

한국관광 데이터 분석을 통한 주요 영향 요인에 관한 연구

양소희¹, 나현지², 정홍주³
¹동덕여자대학교 세무회계학과 학부생
²동덕여자대학교 컴퓨터학과 학부생
³온스페이스 컨설팅사업부 이사

skg703@naver.com, again_july@naver.com, HongjuJung@hotmail.com

A Study on Key influencers through Analysis of Korea Tourism Data

So-Hee Yang¹, Hyeon-Ji Na², Hong-Ju Jung³
¹Dept. of Tax Accounting, Dongduk Women's University
²Dept. of Computer Engineering, Dongduk Women's University
³Dept. of Consulting Service, ONSPACE

요 약

COVID-19가 종식되어 관광산업이 활발해짐에 따라 외래관광객 데이터를 분석하여 활용할 필요가 있다. 한국관광 데이터랩의 외래관광객 조사 보고서는 설문문을 통해 수집한 데이터에 대해 각 요인별 통계를 제공하고 있지만 복잡적 요인으로 분석하기 어렵다. 본 연구는 2017년부터 2022년까지 한국관광 데이터랩의 외래관광객 조사 데이터를 Microsoft Power BI를 통해 분석하여 여성 20대, 서구 관광객의 방문 활동과 방문 고려 요인 등 복잡적 주요 영향 요인을 도출하고 그에 따른 발전 방향을 결론에서 제시한다.

1. 서론

엔데믹에 따라 관광산업이 활발해지고 있으며, 그에 따라 외래관광객 데이터를 분석하여 활용할 필요가 있다. 한국관광 데이터랩의 외래관광객 조사 보고서는[1] 설문문을 통해 수집한 데이터에 대해 각 요인별 통계를 제공하고 있으나, 여러 요인을 결합하여 복합적인 인사이트를 도출하는 데 한계점을 가지고 있다. 본 연구는 2017년부터 2022년까지의 한국관광 데이터를 수집하고 분석하여 복잡적 주요 영향요인을 도출하고, 발전방향을 제시하고자 한다.

2. 관련 연구

한국관광공사에서는 관광 빅데이터 분석 가이드라인을 수립하는 연구가 진행되었다[2]. 한국관광 데이터랩의 2022년 외래관광객 조사 보고서는[1] 2016년부터 입국자 추이 등을 제공하고 있다. 부산관광공사의 2022 부산 방문관광객 실태 조사보고서에서는 각 설문 문항 별 기초통계 분석 결과를 제공하고 있다[3]. 관광 데이터와 공공데이터를 결합하여 융합 분석을 통한 주요 원인을 파악하는 연구 모형을 제시하고 있다[4]. COVID-19로 인하여 관광소비 지출액은

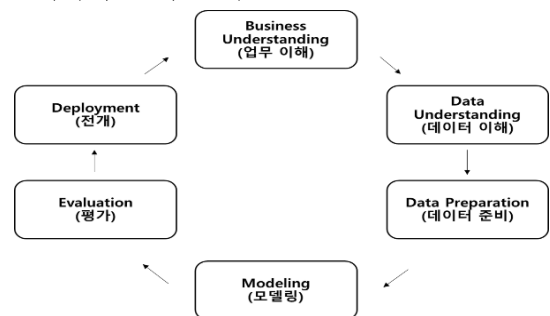
전년 동월 대비 약 30% 감소 폭을 보였으며[5], 그 경제적 구조적 원인과 특징을 살펴보고 회복 과정을 설명하는 연구가 진행되었다[6].

3. 한국관광 데이터 분석 방법

3.1 한국관광 데이터

외래 방문객의 주요 영향 요인 도출을 위한 데이터는 2017년부터 2022년까지 한국관광 데이터 랩의 외래관광객 조사 원자료를 이용한다[1]. 데이터는 연도별 질문 항목이 상이하며 응답의 세부 값도 상이하나 공통적인 항목은 입국 년월, 국가, 성별, 나이, 방문 활동, 방문 고려 요인, 체류 기간과 비용 등이다.

3.2 데이터 분석 절차



(그림 1) CRISP-DM

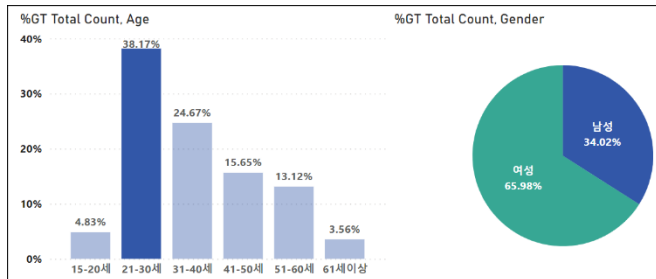
본 연구에서는 CRISP-DM을 기반으로 데이터 분석을 진행한다[7]. 업무/데이터 이해 단계에서는 스토리텔링을 위한 필요 데이터 항목을 도출한다. 데이터 준비와 모델링 단계에서는 연도별 한국관광 데이터에서 필요 컬럼을 정의하여 데이터를 변환하고, 차원 테이블을 생성한다. Microsoft Power BI에서 제공하는 DAX(Data Analysis Expressions)를 통해 측정값을 생성하며 시각화에서 컨텍스트에 해당하는 응답자 수를 제공한다[8]. 평가 단계에서는 보고서 시각화를 통해 슬라이서와 필터를 이용하여 주요 영향 요인을 파악하고 인사이트를 도출한다.

4. 한국관광 주요 영향 요인 도출

Microsoft Power BI를 통해 한국관광 데이터랩의 외래관광객 조사 데이터에서 [1] 코로나 중(2020년-2021년 10월)시기의 분석은 제외하며 서구는 북미, 유럽, 호주, 러시아로 한정한다.

4.1 방한객

코로나 전(2017-2019년)과 위드 코로나(2021년 11월-2022년) 시기 전체적으로 중국과 일본 관광객이 27.75%로 많으며, 여성 비율이 52.98%이며, 연령대는 20대(21-30)의 경우 38.17%로 높게 나타난다. 20대의 경우 여성 비율은 62.37%로 높으며, 특히 일본의 경우 20대 여성 비율은 75.75%로 아주 높게 나타난다.



(그림 2) 여성 20대

4.2 여행 준비

관광객들의 주요 방문 고려 요인(중복응답)은 음식/미식 탐방, 쇼핑 순이며, 주요 방문 목적은 여가, 위락, 휴가로 53.53%로 높게 나타난다. 여성 20대의 경우 주요 방문 고려 요인과 방문 목적은 쇼핑과 여가, 위락 휴가로 나타난다. 중국과 일본 관광객의 경우 쇼핑, 음식/미식 탐방 순으로 나타난 반면, 서구의 경우는 음식/미식 탐방, 역사/문화유적/전통문화체험 순으로 차이를 보인다. 'K-POP/한류스타 관련 팬미팅/촬영지 등 경험'은 일본이 가장 높은 모습을 보인다.

4.3 방한 활동

관광객들의 주요 방문지역은 서울이 가장 많은 것으로 나타난다. 여성 20대 중국과 일본 관광객의 경우 주요 방한 활동(중복응답)은 쇼핑, 식도락 관광, 자연경관 감상 순으로 나타난 반면, 서구의 경우는 쇼핑, 고궁/

역사 유적지 방문, 자연경관 감상 순으로 높게 나타난다.

4.4 여행 형태

여행의 형태는 개별 여행이 85.57%로 많이 나타나며, 체류일은 4-7일로 53.67%로 가장 높다. 서구의 경우 위드 코로나(2021년 11월-2022년)시기에는 체류일이 11-30일이 가장 많이 나타난다. 여성 20대의 경우 개별 여행을 선호하며, 4-7일 체류하는 것을 알 수 있다.

5. 결론

20대 여성들을 타깃으로 K-Culture(K-Pop 댄스 축제 등)를 체험하는 프로그램과 안전을 제공한다면 재방문 비율을 높일 수 있다. 서구 방문객의 방문 고려 요인에 역사/문화유적/전통문화체험이 높은 순위에 있으므로 한국 전통문화를 체험하는 활동을 제공해 이들의 니즈를 충족시켜야 한다. 관광객들의 방문지역이 대부분 서울이므로 지방 방문을 유도하기 위해 지방 특색을 살린 축제나 특산물을 활용한 음식 문화[9]를 활용하여 해당 지역에서만 경험할 수 있는 활동을 제공하는 등 맞춤형 전략이 필요하다. 다만 한국관광 데이터랩의 외래 방문객 조사 데이터의 [1] 항목이 매년 변경되어 연속성을 확보하기 어렵다는 한계점이 있으며 외래 관광객 조사 데이터 외의 다른 카드, 통신 데이터 등을 결합하여 다각도로 분석해야 한다. 또한 데이터 분석 시 여행 제한 등 국가 및 정치적 환경도 고려해야 한다.

참고문헌

[1] 한국관광 데이터랩, <https://datalab.visitkorea.or.kr>
 [2] 한국관광공사, 관광 빅데이터 분석 가이드라인 수립 연구:이동통신 및 신용카드데이터를 중심으로, 2020.
 [3] 부산관광공사, <https://bto.or.kr>
 [4] 한국관광공사, 관광 빅데이터 및 공공데이터 간 융합분석을 통한 내/외국인 방문수요의 주요 원인은?, 데이터엔투어리즘, 21호, 2023.
 [5] 김종욱, 지역별 방문자 수 및 카드 소비 지출 변화, 노동리뷰, 113-116, 2021.
 [6] 홍태희, 코로나 19 팬데믹 경제위기의 특성과 원인 그리고 전망, 경상논총, 38(3), 79-99, 2020.
 [7] 한국데이터산업진흥원, 데이터 분석 전문가 가이드 (ADP)(ADsP), 서울, 한국데이터산업진흥원, 2021.
 [8] Microsoft PowerBI, <https://powerbi.microsoft.com/ko-kr/>
 [9] 김정훈, 장미향, 김대철, 음식 관여도 및 신기성 추구가 음식관광행동과 전주비빔밥축제 재방문의도에 미치는 영향, 관광연구저널, 30(6), 71-84, 2016.

본 프로젝트는 과학기술정보통신부 정보통신창의인재양성사업의 지원을 통해 수행한 ICT 멘토링 프로젝트 결과물입니다.