible to evidence the difference statistically found.

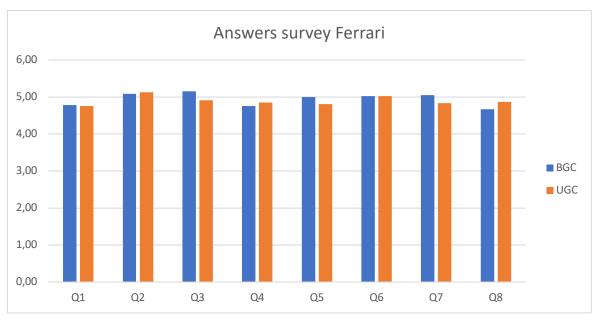


Figure 11 Average answers value for Ferrari

As it can be seen from Figure 11, the perception of authenticity and credibility is almost equivalent for the two channels of Ferrari. The part of authenticity regarding the reliability (Q1 and Q2) shows that people's expectations are met a little more for official channel videos (4.78 vs 4.75 in average), while the viewers believe more in what is said by the influencers (5.08 vs 5.13 in average). The same happens for the integrity (Q3 and Q4): indeed, the brand gives the impression to be more sincere and spontaneous (5.16 vs 4.92 in average), while the influencers care more about people who see the clip (4.75 vs 4.85 in average). These data indicate an equilibrium in the perception of authenticity, placing on the same level the impressions of viewers. The other four statements are related to credibility and show the prevailing of the official channel videos, even if the differences are very small. The only exception is Q8, where it emerges that UGC have more accurate and understandable information (4.67 vs 4.87 in average), while the other sentence related to the expertise (Q7) underlines the higher skills and competences of people who speaks in the BGC clips (5.05 vs 4.83 in average). The aspects regarding the trustworthiness part (Q5 and Q6) highlight a

better design of the official channel tapes (4.99 vs 4.83 in average), meanwhile both the information shared are considered unbiased at the same level (5.03 vs 5.02 in average).

Another reasoning should be done for Tiffany & Co.; as the Figure 12 shows, the BGC average answers values are higher than the UGC ones for every category.

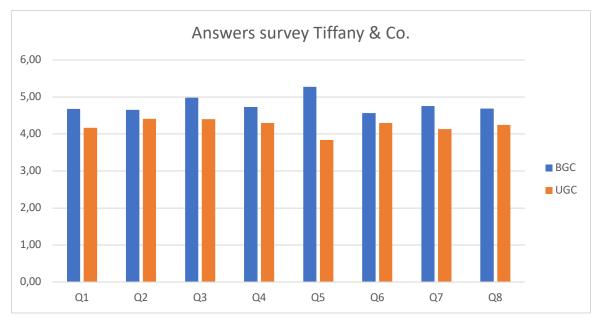


Figure 12 Average answers values for Tiffany

The official channel videos meet better the respondents' expectations (Q1: 4.68 vs 4.17 in average) and the concepts exposed by them are more believable than the ones of the influencers (Q2: 4.65 vs 4.41 in average); these items make the reliability higher for BGC clips.

The integrity part also performs better for official channel videos. The sincerity and the genuineness are better transmitted by BGC (Q3: 4.98 vs 4.40 in average) and it also seems that the brand has a higher care of its customers (Q4: 4.73 vs 4.29 in average).

For these reasons, the authenticity, described by the two precedent variables, is perceived higher in BGC than in UGC.

The most interesting values are in Q5; the design of the videos (sounds, images and influencer's perception) scored 5.27, on average, for BGC, while only 3.84 for UGC. It can be deducted that the clips from the official page are more appreciated and better made and

edited. Q6 ask to evaluate the unbiased information (4.57 vs 4.29 in average) and the results confirm the higher performance of BGC in terms of trustworthiness.

Lastly, Q7 and Q8, expressing expertise level, have higher values in BGC part (4.75 vs 4.14 and 4.68 vs 4.24 in average respectively), indicating that people in BGC videos are perceived more skilled and competent and that the information they share are more accurate and understandable.

Consequently, also the credibility is better perceived by the official page rather than the influencer's channels.

After this analysis it has been tried to find some connection between the 8 statements and the numerical engagement measures. Studying first Ferrari, it was found no correlation between any variables. While considering Tiffany, it was possible to discover three links, in

BGC	Views	Likes	Dislikes	Comments
Q1	-0,05	0,29	0,48*	0,37
Q2	-0,36	-0,11	0,05	-0,04
Q3	-0,32	0,07	0,05	0,15
Q4	-0,15	-0,02	-0,01	-0,03
Q5	-0,21	0,39	0,29	0,44*
Q6	-0,16	0,21	0,06	0,29
Q7	-0,25	0,03	0,14	0,08
Q8	-0,5*	0,03	-0,06	0,1

BGC part, but just one that has a meaning. Indeed, Q5 is related with comments (0,44 with 4,8% significance level). Q5 express how much a video is well made, so the comments increase when sound, editing, images and speakers' impressions have a good quality (Table 23).

Table 23 Tiffany correlation matrix BGC

While the numerical engagement variables show almost no correlation with the 8 statements, the language ones outline more links especially for Ferrari UGC.

Complexity variable is correlated with Q7 and Q8 Ferrari and Tiffany (only UGC), i.e. the complexity increase when the person who expose the concepts in the video is skilled and the information provided are understandable.

The other relationship found are only in Ferrari unofficial channels. Complexity and analytic

UGC	Sentiment	Emotionality	Complexity	Words	Analytic
Q1	0,53**	0,43*	0,63**	0,49*	0,7***
Q2	0,53**	0,46*	0,64***	0,37	0,7***
Q3	0,64***	0,56**	0,48*	0,09	0,62**
Q4	0,52**	0,36	0,51*	0,35	0,6**
Q5	0,4	0,38	0,46*	0,25	0,5*
Q6	0,66***	0,56**	0,77***	0,47*	0,83***
Q7	0,42*	0,31	0,62**	0,5*	0,68***
Q8	0,47*	0,41*	0,58**	0,46*	0,66***

have a connection with all the 8 statements, while number of words results related only with Q1, Q6, Q7 and Q8. Considering the feeling variables sentiment present no correlation with Q5, meanwhile

Table 24 Ferrari correlation matrix UGC

emotionality does not present it also with Q4 and Q7 (Table 24).

4.3.1 Discussion

The analysis of the answers received shows different perceptions in Ferrari and Tiffany: for the first, BGC and UGC are very similar, while for the second they present a different perception, BGC channels result to be better in authenticity and credibility than UGC one. Starting from Ferrari, the majority of UGC clips are made by channels related to the world "Ferrari", made by very competent people. As a consequence, what influencers say is considered more believable and accurate, indicating the high competences of these persons together with the capacity to explain the concepts with the aim to get them understandable to everyone. However, the design of the video probably highlights the presence of an employee dedicated to the video editing, because BGC clips have better images, sounds and speakers' perception. Thanks to the analysis of correlation about Ferrari, it possible to notice in UGC channels that the more authentic and credible is the perception of the videos, more complex and analytic a comment is; while in the BGC one there is no such relation. This is a further confirmation of the previous findings: official channel generates less interaction and the little one created is mostly about the brand. Moreover, it also highlights the fact that who makes the videos are skilled people. Considering sentiment and emotionality in UGC channel, it is clearly visible that when a content is authentic and credible provoke in users a positive feeling that is expressed in the comments.

Moving on to Tiffany, the findings brought to a different evaluation: BGC is perceived much better than UGC, with a peak of one point and a half more in a question. Indeed, differently from Ferrari, the UGC clips are made by people who show the unboxing of products previously ordered or their jewelers' collections, they are probably fan of the company that created a YouTube channel to talk about it, or they do reviews for job so a brand product could be one of this. It is possible to explain in this way the differences in the believability, sincerity and level of expertise, all in favor of the official page. A particular reasoning should be done for the design of the video, that is the statement with the highest difference between the two types of channel; the reasons potentially include the same concept expressed for Ferrari, i.e. Tiffany probably has a video editor that does this work, together with the fact that people register the clip with their phones, leading to a reduction in the quality.

Considering the correlation analysis of Tiffany, the links of complexity with Q7 and Q8 show that people comments more complex when users are more skilled and the information are understandable, showing that the videos are made by capable people.

The study included the vision of all the 93 videos, used in the survey, of these two brands; thanks to this, it is possible to confirm what it is written above: indeed, the UGC videos of Ferrari are made by competent people, that are influencers of Ferrari and aim to describe in details the products with their characteristics; meanwhile, the users of the unofficial channel post video about Tiffany simply because they love its products and show to the world what he/she bought. The correlation previous exposed are a validation of what was seen in the videos for Ferrari, but they are in contrast for Tiffany case.

In general, a content that is perceived more authentic (Lu et al., 2015; Choi et al., 2014) and credible (O'Keefe, 1990; Metzger et al., 2003) should bring to an increment in the popularity; in social media popularity is measured through likes and comments mostly (De Vries et al., 2012). The absence both in Ferrari and Tiffany of the correlation between the previous variables and the authenticity and credibility factors is in contrast with the existing studies. These researches were mainly on Facebook and Twitter, so, it is possible that YouTube popularity factors are others, such as: ratings and views; which are mostly used to compare channel and videos (Chatzopoulou et al., 2010).

After all this dissertation, it is possible to say that a clear difference between BGC and UGC in term of authentic and credible content cannot be done; in fact, in one case they are equally perceived, in the other BGC is better perceived. Two single brands are not enough to derive statics conclusion but can be the example to follow.

5. CONCLUSION

The aim of this study was to perform a deep analysis to highlight the differences between brand-generated content and user-generated content on YouTube.

In order to accomplish the goals, the research required several steps:

- Literature review
- Data exploration
- Data analysis

The first step was to design a structured literature review. In order to do so it was necessary to previously remove the material not relevant for this study.

Eventually, studies among academic journals, whitepapers, websites and books were selected and divided according to seven main topics: social media marketing, YouTube, social media engagement, brand-generated content, user-generated content, authenticity, credibility.

Hence, the first part of literature review focused on exploring the social media world, particularly YouTube one. Here, the principal characteristics and the advantages of using social media were presented.

After this cognitive part, the focus shift to brand and user-generated content. The first one is content made and promoted by a brand, while the second is something made by users that can be paid by brands or done for their personal interest.

Lastly, the two notions of authenticity and credibility were defined. Authenticity, thanks to the findings, was defined by two concepts: reliability and integrity. The first is the brands' willingness and ability to deliver on its promises; the second is the virtue reflected in the brand's intention and in the value it communicates. Credibility, from many years, is defined by five factors, the main ones are: trustworthiness and expertise. The first is the reliability of a person/statement, while the second is the information quality and availability together with its certification.

The data exploration part consisted in the study of the data already available and to the collection of new data through a questionnaire. The survey was thought to analyze people perceptions of authenticity and credibility of YouTube videos.

In order to do so, it was structured in two parts:

1. Videos and some statements to evaluate them;

2. General information about the respondent.

Last part regarding data analysis was centered on the study of the data available and the new ones from the survey. Here the differences between BGC and UGC were highlighted. The first divergences were at video level:

- influencer channel has more subscribers, in average;
- videos last way more in UGC channels;
- BGC clips have more views;
- The interaction level (likes, dislikes, comments) is higher in unofficial channels;

Analyzing the brand level, other differences were found:

- Following the interaction level, also actors and links among actors are more in UGC channels;
- The brand is mentioned more under influencer videos, but the value is more significative for official ones.

The correlation also helped in highlighting some diversity, due to the distinction present between the variables such as views, likes, dislikes and comment. Indeed, the correlation indexes of the numerical engagement measures for BGC are weaker than the UGC ones. Regarding the language engagement measures, the relation between the variables is similar in both channels, for number of words, complexity and analytic. The differences are in the feeling variable, indeed, UGC part present more negative sentiment than BGC ones. Considering the relation between network and SBS with numerical engagement measures, the unofficial channel confirms what was said before about the interaction, having a high correlation between these variables.

Analyzing the data from the questionnaire, a perception of authenticity and credibility about the two channels was outlined. Ferrari present a high perception of the two variables for both channels, while Tiffany & Co. result to be better in BGC than in UGC.

In the end the analysis shows that BGC and UGC have their strength, the first in terms of views and brand mentions, the second in term of interaction. Views for brand channels are driven not only by YouTube but also by other social media where the number of views is also higher and where the visibility for users is not that strong as on YouTube. The findings from the questionnaire did not bring to an objective and statistic evaluation of brand and user

contents due to the low number of brands analyzed. The results can anyway be considered as a starting point for future studies.

5.1 MANAGERIAL IMPLICATIONS

This research provides some useful suggestion for managers of companies thanks to the comparison between Brand-generated and User-generated content. First thing to address is the subscribers, managers have to consider the possibility of partnership with influencers or YouTubers to boost their company, attaching their link official page below the video. Indeed, users want to have everything ready to use and so the immediate possibility for brand to receive a visit and a possible new subscriber. The consumers want more and more to be active part of the product/process development, so the managers can consider also to make them center of their videos on YouTube, stimulating participation and also proposing content closer to UGC one. From this they can achieve higher engagement, increasing especially likes, dislikes and comments metrics that right now are scoring much lower than unofficial channels. In general, the popularity in social media is higher for people than brand so it is not easy to bring these contributions but starting from them, they can, with the time, change the trend.

Such problem is not present with the variable views, so managers, considering the already the big audience they reach trough the social media, have to exploit this advantage to propose the new contents always to achieve better scoring in the numerical engagement measures.

Managers have to focus on comments because they are a clear expression of their fans. They need to pay attention to the rate at which they post marketing messages that increase the number of negative comments, it is better to alternate them with fun post. In fact, firms that mix their marketing campaign with fun post attract fewer negative comments, increasing the positive hype around the brand. Managers need to avoid ignoring or censoring these negative messages, but make them positive, improving the criticized things.

As already said semantic brand score is really important for brand channels, managers must consider them. The first step is to address who mentioned them, once recognized the belonging category of that person, managers need to act accordingly to it. In this way is possible to face the negative and positive ones.

The most important thing is to not consider YouTube as a standalone platform, but managers have to consider a whole ecosystem where all the social media can work together, exploiting the advantages of every single media, giving consumers an experience at 360 degrees.

5.2 LIMITATIONS

Even if this study allowed to find some diversity between BGC and UGC, it is affected by some limitations, mostly related to the collection of the data regarding the perception of authenticity and credibility through the vision of some videos.

To collect the tapes, a time frame of one year is considered. Therefore, there could be difference between the older clips and the younger ones; indeed, the more recent videos may not have reached the highest level of numerical engagement variables (views, likes, dislikes, comments), because they could rise in the future, while the oldest ones should join an almost definitive degree.

On YouTube, since everyone can search whatever he/she wants, the number of the numerical engagement variables can never be definitive: indeed, if a person would like to search an old video, he/she can simply find it, increasing the views and maybe the likes or the comments. This is a difference between YouTube and other social media as Facebook and Twitter, because in the last platforms the new upload are the ones that appear at the beginning, while on YouTube the research could give as output a list of different videos without a link with their publication date.

To create the initial dataset, the companies are selected between the top 100 global brands; this could be a limitation, because the companies that are not listed here could respect the thresholds imposed (activity of the page, uniform and global social media strategy, presence of UGC about the brand on YouTube) and consequently could be considered in the analysis. Moving to the surveys, the first limitation is the further reduction of the brands: the questions are posed only for two of them (Ferrari and Tiffany & Co.), restricting the analysis. People who filled out the questionnaires were also focus on a little part of the total amount of videos; indeed, only 3 out of 25 clips for survey of each type of channel were included, and they were even cut. Indeed, they were too long to ask people to watch such clips, the shortage versions allowed to the video to be seen but with a lot of content removed that could have

changed people perceptions. Probably, the answers might be more accurate if the respondents watched all the whole videos before to complete the surveys.

The sample size could be improved, too. Even if 320 replies are a good number, the bigger are the respondents, the better are the results. However, a different discussion should be done about the age range and the professions of the sample: the majority of participants are students or are just graduated between 20 and 30 years old (about 27 years old on average); a more homogeneous representation of these categories can be useful to measure the perception of the two variables studied into different groups.

Furthermore, the surveys had a high level of subjectivity; so, it can happen that the results change according to the reference sample. As a consequence, the correlations between the eight questions and the numerical engagement measures may be influenced, since they calculate links between a subjective value and an objective one.

Finally, authenticity and credibility are described using the definition found in the already existing literature. However, they were used in different fields than the one of this study. It is possible that changing the definitions of these two variables, also the results could vary.

5.3 FUTURE DEVELOPMENT

Possible future developments are mostly related to the resolution of the limitations previously described.

Firstly, the analysis done was explorative, since the actual state of art did not present any findings about this topic; consequently, this research could be used as a starting point for future and deeper studies.

An update of the dataset previously collected could bring more and detailed information about the already presented brands and new ones that are nowadays in the "100 Best Global Brands". Moreover, in the collection of the data, it could be better to take into consideration just the videos with all the information available.

Concerning the surveys, it could be interesting to modify the number of brands selected. It is possible to start including all the brand already available, amplifying the number of videos that will be evaluated. The more the videos seen and evaluated by people the more the results can be an objective finding.

In addition, the sample could be increased in terms of ages and professions, in order to have a larger view. Furthermore, it may be better to send the same questionnaire indifferently from the gender, avoiding the division and allowing to have opinions from both males and females on the same videos.

Lastly, authenticity and credibility can be defined in different ways, with the aim to change the factors that affect them and therefore find different results.

6. BIBLIOGRAPHY

- Aaker, D. A. (1996). Measuring brand equity across products and markets. *California management reviews*.
- Ahluwalia, R. (2002). How prevalent is the negativity effect in consumer environments? *Journal of consumer research*.
- Ali, H., Hariharan, M., Yaacob, S., & Adom, A. H. (2015). Facial emotion recognition using empirical mode decomposition. *Expert systems with application*.
- Alonso, J. B., Cabrera, J., Medina, M., & Travieso, C. M. (2015). New approach in quantification of emotional intensity from the speech signal: emotional temperature. *Expert systems with applications*.
- Anderegg, W. R., Prall, J. W., Harold, J., & Schneider, S. H. (2010). Expert credibility in climate change.
- Armstrong, C. L., & McAdams, M. J. (2009). Blogs of information: how gender cues ad individual motivation influence perceptions of credibility. *Journal of computer-mediated communication*.
- Ashley, C., & Tuten, T. (2014). Creative Strategies in Social Media Marketing: An Exploratory Study of Branded Social Content and Consumer Engagement. *Psychology & marketing*.
- Asur, S., & Huberman, B. A. (2010). Predicting the future with social media.
- Bagozzi, R. P. (1983). A holistic methodology for modeling consumer response to innovation. *Operations research*.
- Balti, H., & Elmaghraby, A. S. (2014). Emotion analysis from speech using temporal contextual trajectories. *Proceedings of IEEE Symposium on Computers and Communication (ISCC)*.
- Basile, V., & Nissim, M. (2013). Sentiment analysis on Italian tweets. *Proceedings of the*4th workshop on computational approaches to subjectivity, sentiment and social media analysis.
- Berger, B. (2015, May). *People Are Talking About Your Brand Here's How to Engage*. Retrieved from Mention: https://mention.com/blog/engage-with-brand-audience/
- Berlo, D. K., Lemert, J., & Mertz, R. (1969). Dimensions for Evaluating the Acceptability of Message Sources. *Public Opinion Quarterly*.
- Beverland, M. B. (2005). Crafting brand authenticity: the case of luxury wine.
- Bik, H. M., & Goldstein, M. C. (2013). An Introduction to social media for scientists.
- Blackshaw, P., & Nazzaro, M. (2004). Consumer-generated media (CGM) 101: Word-of-mouth in the age of the web-fortified consumer.
- Boyd, D. M., & Hargittai, E. (2010). Facebook privacy settings: who cares?
- Boyd, J. (2019, November 15). *The Most Followed Accounts on Twitter*. Retrieved from Brandwatch: https://www.brandwatch.com/blog/most-twitter-followers/
- Boyd, J. (2019, November 15). *The Most-Subscribed YouTubers and Channels*. Retrieved from Brandwatch: https://www.brandwatch.com/blog/most-subscribed-youtubers-channels/
- Boyle, D. (2003). Authenticity: brands, fakes, spin and the lust for real life.

- Brodersen, A., Scellato, S., & Wattenhofer, M. (2012). YouTube Around the World: Geographic Popularity of Videos.
- Bruce, N. I., Foutz, N. Z., & Kolsarici, C. (2012). Dynamic effectiveness of advertising and word of mouth in sequential distribution of new products. *Journal of marketing research*.
- Bruns, A. (2008). The active audience: trasnforming journalism from gatekeeping to gatewatching.
- Burgess, J. E. (2011). Youtube. Oxford bibliographies online.
- Burgess, J., & Green, J. (2013). YouTube: Online video and participatory culture.
- Burhn, M., Schoenmuller, V., Schafer, D., & Heinrich, D. (2012). Brand Authenticity: Towards a Deeper Understanding of Its Conceptualization and Measurement.
- Cabiddu, F., De Carlo, M., & Piccoli, G. (2014). Social media affordances: Enabling customer engagement. *Annals of tourism research*.
- Capannelli, G. A., & Capannelli, S. (2004). Authenticity: simle strategies for greater meaning and purpose at work and at home.
- Casalò, L. V., Flaviàn, C., & Guinaliu, M. (2008). Promoting consumer's participation in virtual brand communities: a new paradigm in brand strategies. *Journal of marketing communication*.
- Cha, M., Kwak, H., Rodriguez, P., Ahn, Y. Y., & Moon, S. (2009). Analyzing the video popularity characteristics of large-scale user generated content systems.
- Chatzopoulou, G., Sheng, C., & Faloutsos, M. (2010). A first step towards understanding popularity in YouTube. *procedings IEEE INFOCOM*.
- Chavhan, Y., Yelure, B., & Tayade, K. (2015). Speech emotion recognition using RBF kernel of LIBSVM. *In Proceedings of the 2nd International Conference on Electronics and Communication Systems (ICECS)*.
- Chen, J., Xu, H., & Whinston, A. B. (2011). Moderated online communities and quality of user generated content. *Journal of management information systems*.
- Cheng, X., Dale, C., & Liu, J. (2007). Understanding the Characteristics of Internet Short Video Sharing: YouTube as a Case Study.
- Chiou, J. S., & Cheng, C. (2003). Should a company have message boards on its web sites? Journal of interactive marketing.
- Choi, B., & Lee, I. (2017). Trust in open versus closed social media: the relative influence of user-and marketer-generated content in social network services on customer trust. *Telematics and informatics*.
- Choi, H., Ko, E., Kim, E. Y., & Mattila, P. (2014). The Role of Fashion Brand Authenticity in Product Management: A Holistic Marketing Approach.
- Colladon, A. F. (2018). The Semantic Brand Score. Journal of business research.
- Cowie, R., & Cornelius, R. R. (2003). Describing the emotional states that are expressed in speech. *Speech communication*.
- Dawson, V. (2018). Fans, Friends, Advocates, Ambassadors, and Haters: Social Media Communities and the Communicative Constitution of Organizational Identity. *Sage journal*.
- De Vries, L., Gensler, S., & Leeflang, P. S. (2012). Popularity of brand posts on brand fan pages: an investigation of the effects of social media marketing. *Journal of interactive marketing*.

- Dekay, S. (2012). How large companies react to negative Facebook comments. *Corporate Communications: An International Journal*.
- Dellarocas, C. (2003). The digitalization of word-of-mouth: promise and challenges of online feedback mechanisms. *Management science*.
- Desmet, B., & Hoste, V. (2013). Emotion detection in suicide notes.
- Dichter, E. (1966). How word-of-mouth advertising works. *Harvard Business review*.
- Ding, Y., Phang, C. W., Lu, X., Tan, C. H., & Sutanto, J. (2014). The role of marketer- and user-generated content in sustaining the growth of a social media brand community.
- Erdem, T., & Swait, J. (2004). Brand credibility, brand consideration, and choice. *Journal of consumer research*.
- Erdem, T., Swait, J., & Louvriere, J. (2002). The impact of brand credibility on consumer price sensitivity. *International journal of research in marketing*.
- Evert, S. (2005). The statistics of word cooccurences word pairs and collocations.
- Ewing, M. T., & Napoli, J. (2005). Developing and validating a multidimensional. *Journal of Business Research*.
- Eynon, R., & Malmberg, L. E. (2011). A typology of young people's Internet use: Implications for education. *Computers & education*.
- Falls, J. (2010). What is engagement and how do we measure it? Social media explorer.
- Fine, G. A. (2003). Crafting authenticity: the validation of identity in self-taught.
- Fisher-Buttinger, C., & Vallaster, C. (2008). Connective branding: Building brand equity in.
- Flanagin, A. J., & Metzger, M. J. (2000). Perceptions of internet information credibility. Journalism and mass communication quarterly.
- Flanagin, A. J., & Metzger, M. J. (2007). The role of site features user attributes, and information verification behaviors on the perceived credibility of web-based information. *New media & society*.
- Flanagin, A. J., & Metzger, M. J. (2011). From encyclopaedia britannica to wikipedia: generational differences in the perceived credibility of online encyclopedia information.
- Fogg, B. J. (2003b). Prominence-Interpretation theory: explaining how people assess credibility online.
- Foutz, N. Z., & Jank, W. (2010). Prerelease demand forecasting for motion pictures using functional shape analysis of virtual stock markets. *Marketing sci*.
- Freberg, K., Graham, K., McGaughey, K., & Freberg, L. A. (2011). Who are the social media influencers? A study of public perceptions of personality. *Public relations reviews*.
- Frederick, E. L., Lim, C. H., Clavio, G., & Walsh, P. (2012). Why we follow: an examination of para-social interaction and fan motivations for following athlete archetypes on Twitter. *International journal of sport communication*.
- Fuchs, C. (2009). Information and communication technologies and society: a contribution to the critique of the political economy of internet. *European journal of communication*.
- Gass, R., & Seiter, J. (1999). Persuasion, Social Influence and Compliance Gaining.
- Gloor, P. A. (2017). Sociometrics and human relationship: analyzing social networks to manage brands, predict trends, and improve organizational performance. *Emerald publishing limited*.

- Go, A., Bhayani, R., & Huang, L. (2009). Twitter sentiment classification using distant supervision.
- Goh, K. Y., Heng, C. S., & Lin, Z. (2013). Social media brand community and consumer behavior: quantifying the relative impact of user-and marketer-generated content. *Information system research*.
- Goldenberg, J., Libai, B., & Muller, E. (2001). Talk of the network: a complex systems look at the underlying process of word-of-mouth. *Marketing letters*.
- Grayson, K., & Martinec, R. (2004). Consumer perception of iconicity and indexicality and their influence on assessments of authentic market offerings. *journal of consumer research*.
- Gretzel, U., & Yoo, K. (2008). Use and impact of online travel reviews.
- Grewal, L., Lilien, G. L., & Mallapragada, G. (2006). Location, location, location: how network embeddedness affects project success in open source system.

 Management sci.
- Grohs R., R. K. (2016). One pie, many recipes: alternative paths to high brand strenght. Journal of business research.
- Habibi, M. R., Laroche, M., & Richard, M.-O. (2014). The roles of brand community and community engagement in building brand trust on social media. *Computers in Human Behavior*.
- Hanna, R., Rohm, A., & Crittenden, V. L. (2011). We're all connected: the power of the social media ecosystem. *Business horizons*.
- Hargittai, E. (2007). Whose space? Differences among users and non-users of social network sites. *Journal of computer-mediated communication*.
- Harrigan, P., Evers, U., Miles, M., & Daly, T. (2017). Customer engagement with tourism social media brands. *Tourism Management*.
- Harrison, J. (2010). User-generated content and gatekeeping at the bbc hub. *Journalism studies*.
- Heinonen, k. (2011). Consumer activity in social media: Managerial approaches to consumers' social media behavior. *Journal of Consumer Behaviour*.
- Heinonen, K. (2011). Consumer activity in social media: Managerial approaches to consumers' social media behavior . *Journal of consumer behavior*.
- Henning-Thurau, T., Gwinner, K., Walsh, G., & Gremler, D. (2004). Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet? *Journal of interactive marketing*.
- Hofacker, C. F., & Belanche, D. (2016). Eight social media challenges for marketing managers. *Spanish journal of marketing*.
- Hollebeek, L. (2011). Exploring customer brand engagement: definition and themes. Journal of strategic Marketing.
- Hollebeek, L. D., Glynn, M. S., & Brodie, R. J. (2014). Consumer Brand Engagement in Social Media: Conceptualization, Scale Development and Validation. *journal of interactive marketing*.
- Hong, S. (2013). Who benefits from Twitter? Social media and political competition in the U. S. house of representatives. *Govirnement information quarterly*.
- Horton, D., & Wohl, R. R. (1956). Mass communication and para-social interaction: observations on intimacy at a distance. *Psychiatry*.

- Hovland, C. I., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. *The public opinion quarterly*.
- Hovland, C. I., Janis, I. L., & Kelley, J. J. (1953). Communication and Persuasion.
- Hoy, M. G., & Milne, G. (2010). Gender differences in pricavy-related measures for young adult facebook users. *Journal of interactive advertising*.
- Hull, K., & Lewis, N. P. (2014). Why Twitter displaces broadcast sports media: A model. International journal of sport communication.
- Hyung-Park, D., Lee, J., & Han, I. (2007). The effect of on-line consumer reviews on consumer purchase intention: the moderating role of involvement. *International journal of electronic commerce*.
- Johnson, T. J., & Kaye, B. K. (1998). Cruising is believing? comparing internet and traditional sources on media credibility measures. *Journalism of mass communication quarterly*.
- Kaplan, A. M., & Haenlein, M. (2010). User of the world, unite! The challenges and opportunities of social media. *Business Horizons*.
- Karahasanovic, A., Brandtzaeg, P. B., Heim, J., Luders, M., Vermeir, L., Pierson, J., . . . Jans, G. (2009). Co-creation and user-generated content-elderly people's user requirements. *Computers in human behavior*.
- Kates, S. M. (2004). The dynamics of brand legitimacy: an interpretive study in the gay men's community. *Journal of consumer research*.
- Keller, K. L. (1993). Conceptualizing, measuring, and managing customer-based brand equity. *Journal of marketing*.
- Keller, K. L. (2016). Reflections on customer-based brand equity: perspectives, progress, and priorities. *AMS review*.
- Khan, M. L. (2017). Social media engagement: What motivates user participation and consumption on YouTube? *Computers in Human Behavior*.
- Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business horizons*.
- Kim, A. J., & Ko, E. (2010). Impacts of Luxury Fashion Brand's Social Media Marketing on Customer Relationship and Purchase Intention. *Journal of Global Fashion Marketing*.
- Kim, A. J., & Ko, E. (2012). Do social media marketing activities enhance customer equity?

 An empirical study of luxury fashion brand ★. *Journal of Business research*.
- Kim, H., & Jamal, T. (2007). Touristic quest for existential authenticity. *Annals of tourism research*.
- Kirmani, A. (1990). The effect of perceived advertising costs on brand perceptions. *Journal of consumer research*.
- Kumar, A., Bezawada, R., Rishika, R., Janakiraman, R., & Kannan, P. K. (2016). From social to sale: the effects of firm-generated content in social media on customer behavior. *Journal of marketing*.
- Kumar, V., & Mirchandani, R. (2012). Increasing the ROI of Social Media Marketing. *MIT sloan management review*.
- Lambrecht, A., Goldfarb, A., Bonatti, A., Ghoes, A., Goldstein, D. G., & Lewis, R. (2014). How do firms make money selling digital goods online? *Marketing letters*.

- Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Contents lists available at ScienceDirect Journal of Experimental Social Psychology*.
- Li, W., & Xu, H. (2014). Text-based emotion classification using emotion cause extraction. Expert systems with application.
- Li, Y., Oh, L. B., & Wang, K. (2017). Why users share marketer-generated content on social broadcasting web sites: a cognitive-affective involvement perspective. *Journal of organizational computing and electronic commerce*.
- Lim, J. S., Hwang, Y., Kim, S., & Biocca, F. A. (2015). How social media engagement leads to sports channel loyalty: Mediating roles of social presence and channel commitment. *Computers in Human Behavior*.
- Litvin, S. W., Goldsmith, R. E., & Pan, B. (2007). Electronic word-of-mouth in hospitality and tourism management. *Tourism management*.
- Liu, Y., & Shrum, L. J. (2002). What is interactivity and is it always such a good thing? Implications of definition, person and situation for the influence of interactivity on advertising effectiveness. *Journal of advertising*.
- Lu, A. C., Gursoy, D., & Lu, Y. C. (2015). Authenticity perceptions, brand equity and brand choice intention: The case of ethnic restaurants. *International journal of hospitality management*.
- MacInnis, D. J., Rao, A. G., & Weiss, A. M. (2002). Assessing when increased media weight of real-world advertisements helps sales. *Journal of marketing research*.
- Mahajan, V., Muller, E., & Bass, F. M. (1995). Diffusion of new products: empirical generalizations and managerial uses. *Marketing sci*.
- Mangold, W. G., & Faulds, D. J. (2009). Social media: The new hybrid element of the promotion mix. *Business horizons*.
- McAlexander, J. H., Schouten, J. W., & Koenig, H. F. (2002). Building brand community. *Journal of marketing*.
- McCarthy, P. M., & Jarvis, S. (2010). A validation study of sophisticated approaches to lexical diversity assessment. *Behavior research methods*.
- McCroskey, J. C. (1966). Scales for the Measurement of Ethos.
- Mehrabian, A., & Russell, J. A. (1974). An approach to environmental psychology.
- Metzger, M. J., Flanagin, A. J., Eyal, K., Lemus, D. R., & McCann, R. (2003). Bringing the concept of credibility into the 21st century: integrating perspectives on sourrce, message and media credibility in the contemporary media environment. *Communication yearbook*.
- Morhart, F., Malar, L., Guevremont, A., Girardin, F., & Grohmann, B. (2015). Brand authenticity: An integrative framework and measurement scale. *Journal of consumer psychology*.
- Moulard, J. G., Raggio, R. D., & Garretson Folse, J. A. (2016). Brand Authenticity: Testing the Antecedents and Outcomes of Brand Management's Passion for its Products. *Psychology & marketing*.
- Muniz, M., & O'Guinn, C. (2001). Brand community. *Journal of consumers research*.
- Mustak, M., Jaakkola, E., & Halinen, A. (2013). Customer participation and value creation: a systematic review and research implications. *Managing service quality*.

- Nadkarni, A., & Hofmann, S. G. (2012). Why people use facebook? *Personality and individual differences*.
- Napoli, J., Dickinson, S. J., Beverland, M. B., & Farrelly, F. (2014). Measuring consumer-based brand authenticity. *Journal of business research*.
- Nardi, B. A., Schiano, D. J., Gumbrecht, M., & Swartz, L. (2004). Why we blog. communications of the ACM.
- Nassirtoussi, A. K., Aghabozorgi, S., Wah, T. Y., & Ngo, D. C. (2014). Text mining for market prediction: a systematic review. *Expert systems with applications*.
- Nelson, P. (1974). Advertising as information. Journal of political economy.
- O'Keefe, D. J. (1990). Persuasion: theory and research.
- Olsen, K. (2002). Authenticity as a concept in tourism research: the social organization of the experience of authenticity.
- Papacharissi, Z. (2003). The blogger revolution: a uses and gratifications study. Association of internet researchers annual conference.
- Pearce, P. L., & Moscardo, G. M. (1985). The relationship between travellers' career levels and the concept of authenticity. *Australian journal of psychology*.
- Pennebaker, J., Chung, C., Frazee, J., Lavergne, G., & Beaver, D. (2014). When Small Words Foretell Academic Success: The Case of College Admissions Essays. *PLOS ONE*.
- Perloff, R. M. (1993). The Dynamics of Persuasion.
- Peterson, R. A. (1997). Creating country music, fabricating authenticity.
- Rao, A. R., Qu, L., & Ruekert, R. W. (1999). Signaling unobservable product quality thorugh a brand ally. *Journal of marketing research*.
- Reisenger, Y., & Steiner, C. J. (2006). Reconceptualizing object authenticity. *Annals of tourism research*.
- Rieh, S. Y., & Danielson, D. R. (2007). Credibility: a multidisciplinary framework. *Annual review of information science and technology*.
- Robins, D., & Holmes, J. (2008). Aesthetics and credibility in web site design. *Information processing and management*.
- Rubin, A. M., Perse, E. E., & Powell, R. A. (1985). Loneliness, para-social interaction and local television news viewing. *Human communication research*.
- Sajid, S. (2016). Social Media and Its Role in Marketing. *Business and Economics Journal*. Salcito, K. (2009). Online journalism ethics: Gatekeeping.
- Sarason, S. B. (1974). The psychological sense of community: prospects for the community psychology.
- Scheer, K. S., & Stern, L. W. (1992). The effect of influence type and performance outcomes on attitude toward the influencer. *Journal of marketing research*.
- Schmitt, B. (2012). The consumer psychology of brands. *Journal of consumer Psychology*.
- Sen, S., & Lerman, D. (2007). Why are you telling me this? An examination into negative consumer reviews on the web. *Journal of interactive marketing*.
- Senecal, S., & Nantel, J. (2004). The influence of online product recommendations on consumers' online choices. *Journal of retailing*.
- Shifman, L. (2011). An anatomy of a YouTube meme. new media & society.
- Shih, C. (2009). The facebook era: tapping online social networks to build better products, reach new audiences and sell more stuff.

- Smith, A., & Anderson, M. (2018, March 1). *Social media use in 2018*. Retrieved from pewresearch: https://www.pewresearch.org/internet/2018/03/01/social-media-use-in-2018/
- Smith, B. G., & Gallicano, T. D. (2015). Terms of engagement: Analyzing public engagement with organizations through social media. *Computers in Human Behavior*.
- Smith, K. (2019, June 13). 126 amazing social media statistics and facts. Retrieved from Brandwatch: https://www.brandwatch.com/blog/amazing-social-media-statistics-and-facts/
- Smith, K. (2019, July 15). 52 Fascinating and Incredible YouTube Statistics. Retrieved from Brandwatch: https://www.brandwatch.com/blog/youtube-stats/
- Stafford, T., Stafford, M., & Schkade, L. (2004). Determining Uses and Gratifications for the Internet. A journal of the decision sciences institutes.
- Stelzner, M. (2013). The 2013 social media marketing industry report.
- Stepchenkova, S., & Zhan, F. (2013). Visual destination images of Perù: comparative content analysis of DMO and user-generated photography. *Tourism management*.
- Stielglitz, S., & Dang-Xuan, L. (2014). Emotions and Information Diffusion in Social Media—Sentiment of Microblogs and Sharing Behavior. *journal of management nformation systems*.
- Szmigin, I., & Reppel, A. E. (2001). Internet community bonding: the case of mac-news. *European journal of marketing*.
- Tafesse, W. (2016). An experiential model of consumer engagement in social media. Journal of product and brand management.
- Taylor, C. R. (2009). Editorial: the six principles of digital advertising. *International journal of advertising*.
- Thurman, N. (2008). Forums for citizens journalists? Adoption of user generated content initiatives by online news media. *New media & society*.
- Tong, S. T., Van Der Heide, B., Langwell, L., & Walther, J. B. (2008). Too much of a good thing? The relationship between number of friends and interpersonal impressions on facebook. *Journal of computer-mediated communication*.
- Turner, J. R. (1993). Interpersonal and psychological predictors of para-social interaction with different television performers. *Communicaton quarterly*.
- Vermeulen, I. E., & Seegers, D. (2009). Tried and tested: the impact of online hotel reviews on consumer consideration. *Tourism management*.
- Villanueva, J., Yoo, S., & Hanssens, D. M. (2008). The impact of marketing-induced vs word-of-mouth customer acquisition on customer equity. *Journal of marketing research*.
- Wang, N. (1999). Rethinking authenticity in tourism experience. *Annals of tourism research*.
- Whitehead, J. L. (1968). Factors of Source Credibility.
- Wikipedia. (2019, November 8). *List of most-followed Facebook pages*. Retrieved from Wikipedia: https://en.wikipedia.org/wiki/List_of_most-followed_Facebook_pages
- Zhou, R., Khemmarat, S., & Gao, L. (2010). The Impact of YouTube Recommendation System on Video Views. *human factors*.

Zhu, F., & Zhang, X. (2006). The influence of online consumer reviews on the demand for experience goods: the case of video games. *Proceedings of twenty-seven international conference on information systems (ICIS)*.

7. APPENDIX

7.1 SURVEY

SECTION	QUESTION	ANSWER
Brand Familiarity	How familiar/knowledgeable are you with the brand?	Range from 1 to 7
	Videos meet my expectations about the brand I believe in what I have seen and heard in the videos	Range from 1 to 7 Range from 1 to 7
	The brand gives the impression of being sincere	Range from 1 to 7
	The brand cares about the customers	Range from 1 to 7
BGC videos (3 times)	The design of the videos in terms of sound, editing, images and speakers' impression is well done	Range from 1 to 7
	The information shared are fair and do not lead us to believe something that is untrue or irrational	Range from 1 to 7
	The concepts are exposed by someone with the right skills (the competences to talk about the topic)	Range from 1 to 7
	There are quality, accurate and understandable information	Range from 1 to 7

	I feel confident with the	Range from 1 to 7	
	speakers		
	I believe in what I have seen and heard in the videos	Range from 1 to 7	
	Videos seem genuine and		
	spontaneous	Range from 1 to 7	
	Speakers care about the	Range from 1 to 7	
	viewers	rungo rom r vo v	
	The design of the videos in		
	terms of sound, editing,	Range from 1 to 7	
UGC videos	images and speakers'	3	
(3 times)	impression is well done		
	The information shared are		
	fair and do not lead us to	Range from 1 to 7	
	believe something that is	3	
	untrue or irrational		
	The concepts are exposed		
	by someone with the right	Range from 1 to 7	
	skills (the competences to	5	
	talk about the topic)		
	There are quality, accurate		
	and understandable	Range from 1 to 7	
	information		
	How old are you?	Open range	
	What is your gender?	Two options: female or male	
Personal		Nine options: legal occupations;	
information	Which of the following best	teacher; architecture and	
	describes your current	engineering occupations; business	
	occupation?	and financial operations	
		occupations; computer and	
		mathematical occupations;	

management occupations; student;
sales and related occupations; other
(please specify).

7.2 CORRELATION MATRIXES

7.2.1 Numerical engagement p-value matrixes

BGC	Views	Likes	Dislikes	Comments
Views		0,0079	0	0,0705
Likes	0,0079		0	0
Dislikes	0	0		0
Comments	0,0705	0	0	

UGC	Views	Likes	Dislikes	Comments
Views		0	0	0
Likes	0		0	0
Dislikes	0	0		0
Comments	0	0	0	

7.2.2 Language engagement p-value matrixes

BGC	Words	Complexity	Analytic	Sentiment	Emotionality
Words		0,8512	0,0007	0,2955	0,2705
Complexity	0,8512		0	0,3184	0,0012
Analytic	0,0007	0		0,0379	0,212
Sentiment	0,2955	0,3184	0,0379		0,0005
Emotionality	0,2705	0,0012	0,212	0,0005	

UGC	Words	Complexity	Analytic	Sentiment	Emotionality
Words		0,4617	0,0288	0,0376	0,2522
Complexity	0,4617		0	0	0,8649
Analytic	0,0288	0		0	0,0993
Sentiment	0,0376	0	0		0
Emotionality	0,2522	0,8649	0,0993	0	

7.2.3 Numerical vs. language colleration and p-value matrixes

BGC	Views	UG
Words	-0,03	Wo
Complexity	-0,02	Cor
Analytic	0,01	Ana
Sentiment	-0,03	Ser
Emotionality	-0,06	Em

UGC	Views
Words	-0,15*
Complexity	0,08
Analytic	0,08
Sentiment	-0,29***
Emotionality	-0,17*

BGC	Views
Words	0,728
Complexity	0,8041
Analytic	0,8965
Sentiment	0,7183
Emotionality	0,4514

UGC	Views
Words	0,0297
Complexity	0,2638
Analytic	0,2338
Sentiment	0
Emotionality	0.015

Correlation P-value

BGC	Likes
Words	0,1
Complexity	0,06
Analytic	0,04
Sentiment	-0,18*
Emotionality	-0,09

UGC	Likes
Words	-0,16*
Complexity	0,14*
Analytic	0,15*
Sentiment	-0,31***
Emotionality	-0,29***

BGC	Likes
Words	0,1868
Complexity	0,424
Analytic	0,628
Sentiment	0,0138
Emotionality	0,2394

UGC	Likes
Words	0,0191
Complexity	0,0414
Analytic	0,0314
Sentiment	0
Emotionality	0

Correlation

P-value

BGC	Dislikes
Words	0,03
Complexity	0,02
Analytic	-0,01
Sentiment	-0,11
Emotionality	-0,03

UGC	Dislikes
Words	-0,1
Complexity	0,02
Analytic	0
Sentiment	-0,2**
Emotionality	-0,06

BGC	Dislikes
Words	0,6684
Complexity	0,7357
Analytic	0,8928
Sentiment	0,1318
Emotionality	0,646

UGC	Dislikes
Words	0,1616
Complexity	0,7887
Analytic	0,9996
Sentiment	0,0048
Emotionality	0,3751

Correlation

P-value

BGC	Comments	
Words	0,11	
Complexity	0,06	
Analytic	0,05	
Sentiment	-0,18*	
Emotionality	-0,08	

UGC	Comments
Words	-0,02
Complexity	0,09
Analytic	0,11
Sentiment	-0,33***
Emotionality	-0,28***

3GC	Comments
Nords	0,1308
Complexity	0,3846
Analytic	0,5019
Sentiment	0,0118
motionality	0,2607

UGC	Comments
Words	0,7579
Complexity	0,2203
Analytic	0,1317
Sentiment	0
Emotionality	0

Correlation

P-value

7.2.4 Network & SBS p-value matrixes

BGC	Actors	Links among Actors	SBS
Actors		0	0,0033
Links among Actors	0		0,009
SBS	0,0033	0,009	

UGC	Actors	Links among Actors	SBS
Actors		0	0,0848
Links among Actors	0		0,0316
SBS	0,0848	0,0316	

7.2.5 Network vs. Numerical correlation and p-value matrixes

BGC	Views
Actors	0,19
Links among Actors	0,19
SBS	-0,09

UGC	Views
Actors	0,88**
Links among Actors	0,86**
SBS	0,82**

BGC	Views
Actors	0,6274
Links among Actors	0,6257
SBS	0,8162

UGC	Views
Actors	0,002
Links among Actors	0,003
SBS	0,0068

Correlation

P-value

BGC	Likes
Actors	1***
Links among Actors	0,99***
SBS	0,86**

UGC	Likes
Actors	0,96***
Links among Actors	0,94***
SBS	0,68*

BGC	Likes
Actors	0
Links among Actors	0
SBS	0,0031

UGC	Likes
Actors	0
Links among Actors	0,0001
SBS	0.0437

Correlation

P-value

BGC	Dislikes
Actors	0,72*
Links among Actors	0,72*
SBS	0.53

UGC	Dislikes
Actors	0,71*
Links among Actors	0,68*
SBS	0,79*

BGC	Dislikes
Actors	0,0288
Links among Actors	0,0277
SBS	0,1411

UGC	Dislikes
Actors	0,0328
Links among Actors	0,0447
SBS	0,0106

Correlation

P-va	lue

BGC	Comments
Actors	1***
Links among Actors	0,99***
SBS	0,85**

UGC	Comments
Actors	1***
Links among Actors	0,96***
SBS	0,61

BGC	Comments
Actors	0
Links among Actors	0
SBS	0,0035

	UGC	Comments
)	Actors	0
)	Links among Actors	0
;	SBS	0,0787

Correlation

P-value

7.2.6 Network vs. Language correlation and p-value matrixes

BGC	Words
Actors	0,41
Links among Actors	0,4
SBS	0,6

Words
0,24
0,28
0,17

BGC	Words
Actors	0,2785
Links among Actors	0,2915
SBS	0,0846

UGC	Words
Actors	0,5337
Links among Actors	0,4709
SBS	0,6544

Correlation

P-value

BGC	Complexity
Actors	0,33
Links among Actors	0,27
SBS	0,33

UGC	Complexity
Actors	0,66
Links among Actors	0,74*
SBS	0,9***

BGC	Complexity
Actors	0,3842
Links among Actors	0,49
SBS	0,3898

]	UGC	Complexity
2	Actors	0,0553
1	Links among Actors	0,024
3	SBS	0,0009

Correlation

P-value

BGC	Analytic
Actors	0,26
Links among Actors	0,27
SBS	0,26

UGC	Analytic
Actors	0,71*
Links among Actors	0,85**
SBS	0,88**

BGC	Analytic
Actors	0,5028
Links among Actors	0,4825
SBS	0,4999

UGC	Analytic
Actors	0,0327
Links among Actors	0,004
SBS	0,0019

Correlation

P-value

BGC	Sentiment	UGC	Sentiment
Actors	-0,9**	Actors	-0,73*
Links among Actors	-0,9**	Links among Actors	-0,85**
SBS	-0,87**	SBS	-0,71*

BGC	Sentiment
Actors	0,0011
Links among Actors	0,001
SBS	0,0023

UGC	Sentiment
Actors	0,0263
Links among Actors	0,0036
SBS	0,0315

Correlation

P-value

BGC	Emotionality
Actors	-0,57
Links among Actors	-0,5
SBS	-0,62

UGC	Emotionality
Actors	-0,64
Links among Actors	-0,76*
SBS	-0,42

-0,71*

BGC	Emotionality
Actors	0,1084
Links among Actors	0,1735
SBS	0.0757

UGC	Emotionality
Actors	0,0613
Links among Actors	0,0177
SBS	0,2542

Correlation

 $P ext{-}value$

7.2.7 Ferrari correlation and p-value matrixes with numerical engagement measures

BGC	Views	Likes	Dislikes	Comments
Q1	0,14	-0,08	0,08	-0,17
Q2	-0,09	-0,12	-0,17	-0,07
Q3	-0,07	-0,14	-0,1	-0,13
Q4	-0,03	-0,12	-0,13	-0,16
Q5	0,31	0,07	0,25	-0,03
Q6	0,04	0,03	-0,04	0,03
Q7	-0,11	-0,03	-0,2	0,06
Q8	-0,09	-0,11	-0,13	-0,06

UGC	Views	Likes	Dislikes	Comments
Q1	-0,06	-0,25	0,16	-0,12
Q2	0,21	-0,04	-0,04 0,36	
Q3	0,06	-0,09	0,16	0,02
Q4	-0,08	-0,22	0,13	-0,08
Q5	-0,04	-0,2	0,14	-0,15
Q6	-0,1	-0,31	0,12	-0,18
Q7	-0,13	-0,33	0,08	-0,26
Q8	-0,11	-0,35	0,13	-0,24

Correlation

BGC	Views	Likes	Dislikes	Comments
Q1	0,512	0,7128	0,6986	0,4277
Q2	0,684	0,5726	0,4374	0,7489
Q3	0,7289	0,5166	0,6478	0,5324
Q4	0,8921	0,5676	0,5499	0,4658
Q5	0,145	0,7385	0,2346	0,8889
Q6	0,8644	0,8976	0,8482	0,9
Q7	0,5951	0,8915	0,3514	0,7708
Q8	0,6865	0,5962	0,5703	0,7741

UGC	Views	Likes	Dislikes	Comments
Q1	0,7894	0,2451	0,4565	0,5774
Q2	0,327	0,8651	0,0841	0,9892
Q3	0,7948	0,6896	0,451	0,9226
Q4	0,7245	0,2924	0,5456	0,7182
Q5	0,8543	0,3473	0,5105	0,4881
Q6	0,6439	0,1467	0,5906	0,3878
Q7	0,5466	0,1141	0,7118	0,2264
Q8	0,6028	0,093	0,5413	0,2683

P-value

7.2.8 Tiffany correlation and p-value matrixes with numerical engagement measures

BGC	Views	Likes	Dislikes	Comments
Q1	-0,05	0,29	0,48*	0,37
Q2	-0,36	-0,11	0,05	-0,04
Q3	-0,32	0,07	0,05	0,15
Q4	-0,15	-0,02	-0,01	-0,03
Q5	-0,21	0,39	0,29	0,44*
Q6	-0,16	0,21	0,06	0,29
Q7	-0,25	0,03	0,14	0,08
Q8	-0,5*	0,03	-0,06	0,1

UGC	Views	Likes	Dislikes	Comments
Q1	0,34	0,25	0,34	0
Q2	0,33	0,19	0,3	-0,05
Q3	0,34	0,27	0,35	0,003
Q4	0,25	0,22	0,27	0,08
Q5	0,27	0,19	0,27	-0,03
Q6	0,16	0,03	0,13	-0,14
Q7	0,12	-0,06	0,07	-0,28
Q8	0,24	0,11	0,21	-0,08

Correlation

BGC	Views	Likes	Dislikes	Comments
Q1	0,8407	0,2042	0,0294	0,0979
Q2	0,1102	0,6246	0,8329	0,8482
Q3	0,1524	0,759	0,8307	0,5301
Q4	0,5305	0,9207	0,9787	0,9022
Q5	0,3677	0,082	0,1971	0,0486
Q6	0,4785	0,3638	0,7831	0,2033
Q7	0,2742	0,9004	0,5448	0,7186
Q8	0,0205	0,8863	0,8089	0,6706

UGC	Views	Likes	Dislikes	Comments
Q1	0,0995	0,2406	0,9982	0,1027
Q2	0,115	0,3754	0,8047	0,1498
Q3	0,1034	0,1964	0,885	0,0892
Q4	0,2368	0,3078	0,7024	0,2079
Q5	0,1977	0,3624	0,8771	0,2006
Q6	0,4592	0,8981	0,5095	0,5488
Q7	0,5743	0,7781	0,1777	0,7282
Q8	0,2534	0,5973	0,7052	0,3349

P-value

7.2.9 Ferrari correlation and p-value matrixes with language engagement measures

BGC	Sentiment	Emotionality	Complexity	Words	Analytic
Q1	-0,26	-0,06	0,05	-0,25	-0,16
Q2	-0,05	0,18	0,3	-0,16	-0,2
Q3	0,03	0,29	0,35	-0,33	-0,09
Q4	0,07	-0,09	0,13	0,06	-0,08
Q5	-0,26	-0,24	-0,06	-0,02	-0,14
Q6	0,23	0,33	0,38	-0,15	-0,04
Q7	0,24	0,39	0,6**	-0,08	-0,06
Q8	-0,02	0,18	0,41*	-0,06	-0,09

UGC	Sentiment	Emotionality	Complexity	Words	Analytic	
Q1	0,53**	0,43*	0,63**	0,49*	0,7***	
Q2	0,53**	0,46*	0,64***	0,37	0,7***	
Q3	0,64***	0,56**	0,48*	0,09	0,62**	
Q4	0,52**	0,36	0,51*	0,35	0,6**	
Q5	0,4	0,38	0,46*	0,25	0,5*	
Q6	0,66***	0,56**	0,77***	0,47*	0,83***	
Q7	0,42*	0,31	0,62**	0,5*	0,68***	
Q8	0,47*	0,41*	0,58**	0,46*	0,66***	

Correlation

BGC	Sentiment	Emotionality	Complexity	Words	Analytic	UGC	Sentiment	Emotionality	Complexity	Words	Analytic
Q1	0,2205	0,767	0,8317	0,2298	0,4475	Q1	0,0074	0,0361	0,0011	0,0146	0,0001
Q2	0,8048	0,4016	0,1564	0,4435	0,4605	Q2	0,0078	0,0226	0,0007	0,0743	0,0001
Q3	0,8826	0,1655	0,0943	0,1128	0,6603	Q3	0,0008	0,0046	0,0184	0,6698	0,0012
Q4	0,7313	0,6922	0,5403	0,7688	0,6953	Q4	0,0094	0,0828	0,0104	0,096	0,002
Q5	0,2173	0,2682	0,7918	0,9367	0,5268	Q5	0,0511	0,069	0,0241	0,2301	0,0139
Q6	0,2803	0,1165	0,0665	0,4957	0,8432	Q6	0,0004	0,0044	0	0,0208	0
Q7	0,2611	0,0586	0,0018	0,704	0,7894	Q7	0,0392	0,1395	0,0011	0,0132	0,0002
Q8	0,9165	0,396	0,0494	0,7852	0,6867	Q8	0,0291	0,0466	0,0028	0,0242	0,0005

P-value

7.2.10 Tiffany correlation and p-value matrixes with language engagement measures

BGC	Sentiment	Emotionality	Complexity	Words	Analytic	UG
Q1	-0,13	0,06	0,2	-0,1	0,05	Q1
Q2	-0,09	0,45*	-0,01	-0,34	-0,2	Q2
Q3	-0,12	0,23	-0,02	-0,09	-0,2	Q3
Q4	-0,21	0,26	0	-0,11	-0,17	Q4
Q5	-0,38	0,06	0,16	-0,09	-0,12	Q5
Q6	0	0,31	0,03	-0,29	-0,02	Q6
Q7	0,01	0,53*	0,2	-0,43	-0,19	Q7
Q8	-0,12	0,32	0,16	-0,27	-0,1	Q8

UGC	Sentiment	Emotionality	Complexity	Words	Analytic
Q1	-0,26	-0,06	0,05	-0,25	-0,16
Q2	-0,05	0,18	0,3	-0,16	-0,2
Q3	0,03	0,29	0,35	-0,33	-0,09
Q4	0,07	-0,09	0,13	0,06	-0,08
Q5	-0,26	-0,24	-0,06	-0,02	-0,14
Q6	0,23	0,33	0,38	-0,15	-0,04
Q7	0,24	0,39	0,6**	-0,08	-0,06
Q8	-0,02	0,18	0,41*	-0,06	-0,09

Correlation

	T				
BGC	Sentiment	Emotionality	Complexity	Words	Analytic
Q1	0,5756	0,7991	0,3842	0,6718	0,8286
Q2	0,6825	0,0428	0,9726	0,1317	0,3753
Q3	0,5986	0,3101	0,9419	0,6876	0,3741
Q4	0,3685	0,253	0,9998	0,6294	0,4553
Q5	0,0856	0,7846	0,4757	0,712	0,6138
Q6	0,9834	0,1645	0,8824	0,1951	0,9155
Q7	0,9811	0,013	0,3822	0,0516	0,4166
Q8	0,5976	0,1619	0,482	0,2428	0,6583

UGC	Sentiment	Emotionality	Complexity	Words	Analytic
Q1	0,2205	0,767	0,8317	0,2298	0,4475
Q2	0,8048	0,4016	0,1564	0,4435	0,3605
Q3	0,8826	0,1655	0,0943	0,1128	0,6603
Q4	0,7316	0,6922	0,5403	0,7688	0,6953
Q5	0,2173	0,2682	0,7918	0,9367	0,5268
Q6	0,2803	0,1165	0,0665	0,4957	0,8432
Q7	0,2611	0,0586	0,0018	0,704	0,7894
Q8	0,9165	0,396	0,0494	0,7852	0,6867

P-value

7.3 T-TEST

Views	Variabile 1	Variabile 2
Media	751398,195	261734,008
Varianza	7,025E+12	5,4608E+11
Osservazioni	266	266
Correlazione di Pearson	-0,063366	
Differenza ipotizzata per le medie	0	
gdl	265	
Stat t	2,85597356	
P(T<=t) una coda	0,00231513	
t critico una coda	1,65062398	
P(T<=t) due code	0,00463027	
t critico due code	1,96895628	

Likes	Variabile 1	Variabile 2
Media	1134,44361	5450,62406
Varianza	21046764,8	296517012
Osservazioni	266	266
Correlazione di Pearson	0,08352232	
Differenza ipotizzata per le medie	0	
gdl	265	
Stat t	-4,0349787	
P(T<=t) una coda	3,5748E-05	
t critico una coda	1,65062398	
P(T<=t) due code	7,1495E-05	
t critico due code	1,96895628	

Dislikes	Variabile 1	Variabile 2
Media	50,1382114	381,939024
Varianza	23409,5318	3004418,34
Osservazioni	246	246
Correlazione di Pearson	0,01329618	
Differenza ipotizzata per le medie	0	
gdl	245	
Stat t	-2,9942342	
P(T<=t) una coda	0,00151659	
t critico una coda	1,65109682	
P(T<=t) due code	0,00303319	
t critico due code	1,96969392	

Comments	Variabile 1	Variabile 2
Media	63,7638889	439,486111
Varianza	38188,9626	794643,73
Osservazioni	216	216
Correlazione di Pearson	0,16188655	
Differenza ipotizzata per le medie	0	
gdl	215	
Stat t	-6,2667513	
P(T<=t) una coda	9,9302E-10	
t critico una coda	1,65197175	
P(T<=t) due code	1,986E-09	
t critico due code	1,97105912	

Sentiment	Variabile 1	Variabile 2
Media	0,53694906	0,52404857
Varianza	0,0242992	0,01612711
Osservazioni	175	175
Correlazione di Pearson	0,09515535	
Differenza ipotizzata per le medie	0	
gdl	174	
Stat t	0,89132333	
P(T<=t) una coda	0,1869933	
t critico una coda	1,65365802	
P(T<=t) due code	0,3739866	
t critico due code	1,97369144	

Emotionality	Variabile 1	Variabile 2
Media	0,28667656	0,28353227
Varianza	0,00467172	0,00094372
Osservazioni	175	175
Correlazione di Pearson	-0,0019984	
Differenza i potizzata per le medie	0	
gdl	174	
Stat t	0,55465875	
P(T<=t) una coda	0,28992018	
t critico una coda	1,65365802	
P(T<=t) due code	0,57984035	
t critico due code	1,97369144	

Complexity	Variabile 1	Variabile 2
Media	5,84737714	5,79119105
Varianza	0,54355044	0,27181234
Osservazioni	175	175
Correlazione di Pearson	0,13617152	
Differenza ipotizzata per le medie	0	
gdl	174	
Stat t	0,881679	
P(T<=t) una coda	0,18958337	
t critico una coda	1,65365802	
P(T<=t) due code	0,37916673	
t critico due code	1,97369144	

Number of words	Variabile 1	Variabile 2
Media	7,26314553	10,4962174
Varianza	21,7780668	13,9171885
Osservazioni	176	176
Correlazione di Pearson	0,07798203	
Differenza i potizzata per le medie	0	
gdl	175	
Stat t	-7,4687227	
P(T<=t) una coda	1,8426E-12	
t critico una coda	1,65360744	
P(T<=t) due code	3,6852E-12	
t critico due code	1,97361246	

Analytic	Variabile 1	Variabile 2
Media	63,8647802	59,128022
Varianza	623,913456	222,973425
Osservazioni	182	182
Correlazione di Pearson	0,05942292	
Differenza i potizzata per le medie	0	
gdl	181	
Stat t	2,25568427	
P(T<=t) una coda	0,01264369	
t critico una coda	1,65331576	
P(T<=t) due code	0,02528738	
t critico due code	1,97315704	

Actors	Variabile 1	Variabile 2
Media	1140,66667	8641,11111
Varianza	2612115,75	77994901,1
Osservazioni	9	9
Correlazione di Pearson	0,91557388	
Differenza i potizzata per le medie	0	
gdl	8	
Stat t	-3,04879885	
P(T<=t) una coda	0,00792497	
t critico una coda	1,85954804	
P(T<=t) due code	0,01584993	
t critico due code	2,30600414	

Links among actors	Variabile 1	Variabile 2
Media	477,888889	3264,22222
Varianza	701382,861	11442593,7
Osservazioni	9	9
Correlazione di Pearson	0,85948092	
Differenza i potizzata per le medie	0	
gdl	8	
Stat t	-3,09927923	
P(T<=t) una coda	0,00734089	
t critico una coda	1,85954804	
P(T<=t) due code	0,01468178	
t critico due code	2,30600414	

SBS	Variabile 1	Variabile 2
Media	74,5644534	87,0244576
Varianza	2765,59752	2842,05368
Osservazioni	9	9
Correlazione di Pearson	0,74561625	
Differenza ipotizzata per le medie	0	
gdl	8	
Stat t	-0,98956678	
P(T<=t) una coda	0,17568383	
t critico una coda	1,85954804	
P(T<=t) due code	0,35136766	
t critico due code	2,30600414	