Topic modelling of festival-visitors' opinions: A case study of Sziget Festival

Mátyás Hinek [1]* and Árpád Papp-Váry [2]

[1] Budapest Business University, Hungary. Email: hinek.matyas@uni-bge.hu
[2] Budapest Business University, Hungary. Email: papp-vary.arpad@uni-bge.hu

*Corresponding author

Abstract
The Sziget Festival is one of Europe's biggest music events, held every August on the Óbuda Island in Budapest. The elements of festival experiences were analysed through social media opinions of Sziget Festival attendees published between 1 January 2014 and 31 July 2022, using structural topic model (STM) computer algorithm and latent Dirichlet allocation (LDA). Our findings support existing literature, showing that festival experiences are cognitive-affective in nature with key elements of socialisation, environment, and festival value, some of which are inter-connected. Although this kind of research is rare in tourism, our paper argues for its usefulness and provides an example of application, using consumer data to facilitate decision-making and support the festival industry. The paper concludes with recommendations for future research, as well as incorporating LDA with other qualitative and quantitative research.

Keywords: social media analysis, Latent Dirichlet allocation, topic modelling, customer experience, festival experience

1. Introduction

Festivals are an essential part of the tourism industry. They help build the image of the host country and can even be used effectively in the country's promotional materials. It is a well-known fact that this in turn can affects the economy, including additional revenue and employment growth (Bruwer & Kelley, 2015; Cianga, 2020; Crompton *et al.*, 2001; Duarte *et al.*, 2018; Felsenstein & Fleischer, 2003; Gamage & Higgs, 1996; Gursoy *et al.*, 2004; Kim & Petrick, 2005; Rivera *et al.*, 2015; Tohmo, 2005; Vecco & Srakar, 2017). However, these economic benefits may be overestimated as the majority of spending happens within the festival space that is often not operated by the local residents (Long & Perdue, 1990). Overall, festivals have a complex impact on the locals with both positive and negative effects. Besides economic benefits, positive effects include cultural enrichment, the growth of national pride and identity, a strengthening of the sense of community, and even the involvement of locals in the festival (Arcodia & Whitford, 2007; Bakas *et al.*, 2019; Gursoy *et al.*, 2004; Kwiatkowski *et al.*, 2020; Moscardo, 2007; Negruşa *et al.*, 2016; Small, 2007; Yolal *et al.*, 2016). The negative impacts found were social (for example, rise in crime rate, traffic, conflicts between festival-goers and locals) or environmental costs (for example, noise pollution, litter, damage to the natural environment) (Deccio & Baloglu, 2002; Deery & Jago, 2010; Gursoy *et al.*, 2004; Pavluković *et al.*, 2017; Sulyok & Sziva, 2009). As Moisescu (2020) highlighted, these issues need to be taken seriously and minimised as they affect the hosts' quality of life and well-being. National culture was found to influence the locals' perception of the impact of festivals, but as cultural differences were measured only using Hofstede's cultural dimensions scale, this finding needs further research (Pavluković *et al.*, 2017; Venaik & Brewer, 2016).

Customer experience in events and festivals is becoming a popular academic research topic, which is now seen as an important element of customer behaviour (Brown & Sharpley, 2019; Geus *et al.*, 2016). The majority of articles on the topic have focused on the connection between festival experiences and the satisfaction of the festival-goers (Christou, *et al.*, 2018; Kim, 2015; Oklevik *et al.*, 2022) quality of life (Lee, 2017), subjective well-being (Tan *et al.*, 2020) and behaviour intent (Borges *et al.*, 2021; Giaccone & Galvagno, 2021; Lee *et al.*, 2008; Pivac *et al.*, 2019; Tanford & Jung, 2017; Zhang *et al.*, 2019). Some found a direct link between festival experiences and behaviour intent (Cole & Illum, 2006; Dalgıç & Birdir, 2020; Ding & Hung, 2021; Drengner *et al.*, 2018; Kim *et al.*, 2017). Others found a link between festival experiences and aspects of well-being (Ballantyne *et al.*, 2014; Kinnunen *et al.*, 2020). Moreover, several studies supported the idea that an overall positive experience has a positive impact on how memorable the experience will be (Ding & Hung, 2021; Manthiou *et al.*, 2014; Rivera *et al.*, 2015; Semrad & Rivera, 2018; Wood & Kinnunen, 2020). This in turn positively impacts their future behaviour (including future intention to revisit, saying positive things, recommendations, electronic word-of-mouth or eWOM, loyalty) (Ding & Hung, 2021; Manthiou *et al.*, 2014; Rivera *et al.*, 2015; Semrad & Rivera, 2018; Wood & Kinnunen, 2020). This makes sense as it is established that people are more likely to remember emotional events than more neutral ones (Kensinger & Corkin, 2003).

One of the desired behavioural outcomes is providing online recommendations or eWOM. Semrad and Rivera (2018) found that memorable experiences had an impact on providing positive eWOM (including the posting of positive comments and photos on Facebook, Instagram, and Twitter) for the Aruba Electric Festival. Süli and Martyin-Csamangó (2017) acknowledged that Facebook, Instagram, and Snapchat are the main platforms where festival-goers share their festival experiences. Iványi and Bíró-Szigeti (2020) added that the Generation Z audience of Hungarian music festivals does not only share their experiences to create memories but also to help others or become a social media celebrity. This behaviour implies that music festivals not only have a self-experience utility for festival-goers, they are also useful to signal their identity to others (Bronner & de Hoog, 2021). This conspicuous consumption is in line with the consumer behaviour trend of moving away from collecting and displaying material purchases to seeking unique experiences. According to Bronner and de Hoog (2019), this trend shift has been impacted by the rise of social media. EWOM is a desired behaviour due to several reasons. First,

from the festival-goers' perspective, it is very useful in the initial decision-making phase about festival attendance. As tourism services cannot be evaluated in advance of consumption, WOM is one of the primary sources of information in tourism, and its electronic version, eWOM, provides this at scale, primarily on social media platforms (Iványi, 2021; Litvin *et al.*, 2008; Park & Lee, 2009). This is particularly true for the younger Generation Y and Z consumers who make up the core users of these platforms (Bolton *et al.*, 2013; Turner, 2015). Additionally, the high cost of festival attendance (tickets as well as transportation) means that these are high-risk purchases, making reference group evaluations even more important (Litvin *et al.*, 2008). Second, these social media platforms are used not just to gather information during decision-making, but they also help with information for planning and preparations as well as helping to build positive emotions and a sense of communitas ahead of the event (Brown *et al.*, 2020; Hudson *et al.*, 2015). While this kind of social media communication is reduced during the live experience (as it moves offline), it continues after the event. Third, from a marketing perspective, eWOM was found to strengthen the music festival – customer relationship with young consumers (Strand & Robertson, 2020). Finally, eWOM provides marketing managers with an opportunity to extend their communications reach in a cost-effective, and even more importantly, a trustworthy way (Ismagilova *et al.*, 2017).

Additionally, these eWOM activities generate vast amounts of data; with information becoming electronic and available on the internet. This provides access to huge amounts of data in written, visual, and other formats. In addition to information available in a structured form, a large amount of unstructured information is generated every day; however, its availability, retrievability and processability are often limited, especially in the case of posts on social media and other platforms. In recent years, a whole range of computer technologies – including internet search engines and statistical methodologies supporting them, machine learning algorithms and artificial intelligence solutions – have been developed to collect and manage 'big data' in text, image and other formats and extract useful information from them, revolutionising the way information is processed.

Previous research examined the elements of festival experiences as they appear in eWOM through manual analysis. As an immediate antecedent to the present research, Hinek and Kulcsár (2019) investigated the experiential dimensions of the Sziget Festival in this way. In their research they conducted a qualitative text analysis of guest reviews on the Sziget Festival Official website. In doing so, they identified several topics (different factors of festival satisfaction) in the comments written by visitors, which they organised using a hierarchical mind map. Only a few articles within tourism, leisure, and hospitality literature have analysed customer reviews with eWOM data and computer modelling (see Calheiros *et al.*, 2017; Hu *et al.*, 2019; Kirilenko *et al.*, 2021; Smith *et al.*, 2020). Several of these articles used latent Dirichlet allocation or structural topic model for topic modelling (e.g., Hu *et al.*, 2019; Korfiatis *et al.*, 2019). The authors often used different data sources (e.g., TripAdvisor, Park & Ha; 2017) and contexts (e.g., hotels, Gao *et al.*, 2022; Guo *et al.*, 2017; or Korea, Sutherland *et al.*, 2020), but none of these studies used Facebook as the source, and they did not focus on festival experience. Our study proposes to fill this specific research gap. We analyse written festival guest reviews available on Facebook using computer-based methods to identify which factors of the experience phases are most important in determining satisfaction. We investigate what topics can be identified in the guest reviews and how these topics relate to the festival visitor experience analysed in the literature. First, the recent key literature that is relevant for the analysis of music festival experiences and in particular the in-depth breakdown of their key elements will be discussed. The study includes a range of music festivals, from the United Kingdom and the Netherlands to China. It begins with the definition of festival experiences and their three phases, then the main part of the review analyses the elements of the music festival experience. Finally, selected feedback on the Sziget Festival will be analysed with latent Dirichlet allocation (LDA) and its results compared with previous literature findings.

2. Elements of music festival experiences

Festival experiences focus on the particular space and time of the event, and studies often focus on these experiential attributes. In contrast, only a few studies have analysed a more diverse range of experiences that happen before and after the festival, in the pre- and post-liminal phases. In this sense, the concept of liminality is understood as transformative,"out-of-the-ordinary" experiences, those that are rarely experienced in normal life (Park, 2016; Szmigin *et al.*, 2019; Wood *et al.*, 2020). Therefore, the paper divides festival experiences into three main phases, pre-liminal, liminal, and post-liminal phases. First, experiences during the pre- and post-liminal phases will be examined.

2.1. The pre- and post-liminal phases

Anticipation begins approximately one month before the festival; the netnographic research done undertaken by Wood *et al.* (2020) showed that the amount of group communication on WeChat significantly increased from around 20-50 to 500-1,600 conversations per day as the event start date was approaching. The increasing presence of emotions follows a similar pattern. Conversations were about the logistics of preparing for the upcoming event, for example, travelling, tents, and outfits as well as a sense of excitement and a discussion of memories of past festival experiences. They even noted a daily countdown. The study was undertaken in China, based on the Midi Festival, the largest music festival in China, where such festivals are rare and therefore heightened emotions are expected. Other researches drew similar conclusions, e.g. Wood and Moss (2015) highlighted that these emotions are oscillating and multifaceted, from great excitement about the upcoming festival to some anxious thoughts about their responsibilities to complete before the event. This pre-liminal phase also includes the journey to the festival when festival-goers have left their home and everyday lives behind. This physical distancing from their normal lives is filled with excitement and happiness, with some communitas starting to form among the festival-goers (Wood *et al.*, 2020).

Once the festival has ended, the post-liminal phase is important because choices are made based on memories of events (Kahneman, 2011). Right at the end of the festival emotions are at a peak with intense feelings of a sense of loss, deep attachment, and communitas. When examining its long-term effects, these intense feelings linger for a while even after participants return to their everyday lives. Many festival-goers experienced communitas (shared taste in music, symbols, and experiences) through online groups long after the festival ended. Attendees give names to this group identity, for example, attendees of the Midi Festival call themselves "Midier". These groups facilitate friendships that make people chat and even meet in between festivals, thus increasing the participants' social capital. One of the common ways that festival management can encourage the formation of these groups is through official fan clubs where they can also join in on the conversation and help keep the conversation going (Barrière & Finkel, 2022). Material objects, such as festival T-shirts – apart from their direct sales benefits – are symbolically charged in that they help display their social identity and maintain the communitas after the festival (Barrière & Finkel, 2022).

The pre- and post-liminal phases are under-researched but important phases. In the next section, the core festival experience and the elements of the atmosphere of the festival itself will be discussed in depth: the environment, the social aspect, the various forms of entertainment, the value and the festival image.

2.2. Environment

Kinnunen *et al.* (2020) established that the environment is one of the key elements of the festival experience. As mentioned in the concept of liminality, festivals exist in an environment that is limited in a time and space that is different from everyday life. It is argued that this physical environment creates a new aesthetic experience, a different atmosphere (Tomazelli, 2017). There are different points of view on what this includes. The literature defines the environment such as food and drinks, staff

professionalism, cleanliness, restroom, information, parking, accessibility, and other amenities and facilities (Alvarado, 2022; Bohn D., de Bernardi C. 2022; Tanford & Jung, 2017; Tan *et al.*, 2020; Özdemir & Çulha, 2009: Lee *et al.*, 2008; Tan *et al.*, 2020; Welthagen & Lötter, 2020; Yoon *et al.*, 2010). These can be classified as Herzberg's hygiene factors that are necessary to be met to avoid dissatisfaction, but on their own, they are not sufficient to cause satisfaction. Satisfaction is instead achieved by other, motivator attributes (Crompton, 2003). Another perspective includes the performances' venue characteristics (including layout, stage setting, and acoustics), as well as the weather conditions (Sun *et al.*, 2019; Wilders & Rusch, 2020). In the case of outdoor festivals, people are more enthusiastic and active during nice weather conditions (Sun *et al.*, 2019).

A study about the SummerJazzCycleTour found that the surroundings and transport methods of the festival were also integral parts of the experience as participants said that cycling from one concert to the next helped them discover new landscapes, allowed them to socialise, and helped them mentally prepare for the concert ahead (Wilders & Rusch, 2020). Another location-based factor, proximity to home is also seen as a reason for attendance but it is not one of the main ones (Welthagen & Lötter, 2020). Iványi (2021) suggested that for the Hungarian members of Generation Z, the surrounding city or place was not so important – this may indicate that this group of festival-goers do not intend to spend much time outside of the festival site.

The term"festivalscape", based on the concept of servicescape, establishes that the environment impacts people's behaviour (Duffy *et al.*, 2011; Sun *et al.*, 2019). It is understood as a place where festival attendees can meet and engage with each other, and this then defines the nature of their social interactions. In physical proximity, there is a balance to be achieved because on one hand – depending on individual differences – too much crowding may not be perceived as optimal (Rasoolimanesh *et al.*, 2019). However, too few people with too much distance between them would not achieve a good atmosphere in which emotions are contagious and spontaneous interactions can easily occur (Sun *et al.*, 2019). The right kind of tourist interaction can therefore be influenced and encouraged by festival management by careful planning of the environment (Sun *et al.*, 2019).

2.3. Socialisation
Festivals are inherently social places. The social aspect of the environment has been mentioned already, and compared to the physical environmental elements academics see it as more important in the festival experience (Jang *et al.*, 2015). Sun *et al.* (2019) and Iványi (2021) argued that socialisation is the main motivator for some festival-goers, especially for first-timers. Research argued that the most memorable part of the experience was the points of interaction, both with others and the performers (Marline & Wilders, 2020; Zhang *et al.*, 2019). Mulder and Hitters (2021) compared festivals to concerts and found that festivals are primarily visited for the social aspect, while in contrast, concerts are attended for the music itself. In addition, it is not uncommon that the socialisation aspect is so important for some that they buy their festival tickets before the musical performances are even announced (Iványi, 2021). This socialisation behaviour is affected by individual differences such as extraversion, perceived similarity, and the individual's nationality (Zgolli & Zaiem, 2017). Therefore, socialisation is determined by individual differences as well as the environment, as mentioned above. This social aspect has been conceptualised as the "social festivalscape" and many academics argue that the emotions of other festival-goers greatly impact the individuals' own. Furthermore, as the social facilitation theory states, even the mere presence of others impacts human behaviour (Zajonc, 1965).

A specific concept within socialisation, tourist-to-tourist interaction (TTI) has been discussed in various contexts such as cruising, backpacking, and group tours (see Adam, 2021; Huang & Hsu, 2010; Koç *et al.*, 2022; Levy *et al.*, 2011; Lin *et al.*, 2019; Murphy, 2001; Papathanassis, 2012; Rihova *et al.*, 2018; Sökmen &

Yolal, 2022). Sun *et al.* (2019) found that there are three distinct types of TTI with varying impact on the festival experience: entertainment, mutual assistance, and conflict. Entertainment interactions include musical elements, including interactive play of jumping together during performances and playing, singing, or drinking together. Mutual assistance is enabled by the altered social identities within the festival atmosphere, making people more likely to help others. This is especially present in interactions with first-time visitors to make the most of the festival, create the best shared experience, and build relationships. The third interaction type, conflict arises due to individual differences, for example, unsocial acts such as drunkenness, loud and aggressive behaviour (Sun *et al.*, 2019).

Another categorisation of the social servicescape came from Kinnunen *et al.* (2020) that looked at festival experiences of Finnish adolescents. They separated bonding, bridging, a sense of community, and a sense of belonging. Out of these four categories, sense of belonging was seldom present in the adolescents' music festival experiences; in contrast, it appears in the experiences of adults, in particular during metal or punk events. The other three categories appear in the experiences of adolescents as well as in those of adults. First, bonding and bridging originate from Putnam's (2000) approach to social capital, and these are developed on an individual level (Kinnunen *et al.*, 2020). On one hand, bonding, defined as quality time spent with a known group, induces happiness, strengthens existing relationships, creates shared memories, and ultimately enhances their well-being (Kinnunen *et al.*, 2020). It can be observed at many festivals (Nordwall *et al.*, 2014; Rihova *et al.*, 2015; Wilks, 2011). Bonding is typically practiced with friends and family, and in the case of Kinnunen *et al.* (2020) it happened with friends in particular, which makes sense when considering the young age of the sample. On the other hand, bridging can be defined as more outward-looking, an opportunity to make new friendships. It is important that this goes beyond simple occasional interactions and requires individuals to form lasting connections (Nordwall *et al.*, 2014; Kinnunen *et al.*, 2020). This is seen to be difficult to achieve and Wilks (2011) and Iványi (2021) did not find these bridging practices common at the pop and folk festivals in their studies, despite a general sense of safety and friendliness at these events. Additionally, this aspect of festivals is perceived as more important for younger participants than older ones as it can help expand a limited friendship circle from school by having another group of friends from outside. It also helps with establishing romantic relationships, something that is very important for young people (Kinnunen *et al.*, 2020; Nordwall *et al.*, 2014). Interestingly, the distinction between bonding and bridging may become blurred as a characteristic of long-running festivals is that people return each year, therefore transforming relationships between festivalgoers from bridging into bonding relationships. Quinn and Wilks (2016) proposed a new term to define this phenomenon, called "recurrent bonding social capital". Then the next category of social servicescape after bonding and bridging was the creation of a sense of community. This is a temporary experience concerning a wider group of participants, it develops due to the shared nature of the live event. This sense of community is difficult to articulate and thus festivalgoers might call it the social atmosphere or the live atmosphere of the event. However, it is certainly one of the main motivations for festival attendance (Kinnunen *et al.*, 2020). According to the joint attention model, highly emotional, both enthusiastic and disapproving behaviour from others (cheering, booing, singing, jumping, dancing, etc.) drives collective attention and emotions, becoming fully immersed together in the event (Cochrane, 2009). This is linked to the social proof effect, where behaviour is affected by others' behaviour, which signals what appropriate behaviour looks like (Cialdini, 2001). Furthermore, people at festivals are aware of this joint attention and this mutual awareness creates a sense of community, or "we mode" between them (Cochrane, 2009; Kinnunen *et al.*, 2020).

Finally, the festival environment allows for a more intimate, more self-centred experience as well. The limited space and time of a festival enable festival-goers to escape their everyday lives and thus interact more openly and positively, and enable self-expression of identity that may not be possible outside (Bohn & de Bernardi, 2022; Kinnunen *et al.*, 2020; Szmigin *et al.*, 2017). This opportunity and acceptance

of self-expression are one of the key motivators for festival attendance as it gives a sense of safety and inclusion to participants. This allows festival-goers to reveal their authentic and real selves, such as displaying their tattoos or their deepest selves with the lyrics of the music (Wood *et al.*, 2020). In the end, this then makes festival revisit more likely (Kinnunen *et al.*, 2020).

2.4. Entertainment
The primary part of the program is the musical line-up and the headliners. However, there is a trend of diversification recently as other entertainment is becoming more widespread and diverse (see Bohn & de Bernardi, 2022; Brown & Sharpley, 2019; Szabó, 2015). Cole and Chancellor (2009) argued that a downtown festival's quality of live entertainment and live bands is more important than other programs (for example, various booths and exhibits) and amenities (for example, accessibility, restrooms, food and beverage, hygiene). Wen-Chiang *et al.* (2012) came to a similar conclusion, with the performance programming found to be more important for festival quality than the staff, the environment, accessibility, information, cleanliness, and the restrooms.

Of course, music festivals would not exist without music but the importance of it varies between individuals (Bowen & Daniels, 2005; Brown & Sharpley, 2019). For example, Iványi (2019) noted that Hungarian Generation Z music festival-goers saw music as the second most important element of their experience, overtaken only by socialising. Listening to music has clear benefits, for example, in the regulation of emotions and relationship building, in particular in a live music setting with an added shared social aspect (Kinnunen *et al.*, 2020; MacDonald, 2013; Papinczak *et al.*, 2015). When diagnosing music festivals, programmes were found to have the most substantial and positive significant relationship with festival value and together with information services and food they accounted for almost half of the festival value (Tan *et al.,* 2020). These findings are in line with findings from other festivals, for example, Vinnicombe and Sou (2017) and Yoon *et al.* (2010). Welthagen and Lötter (2020) even attested that the primary reason for attending the Park Acoustics Live Music Festival was the variety of entertainment and socialising came second.

Additional programming includes every other entertainment that is not part of the core offering but reflects the central theme of the festival (Tan *et al.*, 2020). Interestingly, Brown and Sharpley (2019) found that out of the analysed elements of music festival experiences, entertainment, added value and music were the most important, while services, engagement, the festival image, and ethics did not significantly impact the perceived experience. Iványi (2021) also identified additional programming as an element of the festival experience, but one of the less important ones.

2.5. Value
Perceived value is obviously an important factor in any purchase; Welthagen and Lötter (2020) found that this was the third most important factor of music festivals, following socialisation and entertainment. Meanwhile, Iványi (2021) put it in fourth place, the final one within the most important elements, following socialisation, programming, and the atmosphere. Yoon *et al.* (2010) claimed that a festival's value impacts visitor satisfaction and loyalty. The quality of the festival is found to impact the amount festival-goers are willing to pay for the event (Baker & Crompton, 2000).

Some research has been done on added value (for example, VIP upgrades and personalised experiences), with promising results, but more research is needed. Brown and Sharpley (2019) suggested that added value was one of the important elements of perceived value of music festivals. This aligns with previous research findings of Choo *et al.* (2016), Özdemir and Çulha (2009), and Yoon *et al.* (2010).

2.6. Festival image

Brown and Sharpley (2019) found that a festival's image, services, ethics, and environmental concerns are factors in festival experiences, but these are less important than music, added value, or entertainment. In contrast, Leenders (2010) emphasized that brand image as well as the festival atmosphere are more important for loyalty than the actual value of the music festival. Ding and Hung (2021) argued that what is important is a congruence between the festival's image and the festival-goer's self-image.

2.7. Specific kinds of music festival experiences

Two distinct but related concepts are the flow experience and authentic experiences. Their key factors and how these link to the elements mentioned earlier will be discussed. Flow experiences have been researched in several tourism contexts (see Chang, 2014; Chen *et al.*, 2017; Cheng & Lu, 2015; Fu *et al.*, 2017; Kim & Thapa, 2018; Wu & Liang, 2011; Zhang *et al.*, 2019). Ding and Hung (2021) defined the flow experience as the quality of performance, ambiance, self-congruence, other consumers' passion, and consumer-to-consumer interaction. The quality of performance is similar to the entertainment aspect, and ambiance includes the sensory elements (for example, smell, lighting, temperature, colour, cleanliness), which is similar to the environmental aspect. Self-congruence implies a match of the self-concept with the festival image. The final two, other consumers' passion and consumer-to-consumer interaction would fit under socialisation, focusing on the emotional and behavioural aspects respectively. Szmigin *et al.* (2017) argued that big commercial festivals can also create authentic experiences as the size and exploratory setting of big festivals allow people to experiment with various performances and thus contributing to an authentic experience. Place and social interactions are key in authentic experiences. The physical liminality of festivals separates festivalgoers from their routine, joins them with like-minded people, and allows them to be free. This shows links to the environmental element. The resulting social element combines intra-personal (self-identity, self-discovery, self-expression, self-making) and interpersonal (social belonging, bonding, communitas, bonding) elements.

2.8. Summary of the literature on festival experiences

Therefore, based on the literature detailed in Section 2, the following prioritisation of the festival elements can be established. Socialisation, entertainment, festival value and image are seen as the core elements of having positive festival experiences. The 'right' environment and atmosphere are necessary to avoid negative experiences.

3. Categorisation

It is particularly important that festival experiences and their elements are well understood. However, there is more than one way of categorising elements and different ways of measuring (Tan *et al.*, 2020). Some researchers focused on the "festivalscape", while others split the experience into several categories. A popular approach is from Pine and Gilmore (1998), who developed the experience economy model with four factors relating realms of experience: entertainment (passive experiences, for example, watching TV), educational (more active participation, for example attending a class), escapist (more immersed with active participation, for example, acting), and aesthetic (immersed but not active, for example, visiting a gallery). They argued that the richest experiences encompass all four factors and as a result create memorable experiences (Pine & Gilmore, 1988). Aşan *et al.* (2020) used these four factors to analyse the rock music festival Kuzey Fest in Turkey and argued that aesthetic and entertainment elements are the most important, followed by escapist and education elements. Furthermore, this experience economy model was extended by a fifth element, economic value by Rivera *et al.* (2015), creating the 5E model in a music festival context. Based on the 5E model, Iványi and Bíró-Szigeti (2020) found that in the context of Hungarian music festivals several of these elements are interconnected. Education, entertainment, and escapist elements are influenced by aesthetic elements

and economic value is influenced by the other four elements (Iványi & Bíró-Szigeti, 2020). Finally, what is important to remember is that these elements are intangible and subjective and therefore difficult to define (Brown & Sharpley, 2019).

Additionally, as there is no given categorisation and no single measurement tool either, researchers used and adapted a range of scales in their studies, for example, Baker and Crompton (2000); Bowen and Daniels (2005); Chen *et al.* (2012); Cole and Chancellor (2009); Geus *et al.* (2016); Hudson and Hudson (2013); Kim *et al.* (2018); Lee *et al.* (2008); Manthiou *et al.* (2014); Morgan (2008); Oh *et al.* (2007); Parasuraman *et al.* (1988); Yoon *et al.* (2010). This means that there is no comprehensive list of festival experience attributes, and the selection is dependent on the nature of the study, and in particular, its context and individual differences. Several studies highlighted that individual differences exist in festival experience perceptions based on sociodemographic, psychologic, or behavioural factors (for example, age, gender, education, marital status, personal values, subjective well-being, previous experiences, etc.) (see Aguado *et al.*, 2021; Altunel & Koçak, 2017; Brown & Sharpley, 2019; Saragih & Amelia, 2020). As mentioned earlier, there are several possible categorisations of festival experiences, such as the experience economy or the 5E model.

4. The scientific background of computer-based topic modelling

For computer-based topic modelling, the Latent Dirichlet allocation (LDA) method was used. LDA is a so-called unsupervised machine learning algorithm that identifies topics as latent (hidden) information in large collections of documents. It is unsupervised in the sense that it does not use a pre-built vocabulary or glossary to identify topics, the algorithm is not "trained" in advance by the researcher.

Latent Dirichlet allocation (LDA) is based on the hypothesis that an author writes a document with certain topics in mind. For a given topic, a given word is chosen with a certain probability from a set of words specific to that topic. The document as a whole can be described as a mixture of different topics. LDA can be interpreted as a "backpropagation" of topics; its modelling process can be described as determining a mixture of topics in documents (in this case, social media comments) as a probability distribution, and then assigning to each topic the terms (words) specific to that topic, which assignment also follows a probability distribution (Krestel *et al.*, 2009). LDA thus assumes that each document in a document collection (corpus) can be represented as a probability distribution of K number of hidden topics, and each topic can be defined as a multinomial distribution of words that make up the vocabulary of the documents (Blei *et al.*, 2003).

The text collection (corpus) contains M pieces of documents. Each document consists of N_i number of words and $w_{d,n}$ denotes the n_{th} word of document d. All words in the corpus are grouped into K number of topics. V denotes the individual words of the corpus, Z denotes the topic to which the words forming a document belong. Each topic can be interpreted as a multinomial distribution over the set of words making up the document. α and β are priori Dirichlet hyperparameters, their values are typically low, based on the assumption that the number of topics and the number of words belonging to each topic are limited, i.e., "sparse". The model has a single observed variable, the words that make up the corpus.

The hidden variables to be inferred from the model are:
- θ_d, the probability distribution of subjects over document d,
- ϕ_k is the probability distribution of the words associated with topic k,
- z_{dn}, is the "topic label", indicating the topic to which each word in the documents belongs.

The generative process of LDA is as follows (Balogh, 2015):
1. For each k topic, choose a word distribution $\phi k \sim Dir(\beta)$.

2. For each document
   a. choose a topic distribution $\theta_d \sim \text{Dir}(\alpha)$.
   b. For each word
      i. choose a topic assignment $z_{dn} \sim \text{Mult}(\theta_d)$, where $z_{dn} \in 1....K$;
      ii. choose a word $w_{d,n} \sim \text{Mult}(\phi_{zd,n})$, where $w_{d,n} \in 1, \ldots, V$.

The graphical representation of LDA is shown in Figure 1.
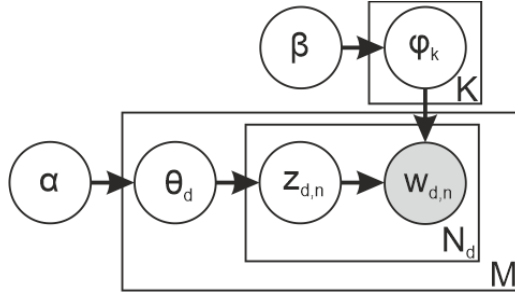
Source: Blei *et al.*, 2003; Balogh, 2015

The outer tray represents the documents, while the inner tray represents the repeated selection of topics and words in the document

Figure 1: Relationship between variables in the LDA model using a "tray" representation

The posteriors of the various distributions of the model, i.e., the set of topics, the probabilities of the words associated with them, the topics associated with each word, and the specific topic mixes of each document, are obtained by statistical inference. Several methods can be used, but the most commonly used is Gibbs sampling, which is highly algorithmic.

A further feature of LDA is that it does not take into account the position of words, the order of words in documents, and the order of documents themselves. For this reason, LDA is often described as a "bag of word" model (Blei *et al.* 2003). It is also important to note that LDA is a mixed-membership model, i.e. it is not a traditional classification, and the words identified for the topics as a result of the modelling will not be exclusive, the same word may appear in several topics with different weights and probabilities (Airoldi *et al.*, 2014).

It is also important to note that LDA is a mixed model, i.e., it is not a traditional classification, and the words identified for the topics as a result of the modelling will not be exclusive, the same word may appear in several topics with different weights and probabilities (Airoldi *et al.*, 2014).

Several researchers have proposed improvements of the original model, such as improving the predictive ability of the topic models; see for example, Paul and Dredze, 2015; Gerrish and Blei, 2012; Weinshall *et al.*, 2013; Wilson and Chew, 2010; Einstein *et al.*, 2011; Lafferty and Blei, 2006, etc. The present study makes use of Roberts and co-authors (2014), who proposed an improvement of the original LDA model. The structural topic model (STM) developed by them is suitable for processing texts for which document-level metadata are available. These metadata can be arbitrary, for example, the document's date of creation, the nationality of the commenter, possibly the number of likes, etc. They can be continuous (for example, the acceleration metric in vehicle data sheets) or discrete (for example, a rating on a five-point scale). The outputs of the model can also be used to test a hypothesis or hypotheses describing the relationship between metadata and topics (Roberts *et al.*, 2014).

5. Research method
5.1. Research context
Since the end of the Soviet era, the Hungarian festival market has experienced fast growth, and the year 2015 was outstanding both in terms of visitor numbers and revenues (Heineken, n.d.). Szabó (2015), an expert in Hungarian festivals sees that further growth will come from constant innovations. This is what can provide the necessary competitive advantage, by continuously building on a clear concept and mission.

The flagship event of the Hungarian festival industry is the Sziget Festival. The first Festival was organized in 1993 and since then only the events in 2020 and 2021 were cancelled due to the Covid-19 pandemic. Sziget means "island" in Hungarian, and it is located in Budapest, on Óbuda Island. Because of this, the festival is often called the Island of Freedom, reinforcing its sense of being in a bubble, an enclosed space. It welcomes hundreds of thousands of festival-goers or "Szitizens" each year with a variety of genres across 60 stages, over 1,000 shows for 6 days in August. Each year, the biggest names in music appear on its stage, for example, in 2014, The Prodigy, Calvin Harris, and Bastille; in 2015, Robbie Williams, Ellie Goulding, and Avicii; in 2016, Rihanna, David Guetta, and Sia. Although it is primarily a music festival, it also includes other entertainment including art, talks, theatre, and circus experiences. Sziget has won several awards. In 2022, the Evening Standard named Sziget Festival as one of the best European festivals to visit (Embley, 2022). Recently it won the UK Festival Awards in the Best Overseas Festival category (UK Festival Awards, n.d.). It is appealing to foreigners with 100 nationalities present, and 7 out of 10 visitors coming from abroad. This may be because the costs of tickets are half of the average of festivals of similar size in Western Europe (Heineken, n.d.; Menegaki *et al.*, 2020; Sziget Festival, n.d.; Sándor, 2022). Festivals are more popular amongst younger tourists in Budapest (Smith *et al.*, 2022).

5.2. Data collection
Despite a wealth of research, most of the articles examined festival experiences with questionnaires or other qualitative methods such as semi-structured interviews or observations. Social media as data source in music festival research is less common (see Brown *et al.*, 2020; Carah & Angus, 2018; Danielsen & Kjus, 2019; de Lira *et al.*, 2019; Gilstrap *et al.*, 2021; Hudson & Hudson, 2013; Hudson *et al.*, 2015; Strand & Robertson, 2020). The data was collected from the guest reviews section publicly available on the Sziget Festival Official Facebook page. The survey was conducted considering reviews written between 1 January 2014 and 31 July 2020. During this period, thousands of posts were found, but most of them did not contain any textual information or opinions, only a rating of the festival on a scale of 1-5. About 8-10% of the posts contained textual information, the frequency of which varied from year to year. The highest number of text entries was recorded in 2017, followed by 2016, 2015 and 2014. 2018 was the last year with a text entry in the reviews section, with only numerical ratings in the following years. The festival was cancelled in 2020 and 2021 due to the pandemic and held again for the first time in 2022. In those years, between 2019 and 2021, no relevant guest reviews were published on the Sziget Festival Official Facebook page. The collection of data ended in July 2022, before the 2022 Sziget Festival.

It is also important to note that Facebook has changed the review section. For a long time, guests could rate the festival on a scale of 1 to 5, but in recent years the scale has disappeared, and reviewers can simply recommend or not recommend the Sziget Festival by pressing the yes or no buttons.

5.3. Data analysis
Since Facebook prohibits automated data collection and web scraping, and quickly disables profiles that request them, not even allowing the page owner to download guest reviews, the data were collected manually. In addition to the reviews, the name (profile name) of the reviewer was recorded, their

country of origin (if identifiable), the total number of likes and other emoticons, the posting person's rating on a scale of 1-5, and the date of the post. The comments posted under each post were also collected, as they also contained valuable information for our analysis, but the date of posting was not available.

In the next step, the cleaning of primary data was performed. The most important decision was to keep the English entries only. Although topic modelling can be conducted in any language, the statistical nature of the methodology means that the analysis can only be carried out at the same time efficiently on a corpus of the same language, otherwise terms with the same meaning but in different languages and the topics modelled from them will not be coherent. In any case, around 60-70% of the entries were in English, especially those dealing with the 'substantive' aspects of the festival (music, catering, ancillary services, visitor experience), while the entries in Hungarian were often not written by festival visitors.

During the further data cleaning process, we excluded those entries that were included as spam (for example, entries about the noise of the festival, several times in a row with exactly the same content) and filtered out the responses written by Sziget Festival Official.

Prior to the linguistic preparation, the corpus of 752 entries contained a total of about 37,000 words and 3,800 individual words (phrases), totalling about 175,000 characters. Compared to the research history, a small corpus was created, which may have not necessarily led to satisfactory results in machine-based statistical probabilistic processing, especially since the procedure determines topics based on co-occurring terms/words, so a larger corpus may be preferable (Crossley *et al.*, 2017; Syet & Spruit, 2017). There were some reviews that were several hundred words long, but many entries were short, and there were even some that contained only one word (for example, "Super!"). Since LDA treats each document as a mix of topics and topics are formed from the distribution of words that make up the documents, very short entries are not ideal for topic modelling. There are studies that exclude entries shorter than 7-8 words from topic modelling (Wang *et al.* 2013), other studies do not emphasise this (see the Twitter-based topic models presented earlier, where entries are up to 140 characters long). Considering the size of the corpus, it was decided to keep short entries as a first step, given that in the next step, language pre-processing, many words or documents were dropped from the initial corpus.

For processing, the R (R Core Team, 2022) STM package was used (Roberts *et al.*, 2019), and the analysis was performed entirely with R. In preprocessing the raw data, punctuation, numbers and words that occur frequently but do not carry meaning in the text (so-called "stop words", for example, pronouns and auxiliaries) were removed from the corpus. The same word can have multiple forms, which can bias the analysis, so the words were lemmatized, i.e., word fragments of the conjugated and formed forms (for example amazing → amaz, beautiful → beauti) remained.

In the next step, in order to improve the accuracy of the analysis, rare words were excluded from the analysis, retaining only those that appeared in at least three documents. Similarly, very frequent words that occur in many documents can be omitted (in this case, the words "island" and "festival"), but after trial runs, it was decided to keep them because they do not bias the analysis.

Using the capabilities of the STM R package, two document-level metadata were included in the analysis: the year of entry and the number of tags. The additional document-level metadata (for example, nationality of the commenter, rating on the five-point scale, etc.) had to be excluded from the analysis because they were not available for each post. After the preparation steps, 736 documents and a dictionary of 751 unique terms were used to perform the topic modelling.

An important feature of topic modelling is that the number of topics (K) must be predefined. Two indicators were examined using the options offered by the STM package: semantic coherence, which indicates the extent to which the topics are composed of related words that facilitate the clear identification of the topics, and exclusivity, which expresses the extent to which different words constitute each topic (Figure 2). If the model is running with topics (see bottom right corner of the figure), this would result in the highest semantic coherence value, but the exclusivity of the terms that make up the topics would be low. In this respect, 11 or 12 topics would perform best, but in that case the semantic coherence value would be low. (Somewhat) balanced topic numbers for the two indicators would be 8, 9, 10, 11, 14, although no truly favourable values are found for either indicator. After a short experiment with topic numbers, the 9-topic solution was found to be best interpreted.
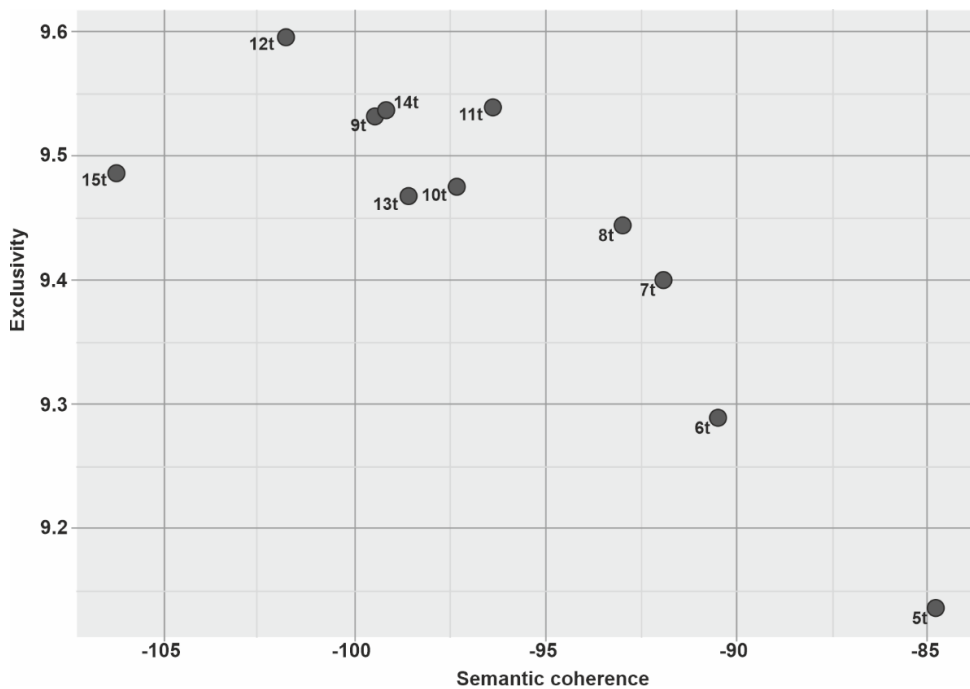

(For both indicators, the higher value is more favourable. In the figure this means the points closer to the top right corner.)

Source: own edit using the STM and ggplot2 R software packages

Figure 2: Exclusivity and semantic coherence for different numbers of topics in the topic modelling of visitors' opinions on the Sziget Festival

6. Results

Based on the statistical probabilities, the algorithm determines the most likely words for each topic and the proportion of topics in each document (comments) and in the corpus (see Figure 3). Topic 5 appears with the highest estimated frequency in the visitor reviews, with a share of more than 18%, while the other topics have a probability of 6-14% according to the algorithm.
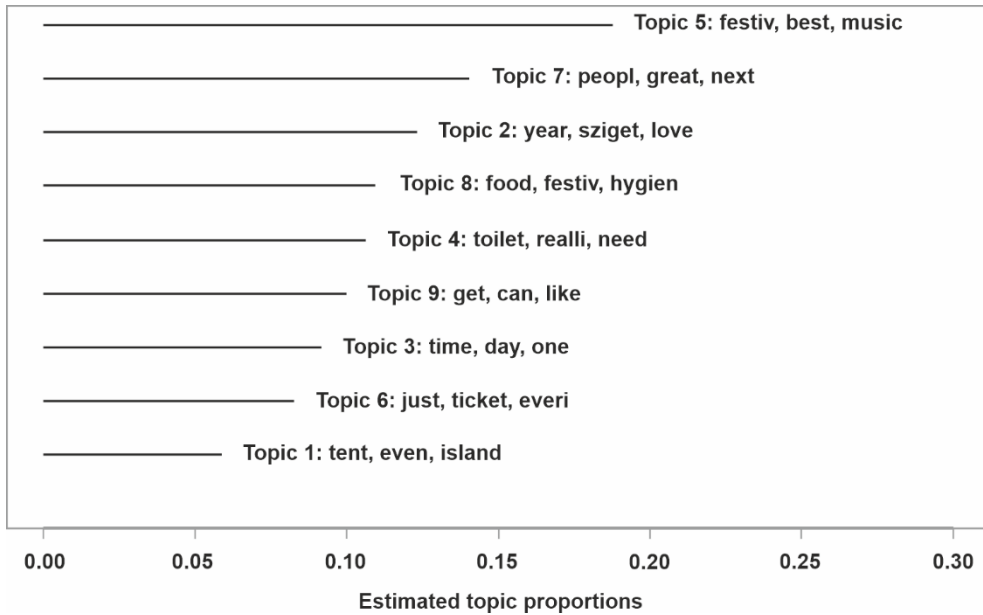
Source: own edit using the STM R software package
Figure 3: Main words and estimated proportions of topics in the comments

The STM package also provides the words that are both frequent and exclusive. These are the words that distinguish topics from each other (FREX). The "lift" and "score" indicators measure the weighted frequency of a word in each topic. The weights are produced so that words that occur less frequently in other topics are given a higher weight.

Identifying topics by words is already a researcher's task; see Table 1.

It is sometimes difficult to identify the actual content of a topic from the words alone, especially when the semantic link between words is weak, so for several topics, the researchers looked at which specific opinions they were most likely to appear in. For example, topic 3 was most likely to appear in two closely related posts on Facebook about whether it is worth coming to the Sziget Festival with children, and how much the commenters agreed. In other comments specific to the topic, commenters also expressed an opinion on a factor ("I think") or supported the opinion of other commenters ("I agree").

The emotional Sziget experience is reflected in several topics. These are topics 2, 5 and 7, which express (almost) the same thing in different words, namely that the Sziget Festival is a great experience. These topics are also correlated with each other, and the algorithm calculated the correlation matrix of the topics (Figure 4). In the figure, the peaks represent the topics, and their size represents the estimated proportions of each topic in the corpus. The thickness of the edges connecting the peaks indicates the strength of the correlation between the topics. The most significant correlation – but only medium in strength – is between topic 2 and topic 5 (Pearson's correlation coefficient of 0.47), while the correlation measure between topics 5 and 7 is 0.27 and between topic 7 and topic 2 is 0.21, indicating a weak to medium correlation.

Table 1: *Words representing the topics*

| Topics	Lemmatised words by different indicators	Topic identification
1.	*Most likely:* tent, even, island, peopl, use, complet, mani *FREX:* past, tent, worth, horribl, complet, even, use *Lift:* direct, gone, leak, worth, yet, across, auchan *Score:* tent, refund, wash, station, often, continu, leak	Experiences with the Sziget campsite
2.	*Most likely:* year, sziget, love, experi, thank, come, still *FREX:* love, thank, year, experi, stay, come, absolut *Lift:* air, heaven, unbeliev, altern, blast, contact, epic *Score:* year, love, thank, absolut, welcom, fun, experi	A great festival experience on the island of the Sziget Festival
3.	*Most likely:* time, day, one, thing, think, line, first *FREX:* agre, think, line, total, guy, idea, thing *Lift:* agre, comparison, total, breakfast, except, front, guest *Score:* guest, time, one, thing, day, line, think	Opinions and thoughts on various aspects of the festival (for example, can children be brought to Sziget?)
4.	*Most likely:* toilet, realli, need, water, price, camp, last *FREX:* water, queue, price, arriv, disappoint, eat, fest *Lift:* bottl, min, queue, tap, add, arriv, belie *Score:* toilet, water, suck, enough, need, problem, price	Opinions on basic sanitation and other services (for example, pricing of bottled water, public taps)
5.	*Most likely:* festiv, best, music, amaz, back, ever, life *FREX:* best, music, ever, life, beer, world, amaz *Lift:* beer, cultur, die, england, ever, excit, greatest *Score:* best, music, amaz, festiv, ever, life, world	The best festival, the greatest experience
6.	*Most likely:* just, ticket, everi, way, stage, feel, crowd *FREX:* crowd, act, budapest, play, push, way, everi *Lift:* act, anniversari, budapest, corner, bare, beach, beat *Score:* ticket, crowd, lineup, budapest, play, way, broke	Opinions on the crowd and the music selection
7.	*Most likely:* peopl, great, next, will, see, place, much *FREX:* great, awesom, look, next, everyon, will, far *Lift:* awesom, alcohol, bin, control, dirti, drunk, environ *Score:* great, next, awesom, peopl, will, simpl, look	Opinions on great festivalgoers, venue, excellent organisation and services
8.	*Most likely:* food, festiv, hygien, staff, got, sziget, realli *FREX:* hygien, staff, dust, visitor, trash, right, food *Lift:* age, bacteria, drive, favourit, feedback, mate, result *Score:* food, hygien, poison, sick, staff, favourit, bar	Experience regarding catering and food hygiene
9.	*Most likely:* get, can, like, sziget, stage, dont, mani *FREX:* concert, sound, live, perform, phone, band, get *Lift:* move, park, phone, sound, steve, aoki, assum *Score:* sound, get, concert, telekom, assum, aoki, chainsmok	Opinions on concerts, artists, performers, Sziget stages and venues

Source: own edit using the STM R software package
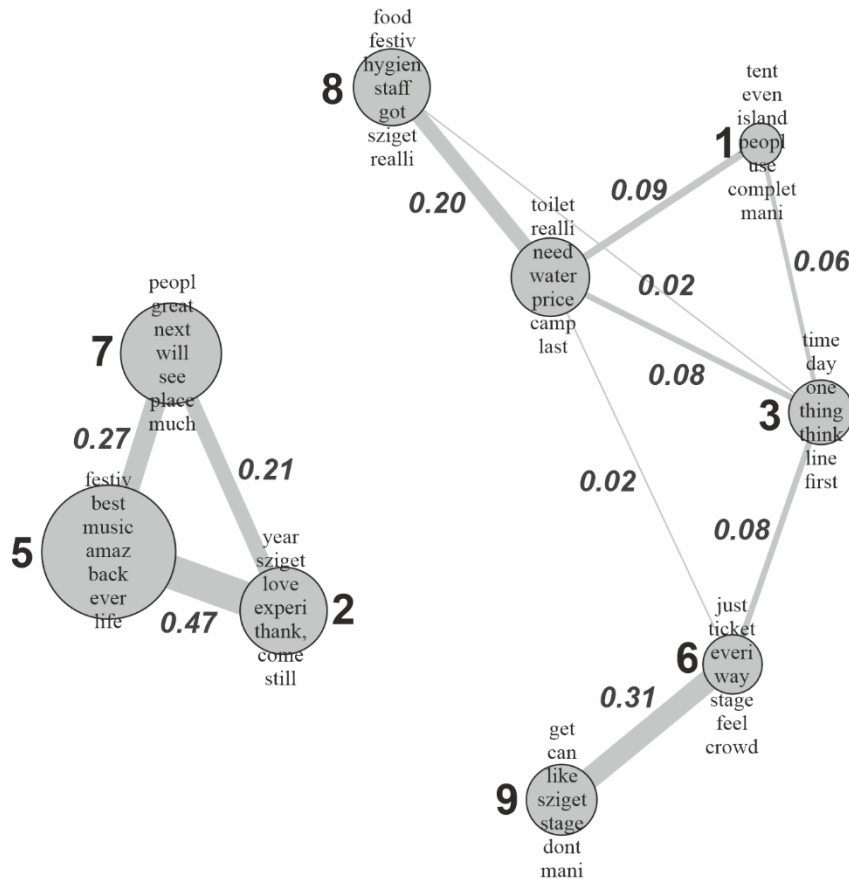
Source: own edit using the STM and igraph R software packages
Figure 4: The weight of the topics and the correlations between them

The figure clearly shows that the topics related to the experience do not correlate with the topics related to the various service factors of the festival. Of these, topic 8 has an average weight of 11% in the comments, with experiences related to catering and food hygiene appearing here. However, topic 8 is slightly correlated with topic 4, which represents opinions on basic hygiene and other services. Also, in this "bush", experiences connected with camping, opinions about performers and line-up, the crowd can be seen – see topics 6 and 9, which are moderately correlated.

Finally, the role of covariates was also examined. In the structural topic model, two metadata were incorporated, the year of the comment and the total number of likes and other tagged emoticons associated with the post. Using the STM package, the effect of covariates on the distribution of certain topics in the comments was investigated (see Figure 5).

The effect of the year of post as a discrete covariate had a particular impact on the presence of Topic 8. Catering and food hygiene had the highest share of comments in 2017, with more than 15%. (In the figure, the dot shows the value of the estimate, and the lines on the right and left indicate the confidence interval of the estimate.) Reviewing the comments, a long stream of comments was found in this year, where several people complained of diarrhoea and food poisoning, because food often spoiled in the summer heat when stored in the hot sun.

(Covariate level: the level of covariate represented by the different years)
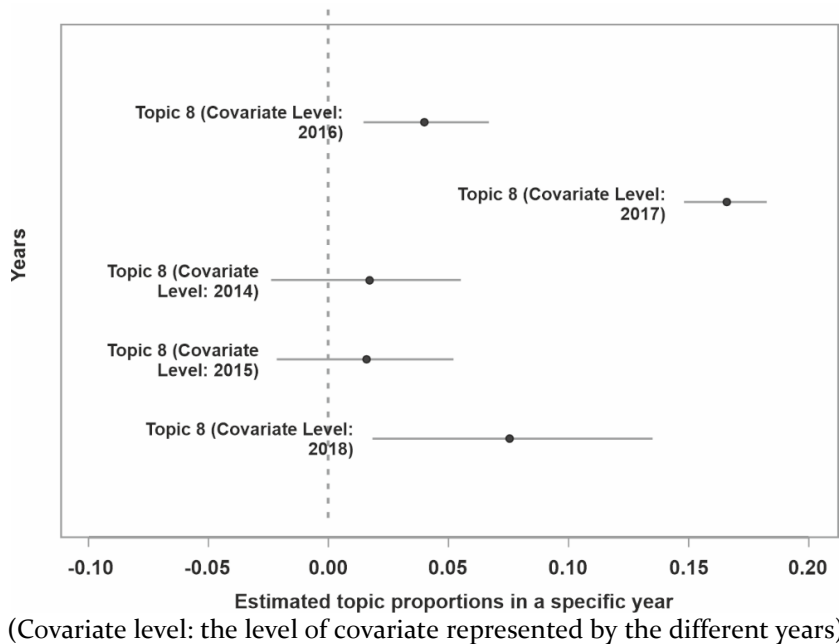
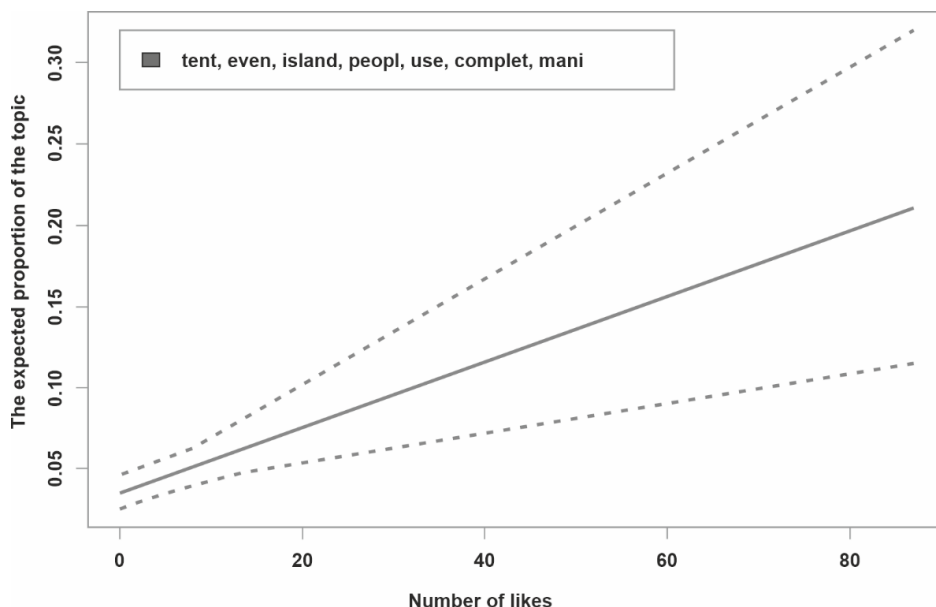Source: own edit using the STM R software package

Figure 5: Effect of discrete covariate (year of comment) on the share of topic 8 (opinions on catering and food hygiene)

The relationship between the number of likes (tags) and the expected share of topics is represented in Figure 6, which shows how the share of the appearance of the topic of camping services in comments changed in accordance with the number of likes. The figure indicates that as the number of likes increases, the share of the topic in the comments increases (or would increase), with widening lower and upper confidence intervals, that is, the uncertainty of the estimate is large. It is the view of the researchers that, this covariate does not have much predictive power. Although there are some topics that generate a higher number of likes, the increase in the number of likes does little to explain the increase in the expected share of a given topic in comments.

7. Discussion

In the opinions of festival visitors, the overall experience of the festival was the most important topic, complemented with a strong emotional charge (Topic 5). This emotional experience appeared in Topic 2 and 7 as well, thus forming the first group of topics. The combined share of these three topics was a significant 45 percent, supporting that the emotional factor is a critical element for music festival attendees (Del Bosque & San Martin, 2008; Jun *et al.*, 2001; Van Dolen *et al.*, 2004). In particular, the research gave further support to the importance of the social aspect (Section 2.3) of music festivals covered in Topic 7 as this was the only element that appeared in the same group with the overall festival experience (Jang *et al.*, 2015; Sun *et al.*, 2019; Iványi, 2021; Marline & Wilders, 2020; Zhang *et al.*, 2019). The other group of topics includes reviews that are perhaps more rational, including festival layout, general and food hygiene, all previously identified as elements of the environment (Section 2.2), entertainment (Section 2.4), and value (Section 2.5). The environmental topics included several negative phrases (for example, horrible, suck, disappoint, problem), supporting the fact that these elements act as hygiene factors that rather have an impact on dissatisfaction than satisfaction (Crompton, 2003). The findings in this study supported Iványi (2021) considering that a variety of programmes and additional cultural programming were sometimes mentioned, but these were not one

of the primary elements; instead, music was the primary factor of entertainment. In terms of value, Sziget Festival is one of the most affordable music festivals of its size in Europe; however, the cost was mostly mentioned in relation to food.

Source: own edit using the stm R software package

Figure 6: Effect of a continuous covariate (number of likes of posts) on the share of Topic 1 (opinions on camping services)

An interesting theme was raised in Topic 3 which covered practical and planning-related questions about the festival. This could be a typical topic of the pre-liminal phase as Wood *et al.* (2020) suggested this phase is centred around the logistics of preparing for the upcoming event. Future research could separate the reviews based on timings into the three phases (pre-liminal, liminal, and post-liminal) to provide further evidence for this theory. This research setup can be based on Fischer-Preßler *et al.* (2019) that tracked the changes in tweets in the days following a terrorist attack. Furthermore, there were seldom mentions of feeling free, the escape from everyday life, and living their authentic selves, despite the core theme of Sziget being the" Festival of Freedom". Additionally, this research supports the finding of Iványi and Bíró-Szigeti (2020) that the elements of Hungarian music festivals are interconnected as several medium and medium to weak relationships have been identified (between Topic 5 and Topic 2, Topic 7 and Topic 2, Topic 6 and Topic 9, and Topic 8 and Topic 4). Based on these findings, future research could investigate other contexts to understand the differences; for example, Mulder and Hitters (2021) found differences between festivals and concerts. This future research may analyse festivals based on size, theme, audience or location.

8. Conclusion

This paper generates several theoretical contributions. First, it provides an overview of existing research on the elements of festival experiences and their interconnected nature. Additionally, the main theoretical finding is in the novel methodology, highlighting that topic modelling is a useful method for the descriptive analysis of unstructured social media datasets, and the results of a topic model can also lead to deeper qualitative analysis (Rodriguez & Stoker, 2019).

Despite the unique nature and above-mentioned strengths, LDA and STM has its limitations, just as any other research method. The key limitations of LDA have been mentioned before, including its "bag of words" and mixed membership model (Blei *et al.*, 2003; Airoldi *et al.*, 2014). Additionally, although the use of social media data allowed a broader analysis of all phases of festival experiences, it fails to clearly differentiate between them. Furthermore, due to the nature of social media comment data, the research is not representative of festival attendees' opinions as people tend to contribute to eWOM when highly satisfied or dissatisfied with the experience. Therefore, it is best practice to use a mixed method design to be able to compare the findings and to see if the findings corroborate (Saunders *et al.*, 2019).

A clear limitation of the study was the size of the corpus, that is, the number of entries written by visitors. From the point of view of computer topic modelling, the aim was to analyse a small corpus using machine methods, but this seems to have been partly at the expense of semantic coherence. No matter how many topics we experimented with during the runs, the quality of the results, i.e. the comprehensibility of the topics generated by the algorithm, improved very little. This limited corpus resulted in a lack of detail and nuances in the results, such as the socialisation element presented as one topic instead of split into either bridging, bonding, and a sense of community or entertainment, help, and conflict. A much larger corpus is likely to be needed for such studies in the future. In this sense, this study can be considered only partially successful, while in other respects the visitors' social media posts can be effectively analysed using computer-based methods, so that a quantitative and cost-effective method can be used to identify the factors that determine the visitors' satisfaction.

The corpus can be expanded in several ways. As this research included data from recent years, it is recommended to increase the corpus horizontally by adding new platforms. Twitter would be a good option because its database is open access, unlike that of Facebook.

Besides the theoretical contributions mentioned above, this paper also provides managerial contributions. The insight into the festival experience elements, their interconnections, and priorities can help managers make future planning decisions. This opportunity is especially useful for managers now as they look to reinvigorate the sector by gaining a more multi-faceted view of the consumers' thoughts and feelings.

A deep understanding of the nuances of festival experiences is essential for the host country as it may derive several benefits from successful events, including economic growth, and these experiences enhance the image of the country (or destination) as well. However, these benefits can only be achieved if festival attendees have a satisfying experience, therefore this has become a popular research topic in academia. Despite recognising its importance, only a few studies have analysed festival experiences and their elements in the context of music festivals. Moreover, there is a need to analyse not just the liminal phase but what happens before and after the festival event as well.

Finally, topic modelling is rarely used in tourism and especially in festival research despite its popularity in other subjects. Topic modelling is a promising method in the research of music festivals as attendees tend to be younger and typically active users of social media platforms that topic modelling applies as data sources. These are the gaps that this paper attempted to address.

References

Adam, I. (2021). Negative tourist-to-tourist interactions, value destruction, satisfaction, and post consumption behavioral intention. *Journal of Destination Marketing & Management, 20,* 100557. https://doi.org/10.1016/j.jdmm.2021.100557

Aguado, L. F., Arbona, A., Palma, L., & Heredia-Carroza, J. (2021). How to value a cultural festival? The case of Petronio Álvarez Pacific Music Festival in Colombia. *Development Studies Research, 8*(1), 309-316. https://doi.org/10.1080/21665095.2021.1979417

Airoldi, E. M., Blei, D. M., Erosheva, E. A., & Fienberg, S. E. (2014). Introduction to Mixed Membership Models and Methods. In E. M. Airoldi, D. M. Blei, E. A. Erosheva, & S. E. Fienberg (Eds)., *Handbook of mixed membership models and their applications* (pp. 3-14). CRC Press. https://doi.org/10.1201/b17520-5

Altunel, M., & Koçak, Ö. (2017). The roles of subjective vitality, involvement, experience quality, and satisfaction in tourists' behavioral intentions. *European Journal of Tourism Research*, 16, 233-251. https://doi.org/10.54055/ejtr.v16i.287

Alvarado, K. G. (2022). Accessibility of music festivals: A British perspective. *International Journal of Event and Festival Management, 13*(2), 203-218. https://doi.org/10.1108/IJEFM-12-2020-0082

Arcodia, C., & Whitford, M. (2007). Festival attendance and the development of social capital. *Journal of Convention & Event Tourism, 8*(2), 1-18. https://doi.org/10.1300/J452v08n02_01

Aşan, K., Kaptangil, K., & Kınay, A. G. (2020). Mediating role of perceived festival value in the relationship between experiences and satisfaction. *International Journal of Event and Festival Management, 11*(2), 255-271. https://doi.org/10.1108/IJEFM-11-2019-0058

Bakas, F. E., Duxbury, N., Remoaldo, P. C., & Matos, O. (2019). The social utility of small-scale art festivals with creative tourism in Portugal. *International Journal of Event and Festival Management, 10*(3), 248-266. https://doi.org/10.1108/IJEFM-02-2019-0009

Baker, D. A., & Crompton, J. L. (2000). Quality, satisfaction and behavioral intentions. *Annals of Tourism Research, 27*(3), 785-804. https://doi.org/10.1016/S0160-7383(99)00108-5

Ballantyne, J., Ballantyne, R., & Packer, J. (2014). Designing and managing music festival experiences to enhance attendees' psychological and social benefits. *Musicae scientiae, 18*(1), 65-83. https://doi.org/10.1177/1029864913511845

Balogh, K. (2015). A látens Dirichl*et al*lokáció társadalomtudományi alkalmazása. A kuruc.info romaellenes megnyilvánulásainak tematikus elemzése. Szakdolgozat ("Applications of Latent Dirichl*et al*location in social science. A thematic analysis of anti-Roma manifestations on kuruc.info. Thesis paper"), *Eötvös Lóránd University, Faculty of Social Sciences, Master of Arts.* Retrieved from https://tas.precognox.com/labs/kuruc-info-visualization/A_latens_Dirichlet_allokacio_tarsadalomtudomanyi_alkalmazasa_Balogh_Kitti.pdf (Accessed on 08.07.2020)

Barrière, L., & Finkel, R. (2022). The material culture of music festival fandoms. *European Journal of Cultural Studies, 25*(2), 479-497. https://doi.org/10.1177/1367549420973211

Bíró, I., Siklósi, D., Szabó, J., & Benczúr, A. A. (2009). Linked latent Dirichl*et al*location in web spam filtering. *Proceedings of the 5th International Workshop on Adversarial Information*. 37-40. https://doi.org/10.1145/1531914.1531922

Bíró, I., & Szabó, J. (2009). Latent dirichl*et al*location for automatic document categorization. In: W. Buntine, M. Grobelnik, D. Mladenić, & J. Shawe-Taylor (Eds.), *Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2009. Lecture Notes in Computer Science, Vol. 5782.* Springer. https://doi.org/10.1007/978-3-642-04174-7_28

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichl*et al*location. *Journal of Machine Learning Research*, 3. 993-1022.

Bohn, D., & de Bernardi, C. (2022). Celebrating 30 years louder than hell: Exploring commercial and social 'Host Event Zone' developments of the heavy metal festival Wacken Open Air. *Annals of Leisure Research, 25*(1), 116-137. https://doi.org/10.1080/11745398.2020.1825972

Bolton, R. N., Parasuraman, A., Hoefnagels, A., Migchels, N., Kabadayi, S., Gruber, T., Loureiro, Y. K., & Solnet, D. (2013). Understanding Generation Y and their use of social media: A review and research agenda. *Journal of Service Management, 24*(3), 245-267. https://doi.org/10.1108/09564231311326987

Borges, A. P., Cunha, C., & Lopes, J. (2021). The main factors that determine the intention to revisit a music festival. *Journal of Policy Research in Tourism, Leisure and Events, 13*(3), 314-335. https://doi.org/10.1080/19407963.2020.1817051

Bowen, H. E., & Daniels, M. J. (2005). Does the music matter? Motivations for attending a music festival. *Event Management, 9*(3), 155-164. https://doi.org/10.3727/152599505774791149

Bronner, F., & de Hoog, R. (2019). Comparing conspicuous consumption across different experiential products: Culture and leisure. *International Journal of Market Research, 61*(4), 430-446. https://doi.org/10.1177/1470785318799898

Bronner, F., & de Hoog, R. (2021). Conspicuous leisure: The social visibility of cultural experiences. *International Journal of Market Research, 63*(3), 300-316. https://doi.org/10.1177/1470785319880715

Brown, A. E., & Sharpley, R. (2019). Understanding festival-goers and their experience at UK Music Festivals. *Event Management. 23*(4-5), 699-720. https://doi.org/10.3727/152599519X15506259855733

Brown, A. E., Donne, K., Fallon, P., & Sharpley, R. (2020). From headliners to hangovers: Digital media communication in the British rock music festival experience. *Tourist Studies, 20*(1), 75-95. https://doi.org/10.1177/1468797619885954

Bruwer, J., & Kelley, K. (2015). Service performance quality evaluation and satisfaction in a USA wine festivalscape. *International Journal of Event and Festival Management, 6*(1), 18-38. https://doi.org/10.1108/IJEFM-04-2014-0009

Carah, N., & Angus, D. (2018). Algorithmic brand culture: participatory labour, machine learning and branding on social media. *Media, Culture & Society, 40*(2), 178-194. https://doi.org/10.1177/0163443718754648

Calheiros, A. C., Moro, S., & Rita, P. (2017). Sentiment classification of consumer-generated online reviews using topic modeling. *Journal of Hospitality Marketing & Management*, *26*(7), 675-693. https://doi.org/10.1080/19368623.2017.1310075

Chang, K. C. (2014). Examining the effect of tour guide performance, tourist trust, tourist satisfaction, and flow experience on tourists' shopping behavior. *Asia Pacific Journal of Tourism Research, 19*(2), 219-247. https://doi.org/10.1080/10941665.2012.739189

Chen, J. V., Htaik, S., Hiele, T. M., & Chen, C. (2017). Investigating international tourists' intention to revisit Myanmar based on need gratification, flow experience and perceived risk. *Journal of Quality Assurance in Hospitality & Tourism, 18*(1), 25-44. https://doi.org/10.1080/1528008X.2015.1133367

Chen, W., Lee, C., & Lin, L. (2012). Investigating factors affecting festival quality: A case study of Neimen Song Jiang Jhen Battle Array, Taiwan. *African Journal of Marketing Management, 4*(2), 43-54. https://doi.org/10.5897/AJMM10.038

Cheng, T. M., & Lu, C. C. (2015). The causal relationships among recreational involvement, flow experience, and well-being for surfing activities. *Asia Pacific Journal of Tourism Research, 20*(sup1), 1486-1504. https://doi.org/10.1080/10941665.2014.999099

Choo, H., Ahn, K., & Petrick, J. F. (2016). An integrated model of festival revisit intentions: Theory of planned behavior and festival quality/satisfaction. *International Journal of Contemporary Hospitality Management, 28*(4), 818-838. https://doi.org/10.1108/IJCHM-09-2014-0448

Christou, P., Sharpley, R., & Farmaki, A. (2018). Exploring the emotional dimension of visitors' satisfaction at cultural events. *Event Management, 22*(2), 255-269. https://doi.org/10.3727/152599518X15173355843389

Cialdini, R. B. (2001). *Influence: Science and practice*. Boston, MA: Allyn & Bacon.

Cianga, I. (2020). Festivals, place-making and local economic development: The untold festival in Cluj. *Journal of Balkan and Near Eastern Studies, 22*(4), 506-520. https://doi.org/10.1080/19448953.2020.1775402

Cochrane, T. (2009). Joint attention to music. *The British Journal of Aesthetics, 49*(1), 59-73. https://doi.org/10.1093/aesthj/ayn059

Cole, S. T., & Chancellor, H. C. (2009). Examining the festival attributes that impact visitor experience, satisfaction and re-visit intention. *Journal of Vacation Marketing, 15*(4), 323-333. https://doi.org/10.1177/1356766709335831

Cole, S. T., & Illum, S. F. (2006). Examining the mediating role of festival visitors' satisfaction in the relationship between service quality and behavioral intentions. *Journal of Vacation Marketing, 12*(2), 160-173. https://doi.org/10.1177/1356766706062156

Crompton, J. L. (2003). Adapting Herzberg: A conceptualization of the effects of Hygiene and motivator attributes on perceptions of event quality. *Journal of Travel Research, 41*(3), 305-10. https://doi.org/10.1177/0047287502239039

Crompton, J. L., Lee, S., & Shuster, T. J. (2001). A guide for undertaking economic impact studies: The Springfest example. *Journal of Travel Research, 40*(1), 79-87. https://doi.org/10.1177/004728750104000110

Crossley, S., Dascalu, M., & Mcnamara, D. (2017). How important is size? An investigation of corpus size and meaning in both Latent Semantic Analysis and Latent Dirichl*et al*location. *FLAIRS 2017 - Proceedings of the 30th International Florida Artificial Intelligence Research Society Conference* (pp. 293-296). AAAI Press.

Dalgiç, A., & Birdir, K. (2020). Key success factors on loyalty of festival visitors: The mediating effect of festival experience and festival image. *Tourism & Management Studies, 16*(1), 28-38. https://doi.org/10.18089/tms.2020.160101

Danielsen, A., & Kjus, Y. (2019). The mediated festival: Live music as trigger of streaming and social media engagement. *Convergence, 25*(4), 714-734. https://doi.org/10.1177/1354856517721808

de Lira, V. M., Macdonald, C., Ounis, I., Perego, R., Renso, C., & Times, V. C. (2019). Event attendance classification in social media. *Information Processing & Management, 56*(3), 687-703. https://doi.org/10.1016/j.ipm.2018.11.001

Deery, M., & Jago, L. (2010). Social impacts of events and the role of anti-social behaviour. International Journal of Event and Festival Management. 1(1), 8-28. https://doi.org/10.1108/17852951011029289

Del Bosque, I. R., & San Martín, H. (2008). Tourist satisfaction a cognitive-affective model. *Annals of Tourism Research, 35*(2), 551-573. https://doi.org/10.1016/j.annals.2008.02.006

Ding, H. M., & Hung, K. P. (2021). The antecedents of visitors' flow experience and its influence on memory and behavioral intentions in the music festival context. *Journal of Destination Marketing & Management, 19*, 100551. https://doi.org/10.1016/j.jdmm.2020.100551

Duarte, P., Folgado-Fernández, J. A., & Hernández-Mogollón, J. M. (2018). Measurement of the impact of music festivals on destination image: The case of a womad festival. *Event Management, 22*(4), 517-526. https://doi.org/10.3727/152599518X15264726192451

Duffy, M., Waitt, G., Gorman-Murray, A., & Gibson, C. (2011). Bodily rhythms: Corporeal capacities to engage with festival spaces. *Emotion, Space and Society,* 4(1), 17-24. https://doi.org/10.1016/j.emospa.2010.03.004

Eisenstein, J., Ahmed, A., & Xing, E. P. (2011). Sparse additive generative models of text. *Proceedings of the 28th International Conference on Machine Learning* (pp. 1041–1048).

Embley, J. (2022). European music festivals 2022: The best events to travel to, from Primavera Sound to Sziget. *Evening Standard.* Retrieved from https://www.standard.co.uk/culture/music/european-music-festivals-2022-summer-b983146.html (Accessed on 12.07.2022)

Felsenstein, D., & Fleischer, A. (2003). Local festivals and tourism promotion: The role of public assistance and visitor expenditure. *Journal of Travel Research, 41*(4), 385-392. https://doi.org/10.1177/0047287503041004007

Fischer-Preßler, D., Schwemmer, C., & Fischbach, K. (2019). Collective sense-making in times of crisis: Connecting terror management theory with Twitter user reactions to the Berlin terrorist attack. *Computers in Human Behavior, 100,* 138-151. https://doi.org/10.1016/j.chb.2019.05.012

Fu, X., Kang, J., & Tasci, A. (2017). Self-congruity and flow as antecedents of attitude and loyalty towards a theme park brand. *Journal of Travel & Tourism Marketing, 34*(9), 1261-1273. https://doi.org/10.1080/10548408.2017.1343704

Gamage, A., & Higgs, B. (1996). Economics of venue selection for special sporting events: With special reference to the 1996 Melbourne Grand Prix. *Asia Pacific Journal of Tourism Research, 1*(2), 15-26. https://doi.org/10.1080/10941669708721972

Gao, B., Zhu, M., Liu, S., & Jiang, M. (2022). Different voices between Airbnb and hotel customers: An integrated analysis of online reviews using structural topic model. *Journal of Hospitality and Tourism Management, 51*, 119-131. https://doi.org/10.1016/j.jhtm.2022.03.004

Gerrish, S. & Blei, D. M. (2012). How they vote: Issue-adjusted models of legislative behavior. *Advances In Neural Information Processing Systems 25. (NPIS 2012),* 2753-2761.

Geus, S. D., Richards, G., & Toepoel, V. (2016). Conceptualisation and operationalisation of event and festival experiences: Creation of an event experience scale. *Scandinavian Journal of Hospitality and Tourism, 16*(3), 274-296. https://doi.org/10.1080/15022250.2015.1101933

Giaccone, S., & Galvagno, M. (2021). Small-scale cultural festivals in Southern Europe: A motivation-based segmentation study. *European Journal of Tourism Research, 29,* 2902. https://doi.org/10.54055/ejtr.v29i.2413

Gilstrap, C., Teggart, A., Cabodi, K., Hills, J., & Price, S. (2021). Social music festival brandscapes: A lexical analysis of music festival social conversations. *Journal of Destination Marketing & Management, 20*, 100567. https://doi.org/10.1016/j.jdmm.2021.100567

Guo, Y., Barnes, S. J., & Jia, Q. (2017). Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichl*et al*location. *Tourism Management*, *59*, 467-483. https://doi.org/10.1016/j.tourman.2016.09.009

Gursoy, D., Kim, K., & Uysal, M. (2004). Perceived impacts of festivals and special events by organizers: An extension and validation. *Tourism management, 25*(2), 171-181. https://doi.org/10.1016/S0261-5177(03)00092-X

Hinek, M., & Kulcsár, N. (2019). A fesztiválélmény megjelenése a közösségi médiában: a Sziget Fesztivál példája ("The festival experience as represented in social media: the example of the Sziget Festival"). *Turizmus Bulletin, 19(3),* 4–12. https://doi.org/10.14267/TURBULL.2019v19n3.1

Heineken (n.d). Tisztul és nő a fesztiválpiac ("The festival market is clearing and growing"). Retrieved from https://www.heinekenhungaria.hu/sajtoszoba/sajtokozlemenyek/149-tisztul-es-no-a-fesztivalpiac.html (Accessed on 05.07.2022)

Heinonen, K., Jaakkola, E., & Neganova, I. (2018). Drivers, types and value outcomes of customer-to-customer interaction: An integrative review and research agenda. *Journal of Service Theory and Practice. 28*(6), 710-732. https://doi.org/10.1108/JSTP-01-2017-0010

Huang, J., & Hsu, C. H. (2010). The impact of customer-to-customer interaction on cruise experience and vacation satisfaction. *Journal of Travel Research, 49*(1), 79-92. https://doi.org/10.1177/0047287509336466

Hudson, S., & Hudson, R. (2013). Engaging with consumers using social media: A case study of music festivals. International *Journal of Event and Festival Management, 4*(3), 206-223. https://doi.org/10.1108/IJEFM-06-2013-0012

Hudson, S., Roth, M. S., Madden, T. J., & Hudson, R. (2015). The effects of social media on emotions, brand relationship quality, and word of mouth: An empirical study of music festival attendees. *Tourism Management, 47*, 68-76. https://doi.org/10.1016/j.tourman.2014.09.001

Hu, N., Zhang, T., Gao, B., & Bose, I. (2019). What do hotel customers complain about? Text analysis using structural topic model. *Tourism Management*, *72*, 417-426. https://doi.org/10.1016/j.tourman.2019.01.002

Ismagilova, E., Dwivedi, Y. K., Slade, E., & Williams, M. D. (2017). *Electronic word of mouth (eWOM) in the marketing context: A state of the art analysis and future directions.* Cham: Springer International Publishing.

Iványi, T. (2021). Analysis of Sources for Information Gathering before Visiting a Festival in the Case of Generation Z. *Periodica Polytechnica Social and Management Sciences, 29*(2), 196-208. https://doi.org/10.3311/PPso.15947

Iványi, T., & Bíró-Szigeti, S. (2020). Understanding internal connections of music festivals'experience dimensions. *Tourism and Hospitality Management, 26*(2), 437-454. https://doi.org/10.20867/thm.26.2.9

Jang, Y., Ro, H., & Kim, T. H. (2015). Social servicescape: The impact of social factors on restaurant image and behavioral intentions. *International Journal of Hospitality & Tourism Administration, 16*(3), 290-309. https://doi.org/10.1080/15256480.2015.1054758

Jun, S., Hyun, Y., Gentry, J., & Chang-Seok, S. (2001). The relative influence of affective experience on consumer satisfaction under positive versus negative discrepancies. *The Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior, 14,* 141-153. https://doi.org/10.1016/S0148-2963(02)00277-1

Kahneman, D. (2011). *Thinking, fast and slow.* Macmillan.

Kay Smith, M., Pinke-Sziva, I., Berezvai, Z., & Buczkowska-Gołąbek, K. (2022). The changing nature of the cultural tourist: Motivations, profiles and experiences of cultural tourists in Budapest. *Journal of Tourism and Cultural Change, 20*(1-2), 1-19. https://doi.org/10.1080/14766825.2021.1898626

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words?. *Memory & Cognition, 31*(8), 1169-1180. https://doi.org/10.3758/BF03195800

Kim, B. (2015). What facilitates a festival tourist? Investigating tourists' experiences at a local community festival. *Asia Pacific Journal of Tourism Research, 20*(9), 1005-1020. https://doi.org/10.1080/10941665.2014.957222

Kim, K., (A.), Byon, K. K., & Baek, W. (2020). Customer-to-customer value cocreation and co-destruction in sporting events. *Service Industries Journal, 40*(9-10), 633-655. https://doi.org/10.1080/02642069.2019.1586887

Kim, M., & Thapa, B. (2018). Perceived value and flow experience: Application in a nature-based tourism context. *Journal of Destination Marketing & Management, 8,* 373-384. https://doi.org/10.1016/j.jdmm.2017.08.002

Kim, S. S., & Petrick, J. F. (2005). Residents' perceptions on impacts of the FIFA 2002 World Cup: The case of Seoul as a host city. *Tourism Management, 26*(1), 25-38. https://doi.org/10.1016/j.tourman.2003.09.013

Kim, S., Lee, Y.-K., & Lee, C.-K. (2017), The moderating effect of place attachment on the relationship between festival quality and behavioral intentions. *Asia Pacific Journal of Tourism Research, 22*(1), 49-63. https://doi.org/10.1080/10941665.2016.1176060

Kinnunen, M., Homi, H., & Honkanen, A. (2020). Social sustainability in adolescents' music event attendance. *Sustainability, 12*(22), 9419. https://doi.org/10.3390/su12229419

Kirilenko, A. P., Stepchenkova, S. O., & Dai, X. (2021). Automated topic modeling of tourist reviews: Does the Anna Karenina principle apply?. *Tourism Management, 83,* 104241. https://doi.org/10.1016/j.tourman.2020.104241

Koç, B., Küçükergin, K. G., & Dimanche, F. (2022). How destructive are negative tourist-to-tourist interactions despite the mitigating effect of optimism?. *Journal of Destination Marketing & Management, 23,* 100693. https://doi.org/10.1016/j.jdmm.2022.100693

Korfiatis, N., Stamolampros, P., Kourouthanassis, P., & Sagiadinos, V. (2019). Measuring service quality from unstructured data: A topic modeling application on airline passengers' online reviews. *Expert Systems with Applications, 116,* 472-486. https://doi.org/10.1016/j.eswa.2018.09.037

Krestel, R., Fankhauser, P., & Nejdl, W. (2009). Latent dirichl*et al*location for tag recommendation. *Proceedings of the third ACM conference on Recommender systems.* 61-68. https://doi.org/10.1145/1639714.1639726

Kwiatkowski, G., Oklevik, O., Hjalager, A. M., & Maristuen, H. (2020). The assemblers of rural festivals: Organizers, visitors and locals. *European Planning Studies, 28*(2), 255-272. https://doi.org/10.1080/09654313.2019.1651829

Lafferty, J. D. & Blei, D. M. (2006). Correlated topic models. *Advances In Neural Information Processing Systems,* 147-154.

Ledesma, J. (2020). Why do people write reviews? What our research revealed. Retrieved from https://business.trustpilot.com/reviews/learn-from-customers/why-do-people-write-reviews-what-our-research-revealed (Accessed on 31.07.2022)

Lee, K-J. (2017). Cultural capital of visitors, classical music festival satisfaction, and quality of life. *Global Business & Finance Review, 22*(4), 1-14. https://doi.org/10.17549/gbfr.2017.22.4.1

Lee, Y., Lee, C., Lee, S., & Babin, B. J. (2008). Festivalscapes and patrons' emotions, satisfaction, and loyalty. *Journal of Business Research, 61*(1), 56-64. https://doi.org/10.1016/j.jbusres.2006.05.009

Leenders, M. A. (2010). The relative importance of the brand of music festivals: A customer equity perspective. *Journal of Strategic Marketing, 18*(4), 291-301. https://doi.org/10.1080/09652541003768061

Levy, S. E. (2010). The hospitality of the host: A cross-cultural examination of managerially facilitated consumer-to-consumer interactions. *International Journal of Hospitality Management, 29*(2), 319-327. https://doi.org/10.1016/j.ijhm.2009.10.003

Levy, S. E., Getz, D., & Hudson, S. (2011). A field experimental investigation of managerially facilitated consumer-to-consumer interaction. *Journal of Travel & Tourism Marketing, 28*(6), 656-674. https://doi.org/10.1080/10548408.2011.603633

Lin, H., Zhang, M., Gursoy, D., & Fu, X. (2019). Impact of tourist-to-tourist interaction on tourism experience: The mediating role of cohesion and intimacy. *Annals of Tourism Research, 76*, 153-167. https://doi.org/10.1016/j.annals.2019.03.009

Litvin, S. W., Goldsmith, R. E., & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism management, 29*(3), 458-468. https://doi.org/10.1016/j.tourman.2007.05.011

Long, P. T., & Perdue, R. R. (1990). The economic impact of rural festivals and special events: Assessing the spatial distribution of expenditures. *Journal of Travel Research, 28*(4), 10-14. https://doi.org/10.1177/004728759002800403

MacDonald, R. A. (2013). Music, health, and well-being: A review. *International Journal of Qualitative Studies on Health and Well-being, 8*(1), 20635. https://doi.org/10.3402/qhw.v8i0.20635

Manthiou, A., Lee, S., Tang, L., & Chiang, L. (2014). The experience economy approach to festival marketing: Vivid memory and attendee loyalty. *Journal of Services Marketing, 28*(1), 22-35. https://doi.org/10.1108/JSM-06-2012-0105

Menegaki, A., Tiwari, A., & Agiomirgianakis, G. (2020). Asymmetries in European inbound and outbound tourism: Normal, luxury or inferior good? Fresh evidence from a quantile regression. *European Journal of Tourism Research, 25*, 2508-2508. https://doi.org/10.54055/ejtr.v25i.423

Moisescu, O. I., Gică, O. A., Coroş, M. M., & Yallop, A. C. (2019). The UNTOLD story: Event tourism's negative impact on residents' community life and well-being. *Worldwide Hospitality and Tourism Themes, 11*(5), 492-505. https://doi.org/10.1108/WHATT-06-2019-0036

Morgan, M. (2008). What makes a good festival? Understanding the event experience. *Event Management, 12*(2), 81-93. https://doi.org/10.3727/152599509787992562

Moscardo, G. (2007). Analyzing the role of festivals and events in regional development. *Event Management, 11*(1-2), 23-32. https://doi.org/10.3727/152599508783943255

Mulder, M., & Hitters, E. (2021). Visiting pop concerts and festivals: Measuring the value of an integrated live music motivation scale. *Cultural Trends, 30*(4), 355-375, https://doi.org/10.1080/09548963.2021.1916738

Murphy, L. (2001). Exploring social interactions of backpackers. *Annals of Tourism Research, 28*(1), 50-67. https://doi.org/10.1016/S0160-7383(00)00003-7

Negruşa, A. L., Toader, V., Rus, R. V., & Cosma, S. A. (2016). Study of perceptions on cultural events' sustainability. *Sustainability, 8*(12), 1269. https://doi.org/10.3390/su8121269

Niloofar, I., Meysam, F., & Sara, F. (2012). The effect of customer-to-customer interactions on satisfaction with the firm, loyalty to the firm and firm word-of-mouth: The case of Iran Air Company. *African Journal of Business Management, 6*(39), 10427-10437. https://doi.org/10.5897/AJBM12.706

Nordvall, A., Pettersson, R., Svensson, B., & Brown, S. (2014). Designing events for social interaction. *Event Management, 18*(2), 127-140. https://doi.org/10.3727/152599514X13947236947383

Oh, H., Fiore, A.M., & Jeoung, M. (2007). Measuring experience economy concepts: Tourism applications. *Journal of Travel Research, 46*(2), 119-132. https://doi.org/10.1177/0047287507304039

Oklevik, O., Kwiatkowski, G., Preuss, H., & Kurdyś-Kujawska, A. (2022). Contextual engagement in event visitors' experience and satisfaction. *Scandinavian Journal of Hospitality and Tourism, 22*(1), 58-79. https://doi.org/10.1080/15022250.2021.1927829

Özdemir, G., & Çulha, O. (2009). Satisfaction and loyalty of festival visitors. *Anatolia, 20*(2), 359-373. https://doi.org/10.1080/13032917.2009.10518914

Papathanassis, A. (2012). Guest-to-guest interaction on board cruise ships: Exploring social dynamics and the role of situational factors. *Tourism Management, 33*(5), 1148-1158. https://doi.org/10.1016/j.tourman.2011.11.016

Papinczak, Z. E., Dingle, G. A., Stoyanov, S. R., Hides, L., & Zelenko, O. (2015). Young people's uses of music for well-being. *Journal of Youth Studies, 18*(9), 1119-1134. https://doi.org/10.1080/13676261.2015.1020935

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer percepions of service quality. *Journal of Retailing, 64*(1), 12-40.

Park, C., & Lee, T. M. (2009). Information direction, website reputation and eWOM effect: A moderating role of product type. *Journal of Business Research, 62*(1), 61-67. https://doi.org/10.1016/j.jbusres.2007.11.017

Park, H.Y. (2016). Tourism as reflexive reconstructions of colonial past. *Annals of Tourism Research, 58*, 114-127. https://doi.org/10.1016/j.annals.2016.03.001

Park, K., & Ha, S. H. (2017). Customer Service Evaluation based on Online Text Analytics: Sentiment Analysis and Structural Topic Modeling. *The Journal of Information Systems, 26*(4), 327-353. https://doi.org/10.5859/KAIS.2017.26.4.327

Paul, M. J. & Dredze, M. (2015). SPRITE: Generalizing topic models with structured priors. *Transactions of the Association for Computational Linguistics, 3,* 43-57.

Pavluković, V., Armenski, T., & Alcántara-Pilar, J. M. (2017). Social impacts of music festivals: Does culture impact locals' attitude toward events in Serbia and Hungary?. *Tourism Management, 63*, 42-53. https://doi.org/10.1016/j.tourman.2017.06.006

Pine, B. J., & Gilmore, J. H. (1998). Welcome to the experience economy. *Harvard Business Review*, Vol. 76, No. 6, 97-105.

Pivac, T., Blešić, I., Kovačić, S., Besermenji, S., & Lesjak, M. (2019). Visitors'satisfaction, perceived quality, and behavioral intentions: The case study of EXIT festival. *Journal of the Geographical Institute "Jovan Cvijić" SASA, 69(2)*, 123-134. https://doi.org/10.2298/IJGI1902123P

Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community.* New York: Simon & Schuster.

R Core Team (2020). *R: A language and environment for statistical computing, R Foundation for Statistical Computing.* Vienna, Austria. Retrieved from https://www.R-project.org/ (Accessed on 20.07.2022)

Rasoolimanesh, S. M., Jaafar, M., Marzuki, A., & Abdullah, S. (2019). Examining the effects of personal factors and travel characteristics on tourists' perceived crowding. *European Journal of Tourism Research 22*, 5-20. https://doi.org/10.54055/ejtr.v22i.371

Rihova, I., Buhalis, D., Gouthro, M. B., & Moital, M. (2018). Customer-to-customer co-creation practices in tourism: Lessons from customer-dominant logic. *Tourism Management, 67*, 362-375. https://doi.org/10.1016/j.tourman.2018.02.010

Rihova, I., Buhalis, D., Moital, M., & Gouthro, M. B. (2015). Social constructions of value: Marketing considerations for the context of event and festival visitation. In O. T. Moufakkir & T. Pernecky (Eds.), *Ideological, social and cultural aspects of events* (pp. 4-85). Wallingford, UK: CABI International.

Rivera, M., Semrad, K., & Croes, R. (2015). The five E's in festival experience in the context of Gen Y: Evidence from a small island destination. *Revista Española de Investigación de Marketing ESIC, 19*(2), 95-106. http://doi.org/10.1016/j.reimke.2015.06.001

Roberts, M. E., Stewart, B. M., & Tingley, D. (2014). stm: R package for structural topic models. *Journal of Statistical Software, 10*(2). 1-40. https://doi.org/10.18637/jss.v091.i02

Rodriguez, M. Y. & Storer, H. (2020). A computational social science perspective on qualitative data exploration: Using topic models for the descriptive analysis of social media data. *Journal of Technology in Human Services, 38*(1), 54-86. https://doi.org/10.1080/15228835.2019.1616350

Saragih, H. S., & Amelia, N. (2020). Segmentation of music festival visitors by values of hedonia, life satisfaction and eudaimonia. *International Journal of Event and Festival Management, 11*(4), 453-472. https://doi.org/10.1108/IJEFM-03-2020-0016

Saunders, M., Lewis, P., & Thornhill, A. (2019). Research methods for business students (5th edition). Harlow, UK: Pearson Education Limited.

Semrad, K. J., & Rivera, M. (2018). Advancing the 5E's in festival experience for the Gen Y framework in the context of eWOM. *Journal of Destination Marketing & Management, 7*, 58-67. https://doi.org/10.1016/j.jdmm.2016.08.003

Sheth, J. (2020). Impact of Covid-19 on consumer behavior: Will the old habits return or die?. *Journal of Business Research, 117*, 280-283.

Small, K. (2007). Social dimensions of community festivals: An application of factor analysis in the development of the social impact perception (SIP) scale. *Event Management, 11*(1-2), 45-55. https://doi.org/10.3727/152599508783943219

Smith, M. K., Jancsik, A., & Puczkó, L. (2020). Customer satisfaction in post-socialist Spas: a case study of Budapest, City of Spas. *International Journal of Spa and Wellness*, *3*(2-3), 165-186. https://doi.org/10.1080/24721735.2020.1866330

Sökmen, S., & Yolal, M. (2022). Tourist-to-Tourist Interaction (TTI): A social distance perspective. In D. Gursoy & S. Çelik (Eds). *Routledge Handbook of Social Psychology of Tourism*. (pp. 216-227). London, UK: Routledge.

Strand, A., & Robertson, M. (2020). An attitudinal impacts analysis of social media platforms and brand relationship quality at music festivals. *Event Management, 24*(6), 769-788. https://doi.org/10.3727/152599520X15894679115538

Sulyok J., & Sziva I. (2009). A fesztiválturizmus nemzetközi és hazai tendenciái. (International and local trends in festival tourism). *Turizmus Bulletin, 13*(3), 3-13. (in Hungarian)

Sun, H., Wu, S., Li, Y., & Dai, G. (2019). Tourist-to-tourist interaction at festivals: A grounded theory approach. *Sustainability, 11*(15), 4030. https://doi.org/10.3390/su11154030

Sutherland, I., Sim, Y., Lee, S. K., Byun, J., & Kiatkawsin, K. (2020). Topic modeling of online accommodation reviews via latent dirichl*et al*location. *Sustainability*, *12*(5), 1821. https://doi.org/10.3390/su12051821

Süli, D., & Martyin-Csamangó, Z. (2017). A közösségi média szerepe a fesztiválturizmusban - a Szegedi Ifjúsági Napok és a Green Future példáján. (The role of social media in the case of festival tourism -

The example of Youth Days in Szeged and Green Future). *Turizmus Bulletin, 17*(3-4), 24-32. (in Hungarian)

Syed, S. & Spruit, M. (2017). Full-text or abstract? Examining topic coherence scores using latent dirichl*et al*location. In *2017 IEEE International conference on data science and advanced analytics (DSAA)* (pp. 165-174). IEEE. https://doi.org/10.1109/DSAA.2017.61

Sziget Festival (n.d.). Retrieved from https://szigetfestival.com/en/ (Accessed on 12.07.2022)

Szmigin, I., Bengry-Howell, A., Morey, Y., Griffin, C., & Riley, S. (2017). Socio-spatial authenticity at co-created music festivals. *Annals of Tourism Research, 63,* 1-11. https://doi.org/10.1016/j.annals.2016.12.007

Tan, K. L., Sim, A. K., Chai, D., & Beck, L. (2020). Participant well-being and local festivals: The case of the Miri country music festival, Malaysia. *International Journal of Event and Festival Management, 11*(4), 433-451. https://doi.org/10.1108/IJEFM-02-2020-0007

Tanford, S., & Jung, S. (2017). Festival attributes and perceptions: A meta-analysis of relationships with satisfaction and loyalty. *Tourism Management, 61,* 209-20. https://doi.org/10.1016/j.tourman.2017.02.005

Tohmo, T. (2005). Economic impacts of cultural events on local economies: An input—output analysis of the Kaustinen folk music festival. *Tourism Economics, 11*(3), 431-451. https://doi.org/10.5367/000000005774352980

Tomazelli, J., Broilo, P. L., Espartel, L. B., & Basso, K. (2017). The effects of store environment elements on customer-to-customer interactions involving older shoppers. *Journal of Services Marketing, 31*(4-5), 339-350. https://doi.org/10.1108/JSM-05-2016-0200

Turner, A. (2015). Generation Z: Technology and social interest. *The Journal of Individual Psychology, 71*(2), 103-113. https://doi.org/10.1353/jip.2015.0021

Sándor, T. (2022). Sűrű lesz a nyár a hazai fesztiválpiacon, így tér vissza a Sziget ("The summer will be a busy one for the Hungarian festival market, this is how Sziget Festival returns."). *VG.hu.* Retrieved from https://www.vg.hu/vilaggazdasag-magyar-gazdasag/2022/04/suru-lesz-a-nyar-a-hazai-fesztivalpiacon-igy-ter-vissza-a-sziget (Accessed on 05.07.2022)

UK Festival Awards (n.d.). Winners 2019. Retrieved from https://www.festivalawards.com/winners-2019/ (Accessed on 12.07.2022)

Van Dolen, W., De Ruyter, K., & Lemmink, J. (2004). An empirical assessment of the influence of customer emotions and contact employee performance on encounter and relationship satisfaction. *Journal of Business Research, 57*(4), 437-444.

Vecco, M., & Srakar, A. (2017). Blue notes: Slovenian jazz festivals and their contribution to the economic resilience of the host cities. *European Planning Studies, 25*(1), 107-126. https://doi.org/10.1080/09654313.2016.1272548

Venaik, S., & Brewer, P. (2016). National culture dimensions: The perpetuation of cultural ignorance. *Management Learning, 47*(5), 563-589. https://doi.org/10.1177/1350507616629356

Vinnicombe, T., & Sou, P. U. J. (2017). Socialization or genre appreciation: The motives of music festival participants. *International Journal of Event and Festival Management, 8*(3), 274-291. https://doi.org/10.1108/IJEFM-05-2016-0034

Wang, Y. C. – Burke, M. – Kraut, R. E. (2013). Gender, topic, and audience response: an analysis of user-generated content on Facebook. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 31-34). https://doi.org/10.1145/2470654.2470659

Weinshall, D., Levi, G., & Hanukaev, D. (2013). LDA topic model with soft assignment of descriptors to words. *International Conference on Machine Learning, 28*(3), 711-719.

Welthagen, L., & Lötter, M. J. (2020). Comparing the satisfaction level of Park Acoustics Live Music Festival attendees'. *African Journal of Hospitality, Tourism and Leisure, 9*(4). https://doi.org/10.46222/ajhtl.19770720-xxx

Wen-Chiang, C., Cheng-Fei, L., & Ling-Zhong, L. (2012). Investigating factors affecting festival quality: A case study of neimen song jiang jhen battle array, Taiwan. *African Journal of Marketing Management, 4*(2), 43-54. https://doi.org/10.5897/AJMM10.038

Wilders, M. L., & Rusch, L. (2020). Moving the audience: Dutch landscape experienced through the SummerJazzCycleTour, *International Journal of Heritage Studies, 26*(6), 558-570. https://doi.org/10.1080/13527258.2019.1583271

Wilks, L. (2011). Bridging and bonding: Social capital at music festivals. *Journal of Policy Research in Tourism, Leisure and Events, 3*(3), 281-297. https://doi.org/10.1080/19407963.2011.576870

Wilks, L., & Quinn, B. (2016). Linking social capital, cultural capital and heterotopia at the folk festival. *Journal of Comparative Research in Anthropology & Sociology, 7*(1), 23-39.

Wilson, A., & Chew, P. A. (2010). Term weighting schemes for latent dirichl*et al*location. In *human language technologies: The 2010 annual conference of the North American Chapter of the Association for Computational Linguistics* (pp. 465-473) Association for Computational Linguistics.

Wood, E. H., & Moss, J. (2015). Capturing emotions: Experience sampling at live music events. *Arts and the Market, 5*(1), 45-72. https://doi.org/10.1108/AM-02-2013-0002

Wood, E., & Kinnunen, M. (2020). Emotion, memory and re-collective value: Shared festival experiences. *International Journal of Contemporary Hospitality Management, 32*(3), 1275-1298. https://doi.org/10.1108/IJCHM-05-2019-0488

Wood, E., Li, Y. N., & Senaux, B. (2020).). Liminality and festivals - Insights from the East. *Annals of Tourism Research, 80*, 102810. https://doi.org/10.1016/j.annals.2019.102810

Wu, C. H. J., & Liang, R. D. (2011). The relationship between white-water rafting experience formation and customer reaction: A flow theory perspective. *Tourism Management, 32*(2), 317-325. https://doi.org/10.1016/j.tourman.2010.03.001

Yolal, M., Gursoy, D., Uysal, M., Kim, H. L., & Karacaoğlu, S. (2016). Impacts of festivals and events on residents' well-being. *Annals of Tourism Research, 61*, 1-18. https://doi.org/10.1016/j.annals.2016.07.008

Yoon, Y.-S., Lee, J.-S., & Lee, C.-K. (2010). Measuring festival quality and value affecting visitors' satisfaction and loyalty using a structural approach. *International Journal of Hospitality Management, 29*(2), 335-42. https://doi.org/10.1016/j.ijhm.2009.10.002

Zajonc, R. B. (1965). Social facilitation: A solution is suggested for an old social psychological problem. *Science, 149*(1), 269-274. https://doi.org/10.1126/science.149.3681.269

Zgolli, S., & Zaiem, I. (2017). Customer-to-customer interaction in tourism experience: Moderating role of nationality. *Arab Economic and Business Journal, 12*(1), 44-56. https://doi.org/10.1016/j.aebj.2017.03.001

Zhang, C. X., Fong, L. H. N., & Li, S. (2019). Co-creation experience and place attachment: Festival evaluation. *International Journal of Hospitality Management, 81*, 193-204. https://doi.org/10.1016/j.ijhm.2019.04.013

Zhang, S. N., Li, Y. Q., Liu, C. H., & Ruan, W. Q. (2019). How does authenticity enhance flow experience through perceived value and involvement: The moderating roles of innovation and cultural identity. *Journal of Travel & Tourism Marketing, 36*(6), 710-728. https://doi.org/10.1080/10548408.2019.1625846