

외래 관광객의 관광 주목적별 관광지 선택 연관성 규칙 분석: 2013년 외래 관광객의 휴가레저 및 비즈니스 집단별 비교 연구*

Analyzing associative rules of destination selections among foreign tourists:

A comparative study of associative rules between leisure and business tourist groups

이유안** 김인신***

Yuan Lee · Insin Kim

ABSTRACT

Applying data mining method to tourist spatial behaviors, the present study were to identify association rules of leisure and business tourist groups using dataset of 2013 international visitor survey. Associative rule mining is the most frequently used method in data mining, and this method is useful to explore an associative relationship between two items out of massive volume of data. In retail setting, this method is often used to demonstrate new findings of consumer behavior patterns. By utilizing SAS E-Miner, this study distinguish associative rules between leisure and business tourist groups with support, confidence, and lift of rules. According to the survey collected by the Ministry of Culture, and Tourism in Korea, leisure and business tourist groups were highly occupied among international tourists from the main purpose of foreign tourists. This study revealed that two tourist groups had a different tendency in visiting tourist attractions in Korea. The academic and practical implications of this study were discussed in this present study.

Key words : Data Mining. Associative Rule Mining. Travel Purpose. Business travellers. Destination Choice Behavior.

* 이 논문은 2015년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2015S1A5AB013532)

** 부산대학교 관광컨벤션학과 박사과정(totitre@hanmail.net)

*** 교신저자: 부산대학교 관광컨벤션학과 조교수(insinkim@pusan.ac.kr)

I. 서 론

다목적지 여행은 하나의 뚜렷한 관광 현상으로서, 오늘날 많은 관광객들은 한 번의 여행에서 여러 관광지를 방문하고 있으며, 이는 관광객의 공간 행동의 두드러진 특징 중의 하나이다(김민형 외 2인, 2011). 다목적지 관광 이동 연구를 통해서 관광 개발 측면에서 정책 형성자는 관광 이동 수요를 정확하게 이해함으로써 효과적인 지역 관광 개발을 도모할 수 있다(Lau and McKercher, 2006). 그 뿐만이 아니라, 관광 마케팅 측면에서 다목적지 여행 패턴을 이해함으로써 통해서 관광객 이동 특성을 반영한 관광 상품을 제공하고, 이에 따라서 관광객을 세분화할 수 있으므로 연구 필요성이 있다(Law et al. 2011). 그러나 위와 같은 관광객 이동 분석의 중요성에도 불구하고, 자료 수집 및 방법론의 한계로 인해서 관광객의 이동에 관한 경험적 사실을 발견하려는 노력은 몇몇 소수의 연구들을 제외하고 상대적으로 미흡한 실정이다.

관광객 다목적지 여행 패턴에 관한 국내 선행 연구로서 이혜진과 송운강(2010)은 설악산 국립공원 방문객들을 대상으로 로짓 모형을 이용하여 다목적지 관광객들이 단일 목적지 관광객 보다 평균 여행 거리, 체재 일수가 더 길고, 동반 인원수가 더 많고, 숙박율과 자가용 이용률도 더 높은 경향성을 발견하였다. 그리고 김민형 등(2012)는 빈도 분석, 교차 분석, 이항 로짓 분석 방법을 사용하여 중국 국적 관광객의 이동 패턴을 분석하였다. 이 연구에 따르면 수도권 지역만을 방문할 경우를 단일 목적지 여행으로 규정하고, 수도권 및 다른 광역시/도 지역을 방문하였을 경우를 다목적지 여행으로 정의하고 있다.

위와 같은 관광객 이동에 관한 소수의 국내 선행 연구들은 관광객의 다목적지 관광 패턴의 영향 관계를 파악하는 측면에서 중요성이 있으나, 특정 목적지에 한정되어 있거나, 또는 다목적지 여행에 관한 정의상의 차이로 인해 관광객의 세부적인 목적지 간의 연계성을 검증하지 못한다는 점에서 아쉬움이 있다(김민형 외, 2012; 이혜진·송운강, 2010). 따라서 인구 통계학적 특성, 이동 거리, 자가용 이용률 등의 관점이 아닌 다른 관점과 새로운 접근법에 의해서 관광객 이동에 관한 연구가 실시되어 보다 정교한 시장 세분화의 필요성이 있다.

위와 같은 배경에서 본 연구는 데이터 마이닝 방법론의 하나인 연관성 규칙 마이닝 기법을 활용하여 방한 외래 관광객의 관광 주목적별 관광지 선택 행동의 연관성 규칙 도출 및 분석에 목적을 두고 실시되었다. 데이터 마이닝은 고객 관점의 시장 발

견을 위한 새로운 방법론으로서 부상하고 있는 방법론으로서 데이터베이스에 저장된 대용량의 데이터를 조사함으로써 기존에 알려지지 않은 의미 있는 새로운 상관관계, 특성, 트렌드 등 유용한 정보 발견을 의미한다(정유경, 2003). 본 연구는 데이터 마이닝 중에서 대표적으로 사용되는 연관성 규칙 마이닝 절차에 따라서 방한 외래 관광객의 관광 주목적별 관광지 집단별로 차별적으로 형성되는 관광지간의 연관성 규칙을 도출함으로써, 관광 주목적 관광객 집단에 따라서 관광 이동 행태 및 관광지 선택 행동에서 어떠한 차이가 발생하는지 검토하고자 한다.

본 연구는 다음과 같은 구체적인 연구 목적에 초점을 두고 있다. 2013년 방한 외래 관광객의 관광 주목적 기준에 따른 가장 큰 비중을 차지하는 휴가·레저 관광객과 비즈니스 관광객 집단을 대상으로 관광지 사이에 형성되는 연관성 규칙 차이를 비교하고 검증하는 것이다. 그리고 본 논문은 연관성 규칙 마이닝 기법을 활용하고 있는 선행 연구(Law et al. 2011; Versichele et al. 2014; 정유경, 2003)의 논문 서술 순서 양식을 참조하여 방법론에 관하여 우선적으로 설명하고, 이러한 방법론을 적용한 문헌 연구들을 검토한 뒤, 도출된 연관성 규칙 분석 및 해석한 다음, 끝으로 본 연구의 이론적, 실무적 시사점을 제시하였다.

II. 이론적 배경

1. 데이터 마이닝

1) 데이터 마이닝의 개념

데이터 마이닝이란 대용량 베이스로부터 변수들 간 연계적 관계를 밝혀내는 기법으로서 기존의 방대한 데이터 변수간의 유용한 상관관계, 경향성과 같은 새로운 정보를 발견하고, 이를 토대로 미래에 실행할 수 있는 정보를 추출하는 것을 의미한다(남기성 외 2인, 2002). 국내에서 데이터 마이닝 기법이 실제 비즈니스 업무에 적용된 최초의 사례는 1977년 BC 카드사의 부정 사용자 적발이며, 이후 통신, 금융 회사의 가입 고객 이탈 방지, 카드 혹은 휴대폰의 도난 도용 방지, 유통업의 상품 진열 배치 등을 위해서 사용되었다(채경희·김상철, 2010). 데이터 마이닝의 세부 기법은 인공 신경망 분석, 의사결정 나무, 분류, 유사성 집단화, 연관성 규칙, 군집화, 프로파일링, 시퀀셜

패턴 마이닝 등으로 분류될 수 있다(Versichele et al. 2014).

특히, 데이터 마이닝은 데이터베이스에 축적된 데이터를 마케팅 목적에 맞도록 정제하여 기업의 고객 관계 관리(customer relationship management) 방안을 결정할 때 근거가 되는 자료를 제공해준다. CRM은 영업 자동화와 같은 운영 CRM, 콜센터 자동화 등 고객 접점과 관련된 협업 CRM, 고객 데이터 분석 활용과 관련된 분석 CRM으로 분류할 수 있다(양승정·이종태. 2007). 특히, 데이터 마이닝 기법은 분석 CRM의 도구로써 활용 가능성이 크며 보유 고객의 데이터를 기반으로 고객 구매 행태 예측, 고객의 이용 특성, 고객 가치 분석을 이해하는데 핵심을 두고 있으며, 고객 특성에 따른 차별적 마케팅 전략 설계 시에 데이터 마이닝 기법이 적극적으로 이용되고 있다(Rygielski et al. 2002). 다시 말해 데이터 마이닝은 데이터베이스를 활용한 마케팅에 적극적으로 활용되어 기존 고객 유지 및 신규 고객 창출에 기여할 수 있다(Ngai et al. 2009).

2) 관광학 분야의 데이터 마이닝 선행 연구

관광학 분야에서 연관성 규칙 마이닝 이외의 다른 데이터 마이닝 기법을 활용한 선행 연구들은 다음과 같다. 김민정(2001)은 클러스터링, 연관성 규칙, 분류 기법을 적용하여 고객 웹 로그 및 웹 콘텐츠를 분석하여 고객 요구 사항과 고객 웹 데이터 사용과의 상관관계를 검토하였다. 즉, 첫째, 클러스터링 기법을 사용하여 사용자 특성을 분류하고, 둘째, 연관성 규칙을 통해서 웹 사이트 이용 유사도를 검토하고, 셋째, 분류 기법을 활용하여 그룹 특성에 따른 이용도 예측 모형 구축함과 더불어 웹 사이트 이용 고객 관리 방안을 제안하였다.

김완섭 등(2010)은 의사 결정 트리 모델을 활용하여 신혼여행 상품 시장 세분화를 시도하였다. 이 연구는 여행사 신혼여행 상품 구매 데이터 5,482건의 자료를 사용하여 신혼여행 상품 구매 특성을 분석하였다. 이를 통해 신혼여행 상품 구매의 주요인이 단체 판매가와 영업 담당 부서임을 도출하였다. 이를테면 판매가가 86만 원 이하인 경우 제주도 여행 비율이 높은 경향성을 발견하였다. 한편, 김창현 등(2003)은 데이터 마이닝 기법 중에서도 신경망 모형과 의사 결정 나무 분석을 적용하였다. 이 선행 연구 결과에 따르면 서울 시내 특정 호텔의 식품 및 음료 구매를 이용한 13,445명의 고객들의 구매 이력을 활용하여 인구 통계학적 특성 및 구매 특성에 따라서 개별 고객의 향후 이용 금액 예측 모델을 구축하였다.

변우희 등(2012)는 데이터 마이닝 및 상황 평가 방법론을 활용하여 지역 관광 수요를 반영한 관광 자원 정보 융합 서비스 모형을 구축하였다. 즉, 데이터 마이닝 기법을 활용하여 다양한 지역 관광 자원의 잠재 가치가 평가되고 재배치되는 관광 정보 서비스 통합 모형을 개발하였다. 다시 말해 이 연구는 데이터 마이닝 프로세스를 통해서 지역 자원의 특수성이 융합된 정보가 효율적으로 관광객에게 전달될 수 있도록 정보 관리 청사진을 제시하여 관광 정책 효과성을 증진시키고자 하였다.

외식 경영에서 데이터 마이닝 기법을 적용한 연구로는 김영훈과 안성식(2006)은 데이터 마이닝 기법을 적용하여 패스트푸드 이용 고객의 특성을 파악하는 연구를 실시하였다. 롯데리아, 맥도날드, 기타 브랜드 사용자 집단 특성을 의사 결정 나무 모형에 의해 1인 지출 비용을 분석한 결과, 롯데리아 선호 고객은 1인 평균 지출 비용이 5천원~7천원 미만인 반면, 맥도날드 선호 고객은 1인 평균 지출 비용이 3천원 미만 이면서 패스트푸드 레스토랑을 거의 매일 이용하는 월평균 외식비용이 5만 원 미만인 고객들이라는 사실을 도출하였다. 또한 박기용과 이훈영(2003)은 의사 결정 나무의 CART(classification and regression tree) 기법을 이용하여 패밀리 레스토랑 추천 의사 증진을 위한 이용 고객 행태를 분석하였다. 즉, 인구 통계적 특성에 따른 의사 결정 나무 모형에서 패밀리 레스토랑 추천 의도에 영향을 미치는 가장 핵심 변수는 결혼의 유무임을 발견하였고, 인구 통계적 변수, 편익 변수, 그리고 레스토랑 선택 변수를 동시에 반영한 패밀리 레스토랑 의사 결정 모형을 제시하고 있다.

2. 연관성 규칙 마이닝

1) 연관성 규칙 마이닝의 개념

본 연구는 데이터 마이닝의 대표적인 기법 중의 하나인 연관성 규칙 마이닝을 적용하였는데 이 기법은 Agrawal et al.(1993)에 의해 도입된 방법으로서 연구 결과에 대한 선형적 가정 없이 데이터에 변수들 간 숨겨진 연관성을 찾는 기법이다. 연관성 규칙 마이닝은 사용자 구매경향성 파악에 빈번히 활용되므로 장바구니 분석으로도 불린다(조광현·박희창. 2012) 즉, 어떤 항목이 또 다른 항목을 수반하는가를 탐색하기 위한 방법으로 항목 X와 항목 Y 사이의 X-> Y 형태의 규칙을 찾아내어 다수의 다양한 사건들 중 두 가지 이상의 사건이 동시에 일어날 가능성을 발견하는 것이다(노미정 외 2인. 2010). 즉, 연관성 규칙은 If X Then Y 형식의 X가 존재하면, Y도 존재할

가능성이 높다는 연관성을 의미한다. 연관성 규칙은 탐색적이고 비목적성 분석이며, 기존의 데이터를 특별한 변형 없이 계산이 용이하게 사용 가능하다는 장점을 가지고 있다(조광현·박희창, 2012). 연관성 분석은 선형성, 정규성, 등분산성 가정을 필요로 하지 않으므로, 두 품목간의 관계를 명확히 수치화함으로써 두 개 이상의 품목간의 관련성을 표시하여 주기 때문에 현업에서 다양한 분야에서 활용되고 있다(노미정 외 2인, 2010). 연관성 규칙을 발견하는 알고리즘 중에서 가장 먼저 개발되었고, 가장 많이 사용되고 있는 것은 Apriori 알고리즘이며 본 연구도 이 알고리즘을 활용하였다(문영수, 2005). Apriori 알고리즘은 두 가지 단계로 구성된다. 첫 번째 최소 향상도 및 지지도 설정 값에 따라 빈도수가 높은 항목의 집합들을 찾아내고, 이들 집합들 중에서 주어진 신뢰도 또는 지지도를 만족하는 연관성 규칙을 찾아낸다. 지지도, 신뢰도, 향상도 기준에 의해 의미 있는 연관성 규칙을 판단할 수 있다(문영수, 2005). 첫째, 지지도(support)는 연관성 규칙의 기초를 이루는 것으로써 전체에 대한 A 항목과 B 항목이 동시에 일어날 수 있는 확률을 의미한다(임제순 외 2인, 2010). 즉 두 항목 A와 B의 지지도는 전체 거래 항목 중에서 A와 B를 동시에 포함하는 거래의 비율을 말한다. 어떤 규칙의 지지도가 10%라면 전체 거래 내역 중 그 규칙에 의한 거래 내역이 10%의 비율을 차지한다는 의미이다. $Support(A \rightarrow B)$ 은 전체 케이스 수 값을 A 항목과 B 항목을 동시적으로 포함하는 케이스 수로 나눈 값을 의미하고, $Support(A \rightarrow B)$ 및 $Support(B \rightarrow A)$ 가 대칭적으로 같은 값을 지니며, $Pr(A \cap B)$ 의 값으로 표기할 수 있다(임제순 외, 2010). 둘째, 신뢰도(confidence)란 항목 A의 거래 중에서 항목 B가 포함된 거래의 비율을 뜻한다. 신뢰도란 지지도의 단점을 보완할 수 있는 지표로써, A가 발생한 경우 중에서 B가 발생할 경우의 조건부 확률을 뜻한다(임제순 외, 2010). $Confidence(A \rightarrow B)$ 란 항목 A를 포함하는 케이스 수 값을 항목 A와 B를 동시에 포함하는 케이스 수 값으로 나눈 값을 의미하며, $Pr(A \cap B)$ 와 $Pr(A)$ 의 곱으로 표기된다(임제순 외 2인, 2010). 셋째, 향상도는 $A \rightarrow B$ 의 연관 규칙에서 임의로 B가 구매되는 경우에 비해 A와의 관계가 고려되어 구매되는 경우의 비율을 의미한다(임제순 외 2인, 2010). $Lift(A \rightarrow B)$ 는 항목 A를 포함하는 케이스와 항목 B를 포함하는 케이스 수를 곱한 값을 항목 A와 B를 동시에 포함하는 케이스 수와 전체 케이스 수의 곱한 값으로 나눈 값을 의미하며, $Pr(A \cap B)$ 을 $Pr(A) Pr(B)$ 으로 나눈 값으로 표기된다(임제순 외 2인, 2010).

2) 관광학 분야의 연관성 규칙 마이닝 선행 연구

관광학 분야에서 데이터 마이닝 기법 중 특히, 연관성 규칙 마이닝이 적용된 연구는 비교적 소수이나, 이를 소개하면 다음과 같다. 정유경(2003)은 호텔 레스토랑 고객 이용 행태를 분석하기 위해 연관성 규칙 마이닝을 활용하였고, 연회와 클럽을 이용하는 고객 행태에서 낮은 지지도를 나타내며, 이를 통해 클럽을 자주 방문하는 고객들은 다른 레스토랑을 동시에 이용하지 않는다는 사실을 발견하였다. 한편 지지도와 리프트 값이 높은 연관성 규칙을 통해서 이탈리아 레스토랑에 자주 가는 고객들은 로비라운지와 뷔페에도 자주 간다는 사실 등을 발견하였다. 그리고 김창현 등(2003)은 호텔 이용 고객의 이용 자료로부터 특정 호텔의 식음료장 이용 고객 구매 특성에 따른 데이터베이스 마케팅 전략을 수립하였다. 즉, 호텔 이용 고객의 인구 통계학적 특성, 동행인 유무, 구매 금액, 구매 목록, 지불 방식을 고려한 마케팅 전략 수립을 위한 데이터베이스 모형을 제안하고 있다.

Versichele et al.(2014)은 벨기에 서북부에 위치하고 있는 항구 도시인 겐트(Gent) 시 방문객의 이동에 관한 연구를 실시하였다. Versichele et al.(2014)의 연구는 블루투스 이동 데이터를 활용하여 관광객의 순차적 이동에 관한 연관성 규칙들을 제시하였다. Versichele et al.(2014)는 입장을 위한 티켓 구매가 필요한 비개방형 관광지와 입장료를 요구하지 않는 개방형 관광지와 같이 두 가지 특성으로 관광지를 분류하고 있다. 연구 분석 결과, 개방형/비개방형 관광지를 방문하는 관광객의 관광지 특성 선택의 일관된 경향성이 발견되었다. 즉, 개방형 관광지를 선호하는 관광객의 80%가 비개방형 관광지를 방문하지 않았고, 비개방형 관광지를 선호하는 관광객의 56%는 개방형 관광지를 방문하지 않았으며, 단지 16%의 관광객만이 두 특성의 관광지를 혼합해서 모두 방문하는 경향이 있는 것으로 연구 결과에서 나타났다. Rong et al.(2012)는 연관성 규칙 마이닝 프로세스를 통해서 마카오 지역 주민의 아웃 바운드 관광객 이동 추세 변화를 비교하였다. 즉, 2005년과 2009년의 마카오 지역 주민을 대상으로 관광 동기, 관광 목적지, 향후 관광 의도라는 변수들을 사용하여 이들의 아웃 바운드 관광 패턴을 검토하였다. 즉, 관광 동기 측면에서 휴가·레저, 친지 방문, 비즈니스, 지식 추구 관광 동기 관광객 등으로 분류하고, 관광 목적지는 중국 본토(mainland)와 중국 외 해외 다른 국가(overseas)로 구분하였다. 해당 연구에서 발견된 연관성 규칙 일부를 설명하면 2005년 휴가·레저 관광 동기 관광객의 32.61%가 관광객이 향후 아웃 바운드 관광 의향이 있으며, 2009년에는 그 비율이 82%로 증가하였다. 한편 지식 추구 동기 관광객의 중국 본토 방문 비율은 2005년에서 2009년 사이에 감소하였다.

이처럼 연관성 규칙 마이닝은 기존에 발견되지 않은 변수간의 상관성을 발견하는데 유용하다. 예를 들어 미국 월마트는 고객의 구매 목록 데이터베이스를 수집, 분석하여 상품간의 관계를 파악한 결과, 맥주를 구매한 고객이 기저귀를 함께 구매하는 빈도가 높다는 경향성이 드러났다. 맥주와 기저귀는 연관성이 없어 보이는 상품 조합이나 기저귀를 구매하기 위해 마트에 온 남편들이 맥주 구매 의도가 높아서 두 상품의 공동 구매 경향성이 발견된 것이다. 이후 두 상품을 근접한 위치에 전시하여 매출의 동반 상승효과를 달성할 수 있었다(정유경, 2003). 이처럼 관광객 이동 분야에서 연관성 규칙 마이닝 방법론은 Rong et al.(2012) 또는 Versichele et al.(2014)와 같은 국외 연구자들에 의해서 적용된 바가 있다. 이러한 선행 연구들은 도출된 연관성 규칙을 토대로 해당 지역의 관광객의 이동 특징 및 추세를 분석하고 있으며, 관광 개발 및 마케팅 측면에서 시사점을 제공하고 있다.

Ⅲ. 연구설계

1. 연구 가설 및 수행 절차

본 연구 핵심 목표는 방한 외래 관광객의 관광 주목적별 관광객 집단에 따라 차별적으로 형성되는 방문한 관광지간의 연관성 규칙을 분석하고, 이를 통해 방한 외래 관광객의 이동 행태 및 관광지간의 관계적 특성을 검토하는 것이다. 즉, 본 연구는 휴가·레저 및 비즈니스 관광객과 같은 관광 주목적 특성에 따라 관광 이동 측면의 차이가 발생하였음을 검증한 Rong et al.(2012)과 Lau and Mackercher(2006)의 선행 연구 결과를 토대로 하여, 두 관광객 집단 간 관광지 연관성 규칙이 다르게 형성될 것이라고 가정하고, [그림 1]과 같은 연구 수행 절차에 따라서 이를 검토하였다.

2013년 외래 관광객 실태조사에 의하면 관광 주목적에 관하여 휴가/레저, 뷰티 건강 및 치료, 종교 및 순례, 쇼핑, 친구/친지 방문, 사업/전문 활동, 교육으로 구분하고 있다. 특히, 본 연구에서 특히 휴가·레저, 비즈니스 관광객이라는 관광 주목적별 집단을 대상으로 분석을 실시한 이유는 관광 주목적 집단 중에 비중이 가장 높은 비율을 차지하고 있기 때문이다. 즉, 2013년 외래 관광객 실태 조사 설문 문항 중 한국 방문의 주목적 문항과 한국에서 방문한 관광지 체크 문항에 주목하여, 관광지간의 관계형 데이터의 형태로 재구성하여 연관성 규칙을 도출하였다. [그림 1]과 같이 연구 수

행 절차에 따라서 연구가 실시되었다.



[그림 1] 연구 수행 절차

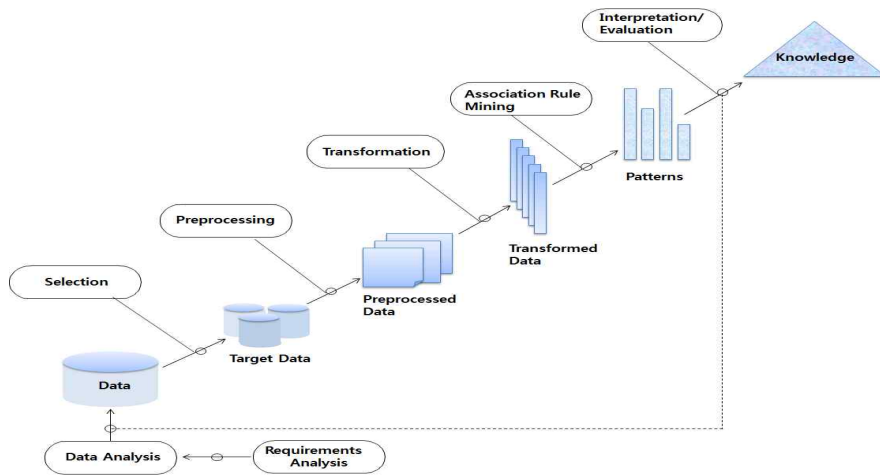
2. 자료의 특성

본 연구의 목적을 달성하기 위하여 2013년 외래 관광객 실태 조사의 설문 문항 중 관광 주목적과 방문한 관광지에 관한 응답 자료를 활용하였다. 외래 관광객 실태 조사의 데이터를 활용한 이유는 해당 조사가 통계청이 승인하는 조사로서 한국 문화 관광 연구원에 의해서 체계적으로 자료가 수집되며, 관광지 방문 항목과 더불어 조사 응답자 특성에 관한 설문 문항을 포함하고 있기 때문이다. 즉, 자료 수집을 위해 인천, 김포, 김해 국제공항과 국제 여객 터미널에서 출국 대기 중인 15세 이상의 외래 관광객들을 대상으로 매달 1000명의 설문지가 외국어 활용이 가능한 면접원에 의해서 수집된다. 이 조사는 외국인 및 교포를 모집단으로 하며, 체류기간이 1박 이상 90박 이하인 15세 이상의 외래 관광객을 표본 대상으로 하고 있다. 그리고 국가별 유의 할당 및 비례 배분 할당과 같은 표본 추출법을 적용하여 한 국가에 편중된 자료 수집을 지양(止揚)하고 있어 자료의 신뢰성이 있다. 본 연구의 분석이 실시된 2015년 여름 당시 관광 정보 데이터 시스템에 업로드된 최신의 자료인 2013년의 응답 자료를 활용하여 관광 이동에 관한 탐색적 연구가 시행되었다.

3. 자료 분석 방법

본 연구는 [그림 2]의 절차를 거쳐 자료 분석을 시행하였다. 첫 번째, 연구 목적 달성에 필요한 데이터베이스를 선정하고 필요한 데이터만 선택하여 추출하고 저장한다.

두 번째, 데이터 전처리 단계에서 아이디 부여하고 데이터의 중복을 제거하는 등 데이터 클리닝을 실시한다. 즉, 누락된 데이터에 대해서는 변수의 평균값, 중위수, 혹은 중간값 등으로 대체시킨다(김기범·안병구. 2013). 셋째, 전치 단계에서 연관성 규칙 마이닝 실시를 위해 적합한 형식으로 변환하고 분리하고 통합한다. 즉, 대용량 데이터를 엑셀 및 SPSS 또는 SAS를 사용하여 사용자가 파악하고 싶은 연결 관계 유형에 따라서 링크 데이터로 가공할 필요성이 있는지 판단한다(홍순만. 2013). 넷째, 연관성 분석 마이닝을 통해 데이터베이스에서 동시 출현 빈도가 빈번한 항목 규칙들을 추출하고 해석한다. 본 연구는 연관성 규칙 마이닝의 도출된 규칙을 판단하는 일반적인 기준인 최소 향상도인 1의 값을 넘어서고, 지지도가 높은 규칙들을 중심으로 검토하였다(Park. 2010).



[그림 2] 연관성 분석 수행 절차

출처: Fayyad, Piatestsky, Symth(1996)을 참고하여 재작성.

IV. 분석결과

1. 조사 대상자의 인구 통계학적 특성

2013년 외래 관광객 중 조사 대상자의 인구 통계학적 특성은 <표 1>과 같다. 성별은 남자 47%(5652명)와 여자 53%(6378명)로서 여자가 남자보다 조금 더 높은 비율을 차지하고 있다. 연령은 전체적으로 10대에서 60대까지 다양하게 나타났다. 15세 이상

외래 관광객의 관광 주목적별 관광지 선택 연관성 규칙 분석

~20세 이하가 5.4%(653명), 20세 이상~29세 이하가 33.3%(4011명), 30세 이상~39세 이하가 26.9%(3237명), 40세 이상~49세 이하가 18.5%(2231명), 50세 이상~59세 이하가 10.8%(1298명), 60세 이상이 4.7%(566명)으로 구성되어 있다. 방한 주목적 측면에서 휴가·레저 관광객은 48.2%(5798명), 비즈니스 관광객은 23.4%(2819명)을 차지한다. 또한 쇼핑 주목적 관광객은 10.6%(1281명), 친지 방문은 9.7%(1166명), 교육 주목적은 5.9%(714명), 건강 치료 주목적은 1.0%(118명), 종교 주목적은 1.0%(122명)의 비율을 구성하고 있다. 또한 최초 방문객은 60%, 2회 방문객은 14%, 3회 방문객은 8%, 4회 이상 방문한 관광객은 17.7%의 비율을 차지하고 있다.

<표 1> 자료의 인구통계학적 특성

Characteristic		N(%)	Characteristic		N(%)	
성별	남자	5652(47)	직업	공무원/ 군인	23(6.7)	
	여자	6378(53)		기업인/ 경영직	69(20.1)	
연령	15-20	653(5.4)		사무/ 기술직	26(7.6)	
	21-30	4011(33.3)		판매/ 서비스직	5(1.5)	
	31-40	3237(26.9)		전문직	31(9.0)	
	41-50	2231(18.5)		생산/기능/ 노무직	12(3.5)	
	52-60	1298(10.8)		자영업자	34(9.9)	
	60세 이상	566(4.7)		학생	6(1.7)	
	모름/무응답	34(0.3)		주부	3(0.9)	
방한 주목적	휴가/레저	5798(48.2)		은퇴자	104(30.3)	
	뷰티 건강 및 치료	118(1.0)		무직	6(1.7)	
	종교 및 순례	122(1.0)		기타	18(5.2)	
	쇼핑	1281(10.6)		무응답	5(1.5)	
	친구/ 친지 방문	1166(9.7)		방한 횟수	최초	7,257(60.3)
	사업/ 전문 활동	2819(23.4)			2회	1,676(13.9)
	교육	714(5.9)			3회	963(8.0)
기타	12(0.1)	4회 이상			2,134(17.7)	
여행 형태	개별 여행	7,964(66.2)		국적	중국	4402(36.6)
	단체 여행	3,327(27.7)			일본	3135(26.1)
	Air-Tel	739(6.1)	미국		773(6.4)	
동반자 유무	동반자 있음	68.3	대만		620(5.1)	
	혼자 왔음	31.7	홍콩		455(3.8)	
	1명	44.1	태국		362(3.0)	
동반 인원수	2명	19.3	말레이시아		217(1.8)	
	3명 이상	36.6	싱가포르		183(1.5)	
	가족/친지	45.7	캐나다		140(1.2)	
	친구/연인	39.2	중동		137(1.1)	
동반 유형	직장동료	19.6	러시아		136(1.1)	
	기타	0.2	영국		123(1.0)	
	전체	12,030(100)	기타		1347(11.3)	
			전체		12,030(100)	

2. 휴가·레저 관광객 연관성 규칙 분석 결과

일반적으로 향상도가 일 이상이고, 최저 지지도를 만족하는 규칙들 중에서 최저 지지도 기준을 초과하는 경우에 일반적으로 연관성 규칙이 생성되는 것으로 간주한다 (Park, 2010). 즉, 향상도 일 이상일 경우, 해당 연관성 규칙이 기본적인 신뢰 조건을 충족시키고 있으며, 지지도가 높을수록 전체 케이스 수에서 비중이 있는 연관성 규칙임을 판단할 수 있다(김용, 2012). 따라서 본 연구에서는 향상도 일 이상이면서 지지도가 가장 높은 각 집단별 연관성 규칙을 제시하였다.

<표 2>는 휴가·레저 관광객 집단의 지지도가 높은 연관성 규칙들을 제시하고 있다. <표 2>의 규칙 1, 규칙 2의 동대문 시장 -> 남대문 시장이 의미하는 바는 동대문 시장 -> 남대문 시장의 지지도가 25.65%는 휴가·레저 관광객들이 방문한 관광지들 중에서 동대문 시장, 남대문 시장 두 장소 모두 방문한 확률을 의미한다. 그리고 규칙 3의 동대문시장 -> 남대문시장의 신뢰도 43.86%는 휴가·레저 관광객들이 동대문 시장을 방문한 경우 남대문 시장도 방문한 확률을 뜻한다. 그리고 <표 2>의 규칙 5, 6, 13, 14, 15, 16, 17, 18에서 알 수 있듯이 남산 N 서울 타워는 고궁, 동대문 시장, 남대문 시장, 롯데월드, 박물관 장소에 관광지 연관성을 형성하고 있음을 알 수 있다. 즉, 본 연구의 지지도 상위 규칙 결과에 따르면 남산 N타워는 고궁, 동대문 시장, 남대문 시장, 롯데월드, 박물관과 연관성 규칙을 형성하며 관광지 사이의 중개자 역할을 수행하고 있다는 특징이 있다.

<표 2> 휴가·레저 관광객의 연관성 규칙

구분	규칙	신뢰도	지지도	향상도	방문자수
규칙 1	동대문시장 -> 남대문시장	43.86	25.65	1.33	1303
규칙 2	남대문시장 -> 동대문시장	77.65	25.65	1.33	1303
규칙 3	롯데월드 -> 동대문시장	75.32	23.08	1.29	1172
규칙 4	동대문시장 -> 롯데월드	39.45	23.08	1.29	1172
규칙 5	남산N서울타워 -> 고궁	62.05	21.60	1.44	1097
규칙 6	고궁-> 남산N서울타워	50.11	21.60	1.44	1097
규칙 7	신촌/홍대 -> 명동	90.05	21.56	1.28	1095
규칙 8	명동-> 신촌/홍대	30.61	21.56	1.28	1095
규칙 9	박물관-> 고궁	77.26	21.34	1.79	1084
규칙 10	고궁-> 박물관	49.52	21.34	1.79	1084
규칙 11	동대문시장-> 광화문광장	29.22	17.09	1.34	868
규칙 12	광화문광장 -> 동대문시장	78.34	17.09	1.34	868
규칙 13	남대문시장->남산N서울타워	51.55	17.03	1.48	865
규칙 14	남산N서울타워->남대문시장	48.93	17.03	1.48	865
규칙 15	롯데월드->남산N서울타워	45.95	14.08	1.32	715
규칙 16	남산N서울타워-> 롯데월드	40.44	14.08	1.32	715
규칙 17	박물관 -> 남산 N서울타워	50.53	13.94	1.45	708

외래 관광객의 관광 주목적별 관광지 선택 연관성 규칙 분석

규칙 17	박물관 -> 남산 N서울타워	50.53	13.94	1.45	708
규칙 18	남산 N 서울타워 -> 박물관	50.46	13.94	1.45	708
규칙 19	북촌 -> 명동	92.32	13.01	1.31	661
규칙 20	명동 -> 북촌	18.48	13.01	1.31	661

둘째, 2013년 휴가·레저 관광객 집단의 연관성 규칙에 의해 출현된 관광지 연관성 규칙에 의하면 비즈니스 관광객 결과와 차이점을 검토한 결과, 가장 주된 특징은 <표 3>과 같이 제주 지역 중심의 고립된 관광지 클러스터가 등장한 점이다. 즉, 한반도 안의 어떤 관광지와의 연관성이 전혀 발생하지 않은 상태로서, 제주도지역 안에서만 형성된 관광지 연관성 규칙이 등장하였다. 다시 말해, 제주 지역의 관광지 연관성 규칙은 비즈니스 관광 동기 집단에는 출현하지 않은 관광지 클러스터가 등장하였다.

<표 3> 휴가·레저 관광객의 제주 지역의 연관성 규칙

구 분	규칙	신뢰도	지지도	향상도	방문자수
규칙 1	성산일출봉 -> 제주민속촌	68.60	8.86	6.27	450
규칙 2	제주민속촌 -> 성산일출봉	80.94	8.86	6.27	450
규칙 3	성산일출봉 -> 용두암	63.87	8.25	6.29	419
규칙 4	용두암 -> 성산일출봉	81.20	8.25	6.29	419
규칙 5	제주민속촌 -> 용두암	66.73	7.30	6.57	371
규칙 6	용두암 -> 제주민속촌	71.90	7.30	6.57	371
규칙 7	한라산 -> 성산일출봉	74.94	6.60	5.80	335
규칙 8	성산일출봉 -> 한라산	51.07	6.60	5.80	335
규칙 9	한라산 -> 제주민속촌	65.32	5.75	5.97	292
규칙 10	제주민속촌 -> 한라산	52.52	5.75	5.97	292
규칙 11	한라산 -> 용두암	61.07	5.38	6.01	273
규칙 12	용두암 -> 한라산	52.91	5.38	6.01	273
규칙 13	중문관광단지-> 성산일출봉	73.10	4.92	5.66	250
규칙 14	성산일출봉 -> 중문관광단지	38.11	4.92	5.66	250
규칙 15	중문관광단지 -> 용두암	59.94	4.04	5.90	205
규칙 16	용두암 ->중문관광단지	39.73	4.04	5.90	205
규칙 17	중문관광단지-> 제주민속촌	56.14	3.78	5.13	192
규칙 18	제주민속촌 -> 중문관광단지	34.53	3.78	5.13	192

2. 비즈니스 관광객 분석 결과

2013년 비즈니스 관광 주목적 외래 관광객 집단으로부터 도출된 상위 빈도 방문지 연관성 규칙은 <표 4>의 결과와 같다. <표 4>의 규칙 1에서 보듯이, 비즈니스 관광객

집단 중 지지도가 가장 높은 연관성 규칙은 고궁과 인사동간의 연관성 규칙이다. 그리고 <표 4>의 비즈니스 관광객 집단과 <표 2>의 휴가·레저 관광객 집단의 연관성 규칙을 대조하였을 때, <표 4>의 규칙 1, 2, 9, 10, 13, 14, 18, 19에서 보듯이 인사동과 고궁, 인사동과 이태원, 이태원과 박물관, 신촌/홍대와 이태원과 같은 연관성 규칙은 비즈니스 관광객 집단에서만 차별적으로 형성되었다. 또한 휴가·레저 관광객 집단의 최상위 지지도 연관성 규칙 결과 중에는 한옥 마을은 등장하지 않았으나, 한옥 마을과 고궁, 한옥 마을과 인사동 간의 연관성 규칙은 비즈니스 관광객 집단에서만 등장하였다.

또한 비즈니스 관광객 집단의 지지도 상위 연관성 규칙에서 휴가·레저 관광객 결과에서 등장한 롯데월드->남산 N 서울 타워와 롯데월드 -> 동대문 시장이 드러나지 않았고, 이는 시간 제약성 및 지리적 근접성 관점에서 이해할 수 있을 것이다. 일반적으로 시간 제약 조건에 구애받지 않는 관광객 집단은 희소하고(Hinch and Jackson, 2000), 특히 비즈니스 관광객 집단은 시간 제약적 조건에 민감한 경향성이 있다(Gelhausen, 2011). 이를테면 2013년 외래 실태 조사 보고서에 의하면, 휴가·레저 관광객 집단 중 1-2일 체류 관광객은 3.2 %인 반면, 비즈니스 관광객은 8.6% 이다. 그리고 3-4일 체류하는 휴가·레저 관광객은 36.1%, 비즈니스 관광객집단은 34.3 %으로 유사하게 나타났으나, 5-6일 휴가·레저 관광객 집단은 44.5%이고 가장 높은 비율을 차지하였으나 비즈니스 집단의 경우 20.8%를 차지하고 있다. 한편 비즈니스 관광 주목적 관광객 집단에게 최우선적인 업무는 비즈니스 과업 수행이며, 비즈니스 관광객은 해당 국가의 체류일수를 근무 시간의 연장성으로 간주한다는 점이 제시되었다(Beaverstock et al. 2009). 그리고 Gelhausen(2011)는 비즈니스 관광객들은 시간 절약을 위해 공항 접근성을 좋은 구매 대안을 선호함을 설명하였고, Balke and Wagner(2003)는 시간에 민감한 비즈니스 관광객의 특성을 반영하여 비즈니스 관광객을 위한 관광지 정보 웹 서비스를 개발하였다.

<표 4> 비즈니스 관광객 연관성 규칙

	규칙	신뢰도	지지도	향상도	방문자수
규칙 1	인사동 -> 고궁	65.44	10.89	2.74	409
규칙 2	고궁 -> 인사동	45.65	10.89	2.74	409
규칙 3	남대문시장->동대문시장	61.84	10.57	3.02	397
규칙 4	동대문시장->남대문시장	51.03	10.57	3.02	397
규칙 5	박물관 -> 고궁	71.99	10.33	3.02	388
규칙 6	고궁-> 박물관	43.30	10.33	3.02	388

외래 관광객의 관광 주목적별 관광지 선택 연관성 규칙 분석

규칙 7	남산/ N타워 ->명동	58.63	9.13	2.35	343
규칙 8	명동 -> 남산/ N타워	36.61	9.13	2.35	343
규칙 9	인사동 -> 이태원	46.72	7.77	2.54	292
규칙 10	이태원 -> 인사동	42.20	7.77	2.54	292
규칙 11	남산 N서울타워 -> 박물관	47.50	6.81	3.05	256
규칙 12	박물관 -> 남산 N 서울타워	43.76	6.81	3.05	256
규칙 13	박물관 -> 이태원	44.34	6.36	2.41	239
규칙 14	이태원 -> 박물관	34.54	6.36	2.41	239
규칙 15	한옥마을 -> 고궁	23.85	4.13	3.04	155
규칙 16	고궁 -> 한옥마을	17.30	4.13	3.04	155
규칙 17	신촌/홍대 -> 이태원	58.85	3.81	3.19	143
규칙 18	이태원 -> 신촌/홍대	20.66	3.81	3.22	143
규칙 19	한옥마을 -> 인사동	57.48	2.02	3.51	123
규칙 20	인사동 -> 한옥마을	19.68	2.02	3.51	123

V. 결 론

본 연구는 2013년 외래 관광객 실태 조사의 응답 자료를 활용하여 휴가·레저 및 비즈니스 관광객을 대상으로 하여 관광객 집단별로 방문 관광지의 연관성 규칙 분석을 실시하여 각 집단별 대조적인 특징을 검토하였다. 즉, 2013년 휴가·레저 관광객의 관광지간 연관성 규칙 중 지지도 비율이 상위인 분석 결과에 따르면 남산 N 서울타워가 다른 관광지에 비해 빈번하게 나타났다. 즉, 남산 N 서울타워->고궁, 남산 N 서울타워->남대문 시장, 남산 N서울타워->박물관, 남산 N 서울타워->롯데월드와 같은 관광지간의 연관성 규칙들이 그 예시가 된다. 즉, 2013년 휴가·레저 관광객 중 고궁, 남대문 시장, 박물관, 롯데월드를 방문한 관광객들은 남산 N 서울 타워를 방문하는 경향성이 있다는 점이 본 연구를 통해서 새롭게 발견된 사실 중 하나이다. 이러한 연관성 규칙을 반영하여 휴가·레저 관광객을 타겟팅하여 홍보 마케팅 전략을 수립할 때 관광 여행사는 다른 관광지보다 남산 N 서울 타워에 근거지를 두고 관광 상품에 대한 광고 캠페인을 실행한다면 정보 확산 및 파급력을 증가시킬 수 있을 것이다.

비즈니스 관광객 집단의 분석 결과에 따르면 인사동->고궁, 인사동->이태원, 이태원->박물관, 신촌/홍대->이태원과 같은 연관성 규칙은 휴가·레저 관광객 집단에는 등장하지 않은 대조적인 연관성 규칙으로서 비즈니스 관광객의 관광 이동 흐름의 다른 특징을 발견할 수 있었다. 이를 토대로 비즈니스 관광객에게 연계 관광 코스를 제안할 때 인사동과 함께 고궁 및 이태원 등의 관광지의 매력을 부각시켜 관광 상품 정보를 전달한다면 그 호소력이 높일 수 있을 것이다. 한편, 남대문 시장과 동대문 시장

의 연관성 규칙은 휴가·레저 관광객 및 비즈니스 관광객 집단 모두 공통적으로 지지도 비중이 높았다. 그리고 고궁과 박물관의 연관성 규칙도 두 집단에서 공통적으로 등장한 규칙이나, 비즈니스 관광객 집단에서 상대적으로 지지도 비중의 순위가 더 높게 나타났다.

본 연구의 이론적 시사점은 다음과 같다. 본 연구는 관광객 이동에 관한 빅 데이터를 분석하여 외래 관광객 이동 특성에 대한 가치 있는 정보 추출을 시도하였다. 국가 정보화 전략 위원회에 따르면, 빅 데이터란 대용량 데이터를 활용하고 분석하여, 가치 있는 정보를 추출하며, 생성된 지식을 바탕으로 능동적으로 대응하거나 변화를 예측하기 위한 정보화 기술로 정의하고 있다(이상윤·윤홍주, 2012). 다시 말해 본 연구는 만 명 이상의 외래 관광객들이 전국 대표 관광지 중에서 방문한 장소에 대한 자료를 재분석하여 연구 결과의 타당성을 증진시켰다. 또한, 경영 과학에 그 이론적 기반을 두고 있는 소매업의 공동 구매 경향성 분석 기법인 연관성 규칙 마이닝을 2013년 외래 관광객의 관광지 선택 및 이동 주제에 도입한 점에서 이론적 의의가 있다. 데이터 마이닝 방법론은 빅데이터 분석과 관련하여 다양한 분야에 적용할 수 있는 방법론으로서(이만재, 2012) 본 연구는 관광객 실제 이동에 의해 형성되는 연관성 규칙 마이닝을 적용하여 관광지 관계성을 분석한 점에서 의의가 있다. 특히, 관광 주목적에 따른 관광객 집단별로 차별적으로 형성되는 관광객 이동 흐름을 대조함으로써 관광객 이동에 관한 세분화 연구를 확장시킨 점에서 이론적 의의가 있다.

본 연구 결과의 실무적 시사점은 다음과 같이 제시된다. 첫째, 관광 분야의 실무자는 관광 목적별 연관성 규칙을 고려하여 관광 패키지 상품 구성 및 관광 상품 교차 판매를 실시할 수 있다는 점이다. 오늘날 불특정 다수를 겨냥한 매스 마케팅보다 세분화된 소집단을 겨냥한 세그먼트 마케팅이 효과적인 마케팅 기법으로 사용되고 있다(김중호, 2012). 이를테면 휴가·레저 관광 주목적 집단을 타겟팅하여 남산 N 서울 타워를 중심으로 하여 고궁, 남대문 시장, 박물관, 롯데월드와의 연관성 규칙을 반영한 관광 상품 교차 판매 전략을 실시할 수 있을 것이다. 또는 비즈니스 관광객 연관성 규칙 분석 결과를 토대로 비즈니스 관광객에게 관광 정보를 무작위적 제공할 것이 아니라 <표 4>의 연관성 규칙에 등장한 관광지를 대상으로 특화된 비즈니스 연계관광 프로그램을 개발하고, 휴가·레저 관광객과 구별되는 관광 상품 홍보할 때 관광 경영 효율성이 증진될 수 있을 것이다. 둘째, 본 연구가 제공하는 관광객 이동 정보를 반영하여 관광 정책 실무자는 허비 앤 스포크(Hub and Spoke) 관광 개발을 도모할 수 있다는 점이다. 네트워크형 관광 개발은 허브로 간주되는 거점 관광지와 소규모의 신규 관광지와의 접근성을 향상시켜, 시설 투자 규모가 제한적인 신규 관광지의 한계점을 극복할 수 있다는 장점이 있다. 즉, 본 연구 결과를 고려하여 연관성 규칙의 지지도의 비율이 낮은 신규 관광지들은 지지도의 비율이 높은 거점 관광지와의 연결을 강화시

키는 활동을 통해서 해당 지역으로의 관광객 유인력을 증대시킬 수 있을 것이다(김남조, 2004). 다시 말해 관광지 연계를 전제하고 있는 허브 앤 스포크 관광 개발이 선호되는 현 시점에서(박경옥 외 2인, 2014) 본 연구가 제시하는 관광지 연관성 규칙은 관광지 관계성을 고려한 관광 개발 정책 수립 시에 유용한 정보로 활용될 수 있을 것이다. 셋째, 본 연구 결과를 고려해서 관광 교통과 관련된 실무자들은 관광 시티 투어 버스와 같은 관광 교통 시스템 기반 구조의 최적화를 도모할 수 있다는 것이다. 즉, 본 연구에서 등장한 연관성 규칙을 토대로 관광지를 중심으로 시티 투어 버스의 관광지 코스를 재설계한다면 사용자 이용률이 증가할 가능성이 크다. 예를 들면 본 연구 결과에서 2013년 휴가·레저 관광객과 비즈니스 관광객 두 집단 모두 신촌/홍대 지역을 방문하는 경향성이 발견되었다. 이 점을 고려하여 서울 시티 투어 버스의 이동 코스를 확대한다면 현재 포함되어 있지 않는 코스인 신촌/홍대 지역을 추가할 수 있을 것이다. 연관성 규칙 마이닝을 통해서 서울 시민 통행량을 분석하여 서울 심야 버스 운행노선 및 배차 주기를 변경하여 만족도를 증가시킨 사례가 이를 지지해준다(마경근, 2015).

본 연구는 관광 주목적에 따른 방문한 관광지간의 연관성 규칙을 도출하고, 방한 외래 관광객의 이동 행태와 관련된 시사점을 제시하였으나, 다음과 같은 연구 한계점이 있다. 본 연구는 관광 이동에 대한 탐색적 연구로서, 2013년의 관광객 이동 자료라는 시간적 범위의 제한성 측면에서 연구의 가장 큰 한계점이 있다. 따라서 2013년 이후의 자료를 활용하여 수년에 걸쳐서 외래 관광객의 관광지 선택 행동 변화를 대조 분석하는 연구를 실시할 수 있을 것이다. 외래 관광객의 이동 변화에 관한 심화 연구를 통해서 관광객의 관광지 선택 행동의 동적 특성을 확인하여 연구 결과의 신뢰성 및 실무적 활용성을 증진시킬 수 있을 것이다. 또한 본 연구는 단체 여행객과 개별 자유 여행객을 구분하지 않았으나 관광객의 이동의 정교한 시장 세분화를 위해서, 단체 여행객을 제외한 개별 자유 여행객을 대상으로 한 연구를 실시할 수 있을 것이다. 그 밖에 본 연구는 한국 관광 지식 정보시스템에 적재되어 있는 2013년의 공공 부문의 정형화된 데이터를 활용하였으나, 향후 연구에서 민간 부분의 데이터를 융합하거나 웹로그 기록, 소셜 네트워크 정보, 텍스트 정보, 사물 통신 데이터와 같은 비정형화된 데이터를 포괄하여 데이터 마이닝 연구 결과의 신뢰성을 증진시킬 수 있을 것이다.

참고문헌

- 김기범·안병구(2013). “공간 데이터 마이닝 분석을 통한 데이터의 효과적인 활용”. 『한국인터넷방송통신학회 논문지』. 13(3). pp.157-163.

- 김남조(2004). “관광거점의 개발방향과 과제”. 『관광연구논총』 . 16(1). pp.37-56.
- 김민정(2001). “데이터 마이닝 기법을 이용한 여행사 고객 관계 관리”. 박사학위논문.
이화여자대학교 과학기술대학원.
- 김민형·추승우·황영현(2011). “방한 중국인 관광객의 다목적지 여행 영향요인에
관한 연구: 수도권과 기타 지역의 방문을 기준으로”. 『관광학연구』 . 35(9).
pp.83-296.
- 김영훈·안성식(2006). “데이터 마이닝을 이용한 패스트푸드 레스토랑 이용고객
특성에 관한 연구”. 『관광·레저연구』 . 18(2). pp.191-209.
- 김완섭·강미라·류시영(2010). “데이터 마이닝을 이용한 신혼여행상품의 시장
세분화”. 『관광연구저널』 . 24(3). pp.173-187.
- 김용(2012). “Apriori 알고리즘 기반의 개인화 정보 추천시스템 설계 및 구현에 관한
연구”. 『한국비블리아학회지』 . 23(4). pp.283-308.
- 김종호(2012). “고객 지향 세분시장 획득을 위한 데이터 마이닝 기법
적용방안”. 『한국디지털콘텐츠학회논문지』 . 13(3). pp.385-392.
- 김창현·김우곤·최용호(2003). “데이터 마이닝 기법을 이용한 호텔산업의
데이터베이스 마케팅에 관한 연구”. 『호텔관광연구』 . 11. pp.33-53.
- 남기성·김희재·오재환(2002). “데이터 마이닝의 연관성 규칙을 이용한 동시행동
분석”. 『사회조사연구』 . 17. pp.137-156.
- 노미정·김진화·이재범(2010). “연관 관계 분석을 통한 스마트폰과 기능의 융합
분석”. 『한국경영정보학회 학술대회』 . pp.254-259.
- 마경근(2015). “지역 속으로: 자치단체 정보화사업; 서울시 정보화 우수사례: 정보화
기획단 빅데이터를 활용한 과학적 행정 구현”. 『지역정보화』 . 91. pp.14-19.
- 문화체육관광부(2014). 『2013 외래관광객 실태조사 보고서』 .
- 박경옥·김인신·류지호(2014). “Hub-and-spoke 관광개발을 위한 Networks
분석”. 『관광연구』 . 29(3). pp.1-20.
- 박기용(2006). “의사결정나무모형을 이용한 레스토랑 프랜차이즈 가맹자의
선택요인에 관한 연구”. 『호텔경영학연구』 . 15(4). pp.105-117.
- 변우희·이정열·변성희(2012). “지역 관광 수요에 대응한 관광 자원 정보 서비스
융합 모형 연구: 데이터 마이닝과 상황평가방법론 적용”. 『관광학연구』 . 36(5).
pp.221-241.
- 양승정·이종태(2007). “퍼지의사결정나무 개선방법을 이용한 CRM 적용 사례”.
『한국콘텐츠학회논문지』 . 7(8). pp.13-20.

- 이만재(2012). “빅 데이터와 공공 데이터 활용”. 『정보과학회지』. 2(2). pp.47-64.
- 이상윤·윤홍주(2012). “공공 데이터를 활용한 국가정보화 전략 연구”. 『한국전자통신학회논문지』. 7(6). pp.1259-1273.
- 이혜진·송운강(2010). “다목적지 여행의 결정요인 분석”. 『관광학 연구』. 34(4). pp.127-142.
- 임제순·이경준·조영석(2010). “발생빈도를 고려한 연관성분석 연구”. 『한국데이터정보과학회지』. 21(6). pp.1061-1069.
- 정유경(2003). “연관 규칙을 이용한 호텔 레스토랑 고객의 이용 행태 분석”. 『호텔경영학연구』. 12(2). pp.1-22.
- 채경희·김상철(2010). “의사결정나무 기법을 활용한 백화점의 고객세분화 사례연구”. 『유통과학연구』. 8(1). pp.13-19.
- 홍순만(2013). 『소셜 네트워크의 세계와 빅데이터 활용 사람 사이를 보는 힘』. 서울: 파워북.
- Agrawal, R., Imieliński, T. and A, Swami(1993). Mining association rules between sets of items in large databases. *ACM SIGMOD*. 22(2). pp.207-216.
- Balke, W. T. and M, Wagner(2003). *Towards personalized selection of web services*. In Proceedings of the 12th International World Wide Web Conference (WWW 2003). Budapest, Hungary, May 2003.
- Beaverstock, J. V., Derudder, B., Faulconbridge, J. R. and F, Witlox(2009). International business travel: some explorations. *Human Geography*. 91(3). pp.193-202.
- Fayyad, U., Piatetsky-Shapiro, G. and P, Smyth(1996). From data mining to knowledge discovery in databases. *AI Magazine*. 17(3). pp.37-54.
- Gelhausen, M. C(2011). Modelling the effects of capacity constraints on air travellers' airport choice. *Journal of Air Transport Management*. 17(2). pp.116-119.
- Hinch, T. D. and E. L. Jackson(2000). Leisure constraints research: Its value as a framework for understanding tourism seasonability. *Current Issues in Tourism*. 3(2). pp.87-106.
- Law, R., Rong, J., Vu, H. Q., Li, G. and H, A. Lee(2011). Identifying changes and trends in Hong Kong outbound tourism. *Tourism Management*. 32(5). pp.1106-1114.
- Lau, and McKercher(2006). Understanding tourist movement patterns in a destination: A GIS approach. *Tourism and Hospitality Research*. 7(1). pp.39-49.

- Ngai, E. W., Xiu, L. and D, Chau(2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications*. 36(2). pp.2592-2602.
- Park, H. C(2010). Standardization for basic association measures in association rule mining. *Journal of the Korean Data and Information Science Society*. 21(5). pp.891-899.
- Rong, J., Vu, H. Q., Law, R. and G, Li(2012). A behavioral analysis of web sharers and browsers in Hong Kong using targeted association rule mining. *Tourism Management*. 33(4). 731-740.
- Rygielski, C., Wang, J. C., and D. Yen(2002). Data mining techniques for customer relationship management. *Technology in Society*. 24(4). pp.483-502.
- Versichele, M., de, Groote. L., Claeys, Bouuaert. M., Neutens, T., Moerman, I. and N, Van. de. Weghe(2014). Pattern mining in tourist attraction visits through association rule learning on Bluetooth tracking data: A case study of Ghent, Belgium. *Tourism Management*. 44(1). pp.67-81.

2016년 7월 19일 원고접수. 2016년 8월 25일 1차수정본 접수.

2016년 10월 17일 최종수정본 접수.

3인 익명심사 畢