



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

경영학박사 학위논문

관광지 여행상품의 구성요소와
지각된 가치에 대한
사회연결망분석

- 부산 여행을 중심으로 -



2021년 2월

부경대학교 대학원

경영컨설팅 협동과정

정 금 필

경영학박사 학위논문

관광지 여행상품의 구성요소와
지각된 가치에 대한
사회연결망분석

- 부산 여행을 중심으로 -

지도교수 설 훈 구 · 전 중 옥

이 논문을 경영학박사 학위논문으로 제출함

2021년 2월

부경대학교 대학원

경영컨설팅 협동과정

정 금 필

정금필의 경영학박사 학위논문을 인준함.



위 원 장	경 영 학 박사	이 진 수
위 원	경 영 학 박사	전 중 옥
위 원	경 영 학 박사	배 상 옥
위 원	경 영 학 박사	원 천 식
위 원	관 광 경 영 학 박사	설 훈 구



목 차

I. 서론	1
1. 연구배경 및 목적	1
가. 연구배경	1
나. 연구목적	3
2. 연구구성 및 방법	4
가. 연구구성	4
나. 연구방법	6
II. 이론적 배경	7
1. 여행상품	7
가. 여행상품의 개념	7
나. 여행상품의 구성요소	12
2. 지각된 가치	17
가. 지각된 가치의 개념	17
나. 지각된 가치의 구성요소	22
3. 빅데이터와 사회연결망분석	28
가. 빅데이터의 개념	28
나. 빅데이터의 특징과 역할	30
다. 빅데이터와 사회연결망분석 기술과 기법	35
라. 사회연결망분석을 활용한 관광관련 연구	44
III. 연구방법	49
1. 연구문제	49

2. 연구절차	51
3. 데이터 수집대상과 범위 선정	54
가. 데이터 수집대상과 키워드 선정	54
나. 수집 데이터의 분류 범주 선정	56
4. 데이터 수집과 분석방법	59
가. 텍스트마이닝(Text Mining)	59
나. 사회연결망분석(Social Network Analysis)	63
IV. 연구결과	71
1. 데이터 수집결과	71
2. 연구문제 I 분석결과	72
가. 부산 여행에 영향을 미치는 텍스트 분석	72
나. 부산 여행에 영향을 미치는 요소 분석	76
3. 연구문제 II 분석결과	78
가. 지각된 가치에 영향을 미치는 텍스트 분석	79
나. 지각된 가치에 영향을 미치는 요소 분석	81
4. 연구문제 III 분석결과	85
가. 부산 여행 관련 텍스트의 시멘틱 네트워크 구조 분석	86
나. 부산 여행 관련 텍스트의 연관 단어 영향력 분석	94
다. 부산 여행 관련 텍스트의 유사군집 분석	102
V. 결론	106
1. 연구의 요약	106
2. 연구의 논의 및 시사점	109
3. 연구의 한계점 및 향후 연구의 방향	112

참고문헌..... 114
부록..... 123



표 목차

<표 2-1> 여행상품의 개념	9
<표 2-2> 여행상품의 분류	11
<표 2-3> 여행상품의 구성요소	14
<표 2-4> 여행지 선택속성의 개념	16
<표 2-5> 지각된 가치의 개념	21
<표 2-6> 지각된 가치측정의 구성요소	27
<표 2-7> 빅데이터의 개념	29
<표 2-8> 빅데이터의 4가지 구성요소	32
<표 2-9> 빅데이터의 역할	34
<표 2-10> 빅데이터 분석기술 과정	35
<표 2-11> 빅데이터 분석기법	43
<표 2-12> 사회연결망분석 활용방식	45
<표 2-13> 관광분야 사회연결망 연구현황	48
<표 3-1> 여행상품 관련 연구 수집 데이터의 분류 범주	57
<표 3-2> 지각된 가치 관련 연구 수집 데이터의 분류 범주	58
<표 3-3> 동일한 의미로 수정된 명사 텍스트의 예	62
<표 3-4> 동일한 의미로 수정된 형용사 텍스트의 예	62
<표 3-5> 데이터 분석방법 및 내용	64
<표 4-1> 수집 데이터양	71
<표 4-2> 명사 데이터의 주요 텍스트 빈도분석	74
<표 4-3> 명사 데이터의 여행상품 요소별 분류	77
<표 4-4> 형용사 데이터의 주요 텍스트 빈도분석	80

<표 4-5> 형용사 데이터의 지각된 가치 요소별 분류	84
<표 4-6> 부산 여행 명사의 밀도와 중심화	87
<표 4-7> 부산 여행 형용사의 밀도와 중심화	89
<표 4-8> 부산 여행 명사의 유의성 검정 결과	92
<표 4-9> 부산 여행 형용사의 유의성 검정 결과	93
<표 4-10> 명사, 형용사의 네트워크 데이터 유의성 검정 결과	94
<표 4-11> 명사 데이터의 연결중심성 분석	97
<표 4-12> 형용사 데이터의 연결중심성 분석	101
<표 4-13> 각 클러스터의 텍스트 분류	104

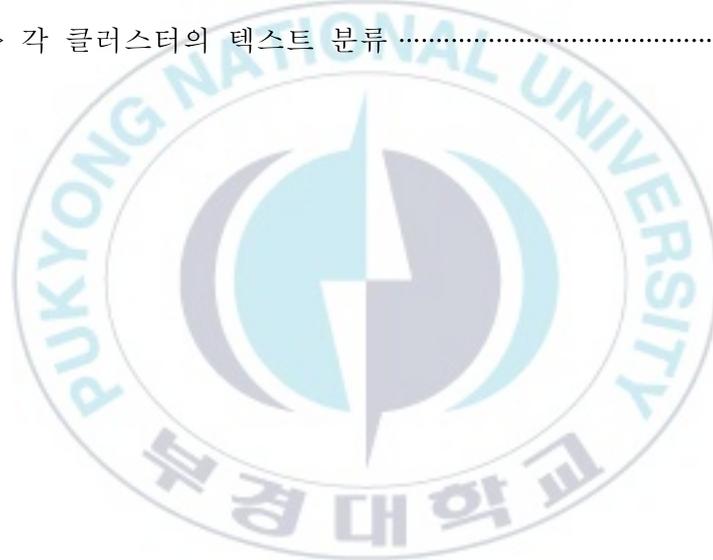
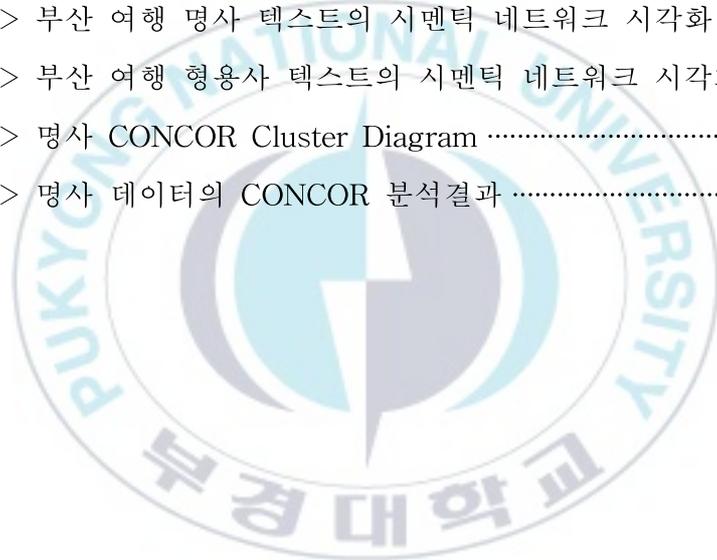


그림 목차

<그림 1-1> 연구의 구성도	5
<그림 2-1> 빅데이터의 3V 특징	32
<그림 3-1> 연구절차	53
<그림 3-2> UCINET6 분석 예시	65
<그림 3-3> NetDraw 분석 예시	65
<그림 4-1> 부산 여행 명사 텍스트의 시맨틱 네트워크 시각화	88
<그림 4-2> 부산 여행 형용사 텍스트의 시맨틱 네트워크 시각화	90
<그림 4-3> 명사 CONCOR Cluster Diagram	103
<그림 4-4> 명사 데이터의 CONCOR 분석결과	105



Social Network Analysis in Components of Travel Products and Perceived Value of Tourist Destination : Focus on Busan Travel

Jeong, Keum Phil

Department of Tourism Management, Graduate School
Pukyong National University

Abstract

As the economy grows and the appetite for travel increases, the number of travelers continues to increase gradually. Even with the growth of travelers, quality enhancement of the domestic travel industry is insufficient. The purpose of this study is to find the components of travel products and the value perceived by travelers and to vitalize the tourism industry by enhancing the quality of travel products.

Big data was collected from major domestic portal sites such as 'Naver' and 'Daum', focusing on the key word, 'Busan travel.' The social network analysis indicated the components of Busan travel products, cognitive values, and the clusters of these factors. The results were summarized as follows.

First, travel to Busan was identified as 'Haeundae', 'activity', 'tourist site', 'travel service', 'transportation', 'accommodation', 'food', 'shopping' and so on. Second, tourism activity such as 'eating', 'riding', and tourist attraction such as 'good', 'recommending', 'many', 'none' were identified as the main attributes in Busan travel products. Third, the perceived value formed in the Busan trip showed positive and emotional value factors. Fourth, the cluster formed for the travel to Busan provided four clusters including 'before traveling to Busan', 'package-related', 'food-related', and 'experience-related.' These most influential attributes and clusters as the components of Busan travel products and perceived value could be seen critical aspects to enhance the quality overall.

The results of this study provided a method of making travel products and using them as marketing tools by investigating the cognitive values of Busan

travel products and tourists. This study was able to measure the factors and cognitive values measured in big data, instead of using the survey method. This study contributed to the body of knowledge utilizing a fairly new methodology, social network analysis, by analyzing the components of travel products and the cognitive value of travelers based on previous research results.

Keywords: Busan, Traveler, Travel Product, Perceived Value, Tourist Destination, Big Data, Text Mining, Social Network Analysis



I. 서론

1. 연구배경 및 목적

가. 연구배경

관광산업은 지역경제에 있어서 지역 재정과 고용창출에 기여하고, 그와 관련된 다양한 산업의 발전에 기여한다. 이에 따라 다수의 지역과 국가들은 경제의 활성화와 경쟁력 확보를 위한 관광산업의 활성화에 노력하고 있다.

지난해 부산을 찾은 관광객은 2천802만 명으로 전년대비 0.4% 증가하였다. 이는 부산시가 휴대전화와 신용카드의 사용에서 발생하는 빅데이터 등을 기반으로 조사한 ‘2019년 부산 관광산업 동향 분석’에 따른 것이다.

관광객은 외국인 268만 명과 내국인 2천534명으로 신용카드 지출액은 전년보다 3.6% 증가한 4조5천605억 원으로 집계되었다(연합뉴스, 2020.5.7.).

또한 팬데믹의 재확산 중에도 국내 여행객들은 꾸준히 증가하는 것으로 파악되었다. 한국항공협회(KCA)에 따르면 항공을 수단으로 한 올해 국내선 탑승객은 1월의 276만4475명에서 코로나19가 확산된 3월에는 110만3206명으로 대폭 감소했다. 그러나 그 이후에는 지속적인 회복세를 보이며 7월 249만263명까지 늘었다(세계일보, 2020.9.1).

현재까지 대다수의 관광분야 연구는 주로 관광객들의 관광수요에 대한 지속적인 증가와 더불어 관광객들의 행태파악을 위한 ‘국민여행조사’, ‘외래관광객조사’와 같은 다양한 설문조사 중심의 정량적 방법을 이용한 조사

위주로 이루어지고 있으나, 조사통계의 특성상 결과에 대한 수집 확보가 낮은 편이다. 또한 조사방식에 따른 정확한 자료량 추측에 대한 문제로 인해 정책이나 마케팅 수립의 기초 자료로서 활용하기에는 많은 추가 검토가 필요한 실정이다.

이에 관광분야에 있어서 기존의 조사통계에 대한 보완과 함께 객관성을 골고루 갖춘 통계를 생산하기 위해 빅데이터의 종합적인 활용과 함께 검증에 대한 연구가 필요하다. 빅데이터를 활용한 온라인 데이터의 내용 분석은 샘플링과 코딩의 한계를 극복할 수 있다는 잠재성을 제공하기도 한다(Lewis, Zamith & Hermida, 2013). 그러나 빅데이터 활용 시 객관성과 신뢰성이 확보되어야 하는 기본적인 문제를 포함하고 있으므로 데이터를 사용함에 있어 다양한 변수를 고려해야만 한다.

최근에는 스마트폰, SNS(social network service) 등의 활성화로 짧은 시간에 많은 양의 데이터가 생성되고 있다. 웹사이트의 댓글이나 이미지 등과 같은 비정형 데이터는 기존의 연구에서 의미 없는 것에 불과하였으나, 빅데이터 활용 방법의 발전으로 인해 산업의 정책과 산업발전의 방향을 결정하는데 있어서 매우 중요한 정보가 되고 있다.

여행객들은 여행을 떠나기 전에 인터넷의 다양한 채널들을 통해서 관광 정보들을 적극 활용하고 있으며, 여행지에 대한 정보를 다양한 SNS 채널을 통해 많은 양의 여행 정보를 쉽게 공유한다.

따라서 본 연구를 통해 부산 여행에 대한 특정 단어의 사용빈도와 그 단어들의 관계를 통하여 중심이 되는 단어를 찾고, 그런 단어들과의 결합과 상관관계를 통하여 부산 여행상품의 구성요소와 부산 여행에서 형성된 지각된 가치를 빅데이터와 사회연결망분석을 통하여 부산 여행업계의 활성화 방안을 수립하고자 하며, 부산 여행상품의 마케팅 실행을 위한 실무적인 시사점을 제공하고자 한다.

나. 연구목적

본 연구의 목적은 비정형 데이터들을 텍스트마이닝(Text Mining)과 사회연결망분석을 활용하여 부산 여행상품과 부산 여행에서 형성된 지각된 가치를 분석하고 여행업계에서의 부산 여행상품 개발 방안에 대한 시사점을 제시하고자 하는 것이다.

본 연구의 목적은 다음과 같다.

첫째, 포털 사이트인 네이버와 다음의 블로그, 카페, 뉴스 등 3개 채널을 대상으로 부산 여행상품의 구성요소와 지각된 가치를 파악할 수 있는 텍스트 데이터들을 수집하고 분석하여, 부산 여행상품의 구성요소와 지각된 가치에 영향을 미치는 텍스트 요소들을 파악하고자 한다.

둘째, 부산 여행과 관련된 텍스트의 분석과 부산 여행상품에 영향을 미치는 여행상품의 요소 분석을 하여 텍스트의 영향력을 파악하고자 한다.

셋째, 부산 여행과 관련된 지각된 가치의 텍스트 분석과 부산 여행상품에 영향을 미치는 지각된 가치의 요소 분석을 하여 텍스트의 영향력을 파악하고자 한다.

넷째, 부산 여행상품의 구성요소와 관련된 텍스트들을 유사군집 분석으로 부산 여행상품 구성요소의 형성 구조를 분석하고자 한다.

다섯째, 부산 여행상품과 지각된 가치의 텍스트 분석결과를 통해서 우리나라 여행업의 여행상품 개발 및 마케팅에 대한 시사점을 도출하고자 한다.

2. 연구구성 및 방법

가. 연구구성

본 연구는 다음과 같은 범위를 구분하여 수행하였으며, 이 과정을 통하여 연구의 목적을 달성하고자 하였다. 공간적 범위로는 국내 최대의 온라인 포털사이트 네이버와 다음을 중심으로 블로그, 카페, 뉴스의 3개 채널을 대상으로 ‘부산 여행’ 키워드를 적용하여 국내 이용자들을 대상으로 텍스트를 수집하였다. 시간적 범위는 많은 문헌연구에서 소비자의 최근 동향을 파악하는데 주로 사용하는 1년의 기간을 기준으로 하였으며, 2019년 1월 1일을 시작으로 2019년 12월 31일까지의 1년간 데이터로 한정하였다. 연구는 서론, 이론적 배경, 연구방법, 연구결과, 결론의 5장으로 구성했으며, 구체적인 내용은 다음과 같다.

I. 서론에서는 연구배경 및 연구목적, 연구구성과 연구방법에 대해 기술하였다.

II. 이론적 배경에서는 여행상품, 지각된 가치, 빅데이터를 통한 사회연결망분석 등의 관련된 문헌 및 선행연구의 분석을 정리하였다.

III. 연구방법에서는 문헌연구를 바탕으로 하여 연구문제를 제시하고, 연구설계를 통하여 텍스트 수집과 분석방법에 대한 연구절차를 기술하였다.

IV. 연구결과에서는 데이터의 수집 및 분석결과를 표출한 다음, 시각화를 통해 데이터의 결과를 보기 쉽게 제시하여 총체적 결과해석을 기술하였다.

V. 결론에서는 연구의 요약과 논의 및 시사점, 연구의 한계점과 향후 연구 방향을 제시하였다.



<그림 1-1> 연구의 구성도

2. 연구방법

본 연구는 고찰을 통한 선행연구의 문헌연구와 온라인 빅데이터 수집과 사회연결망분석을 통한 실증연구를 실행하였다.

첫째, 연구의 목적을 달성하기 위해 연구의 배경과 목적을 제기하고, 선행연구 고찰을 통한 이론적 체계를 수립하기 위해 문헌 연구를 실행하여 연구문제와 연구방법을 제시하였다. 이를 위해 여행상품, 지각된 가치, 빅데이터와 사회연결망분석 등에 대한 선행연구를 고찰하고 연구문제 해결에 적합한 데이터의 수집과 분석방법을 제시하였다. 문헌연구를 위해 국내외 논문, 학회지, 인터넷, 관련기관 보도자료, 기타 각종 간행물 및 통계자료들을 활용하였다. 둘째, 연구문제를 해결하기 위하여 빅데이터 분석 솔루션인 TEXTOM을 활용하여 키워드를 선정하고, 네이버와 다음의 온라인 포털사이트를 중심으로 데이터 수집을 위해 텍스트마이닝을 실행하여, 부산 여행상품의 구성요소와 지각된 가치에 관한 텍스트 데이터를 수집하였다.

셋째, 텍스트마이닝을 통해 수집된 텍스트 데이터 중에서 부산 여행상품 구성요소와 지각된 가치의 요소에 관련된 텍스트들을 선택하여 정제하였다. 최종 정제된 텍스트들의 동시출현 주요 텍스트들을 추출하여 부산 여행상품과 지각된 가치에 영향을 미치는 텍스트들을 각 요소별로 분류하여 파악하였다. 넷째, 동시출현 빈도분석으로 추출된 주요 텍스트들을 엑셀파일 매트릭스로 변환시켜 소셜 네트워크 분석도구인 UCINET6를 통해 통계분석을 실행하였으며, UCINET6의 NetDraw 기능을 통해 결과를 시각화하였다. 다섯째, 텍스트마이닝과 사회연결망분석을 통해 도출된 결과를 분석하여 국내 여행업계의 부산 여행상품 개발 및 마케팅 방안에 대한 시사점을 도출하였다.

Ⅱ. 이론적 배경

1. 여행상품

가. 여행상품의 개념

여행이란 ‘일이나 유람을 목적으로 다른 고장이나 외국으로 가는 일, 자기 거주지를 떠나 객지로 다니는 것’이라는 사전적 의미를 갖는다(윤대순, 1997). 여행이란 관광객이 집을 떠나서 다시 돌아올 때까지의 모든 경험과 추억들을 의미하는 것이며, 이러한 것들을 상품화시킨 것을 여행상품이라고 정의하였다(주상욱, 2000). 또한 여행은 인간이 자신의 생활권을 떠나서 특정한 여행의 목적들을 정해 놓고 일정기간 동안 이동하였다가 다시 자신의 생활권으로 돌아오는 이동행위이다(김현철, 2002). 이러한 여행을 상품화한 여행상품은 현재 다양한 형태로 존재하며 판매되고 있다. 기본적으로 여행상품은 서비스 상품으로 그 상품이 가지고 있는 기본적인 특성을 포함하고 있다. 그리고 여행상품은 여행객이 여행을 하면서 필요로 하는 서로 다른 관광 사업체가 제공해주는 시설물과 모든 서비스를 포함한 종합적인 상품과 부분적 상품이라 정의할 수 있다(박정민, 2012).

WTO는 여행상품을 여행 목적지, 교통수단, 숙박, 보조서비스의 관광매력을 결합시킨 것이라고 정의했으며, 정익준(1997)은 결국 여행상품이란 수많은 여행 소재들의 결합에 의해 생산된다는 것을 확인하게 된다고 정의하였다.

여행상품이란 각각의 많은 여행 구성요소들의 합리적인 결합으로 생산되

는 것이며, 단순히 여행 구성요소들이 개개인의 예약만으로 여행상품이 될 수 없고 여행을 상품화하기 위해서는 우선적으로 여행 목적지나 관광 대상을 설정하고 이들의 여행 구성요소를 출발해서 도착할 때까지의 일관된 종합적인 것으로 구성하여 부가가치를 추구해야하는 것이라고 정의하였으며(천덕희, 2005), 이러한 부가가치의 창출은 여행사가 독자적으로 가격을 설정하여 특색 있고, 자기주장과 책임을 가진 것이어야 한다고 하였다. 박시점 외(2010)는 여행상품은 여행 장소, 매력성, 서비스 및 모든 과정 등이 결합되어 거래되는 대상물들로 정의하였다. 박시점·이병열·홍영호·서정원(2010)은 여행상품은 여행사가 여행객들의 욕구를 유발시켜 각종 서비스를 제공해서 생산된 유형·무형의 상품이라고 정의하였다.

이성철(2013)의 연구에서는 여행상품에 대해 여행의 출발에서 마지막까지의 모든 여행 과정, 숙박 장소, 교통 등과 같은 여행관련 시설물과 관광지 등의 여행요소들을 잘 결합시켜 그것들에 의해 제공 되어지는 서비스와 총체적으로 이루어진 유형·무형의 상품이라고 정의하였다.

여행상품에 대한 주요 선행연구를 <표 2-1>과 같이 정리하였다.

<표 2-1> 여행상품의 개념

연구자	여행상품의 개념
주상욱 (2000)	관광객이 집을 떠나 다시 돌아올 때까지의 모든 추억과 경험들을 의미하고 이것들을 상품화한 것
정익준 (2001)	여행 기업이 생산하는 모든 용역과 재화로 유형·무형적인 상품
최승이 (2001)	관광객의 욕구를 충족시켜주는 유형·무형의 재화와 모든 서비스의 결합물. 여행상품의 8가지의 특성
이선희 (2005)	숙박, 식사, 안내, 관광, 운송 등의 자동화된 모든 요소가 기능적, 시간적으로 구성되어져 효과를 나타내어서 만족하는 것
천덕희 (2005)	각각의 다양한 여행 구성요소들의 결합이 합리적으로 생산되는 것 여행을 상품화하기 위해서는 우선적으로 관광대상이나 여행목적지를 설정하고 이들의 여행구성요소를 출발에서부터 도착까지의 일관되게 종합적으로 구성하여 부가가치를 추구해야하는 것
송수원 (2010)	유형적인 특성으로 구성되어진 유형의 상품과는 다르게 무형적 특성인 서비스가 주체가 되고 무형성·소멸성·이질성·동시성·계절성의 5가지 특성
박시점 등 (2010)	여행지, 매력성, 모든 서비스 및 모든 과정들이 결합되어진 거래의 대상물이며, 관광기업이 관광객들의 욕구를 끌어내어서 좋은 여행 대상을 바탕으로 각종 서비스를 제공해서 생산된 유형·무형의 상품
박정민 (2012)	관광객들이 여행을 함에 있어서 필요한 각각의 관광기업체가 제공하는 모든 시설 또는 서비스를 포함한 종합상품과 부분적 상품
이성철 (2013)	여행의 시작에서 끝나는 모든 과정의 교통, 호텔, 그 밖의 여행과 관련된 관광자원과 시설 등의 여행상품들을 잘 결합해서 제공 되어지는 모든 서비스와 종합적으로 이루어진 유형·무형의 상품
한국관광공사 (2014)	여행자가 일정한 목적지와 장소에서 관광할 수 있도록 여행 기업이 제공하는 모든 재화와 함께 서비스로 판매할 것을 약속하여 상품화한 것

자료: 선행연구를 바탕으로 연구자 재구성

여행상품에는 국내여행상품, 외국인의 국내여행을 안내하는 외래여행상품(inbound tours), 내국인의 국외여행상품(outbound tours)이 있으며, 어느 곳에서 숙박을 할 것인가 하는 숙박상품, 어떤 교통편을 이용할 것인지에 해당되는 교통수단, 여행 중 어떤 음식을 이용할 것인지에 해당되는 식사상품과 토산품, 특산품, 기념품과 관련되는 쇼핑상품 이외에 관광을 행하는 장소에서 이루어지는 관광 입장권, 사진 등의 상품이 함께 이루어진다(김옥제, 2001).

내국인의 국외여행상품에 속하는 여행상품은 관광객이 여행을 할 때 여행정보와 이용시설 및 여행에 필요한 모든 재화와 서비스를 구매하는 거래의 대상임을 나타낸다. 이러한 개별의 대상들을 결합한 상품 또는 개별의 상품을 총칭하며 여행사에서 다루고 있는 모든 업무와 서비스를 포함하는 넓은 의미의 여행상품 개념과 함께 여행사가 판매하고 있는 패키지 상품과 같은 좁은 의미의 개념으로 나뉘어 파악될 수 있다(최승일, 2007). 다시 말해, 여행상품이란 관광지로 하여 각각의 여행 구성요소들이 통합적으로 생산되는 상품이며 여행의 출발에서부터 도착에 이르는 모든 과정 속에 소요되어지는 여행관련 시설과 여행 서비스를 결합한 것으로 여행자들을 만족시키기 위한 여행의 가치를 포함한 종합이라고 할 수 있다(이종철, 2016).

여행의 상품은 참가형태, 판매시장 상품구성, 여행 형태, 참가 계층, 항공운임, 여행 주체에 따라 8가지로 분류될 수가 있으며, 참가형태에 따라 FIT형, 개인 참가형, 단체 조직형, 판매시장에 따라 일반 모집형, 인센티브형, 어퍼니티형, 경로여행으로 나뉘며, 목적에 의해서는 관광여행, 연수여행, 업무여행, 시찰여행, 국제회의로 분류된다. 상품구성에 의해서는 단체여행, 패키지투어, 여행 형태에 의해서는 DIY형, SIT형, 디럭스형, 주유형, 체재형으로 나뉘고 있다. 그리고 참가 층에 의한 분류는 숙련여행, 실버여행, 허니문(honeymoon), 여행주체에 의한 분류는 공동주최여행, 주최여행,

초대여행, 수배여행, 마지막으로 항공운임에 의한 분류는 보통 개인 운임형, GIT형, 차터(charter)형, 에어온리(aironly)형이다(정익준, 1997). 여행상품의 분류에 대한 주요 선행연구를 <표 2-2>와 같이 정리하였다.

<표 2-2> 여행상품의 분류

여행상품의 분류	여행상품
참가형태에 의한 분류	FIT형 · 개인참가형 · 단체조직형
판매시장에 의한 분류	일반모집형 · 인센티브형 · 어퍼니티형, 경로여행
목적에 의한 분류	관광여행 · 연수여행 · 업무여행 · 사찰여행 · 국제회의
상품조성에 의한 분류	단체여행 · 패키지 투어
여행형태에 의한 분류	DIY형 · SIT형 · 디럭스(deluxe)형 · 주유형 · 체재형
참가층에 의한 분류	숙련여행 · 실버여행 · 허니문(honeymoon)
여행주체에 의한 분류	공동주체여행 · 주최여행 · 초대여행 · 수배여행
항공운임에 의한 분류	보통 개인 운임형 · GIT형 · 차터(charter)형 에어온리(aironly)형

자료: 정익준(1997), 최신 관광사업론, 형설출판사, 292-293.

나. 여행상품의 구성요소

여행상품은 이미 완성되어진 다양하고 이질적인 여행요소를 생산자인 여행사가 결합하여 하나의 상품으로 만든 것을 의미한다(송호철, 1992).

박정민(1995)은 여행상품이란 여행을 하는데 있어서 관광객이 필요한 모든 시설 및 대상인 것으로 정의하였다. 이러한 여행상품의 개념이 보급되어 일반적으로 활용되기 시작한 것은 최근의 일이다.

여행업의 발전사적인 관점에서 보면 이는 대중관광이 성립된 지금에 이르러서야 비로소 여행상품이라는 개념이 정착되어 사용하게 되었다. 즉, 여행사가 여행객 개인별 요구에 따라 숙박시설과 교통수단을 수배하거나 또는 주공급자의 요청에 의해 고객을 소개하던 시대로부터 여행사가 미리 여행객의 욕구와 비용을 고려하고, 예상 여행객수를 추산하여 여행소재와 관광목적 등을 구매하여 패키지 상품을 만들어 판매하는 시대에 돌입해서야 여행상품이라는 용어가 탄생하게 되었다.

비교적 여행목적지가 원거리에 있고 체류기간도 길며 다양성을 추구하는 여행일 경우에는 더 많은 여행소재들이 개입되기 쉬운데, 이처럼 다양한 여행 소재들을 결합시킨다는 의미에서 여행 산업을 장치 산업이라고도 한다. 여행상품은 숙박시설, 교통수단, 요식시설, 지원시설, 관광대상 목적물, 기념품, 여행서비스 총 7개로 구성될 수 있다. 여행상품 구성요소 중의 하나인 교통수단은 항공기, 열차, 버스, 선박으로 예를 들 수 있으며, 숙박시설로는 호텔, 모텔, 여관 등이 포함되며, 요식시설에는 레스토랑, 요리식당 등이 포함된다. 그리고 관광대상 목적물은 자연자원과 인문자원으로 나눌 수 있으며, 자연자원은 기후 및 경관이 포함되며 인문자원은 인간문화재와 및 전통행사가 포함된다. 또한 지원시설은 미디어를 예를 들 수 있다(정의

준, 2000).

여행상품은 교통, 숙박시설, 유흥, 안내, 기타 서비스 등의 복합적 요소를 시간적, 기능적으로 구성한 것이라고 하였으며 구성요소들은 여행목적지로서의 관광지, 한 지점에서 다른 한 지점으로 이동시키는 모든 교통수단, 숙박시설, 관광활동, 여행 중의 식사문제, 쇼핑활동, 여행사, 선택 관광·야간 관광, 여행 안내원으로 총 9가지로 정의하였다(천덕희, 2005).

문명희(2003)는 숙박, 교통, 식사 및 관광지를 여행상품의 필수적인 구성요소로 정의했으며, 호텔과 항공, 식당, 쇼핑, 관광활동, 여행사, 여행사 안내원으로 총 6가지 요소를 사용하였다.

조태진(2004)은 여행상품의 구성요소는 여행상품의 종류와 형태에 따라서 그 범위와 구성요소가 달라질 수도 있으며, 또한 그 시대의 문화, 사회적인 영향에 따라서도 달라질 수 있다고 하였다. 그러나 여행상품의 구성요소에는 관광 자원과 대상을 비롯하여 여행에 있어서 필요한 모든 인적, 물적 서비스가 포함된다 할 수 있으며, 그 구성요소는 여행목적지로서 관광활동, 관광지 관련사항, 숙박시설, 교통수단, 식사문제, 쇼핑활동, 여행사, 관광안내원으로 구성하였다.

정대봉(2008)은 여행상품은 여행객이 생산하는 일체의 서비스와 재화를 말하며 여행객을 만족시킬 수 있는 유형과 무형의 상품으로서, 여행상품의 구성요소로는 음식, 기념품, 토산품, 특산품 등과 같은 유형의 상품과 운송, 숙박, 여행안내, 오락 등과 같은 무형의 상품이 있다고 하였다.

김기석(2017)은 패키지여행 상품 구성요소를 관광지, 교통, 숙박, 식사 및 쇼핑, 여행사로 구분하여 설명하였다.

<표 2-3> 여행상품의 구성요소

연구자	여행상품의 구성요소
정익준(2000)	관광대상 목적물, 숙박시설, 교통수단, 요식시설, 지원시설, 여행 관련 서비스
문명희(2003)	쇼핑, 항공, 식당과 호텔, 관광활동, 여행안내원
조태진(2004)	관광지 관련사항, 교통수단, 숙박시설, 식사문제, 관광활동, 쇼핑 활동, 여행사, 관광안내원
천덕희(2005)	숙박시설, 교통수단, 관광활동, 여행사, 여행안내원, 여행 중의 식사문제, 쇼핑활동, 선택 관광·야간관광,
정대봉(2008)	음식, 기념품, 특산품과 같은 유형상품과 숙박, 운송, 여행안내, 오락과 같은 무형상품
이미혜(2009)	관광자원과 관광지, 숙박시설, 교통시설, 식음료시설, 정보서비스, 쇼핑, 출입국서비스, 금융서비스
황명선(2014)	항공사, 호텔, 식당, 관광지
김기석(2017)	관광지 숙박, 식사 및 쇼핑, 교통, 여행사
Medlik & Middleton (1973)	관광자원, 시설, 접근성, 이미지, 가격
Stephen(1994)	물리적 환경, 서비스시설, 환대성, 여행자활동, 편익

자료: 선행연구를 바탕으로 연구자 재구성

여행상품 개념에 관광지 선택속성과 태도가 포함되어 있으며, 관광지 선택속성은 관광객이 관광목적지에 대한 태도를 측정하게 하는 틀을 의미하며, 관광객은 다양한 관광지 속성을 비교, 평가하여 관광지를 선택하고 관광태도를 결정하게 된다(Mayo & Jarvis, 1981). 관광지 선택속성은 관광객이 관광지를 선택할 때 중요하게 고려되는 요인으로서, 관광을 하는 동안 체험하는 각각의 요소들의 개별속성에 대해 개인이 지니는 선택정도를 의

미하고, 속성 중에서는 선택에 결정적인 영향을 미치는 것과 그렇지 못한 것이 있기 때문에 효과적인 마케팅 전략수립을 위해서는 이러한 속성별 특징을 파악하는 것이 중요하다(김진탁·김계섭·공기열, 2002).

관광지 선택속성은 관광객들의 관광의사 결정에서 중요한 역할을 하지만 (Gronroos, 1984; Kotler, Haider & Rein; 1993), 일반적으로 관광지는 관광객의 요구나 욕구를 충족시킬 수 있도록 효과적으로 개발하여 관광시장에 제공됨으로써 매력적인 장소가 될 수 있을 것이다(곽재용, 2005). 따라서 관광지 선택속성은 관광지를 하나의 관광 상품으로 보는 관광객들이 이를 선택하는데 있어 가장 중요하다고 생각하고 선택하는 속성으로서, 관광객들이 선택이나 의사결정에 있어 영향을 미치며, 최종 방문지로 이어지는 핵심적 역할을 한다(Goodrich, 1997). 이와 같이 관광지 선택속성은 관광객의 목적지 선택기준을 포함하는 다양한 의사결정과정으로, 관광객들은 관광지의 속성비교와 평가를 통해 관광지를 선호하고 관광행동을 결정하게 된다(김오성, 2010).

관광지 선택속성들은 관광객들의 관광 목적지 선택 평가기준, 관광지선택의 결정요인으로 작용할 수 있으며, 사람들의 모든 활동들이 그들의 선택과 관련이 있듯이 모든 관광행동은 선택속성을 포함하고 있다(최용훈, 2008). 즉 관광지 선택속성이란 관광지에 대한 개인의 주관적인 선택기준을 의미하므로 국가별 관광지 선택에 대한 의사결정은 대부분 관광 참가자의 만족 및 방문의도에 영향을 주게 되고, 또한 영향을 받게 되는 상호보완적인 관계라고 할 수 있다(유재충·노재현, 2010). 관광지의 선호도와 선택속성은 관광객의 의사결정이나 선택에 영향을 미치고 궁극적으로 방문행동으로 연결된다(류인평·김정준, 2010).

관광지 선택속성이란 관광지를 구성하는 다양한 요소들의 개별속성에 대해 개인이 지니는 선택정도라고 할 수 있다(최용훈, 2008). 따라서 관광객

들이 가고자 하는 관광지를 선택할 때 중요하게 고려되는 요인으로서, 관광을 체험하는 각 요소들의 개별속성에 대하여 개인이 가지고 있는 선택정도를 의미한다(노정연·조우제, 2011).

<표 2-4> 여행지 선택속성의 개념

연구자	여행지 선택속성의 개념
김진탁·김계섭·공기열(2002)	여행을 하는 동안 경험하는 각각의 요소들의 개별 속성에 대해 개인이 가지는 선택정도를 의미
공기열(2003)	유형의 시설과 무형의 인적서비스가 결합되어 그 유형과 형태가 일반적 소비상품에 비하여 매우 복잡하고 다양한 특성을 지니고 있어 일반상품과는 다르게 고객 의사결정에 중요한 영향을 미치는 요소
최용훈(2008)	관광객들의 관광목적지선택에 있어서 관광지 선택과 평가기준의 결정 요인으로 작용할 수 있으며, 모든 활동들이 관광객들의 선택과 관련이 있는 것 처럼 모든 관광행동은 선택속성을 포함
유재충·노재현(2010)	관광지에 대한 주관적인 선택의 기준을 의미하며, 국가별 관광지 선택에 대한 의사결정은 관광객 대부분의 관광 만족 및 방문의도에 영향을 주고, 받게 되는 상호보완적인 관계
노정연·조우제(2011)	관광지를 선택 시 중요하게 생각되는 요인, 관광을 체험하는 각 요소들의 개별속성에 대해 개인이 가지고 있는 선택정도를 의미
송학준(2012)	관광객이 관광활동을 위해 선택에 있어서 관광지는 역사적 장소, 자연경관, 서비스, 이벤트, 시설 등 방문객의 관광욕구와 관광동기를 만족시킬 수 있는 다양한 속성들로 구성되어, 관광지가 지니고 있는 고유한 매력의 요소들을 관광지 선택 속성이라 함
박종진(2012)	도시가 가지고 있는 여러 가지 요소들과 자원 및 도시만의 차별화된 매력성

자료: 선행연구를 바탕으로 연구자 재구성

2. 지각된 가치

가. 지각된 가치의 개념

지각된 가치는 서비스 산업에서 중요한 개념이며, 이는 고객이 지각하는 이익이나 혜택으로 볼 수 있다. 이유재(1999)는 가치란 일반적인 의미에서 사람들이 제품이나 서비스를 통해서 기대하는 이익이나 혜택으로 정의하였다. 지각된 가치는 본질적으로 상대적이고 비교적이며, 지각된 경험 가치는 지속적으로 변화하는 독특하고 개인적인 요구를 다룬다고 하였다(Addis & Holbrook, 2001). 그러므로 소비자가 경험한 가치는 공동으로 만들어지는 상황에서 개인적으로 나타나고(Gronroos, 2011), 그 지각된 가치는 다양한 가치들로 분류되어, 소비자의 선택에 영향을 미칠 수 있다고 하였다(Sheth et al., 1991).

Woodruff(1997)는 고객이 재화이나 서비스를 구매하고자 할 때와 사용하는 동안이나 사용 후의 가치를 다르게 지각한다고 하였다.

Sweeney & Soutar(2001)는 지각된 가치란 구매의 여러 단계에서 사전에 발생하지만 일반적으로 만족도는 구매 후에 사후 평가되므로, 결과적으로 제품이나 서비스를 구매하거나 사용하지 않아도 지각된 가치를 얻을 수 있지만, 만족은 재화나 서비스를 사용한 경험에서 얻어진다고 하여 지각된 가치와 만족도의 차이를 보여주었다.

지각된 가치는 소비자에게 제공되는 것과 소비자가 얻는 것에 대한 지각을 근거로 상품 효용에 대한 소비자의 종합적 평가라고 하였다.

지각된 가치란 서비스산업에서 고객의 행동을 이해하기 위한 중요한 개

념의 하나로 사용되며(Ha & Jang, 2010; Jensen, 1996), Zeithaml(1988)은 가치의 의미를 다음의 4가지로 정의하였다.

첫째, 가치란 저렴한 가격을 뜻한다. 이는 서비스구매에 있어서 주로 금전적인 가격에 맞춘 것으로 소비자들은 세일제품을 구매하거나 쿠폰 혹은 제공된 다른 서비스를 이용하여 구매한 경우 그것을 가치가 있다고 생각한다.

둘째, 가치란 서비스로부터 소비자가 얻고자하는 모든 것들을 의미한다. 이는 주로 소비자가 서비스 구매로부터 얻게 되는 이익에 중점을 둔 것이다. 이런 정의에서 소비자들은 그들이 원하는 수준의 품질을 가지게 된다면 가격에는 그다지 신경 쓰지 않는다는 것이다.

셋째, 가치란 지불한 가격에 대비해서 소비자가 제공받은 품질이다. 이것은 소비자가 지불한 가격과 서비스 품질과의 사이에 있어서 교환관계를 파악하여, 준 것만큼 얻었다고 생각된다면 그것은 가치가 있다고 판단하는 것이다.

넷째, 가치란 소비자가 지불한 가격에 대해 제공받는 것이다. 이것은 소비자가 지불한 가격을 금전적 요소에만 한정한 것이 아니라 시간과 노력 등과 같은 고려해서 얻게 되는 모든 것을 총칭하여 가치라고 정의하는 것이다.

주영환(2007)은 소비자가 해당 제품과 서비스 이용을 위해 지불한 가격에 대한 금전적 가치, 대인 관계에 대한 사회적 가치, 이용 시에 즐거움에 대한 쾌락적 가치를 지각된 가치로 정의하였다.

주현식(2008)은 지각된 가치란 비용과 편익, 가격에 대한 서비스평가와 서비스에 대한 가격평가로 정의하였다. 이와 같이 지각된 가치는 소비자 효용과 희생에 대비되는 지각된 편익, 적합함, 심리적 가격, 그리고 품질 등과 같이 다양하게 정의되고 있다.

특히, 지각된 가치는 인간의 개인행동에 큰 영향을 미치는 기본적인 것이다. 이는 믿음, 관심, 태도, 의견 등의 개념보다 포괄적인 개념으로서 동일한 행동을 평가할 때 개인의 행동에 대해 설명력이 높은 개념으로서 평가받는다(Zeithaml, Gallarza & Saura, 2004).

Day(1990)는 지각된 가치란 고객이 지각한 편익과 비용간의 차이를 나타내는 것이라 정의하였으며, Monroe(1990)는 구매 고객이 인지하는 가치란 제품구매 시 지불한 비용에 비해서 얻게 되는 제품의 질과 효용의 상쇄 정도를 의미한다고 하였다. Woodruff(1997)는 고객이 얻는 것과 제품이나 서비스를 받아 사용하는 것에 대해 반대로 포기해야 하는 것의 상반관계를 의미한다고 하였고, Kashyap & Bojanic(2000)은 서비스 품질의 전반적 대비 가격 지불을 서비스 가치의 비교 관점으로 지각된 가치를 정의하였다.

주영환(2007)은 특히 무형의 서비스를 제공하는 관광분야에서 소비자의 구매행위를 유발하기 위해서는 지각하는 가치에 대한 올바른 이해가 필요하며, 지각된 가치의 제고는 해당 기업이 경쟁우위를 선점하기 위한 필수요소라고 주장하였다.

마케팅 실무자와 연구자들은 시장세분화, 상품 차별화, 포지셔닝 정책과 같은 마케팅 전략수립을 위하여 지각된 가치의 차원을 밝히고 측정하는데 관심을 가지고 있다(Gallarza & Saura, 2004). 가치는 지각된 가격이나 효용성에 대한 평가에서만 판단되어지는 것이 아니라 구입시점과 서비스 접점에 있어서 주변 환경적인 요인에 의해서도 영향을 크게 받는다.

서비스나 상품의 품질이 아무리 좋다고 하더라도 고객들이 지각하는 경제적 비용이 높을 때 기대치는 높아지고 고객들이 지각하는 가치가 낮아져서 고객은 만족하지 않는다(Zeithaml et al., 2006).

이러한 선행연구의 결과에서 지각된 가치는 유형의 상품 또는 무형의 서비스에 관계없이 고객의 만족과 재구매 의도, 선호도와 같은 행동의도에

결정적 영향을 미치고 있다는 사실을 나타내고 있다.

최근에 연구자들이 지각된 가치에 큰 관심을 갖는 것은 지각된 가치가 직접적으로 개인의 행동과 관련 있는 것으로 평가되기 때문이며, 가치에 대한 이해는 제품과 서비스를 공급하는 기업의 경쟁우위를 선점하기 위한 전략도출에 있어서 필수적이라 할 수 있다. 이에 본 연구에서의 지각된 가치란 고객의 인지에서부터 오는 고객가치라는 측면으로 접근하고자 한다. 소비자들이 제품이나 서비스를 통해 기대하는 이익이나 혜택을 의미하며, 이는 구매를 결정하는데 있어서 가격보다 지각된 가치가 더 중요한 요인으로 작용한다(Perkins, Reynolds & Zeithaml, 1988; Shethetal, 1991).

본 연구에서는 선행연구를 종합해 본 결과, 가치의 광범위한 개념인 지각된 가치를 연구한다. 지각된 가치는 고객들이 여행상품을 이용하기 위한 편익과 비용간의 상쇄에 의해서 고객들이 지각하는 서비스의 품질에 대한 전체적인 평가 즉, 가격, 시간, 노력을 고려한 품질에 대한 고객의 지각으로부터 형성되는 것으로 정의할 수 있다. 다시 말하면 관광객이 서비스를 제공받기 위해 지불한 전체 비용에 대해서 고객이 받은 서비스와 편익에 대한 가격평가와 가격에 대해 고객들이 서비스 평가하는 것으로 소비자의 서비스 평가에 있어서 가장 핵심적인 요소인 것이다.

<표 2-5> 지각된 가치의 개념

연구자	지각된 가치의 개념
Zeithaml(1988)	소비자의 지각된 가치는 상품 및 품질의 속성인 내재적, 외재적인 것들에 대한 고객들의 지각에 의해서 이루어지는 것 제공받은 것과 지불한 것에 대한 인식을 바탕으로 하여 제품의 효용성에 대한 소비자의 종합적인 평가
Newman, Gross & Sheth(1991)	정서적, 사회적, 인지적, 기능적, 상황적 가치에 대한 지각된 효용성에 의해서 구매가 결정
Woodruff (1997)	고객의 지불(가격, 희생)과 고객의 획득(품질, 혜택, 효용)간의 교환 관계에 대한 인식
Lapierre(2000)	경쟁우위에서 고객들이 지각하는 높은 가치
Gordon H.G. McDougall(2000)	소비자에 의해 사용된 전체비용과 관련된 혜택과 결과
Sweeney & Soutar(2001)	지각된 가치는 효율성이 한 가지가 아닌 다차원 개념
Zeithaml(2004)	의견·관심·믿음·태도 등을 포함하고 있는 개념으로서 개인행동에 대한 설명력으로서 우수한 것으로 평가
Gallarza(2004)	마케팅 연구자들은 포지셔닝 정책, 상품 차별화, 시장 세분화 등과 같은 마케팅 전략을 수립하기 위해 지각된 가치의 차원을 밝힘
연승호·엄서호 (2004)	관광객의 화폐적 또는 비화폐적인 사용과 관광 상품에 의해 제공된 혜택간의 비교
주영환(2007)	기업이 경쟁우위를 선점하기 위해 지각된 가치가 필수 요소로 제고됨
여호근·박경태 (2007)	관광객들의 투자 비용과 관광한 후의 느낌을 비교하여 긍정적인 효용을 인지하는 것

자료: 선행연구를 바탕으로 연구자 재구성

나. 지각된 가치의 구성요소

가치는 맥락에 따라서 다른 의미를 갖는 추상적 개념이지만 마케팅에서는 가치를 고객의 관점에서 정의하면서(Patterson & Spreng, 1997), 기능적, 상황적, 사회적, 감성적, 효율적인 속성을 가진 소비행동의 동기로서 품질과 가격의 상쇄관계라는 협의의 관점에서 해석하고 있다(이미혜, 2009).

Oh(2000)는 가치의 구성요소로 제품과 서비스에 대한 금전적 가치, 가격 비교, 가격, 가격에 상응하는 무엇에 대한 기대가치를 측정하였다.

Naylor & Frank(2001)는 전반적 금전, 기대된 금전적 원가, 금전적 원가, 기대라는 전체 가격의 성과, 전체가격, 품질을 독립변수로 해서 가격이 고객의 지각된 가치에 미치는 영향에 대한 상관관계를 연구하였다.

Sweeney & Soutar(2001)는 기존 선행연구를 바탕으로 각각의 요인들을 측정하여 지각된 가치를 감성적 가치, 기능적 가치, 사회적 가치, 금전적 가치의 4가지 요인으로 분류하였다. Ledden et al.(2007)은 지각된 가치를 기능적 가치, 사회적 가치, 인식적 가치, 감성적 가치, 조건적 가치로 구분하고, Cengiz and Kirbir(2007)는 지각된 가치를 기능적 측면과 정서적 측면으로 구분하였다(장병주·정구점, 2011).

Lee, Yoon & Lee(2007)는 감성적 가치, 기능적 가치, 총체적 가치의 3가지로 지각된 가치를 구분하였다.

박동균(2003)은 금전적 가치와 비금전적 가치로 지각된 가치를 구성하여 만족, 선호도, 재이용의도간의 관계를 규명하였다.

박경대(2007)는 지각된 가치를 3가지의 소비시간 대비, 지출비용 대비, 방문 노력 대비 관광의 가치를 중심으로 영향관계를 연구하였으며, 양성수(2009)는 역시 투입한 시간, 노력, 기회비용, 비용 등에 대한 지각된 가치를

중심으로 연구를 진행하였다. 김홍미(2008)는 한국관광의 가치를 감성적, 기능적, 총체적 가치의 3가지로 분류하였다. 또한 김판영과 김문성(2010)은 지각된 가치의 측정을 감성적, 기능적 그리고 사회적 차원의 3가지 가치로 측정하였다.

Sheth et al.(1991)은 소비가치는 어떤 주어진 상황에 차등적으로 기여하고, 독립적이며, 소비자의 선택은 여러 가지 소비가치들의 기능이라 하였으며, 총 5가지로 구성한다고 하였다.

(1) 기능적 가치

기능적 가치(functional value)는 실용적이고 기능적이며 물리적인 성과에 형성된 지각된 효용이다.

관광 분야에서는 Williams & Soutar(2009)는 관광에 대한 흐름을 파악하는 정보에 대한 필요성을 느끼고, Sweeney & Soutar(2001)의 구성요인들을 가지고 연구하였다. 그리하여 관광산업에서는 위험과 안전문제를 최소화하는데 필요한 기능적 가치가 중요하며 여행안내자는 편리함, 연락, 속도, 효율성 및 행정적인 도움으로 기능적 가치를 여행객에게 제공할 수 있다고 하였다.

기능적 가치는 지각된 품질 또는 서비스를 소비자가 사용하는데 있어서 예상되는 성과로부터 얻어진 유용성이며, 가격과 품질로 두 가지로 나누어 가격은 지각된 단기적이고 장기간의 비용의 인하로 인해 나타내며, 품질은 제품의 지각된 품질과 예상된 성과로부터 나타난다고 하였다(Sweeney & Soutar, 2001). Lee et al.(2005)은 DMZ 관광에 있어서 기능적 가치는 방문에 대한 가격적인 합리성이라고 하였다.

(2) 사회적 가치

사회적 가치(social value)는 구체적인 사회적 그룹들의 연합이며, 그 사회에서 부정적이고 긍정적으로 지각되는 효용이다. 그리고 사회적 가치는 소비자가 서비스를 이용할 때 다른 사람들과 연결되어있는 느낌을 받을 때 얻어지는 가치라고 정의하였다(Sheth, Newman & Gross, 1991).

Warner & Lunt(1941)는 사회적 가치란 사회계층에 관한 연구에서 영향을 받은 개념으로서, 사회적 가치란 제품을 소비하는 특정집단과 관련된 소비가치로서 선택대안이 가지는 전형적인 인구 통계적 특성 집단, 사회경제적인 차별 집단, 문화인류학적인 차이를 지닌 집단에 대한 귀속감과 소속감을 표출시켜 줄 수 있는 가치이다.

(3) 감성적 가치

감성적 가치(emotional value)는 감성적으로 유발시키거나 감성 상태에서 영향을 미침으로써 나타나는 지각된 효용이다.

감성적 가치는 소비자에게 서비스가 발생시키는 감성적·정서적 상태로서 경험으로 인하여 의도치 않게 발생되어진 무의식적인 감성은 소비자의 의사결정 과정에 있어 매우 중요한 역할을 수행하는 것이다(Sweeney & Soutar, 2001).

감성적 가치는 모험관광 체험에서 흥분이나 기분 전환을 이끌어 내지만 종종 두려움과 망설임, 걱정과 같은 부정적인 부분도 이끌어낸다. 따라서 모험관광 체험에서 감성적 가치는 핵심요소가 될 가능성이 높다고 하였다. 또한 사회적 가치는 관광 여행에서 사람들 간의 상호 작용과 같은 사회적

가치를 창출할 수 있다고 하였다.

그리고 감성적 가치는 제품에서 나타나는 느낌이거나 정서적인 상태를 말한다. 사회적 가치는 사회적 자아개념을 강화해주는 제품의 능력으로 인해 나타난다고 하였다(Sweeney & Soutar, 2001). Lee et al.(2005)은 DMZ 관광에 감성적 가치는 방문의 기쁨, 다른 관광지와 비교한 느낌, 즐거운 느낌으로 설명하였다.

(4) 지적 가치

지적 가치는 지식에 대한 욕구가 호기심을 유발하거나 신기함을 제공해서 만족함으로서 나타나는 지각된 효용이다. 지적가치는 관광에서 신기성과 새로운 지식 추구는 모험 여행에 대한 중요한 동기이며, 탐험적이고 신기성 및 다양성 추구행동에 대한 관광객의 욕구로 인해 많은 모험 관광 상품에서 핵심 요소이다(Sheth et al., 1991).

(5) 조건적 가치

조건적 가치(Conditional Value)는 소비자가 어떤 특수한 상황이나 환경으로 인해 선택에 직면했을 때 구체적인 상황이 아니면 상황들의 설정에 따라 나타나는 결과에 대한 지각된 효용이다(Sweeney & Soutar, 2001).

이러한 상황이나 환경들은 소비자의 구매선택 행동에 영향을 미쳐 소비자로부터 평소 계획했던 제품이 아닌 다른 제품을 구매하도록 부추기기도 한다. 이런 조건적 가치는 사회적 가치와 기능적 가치를 포함하고 있는데, 조건적 가치는 제품 속성에 대한 소비자의 중요도 평가와 신념을 일시적으로

로 변화시키려는 상황 때문에 발생한다.

Sheth et al.(1991)과 Sweeney & Soutar(2001)는 5가지의 구성요소들은 지각된 가치의 이론적 구조를 발전시키는데 기여하였으며, 소비자 의사결정의 과정과 선택행동을 이해하기 위해 구매 전과 구매 후에도 소비자가 제품에 대한 가치의 인식을 별도로 측정해야 한다고 하였다.

이를 바탕으로 최근 연구에서는 기능적, 정서적, 사회적 이점으로서 비즈니스 서비스에서 하나의 지각된 가치 구성 요소로 사용하여 비즈니스 시장에 체계를 확장하고 있다(Candi & Kahn, 2016).

정윤희(2008)는 패키지관광 상품에 대한 지각된 가치를 Sweeney & Soutar(2001) 및 Lee et al.(2005)의 지각된 가치의 구성요소를 이용하여 만족도와 재구매 의도를 설명하였다. 이종철(2016)은 여행상품에 대한 지각된 가치를 감성적 가치, 인지적 가치, 사회적 가치로 구분하여 브랜드 충성도와 브랜드 인지도 구매행동에 미치는 영향에 관해 연구하였다. 이상의 선행연구를 중심으로 지각된 가치 측정차원은 다음의 <표 2-6>과 같다.

<표 2-6> 지각된 가치측정의 구성요소

연구자	지각된 가치측정의 구성요소
박경대(2007)	방문 노력·소비시간·지출비용 대비 관광의 가치
김홍미(2008)	감성적·기능적·총체적 가치
서현 외(2008)	획득·금전적·감성적 가치
정윤희(2008)	기능적·사회적·감성적·지적·조건적·전반적 가치
양성수(2009)	투입한 시간·노력·기회비용·비용 등에 대한 지각된 가치
김관영·김문성(2010)	기능적·감성적·사회적 차원
전형진 외(2011)	정서적·기능적 품질·기능적 가격·사회적 가치
이종설(2016)	사회적·인지적·감성적 가치
Zeithaml(1988)	저렴한 것, 얻고자하는 서비스, 지불한 가격에 대해 받은 것, 지불한 가격에 비해서 고객이 제공받은 품질,
Sheth et al.(1991)	기능적·사회적·감성적·지적·조건적 가치
Oh(2000)	가격비교, 금전적·가격·기대 가치
Sweeney & Soutar(2001)	기능적·감성적·금전적·사회적 가치
Lee et al.(2005)	감성적·기능적·전반적 가치
Ledden et al.(2007)	기능적·사회적·인식적·감성적·조건적 가치
Lee,Yoon & Lee(2007)	감성적·기능적·총체적 가치
Cengiz and Kirbir(2007)	기능적·정서적 측면
Williams & Soutar(2009)	기능적·사회적·감성적·지적·조건적 가치
Candi & Kanh(2016)	기능적·정서적·사회적 가치

자료: 선행연구를 바탕으로 연구자 재구성

3. 빅데이터와 사회연결망분석

가. 빅데이터의 개념

현대사회에 정보통신기술의 발달과 함께 인터넷의 발전은 각종 온라인 매체와 소셜 미디어, SNS 등의 등장으로 인하여 데이터의 양이 폭증하는 빅데이터의 시대가 도래되었다(송은지, 2013). 온라인 사용자들은 인터넷에서 시공간을 초월하여 다양하고 대용량의 정보들을 텍스트, 사진, 동영상과 같은 수많은 다양한 형태로 데이터를 생성하고 저장하고 이를 사용자들 간에 공유하게 되었다. 이러한 현상은 대중들이 만들어 놓은 자료와 정보를 사용자의 필요와 목적에 맞춰 데이터화하고, 그 데이터를 분석하여 많은 사람들의 행동과 심리를 예측할 수 있는 단계까지 이르렀다(박영주, 2014).

빅데이터의 개념은 빅데이터에 대한 범위가 중요하게 작용한다. 정의가 상이할 경우 접근 방법에 있어 차이가 발생하므로 빅데이터를 정의할 때에는 범위를 명확하게 할 필요성이 요구된다(이만재, 2012). 또한, 빅데이터의 패러다임은 기업이나 공공기관뿐만 아니라 개인에게도 영향을 미치게 되고 이러한 영향은 미래사회에서도 영향을 줄 수 있기 때문에 중요성이 높아지고 있다(이영진·윤지환, 2014). Manyika & Chui(2011)는 빅데이터의 개념을 전통적인 데이터베이스 소프트웨어를 통해 저장, 관리, 분석할 수 있는 규모를 초과하는 데이터로 정의하였으며, 매킨지 보고서는 빅데이터에 대한 이슈를 적극적으로 제기하고 있으며, 데이터의 크기가 너무 커서 기존의 데이터베이스 도구로는 처리, 저장, 관리, 분석이 불가능한 데이터의 세트라고 했다(WIKI, 2009).

Douglas(2012)는 빅데이터를 실시간으로 다양한 형태의 거대한 양의 데이터가 생성되는 대용량, 속도, 다양성의 특징을 지닌 정보 자산이라고 정의하였다. Gartner(2011)는 빅데이터 요소를 Volume(규모), Velocity(속도), Variety(다양성)와 Complexity(복잡성)로 정의하였다. Philip(2011)은 빅데이터의 개념을 Volume, Velocity, Variety 3가지로 정의하였다.

이처럼 빅데이터의 정의가 다양한 형태로 이루어지고는 있으나 공통적으로 제시하는 것은, 기존의 기업에서 분석하는 데이터와 비교해 보면, 그 규모가 훨씬 방대한 형태의 데이터들의 집합이자 다양한 형태로 실시간 데이터가 수집되고 있고, 기존의 방식으로는 분석이 어려우므로 새로운 방식으로 분석을 해야 하는 데이터로 정의된다고 할 수 있다.

<표 2-7> 빅데이터의 개념

연구자	빅데이터의 개념
Manyika & Chui(2011)	데이터베이스 소프트웨어를 통해 데이터를 분석·관리·저장할 수 있는 규모를 초과하는 데이터로 정의
Gartner(2011)	빅데이터의 개념을 Volume(규모), Velocity(속도), Variety(다양성), Complexity(복잡성)로 정의
Philip(2011)	빅데이터의 개념을 Volume, Velocity, Variety로 정의
IDC(2011)	저렴한 비용으로 다양한 종류의 대용량 데이터로부터 가치를 추출하고 데이터의 초고속 발굴, 수집, 분석을 지원하도록 개발된 차세대 기술 및 아키텍처라고 정의
Douglas(2012)	다양한 형태의 방대한 양의 데이터가 실시간으로 생성되는 대용량, 속도, 다양성의 특징을 가진 정보 자산이라고 정의
송은지(2013)	빅데이터는 기존 데이터에 비해 데이터양이 너무 많아 기존의 방법이나 도구로 검색, 수집, 분석, 저장, 시각화 등이 어려운 비정형 또는 정형 데이터

자료: 선행연구를 바탕으로 연구자 재구성

나. 빅데이터의 특징과 역할

(1) 빅데이터의 특징

일반적으로 빅데이터는 데이터의 규모, 속도, 다양성의 복합적인 변화를 특징으로 하며, 기본적 특징인 3V(Volume, Velocity, Variety)를 바탕으로 하여 각기 다른 특징들을 추가하여 설명하고 있다(서정아, 2015).

첫 번째 특징인 데이터 규모는 데이터 관리나 분석에 대한 개념이며, 대략적으로 빅데이터는 적게는 테라바이트(TB)에서 많게는 페타바이트(PB) 정도 크기의 데이터 집합으로 지칭하는 것이 일반적이다(윤홍근, 2013). 그러므로 미디어, 위치 정보, 동영상 등과 같은 데이터 크기를 나타내는 것으로, 물리적인 크기뿐만 아니라 현재의 기술로서 처리 가능한 양인지 아닌지에 따라 빅데이터 인지 아닌지를 판단한다(강만모 외, 2012).

두 번째 특징인 데이터 속도는 데이터를 생성, 수집, 통합, 분석하고 활용하는 모든 단계에 있어서 속도가 중요하며, 궁극적으로 빅데이터에서는 분석결과를 실시간으로 활용하는 것을 추구한다는 것이 유사한 기술 트렌드와 빅데이터를 구별하는 것의 가장 큰 특징이자 성격이라 할 수 있다(박현길, 2012). 이는 사물정보와 스트리밍 정보 등 실시간성 정보의 증가와 이러한 실시간성으로 인한 데이터의 생성과 유통의 속도가 증가하였고, 이러한 데이터 활용을 위해서 데이터 처리 및 분석 속도가 중요시 되고 있다(윤홍근, 2013).

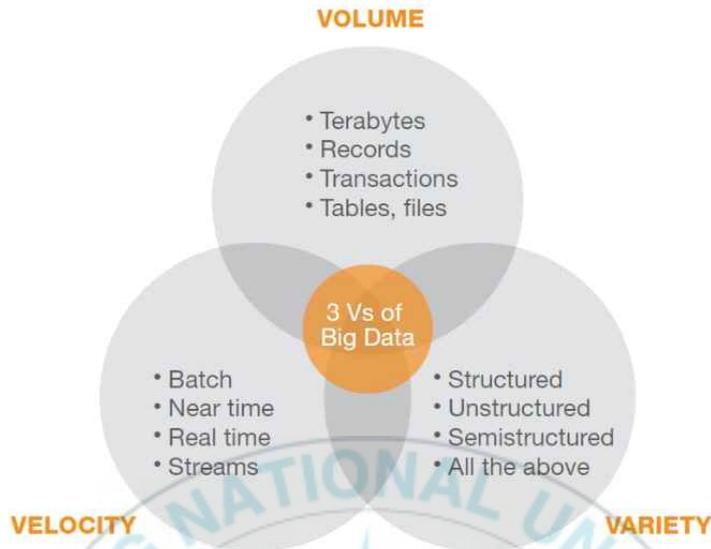
세 번째 특징인 데이터의 다양성은 빅데이터의 정형화 정도에 따라 분류되며, 정형 데이터란 고정된 필드로부터 저장된 데이터를 말하고, 반정형 데이터란 고정된 필드가 아닌 스키마를 포함하는 데이터를 말한다. 그리고 비정형데이터는 고정된 필드에 저장되어 있지 않은 데이터(텍스트, 이미지,

동영상 등)를 말한다(강만모 외, 2012).

IBM(2011)은 데이터 규모에 대해 여러 데이터(환경데이터, 금융데이터, 의료데이터, 감시데이터 등)를 현재 우리가 저장하고 있으며, 조직은 대용량의 데이터에 직면하고 있고, 데이터를 처리 비율은 점점 줄어들어 조직은 이런 데이터 관리 방법을 몰라 이 데이터에 압도될 것으로 판단되어지니, 적절한 기술 플랫폼을 통해 더 나은 정보를 얻기 위해 고객 및 시장, 비즈니스에 대한 모든 데이터를 분석해야한다고 하였다. 또한 증가하는 데이터양을 빠르게 처리하기 위해서는 데이터의 속도가 중요하며, 빅데이터를 효율적으로 사용하려면 앞서 언급한 데이터의 특징인 데이터 다양성 및 데이터 규모가 정지 상태가 아닌 이동 중에도 분석할 수 있어야 한다고 주장하였다.

데이터의 다양성은 스마트, 센서 및 사회적 협업 기술이 빠른 속도로 증가함에 따라 기업이 가진 데이터는 전통적 유형인 관계형 데이터뿐만 아니라 웹 로그(web log), 웹 페이지(web page) 파일과 같은 반정형 데이터 및 전자 메일, 검색 기록, 소셜 미디어와 같은 비정형 데이터까지 모두 포함하기 때문에 복잡해졌으며, 이에 따라 기업은 빅데이터의 기회를 이용하려면 모든 데이터의 유형을 분석할 수 있어야 한다고 하였다.

이후에는 Gartner(2012)는 데이터 복잡성을 추가하여 외부 데이터의 활용, 데이터 종류의 확대로 관리 대상이 증가하고 처리 및 관리의 복잡성이 심화되고 있다고 하였다.



출처 : Google

<그림 2-1> 빅데이터의 3V 특징

<표 2-8> 빅데이터의 4가지 구성요소

구 분	주요 내용
규모의 증가 (Volume)	기술 발전과 IT의 일상화가 진행되면서 매년 디지털 정보량이 폭증하여 제타바이트 시대로 진입
속도 증가 (Velocity)	스트리밍 정보, 사물정보(모니터링, 센서) 등 실시간성 정보 증가 데이터의 생성, 이동(유통) 속도의 실시간성 증가 대규모 데이터를 처리하고 가치 있는 실시간 현재정보 활용을 위해 데이터 처리와 분석 속도가 중요
다양성 증가 (Variety)	위치, 로그기록, 소비, 소셜, 현실데이터 등 데이터의 종류가 증가 멀티미디어 등의 텍스트 이외의 비정형화된 데이터 유형의 다양화
복잡성 증가 (Complexity)	데이터 저장방식의 차이, 구조화되지 않은 데이터, 중복성문제 등 외부 데이터 활용으로 데이터 종류의 확대, 관리대상의 증가, 데이터 처리 및 관리의 복잡성이 심화되고 새로운 기법 요구

출처: 정지선(2011), 빅데이터의 새로운 가능성과 대응 전략

(2) 빅데이터의 역할

정지선(2011)은 미래 환경에서의 빅데이터 역할이라는 선행연구에서 4가지 요인을 제시하였다.

첫째, 융합에 의한 창조력으로 다른 분야와의 융합을 통한 새로운 가치 창출과 컨버전스 분야의 데이터 분석으로 안정성 확보와 시행착오의 최소화, 빅데이터 활용으로 새로운 융합시장 창출을 기대효과로 나타내었다. 둘째, 리스크에 대한 대응력으로 정보의 패턴분석으로 이상신호와 위험징후를 사전에 포착할 수 있고, 최근의 이슈와 트렌드를 사전에 인지하고 분석하여 신속한 의사결정과 대응을 할 수 있게 되었다. 셋째, 스마트화에 대한 경쟁력으로 방대한 양의 빅데이터 분석으로 상황의 인지와 인공지능 서비스, 소셜 등의 오피니언분석 기반으로 사용자가 최적선택을 할 수 있도록 지원해 준다. 이는 제품 경쟁력 확보와 미래 트렌드 변화를 분석하는 데 유리하다. 넷째, 불확실성에 대한 통찰력으로 현실세계의 사회현상을 빅데이터를 통한 패턴분석과 미래를 전망하였고, 여러 가지 상황에 대한 시뮬레이션과 다각적으로 분석한 통찰력을 제시한다고 하였다.

빅데이터는 모바일 스마트혁명의 핵심 자원이다. 혁신과 경쟁력 강화 및 생산성 축진의 4차 산업의 품질경영활동으로서 빅데이터 가치창출 효과로 기대하고 있다.

나아가 빅데이터의 수집, 분석, 가공능력이 경쟁력과 경제적 자산의 주요 척도로서 부각되고 있다. 사회발전 속도의 가속화와 위험요인과 복잡성 증가로 신속히 환경변화를 감지할 필요성이 미래 환경에서 빅데이터는 신 가치창출의 역할로서 볼 수 있는 융합, 리스크, 스마트, 불확실성 등에 대응하는 역할과 기회창출을 품질경영활동에 적용하고 있다.

<표 2-9> 빅데이터의 역할

미래사회의 특성	빅데이터의 역할	
불확실성	통찰력	패턴분석과 미래전망(현실세계, 사회현상의 데이터 기반) 다양한 가능성에 대한 시나리오 시뮬레이션 다양한 시나리오로 상황 변화에 유연하게 대처 다각적으로 상황이 고려된 통찰력제시
리스크	대응력	소셜, 환경, 모니터링 정보의 패턴분석 통한 이상신호, 위험징후 포 착. 이슈를 사전에 인지하고 분석하여 빠르게 의사결정해서 실시간 대응지원. 기업과 국가의 경영 투명성 제고 및 낭비요소 절감
스마트	경쟁력	방대한 데이터 분석을 통한 상황인지, 개인화 지능화 서비스, 인공지 능 서비스 가능 제공확대 분석을 통하여 트렌드 변화 제품의 경쟁력 확보 소셜분석, 신용, 평가, 평판분석을 통한 최적의 선택을 지원
융합	창조력	다른 분야와의 결합을 통하여 새로운 가치 창출(의료정보, 환경정보, 자동차정보, 건물정보 등) 복잡한 상관관계, 인과관계 컨버전스 분야의 데이터 분석을 통해 시 행착오 최소화, 안전성 향상 방대한 데이터를 활용하여 새로운 융합시장 창출

자료: 선행연구를 바탕으로 연구자 재구성

다. 빅데이터와 사회연결망분석 기술과 기법

(1) 빅데이터 분석기술

대용량의 데이터를 가지는 빅데이터의 처리과정 흐름은 [생성-수집-저장-분석-표현]의 5가지 처리 전 과정을 거친다(송태민·송주영, 2013). 빅데이터 처리 기술과정은 다음의 <표 2-10>과 같다.

<표 2-10> 빅데이터 분석기술 과정

흐름	영역	개요
소스	내부데이터	파일 관리시스템, 데이터베이스
	외부데이터	파일 멀티미디어 스트리밍
수집	크롤링(Crawling)	데이터의 수집(검색엔진 이용)
저장	Extraction Transformation Loading	소스데이터 적재, 변환, 전송, 추출
	NoSQL 데이터베이스	비정형데이터의 관리
	스토리지(Storage)	빅데이터의 관리
처리	서버(Servers)	초경량 서버처리
	맵 리듀스(Map Reduce)	데이터의 추출
분석	프로세싱(Processing)	다중 업무처리 분석
	Nero Linguistic Programming	자연어의 처리
	Machine Learning	기계학습을 통해 데이터의 패턴발견
	Serialization	데이터 간의 순서화표현
표현	Visualization	데이터를 도표나 그래프로 표현
	Acquisition	데이터 획득 및 재해석

출처: 송태민·송주영(2013). 빅데이터 분석방법론

(2) 빅데이터 분석기법

빅데이터의 분석기법들은 기계학습, 통계학, 데이터 마이닝 분야에서 사용되던 기법들이다.

이 분석기법의 알고리즘을 대용량 데이터를 위한 처리에 맞게 개선하여 빅데이터 처리에 적용시켰다(김휴찬·김미정·신현정, 2014).

소셜 네트워크나 SNS에 따른 비정형 데이터의 증가로 인해 빅데이터 분석기법들은 데이터마이닝(Data Mining), 텍스트마이닝(Text Mining), 오피니언마이닝(Opinion Mining), 사회연결망분석(Social Network Analysis), 군집분석(Cluster Analysis), 현실마이닝(Reality Mining)으로 구분될 수 있다. 빅데이터 분석기법들은 다음의 <표 2-11>과 같다.

(가) 데이터마이닝

데이터마이닝은 정형화된 데이터를 분석하는 빅데이터 분석에 있어서 가장 대표적인 분석기법이다. 데이터마이닝은 자료로부터 숨겨진 연관성을 찾아 파악하고 새로운 가치 있는 정보를 만드는 분석기법으로 의사결정과정에서 적용할 수 있다(박영주, 2014). 데이터마이닝은 통계, 수학적 기술뿐만 아니라 패턴 인식 기술들을 이용하여 데이터 저장소에 저장된 대용량의 데이터를 조사 분석하여, 의미 있는 새로운 상관관계, 패턴, 트렌드 등을 발견하는 과정이라 할 수 있다(윤홍근, 2013).

(나) 텍스트마이닝

텍스트마이닝은 텍스트 분석 기법의 기본인 키워드(keyword) 분석에서 시작되며, 이를 기본으로 보다 발전된 형태의 분석을 수행하여 기존에 발견되지 못한 숨은 지식이나 정보를 찾아내는 특징이 있다. 텍스트마이닝은 문서군집, 문서분류, 정보추출, 문서요약 등에 사용된다. 이는 비정형 텍스트 데이터를 수집한 후 자연어 처리기술을 적용해 해당 텍스트 데이터에서 필요한 정보를 추출하고 또한 유의미한 정보로 가공하는 것이 가능해졌다. 이 기술은 아주 큰 용량의 텍스트 데이터에서 의미 있는 정보를 빼내고 해당하는 정보와 연관된 정보를 분석하는 등의 텍스트가 가지는 단순한 정보 이상의 의미 있는 결과를 얻을 수 있다는 장점이 있다(WIKI, 2014).

다른 정보와의 연계성을 파악하고 분류 또는 군집화하며 요약화 함으로써, 빅데이터에 숨겨진 의미 있는 정보를 발견하는 분석기법이다(송태민·송주영, 2013). 빅데이터 분석기법 중 텍스트마이닝은 비정형 혹은 반정형의 특징을 지닌 자연어 텍스트를 특정한 목적에 유용한 정보 추출을 위해 분석하는 과정을 의미한다(Witten, 2005; Jusoh & Alfawareh, 2012; Inzalkar & Sharma, 2016). 경우에 따라서 분석 과정에 나타나는 어휘나 언어 사용 방식을 찾아내는 시스템을 의미하기도 한다(Sebastiani, 2002).

대표적인 비정형 데이터인 텍스트 데이터는 도서, 문서, 편지, 이메일, 신문 기사, 연설문, 광고문, 블로그, SNS, 웹페이지로부터 데이터를 얻을 수 있고, 텍스트마이닝은 이 데이터를 분석대상으로 정한다. 기본적으로 자연어 처리기술을 기반으로 텍스트에 나타나는 단어를 이용하여 분석하는데, 텍스트 요약, 문서검색, 정보검색과 같은 문서사용의 효율성을 위한 정보추출이 가장 기본적으로 텍스트마이닝을 통해 수행되고, 텍스트 분류, 문서

군집화, 언어 인식, 핵심문구 식별 등 문서의 유사성을 측정하는 방식으로 확대되기도 한다(Khan & Kanth, 2016).

텍스트마이닝은 대량의 텍스트를 활용할 수 있어, 기존 연구들이 연구자의 역량에 의존하여 일부의 텍스트에 대해 제한적으로 접근하므로 이전에는 볼 수 없었던 숨은 지식을 찾아내는 것에 연구방법으로서의 핵심적인 가치가 있다(이영재·이성수, 2011; 김미아·송민, 2012; 정다미 외, 2013; Mahgoub, Rösner, Ismail & Torkey, 2008; Wisnowski, Rushing & Karl, 2015). 특히 정확하게 파악하기 어려운 사회현상 및 이슈와 관련된 다양한 논의 및 그 변화를 파악하는데 유용하게 사용된다(정다미 외, 2013; 원진영·김대곤, 2014; Hale, 2005; Jusoh & Alfawareh, 2012).

(다) 오피니언마이닝

오피니언마이닝은 특정 상품이나 서비스에 대한 소비자 반응, 시장규모 예측, 구전효과 분석 등에 이용되는 기술이다. 소셜 미디어에 게시되어 있는 정형 및 비정형 텍스트의 성향을 판별하는 기술이다. 오피니언이 긍정, 부정, 중립인지의 선호도를 표현하여 데이터 마이닝의 정확성을 발전시킨다(WIKI, 2014). 웹사이트와 소셜 미디어에 나타난 여론과 의견을 분석하고 실질적으로 유용한 정보로 재가공하는 분석기법으로 여론분석을 목적으로 하는 빅데이터 분석기법이다(윤홍근, 2013).

(라) 사회연결망분석(Social Network Analysis)

사회연결망분석은 연결구조와 연결강도 등을 바탕으로 사용자의 명성 및 영향력을 측정하여 사회연결망상의 연결중심이나 허브 역할의 사용자를 발견하여 메시지의 전파경로 파악과 영향을 받는 사용자 발견 등에 활용된다(김휴찬 등, 2014). 또한 소셜 미디어에 올라오는 글과 사용자를 분석해 소비자의 흐름이나 패턴 등을 분석하고, 판매나 홍보에 적용한다(윤홍근, 2013). 사회연결망분석은 텍스트마이닝 기법에 의해 이루어지고 있으며, 확산된 내용과 함께 연결의 맥락을 파악하는 것이 중요하다(윤홍근, 2013).

사회연결망분석(SNA: social network analysis)이란 다양한 행위자들의 상호작용에 의해서 형성된 관계 연결망을 통해 인간의 행위나 사회의 구조 효과를 설명하고자하는 방법론으로, 개체들 간의 관계를 시각적으로 표현하여 그것들의 연결 구조 및 연결 상태의 특성을 계량적으로 분석하는 기법이다. 사회연결망 구조를 구성하는 링크의 수와 노드, 연결 유무 등을 바탕으로 생성된 행렬을 기반으로 하여 여러 가지 분석방법으로 연결망 분석을 실행할 수 있다.

사회연결망은 매트릭스(matrix)로 분석되는데, 행과 열이 만나는 셀에 특정 값을 나타내어 관계를 표시하는 방법이다. 서로간의 관계가 성립하면 '1', 관계가 없다고 보면 '0'으로 표시한다(원효진, 2014).

사회연결망분석은 연결망 분석수준에 따라서 네트워크 수준의 분석, 노드 수준의 분석, 네트워크의 내재된 속성 분석, 중심성 분석, 하위집단 수준의 분석, 예고 네트워크 수준의 분석이라는 6가지 형태로 접근이 가능하다(이수상, 2012). 이 중에서 국내 사회연결망 연구들에서 주로 사용되고 있는 주요한 측정 지표에는 밀도 분석(density analysis), 중심성 분석

(centrality analysis), 집중도 분석(centralization analysis) 등이 있다(김성희·장로사, 2010).

첫째, 밀도가 높은 연결망일수록 노드 간의 연결 관계가 높으며 이는 서로 간의 교류가 많다는 것을 의미한다. 즉, 연결망상에서 사람들이 서로 알거나 혹은 연결되어 있는 정도를 의미하는 것으로(Brass, 1995), 한 연결망에 참여한 사람들 간의 관계응집 정도이다. 이러한 밀도 값은 연결 관계의 복잡성, 결속도, 응집성 등을 나타낸다. 밀도 값은 0과 1의 사이에 위치하며, 밀도 값이 1로 나타난다는 것은 모든 노드와 연결되어 있다는 의미로 밀도의 값이 1에 가까워질수록 연결망에서 연결 관계의 복잡성, 결속도, 응집성 등이 높다고 할 수 있다. 연결정도는 한 점과 직접적으로 연결된 점들의 수를 말하며, 포괄성은 한 연결망 내에서 서로 연결된 행위자들의 수를 뜻하고, 한 연결망의 전체 수의 점에서 격리된 점들의 수를 제외하고 남은 수로 계산된다(김성희·장로사, 2010).

둘째, 중심성(centrality)이란 중앙성이라고도 불리며, 하나의 행위자가 연결망 전체에서 중심에 위치하는 정도를 표현하는 지표이다.

중심성은 다양한 방법으로 측정되고 있는데, 대표적인 것으로 연결정도 중심성(degree centrality), 근접 중심성(closeness centrality), 매개 중심성(betweenness centrality)이 있으며(Freeman, 2005), 최근 더욱 발전시킨 지표로써 위세 중심성(prestige centrality), 페이지 랭크(page rank) 등의 중심성 지표를 활용하고 있다(이수상, 2012; 김용학, 2016). 연결정도 중심성은 다른 점과의 연결된 정도를 중요시하며, 연결망에서 한 점에 연결되어 있는 점들의 합계를 말한다.

셋째, 집중도(centralization)란 중심화라고도 하며, 전체 연결망의 형태가 어느 정도 중앙에 집중되었는지를 보여주는 개념이다(손동원, 2013). 중심성이 개별 노드의 중심적 위치에 초점을 둔다면, 집중도는 연결망 전체에

초점을 맞추면서 하나의 연결망이 얼마나 중앙 집중적인 구조를 가졌는지에 주목한다(손동원, 2013). 일반적으로 집중도를 측정하는 방법은 가장 중심에 위치한 점과 다른 모든 점들이 갖는 중심성 점수들 간의 차이를 각각 구하여 이를 모두 합한 후, 이것을 논리적으로 가능한 최대값으로 나누는 것이다. 연결망 분석에서는 이를 %로 표시하여 제시한다(이상균, 2003).

이와 같은 집중도 측정의 주요 지표에는 매개 집중도, 연결정도 집중도, 군집 집중도 등이 있다(김용학, 2014).

사회연결망은 점과 선을 통해 시각화할 수 있으며, 그 의미를 탐색하는데도 이러한 시각화는 매우 효과적이다(Krackhardt, 1992).

시각화는 숫자에 의해 그 값이 측정되어지기 때문에 다양한 분석도구들을 통해 단시간에 복잡한 구조 분석에 대한 결과 값을 얻을 수 있으면서, 시각적으로도 표현이 가능해 바로 직관적인 확인이 가능하다. 이를 위해 사회연결망분석 도구를 통해 의미 탐색을 위한 사회연결망분석 값을 도출하고, 시각화하는 도구들을 다수 활용할 수 있다. 이는 크게 수집된 데이터 분석을 위한 분석도구와 분석데이터를 시각적으로 표현하는 시각화 도구로 구분된다(이수상, 2012).

본 연구에서 채택되고 있는 UCINET6는 프리먼(Freeman)이 개발한 연결망 분석기법들을 다양하게 활용할 수 있는 종합프로그램이다. UCINET6는 자아중심 연결망, 중심성은 물론이고 응집력, 하위집단, 역할 및 위치, 중앙과 변방 등 대부분의 요소들을 분석할 수 있고 자유롭게 자료를 변환할 수 있는 기능도 제공한다(손동원, 2013). 또한 2-모드 스켈링, 유사도 및 비유사도, 군집(Cluster) 등 연결망 분석과 밀접히 연관된 통계기법도 사용할 수 있다(김용학, 2014).

(마) 군집분석

비슷한 특징을 가지는 개체를 합쳐가며 최종적으로 유사 특징의 군집을 발견하는데 사용된다(김휴찬, 2014). 군집분석은 각 유사성을 측정하여 유사성이 높은 대상 집단을 분류하고, 군집에 속한 객체들의 유사성과 서로 다른 군집에 속한 객체 간의 상이성을 규명하는 통계기법으로 명확한 기준이 존재하지 않거나 밝혀지지 않은 상태에서 다양한 특성이 지닌 대상을 집단으로 분류하는데 사용된다(윤홍근, 2013).

(바) 현실마이닝

사람들의 행동패턴을 예측하기 위하여 관련된 정보를 휴대폰이나 GPS를 통해 얻고 분석하는 기법으로 휴대폰과 같은 모바일 기기들이 제공하는 정보를 기반으로 인간관계와 행동패턴을 유추할 수 있다는 장점이 있다(윤홍근, 2013).

<표 2-11> 빅데이터 분석기법

구분	내용
데이터마이닝 (Data Mining)	데이터베이스, 대용량의 데이터 등에서 숨겨진 지식, 새로운 규칙, 기대하지 못했던 경향 등의 유의미한 정보를 찾아내는 과정 정보의 연관성을 데이터마이닝을 통해 파악하여 가치 있는 정보를 생성하여 의사결정에 적용
텍스트마이닝 (Text Mining)	자연어(NLP)로 구성된 비정형 텍스트 데이터에서 패턴이나 관계를 추출하여 의미와 가치 있는 정보를 발견하는 마이닝기법 자연어 처리기술에 기반한 반정형, 비정형 텍스트로부터 사용가능한 정보추출분석기법
오피니언마이닝 (Opinion Mining)	소셜 미디어와 웹사이트에 나타난 의견과 여론을 분석하여 실제로 유용한 정보로 재가공하는 기술
사회연결망분석 (Social Network Analysis)	마케팅 분야뿐만 아니라 사회의 흐름, 여론과 트렌드 변화의 추이를 파악하는 소셜 미디어 시대의 새로운 마이닝기법 소셜 미디어의 글과 사용자를 분석해 소비자의 패턴이나 흐름 등을 분석하고 판매나 홍보에 적용 주요한 측정 지표에는 밀도 분석(density analysis), 집중도 분석(centralization analysis), 중심성 분석(centrality analysis)
군집분석 (Cluster Analysis)	대량의 데이터 군 큰 묶음 하나를 통계기법을 활용하여 유사성을 지닌 작은 여러 개의 묶음으로 분류하는 방법
현실마이닝 (Reality Mining)	소비자의 행동패턴을 예측하기 위해서 기기를 통해 사회적 행동과 관련된 정보를 획득하고 분석하는 기법

출처: 윤홍근(2013), 문화마케팅 입문

이처럼 다양한 분석기법들이 있지만 본 연구에서는 비정형 데이터들 중에서 텍스트를 연구 자료로 사용하므로 텍스트마이닝을 이용한 사회연결망 분석에서 시멘틱 네트워크 분석 기법을 활용하여 연구결과를 도출하였다.

라. 사회연결망분석을 활용한 관광관련 연구

사회연결망분석은 국가, 정부, 지방자치단체, 기업, 서비스, 공공분야 등에서 다양하게 활용되고 있으며, 정치, 경제, 문화, 사회, 의료와 같은 다양한 영역에 걸쳐 주목한 만한 성과를 나타내고 있다(김동완, 2013).

사회연결망분석은 우리 사회 다양한 분야에서 여러 가지 방법으로 활용되고 있다. 빅데이터는 기업들이 고민해온 문제에 대한 해결책을 제공해주며, 프로세스와 조직, 산업 전반뿐만 아니라, 사회 자체를 변화시킬 수 있는 새로운 방법까지 제시하고 있다(김동완, 2013). 빅데이터 활용의 장점은 기존의 저장된 자료와 실시간 저장되어지는 데이터의 분석이 가능하며, 비정형 데이터는 기존의 연구로 생각하지 못했던 다양한 분야에서 연구를 가능하게 해준다(박영주, 2014).

최근 사회연결망분석은 트위터, 페이스북과 같은 SNS 상에서 수집되는 정보를 통해 소비자 심리를 분석하는 기법으로 활용되어 기업의 매출과 수익을 늘려주는 마케팅 수단과 기업의 위험을 감소시키는 위기관리 수단으로 주목받고 있다(도해용, 2014). <표 2-12>는 많은 산업분야에서 트렌드 분석, 이슈분석, 평판분석, 이미지분석, 위기관리분석, 타겟 마케팅 등에 빅데이터를 통한 사회연결망분석이 잘 활용되고 있다는 것을 보여준다.

사회연결망분석의 성공적인 활용에 있어서의 관건은 양질의 데이터를 확보하는 것이며, 뛰어난 데이터의 수집과 분석기술을 통하여 많은 양의 데이터를 수집하였다 하더라도 양질의 데이터가 아니면 활용하기 위한 정보 획득을 할 수 없기 때문이다. 빅데이터 활용의 활성화를 위해서는 수집된 데이터의 필요성과 가치를 중시하는 공감대 형성과 적극적인 활용이 필요하다(서정아, 2015). 관광산업 분야에서도 빅데이터의 중요성이 강조되면서

빅데이터의 활용방안에 대하여 관심이 증가하고 있다. 2013년부터 ‘한국관광학회’와 ‘한국관광문화연구원’ 등에서도 관광객의 실태조사에 대한 데이터를 분석하여 관광산업에 대한 정책수립과 마케팅을 위해 관광 데이터 마이닝 대회를 개최하고 있다. 한국관광공사는 설문조사와 더불어 전문가의 의견수렴을 통한 온라인 빅데이터 분석으로 2015년 ‘한국관광명소 100선’을 선정하였다.

<표 2-12> 사회연결망분석 활용방식

구 분	내 용
이미지 분석 (브랜드 분석)	특정한 브랜드나 제품에 대한 대중의 인식 분석 대중들의 제품이나 서비스에 대한 부정적, 긍정적 의견, 니즈 분석을 통한 이미지 포지셔닝 방안 제시
트렌드 분석	실시간 생성된 데이터 활용을 통한 대중들의 유행, 욕구, 기호 분석 제품이나 서비스의 경쟁력 제언에 활용
위기관리	다양한 논란 제기가 될 수 있는 제품, 서비스, 기업 등의 문제점을 분석하여 사전에 위기 방지와 피해의 최소화 소셜 미디어 모니터링을 통한 위험요소를 발견 후 체계적이고 신속하게 대응 가능
스토리텔링	스토리텔링을 통한 대중들의 유망 콘텐츠와 관심 콘텐츠의 유형분석 문화산업 분야, 제품, 서비스에서 대상에 관한 스토리 요소 보충 가능
타겟 마케팅	고객 데이터 분석과 기업 활동을 통한 표적시장 공략에 이용 선호도 조사를 통한 타겟 마케팅, 고객의 제품과 서비스 이용

자료: 선행연구를 바탕으로 연구자 재구성

사회연결망분석에 관련 연구는 몇 가지 사례를 보면 다음과 같다. 김재성(2012)은 SNS를 통하여 빅데이터를 수집한 후, 스마트 폰의 선택속성을

다차원 분석을 기반으로 텍스트마이닝으로 도출하여 선호도를 분석하였다.

공효순(2013)은 기존 CRM의 한계를 살펴보고자 호텔의 데이터베이스를 이용한 빅데이터를 활용하여 보다 효율적인 CRM을 제시하였다. 공효순·송은지(2015)는 사회연결망을 이용해 호텔이용 고객의 평판 분석하였으며, 실제 서울 시내 5성급 3개의 호텔을 분석하여 고객의 소리를 긍정과 부정의 평가요소를 가지고 호텔기업의 고객 평판의 개선안을 찾아내고자 하였다.

김용범(2015)의 연구에서 호텔 이용 고객의 경험과 선택속성의 중요도와 만족의 사이 관계는 리뷰를 통해 본질적으로 소비자 행동에 연결되어 있음을 확인하였으며, 사회연결망을 활용하여 호텔의 선택속성도 도출될 수 있음을 밝히고 있다. 또한 고객의 평판도 호텔 선택속성의 키워드인 만족과 불만족으로 만족도를 도출하였다.

관광학계에서도 온라인의 리뷰를 통한 감성분석, 사회연결망분석을 활용한 버즈모니터링, 정보탐색 연결망 연구 등의 다양한 사회연결망분석 활용 연구들이 시도되고 있다. 그러나 아직까지 국내 관광학 분야에서의 사회연결망분석 활용은 기초적 수준에 그치고 있기 때문에 좀 더 다양하게 온라인 빅데이터 수집과 분석도구를 활용한 실증적인 연구들이 필요하다.

사회연결망 환경은 정부를 포함한 공공부문에서의 혁신을 수반하는 대변화를 의미할 뿐만 아니라 기업 경영활동이며, 오늘날의 관광사업이나 관광 정책에 있어서 사회연결망분석을 얼마나 잘 활용하는지가 사업이나 정책에 있어 핵심적인 경쟁우위 요소이다. 그리고 국내관광 산업의 규모가 점점 커지면서, 관광기업에서 얻어지는 데이터도 지속적으로 축적되고 있는 실정이다. 따라서 이제는 이렇게 축적된 대용량 데이터를 활용하도록 할 수 있다(박상곤, 2014).

트립어드바이저(Tripadvisor)는 매달 5,000만 명의 여행객들이 방문하며, 여행객들의 100% 자발적인 리뷰와 참여로 유지되고 있는 소셜 사이트이

다. 트립어드바이저는 개개인의 이미지와 텍스트 리뷰, 즉 다양한 비정형 데이터 분석을 통해 여행의 지역별, 목적별, 취향별 등에 대한 개인이 선호하는 스타일을 분석해낸다. 비슷한 지역의 비슷한 연령대가 선호하는 여행 형태도 쉽게 파악할 수 있어 빅데이터 분석을 통해 각 여행자의 기호에 맞는 여행 상품을 제안하고 자세한 여행 정보를 함께 제공할 수 있게 되었다(삼정 KPMG경제연구원, 2012).

피서객이 몰리는 여름 휴가철마다 이슈가 되어 온 부산의 해운대, 송정 해수욕장의 피서객 수는 SK텔레콤이 휴대전화 사회연결망을 활용해 더욱 정밀하게 집계하였다(연합뉴스, 2017).

이와 같이 빅데이터를 활용하여 관광과 관련된 사업에 접목시키고 있으며, 최근에는 관광과 관련된 연구에도 빅데이터를 활용하고 있다. 오익근(2015)은 관광객들의 관광에 대한 인식을 SNS 및 빅데이터를 통해 분석하였으며, 조완섭 외(2015)는 온라인에서 충북관광에 관한 글들을 수집하고, 이를 텍스트마이닝 기법을 활용하여 분석함으로써 충북지역의 관광 활성화를 위한 홍보, 마케팅 전략과 관광 진흥정책을 수립의 방안을 제시하였다.

그리고 네트워크 텍스트 분석 기법을 적용하여 5성급 호텔 패키지 이용에 대한 태도 분석(안명숙 외, 2015), 소셜 빅데이터를 이용한 여행사 평가(공효순 외, 2015), 빅데이터를 활용한 템플스테이 키워드 분석(김보경 외, 2015), 온라인 빅데이터를 중심으로 사회연결망분석을 활용한 대구의 관광지 이미지 분석(서정아, 2015), 소셜 미디어 빅데이터 분석을 통한 캠핑에 대한 정부정책과 국민들의 인식 변화 분석(장미화 외, 2016), 빅데이터 분석을 통한 전통마을 관광 개선방안(정동현, 2016), 전라북도 중심으로 빅데이터 분석을 통한 지역관광 활성화 방안(박로운, 2017), 빅데이터 분석을 통한 중국 관광시장의 한국 관광인식 분석(한혜림, 2017) 등과 같이 빅데이터를 활용한 연구들이 진행되고 있다.

<표 2-13> 관광분야 사회연결망 연구현황

연구자	내 용
안명숙(2015)	특급 호텔 패키지 이용에 대한 태도 분석
공효순(2015)	소셜 빅데이터를 이용한 여행사 평가에 관한 연구
김보경(2015)	빅데이터를 활용한 템플스테이 키워드 분석
김용범(2015)	호텔 고객 경험의 선택속성 중요도와 만족 사이의 관계는 소비자 행동에 본질적으로 연결되어 있음을 확인
서정아(2015)	사회연결망분석을 활용한 대구 관광지 이미지분석
정동현(2016)	빅데이터 분석을 통한 전통마을 관광 개선방안
장미화(2016)	소셜 미디어 빅데이터 분석을 통한 캠핑에 대한 정부 정책과 국민들의 인식 변화연구
정동현(2016)	빅데이터 분석을 통한 전통마을 관광 개선방안
박로운(2017)	전라북도 중심으로 빅데이터 분석을 통한 지역관광 활성화 방안
한국문화관광연구원(2017)	‘외래관광객 실태조사’에서 빅데이터 분석을 통해 관광산업 진행을 위한 정책 제안
한혜림(2017)	빅데이터 분석을 통한 중국 관광시장의 한국 관광인식 분석
해운대구청(2017)	빅데이터 분석해 관광객의 인식과약 후 관광정책에 반영

자료: 선행연구를 바탕으로 연구자 재구성

Ⅲ. 연구방법

1. 연구문제

소비자의 요구에 부응하는 여행상품을 개발에 있어서 이용 대상별, 여행 목적별로 구분하여 개발을 하는 것을 고려해야 하며(김영길, 1998), 여행객의 여행상품 선택에 영향을 주는 여행상품의 구성요소는 중요하다(문명희, 2003). 그리고 인구통계학적 특성뿐만 아니라 여행 동기와 라이프 스타일에 따른 여행상품 선택요인의 차이가 있다고 하였다(천덕희, 2005). 또한, 선택요인은 관광객의 만족에 통계적으로 유의한 영향을 미치며 만족 수준이 높을수록 행동의도에 정적인 영향을 미치는 결과가 나타났다(김기석, 2017). 지각된 가치와 관련된 연구에서는 관광지가 관광객의 지각된 가치인 쾌락적 가치와 효용적 가치에 영향을 미쳤으며, 쾌락적 가치와 효용적 가치는 관광객의 관광 만족에 유의한 영향을 미쳐 관광객의 관광 만족이 행동의도에 유의한 영향을 주었다. 해외여행을 다녀온 관광객의 지각된 가치도 역시 만족도와 재구매 의도에 영향을 미치는 결과가 나타났다(전형규, 2008; 정윤희, 2008). 그리고 행동의도 중 하나인 온라인 구전은 구매의도에 주요한 영향력을 미치는 것으로 나타났다(이원희, 2009).

이러한 결과를 바탕으로 본 연구에서는 기존 연구의 결과에서 나타난 행동의도 중 하나인 온라인 구전을 통해서 나타난 빅데이터를 가지고 역으로 접근하여 부산 여행을 중심으로 여행상품 구성요소 및 지각된 가치를 분석한 후 국내 여행사들의 마케팅 방안에 대한 시사점을 도출하고자 하였다.

본 연구의 연구문제는 다음과 같다.

연구문제 I

부산 여행상품 형성에 영향을 미치는 명사 텍스트들과 여행상품의 요소들은 무엇인가?

1-1. 부산 여행상품 형성에 영향을 미치는 명사 텍스트들은 무엇인가?

1-2. 부산 여행상품 형성에 영향을 미치는 명사 텍스트들과 연관된 여행상품의 요소들은 무엇인가?

연구문제 II

부산 여행에 영향을 미치는 형용사 텍스트들과 지각된 가치의 요소들은 무엇인가?

2-1. 부산 여행에 영향을 미치는 형용사 텍스트들은 무엇인가?

2-2. 부산 여행에 영향을 미치는 형용사 텍스트들과 연관된 지각된 가치의 요소들은 무엇인가?

연구문제 III

부산 여행상품의 구성요소 및 지각된 가치에 연관된 실제 텍스트들의 네트워크 구조와 특성을 통해서 파악되는 부산 여행상품의 구성요소 및 지각된 가치의 형성구조는 어떠한가?

3-1. 부산 여행상품 및 지각된 가치의 네트워크 구조와 특성을 통해서 부산 여행상품의 구성요소 및 지각된 가치의 형성구조는 어떠한가?

3-2. 부산 여행상품 및 지각된 가치와 연관된 주요 텍스트들의 네트워크 구조에서 가장 영향을 미치는 텍스트들과 요소들은 무엇인가?

3-3. 부산 여행상품의 구성요소를 형성하는 유사집단 구분은 어떠한가?

2. 연구절차

본 연구는 빅데이터의 분석기법 중 하나인 텍스트마이닝과 사회연결망분석을 활용하여 부산 여행상품과 지각된 가치의 분석을 위하여 다음과 같이 연구절차를 진행하였다.

첫 번째로 키워드를 선정하는 방식이다. 빅데이터를 활용한 사회연결망 분석을 할 때 키워드 선정은 아주 중요하다. 데이터를 검색하고 수집할 때 키워드에 따라 연구자가 원하는 데이터를 얻을 수 있기 때문이다. 이를 위해 ‘부산 여행’, ‘부산 관광’, ‘부산 투어’를 선정하여 데이터양과 내용을 사전 검토한 결과, 본 연구문제의 해결에 데이터 수집이 가능하다고 판단되는 ‘부산 여행’으로 키워드를 선정하였다.

두 번째는 데이터의 수집대상과 데이터의 수집범위를 선정하는 것이다. 데이터가 공개되어 있고 양질의 데이터 수집이 가능한 포털 사이트인 네이버(Naver)와 다음(Daum)의 블로그, 카페, 뉴스 등 3개의 채널을 선정하였으며, 데이터 수집도구 중 하나인 TEXTOM을 활용하여 수집된 텍스트에 대한 키워드 분석을 실행하여 데이터를 수집하였다.

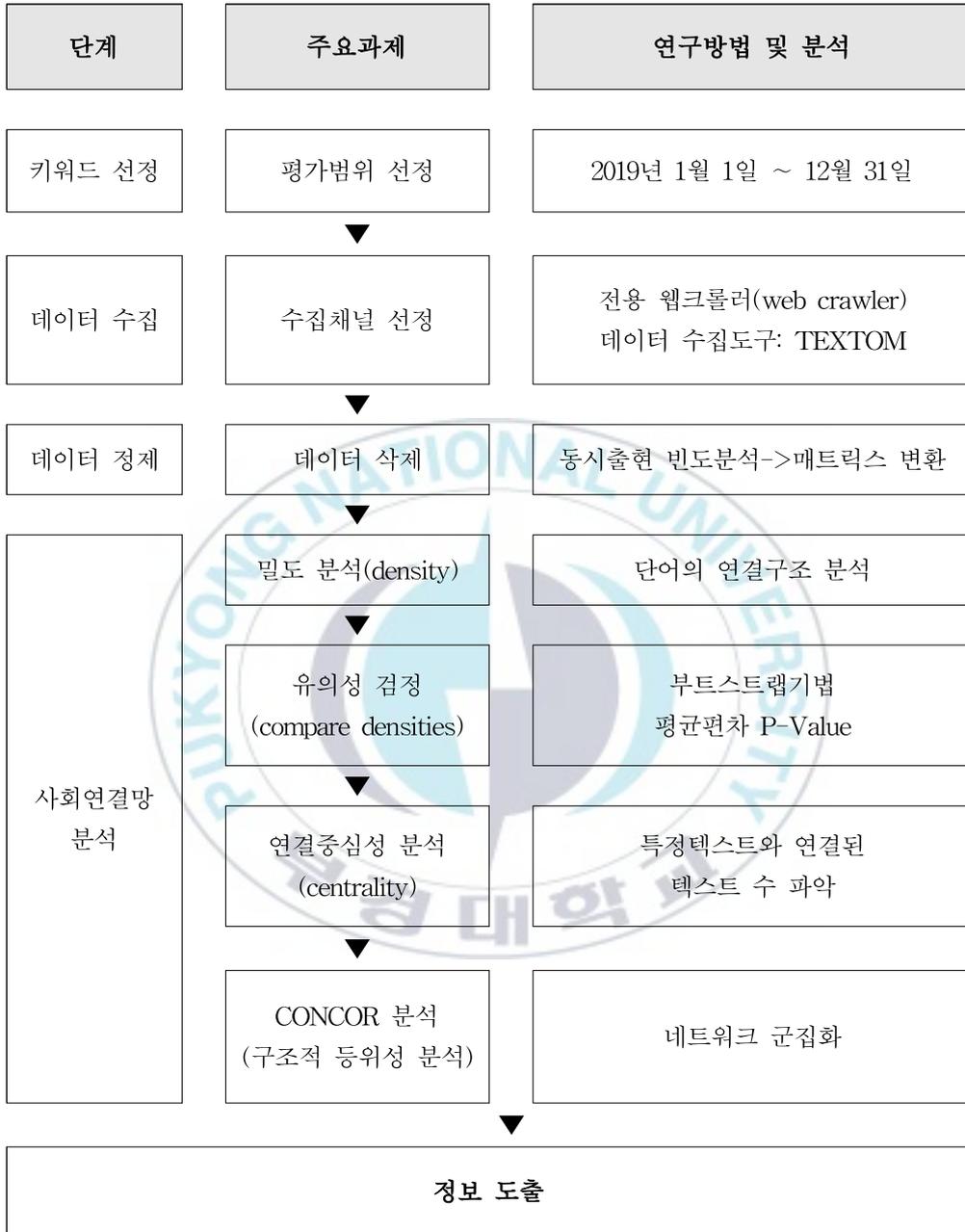
세 번째는 데이터의 정제이다. 수집된 데이터 중에서 연구 목적에 부적합하다고 판단되는 기호나 텍스트 등의 데이터를 삭제하였으며, 비정형 데이터의 형태로 데이터가 구성되어있기 때문에 형식이 따로 정해져 있지는 않았다. 그래서 가장 많은 빈도수를 가진 키워드를 중심으로 데이터를 코딩하였다. 또한, 여행상품의 특징 및 구성요소와 지각된 가치의 특징 및 구성요소에서 본 연구에 유용한 결과를 도출할 수 있는 텍스트를 선택하여 정제를 실행하였다. 최종적으로 정제된 데이터는 동시출현 빈도분석을 실행하여 상위 주요 텍스트들을 추출하여, 매트릭스로 변환시켜 사회연결망

분석이 가능하게 만들었다.

네 번째는 수집한 데이터의 사회연결망분석 및 시각화이다. 빈도 매트릭스를 활용해 UCINET6 프로그램으로 통계적 분석을 실행하고, NetDraw를 통해 분석결과를 시각화하였다.

다섯 번째는 연구결과에 따른 정보를 도출하여 연구의 결론 및 시사점을 제시하였다. 본 연구의 절차는 다음의 <그림 3-1>과 같다.





<그림 3-1> 연구절차

3. 데이터 수집대상과 범위 선정

가. 데이터 수집대상과 키워드 선정

본 연구의 연구문제를 해결하기 위한 데이터 수집 대상은 우리나라의 대표적인 온라인 포털 사이트인 네이버와 다음으로 선정하였다. 선정의 이유는 카카오톡(Kakaotalk), 인스타그램(Instagram), 페이스북(Facebook)과 같은 SNS는 데이터를 비공개하여 데이터 수집이 불가능하였으며, 데이터가 공개되어 있는 트위터(Twitter)는 본 연구에 대한 데이터가 수집 제공되지 않은 반면, 네이버와 다음은 한국인이 부산 여행의 정보를 검색하고 공유할 때 가장 많이 활용하는 일반적인 웹사이트로 판단되었기 때문이다. 그리고 부산을 다녀온 여행객이라면 누구나 온라인 구전활동을 할 수 있는 채널인 네이버와 다음을 선정하였다.

포털 사이트 네이버와 다음의 채널별 데이터 수집 기간은 많은 선행 문헌들이 소비자의 최근 동향을 파악하기 위한 기준을 1년으로 정하고 있어, 이를 바탕으로 2019년 1월 1일부터 12월 31일까지 1년간의 데이터로 한정하였다.

본 연구에 대한 데이터를 수집하기 위한 키워드는 ‘부산 여행’으로 선정하였다. 연구 목적에 적합한 결과를 도출하기 위해서는 키워드 선정이 중요한데, 이를 위해 선행논문을 참고하여 ‘부산 여행’, ‘부산 관광’, ‘부산 투어’를 시범적으로 검색하였다. ‘부산 관광’은 데이터양은 많았지만, 부산 여행에 다녀온 관광객들의 글 외에도, 부산 정부정책이나 학문적인 글이나 광고성 글들도 많았다. 그리고 ‘부산 투어’는 여행상품의 전반적인 각 구성요소를 포함하기 보다는 부산에서 관광활동 부분으로 많이 사용하고 있었으며, 많은 데이터가 수집되지 않았다.

결과적으로 부산 여행상품과 부산 여행객의 지각된 가치는 여행이라는 행위와 깊이 관련되어 있으며, 다른 논문들과 비교해 데이터양도 충분하다고 판단되어 ‘부산 여행’으로 선정하였다.

데이터 수집도구는 빅데이터 수집도구를 설명하는 빅데이터를 활용한 논문(오익근, 2015; 안명숙, 2015; 김보경, 2015; 서정아, 2015; 장미화, 2016; 한혜림, 2017)을 참고하여 대용량의 웹 또는 SNS 데이터를 효율적으로 저장, 처리, 분석할 수 있는 데이터 기반의 수집, 분석의 플랫폼이며, 데이터를 채널 별로 자동 수집하여 정제, 매트릭스 생성까지 일괄처리 해주는 데이터 처리 솔루션으로 크롤링(crawling)방식을 이용하는 TEXTOM을 활용해서 키워드 분석을 실행하여 데이터를 수집하였다.



나. 수집 데이터의 분류 범주 선정

본 연구에서는 기존의 연구에서 생성된 요소들을 활용하여 키워드 분석을 실행한 선행연구 논문을 참고하여, 포털사이트 데이터를 통해 부산 여행상품의 요소들과 부산 여행에서 생성된 지각된 가치의 요소들 사이의 관계를 알아보고자 하였다.

여행상품은 여행상품의 생산자인 여행사가 이미 완성된 다양한 이질적인 요소들을 결합하여 하나의 상품으로 생산하기 때문에 명사로 구성되어 있는 요소들이다. 그리고 지각된 가치는 무엇을 주었고 무엇을 받았는지에 대한 지각들을 기반으로 하여 제품의 효용성에 대한 여행객의 종합적인 평가이며, 여행상품의 요소들처럼 명사로 설명하는 것보다는 여행객의 느낌을 가치로 표현하는 것이기 때문에 형용사로 구성되어있는 요소들이라고 할 수 있다.

본 연구는 포털사이트의 명사 텍스트를 가지고 부산 여행상품의 요소를 찾아보았으며, 여행상품과 관련된 선행연구에서 사용된 구성요소들을 모두 참고하여 여행상품의 분류 범주를 관광지 관련사항, 숙박시설, 교통수단, 식사문제, 관광활동, 쇼핑활동, 여행관련 서비스 등 총 7개의 요소로 분류하여 연구하였다. 그리고 형용사 텍스트는 부산 여행에서 형성된 지각된 가치를 알아보았으며, 지각된 가치의 요소 분류는 앞서 선행연구를 바탕으로 Sweeney & Soutar(2001)의 구성요소를 가지고 관광 분야에서 지각된 가치에 대한 연구를 한 Williams & Soutar(2009)의 지각된 가치의 구성요소인 기능적 가치, 감성적 가치, 사회적 가치, 지적 가치, 조건적 가치의 5가지 요소로 분류하여 연구하였다.

<표 3-1> 여행상품 관련 연구 수집 데이터의 분류 범주

구분	세부적인 요인
관광지 관련사항	관광지 편의성, 관광지 매력성, 관광지 접근성, 접대, 환대 등을 포함한 속성들이며 역사, 문화, 유물·유적, 기후, 예술, 오락, 스포츠, 동식물, 해양·해안, 지역특정산업 등
교통수단	한 곳에서 다른 한 곳으로 이동시키는 모든 교통수단으로 버스, 철도, 선박, 항공기 등
숙박시설	숙박시설 내 각종 부대시설. 관광객들이 숙박할 수 있는 부대시설 및 시설이며 유스호스텔·콘도·호텔 등 객실
관광활동	여행목적에 반영하는 관광객이 여행 중의 행위들과 관련된 것들을 말하며 견문확대, 견학, 시찰, 친목회, 연수여행, 신혼여행, 가족여행, 체육대회 참관, 골프, 낚시, 온천, 해수욕장, 휴양, 국제회의 참가, 역사적 유적지 방문, 종교행사 등
식사문제	음식을 먹을 수 있는 여행 상품, 관광지 내에서 식사를 할 수 있는 음식 및 시설이며, 전문분야 음식점, 유명음식점, 카페테리아, 커피전문점, 노천카페, 라운지와 같은 식음료 시설 및 동양식, 서양식, 보양식, 기호식품과 같은 음식이나 식사와 함께하는 쇼 관람 등
쇼핑활동	관광지에서 하는 활동이며 기념품, 지역특산물, 유명상품 등
여행관련 서비스	여행과 관련된 무형의 서비스를 말하며, 특정한 장소나 지역을 안내하는 안내원 및 여행관리자, 무형의 상품을 만들어 여행객들에게 제공하는 여행사, 출입국수속 서비스, 브로셔, 카탈로그, 지도책, 전문잡지, 전문여행지 등의 정보 안내지, 금융서비스, 인터넷, 대중매체 등을 포함한 정보제공 시스템 등

자료: 선행연구를 바탕으로 연구자 재구성

<표 3-2> 지각된 가치 관련 연구 수집 데이터의 분류 범주

구분	세부적인 요인
기능적 가치	여행에서 접할 수 있는 관광지의 수, 투어시간 이행, 편안한 좌석, 안전 기록 및 가격 등의 가치 인식에 영향을 미치는 요인
사회적 가치	여행을 통해 얻은 개인의 위신 또는 인지도, 관광에서 사람들 간의 상호 작용, 관광객과 관광 안내원 간의 관계와 같은 요인
감성적 가치	여행에서 나타나는 감성이거나 정서적인 상태이며 예로 모험 관광 체험에서 느끼는 기분 전환과 흥분 및 망설임, 두려움, 걱정과 같은 요인
지적 가치	관광에서 새로운 지식 추구하고 참신함, 탐험적이며 신기성 추구, 다양성 추구 행동과 같은 요인
조건적 가치	여행에서 선택이라는 것에 직면했을 때 상황들의 설정이나 구체적인 상황에 따라 도출되는 결과에 대한 지각된 효용 및 관련된 요인

자료: Sheth et al.(1991), Williams & Soutar(2009)의 선행연구를 바탕으로 연구자 재구성

4. 데이터 수집과 분석방법

본 연구는 네이버와 다음의 포털사이트 데이터를 추출해서 나온 텍스트를 가지고 부산 여행상품의 구성요소들과 부산 여행에서 형성된 지각된 가치를 분석하고 결과 도출을 위하여 비정형 데이터를 수집하고 텍스트마이닝(Text Mining)을 통해 데이터를 정제하였으며, 거기에서 정제된 텍스트들 간의 관계를 파악하기 위하여 사회연결망분석 기법 중 하나인 시맨틱 네트워크 분석(semantic network analysis)을 실행하였다.

가. 텍스트마이닝(Text Mining)

IBM 보고서(2015)에 의하면 전 세계의 데이터 중 80%가 텍스트와 같은 비정형 데이터(unstructured data)로 구성되어져있다. 텍스트마이닝은 비정형, 반정형 텍스트 데이터에서 자연어처리 기술에 기초하여 유의미한 정보를 추출하고 가공하는 것을 목적으로 하는 기술이다. 텍스트마이닝 기법을 통해 대용량 텍스트에서 유의미한 정보를 추출해 내고, 다른 정보와의 연관관계를 파악하며, 텍스트가 가진 카테고리를 찾아내거나 간단한 정보 검색이 아닌 그 이상의 결과를 도출해 낼 수 있다. 컴퓨터가 자연어(인간이 사용하는 언어)를 분석하고 내면에 숨겨진 정보를 얻기 위해 대용량 언어 자원과 통계적이며 규칙적인 알고리즘이 사용되고 있다. 주요 응용분야로 문서분류, 문서군집, 정보추출, 문서요약 등이 있다(강만모 외, 2012).

본 연구는 TEXTOM에서 네이버와 다음의 키워드를 통한 데이터 수집을 하였으며, 수집한 데이터를 정제하였고, 정제된 데이터를 바탕으로 텍스

트의 동시출현 빈도분석을 시행하였다. 그리고 부산 여행의 요소들과 지각된 가치의 요소들을 분류하여 상위에 있는 주요 텍스트들을 구분하여, 시맨틱 네트워크 분석을 위한 빈도 매트릭스 변환시켜 연구를 진행하였다.

(1) 키워드 분석을 통한 데이터 수집

키워드 분석은 정제되지 않은 방대한 데이터에서 유용한 텍스트를 추출하기 위해서 자주 활용되고 있으며, 동시출현 빈도분석은 키워드 검색 대상인 텍스트들과 함께 상대적으로 출현빈도가 높은 텍스트들을 기반으로 유용한 텍스트를 추출해내는 기법이라고 할 수 있다. 기존의 텍스트 데이터를 활용한 대다수의 연구들은 동시출현 빈도가 높은 텍스트들이 가장 유의미한 텍스트들로 간주하고 있다(서정아, 2015).

데이터 수집을 위해 TEXTOM 웹사이트의 수집하기(collecting)에서 수집채널은 네이버와 다음의 블로그, 카페, 뉴스 총 3개를 선택하여 키워드 검색창에 '부산 여행'을 입력한 후, 부산 여행에 관련된 명사와 지각된 가치 텍스트 데이터를 추출하기 위해 형용사를 선택하여 2019년 1월 1일부터 2019년 12월 31일까지 해당되는 날짜의 데이터를 수집하였다.

(2) 수집데이터의 정제

포털사이트 네이버와 다음의 블로그, 카페, 뉴스 3개 채널은 비정형 데이터로 형식이 따로 정해져 있지 않기 때문에 추가적인 수정 작업이 필요하다. 텍스트와 엑셀파일로 변환된 데이터를 가지고 각종 조사들과 특수문자 코드, 부호 등을 제거하였으며, 띄어쓰기가 필요한 부분들은 띄어쓰기를 실

행하였다. 그리고 동일한 의미를 가진 명사 텍스트들은 빈도수가 가장 많은 키워드로 통일하여 동일하게 수정하였다. 동일한 의미로 수정된 명사 텍스트의 예는 <표 3-3>과 같다.

형용사는 어미가 다른 동일한 뜻의 여러 단어들에서 어근은 그대로 두고 어미를 제거하여 정제한 후, 하나의 동사 단어로 변경하여 가치를 나타내도록 동일하게 정제작업을 추가하였다. 동일한 의미로 수정된 형용사 텍스트의 예는 다음의 <표 3-4>와 같다.

위와 같은 과정을 여러 번 반복 실행한 후 ‘부산 여행’ 데이터들을 여행 상품과 관련된 연구에서 사용된 여행상품의 구성요소인 관광지 관련사항, 숙박시설, 교통수단, 식사문제, 관광활동, 쇼핑활동, 여행 관련 서비스에 관련된 7가지 텍스트로 선택하고 정제하였으며, 지각된 가치의 구성요소인 Willams & Soutar(2009)의 구성요소 즉 감성적 가치, 기능적 가치, 사회적 가치, 지적 가치, 조건적 가치와 관련된 5개의 텍스트들을 선택 및 정제하였다. 그리고 시멘틱 네트워크 분석을 실행하기 위해 여행상품과 지각된 가치의 구성요소의 빈도 매트릭스 파일을 각각 생성했다.

<표 3-3> 동일한 의미로 수정된 명사 텍스트의 예

동일한 의미의 유사 텍스트 통합 전 명사 텍스트	동일한 의미의 텍스트로 통합 후 명사 텍스트
1박, 이틀, 1박 2일	1박 2일
흰여울, 흰여울마을	흰여울문화마을
해수욕, 바닷가, 바다	해수욕장
케이블카, 송도해상관광케이블카	송도케이블카
해동용궁, 용궁, 용궁사	해동용궁사
이기대스카이워크, 오륙도스카이워크	스카이워크
당일치기, 당일, 하루	당일
자갈치시장, 자갈치, 태종, 태종대	자갈치시장 태종대
코스, 투어, 루트	코스

<표 3-4> 동일한 의미로 수정된 형용사 텍스트의 예

동일한 의미의 유사 텍스트 통합 전 형용사 텍스트	동일한 의미의 텍스트로 통합 후 형용사 텍스트
분위기 좋고, 위치 좋은, 좋은 방법, 코스로 좋은데요	좋다
먹을, 먹을지, 먹고 싶은, 먹으러, 먹었던	먹다
추천합니다, 추천하는, 추천해서, 추천할	추천하다
기차, KTX, 모노레일, 송케이블카, 버스 타고, 타러	타다
없고, 없을, 손색이 없을, 높일 리 없는	없다
즐기는, 즐길만한 곳, 즐기고 싶은	즐기다
처음으로 와 봤는데, 부산 여행 처음인데, 처음갔을때	처음이다
버스를 이용해서, 스파를 이용하니, 이용후기	이용하다
맛있는 브런치, 멋있게 먹은, 맛있는 식당	맛있다
숙소, 차량, 호텔을 예약하고, 예약하는, 예약했다	예약하다
소개하는, 소개된 맛집, 소개하는 곳	소개하다
행복했던 여행, 행복한 추억, 행복했던 여행	행복하다

나. 사회연결망분석(Social Network Analysis)

사회연결망분석은 수학의 그래프 이론에 근거를 두고 있다. 사회연결망의 연결구조 및 연결강도 등으로 사용자의 영향력 및 명성을 측정하여, 사회연결망 상에서 허브 역할을 하는 사용자나 소문의 중심지를 찾는 데 주로 활용된다. 사회연결망에서 영향력이 있는 사용자를 인플루언서(influencer)라고 부르는데, 인플루언서 모니터링 및 관리는 마케팅 관점에서 매우 중요하다고 할 수 있다(강만모 외, 2012).

그리고 사회연결망분석의 종류 중 하나인 시멘틱 네트워크 분석은 텍스트가 전달하려고 하는 의미를 분석의 대상으로 정하며, 텍스트의 맥락까지 고려하여 텍스트들 간의 상호 관계를 측정하고 분석하는 방법이다.

시멘틱 네트워크 분석은 텍스트들을 사회연결망으로 구성하는 노드를 사용하고, 텍스트들의 연결망을 통해서 텍스트들 간의 상호 관계성을 파악하여 유의미한 정보를 도출하며, 다양한 중심성 분석을 통해 텍스트들과 관계에서 역할분석, 텍스트들의 패턴분석도 가능하다(서정아, 2015).

사회연결망분석과 마찬가지로 네트워크의 연결 구조 특성을 파악하기 위해 중심성 분석(centrality analysis), 밀도 분석(density analysis), 구조적 등위성 분석(CONCOR 분석) 등과 같은 분석지표를 사용한다(이수희 등, 2018; 김두환, 2020).

실질적으로 가장 많이 활용되는 분석으로는 중심성 분석이며, 노드가 중심에 어느 정도로 근접한지를 분석하는 방법으로 이것은 네트워크에서 핵심 역할을 하는 텍스트가 무엇인지를 판별할 수 있다.

중심성 분석의 가장 대표적인 측정 방법으로는 연결정도 중심성(degree centrality), 근접 중심성(closeness centrality), 매개 중심성(betweenness

centrality), 위세 중심성(eigenvector centrality) 총 4가지 종류가 있다(한지연·김홍범, 2017).

텍스트마이닝을 통해서 정제된 데이터 부산 여행상품의 요소와 지각된 가치의 요소는 동시출현 분석을 통해서 만들어진 텍스트들로 빈도 매트릭스로 변환하였으며, 그 매트릭스를 시멘틱 네트워크 분석을 통해 실행하기 위해 이진법으로 부호화한 후 UCINET6를 분석 프로그램으로 선택하였으며, 매트릭스 에디터(matrix editor)의 기능을 통해 UCINET dataset 파일로 변환하였다.

UCINET dataset로 변환시킨 파일을 적용하여 시멘틱 네트워크 분석을 실행하였다. 그리고 UCINET6의 NetDraw기능을 통하여 부산 여행상품과 지각된 가치의 요소를 각각 시각화하였다. 시멘틱 네트워크 분석은 먼저 밀도와 네트워크 중심화 분석, 연결중심성 분석, CONCOR 분석의 순서로 진행하였다.

<표 3-5> 데이터 분석방법 및 내용

분석방법		분석내용	도구
텍스트마이닝 (Text Mining)	키워드 분석	부산 여행에 대한 텍스트 수집	TEXTOM
	동시출현 빈도분석	부산 여행 상품의 구성요소와 연결된 단어들의 동시출현 빈도분석	
의미연결망 분석 (Semantic Network Analysis)	밀도 분석	부산 여행 관련 단어들의 연결구조와 특성 분석	UCINET6 & NetDraw
	전체네트워크 집중분석	부산 여행 특정단어 집중도 분석	
	중심성 분석	부산 여행에 영향력이 높은 단어분석	
	CONCOR 분석	부산 여행에서 실제 단어들의 유사군집 분석	

자료: 선행연구를 바탕으로 연구자 재구성

(1) 밀도와 네트워크 중심화 분석

네트워크 분석에서 밀도(density)란 가능한 총 관계수와 실제로 맺어진 관계수의 비율이며, 일반적으로 관찰된 밀도는 집단의 크기와 반비례한다 (김용학 외, 2014).

예를 들어, 4개의 점이 있다고 가정을 한다면 그 점을 모두 연결 할 수 있는 링크는 총 6개이다. 만약에 연결 할 수 있는 링크에 모두가 다 연결 되어 있다면 그 분석에서의 밀도는 1이며, 연결 링크의 수가 작아질수록 밀도는 작아진다.

$$\text{네트워크 밀도} = \frac{k}{n(n-1)/2}$$

k = 연결된 링크수
 $n(n-1)/2$ = n 개의 노드에서 연결가능한 링크수

방향성이 없는 4개의 점을 네트워크 밀도 식으로 계산한다면 만약 링크의 총 개수가 6개라면 밀도가 1이며, 링크의 수가 4개라면 밀도는 0.67이 된다.

밀도가 높다는 것은 그 만큼 연결망을 구성하는 노드들의 상호 간 도움이나 교류가 많다는 것을 의미하며, 밀도의 값을 통해 네트워크가 형성하는 관계의 응집성, 결속도, 복잡성 등을 설명할 수가 있다(서정아, 2015).

본 연구는 부산 여행상품의 요소와 지각된 가치의 요소 키워드들이 얼마나 상호 간에 도움이나 교류가 얼마나 있는지를 파악하고자 하였다.

네트워크 중심화(centralization)는 연결망 형태가 전체적으로 얼마나 집중되어 있는지를 나타내는 개념이며, Freeman(1979)이 이 지표를 고안했으며, 그래프 중심성(graph centrality)이라고 표현하기도 하였다. 중심화가

가장 높은 연결망은 한 점으로 모든 연결이 연결되어 있는 방사선 형태의 연결망이며, 중심화는 전체적으로 하나의 연결망이 얼마나 중앙 집중적인 구조를 가졌는지, 혹은 연결망이 얼마나 하나의 점을 중심으로 결속되었는가를 측정하는 것이다. 중심화를 측정하는 방법은 가장 중심적인 점과 다른 모든 점들의 중심성 점수들 간의 차이를 각각 구하여 이를 모두 합한 다음, 이것을 논리적으로 가능한 최대값을 나누는 것이다(김용학 외, 2014). 예를 들어 연결정도 중심성에 근거한 중심화 수준은 다음과 같다.

$$C_D = \frac{\sum_{i=1}^n [C_D(p_*) - C_D(p_i)]}{\max \sum_{i=1}^n [C_D(p_*) - C_D(p_i)]}$$

$C_D(p_i)$ = 노드 i 에서의 연결정도중심성
 $C_D(p_*)$ = 연결망에서 연결정도중심성 최대값

$$\max \sum_{i=1}^n [C_D(p_*) - C_D(p_i)]$$

= 한 노드로 연결이 집중되었을 때 (★ 연결망) 즉 논리적으로 한 노드의 중심성과 나머지 노드들의 중심성 차이가 최대가 되는 경우에서 차이의 합

UCINET6의 네트워크 중심화에서 백분율로의 표현식이 전체 메시지 구성의 짜임새와 관련이 있다. 즉 연결중심화 정도가 100%에 가까울수록 주요 핵심텍스트 중심으로 전체 메시지 구조가 짜임새 있게 구성되어있는 것을 의미한다(이은선 외, 2012).

(2) 네트워크 데이터에 대한 통계적 유의성 검정

사회연결망분석은 매트릭스 데이터를 기본으로 하여 각각의 노드들 간 상호 의존적인 패턴을 이해하고 이러한 패턴들의 구조적인 것에 관심을 가진다. 이러한 사회연결망 데이터의 특징으로 인해서 일반적인 통계적 검정 절차를 진행할 수 없는 한계를 가진다. 관측 값들의 독립성 가정을 충족시키지 못하는 네트워크 데이터에 대한 검정 통계량을 계산하기 위해서 표준 오차를 추정할 수 있는 다른 방법이 필요하다(곽기영, 2014).

이 때 활용할 수 있는 방법이 UCINET6를 활용한 부트스트랩 기법이며, 부트스트랩은 이론적 수치에 대한 네트워크 밀도를 비교하는 통계적 검정을 위해 실행한다(서정아, 2015). 본 연구 또한 부산 여행상품의 요소와 지각된 가치의 네트워크 데이터의 유의성 검정을 위해 부트스트랩을 적용한 일 표본 평균검정을 실행하였다. 단일 네트워크 대상으로 하는 밀도 또는 평균 관계 강도에 대한 가설 검정은 전통적 통계적 통계분석기법의 일 표본 평균검정과 같은 의미를 갖는다(곽기영, 2014). 즉, 표본의 평균을 이용하여 관계 데이터 모집단의 평균에 대한 가설을 검정하였다. 부트스트랩기법의 활용을 통해 표본의 평균들로 구성된 표집분포를 직접 도출하여 표집분포의 평균과 표준오차를 산출할 수 있다. 산출된 표준오차를 이용하여 계산한 검정 통계량을 통해서 네트워크의 관계의 유의성 검정이 가능하다(서정아, 2015).

부트스트랩에서 ‘proportion of absolute differences as large as observed’에 대한 확률 값은 유의 수준 5%를 기준으로 하여 유의성을 검정할 수 있으며, 이 확률은 일반적인 통계 검정에 사용되는 p-값에 해당된다(곽기영, 2014).

(3) 중심성 분석

중심성은 권력과 영향력이라는 개념과도 연결이 되어 가장 많이 쓰이는 지표 중의 하나이다. 많은 경험분석 결과에서 중심성이 높은 개인은 특별한 사회적, 경제적 지위에 있고, 중심성이 높은 조직은 조직의 생존율이 높거나 성과가 좋다. 국가 간 교류 연결망에서도 중앙의 국가가 세계 경제에 있어서 가장 커다란 영향력을 행사한다. 중앙의 국가와만 교류하는 변방의 국가들은 교역 상대를 잃지 않기 위해 중앙 국가의 압력에 굴복하기도 한다.

이처럼 중앙에 얼마나 가까이 있는지를 측정하는 것은 매우 중요하며, 따라서 다양한 측정법을 통해 산출되는 여러 지표들이 제안되어 있는데, 이에선 연결정도 중심성(degree centrality), 인접 중심성(closeness centrality), 사이 중심성(betweenness) 등이 있다(김용학 외, 2014).

본 연구는 연결된 노드의 수를 의미하는 연결중심성을 이용하여 부산 여행상품의 요소와 지각된 가치의 요소를 파악하여 키워드의 영향력을 파악한다. 네트워크에서 연결 중심성이 높을수록 영향력이 매우 크다고 할 수 있다. 다시 말해 연결 중심성이 높은 키워드는 부산 여행상품의 요소와 지각된 가치의 요소에 영향력이 매우 높다고 할 수 있다. 또한 대부분의 연결망 분석 프로그램에서 연결 중심성은 밀도와 마찬가지로 최대 가능한 연결 수에 대한 %로 나타나는데, 이 값은 크기가 다른 연결망들을 비교할 수 없는 문제가 있다. 연결망의 크기가 클수록 그 연결망에 속한 점들의 중심성이 낮아지는 통계적 경향이 있기 때문에 크기가 같은 연결망을 비교할 때 쓰면 좋다(김용학 외, 2014).

(4) CONCOR 분석

CONCOR 분석은 상관관계를 이용한 대표적인 구조적 등위성 측정 방법으로, 행위자들 간의 관계 패턴을 도출하기 위해 행위자들 간의 상관관계(correlation)를 사용하는 방법이다. CONCOR는 두 행위자 간의 관계 패턴이 얼마만큼 비슷한지에 관심이 있는 경우에 적절하며, STRUCTURE는 관계의 동일성 여부에 중점을 두는 경우에 적절하다.

구조적 등위성이란 하나의 네트워크에서 다른 행위자들 간에 직접적인 관계는 없지만 동일한 관계패턴을 가지는 경우를 말한다. 또한 구조적 등위성 분석이란 유사한 지위(position)의 행위자들을 그룹화하고, 그 그룹들 간의 관계를 설명하는 것이다.

구조적 등위성 분석의 첫 번째 단계는 구조적 등위성을 측정하는 것이다. 대표적인 측정 방법으로는 상관관계를 이용한 방법(CONCOR)과 유클리디안 거리를 이용한 방법(STRUCTURE)이 있다.

CONCOR는 복잡한 네트워크 집단사이에 숨겨진 하위 집단의 관계 분석에 활용된다(오익근 외, 2015).

본 연구는 UCINET6를 통해 부산 여행상품의 네트워크에서 유사한 군집의 집단들을 파악하여 여행상품을 만들 때 참고가 될 수 있도록 하였다. CONCOR 분석은 매트릭스의 피어슨 상관계수를 계산하고 동일한 과정을 반복하여 새로운 상관관계를 도출하는데, 전 단계에서 생성된 상관계수를 이용하여 다시 상관분석을 수행하는 것과 같다(곽기영, 2014).

IV. 연구결과

1. 데이터 수집결과

본 연구를 위해 포털 사이트 네이버와 다음의 블로그, 카페, 뉴스의 3개 채널에서 ‘부산 여행’ 키워드를 TEXTOM을 이용하여 2019년 1월 1일부터 2019년 12월 31일까지 총 1년의 데이터를 수집하였으며, 그 결과 데이터의 수집량은 총 65,559(22.68MB) 텍스트에 달하였다.

각 채널별로 수집된 데이터양을 살펴보면, 네이버는 블로그(12,000 텍스트), 카페(11,911 텍스트), 뉴스(10,499 텍스트) 순으로 수집되었으며, 다음은 카페(11,816 텍스트), 블로그(10,545 텍스트), 뉴스(8,788 텍스트) 순으로 나타났다.

블로그와 카페가 뉴스에 비해 많은 양으로 보아 개인이 부산 여행을 경험한 후 직접 글을 올리는 경우가 많다고 할 수 있다. 텍스트마이닝을 통해 채널 별로 수집된 데이터양은 <표 4-1>에서 확인할 수 있다.

<표 4-1> 수집 데이터양

포털사이트	블로그	카페	뉴스	합계
네이버	12,000	11,911	10,499	34,410
다음	10,545	11,816	8,788	31,149
합계	22,545	23,727	19,287	65,559

자료: 데이터 수집량을 바탕으로 작성.

주: 전체적인 데이터양 순위 블로그 > 카페 > 뉴스

2. 연구문제 I 분석결과

연구문제 I

부산 여행상품 형성에 영향을 미치는 명사 텍스트들과 여행상품의 요소들은 무엇인가?

1-1. 부산 여행상품 형성에 영향을 미치는 명사 텍스트들은 무엇인가?

1-2. 부산 여행상품 형성에 영향을 미치는 명사 텍스트들과 연관된 여행상품의 요소들은 무엇인가?

가. 부산 여행에 영향을 미치는 텍스트 분석

부산 여행 텍스트의 동시출현 빈도분석을 실행하였으며, 부산 여행상품에 영향을 미치는 명사 텍스트를 분석하고 결과를 도출하고자 하였다. ‘부산 여행’ 키워드로 수집된 데이터에서 상위 100개의 텍스트의 구성은 부산의 지역, 여행 형태, 여행 기간, 여행 시기, 항공, 미디어가 주로 이루어져 있다.

명사 텍스트의 동시출현 빈도분석 결과는 다음과 같다.

첫째, ‘해운대’, ‘광안리’, ‘태종대’, ‘기장’, ‘남포동’, ‘서면’, ‘영도’ 순으로 지역의 빈도수가 많이 나타났다.

둘째, ‘1박2일’, ‘2박3일’, ‘당일치기’의 순으로 여행의 형태가 나타났다. 부산 여행에서 ‘1박2일’이 ‘2박3일’ 여행보다 더 많이 출현하였으며, 빈도수가

높았다. 그리고 ‘친구’, ‘아이’, ‘가족’, ‘엄마’, ‘혼자’ 순으로 빈도가 높게 나타났다. 또한 ‘맛집’, ‘호텔’, ‘숙소’, ‘카페’ 순으로 빈도수가 많이 나타났다.

종합적으로 빈도를 살펴보면, 포털사이트를 사용하는 사람들은 부산 여행을 패키지여행 보다는 자유여행으로 많이 가고, 가족여행의 빈도수가 높은 것으로 나타났다. 그리고 아이와 친구, 부모님 특히 엄마와 같이 여행을 하며 휴가를 가서 호텔에서 휴양하며, 카페와 맛집을 다니고, 부산의 맛집으로는 돼지국밥집을 많이 가는 것으로 나타났다.

셋째, 여행시기와 여행일정과 관련된 텍스트는 ‘여름’, ‘토요일’, ‘겨울’, ‘오후’, ‘저녁’, ‘야경’, ‘휴가’, ‘금요일’, ‘오전’, ‘연휴’, ‘가을’ 순으로 동시출현 빈도가 나타났으며, ‘1박 2일’, ‘2박 3일’, ‘하루’, ‘당일치기’ 순으로 빈도수가 나타났다. 그리고 운송수단은 ‘비행기’, ‘기차’, ‘자동차’, ‘버스’ 순으로 빈도수가 나타났으며, 항공은 ‘에어부산’, ‘티웨이항공’, ‘제주항공’의 저가항공 순으로 빈도수가 나타났다.

넷째, 날씨에 대한 빈도수가 나타내는 것은 부산 여행을 계획하는 관광객들은 부산 여행을 떠나기 전 부산의 날씨에 신경을 쓰는 것으로 나타났다. 명사 데이터의 주요 텍스트 빈도분석은 다음의 표<4-2>와 같이 살펴볼 수 있다.

<표 4-2> 명사 데이터의 주요 텍스트 빈도분석

순위	텍스트	빈도수	백분율(%)	순위	텍스트	빈도수	백분율(%)
1	해운대	6860	0.64315%	26	가족	1428	0.1338%
2	맛집	4827	0.4525%	27	기장	1338	0.1254%
3	추천	3574	0.3631%	28	예약	1334	0.1250%
4	서울	3727	0.3494%	29	아침	1314	0.1231%
5	에어부산	3223	0.3021%	30	주말	1276	0.1196%
6	친구	3162	0.2964%	31	소개	1212	0.1136%
7	호텔	3076	0.2883%	32	날씨	1207	0.1131%
8	사진	2798	0.2623%	33	후기	1198	0.1123%
9	광안리	2776	0.2602%	34	울산	1122	0.1122%
10	노선	2760	0.2587%	35	남포동	1109	0.1039%
11	1박2일	2653	0.2487%	36	국내여행	1106	0.1036%
12	바다	2447	0.2294%	37	차	1103	0.1034%
13	숙소	2446	0.2293%	38	운항	1101	0.1032%
14	카페	2186	0.2049%	39	서면	1094	0.1025%
14	대구	2124	0.1991%	40	여름	1061	0.0994%
16	코스	2057	0.1928%	41	가볼만한곳	1042	0.0997%
17	부산역	1961	0.1838%	42	토요일	995	0.0932%
18	계획	1883	0.1765%	43	감천문화마을	979	0.0917%
19	이벤트	1833	0.1718%	44	대전	960	0.0900%
20	인천	1629	0.1527%	45	투어	959	0.0899%
21	아이	1627	0.1525%	46	엄마	910	0.0853%
22	국내	1568	0.1470%	47	영도	903	0.0846%
23	기차	1507	0.1412%	48	경남	878	0.0823%
24	태종대	1487	0.1394%	49	프로모션	865	0.0810%
25	제주	1436	0.1346%	50	버스	864	0.0810%

순위	텍스트	빈도수	백분율(%)	순위	텍스트	빈도수	백분율(%)
51	겨울	861	0.0807%	76	오륙도	689	0.0645%
52	2박3일	859	0.0805%	77	힐링	676	0.0633%
53	광주	848	0.0795%	78	돼지국밥	673	0.0630%
54	해수욕장	847	0.0794%	79	역사	665	0.0623%
55	오후	839	0.0786%	80	휴가	663	0.0621%
56	프로그램	833	0.0780%	81	김포	657	0.0615%
57	서비스	831	0.0779%	82	금요일	651	0.0610%
58	저녁	828	0.0776%	83	제주항공	646	0.0605%
59	해동용궁사	824	0.0772%	84	속초	619	0.0580%
60	야경	802	0.0751%	85	부산관광공사	603	0.0565%
61	시장	794	0.0744%	86	오전	602	0.0565%
62	축제	794	0.0744%	87	연휴	598	0.0560%
63	비	777	0.0728%	88	남편	595	0.0557%
64	혼자	776	0.0727%	89	자갈치시장	573	0.0537%
65	가족여행	765	0.0717%	90	패키지	572	0.0537%
66	영화	761	0.0713%	91	여수	571	0.0535%
67	제주도	759	0.0711%	92	가을	567	0.0531%
68	가격	756	0.0708%	93	동백섬	563	0.0527%
69	경주	734	0.0688%	94	방송	562	0.0527%
70	숙박	732	0.0686%	95	계절	551	0.0516%
71	국제시장	727	0.0681%	96	전주	546	0.0511%
72	항공권	719	0.0674%	97	거제	522	0.0489%
73	티웨이항공	718	0.0674%	98	호캉스	514	0.0481%
74	광안대교	717	0.0672%	99	경기	513	0.0480%
75	통영	694	0.0650%	100	흰여울문화마을	512	0.0480%

나. 부산 여행에 영향을 미치는 요소 분석

여행상품의 요소에 대해 명사 데이터의 상위 100개의 동시출현 텍스트들을 선행연구를 바탕으로 연구자가 재정립한 7가지 요소로 분류하였다. 분류의 결과, 관광활동과 관광지 관련사항 요소에 포함되는 텍스트가 가장 많았으며, 그 다음으로 여행 관련서비스, 교통수단, 숙박시설, 식사문제, 쇼핑활동 순으로 텍스트가 포함되었다.

첫째, 부산 여행상품을 고려할 때 관광활동의 요소들을 가장 많이 고려한다고 해석할 수 있으며, ‘친구’, ‘아이’, ‘가족’, ‘엄마’, ‘혼자’, ‘가족여행’, ‘패키지여행’, ‘혼자여행’ 등의 여행 형태를 고려하고, ‘1박2일’, ‘2박3일’ 등의 여행 기간, ‘아침’, ‘오후’, ‘야경’ 등의 관광활동과 ‘주말’, ‘여름’, ‘가을’ 등의 여행 시기 등을 고려하는 것으로 나타났다.

둘째, 관광지 관련사항 요소에는 ‘해운대’, ‘광안리’, ‘태종대’, ‘기장’, ‘남포동’, ‘서면’, ‘영도’ 등의 키워드들이 포함되었으며, 부산 여행 관광객들은 부산 관광지 관련사항을 고려할 때 지역의 특성이 있는 관광지에 대해 많이 고려하는 것으로 파악할 수 있으며, 지역에 대한 날씨도 함께 고려한다고 해석할 수 있다.

세 번째, 여행관련 서비스는 ‘추천’, ‘코스’, ‘투어’, ‘계획’, ‘프로모션’, ‘이벤트’ 등의 텍스트가 포함되어 있었다. 부산을 여행하는 관광객들은 여행관련 서비스를 고려할 때 여행관련 ‘방송’, ‘프로모션’을 고려하는 것으로 나타났다.

네 번째, 교통수단 요소에는 ‘에어부산’, ‘부산역’, ‘기차’, ‘자차’, ‘운항’과 같은 텍스트들이 포함되어 있었다. 부산으로 갈 수 있는 교통수단 관련 텍스트들이 다수이며, 국내 항공사의 저비용항공을 이용해 부산을 많이 방문

하는 것으로 파악할 수 있으며, 부산 내에서의 이동은 자차 혹은, 버스와 같은 교통수단을 고려하는 것으로 나타났다.

다섯째, 숙박시설 요소와 관련된 텍스트들은 ‘호텔’, ‘숙소’, ‘숙박’, ‘호캉스’ 등이 포함되어 있어, 민박, 펜션, 게스트 하우스 등과 같은 저가의 숙박 시설보다는 호텔, 리조트와 같은 고가의 숙박 시설을 더 이용하는 것으로 파악할 수 있다. 부산으로 여행을 가는 관광객들이 부산 여행상품을 고려할 때 숙박 시설에 신경을 쓴다는 것을 알 수 있다.

여섯째, 식사문제에서는 ‘맛집’, ‘카페’, ‘돼지국밥’ 등이 포함되어 있다.

일곱째, 쇼핑활동에서는 ‘시장’이 포함되어 있다.

<표 4-3> 명사 데이터의 여행상품 요소별 분류

구 분	포함 텍스트	포함 개수
관광지 관련사항	해운대, 광안리, 바다, 태종대, 기장, 날씨, 남포동, 서면, 감천문화마을, 영도, 해수욕장, 해동용궁사, 비, 국제시장, 광안대교, 오륙도, 자갈치시장, 동백섬, 계절, 흰여울문화마을	20
교통수단	에어부산, 노선, 부산역, 기차, 자차, 운항, 버스, 항공권, 티웨이항공, 김포, 제주항공	11
숙박시설	호텔, 숙소, 숙박, 호캉스	4
관광활동	서울, 친구, 사진, 1박2일, 대구, 인천, 아이, 제주, 가족, 아침, 주말, 울산, 여름, 토요일, 대전, 엄마, 경남, 겨울, 2박3일, 광주, 오후, 저녁, 야경, 축제, 혼자, 가족여행, 제주도, 경주, 통영, 역사, 금요일, 속초, 오전, 연휴, 남편, 여수, 가을, 전주, 거제, 경기, 가볼만한곳, 국내, 국내여행, 휴가	44
식사문제	맛집, 카페, 돼지국밥	3
쇼핑활동	시장	1
여행관련 서비스	추천, 코스, 계획, 이벤트, 예약, 소개, 후기, 투어, 프로모션, 프로그램, 서비스, 영화, 가격, 힐링, 부산관광공사, 패키지, 방송	17

주: 관광활동>관광지관련사항>여행관련서비스>교통수단>숙박시설>식사문제>쇼핑활동

3. 연구문제Ⅱ 분석결과

연구문제Ⅱ

부산 여행에 영향을 미치는 형용사 텍스트들과 지각된 가치의 요소들은 무엇인가?

2-1. 부산 여행에 영향을 미치는 형용사 텍스트들은 무엇인가?

2-2. 부산 여행에 영향을 미치는 형용사 텍스트들과 연관된 지각된 가치의 요소들은 무엇인가?

연구문제Ⅱ를 해결하기 위해 ‘부산 여행’ 키워드를 활용하여 형용사 데이터를 수집하고, 수집된 데이터는 부산 여행에서 형성된 지각된 가치에 적합한 텍스트들만 선택하였다. 적합한 키워드를 가지고 동시출현 빈도분석을 실행한 후 명사의 데이터보다 형용사 데이터의 양이 적기 때문에, 상위 키워드 40개를 선정하여 연구문제Ⅱ를 해결하기 위한 데이터로 활용하였다. 동시출현 빈도분석을 통해 부산 여행에서 형성된 지각된 가치에 영향을 미치는 실제 텍스트들은 무엇이며, 이 텍스트들과 연관된 지각된 가치의 요소들은 무엇인지 파악하고자 하였다.

지각된 가치의 요소로는 Willams at al.(2009)이 관광 연구에 활용한 기능적 가치, 사회적 가치, 감성적 가치, 지적 가치, 조건적 가치(Sweeney & Soutar, 2001)의 총 5개의 구성요소를 활용하였다.

가. 지각된 가치에 영향을 미치는 텍스트 분석

형용사 데이터의 동시출현 빈도분석을 실행하여, 부산 여행의 지각된 가치 형성에 영향을 미치는 텍스트를 분석하였다. 그 결과, 긍정적인 텍스트들이 주로 형성되어 있으며, 부정적인 텍스트도 일부 파악할 수 있다. 형용사 데이터를 통한 동시출현 빈도분석의 결과는 다음과 같다.

첫째, ‘좋다’, ‘먹다’, ‘추천하다’, ‘타다’, ‘많다’, ‘즐기다’, ‘처음이다’, ‘이용하다’, ‘맛있다’, ‘예약하다’, ‘소개하다’, ‘행복하다’, ‘추억이다’, ‘즐겁다’, ‘느끼다’, ‘걷다’, ‘열리다’, ‘아름답다’, ‘짧다’ 순으로 ‘부산 여행’ 키워드와 동시출현한 긍정적인 텍스트가 수집되었다. 긍정적인 텍스트는 상위 40개 중에서 34개나 도출이 되어 부산 여행에서 전반적으로 긍정적 지각된 가치를 얻는 것으로 추측할 수 있었다. 특히, ‘좋다’라는 텍스트는 5266 텍스트로, 40개의 텍스트 중에 동시출현 빈도가 가장 높았으며, 두 번째로 많은 ‘먹다’라는 키워드를 통해 부산 여행의 다양한 먹거리를 볼 수 있어, 부산 여행에 대해서 ‘추천하다’, ‘즐겁다’, ‘맛있다’ 등과 같이 부산을 추천하고 싶고 즐거운 긍정적인 지각된 가치를 얻을 수 있다. 그리고 음식에 대해서는 전반적으로 ‘맛있다’로 나타났으며, 명사 데이터에서 휴양과 관련된 텍스트로는 ‘힐링이다’, ‘편안하다’라는 텍스트를 얻을 수 있다.

둘째, 총 40개 키워드 중에서 34개의 긍정적인 텍스트를 뺀 나머지 6개의 부정적인 텍스트인 ‘없다’, ‘짧다’, ‘힘들다’, ‘춥다’, ‘덥다’, ‘멀다’ 등이 도출되었다. 부산 여행에서 ‘좋다’, ‘먹다’ 만큼 ‘없다’라는 텍스트가 높은 빈도수를 차지하고 있다. 명사 데이터에서 ‘여름’이라는 텍스트가 수집된 것처럼 형용사 데이터에서도 ‘덥다’라는 텍스트가 나타났다. ‘겨울’이라는 텍스트가 수집된 것처럼 형용사 데이터에서도 ‘춥다’라는 텍스트가 나타났다.

부정적인 텍스트는 날씨의 명사 텍스트와 관련된 ‘덥다’, ‘춥다’라는 키워드를 제외하고, 긍정적인 텍스트들 보다 적게 나왔다. 결과적으로, 부산 여행에서 형성된 지각된 가치는 긍정적인 것으로 나타났다. 형용사 데이터의 주요 텍스트 빈도분석은 다음의 <표 4-4>에서 전체적으로 확인할 수 있다.

<표 4-4> 형용사 데이터의 주요 텍스트 빈도분석

순위	텍스트	빈도수	백분율(%)	순위	텍스트	빈도수	백분율(%)
1	좋다	5719	0.76261%	21	기억하다	716	0.0932%
2	먹다	4795	0.6089%	22	힐링이다	676	0.0853%
3	추천하다	3881	0.4863%	23	최고다	662	0.0874%
4	타다	3591	0.4561%	24	기대하다	645	0.0799%
5	없다	3136	0.4047%	25	새롭다	606	0.0791%
6	많다	3064	0.4004%	26	가깝다	579	0.0697%
7	즐기다	2352	0.3103%	27	예쁘다	569	0.0694%
8	처음이다	1796	0.2352%	28	힘들다	478	0.0612%
9	이용하다	1474	0.2892%	29	크다	425	0.0586%
10	맛있다	1413	0.1817%	30	괜찮다	412	0.0568%
11	예약하다	1348	0.1763%	31	여유있다	411	0.0544%
12	소개하다	1217	0.1588%	32	편하다	383	0.0488%
13	행복하다	1005	0.1260%	33	재미있다	322	0.0430%
14	추억이다	962	0.1167%	34	편리하다	258	0.0333%
15	즐겁다	915	0.1162%	35	춥다	240	0.0310%
16	느끼다	845	0.1082%	36	편안하다	208	0.0286%
17	걸다	807	0.0976%	37	다르다	206	0.0282%
18	열리다	799	0.1006%	38	넓다	205	0.0270%
19	아름답다	776	0.1038%	39	덥다	196	0.0270%
20	짧다	747	0.0912%	40	멀다	192	0.0264%

나. 지각된 가치에 영향을 미치는 요소 분석

형용사 데이터의 동시출현 상위 40개 키워드로는 부산 여행에서 지각된 가치 형성에 영향을 미치는 요소들을 분석하였다.

그 결과, 감성적 가치와 기능적 가치가 가장 많이 포함되어 있으며, 그 다음으로 조건적 가치가 많이 포함되어 있었다. 사회적 가치와 지적 가치는 똑같이 2개로 가장 적게 나타났다. 다음은 형용사 데이터인 지각된 가치를 요소별 분류한 결과이다.

첫째, 가장 많은 포함개수를 가진 감성적 가치의 요소를 살펴보면, ‘좋다’, ‘즐기다’, ‘행복하다’, ‘추억이다’, ‘즐겁다’, ‘느끼다’, ‘아름답다’, ‘기억하다’, ‘힐링이다’, ‘기대하다’, ‘예쁘다’, ‘괜찮다’, ‘여유있다’, ‘재미있다’, ‘편안하다’ 등의 텍스트가 포함되어 있었다. 형용사 텍스트로만 보면 판단하기 힘들어 정확히 파악하기 위해 엑셀파일로 저장된 원자료(raw data)를 찾기 기능을 이용하여 해당문장을 찾아보았다. 관련된 문장을 살펴보면 ‘좋다’와 관련된 문장은 “부산 너무 좋다.”, “수서역 SRT를 타고 하루 만에 당일치기 부산 여행 참 좋다.”, “일년에 한 두 번씩 즐겨 찾는 부산 역시 좋다”, “부산 여행 1박2일 코스는 쉽게 찾을 수가 있어 여행보자들 쉽게 떠나기 좋다.”, “여행 가본 사람들은 다 부산이 여행하기 좋다고 하더라구요.” 등 남에게서 들었던 이야기와 “먼저 카페를 갔다가 가는 것이 좋다.” 등의 남들에게 추천을 하는 문장, “나는 부산이 정말 좋다”와 같은 만족에 대한 부산여행이 ‘좋다’와 관련된 글들이 다수였다. ‘힘들다’는 “이제는 더워 힘들다”와 같이 “부산은 더워서 힘들다”라는 날씨에 대한 ‘힘들다’와 “2박 3일 동안 힘들다” 등과 같이 여행 일정에 대한 ‘힘들다’라는 문장이 있었으며, 특히, “아이가 어려서 관광은 힘들다고 하더라”, “아기는 데리고 다니기

힘들다라고 얘기하시는데”라는 문장처럼 아이들과 함께하는 관광이 힘들다는 글이 많이 있었다. ‘즐겁다’, ‘재미있다’, ‘여유있다’, ‘힐링된다’, ‘편안하다’, ‘선호하다’라는 텍스트들은 전반적으로 관광활동에 대해서 많이 나왔으며, ‘예쁘다’, ‘아름답다’라는 텍스트는 관광지에 대한 문장에서 많이 찾아볼 수 있었다.

둘째, ‘먹다’, ‘타다’, ‘없다’, ‘많다’, ‘맛있다’, ‘짧다’, ‘넓다’, ‘편리하다’는 키워드들은 기능적 가치의 요소에 포함시켰다. 그 결과, “부산에는 해운대 말고도 여행을 즐길 만한 곳이 많다.” 등의 긍정적인 문장들이 많이 도출되었다. 그리고 반대의미의 텍스트인 ‘없다’를 검색한 결과 “생각보다 뭐 볼 게 없다는 생각이 든다.”, “감천문화마을 부산 사람들은 별거 없다지만”, “지역 관광업계 어려움은 별다른 지원책이 없다.”, “부산 이동거리가 많았지만 어쩔 수 없다.” 등과 같은 부정적인 문장들이 도출되었지만, “엄마와의 부산 여행 문제없다”, “부산 맛집을 놓칠 수 없다”, “박물관과 아쿠아리움을 꼭 들르는 아들이 이곳을 놓칠 리 없다.”, “부산은 다양한 핫집, 맛집들이 많아 식도락 여행 또한 빠질 수 없다.”, “거가대교의 개통은 부산과 거제를 여행하려는 사람들에게겐 희소식이 아닐 수 없다.”, “부산 여행은 먹으러 온 것과 다름이 없다.”, “부산 여행을 할 때는 바다를 따로 떼어 생각할 수 없다.” 등과 같은 긍정적인 문장도 도출되었다. ‘크다’는 “부산 해운대에 위치한 센텀시티는 크다”, ‘짧다’라는 “일반적으로 다니던 여행에 비하면 너무 짧은 여행인데” 짧은 일정에 대한 문장이 나왔다. 그리고 나머지는 음식이 ‘맛있다’, 호텔과 지역이 ‘넓다’, 물가가 ‘싸다’, ‘비싸다’라는 키워드는 부산의 물가가 비싼 게 아니라 “여름 성수기에는 기차 티켓이 비싸다”라는 문장들이 전반적으로 도출되었다. ‘편리하다’는 교통수단에 대해서 ‘부산 버스는 편리하다’라는 문장이 많았다.

셋째, ‘덥다’, ‘최고다’, ‘멀다’, ‘춥다’, ‘가깝다’, ‘다르다’ 등의 조건적 가치

에 포함된 텍스트 중에서 ‘괜찮다’는 “사람들이 부산 괜찮다고 해서”, “부산이 괜찮다고들 하는데” 등과 같은 남들에게 부산 여행 조건을 추천 받은 글과 “혼자 먹기 괜찮은”, “의외로 생각보다 부산의 음식들이 입맛에 나쁘지 않고, 교통도 괜찮은 편” 등의 조건적으로 괜찮아서 추천하는 글들이 많이 나타났다. ‘최고다’는 “맛은 진짜 최고였다”, “서울에서 3시간이면 닿을 수 있는 부산의 힐링 여행은 최고다” 등과 같은 부산 여행에서 겪은 것에 대한 최고라는 글들이 다수였다. ‘다르다’는 “부산 여행을 다녀왔는데 서울과는 다르게 정말 사람들이 행복하고 즐거워 보이고 여유롭고”, “부산 날씨가 다른 지역과는 다르더라고요” 등과 같은 부산 여행을 갔을 때 다른 곳과 ‘다르다’라는 표현이 다수다. ‘가깝다’는 “서울에서 부산이 가깝다”, “관광지 접근성이 가깝다”라는 문장들이 다수였으며, ‘덥다’, ‘뜨겁다’라는 글은 “다른 지역에 비해서 여름이 덥고 뜨겁다”라는 글과 함께 “2019년도 여행업계와 항공업계가 뜨겁다.” 등과 같이 “부산 여행에 대한 인기가 많다”라는 의미를 ‘뜨겁다’라는 표현으로 쓰기도 하였다.

넷째, 사회적 가치와 관련된 텍스트인 ‘추천한다’는 “부산 여행 추천합니다”, “치안이 양호한 부산으로 추천드립니다”라고 하는 문장처럼 부산 여행을 다녀오고 만족하여 사람들에게 추천을 한다는 문장이 대다수였다.

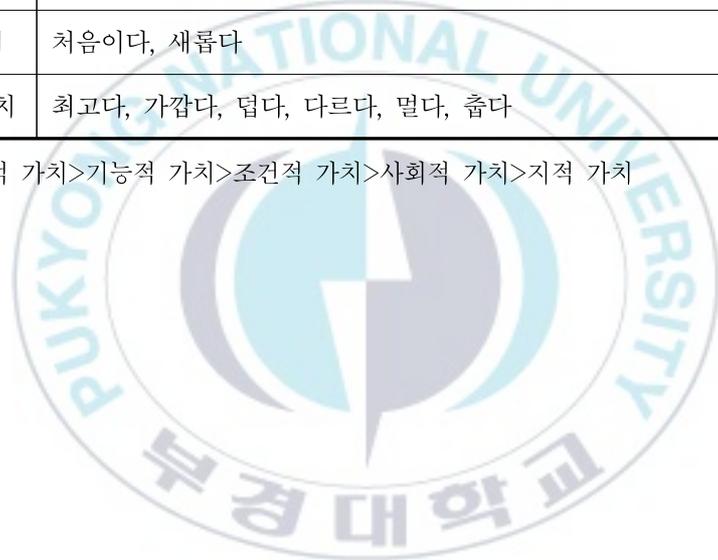
다섯째, 지적 가치요소에 포함되어 있는 ‘새롭다’, ‘매력있다’는 텍스트는 “부산도 충분히 매력적”, “부산이 충분히 매력적인 여행지라는 걸 깨달은 터였다”, “새로운 부산의 모습을 보니”, “부산의 새로운 관광지” 등과 같이 부산의 관광지에 대해서 매력적이며, 새로운 느낌으로 관광에 호기심을 자극하는 것으로 유추할 수 있다.

형용사 데이터의 지각된 가치의 요소별 분류는 다음의 <표 4-5>에서 확인할 수 있다.

<표 4-5> 형용사 데이터의 지각된 가치 요소별 분류

구분	포함 텍스트	포함 개수
기능적 가치	먹다, 없다, 많다, 이용하다, 맛있다, 열리다, 짧다, 긴다, 크다, 편하다, 편리하다, 넓다, 타다	13
사회적 가치	추천하다, 예약하다	2
감성적 가치	좋다, 즐기다, 행복하다, 추억이다, 즐겁다, 느끼다, 아름답다, 기억하다, 힐링이다, 기대하다, 예쁘다, 힘들다, 괜찮다, 여유있다, 재미있다, 편안하다	15
지적 가치	처음이다, 새롭다	2
조건적 가치	최고다, 가깝다, 덥다, 다르다, 멀다, 춥다	6

주: 감성적 가치>기능적 가치>조건적 가치>사회적 가치>지적 가치



4. 연구문제Ⅲ 분석결과

연구문제Ⅲ

부산 여행상품의 구성요소 및 지각된 가치에 연관된 실제 텍스트들의 네트워크 구조와 특성을 통해서 파악되는 부산 여행상품의 구성요소 및 지각된 가치의 형성구조는 어떠한가?

- 3-1. 부산 여행상품 및 지각된 가치의 네트워크 구조와 특성을 통해서 부산 여행상품의 구성요소 및 지각된 가치의 형성구조는 어떠한가?
- 3-2. 부산 여행상품 및 지각된 가치와 연관된 주요 텍스트들의 네트워크 구조에서 가장 영향을 미치는 텍스트들과 요소들은 무엇인가?
- 3-3. 부산 여행상품의 구성요소를 형성하는 유사집단 구분은 어떠한가?

명사와 형용사의 데이터 키워드를 통해 시멘틱 네트워크 분석을 실행하여 연구문제Ⅲ을 해결하고자 하였다. 텍스트마이닝에서 실행한 바와 같이 빈도분석을 실행하여 생성된 명사 텍스트 데이터와 형용사 텍스트 데이터를 TEXTOM에서 네트워크 분석이 가능하도록 빈도 매트릭스로 변환하였다. 그리고 생성된 빈도 매트릭스 파일을 분석 프로그램인 UCINET6에서 .h 파일로 변환시킨 후 밀도와 네트워크 중심화 분석을 실행한다. 그 다음 일 표본 평균 검정을 통해 네트워크 데이터의 유의성 검정을 실행하였으며, 연결중심성을 이용하여 부산 여행상품의 구성요소와 지각된 가치를 파악하여 텍스트의 영향력을 파악하였다. 마지막으로 여행상품을 형성하는 실제 텍스트들의 유사군집 형성을 위해 CONCOR 분석을 실행하였다.

가. 부산 여행 관련 텍스트의 시멘틱 네트워크 구조 분석

(1) 명사 데이터의 주요 텍스트의 밀도와 네트워크 중심화 분석

부산 여행상품을 구성하는 주요 텍스트의 네트워크 구조와 특성을 파악하기 위해 명사 데이터의 핵심 키워드 네트워크 구조의 밀도와 네트워크 중심화 분석을 실행하였다. 네트워크 구조의 밀도는 프로그램 UCINET6를 통해 Network > Whole networks & Cohesion > Density > Density Overall 메뉴에서 .##h 파일을 찾아 불러내어 밀도를 계산할 수 있으며, 네트워크 중심화 분석은 Network > Centrality > Degree를 통해 나온 결과의 제일 하단 Network Centralization 수치로 파악할 수 있다.

밀도 분석의 결과, 연결망의 링크는 총 8,706개로 키워드 하나에 평균 87.060개 연결되어 있었으며, 밀도는 0.879가 나왔다. 그리고 네트워크 중심화 수준은 0.326으로 나타나, 명사 데이터의 네트워크 중심화는 특정 텍스트에 집중되는 텍스트들의 정도가 32.6%라고 할 수 있다. 명사 텍스트의 밀도와 중심화 분석결과는 <표 4-6>에서 확인할 수 있다.

<표 4-6> 부산 여행 명사의 밀도와 중심화

DENSITY / AVERAGE MATRIX VALUE					
Input dataset: N (C:\Users\J\Desktop\2020\N)					
Output dataset : N-density(C:\Users\J\Documents\UCINET data\N-density)					
	1	2	3	4	5
	Density	No. of Ties	Std Dev	Avg Degree	Alpha
1 N	0.879	8706	0.326	87.060	0.999
1 rows, 5 columns, 1 levels.					

그리고 명사 텍스트의 시멘틱 네트워크 분석결과를 시각화한 결과는 <그림 4-1>에서 볼 수 있다.

명사 데이터에서 시각화한 내용을 보면, ‘가족여행’, ‘혼자여행’, ‘자유여행’, ‘친구’의 관광활동 요소와 ‘해운대’, ‘광안리’ 등과 같은 관광지 관련사항 요소들이 중심에 배치되어 있음을 알 수 있다.

(2) 형용사 데이터의 주요 텍스트의 밀도와 네트워크 중심화 분석

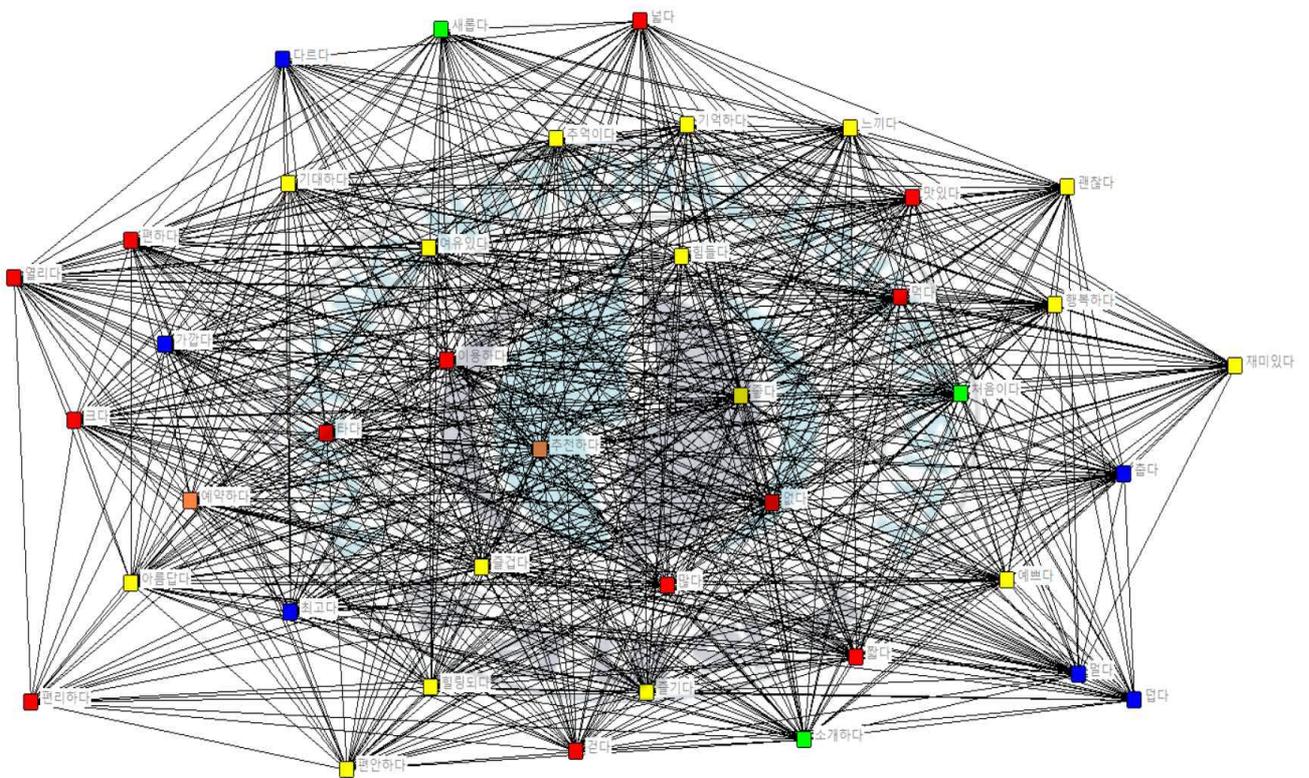
부산 여행에서 형성된 지각된 가치를 나타내는 형용사 데이터의 총 링크 수는 1,424개가 나왔다. 형용사 데이터는 명사 데이터의 양보다 적기 때문에 링크의 총 개수에는 차이가 있었다. 그리고 한 텍스트 당 평균 35.6개 링크가 연결되어 있었으며, 밀도는 0.913으로 확인되었다. 전체 네트워크 중심화 수준은 0.282(28.2%)이다. 형용사 텍스트의 밀도와 중심화 분석결과는 다음의 <표 4-7>에서 확인할 수 있다.

<표 4-7> 부산 여행 형용사의 밀도와 중심화

DENSITY / AVERAGE MATRIX VALUE					
Input dataset: A (C:\Users\J\Desktop\2020\A)					
Output dataset : A-density(C:\Users\J\Documents\UCINET data\A-density)					
	1	2	3	4	5
	Density	No. of Ties	Std Dev	Avg Degree	Alpha
1 A	0.913	1424	0.282	35.600	0.998
1 rows, 5 columns, 1 levels.					

결과적으로 형용사 데이터는 명사를 꾸며주는 단어로 사용되어 명사에 비해 반복적이며 지속적으로 사용되지 않아 명사보다 데이터양이 적을 뿐만 아니라 실제로 연결된 링크 수도 작으며, 특정 텍스트에 집중되어 있지

않아 수평적인 구조로 해석할 수 있다. 형용사 데이터의 주요 텍스트를 시멘틱 네트워크 분석한 후 그 결과를 시각화한 것이 다음의 <그림 4-2>와 같다.



<그림 4-2> 부산 여행 형용사 텍스트의 시멘틱 네트워크 시각화

주: 빨(기능적 가치), 주(사회적 가치), 노(감성적 가치), 초(지적 가치), 파(조건적 가치)

(3) 네트워크 데이터의 유의성 검정

부산 여행상품을 구성하는 요소인 명사 데이터와 지각된 가치의 요소인 형용사 데이터의 유의성 검증을 위해 밀도에 대한 일 표본 평균 검정을 실행하였다.

부트스트랩 기법을 통해 일 표본 평균 검정은 UCINET6에서 Network > Compare densities > Against theoretical parameter 순서대로 클릭 한 후 .###h 파일을 선택하고 Expected Density에는 검정하고자 하는 모수의 값인 0을 입력하고, number of samples에 표본의 개수를 5,000개로 지정하여 값을 산출하였다.

부트스트랩기법은 모집단 내에서 네트워크 밀도가 0이라고 가정하고 표본으로부터 발생한 밀도가 우연인가를 파악하기 위한 유의성 검정방법으로, 표집분포를 직접 도출하여 유의성을 검정한다(서정아, 2015). 명사와 형용사 데이터의 네트워크 데이터 유의성을 검증한 결과는 다음의 <표 4-8>과 <표 4-9>와 같다.

부트스트랩 기법을 통한 네트워크 데이터에 대한 명사 데이터의 유의성 검정 결과, 평균 부트스트랩 밀도(average bootstrap density)는 0.8822이었으며, 표준편차(estimated standard error for density)는 0.0143으로 나왔다. 그리고 Z-score(밀도에 표준편차를 나눈 값)는 61.5554로 나왔다.

일반적인 통계검정에서 사용되고 있는 p-value에 해당되는 관측 값과 절대 값의 차이(proportion of absolute differences as large as observed)는 0.0002이므로 5%유의수준에서 유의하다고 할 수 있다.

<표 4-8> 부산 여행 명사의 유의성 검정 결과

COMPARE DENSITY W/ HYPOTHESIZED VALUE

Random seed: 16500

Parameter value is: 0.0000

Density of C:\Users\J\Desktop\2020\N.##h is: 0.8794

Difference is: 0.8794

Variance of ties for C:\Users\J\Desktop\2020\N.##h: 0.1061

Classical estimate of SE: 0.0033

Number of bootstrap samples: 5000

Estimated standard error for density of

C:\Users\J\Desktop\2020\N.##h: 0.0143

Z-score: 61.5554

Average bootstrap density: 0.8822

Proportion of absolute differences as large as observed: 0.0002

Proportion of differences as large as observed: 0.0002

Proportion of differences as small as observed: 1.0000

부트스트랩 기법을 통한 네트워크 데이터에 대한 형용사 데이터의 유의성 검정 결과, 평균 부트스트랩 밀도(average bootstrap density)는 0.9217이었으며, 표준편차(estimated standard error for density)는 0.0201로 나왔다. 그리고 Z-score(밀도에 표준편차를 나눈 값)는 45.3239로 나왔다.

일반적인 통계검정에서 사용되고 있는 p-value에 해당되는 관측 값과 절대 값의 차이(proportion of absolute differences as large as observed)는 0.0002이므로 5%유의수준에서 유의하다고 할 수 있다.

<표 4-9> 부산 여행 형용사의 유의성 검정 결과

COMPARE DENSITY W/ HYPOTHESIZED VALUE	
Random seed:	8792
Parameter value is:	0.0000
Density of C:\Users\J\Desktop\2020\A.##h is:	0.9128
Difference is:	0.9128
Variance of ties for C:\Users\J\Desktop\2020\A.##h:	0.0796
Classical estimate of SE:	0.0071
Number of bootstrap samples:	5000
Estimated standard error for density of C:\Users\J\Desktop\2020\A.##h:	0.0201
Z-score:	45.3239
Average bootstrap density:	0.9217
Proportion of absolute differences as large as observed:	0.0002
Proportion of differences as large as observed:	0.0002
Proportion of differences as small as observed:	1.0000

<표 4-10> 명사, 형용사의 네트워크 데이터 유의성 검정 결과

구 분	평균 부트스트랩 밀도	표준오차	Z-score	관측값과 절대값 차이
여행상품 (명사)	0.8822	0.0143	61.5554 =밀도(0.8794)/ 표준편차(0.0143)	0.0002*
가치 (형용사)	0.9217	0.0201	45.3239 =밀도(0.9217)/ 표준편차(0.0201)	0.0002*

* .05 수준에서 유의

나. 부산 여행 관련 텍스트의 연관 단어 영향력 분석

동시출현 빈도분석은 빈도를 파악하여 영향을 미치는 실제 텍스트들과의 요소파악은 가능하지만 연구 주제에 대한 형성 요소의 영향력을 파악하기는 힘들다. 이에 따라 본 연구는 부산 여행상품의 요소와 지각된 가치의 요소에 해당하는 실제 텍스트들이 부산 여행상품과 지각된 가치에 미치는 영향을 파악하기 위해 연결중심성 분석을 실행하였다.

(1)부산 여행의 연관 텍스트 연결중심성 분석

부산 여행상품의 요소에 해당하는 명사 데이터를 통해 네트워크상에서 특정한 텍스트와 다른 텍스트와의 연결된 링크 수를 파악하기 위해 연결중심성 분석을 실행하였다. 연결중심성이 높은 텍스트는 관광객들이 부산 여행상품을 선택할 때 많은 영향을 미친다고 할 수 있으며, 결과는 다음과

같다.

첫째, 명사 데이터 연결중심성의 제일 높은 순위는 ‘추천’, ‘서울’, ‘호텔’, ‘사진’, ‘계획’, ‘국내’, ‘여름’, ‘힐링’ 순으로 나타났다. 이들은 연결정도가 높은 것으로 나와 모든 텍스트들과 연결되어 텍스트 간에 영향력이 매우 높다고 할 수 있다. 명사 데이터의 주요 텍스트 빈도수와 비교하면 ‘가족여행’이 ‘해운대’, ‘광안리’보다 빈도수는 적었지만 연결중심성은 더 높게 나왔으며, ‘여름’ 텍스트는 40번째로 빈도수가 많았지만, 연결중심성 값은 첫 번째로 높았다. 결과적으로 ‘여름’, ‘가족여행’은 빈도수에 비해 영향력이 큰 것으로 나타났다.

둘째, 100개 텍스트를 전체적으로 살펴보면 ‘계획’, ‘여름’, ‘힐링’, ‘가족’, ‘겨울’을 비롯하여, ‘토요일’, ‘오후’, ‘가족여행’, ‘오후’, ‘저녁’, ‘투어’, ‘오전’, ‘연휴’ 등의 텍스트가 빈도수에 비해 순위가 두드러지게 나타나 부산 여행상품에 영향력이 높다고 할 수 있다. 그리고 ‘해운대’, ‘맛집’, ‘노선’, ‘에어부산’, ‘운항’ 등의 텍스트는 빈도수에 비해 연결중심성 순위가 낮은 것으로 보아 부산 여행상품에 있어서 영향력이 떨어진다는 것을 확인 할 수 있다.

셋째, 명사 데이터의 연결중심성 값을 1~34위는 상위그룹, 40~66위는 중위그룹, 69~100위는 하위그룹의 세 그룹으로 나누어 각 그룹에 어떤 여행상품의 요소들이 포함이 되어있는지 확인하였다. 상위그룹에는 ‘추천’, ‘서울’, ‘호텔’, ‘사진’, ‘계획’, ‘국내’, ‘여름’, ‘힐링’, ‘친구’, ‘바다’, ‘가족’, ‘겨울’, ‘주말’, ‘날씨’, ‘토요일’, ‘저녁’, ‘해운대’, ‘숙소’, ‘투어’, ‘1박2일’ 등 총 20개의 관광활동 요소에 포함되는 텍스트들이 가장 많았으며, ‘바다’, ‘해운대’, ‘날씨’ 등의 관광지 관련사항 요소가 포함되어있는 텍스트들이 두 번째로 많이 포함되어 있었다. 중위그룹도 역시 관광활동 요소에 포함하는 16개의 텍스트와 관광지 관련사항 요소에 포함되어있는 7개의 텍스트 순으로 나타났다.

마지막으로 하위그룹에는 교통수단 요소에 포함되는 텍스트인 ‘에어부산’, ‘노선’, ‘김포’, ‘운항’, ‘항공권’, ‘제주항공’, ‘티웨이항공’의 총 7개 텍스트와 ‘방송’, ‘부산관광공사’, ‘프로모션’, ‘영화’, ‘돼지국밥’ 등의 텍스트가 포함되어있다. 명사 데이터의 여행상품을 요소별로 분류한 결과와 비교해 볼 때, 가장 많은 텍스트가 포함되어 있는 관광활동 요소와 관광지 관련사항 요소 역시 상위그룹에 많이 포함되어있어 관광객들이 관광활동 요소와 관광지 관련사항 요소에 영향을 많이 받는다고 해석할 수 있다. 하위그룹에서는 제일 적게 포함되어 있는 텍스트들이 교통수단 요소에는 영향을 적게 주는 것을 확인할 수 있지만, 의외로 5번째, 10번째로 많은 텍스트들이 포함되어 있는 요소인 교통수단 요소와 여행관련 서비스 요소의 텍스트들이 2개 포함되어 있어서 교통수단과 여행관련 서비스는 영향을 적게 미치는 것으로 나타났다. 명사 데이터의 연결중심성 분석결과는 다음의 <표 4-11>에서 확인할 수 있다.

<표 4-11> 명사 데이터의 연결중심성 분석

순위	텍스트 (빈도순위)	연결중심성 정도	순위	텍스트 (빈도순위)	연결중심성 정도
1	추천(3)	1.000	26	연휴(87)	0.949
	서울(4)	1.000		1박2일(11)	0.949
	호텔(7)	1.000		카페(14)	0.949
	사진(8)	1.000		인천(20)	0.949
	계획(18)	1.000		예약(28)	0.949
	국내(22)	1.000		아침(29)	0.949
	여름(40)	1.000		국내여행(36)	0.949
	힐링(77)	1.000		버스(50)	0.949
9	친구(6)	0.990	34	맛집(2)	0.939
	바다(12)	0.990		코스(16)	0.939
	가족(26)	0.990		이벤트(19)	0.939
	겨울(51)	0.990		제주(25)	0.939
13	대구(15)	0.980	40	비(63)	0.939
	주말(30)	0.980		금요일(82)	0.939
	날씨(32)	0.980		부산역(17)	0.929
	울산(34)	0.980		후기(33)	0.929
	토요일(42)	0.980		가격(68)	0.929
	오후(55)	0.980		광안리(9)	0.929
	가족여행(65)	0.980		아이(21)	0.919
20	소개(31)	0.970	44	축제(62)	0.919
	저녁(58)	0.970		휴가(80)	0.919
22	해운대(1)	0.960	47	기차(23)	0.909
	숙소(13)	0.960		차(37)	0.909
	투어(45)	0.960		엄마(46)	0.909
	오전(86)	0.960		시장(61)	0.909

순위	텍스트 (빈도순위)	연결중심성 정도	순위	텍스트 (빈도순위)	연결중심성 정도
51	숙박(70)	0.909	76	남편(88)	0.838
	가을(92)	0.909	77	대전(44)	0.828
53	감천문화마을(43)	0.899		영화(66)	0.828
	경남(48)	0.899		돼지국밥(78)	0.828
	해수욕장(54)	0.899		여수(91)	0.828
	야경(60)	0.899	81	자갈치시장(89)	0.798
60	경주(69)	0.899	82	오륙도(76)	0.788
	태종대(24)	0.889		계절(95)	0.788
	혼자(64)	0.899	84	전주(96)	0.778
	62	기장(27)	0.879	85	프로모션(49)
광주(53)		0.879	해동용궁사(59)		0.768
66	가볼만한곳(41)	0.869	88	동백섬(93)	0.768
	서비스(57)	0.869		에어부산(5)	0.758
	제주도(67)	0.869		노선(10)	0.758
	국제시장(71)	0.869		방송(94)	0.758
69	2박3일(52)	0.859	92	거제(97)	0.758
	통영(75)	0.859		현여울문화마을(100)	0.747
	남포동(35)	0.859		93	경기(99)
73	영도(47)	0.848	94	속초(84)	0.687
	프로그램(56)	0.848	95	부산관광공사(85)	0.636
	패키지(90)	0.848	96	김포(81)	0.626
	호캉스(98)	0.848	97	운항(38)	0.616
76	서면(39)	0.838	98	항공권(72)	0.606
	광안대교(74)	0.838	99	제주항공(83)	0.515
	역사(79)	0.838	100	티웨이항공(73)	0.404

(2) 지각된 가치의 연관 텍스트 연결중심성 분석

부산 여행에서 형성된 지각된 가치를 의미하는 형용사 데이터를 통해 텍스트들 간의 영향 관계를 파악하기 위해 연결중심성 분석을 실행하였다.

우선, 상위 5개의 텍스트들은 ‘좋다’, ‘추천하다’, ‘타다’, ‘없다’, ‘많다’, ‘즐거다’, ‘이용하다’, ‘즐겁다’로 나왔는데, 빈도수 순위인 ‘좋다’, ‘먹다’, ‘추천하다’, ‘타다’, ‘없다’와 비교하면 두 번째로 빈도가 높은 ‘먹다’는 연결중심성 상위 순위에 들지 못했다. 결과적으로 ‘좋다’, ‘추천하다’, ‘타다’, ‘없다’, ‘많다’ 순으로 높게 영향을 미치는 것으로 나타났다.

그리고 연결중심성을 전체적으로 분석하면, ‘먹다’와 함께 ‘맛있다’, ‘여유있다’, ‘행복하다’, ‘재미있다’, ‘아름답다’, ‘최고다’, ‘힐링되다’, ‘예쁘다’, ‘괜찮다’, ‘편하다’ 순으로 빈도수에 비해 높은 순위를 나타내고 있으며, 동시출현 텍스트의 개수에 비해 텍스트들 간에 영향력이 높다고 할 수 있다.

‘편리하다’, ‘덥다’, ‘멀다’, ‘다르다’, ‘재미있다’, ‘새롭다’, ‘넓다’, ‘편안하다’ 등의 텍스트는 텍스트 빈도수에 비해 연결중심성은 순위 하락의 폭이 커서 빈도에 비해 텍스트의 영향력은 크지 않다고 할 수 있다.

마지막으로, 30개 텍스트를 17개, 11개, 12씩 나누고 3개의 상/중/하 그룹으로 구분하여 지각된 가치의 요소별로 분류함으로써 요소들의 영향력을 확인하였다. 그 결과, 상위그룹에는 ‘먹다’, ‘타다’, ‘없다’, ‘많다’ 등 총 6개의 기능적 가치에 포함되는 텍스트가 가장 많았으며, 감성적 가치에 포함되는 ‘좋다’, ‘많다’, ‘즐거다’, ‘느끼다’, ‘여유있다’, ‘힘들다’ 등의 텍스트가 5개로 두 번째로 많아 기능적 가치와 감성적 가치의 영향력이 크다는 것을 확인할 수 있다.

그리고 ‘소개하다’, ‘처음이다’는 상위그룹, ‘새롭다’는 하위그룹에 위치하

고 있어 지적가치는 포함된 텍스트의 수는 적지만 영향력은 상당히 높다는 것을 확인할 수 있다. 조건적 가치는 세 번째로 많은 텍스트를 포함하지만, 중위그룹에는 ‘최고다’, ‘가깝다’, ‘출다’가 있고, 하위그룹에는 ‘답다’, ‘다르다’, ‘떨다’ 등 각 3개씩 위치하고 있어 영향력은 낮다고 할 수 있다. 지적 가치와 함께 제일 적은 텍스트가 포함되어 있는 사회적 가치인 ‘추천하다’, ‘예약하다’는 상위그룹에 위치하고 있어 텍스트는 작지만 다른 요소들에 비해 영향력이 크다고 할 수 있다. 분석결과는 다음의 <표 4-12>에서 확인할 수 있다.



<표 4-12> 형용사 데이터의 연결중심성 분석

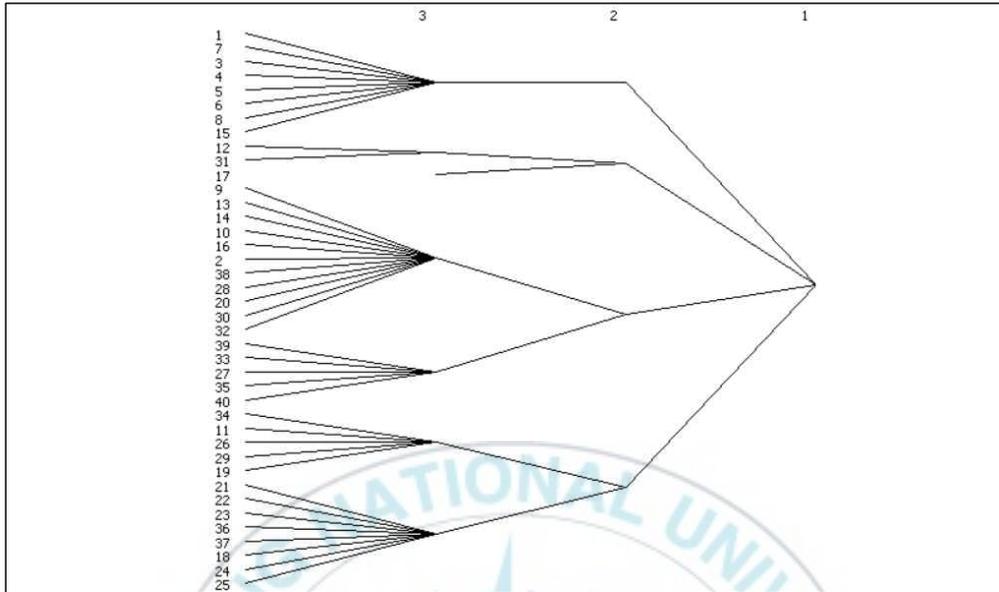
순위	텍스트 (빈도순위)	연결중심성 정도	순위	텍스트 (빈도순위)	연결중심성 정도
1	좋다(1)	1.000	21	기억하다(20)	0.949
	추천하다(3)	1.000		짧다(21)	0.949
	타다(4)	1.000		최고다(22)	0.949
	없다(5)	1.000		힐링되다(23)	0.949
	많다(6)	1.000		가깝다(26)	0.949
	즐기다(7)	1.000	26	기대하다(24)	0.923
	이용하다(8)	1.000	27	크다(29)	0.923
	즐겁다(15)	1.000	28	예쁘다(27)	0.897
9	먹다(2)	0.974	29	괜찮다(30)	0.872
	처음이다(9)	0.974		편하다(32)	0.872
	맛있다(9)	0.974		춡다(35)	0.872
	예약하다(11)	0.974	32	열리다(18)	0.846
	소개하다(12)	0.974		편안하다(36)	0.846
	느끼다(16)	0.974		넓다(38)	0.846
	건다(19)	0.974	35	새롭다(25)	0.795
	힘들다(28)	0.974	36	재미있다(33)	0.769
	여유있다(31)	0.974		다르다(37)	0.769
18	행복하다(13)	0.949	38	멀다(40)	0.744
	추억이다(14)	0.949	39	덥다(39)	0.641
	아름답다(17)	0.949	40	편리하다(34)	0.538

다. 부산 여행 관련 텍스트의 유사군집 분석

부산 여행상품을 설명하는 명사 데이터 네트워크를 유사한 집단으로 군집하여 파악하고, 그 집단 사이에 숨겨진 하위집단의 관계를 분석하기 위해 구조적 등위성 분석 중 하나인 CONCOR 분석을 실행하였다.

UCINET6를 실행 후 Network 메뉴에서 Roles & Position > Structural > Concor > Standard를 선택하여 새로 생기는 창의 Input Dataset에 명사 데이터 .##d 파일을 넣어 실행하면, 각각 .##d, .##h 형태의 3개 아웃풋 파일(ConcorCCPart, ConcorCCPerm, Concor1stCorrd) 총 6개의 파일이 생성된다.

그리고 Netdraw를 실행하고 명사 데이터 파일을 실행하여 매트릭스를 형성시키고 Open > Attribute Data에서 ConcorCCPart 파일을 속성데이터로 실행 한 후 Layout > Group by attribute > Categorical Attribute의 순서대로 메뉴를 실행하였다. 그 다음으로 Attribute to group by의 드롭다운 메뉴를 3으로 설정하여 실행한 후 CONCOR 분석의 속성 값에 따라 키워드를 분류하였다.



<그림 4-3> 명사 CONCOR Cluster Diagram

명사 데이터의 CONCOR 분석을 실행한 결과 총 4개의 대규모 클러스터가 형성되었는데 구체적으로 35개의 텍스트들로 구성된 대규모 클러스터 1개와 각 29개, 27개, 8개로 구성된 3개의 소규모 클러스터가 형성되었다. CONCOR 분석결과를 시각화한 내용은 다음의 <그림 4-4>에서 확인할 수 있다.

4개의 클러스터 중에서 35개의 텍스트들로 이루어진 대규모 클러스터는 관광활동 요소의 텍스트가 총 10개 포함되어 있다. 구체적으로 ‘여름’, ‘겨울’의 여행시기와 ‘1박2일’, ‘2박3일’과 같은 여행기간이 모두 포함되었으며 ‘아이’, ‘기차’, ‘날씨’, ‘비’, ‘계절’, ‘차’, ‘가볼만한곳’, ‘방송’ 등도 포함되어 있어, ‘어느 시기에’, ‘어떤 형태로’, ‘어디에 예약을 할 것인지?’, ‘어디에서 숙박을 할 것인지?’, ‘무엇을 타고 갈 것인지?’에 대한 부산 여행을 떠나기 전에 고려하는 텍스트들이 주로 포함되어 있음을 알 수 있다. 그리고 총 29개로 키워드들이 포함되어 있는 소규모 클러스터를 보면, ‘항공권’, ‘패키

지’, ‘프로모션’, ‘제주항공’ 그리고 ‘가족’을 동행하는 ‘가족여행’도 포함되어 있다.

세 번째로 많은 27개 텍스트로 구성된 클러스터에는 주로 관광지 내에서 이용하는 체험에 관련된 활동들이 포함되어 있으며, 관광지 관련사항 요소도 같이 포함되어 있다.

마지막으로 8개 텍스트들이 포함되어있는 소규모 클러스터에는 ‘추천’, ‘힐링’, ‘호텔’, ‘계획’, ‘여름’과 같은 숙소문제에 포함되어 있는 텍스트 총 8개 중 4개가 이 클러스터에 포함되어 있다. 다음의 <표 4-13>에서 각 클러스터에 포함되어 있는 텍스트를 확인할 수 있다.

<표 4-13> 각 클러스터의 텍스트 분류

구 분	포함 텍스트	포함 개수
클러스터 1 (여행 전 고려사항 관련)	감천문화마을, 방송, 혼자, 영화, 남포동, 국제시장, 기장, 아이, 기차, 날씨, 2박3일, 부산역, 흰여울문화마을, 동백섬, 해동용궁사, 자갈치시장, 광안대교, 가볼만한곳, 친구, 비, 돼지국밥, 태종대, 계절, 통영, 토요일, 차, 서면, 오진, 엄마, 영도, 오륙도, 남편, 야경, 광안리, 카페	35
클러스터 2 (패키지 관련)	부산관광공사, 대구, 항공권, 운항, 인천, 티웨이항공, 연휴, 노선, 예약, 이벤트, 패키지, 울산, 제주, 제주도, 경기, 경남, 소개, 에어부산, 김포, 휴가, 프로모션, 광주, 숙소, 겨울, 제주항공, 서비스, 가격, 오후, 가족	29
클러스터 3 (체험 관련)	시장, 금요일, 축제, 프로그램, 아침, 국내여행, 해운대, 대전, 저녁, 거제, 진주, 숙박, 맛집, 호캉스, 경주, 1박2일, 숙소, 가족여행, 코스, 후기, 가을, 여수, 역사, 버스, 바다, 주말, 투어	27
클러스터 4 (숙소 관련)	추천, 힐링, 호텔, 서울, 계획, 여름, 국내, 사진	8

V. 결 론

1. 연구의 요약

경제의 성장으로 인해 여행객들의 여행에 대한 욕구가 점점 높아짐에 따라 여행객들의 수는 매년 최고치를 갱신해 왔다. 비록 국내여행 산업은 세계적인 팬데믹으로 매우 큰 어려움을 겪고 있지만, 포스트 시대를 준비해야 하는 상황이다. 본 연구는 관광의 욕구가 다양화되고 고도화되어있는 관광 소비자들을 위해 TEXTOM을 활용한 사회연결망분석을 이용하여 어떤 여행상품의 구성요소를 중요시 여기는지 여행지에 대해 어떠한 지각된 가치를 가지고 있는지를 파악함으로써 여행상품의 질적 가치 제고를 통해 여행업계의 활성화에 기여하는데 그 목적이 있다.

본 연구는 국내 주요 포털사이트인 네이버와 다음의 블로그, 카페, 뉴스 3개의 채널 대상으로 키워드 분석을 통한 텍스트마이닝을 실행하여 분석하였다.

‘부산 여행’의 키워드를 활용하여 2019년 1월 1일을 시작으로 2019년 12월 31일까지의 1년간의 데이터로 한정하고 TEXTOM을 활용하여 총 65,559개의 텍스트를 수집하였다. 수집된 데이터는 정제작업을 거쳐 동시출현 빈도분석과 시멘틱 네트워크 분석을 실행하여 부산 여행상품과 지각된 가치에 영향을 미치는 텍스트들과 구성요소를 분류하고 네트워크 구조와 특성을 분석하여 부산 여행상품과 지각된 가치의 형성구조를 파악하였다.

본 연구의 결과는 다음과 같이 요약된다.

첫째, 명사 데이터의 동시출현 빈도분석 결과, 상위 100개의 텍스트를 기준으로 부산의 지역, 여행 형태, 여행 시기, 여행 기간, 교통이 주로 이루어

져 있음을 알 수 있었다. 특히 ‘해운대’는 압도적으로 많은 6,860개의 텍스트가 출현하였으며, 명사 데이터 상위 100개의 동시출현 텍스트들을 여행 상품 요소별로 분류한 결과에서는 관광활동, 관광지 관련사항, 여행관련 서비스, 교통수단, 숙박시설, 식사문제, 쇼핑활동 순으로 부산 여행상품을 구성하는 7가지 주요 요소임을 알 수 있었다.

둘째, 형용사 데이터의 동시출현 빈도분석의 결과 상위 40개의 텍스트 기준으로 긍정적인 텍스트가 주로 형성되었으며, 40개 텍스트 중 상위 주요 10개의 텍스트는 ‘좋다’, ‘먹다’, ‘추천하다’, ‘타다’, ‘없다’, ‘많다’, ‘즐기다’, ‘처음이다’, ‘이용하다’, ‘맛있다’, ‘예약하다’, ‘소개하다’, ‘행복하다’, ‘추억이다’, ‘즐겁다’ 순서로 출현했으며, ‘좋다’는 빈도가 5719번 출현하여 부산 여행에 대해 많은 사람들이 긍정적인 지각된 가치를 가진다는 것을 확인할 수 있었다. 그리고 부산 여행의 지각된 가치에 영향을 미치는 요소로는 감성적 가치가 가장 많은 텍스트들을 포함하고 있었으며, 그 다음으로 기능적 가치, 조건적 가치, 지적 가치, 사회적 가치 순으로 판단할 수 있었다.

셋째, 주요 텍스트의 네트워크 구조 밀도와 중심화 수준을 분석한 결과 명사 데이터에서는 ‘친구’, ‘가족여행’, ‘1박2일’, ‘2박3일’, ‘금요일’, ‘토요일’, ‘아이’, ‘가족’, ‘엄마’, ‘혼자’, ‘연휴’, ‘휴가’ 등의 관광활동 요소와 ‘해운대’, ‘광안리’, ‘태종대’, ‘기장’, ‘남포동’, ‘서면’, ‘영도’ 등의 관광지 관련사항의 요소들이 중심에 있는 것을 알 수 있으며, 네트워크 중심화 수준은 32.6% 정도로 파악할 수 있었다. 형용사 데이터는 명사 데이터 보다 양이 적으므로 링크의 수가 작게 있었으며, 전체 네트워크 중심화 수준은 28.2%로 명사 데이터보다 특정 텍스트에 낮게 집중되어 있다고 판단할 수 있었다.

넷째, 주요 텍스트의 네트워크 연결중심성은 명사 데이터의 경우 ‘추천’, ‘서울’, ‘호텔’, ‘사진’, ‘계획’, ‘국내’, ‘여름’, ‘힐링’ 순으로 나타났으며, 그리고 상위그룹에는 관광활동 요소와 관광지 관련사항 요소가 많이 포함되어 있

으며, 하위그룹에는 쇼핑활동이 포함되어 있는 것으로 파악할 수 있었다. 형용사 데이터의 경우 상위 5개 텍스트들은 ‘좋다’, ‘추천하다’, ‘타다’, ‘없다’, ‘많다’, ‘즐기다’, ‘이용하다’, ‘즐겁다’ 순이며, ‘즐겁다’, ‘이용하다’, ‘즐기다’, ‘많다’, ‘없다’, ‘타다’, ‘추천하다’ 순으로 빈도수보다 연결중심성 순위가 높은 것으로 파악되었다.

다섯째, 명사 데이터의 CONCOR 분석을 실행한 결과, 부산 여행을 가기 전에 고려하는 텍스트들로 이루어진 1개의 대규모 클러스터를 중심으로 패키지 관련, 체험 관련, 숙소 관련 텍스트가 포함되어 있는 3개의 소규모 클러스터 등 총 4개의 클러스터가 형성되었다.



2. 연구의 논의 및 시사점

본 연구 결과에 따른 실무적으로 논의와 시사점은 다음과 같다.

첫째, 부산 여행은 해운대를 중심으로 친구나 가족단위의 자유여행이 많은 것으로 나타났다. 본 연구결과도 앞의 조사에서 같이 ‘친구’와 ‘가족’이라는 텍스트를 통해 가장 많은 여행의 동반자는 친구나 가족으로 유추할 수 있다. 그러나 ‘가족여행’과 비슷한 빈도수를 나타낸 텍스트가 ‘혼자여행’으로 나타나 가족과의 여행의 비율도 높지만, 1인 가구의 증가 등에 따른 혼자 여행의 비율이 증가했으며, 혼자 여행을 떠날 때 가장 중점을 두는 안전에 있어서도 부산이 안전하다는 것을 살펴볼 수 있다.

‘자유여행’은 실속을 생각하는 소비자들이 증가하면서 가격대비 고품질의 상품을 찾으려는 소비자들이 직접 일정을 계획하고 숙소나 항공편 등을 직접 예약하는 여행 형태가 늘어나 자유여행이 패키지여행보다 빈도와 영향력이 높게 나타났다는 것으로 유추할 수 있다.

이상과 같이 볼 때, 관광지인 해운대에서 가족여행의 형태와 자유여행의 패키지 상품 개발에 주력할 필요가 있으며, 늘어나는 혼자여행객을 위한 여행상품의 개발도 적극적으로 강구해야 할 것이다.

둘째, ‘관광활동’과 ‘관광지관련사항’의 요소들이 부산 여행상품을 형성하는 주요 요소로 파악되었다. 이는 부산에서의 활동적인 부분과 관광지에서의 편의성, 관광지의 매력성, 관광지에 대한 위치적 접근성 등과 같은 요인들이 부산 여행상품 형성에 영향을 미친다는 것을 파악할 수 있다. 따라서 여행사는 부산 여행상품을 개발할 때 관광활동 부분과 관광지 관련사항에 대한 부분을 우선적으로 고려할 필요가 있다.

셋째, 부산 여행에서 형성된 지각된 가치는 전체적으로 긍정적인 측면을 나타내고 있어 국내 포털 사이트를 이용하는 사람들이 부산 여행에 전반적으로 긍정적인 태도를 보이고 있음을 알 수 있다. 특히, 감성적 가치의 요소가 가장 많이 포함되어 있어 Williams & Soutar(2009)의 논문에서처럼 여행에서 감성적 가치는 만족과 행동의도에 매우 강력한 관계가 있다는 결과를 뒷받침해주고 있다. 분석결과, 감성적 가치 요소에 포함되는 텍스트들이 많았다. 이는 앞서 나타난 결과와 같이 가족과 함께하는 여행과 인터넷 검색능력 향상으로 인한 자유여행 증가로 인해 나타난 것으로 해석할 수 있다. 결과적으로, 부산 여행상품을 만드는 여행사는 감성적 가치와 기능적 가치를 우선 고려해야 할 것이다.

넷째, 부산 여행을 형성하는 클러스터에는 여행 형태와 여행사, 여행 시기와 여행 기간 등 부산 여행을 떠나기 전에 고려하는 텍스트를 중심으로 대규모 클러스터 1개와 패키지관련 클러스터, 부산 여행의 체험과 관련된 클러스터, 숙소와 관련된 클러스터 3개의 소규모 클러스터가 생성되었다. 패키지관련 클러스터는 숙소의 가격과 연관이 되어있는데, 이는 부산 여행상품, 프로모션이나 교통수단 등을 선택할 때 조금 더 저렴하거나 많은 혜택을 찾는 것으로 나타나, 여행관련 마케팅에서는 부산 여행상품을 기획할 때 가격을 염두하고, 이벤트나 프로모션을 참고할 필요가 있다.

숙소관련 클러스터는 저가의 패키지 관련사항과는 다르게 호텔에서 힐링할 수 있는 사항들에 대한 추천을 고려하기 때문에 고가의 부산 여행상품을 기대하는 고객에 대한 부분도 고려할 필요가 있다.

다섯째, 부산 여행에서 형성된 가장 영향력이 큰 2개의 지각된 가치를 중심으로 여행상품의 구성요소와 연결하여 비교한 결과 전반적으로 긍정적인 부분들이 많음을 확인할 수 있어, 패키지여행의 상품 코스나 콘텐츠의 다양한 개발에 대해 구체적으로 고민할 필요가 있다.

본 연구의 학문적인 시사점으로, 빅데이터를 통해 여행상품과 지각된 가치의 연구에서 요소들을 측정할 점을 들 수 있다. 빅데이터에서 가장 많은 데이터인 비정형 데이터들 중에 하나인 포털사이트는 인터넷 이용자가 원하는 정보를 얻기 위해 먼저 거쳐 가야만하는 플랫폼과 같은 역할을 할 뿐만 아니라 인터넷 사용자들이 본인의 경험과 정보를 올리는 가상의 공간이며, 포털사이트의 이용객은 컴퓨터뿐만 아니라 스마트폰의 사용으로 매년 그 수가 증가하고 있는 실정이다. 네이버 여행의 포털사이트는 엄청난 양의 여행콘텐츠를 제공하고 있으며 여행업계에도 과급력이 크기 때문에, 향후에는 지금보다 훨씬 더 많은 정보량을 통해 여행상품의 요소를 측정할 수 있다.

또한, 시멘틱 네트워크를 통해 텍스트마이닝의 동시출현 빈도분석 뿐만 아니라 각 텍스트의 영향력 정도를 측정할 수도 있었다. 기존의 인과관계를 통한 관광 연구에서 텍스트의 네트워크 분석을 통한 상관관계를 이용하여 유용한 연구 결과를 도출할 수 있어 설문조사가 아닌 사회연결망분석을 활용한 새로운 연구 방법에 대한 방향을 제시할 수 있을 것으로 생각한다.

3. 연구의 한계점 및 향후 연구의 방향

본 연구의 한계점 및 향후 연구 과제는 다음과 같다.

첫째, 본 연구에서는 네이버와 다음의 블로그, 카페, 뉴스 3개의 채널에서만 자료를 수집하여 데이터의 양이 충분하지는 않았다. 페이스북, 인스타그램, 카카오톡과 같은 SNS는 비공개로 되어있어 자료의 수집과 분석이 불가능하였으며, 부산 여행과 연관되는 데이터의 양이 적어 분석하기에는 부족하였다. 따라서 향후 다양한 온라인 데이터를 활용하여 부산 여행에 대한 연구가 수행될 필요가 있다.

둘째, 텍스트마이닝과 네트워크 분석은 데이터를 분석하고 선별해내는 데 충분한 효용을 가지지만, 경험과 환경이 한정되어 도출된 데이터를 해석하는데 있어서 연구자의 주관이 배제되기 어렵다는 한계점이 있다. 또한 시각화를 통한 단어 빈도분석과 의미연결망은 핵심어의 위치와 단어들 간의 관계성을 보여주는 데에는 유용하나, 메시지의 의미를 완벽하게 해석할 수는 없다(김보경 외, 2015). 그러나 메시지 속의 단어는 단어의 영향력과 동시출현 빈도에 영향력을 나타내주기 때문에 충분히 1차적 데이터가 될 수 있으므로, 향후 관광연구를 하는데 도움을 줄 것이다.

셋째, 본 연구는 부산 여행에만 한정으로 연구하여 다른 지역에 대한 여행상품의 구성요소와 지각된 가치는 분석되지 못하였다. 향후에는 우리나라 국내의 다양한 지역뿐만 아니라 중국, 일본, 동남아, 유럽 등의 등 다양한 나라의 여행상품의 구성요소와 지각된 가치를 연구하여 여행업계가 여행상품구성을 계획하거나 마케팅 전략에 유용한 토대를 마련해야 할 것이다.

넷째, 수집기간이 1년으로 한정하여 연구하였기 때문에 단기적인 분석은

가능했으나, 장기적인 관점에서 여행상품과 지각된 가치의 변화를 파악하기 힘들었다. 그러므로 향후에는 좀 더 구체적으로 연도별, 계절별, 월별 등의 기간별 데이터를 수집하여 장기적인 관점의 부산 여행과 지각된 가치의 변화 연구가 수행되어야 할 것이다.

다섯째, 본 연구는 다양한 데이터 중에서 비정형 데이터를 통한 텍스트 분석을 실행하였다. 향후에는 정형 데이터와 반정형 데이터를 이용한 분석이 필요하며, 비정형 데이터도 동영상, 이미지 등의 다양한 종류의 데이터를 활용하여 여행상품과 지각된 가치의 연구뿐만 아니라 관광관련 연구가 추가적으로 필요하다.

본 연구의 의의는 다음과 같다.

기존 선행연구들은 표본을 설정하여 정량적 설문조사를 활용하여 여행상품과 지각된 가치의 요소들을 파악하였지만, 본 연구에서는 사회연결망분석을 활용하여 좀 더 빠르게 변하는 관광의 트렌드나 흐름을 연구하기 위한 새로운 접근을 시도하였다.

또한 온라인 사이트 이용자들이 직접 사용하는 단어들을 활용하여 연구함에 따라, 모든 구매과정에 속하는 소비자들이 직접 사용하는 단어들 간의 상관관계를 바탕으로 하여 여행업계에서 상품개발 및 마케팅 전략 수립에 있어 현실적인 접근이 가능한 분석을 실행하였다.

[참고문헌]

- 강병찬(2012). 명품관광 인식과 지각된 가치가 브랜드 태도 및 행동의도에 미치는 영향. 경희대학교 대학원, 박사학위논문.
- 강만모·김상락·박상무(2012). 빅데이터의 분석과 활용. 정보과학회지, 30(6), 25-32.
- 공효순·송은지(2013). 빅데이터를 이용한 호텔기업 CRM 및 보안에 관한 연구. 융합보안 논문지, 13(4), 69-75.
- 곽기영(2014). 소셜네트워크분석. 서울: 도서출판 청람.
- 곽춘려(2013). 관광 매력속성과 관광제약이 지각된 가치 및 행동의도에 미치는 영향. 호남대학교 대학원, 박사학위논문.
- 김기석(2017). 패키지여행의 상품선택속성과 만족 및 행동의도에 관한 연구. 경기대학교대학원, 석사학위논문.
- 김난도·전미영·이향은·이준영·김서영·최지혜(2016). 트렌드 코리아 2017, 서울: 미래의 창.
- 김동완(2013). 빅데이터의 분야별 활용사례. 동아대학교, 경영문제연구소.
- 김보경·김미경(2015). 빅데이터를 활용한 템플스테이 키워드 네트워크 분석. 관광학연구, 39(5), 27-40.
- 김성희(2009). 디지털도서관 통합 관리를 위한 사회연결망 구조에 관한 연구. 중앙대학교대학원, 박사학위논문.
- 김옥제(2001). 관광객의 성격유형과 여행상품 선택요인간의 관계 연구. 동의대학교대학원, 박사학위논문.
- 김용범(2016). 빅데이터를 활용한 호텔 선택속성의 중요도-만족도에 관한 연구: 5성급 호텔과 중저가 호텔 비교. 대구가톨릭대학교 일반대학원, 박사학위논문.
- 김용학·김영진(2014). 사회 연결망 분석. 서울: 박영사.

- 김재성(2012). SNS 빅 데이터 다차원 분석 기반 스마트폰 선호도 분석. 충북대학교대학원, 석사학위논문.
- 김주일(2017). 빅데이터를 활용한 공유숙박의 선택속성이 만족도·행동의도에 미치는 영향: 문화적응도 조절효과를 중심으로. 목포대학교 대학원, 박사학위논문.
- 김지숙(2013). 빅데이터 활용과 분석기법 고찰. 고려대학교, 석사학위논문.
- 김현성(2012). 빅데이터 시대에서 정보기술 패러다임의 변화와 스마트정부의 공공정보 활용방안. 현대사회와 행정, 22(3), 277-302.
- 김현철(2002). 기획여행상품 품질연구. 경기대학교대학원, 석사학위논문.
- 김현철(2012). 여행상품편익이 고객의 만족, 신뢰, 몰입과 충성도에 미치는 영향 연구. 경기대학교 관광전문대학원, 박사학위논문.
- 김형태(2018). 빅데이터 기반의 데이터 마이닝 기법을 활용한 품질혁신 방법론 연구: 디스플레이 제품의 불량개선 사례 중심으로. 서경대학교 대학원, 박사학위논문.
- 도해용(2014). 빅데이터를 이용한 한국·중국·일본의 레스토랑 브랜드 개선에 관한 연구. 세종대학교 대학원, 박사학위논문.
- 도해용·류기상(2013). 빅데이터 분석기법을 이용한 레스토랑 품질과 고객만족이 추천행동에 미치는 영향에 관한 연구. 외식경영학회, 16(2), 233-255.
- 문명희(2003). 여행상품 구성요소 중요도 연구: 호주 단체 패키지 여행을 중심으로. 경기대학교대학원, 석사학위논문.
- 백수진(2018). 관광 관련 가상현실체험의 현존감과 경험적 가치 및 관광지 방문의도의 구조적 영향관계: 경험적 가치의 매개역할을 중심으로. 세종대학교 대학원, 박사학위논문.
- 박로운(2017). 빅데이터 분석을 통한 지역관광 활성화 방안: 전라북도를 중심으로. 전주대학교대학원, 석사학위논문.
- 박선심(2016). 관광객의 관광지 선택속성이 지각된 가치, 자기효능감, 관광만족 및 재방문의도에 미치는 영향. 동신대학교대학원, 박사학위논문.

- 박상곤(2014). 관광 빅데이터 분석 대회와 활용 증진방안. 한국문화관광연구원, 36.
- 박정민(2012). 국외여행객의 여행상품선택속성에 관한 연구: 부산김해공항 출국 국외여행객을 중심으로. 관광레저연구, 24(8), 49-67.
- 반현정(2019). 빅데이터 분석을 통한 부산의 5성급 호텔과 비즈니스호텔의 의미연결망 분석에 관한 연구. 부경대학교대학원, 석사학위논문.
- 부려(2020). 모바일 여행상품 앱 특성이 경험가치, 지각된 가치, 지속적 이용의도에 미치는 영향: 관계혜택의 조절효과를 중심으로. 동아대학교 대학원, 박사학위논문.
- 사일(2017). 중국 관광객의 한국 전통시장 관광동기, 관광제약 요인이 지각된 가치 및 고객 만족에 미치는 영향. 계명대학교대학원, 박사학위논문.
- 서정아(2015). 사회연결망분석을 활용한 대구의 관광 이미지 분석: 온라인 빅데이터를 중심으로. 계명대학교대학원, 박사학위논문.
- 세계일보(2020.9.1.). 코로나 재확산에도... 국내 여행 다시 늘고 있다.
- 손기준·조인호·김찬우·진채남(2015). 하둡 기반 빅데이터 수집 및 처리를 위한 플랫폼 설계 및 구현. 한국콘텐츠학회, 297-298.
- 손동원(2012). 사회 네트워크 분석. 서울: 경문사
- 송수원(2010). 해외여행 상품의 서비스 품질속성에 대한 여행자의 만족도 분석. 고려대학교대학원, 석사학위논문.
- 송태민·송주영(2013). 빅데이터분석방법론 : 구조방정식모형과 다층모형을 중심으로. 서울: 한나래 출판사.
- 송호철(1992). 관광객의 여행상품 선택행동에 관한 연구. 경기대학교대학원, 석사학위논문.
- 심영석(2016). 텍스트마이닝(Text Mining)을 이용한 관광지 이미지 결정요인에 관한 연구. 세종대학교대학원, 석사학위논문.
- 아마르투스(2013). 관광객의 관광지 선택속성, 지각된 가치, 만족도 및 행동의도 간의 관계 차이연구: 해외관광경험이 있는 한국·몽골 관광객 중심

- 으로. 경희대학교대학원, 박사학위논문.
- 안명숙·오익근(2015). 네트워크 텍스트 분석 기법을 적용한 특급 호텔 패키지 이용에 대한 태도 분석. 관광연구, 30(5), 163-181.
- 연합뉴스(2018.9.3.). 올해 상반기 부산 방문 관광객 7%증가.
- 여호근·박경태(2007). 녹차밭 방문객의 지각된 가치가 만족과 재방문 의도에 미치는 영향. 한국조리학회, 13(2), 110-122.
- 여행신문(2017). “여행시장 호황에 상반기 여행주기도 ‘쭉쭉’.”
http://www.traveltimes.co.kr/bbs/board.php?bo_table=News&wr_id=100883&code=E27
- 연승호·엄서호(2004). 관광상품 가치지각에 관한 연구: 기획여행상품을 대상으로. 관광학연구, 27(4), 219-234.
- 연합뉴스(2017). SK텔레콤 빅데이터로 해운대 피서객 수 정밀 집계한다.
<http://news.naver.com/main/read.nhn?mode=LSD&mid=sec&sid1=103&oid=001&aid=0009449439>.
- 오익근·이태숙·전채남(2015). 빅데이터 분석을 통한 한국관광 인식에 관한 연구. 관광학연구, 39(10), 107-126.
- 오현주(2016). 힐링(Healing)관광 선택속성이 가치, 태도, 만족에 미치는 영향에 관한 연구, 경성대학교대학원, 박사학위논문.
- 윤대순(1997). 관광의 정의 확립에 관한 연구, 관광경영학회지, 1(1), 11-30.
- 윤종식(2016). 데이터분석 준전문가 30일 완성, 서울: 데이터에듀.
- 윤홍근(2013). 문화산업에서 빅데이터의 활용방안에 관한연구. 글로벌문화콘텐츠, 157-179.
- 이만재(2012). 빅 데이터 접근방식의 공공데이터 비주얼라이제이션 사례. 2 한국통신학회, 29(11), 36-42.
- 이상구(2012). 관광상품의 지각된 가치가 관광만족에 미치는 영향 : 한류의 조절효과를 중심으로. 인천대학교대학원, 박사학위논문.
- 이수상(2012). 네트워크 분석방법론. 서울: 논경.
- 이영진·윤지환(2014). 관광분야에서 SNS 빅데이터의 활용 방법 모색: 관광

- 정보 검색 키워드 분석을 중심으로. 한국관광연구학회, 13(3), 5-14.
- 이원희(2010). 온라인 구전이 여행상품 구매의도에 미치는 영향연구. 경기대학교대학원. 석사학위논문.
- 이일주(2019). 패키지여행상품의 편익과 관광지 매력성이 지각된 가치와 재구매에 미치는 영향 연구. 경기대학교대학원, 박사학위논문.
- 이종설(2015). 해외여행상품에 대한 지각된 가치가 브랜드인지도, 구매행동 및 브랜드충성도에 미치는 영향에 관한 연구. 경기대학교대학원, 박사학위논문.
- 이혜미(2018). 빅데이터를 활용한 공유숙박의 지각된 위험 인식에 관한 연구. 경희대학교대학원, 박사학위논문.
- 장미화·윤영일(2016). 소셜 미디어 빅데이터 분석을 통한 캠핑에 대한 정부 정책과 국민들의 인식 변화연구. 관광연구, 31(1), 91-112.
- 장병수(2004). 외래관광객의 여행상품 지각에 의한 관광목적지 가치 평가. 경주대학교대학원, 박사학위논문.
- 전형규(2008). 관광지의 특성, 관광자의 지각된 가치, 만족, 행동의도간의 관계연구. 계명대학교대학원, 박사학위논문.
- 정동현(2016). 빅데이터 분석을 통한 전통마을 관광 개선방안. 숭실대학교대학원, 석사학위논문.
- 정연상·이광열·김영균·유가형(2012). 빅데이터 분석을 통한 기업 미래 가치 창출, 삼정KPMG경제연구원.
- 정유준(2012). 지역관광브랜드 속성, 편익, 지각된 가치, 개발지지 간의 관계. 경기대학교일반대학원, 박사학위논문.
- 정윤희(2008). 패키지 관광상품의 지각된 가치가 관광객 만족도와 재구매의도에 미치는 영향 연구. 경희대학교대학원, 석사학위논문.
- 정익준(1997). 최신 관광사업론, 경기: 형설출판사.
- 정익준(2000). 최신 여행사경영론, 경기: 형설출판사.
- 정지선(2013). 빅데이터로 풀어본 대한민국 IT미래 먹거리. 한국정보화진흥원 빅데이터 전략연구센터.

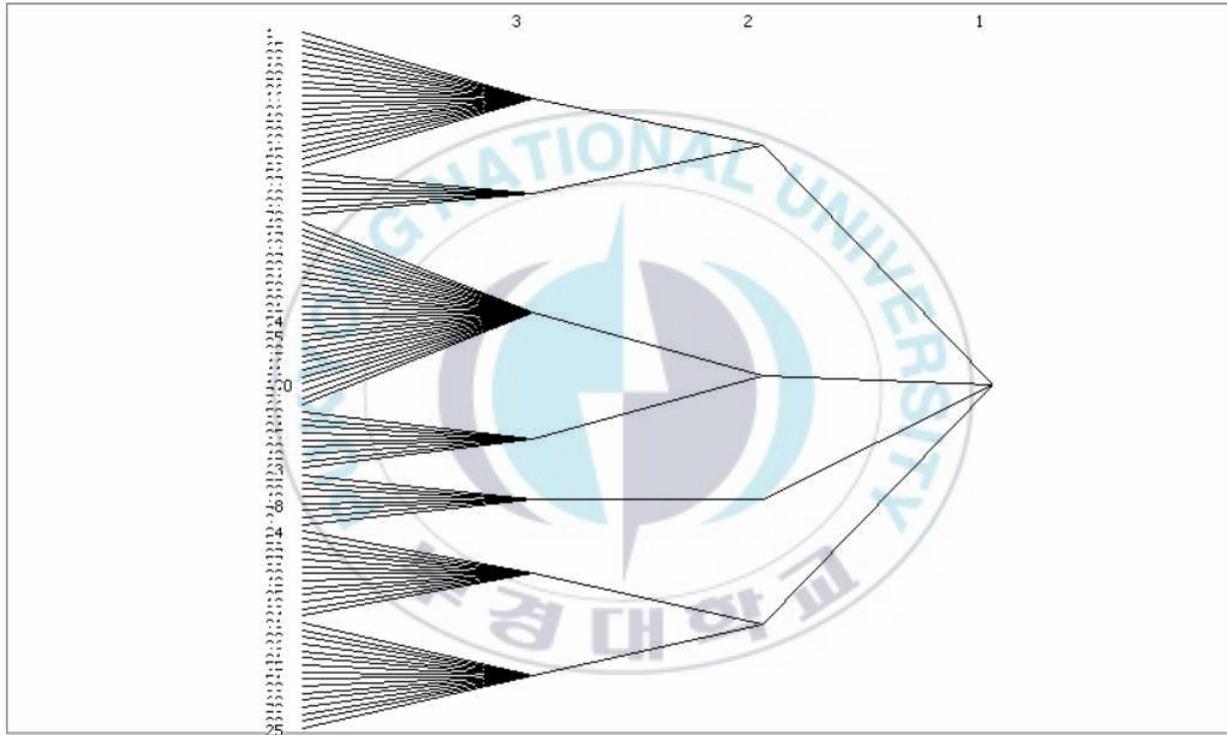
- 정철후(2013). 텍스트마이닝, 복잡계 네트워크, 인터넷 빅데이터를 활용한 미래예측 관련 중요키워드 도출 모델. 한양대학교대학원, 석사학위논문.
- 조완섭·조아·권가은·류관희(2015). SNS 데이터 분석을 통한 스마트 충북관광 구축. 한국데이터정보 과학지, 26(2), 409-418.
- 주상욱(2000). 관광객 유치를 위한 광주·전남지역 관광상품 개발방안에 관한 연구. 호남대학교대학원, 석사학위논문.
- 주영환(2007). 해외 골프관광객의 지각된 가치가 만족과 충성도에 미치는 영향. 한국관광연구학회, 21(2), 325-337.
- 주현식(2008). 호텔고객의 지각된 가치, 고객만족 및 전환비용이 행동의도에 미치는 영향. 한국호텔관광학회, 10(3), 116-129.
- 천덕희(2005). 여행상품 선택요인의 중요도 연구. 경기대학교대학원, 석사학위논문.
- 최승일·김동일(2007). 여행상품 서비스가 고객만족, 재구매의도에 미치는 영향에 관한 연구. 한국콘텐츠학회 종합학술대회 논문집, 5(1), 10-14.
- 한혜림(2017). 빅데이터 분석을 통한 중국 관광시장의 한국 관광 인식 분석. 한양대학교대학원, 석사학위논문.
- 한국관광공사(2017). 2017년 6월 외래객입국·국민해외여행객 및 관광수입·지출 동향.
- 한지연·김홍범(2017). 빅데이터 기반의 사회연결망분석을 이용한 관광지 이미징 인식에 관한 연구. 한국관광협회, 4(8), 91-119.
- 함유근·채승병(2012). 빅데이터 경영을 바꾸다, 삼성경제연구소.

국외 문헌

- Addis, M. & Holbrook, M. B.(2001). On the conceptual link between mass customization and experiential consumption: An explosion of subjectivity. *Journal of Consumer Behaviour*. 1(1), 50-66.
- Baloglu, S., & Mangalolu, Mehmet.(2001). Tourism destination images of Turkey, Egypt, Greece, and Italy as perceived by US-based tour operators and travel agents. *Tour Management*. 22(1), 1-9.
- Candi, M. & Kahn, K. B. (2016). Functional, emotional, and social benefits of new B2B services. *Industrial Marketing Management*. 57(1), 177-184.
- Childs, S. & Merv, A.(2012). Big data challenges for the IT infrastructure team. *Gartner Research*.
- Dodds, W. B.(1991). In search of value: How price and store name influence buyers product perception. *The Journal of Consumer Marketing*. 8(2), 15-24.
- Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., & Bryant, B. E. (1996). The American customer satisfaction index: Nature, purpose, and findings. *Journal of Marketing*, 60(October), 7-18.
- Gallarza(2004). Value dimensions, perceived value, satisfaction and loyalty: an investigation of university students' travel behaviour. *Tourism Management*. 27(3), 437-452.
- Gantz, J. & Reinsel, D.(2011). *Extracting value from chaos*. IDC.
- Gronroos, C.(2011). Value co-creation in service logic: A critical analysis. *Marketing Theory*, 11(3), 279-302.
- Kashyap, R. & Bojanic, DC.(2000). A structural analysis of value, quality, and price perceptions of business and leisure travelers.

- Journal of Travel Research*, 39(1), 45-51.
- Laney, D.(2001). 3D data management: Controlling data volume, velocity and variety. Gartner.
- Lapierre, J.(2000). Customer perceived value in industrial contexts. *Journal of Business & Industrial Marketing*, 15(2), 122-145.
- Ledden, L., Kalafatis, S. P., & Samouel, P.(2007). The relationship between personal values and perceived value of education. *Journal of Business Research*. 60(9), 965-974.
- Lee, C., Yoon, Y. & Lee, S.(2005). Investigating the relationship among perceived value, satisfaction, and recommendation: The case of the Korea DMZ. *Tourism Management*. 28(1), 204-214.
- Lewis. S., Zamith, R., & Hermida, A.(2013). Content analysis in an era of big data: A hybrid approach to computational and manual methods. *Journal of Broadcasting & Electronic Media*, 57(1), 34-52.
- Naylor, G. & Frank, K. E.(2001). The effect of price bundling on consumer perceptions of value. *Journal of Services Marketing*, 15(4), 270-281.
- McDougall, Gordon H. G.(2000). Customer satisfaction with services: putting perceived value into the equation. *Journal of Services Marketing*. 14(5), 392-410.
- Manyika, J. & Chui, M.(2011). (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute.
- Marr, B.(2015). A brief history of big data everyone should read. <https://www.weforum.org/agenda/2015/02/a-brief-history-of-big-data-everyone-should-read>.
- Patterson, P. G. & Spreng, R. A.(1997). Modelling the relationship between perceived value, satisfaction and repurchase intentions in a business to business, services context: an empirical examination.

- International Journal of Service Industry Management*, 8(5), 414-434.
- Prebensen, N. K. & Xie, J.(2017). Efficacy of co-creation and mastering on perceived value and satisfaction in tourists' consumption, *Tourism Management*, 60(1), 166-176.
- Sweeney, J. C. & Soutar, G. N.(2001). Consumer perceived value; The development of a multiple item scale. *Journal of Retailing*, 77(2), 203-220.
- Williams, P. & Soutar, G. N. (2009). Value, satisfaction and behavioral intentions in an adventure tourism context. *Annals of Tourism Research*, 36(3), 413-438.
- Woodruff, R. B.(1997). Customer value: The next source for competitive edge. *Journal of the Academy of Marketing Science*, 25(2), 139 - 153.
- Zeithaml, V. A.(1988). Consumer perceptions of price, quality, and value : A means-end model and synthesis of evidence. *Journal of Marketing*, 52(3), 2-22.
- Zikopoulos, P. & Eaton, C.(2011). *Understanding big data: analytics for enterprise class hadoop and streaming Data*. mcgraw-Hill Osborne Media.



<부록4> CONCOR CLUSTER DIAGRAM