

A Study on the Sentiment Analysis of City Tour Using Big Data

Se-won Jeon*, Gi-Hwan Ryu**

**Ph. D Student, Department of Immersive Content Convergence, Kwangwoon University, Seoul, Korea*

***Professor, Department of Tourism and Food Industry, Graduate School of Smart Convergence, Kwangwoon University, Seoul, Korea*

E-mail: {sw2947, allryu}@kw.ac.kr

Abstract

This study aims to find out what tourists' interests and perceptions are like through online big data. Big data for a total of five years from 2018 to 2022 were collected using the Textom program. Sentiment analysis was performed with the collected data. Sentiment analysis expresses the necessity and emotions of city tours in online reviews written by tourists using city tours. The purpose of this study is to extract and analyze keywords representing satisfaction. The sentiment analysis program provided by the big data analysis platform "TEXTOM" was used to study positives and negatives based on sentiment analysis of tourists' online reviews. Sentiment analysis was conducted by collecting reviews related to the city tour. The degree of positive and negative emotions for the city tour was investigated and what emotional words were analyzed for each item. As a result of big data sentiment analysis to examine the emotions and sentiments of tourists about the city tour, 93.8% positive and 6.2% negative, indicating that more than half of the tourists are positively aware. This paper collects tourists' opinions based on the analyzed sentiment analysis, understands the quality characteristics of city tours based on the analysis using the collected data, and sentiment analysis provides important information to the city tour platform for each region.

Keywords: *Big Data, City Tour, City Tour Bus, Tourism, Sentiment analysis, Text mining*

1. Introduction

City Tour operates a total of 303 routes in 75 local governments across the country[1]. Currently, each local government develops various courses by theme and route, introduces the city's representative resources to tourists efficiently and in a short time, and helps tourists to experience various courses. The number of domestic and foreign city tour users is continuously increasing due to the combination of routes between neighboring municipalities and the expansion of stopovers. Accordingly, the opinions of tourists who have used the city tour are collected on the city tour platform for each region, and the quality characteristics of the city tour are understood based on the analyzed sentiment analysis. Furthermore, it is necessary to establish tourists' opinions

Manuscript Received: March. 7, 2023 / Revised: March. 11, 2023 / Accepted: March. 13, 2023

Corresponding Author: allryu@kw.ac.kr

Tel: *** - **** - **** Fax: +82-2-940-5443

Professor, Department of Tourism and Food Industry, Graduate School of Smart Convergence, Kwangwoon University, Korea

on the platform so that they can develop or even recommend courses for each season. The purpose of this study is to revitalize the use of the city tour by analyzing the tourists' comments on the use of the city tour, adding tourist attractions of the preferred course and supplementing and developing the negative course.

The composition of this thesis consists of Chapter 2, related research related to sentiment analysis, and Chapter 3 describes the research results visualized with data collected through Textom. Chapter 4 describes sentiment analysis and Chapter 5 concludes.

2. Related Work

A city tour is a city tour where you set a certain area to visit famous landmarks or historical sites. You can walk in your own car, but you mostly take a large bus to see all over the city. It refers to a city circular tour that is programmed to tour various cultural sites, tourist attractions, market sites, etc. City tour is a form of cultural tourism that promotes the development of tourism products and tourism activities using various city tourism resources in order to promote tourism convenience for people who tour the city and to activate tourism in the city. It is a key means of city tourism and an alternative that can actively respond to the travel needs of individual travelers [3,4].

2.1 Sentiment Analysis

Online reviews not only have a great influence on consumers' purchase decisions, but also play an important role in determining the success or failure of companies in the market[5]. As for online reviews, studies related to sentiment analysis, which is one of the text mining techniques for extracting users' opinions, attitudes, and emotions, are being actively conducted. Sentiment analysis techniques are broadly classified into dictionary-based techniques and machine learning techniques. The dictionary-based technique is a technique in which a researcher builds an emotion dictionary composed of emotional vocabulary and polarity representing the degree of positive and negative of the vocabulary, and performs emotion using this emotion dictionary[6]. Sentiment analysis is a method of analyzing the emotions of the person who wrote the text on social media. Sentiment analysis is a technique that takes advantage of the propensity of people to express their opinions more honestly in environments where they are not confined to the eyes of others, such as informal or anonymous situations. That is, by extracting the frequency of words that reveal emotions such as positive/negative in documents containing keywords related to a specific topic on social media, it is possible to grasp public opinion about how a specific topic or issue is viewed. In this way, sentiment analysis can use big data data to more effectively grasp the actual opinions or preferences of an unspecified number of people who use it, so in this study, sentiment analysis was conducted on the main keywords of the city tour[7,8].

3. Semantic Network Analysis

This study set a period of several years from January 1, 2018 to June 30, 2022 using textom, and collected data containing the keyword 'city tour' from Naver, Daum, and Google. In relation to the 'City Tour', information in an organized form was collected from text data left by tourists using social media or the web, and the pattern of each keyword was identified to visually express and understand the meaning. Daum and Naver collected news, blogs, cafes, intellectuals, and the web, and Google analyzed the contents collected from news, cafes, and the web. A total of 9,334 posts were collected, and data preprocessing was performed to remove redundant posts and delete meaningless parts of speech such as postposition and numeral. In order to identify the main keywords for the city tour, matrix data was created by calculating word frequencies and TF-IDF using text mining techniques during data analysis. Network analysis, centrality analysis and CONCOR

analysis were performed using UCINET. Big data analysis was conducted to identify consumer perceptions, trends, and trends of city tours over the past five years[9,10].

Figure 1 shows the results of CONCOR analysis by repeatedly performing correlation analysis to find groups with appropriate similarity levels[11]. In this study, As a result of the CONCOR analysis on the city tour, a total of 4 groups were formed: 'Use', 'Region', 'Schedule', and 'Theme' in a clockwise direction from the left as shown in Figure 1. Cluster 1 can be seen in the original data as Kakao Keyword signed with Kakao Mobility and the Korea Tourism Organization as a business agreement for local coexistence and revitalization of local tourism. It was confirmed that the Kakao T platform was used for the purpose of revitalizing local tourism and supporting small businesses through the development of various transportation that enhances the convenience of domestic travel, joint cooperation projects and promotion marketing. In order to revitalize the local economy, it was set as “use” because it supports the growth of tourism content by combining local tourist attractions and city tour buses. Cluster 2 was set as “region” because it consisted of regions with a high rate of use of city tours by tourists, centered on regions where city tours are currently being conducted. Cluster 3 was set as “schedule” in which tourists select and prepare their own time and means of transportation so that they can choose the means of transportation in advance and tour along the course when planning a city tour. Cluster 4 was set as “theme” because it diversifies development and courses through FUSO(Feeling-Up Stress-Off) experience, sunset/night view course, emotional travel that satisfies the five senses such as historical/cultural/environmental/ecological tourist attractions, and various cultures and experiences.

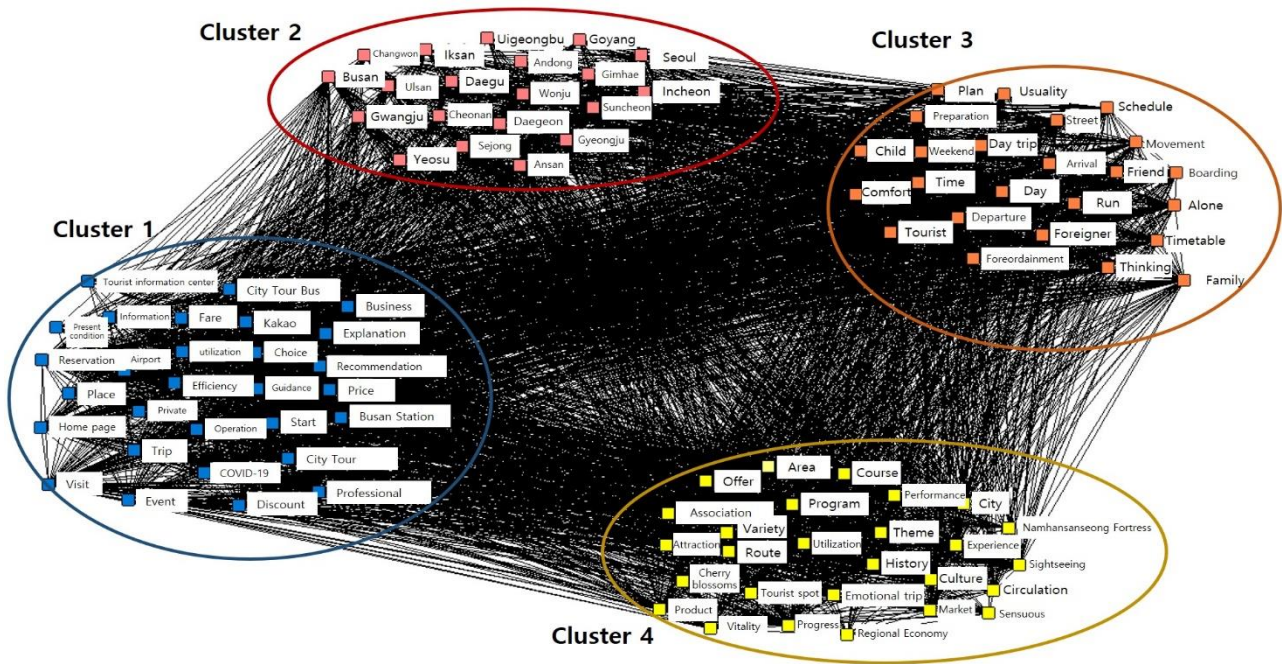


Figure 1. CONCOR Analysis

4. Sentiment Analysis

In this paper, we use a sentiment analysis program to find out the degree of deepened positive and negative sentiment for 'City Tour' and to analyze what emotional words are for each item. Table 1 shows some of the sentiment analysis results for the original data of City Tour.

Table 1 is an actual example of emotional classification, and each positive and negative emotion was

classified into detailed emotions. For example, positive emotions are subdivided into interest, Crush, and pleasure, and the original data classified by subemotional interest are 'It was nice that the guide boarded together and explained the information and major tourist attractions', 'Please recommend a course that you can go with your parents. .', 'It was romantic to enjoy the beautiful night view through the night city tour', 'good', and 'recommended', and the morphemes of emotion words were classified as 'good' and 'recommended'. Next, Table 2 shows a part of sentiment analysis classified as “positive-interest”.

Table 1. Results of sentiment analysis on the original data of the city tour (partial)

Positive	Interest	<ul style="list-style-type: none"> - City Tour Bus restaurant where you can enjoy Seoul city tour and meal together - Wonju City Tour themed course, special course every season, Jeju City Tour night theme operation - It is amazing that the city tour bus is a double-decker bus. - It was difficult for travellers with only public transportation, but they went comfortably with the one-day city tour.
	Crush	<ul style="list-style-type: none"> - It was nice that the guide boarded together and explained the information and major tourist attractions. - Please recommend a course that you can take with your parents. - It was romantic to enjoy the beautiful night view through the night city tour.
	pleasure	<ul style="list-style-type: none"> - Thank you to the guide who worked hard to explain the tourist attractions. - Through the Hwaseong City Tour, I spent a happy day catching clams on the Baekmi-ri tidal flat. - It was a beneficial time to learn history in a fun way, and it was amazing and fun to know the places I had passed by.
negative	Sadness	<ul style="list-style-type: none"> - I think it's a pity that I couldn't go to the course I wanted. - It was difficult because of the busy schedule.
	Rejection	<ul style="list-style-type: none"> - Which means of transportation is most comfortable: a taxi or a city tour bus? - It's a little burdensome because the time is fixed.

Table 2 shows the frequency and emotion intensity of words classified as emotion. Emotion intensity is a value defined in the emotional vocabulary provided by TEXTOM and ranges from 1 to 7, and the higher the value, the stronger the feeling. For example, in the number of frequency cases of the words 'new' and 'want' belonging to 'positive-interest', 'new' is 235 cases and 'want' is 178 cases. However, the emotional intensities of ‘new’ and ‘want’ were 2.7778 and 5.0, respectively. In the positive-interest sensibility, ‘want’ has higher emotional intensity than ‘new’. And looking at the frequency * emotional level, the frequency of want is lower than the frequency of new, but it shows that it is a keyword that similarly affects the emotion of positivity-interest in the entire sentence. In addition, the frequency ratio represents the ratio of the number of occurrences of the corresponding word to the total number of occurrences.

Table 2. Result of ‘positive-interest’ sentiment analysis on the city tour

Emotion Classification	Frequency	Emotional Intensity	Frequency * Emotional Intensity	Frequency Ratio
New	235	2.7778	652.783	2.66
Special	226	3.77778	853.77828	2.56
Want	178	5.0	890	2.02
Amazing	90	5.3333	479.997	1.02
Expect	54	4.66667	252.00018	0.61

Interesting	52	2.6667	138.6684	0.59
Exotic	43	3.6667	157.6681	0.49
Interesting	39	3.4444	134.3316	0.44
Different	32	4.6667	149.3344	0.36
Colorful	26	5.5556	144.4456	0.29
Innovative	7	3.88889	27.22223	0.08
Special	7	4.8889	34.2223	0.08
Unusual	7	4.0	28	0.08

In Table 3, the emotions of tourists can be known by analyzing the emotions shown in the city tour reviews. Out of a total of 8819 documents, there were 7744 positive reviews, accounting for 88.12% of all documents. There were 1075 negative reviews, 11.88%. Through this, it can be expected that the majority of tourists perceive the city tour as a positive image.

Table 3. Results of positive and negative sentiment analysis on the city tour

Division	Frequency (cases)	Frequency ratio (%)	Emotional intensity ratio (%)
Positive	7744	88.12	87.81
Negative	1075	11.88	12.19

As a result of analyzing the detailed emotions of the city tour in Table 4, it was shown in order of crush (68.48%), interest (11.44%), and pleasure (8.2%). It was confirmed that positive emotions accounted for 88.12%, much more than negative emotions. Of the total 8819 cases, 7744 cases were affirmative with a frequency ratio of 88.12%, and 1075 negative cases with a frequency ratio of 11.88%. The emotional intensity ratio was derived as 87.81% and 12.19% for positive and negative, respectively. When emotional intensity was applied, the positive was 0.31 lower than when only simple frequency was considered. In the result of detailed emotion, the image of crush was higher among positive images. This is because words with high frequency and high emotional intensity, such as good, recommended, romantic, and satisfaction, are included. The detailed images of denial sadness, rejection, fear, anger, pain, and surprise were 5.16%, 5.01%, 0.95%, 0.26%, 0.38%, and 0.13%, respectively.

Table 4. Results of detailed sentiment analysis on the city tour

Sensibility in detail	Frequency of detailed sensitivity (cases)	Detailed sensibility ratio (%)
Crush	5996	68.48
interest	1039	11.44
pleasure	709	8.2
sadness	406	5.16
rejection	488	5.01
fear	111	0.95
anger	29	0.26
ache	32	0.38
surprised	9	0.13
TOTAL	8819	100%

5. Conclusion

This study tried to analyze the keyword city tour. Through the text mining technique, the keywords related to the city tour were classified and analyzed as the top 100 keywords, and the results were derived. In this paper, we analyzed the sentiment of tourists' city tours by utilizing big data of portal sites to identify city tour trends and motivations, and looked at the city tours of the last 5 years in general. By using big data frequency analysis, important key words were identified, and sentiment analysis was carried out one step further. As a result, positive and negative elements of tourists were confirmed through positive and negative sentences of the city tour. The keyword that appeared most strongly was crush-good, and it was identified as a factor of satisfaction with the city tour. It is believed that local governments can improve the convenience of use and course development for tourists by paying attention to the reviews of tourists on city tours in each region. This study analyzed the sentiment of the city tour using the city tour review, so it will be helpful for the development of the city tour course by providing tourist information to the city tour platform and analyzing the city tour quality characteristics. It is expected that it will contribute to the development of city tours by continuously analyzing data and recommending customized courses to tourists by converging with more diverse studies based on this, and expanding the platform divided into each region into an integrated platform in the future.

References

- [1] Jung-Hoon Kim, "The Effect of City Tour Service Quality on Tourist' Satisfaction and Behavioral Intention," *Tourism Research*, Vol.44, No.3, Sep 2019.
DOI: <https://doi.org/10.32780/ktidoi.2019.44.3.1>
- [2] Kim Chul-Won and Lee Tae-Sook, "A study on activating strategies of City Tour Bus," *The Korea Academic Society of Tourism and Leisure*, Vol.31, No.1, pp.113-129, Jan 2019.
DOI: <https://doi.org/10.31336/JTLR.2019.1.31.1.113>
- [3] Park Jin-Woo and Seong Eun-Hee, "A Study on the Importance and Satisfaction of Busan City Tour Bus Service: Focusing on IPA Analysis," *Korea MICE Tourism Society*, Vol.23, No.1, pp.167-186, Feb 2023.
DOI: <https://doi.org/10.35176/JMTR.23.1.9>
- [4] You Young-Jun and Lee Seong-Gak, "The Revitalization of Ulsan City Tour as a City Marketing Strategy," *The Korean Urban Geographical Society*, Vol.16, No.2, pp.91-103, 2013.
- [5] Barbara Bickart and Robert M. Schindler, "Internet Forums as Influential Sources of Consumer Information," *Journal of Interactive Marketing*, Vol.15, No.3, pp.31-40, 2001.
DOI: <https://doi.org/10.1002/dir.1014>
- [6] So Jin-Soo and Shin Pan-Seop, "Rating Prediction by Evaluation Item through Sentiment Analysis of Restaurant Review," *Journal of The Korea Society of Computer and Information*, Vol.25, No.6, pp.81-89, June 2020.
DOI: <https://doi.org/10.9708/jksci.2020.25.06.081>
- [7] Sunghoon Kim, "The School Sports Club Big Data Sentiment Analysis," *The Journal of Educational Development*, Vol. 42, No.2, pp.305-323, Oct 2022.
DOI: <https://doi.org/10.34245/jed.42.2.305>
- [8] Lim Young-hee and Kim Hong-bumm, "A study on the sentiment analysis using big data of hotels online review," *Korea Academic Society of Hotel Administration*, Vol.28, No.7, pp.105-123, Oct 2019.
DOI:10.24992/KJHT.2019.10.28.07.105
- [9] Choi Jeong-yun and You Hyung-sook, "A Study on the Change in the Perception of Busan Tourism Using Social Big Data Analysis: Comparison of the periods before and after the COVID-19," *Northeast Asia tourism research*, Vol.17, No.3, pp.175-196, Aug 2021.
DOI: <http://dx.doi.org/10.35173/NATR.17.3.09>
- [10] Deuk-Hee Park, Tae-Goo Kim and Gye-Hee Lee, "An analysis of tourism information network using social big data: Focusing on Seoul and Jeju Island," *International Journal of Tourism and Hospitality Research*, Vol.30, No.8, pp.195-208, 2016.
DOI: <http://dx.doi.org/10.21298/IJTHR.2016.08.30.8.195>
- [11] Se-won Jeon, Seok-Jae Moon and Gi-Hwan Ryu, "Designing the Museum Gamification System," *International Journal of Internet, Broadcasting and Communication*, Vol.14, No.2, pp.219-224, May 2022.
DOI: <http://dx.doi.org/10.7236/IJIBC.2022.14.2.219>