

베스트셀러 순위가 공공도서관 대출에 미치는 영향 분석: 패널자료 분석을 중심으로

The Effects of the Bestseller Ranks on Public Library Circulation: Based on Panel Data Analysis

이종욱(Jongwook Lee)*
강우진(Woojin Kang)**
박종규(Jungkyu Park)***

초 록

본 연구에서는 베스트셀러 목록에 포함된 도서의 순위가 공공도서관에서의 평균 대출 건수에 미치는 영향을 패널분석을 통해 살펴보고자 하였다. 본 연구를 위해 문화 빅데이터 플랫폼을 통하여 국립중앙도서관이 제공하는 데이터를 바탕으로 2018년 1월 1일부터 2019년 12월 29일까지 총 104주 동안의 분석 대상 도서 179권의 공공도서관 대출 데이터 세트를 생성하였고, YES24 웹사이트를 통해 같은 기간 주간 베스트셀러 목록 데이터 세트를 구축하였다. 공공도서관 대출과 베스트셀러 도서 순위 간 정확한 관계를 확인하기 위해 패널자료의 특성을 활용한 분석 방식인 선형회귀모형, 고정효과모형, 확률효과모형 등 세 개의 모형을 비교한 결과, 고정효과모형이 가장 적합한 것으로 나타났다. 순위 데이터 결측값이 47주 미만인 179권의 도서의 자료를 고정효과모형으로 분석한 결과, 도서의 베스트셀러 순위가 한 단계 내려가면 공공도서관에서의 해당 도서 평균 대출 건수가 0.108권 유의미한 수준에서 감소한다는 것을 밝혀내었다. 또한, 베스트셀러 순위가 도서 평균 대출 건수에 미치는 효과가 도서의 내용분류에 따라 상이함을 알 수 있었다. 이 연구는 베스트셀러 순위가 사람들의 도서관 대출행태에 영향을 미치고 있음을 실증적으로 확인한 것으로, 공공도서관에서는 이용자의 요구를 예측하고, 장서 개발 정책 수립에 베스트셀러 목록을 비롯한 사회문화적 맥락을 고려할 필요가 있음을 시사한다.

ABSTRACT

The purpose of this study is to analyze the effects of the bestseller ranks on the book circulations in public libraries. To achieve this goal, the weekly data sets of 179 books' library circulation and bestseller list from January 1, 2018 to December 29, 2019 were constructed based on the data collected from BigData MarketC and YES24. Three methods for analyzing panel data including linear regression, fixed-effect, and random effect models were compared, and it turned out that fixed-effect model was better than other methods. The results show that the average ranks of bestsellers were associated with their public library circulations visually. Also, the analysis of fixed-effect model showed that the single rank decline of a book on the bestseller list decreases its average circulation of 0.108 while the size of effect varied depending on subject of books. The study empirically demonstrated the impact of a bestseller list on people's book circulation behavior, suggesting that public libraries need to reference sociocultural context as well as bestseller book lists to predict library user needs and to formulate collection development policy.

키워드: 공공도서관, 대출 데이터, 대출행태, 베스트셀러
public library, circulation data, circulation behavior, bestseller

-
- * 경북대학교 문헌정보학과 부교수(jongwook@knu.ac.kr) (제1저자)
** 경북대학교 일반대학원 문헌정보학과 석사과정(rkddnwl262@gmail.com) (공동저자)
*** 경북대학교 심리학과 조교수(jkp@knu.ac.kr) (교신저자)
■ 논문접수일자: 2021년 10월 25일 ■ 최초심사일자: 2021년 12월 3일 ■ 게재확정일자: 2021년 12월 18일
■ 정보관리학회지, 38(4), 1-23, 2021. <http://dx.doi.org/10.3743/KOSIM.2021.38.4.001>

※ Copyright © 2021 Korean Society for Information Management
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

1. 서론

베스트셀러(bestseller)에 대한 정의는 사전적 의미, 판매 부수, 내용, 서적 시장 등의 관점에서 다양한 방식(이영희, 1998)으로 내려질 수 있으나 일반적으로 출판계에서는 이를 가장 많이 팔리는 도서로 정의한다. 베스트셀러라는 용어는 미국의 월간 문예지인 “북맨(The Bookman)”의 창간호에서 처음 사용된 것으로 알려져 있으며, 그 당시에는 ‘best selling books’로 불리다가 차츰 ‘베스트셀러’로 표현되기 시작했다(안선희, 2009). 베스트셀러와 관련된 용어로는 스테디셀러(steady seller), 롱셀러(long seller) 등이 있는데, 엄밀하게는 이러한 용어 간의 의미상의 차이가 존재한다. 예를 들어, 스테디셀러나 롱셀러는 일정 기간에 많이 팔린다는 의미를 지닌 베스트셀러와는 달리 오랜 기간 꾸준히 잘 팔리는 도서를 의미한다(최지영, 2006). 이 밖에도 밀리언셀러(million seller), 빅셀러(big seller) 등의 용어가 있으나 본 연구에서는 이러한 용어를 포괄하는 개념으로 대형서점이 판매량에 근거하여 정기적으로 발표하는 베스트셀러 목록에 포함된 도서를 베스트셀러로 정의하였다.

이러한 베스트셀러를 만드는 요인은 크게 출판사의 내적 요인(출판기획 단계의 노력, 출판 이후의 마케팅 요인 등)과 외적 요인(정치, 경제, 사회문화적 요인, 신문 및 출판 보도 요인 등)으로 구분하기도 하며(허연, 2006), 좀더 세분화하여 작품적 요인, 시대상황적 요인, 출판산업 구조적 요인, 출판기획적 요인 등으로 구분(안선희, 2009)하기도 한다. 이처럼 도서가 베스트셀러에 오르는데 영향을 미

치는 요인은 매우 다양하며, 복합적으로 작용하고 있는데 특히, 광고나 마케팅 요인이 베스트셀러에 많은 영향을 미칠 수 있음을 토대로 베스트셀러가 출판계에 부정적인 영향을 끼친다는 견해도 존재한다. 그렇지만, 이러한 부정적 측면에도 불구하고 베스트셀러는 출판사의 내적 및 외적 요인들이 어우러진 일종의 ‘사회문화 현상 또는 산물’(정옥년, 2005; 허연, 2006)로 간주되고 있으며, 사람들의 도서 선택에 유용한 정보원으로 참고되고 있다(김기태, 1997).

도서관법 제2조에 의하면 공공도서관은 공중의 정보이용, 독서활동, 문화활동 및 평생교육을 증진하기 위하여 설치된 도서관으로, 2020년 기준으로 1,172관이 설치되어 운영되고 있으며, 1관당 대출도서 수가 약 10만권(문화체육관광부, 2021)으로 우리나라 국민의 정보, 교육, 문화, 오락 등의 다양한 요구를 충족시키고 있다. 이처럼 공공도서관은 이용자 요구를 파악하고 예측하여 이를 해결하기 위해 정보자원을 비롯한 각종 프로그램을 제공하는 역할이 중요한 부분을 차지하고 있는데, 이러한 요구를 파악하기 위해 주로 이용자 대상 설문조사(이희수, 김기영, 2014)나 면담조사(이정연, 김현애, 2021)를 실시하지만, 최근에는 소셜 미디어 데이터(박태연, 오효정, 2021)나 대출 데이터(심지영, 2021)를 활용한 사례도 존재한다. 이 가운데 도서관 대출 데이터는 이용자의 요구를 간접적으로 제시하는 객관적인 데이터이며, 도서 대출에 영향을 미칠 수 있는 요인을 이해하는 것은 이용자의 정보요구를 예측하여 도서관 장서 개발이나 각종 프로그램 기획에 유용하게 활용할 수 있다.

이에 문헌정보학 분야에서는 이용자들의 공공도서관 대출에 영향을 미칠 수 있는 요인을 식별하는 연구가 꾸준히 이루어지고 있는 상황이다. 이러한 영향 요인들은 크게 개인, 도서관, 지역 차원으로 구분(이종욱, 강우진, 이명, 2021) 할 수 있으며, 연구에 따라 특정 차원 예를 들어, 도서관 요인에 중점(김선애, 2016)을 두거나 지역적 요인을 중심으로 고려(유경중, 박일중, 2009)한 연구들이 있는 반면, 이들 요인들을 종합적으로 다루는 연구(권나현, 송경진, 2014)도 존재한다. 이 밖에 사회문화적 요인도 공공도서관 대출에 영향을 미칠 수 있는데, 예를 들어, 김원중(2015)의 연구에서는 유행성 질병이 공공도서관 대출 책수에 미치는 영향을 조사한 바 있다.

본 연구에서는 사회문화적 맥락을 반영하는 베스트셀러 순위가 국내 공공도서관에서의 베스트셀러 도서에 대한 대출 건수에 미치는 영향을 패널자료 분석 방법을 통해 살펴보고자 한다. 문헌정보학 분야에서 베스트셀러 순위와 공공도서관 대출 간의 관계를 분석한 연구(윤휘영, 2021; 이종엽, 2012)가 수행된 바 있으며, 이들 연구에서 베스트셀러 순위와 공공도서관 대출 간의 관련성이 있음을 밝혀낸 바 있다. 그렇지만 본 연구에서는 주간 베스트셀러 목록에 포함된 여러 주제 분야 도서의 순위와 공공도서관에서의 이들 도서에 대한 주간 대출량 변화를 활용하여 선형회귀모형, 고정효과모형, 확률효과모형을 통한 패널자료 분석을 실시했다는 점에 기존 연구와의 차별성을 확보하고 있다.

2. 선행연구

2.1 베스트셀러 관련 연구

베스트셀러 관련 선행연구는 크게 베스트셀러 목록을 분석하는 연구, 베스트셀러 요인을 분석하는 연구, 베스트셀러의 역할 및 영향력 파악 연구, 베스트셀러에 대한 인식 연구로 구분할 수 있다. 먼저 베스트셀러 목록을 분석한 주요 연구로 이영희(1998)는 1948년부터 1997년까지의 베스트셀러를 분석하였는데, 특히 정치적 격변기에 따라 시대를 구분하여 시기별 출판 상황과 주요 베스트셀러를 살펴보았다. 또한, 베스트셀러 유형을 시대 상황적 유형, 내용적 요인, 기획적 요인, 시장적 요인으로 구분하여 베스트셀러를 분석하였다. 그리고 최지영(2006)은 2005년 1월 첫째 주부터 12월 넷째 주까지의 주간별 베스트셀러 목록을 조사하여 소설과 비소설로 구분하고, 베스트셀러 순위 10위까지의 자료에 대한 점수를 부여하여 스테디셀러를 식별하였다. 이들 스테디셀러를 통해 우리나라 사회동향 즉, 새로운 것에 대한 대중의 반응, 동양학 열풍, 이혼율 상승, 대중매체의 영향, 여성의 처세술이나 인생, 교육열 등을 파악하였다.

다음으로 베스트셀러의 요인을 조사한 연구들은 주로 도서의 유형에 따른 요인을 살펴본 경우가 많았다. 구체적으로 김영조(2006)는 1995년부터 2004년까지의 실용 분야별 교보문고 베스트셀러 10위권 목록을 조사한 후, 이에 대한 출판 내적 및 외적 요인의 영향을 살펴보았다. 출판 내적 요인으로는 기획적, 편집적, 출판산업적 요인이 존재하였으며, 출판 외적 요인으

로는 정치적, 경제적, 사회문화적 요인이 존재하였고, 이들 출판 내외적 요인과 실용서 베스트셀러 간의 관련성이 있음을 밝혀내었다. 다음으로 허연(2006)은 사회 현상을 반영하는 단행본 베스트셀러(2000년 1월부터 2005년 6월까지의 교보문고 종합 베스트셀러 10종)의 특성을 파악하고, 이를 대상으로 베스트셀러 현상 유발요인을 책의 요인(가격, 내용 등), 마케팅적 요인(광고 비용, 이벤트 등), 저널리즘적 요인(매체 노출, 보도 방식 등)으로 구분하여 정리하였다. 또한, 김기태(2015)는 2001년부터 2014년까지의 연도별 베스트셀러 목록에서 국내 문학도서의 추이를 살펴보고, 베스트셀러 요인을 파악하고자 하였다. 그 결과, 종합 베스트셀러 10위권에 속한 문학도서는 모두 소설로 드러났으며, 시간이 지날수록 문학도서에 대한 독자의 관심이 감소하고 각종 미디어와 연계된 도서가 베스트셀러에 오르는 것을 파악할 수 있었다. 한편 안선희(2009)는 특정 도서 유형의 베스트셀러 요인을 분석한 것은 아니지만 IMF 이전과 이후 각 10년간 연간 종합 베스트셀러를 분석하여 IMF 이후 베스트셀러 요인을 시대상황적, 출판산업 구조적, 출판기획적, 작품적 요인으로 구분하여 설명하였으며, 특히 IMF 이후 베스트셀러에 다양성이 저해되고 있음을 지적하였다.

베스트셀러의 역할이나 영향력을 분석하려는 시도로 국내에서는 주로 베스트셀러의 출판계에 대한 영향을 살펴보았다. 먼저 김기태(1997)는 국내 베스트셀러의 변화 양상을 살펴보고, 베스트셀러가 출판계에 미치는 긍정적 및 부정적 영향을 기술하였다. 긍정적인 면으로는 베스트셀러가 독자들에게 유용한 정보 통로로 기능하

고, 양질의 도서 제작을 위한 활력소가 될 수 있으며, 출판계 종사자의 위상을 높이는 데 기여하고, 사람들의 건전한 가치관 정립에 도움을 줄 수 있다는 것이다. 반면 부정적인 면으로는 출판 산업의 상업화가 촉진되고, 출판 유통 구조의 왜곡을 심화시킬 수 있으며, 출판물의 획일화가 이루어질 수 있고, 출판계 종사자들의 심리적 부담이 가중되며, 독자들의 독서가 편향될 수 있다는 것이다.

해외에서는 베스트셀러의 영향력을 파악하기 위한 실질적인 데이터 분석이 이루어진 사례가 있는데, 대표적으로 Sorensen(2007)의 연구에서는 양장본 소설 1,200권의 닐슨 북스캔(Nielsen BookScan) 주간 판매량 데이터를 활용하여 뉴욕 타임즈(New York Times) 베스트셀러 목록이 도서 판매량과 출판 도서 수에 미치는 영향을 분석하였다. 구체적으로 도서의 종류와 시간의 효과를 고정효과 모형으로 통제된 상태에서 뉴욕 타임즈 베스트셀러에 포함 여부와 베스트셀러 순위 등이 도서 판매량에 미치는 영향을 살펴보았다. 이에 더해, 시간 요인과 기타 유명 방송매체 노출의 효과 역시 검정하였다. 결과적으로 베스트셀러 목록에 오르는 것은 평균적으로 도서의 판매량을 완만하게 증가시키며, 해당 효과는 신인 작가의 경우 더욱 큰 것으로 나타났다.

끝으로 베스트셀러에 대한 인식을 조사한 연구로는 베스트셀러에 대한 독자들의 인식을 조사한 사례와 도서관 사서의 인식을 조사한 사례가 있다. 먼저, 김형욱(2004)은 2000~2004년 상반기 동안 20위권에 속한 베스트셀러의 특징을 살펴보고, 베스트셀러에 대한 독자들의 인식 즉, 이용행태와 이용동기, 베스트셀러에 대

한 평가가 출판계에 미친 영향을 조사하였다. 그리고 박옥화(2011)는 공공도서관 사서 78명을 설문조사하여 베스트셀러에 대한 사서의 인식을 살펴보았다. 사서들은 베스트셀러를 비교적 좋은 책 혹은 널리 검증된 양서로 보는 경향이 있었으나, 베스트셀러 목록의 조작 가능성이나 역기능에 대해 인지하고 있는 것으로 나타났다. 즉, 사서들은 대체로 베스트셀러 목록을 수용하면서도 비판적 입장을 견지하고 있었다. 이를 통해 박옥화(2011)는 도서 정보원의 다변화와 도서 유통 전반을 감독하는 공식기구 출현의 필요성을 제안하였다.

2.2 공공도서관 대출 데이터 활용 연구

공공도서관 대출 데이터는 이용자의 도서 이용 패턴을 보여줌과 동시에 정보요구를 반영하는 것으로 이용자의 정보행태에 대한 이해를 높이고, 장서 개발 계획을 수립하고 도서관 서비스 및 프로그램 개발이나 기존 서비스 개선에 활용될 수 있다. 이러한 공공도서관 대출 데이터를 활용한 연구를 살펴보면, 대체로 대출 데이터를 분석하여 도서관 이용 패턴을 분석하거나 대출에 영향을 미칠 수 있는 요인들을 식별하여 도서관 장서개발이나 운영 정책 수립에 참고할 수 있는 시사점을 제시하고 있다.

먼저, 허선과 정연경(2014)은 2012년에 강서구와 양천구 소재 공공도서관 2곳의 이용자 대출 데이터(대출 횟수, 대출자료, 대출시간)와 도서관 내외부 요인들을 포함하여 도서관 이용 분석을 실시하였다. 이용자의 대출 횟수 분석을 통하여 연령대별 이용자의 대출행태를 파악하여, 연령대별 이용자 서비스 개발 및 이용 확대

방안 마련의 가능성을 확인할 수 있었다. 또한, 도서관의 특화장서 서비스가 자료대출에 영향을 미침을 확인하여, 이를 장서개발 정책에 반영할 필요성을 확인할 수 있었다. 이 밖에도 이용자의 선호 이용 시간과 요일을 파악하여 도서관 서비스나 운영에 도움이 되는 기초자료를 확보하였다.

다음으로 김완중(2015)은 사회활동에 많은 영향을 미친 질병 즉, 메르스(MERS) 유행 기간에 이용자의 공공도서관 대출행태에 대해 조사하였다. 공공도서관 빅데이터 수집 및 분석 플랫폼을 통해 수집된 2014년 4월 4일부터 8월 19일까지, 2015년 4월 4일부터 8월 19일까지의 인쇄 단행본의 대출데이터를 각 연간 3개 구간으로 구분하여 질병 유행 기간에 따른 대출 데이터 변화를 파악하고자 하였다. 연구결과, 질병의 유행은 도서관 일평균 대출 권수 감소에 영향을 주었으며, 여름방학은 일평균 도서관 대출 권수 증가에 영향을 준 것을 확인할 수 있었다. 그리고 질병의 유행 기간보다 유행 후에 도서관의 일평균 대출 권수가 감소하여 질병의 유행이 도서관 활동에 장애가 될 수 있음을 알 수 있었다.

2018년에는 국립세종도서관 빅데이터 로그(이용자 정보, 대출 정보, 서비스 이용 정보)를 활용하여 도서관 이용자와 대출 현황을 조사하려는 연구(김태영, 백지연, 오효정, 2018)가 있었다. 해당 연구에서 활용한 대출 데이터는 2014년 1월부터 2017년 10월까지 누적된 국립세종도서관의 대출 이용자 정보와 대출 횟수였으며, 월별, 요일별, 이용자 연령대별 대출 분석이 이루어졌다. 이에 기반하여 저자들은 연령별, 성별에 따른 도서관 운영 정책 세분화 필요성을

제안하였으며, 특히 빅데이터에 기반한 도서관 정책 개발이나 서비스 고도화의 중요성을 제시하였다.

최근에는 심지영(2021)이 서울시 공공도서관 11개관의 대출 데이터를 바탕으로 대출 도서의 속성을 파악하여 도서관별 정보요구를 분석하였다. 구체적으로 이 연구에서는 대출 도서의 이용대상, 이용목적/동기, 관심사/취향, 도서관장르, 주제, 저자 관련 정보요구를 분석하여 유사한 도서 이용 패턴 즉, 정보요구를 갖는 도서관 클러스터를 식별하였다. 또한, 이종욱, 강우진, 이명(2021)의 연구에서는 수도권 지역사회의 사회경제적 박탈 지수를 산출한 후, 이러한 박탈 지수와 공공도서관 대출 책수 간의 관계를 분석하였다. 그 결과, 지역사회의 사회경제적 박탈 수준은 해당 지역의 공공도서관 대출 책수에 부정적인 영향을 미치는 것으로 나타났다.

한편, 문헌정보학 분야에서 공공도서관 대출과 베스트셀러 간의 관계를 조사하려는 시도가 이루어진 바 있다. 구체적으로 이종엽(2012)은 2009년에서 2010년까지 상위 200위 베스트셀러 목록의 주제, 출판년도, 출판사 등을 분석하고, 공공 및 대학도서관에서의 베스트셀러 소장률을 파악하였으며, 나아가 베스트셀러와 공공도서관 베스트대출 목록 간의 상관성을 조사하여 높은 양의 관계가 있음을 확인하였다. 그리고 이호신(2018)의 연구에서는 도서관의 도서 대출에 대해 저작자에게 보상을 제공하는 '공공대출보상권' 도입 주장에 대한 타당성 검토를 위해 도서관의 도서 대출과 서점의 도서 판매 데이터를 비교하였다. 이 과정에 도서관 인기대출도서와 서점 베스트셀러의 목

록을 비교 분석하였는데, 각 목록에서 상위 200위에 속한 도서의 중복 비율이 3년(2014~2016년) 평균 17.7%인 것으로 나타났다. 또한, 두 목록에 속한 도서의 주제 분포와 발행 연도에서도 상당히 다른 양상을 보이는 것으로 드러났다. 이러한 연구 결과를 토대로 이호신(2018)은 도서관에서의 대출과 서점에서의 판매 감소와 연관 짓는 것은 적절하지 않음을 제시한다.

나아가 윤희영(2021)의 연구에서 공공도서관의 과거 대출 데이터와 대형서점 3사의 베스트셀러 목록을 활용하여 공공도서관에서의 효율적인 장서 개발을 위한 이용자들의 요구 즉, 도서 대출빈도를 기계 학습 기반으로 예측하고자 하였다. 구체적으로 2010년부터 2019년까지의 공공도서관 인기대출도서 가운데 문학류 도서의 대출 데이터와 동일 기간 동안의 문학류 베스트셀러 목록을 활용하여 대출 빈도 예측 모델을 구현하였다. 특히, 이 연구에서는 베스트셀러 순위 및 순위의 존재 유무에 따른 공공도서관 대출 빈도 간의 관련성을 분석하였으며, 이를 통해 베스트셀러 순위가 도서관 대출 빈도 예측에 유의미한 변수가 될 수 있음을 언급하였다.

요약하면, 공공도서관 대출 데이터를 분석한 연구는 이용자의 정보행위나 정보요구를 파악하여 도서관 장서 개발이나 서비스 방향을 설정하기 위해 수행되었으며, 이용자의 개인적 특성이나 도서관, 지역, 사회문화적 특성이 도서 대출에 미치는 영향을 살펴보는 시도가 있었다. 공공도서관 대출과 베스트셀러 간의 연관성을 살펴본 연구도 이루어진 바 있으나 윤희영(2021)의 연구를 제외하고는 대부분 영향

관계가 아닌 관련성 분석에 중점을 두었으며, 베스트셀러를 공공도서관 대출에 영향을 미칠 수 있는 사회문화적 특성으로 취급한 경우는 드문 편이었다.

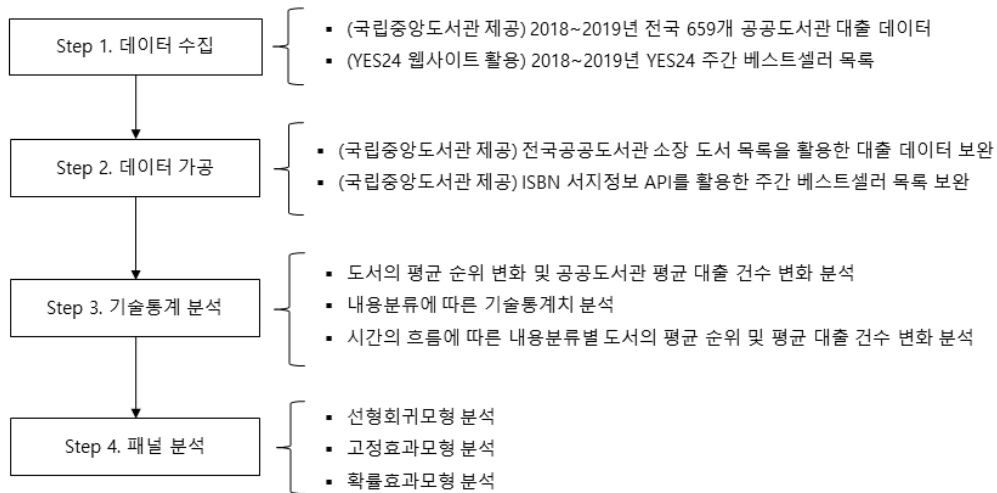
하였다. 이들 데이터를 바탕으로 기술통계 및 패널 분석이 실시되었다. 각 단계에 대한 세부적인 내용은 다음 절부터 기술하였다.

3. 연구방법

연구를 위한 데이터 수집과 분석 과정은 다음의 <그림 1>과 같다. 국립중앙도서관이 제공하는 공공도서관 대출 데이터와 소장 목록을 활용하여 대출 데이터 세트를 생성하였고, YES24 웹사이트를 통해 주간 베스트셀러 목록을 확보

3.1 데이터 수집 및 가공

본 연구에 필요한 데이터 수집 방식과 출처, 유형은 다음 <표 1>과 같다. 먼저, 공공도서관 대출 데이터의 경우, 문화체육관광부 산하 한국문화정보원이 운영하는 문화 빅데이터 플랫폼(<https://www.bigdata-culture.kr>)을 통하여 국립중앙도서관이 제공한 공공도서관 대출 데이터와 공공도서관 소장 도서 목록을 활용하



<그림 1> 데이터 수집 및 분석 절차

<표 1> 데이터 수집 개요

데이터 유형	데이터 수집 방식	데이터 출처
공공도서관 대출 데이터	문화 빅데이터 플랫폼 활용	국립중앙도서관
공공도서관 소장 도서 목록	문화 빅데이터 플랫폼 활용	국립중앙도서관
YES24 주간 베스트셀러 목록	YES24 웹사이트	YES24
ISBN 서지정보 API	ISBN 서지정보 API	국립중앙도서관

였다. 공공도서관 대출 데이터의 경우, 대출일자
와 함께 공공도서관의 코드 및 각 도서관의 도
서 제어번호를 제공하고 있다. 하지만 도서의
ISBN, 제목 등과 같은 서지정보를 제공하지 않
아 어떤 도서가 대출되었는지를 판단할 수 없다.
공공도서관 소장 도서 목록의 경우, 공공도서관
대출 데이터와 같이 공공도서관의 코드 및 도서
관의 도서 제어번호를 제공할 뿐만 아니라 도서
의 ISBN, 제목 등과 같은 서지정보를 제공한다.
따라서 대출 도서에 대한 상세한 정보를 파악하
기 위하여 공공도서관 코드 및 도서 제어번호를
기준으로 공공도서관 소장 도서 목록과 결합한
대출 정보 데이터셋을 생성하였다.

베스트셀러 데이터의 경우, 매출액 기준 온
라인서점 1위이자, 온·오프라인 서점 종합 2
위를 차지하고 있는 YES24를 대상으로 선정
하였다(대한출판문화협회, 2020). YES24 주
간 베스트셀러 목록(<http://www.yes24.com/24/category/bestseller>)의 경우, 1년 52주를
기준으로 주간마다 베스트셀러 목록을 제공할
뿐만 아니라, 주간 베스트셀러 순위를 1위부터
1,000위까지 폭넓게 제공하고 있다. 또한, 주간
베스트셀러 목록을 엑셀을 통해 다운로드를 하
여 쉽게 수집 가능하다는 장점이 있다. 베스트
셀러 목록 데이터는 2018년 1월 1일부터 2019

년 12월 29일까지, 총 104주를 범위로 YES24
의 주간 베스트셀러 목록을 주간마다 1,000위
까지 수집하였다. 부가적으로 도서 대출에 영
향을 미칠 수 있는 도서 정보(즉, 독자 대상, 주
제 분야)를 추가로 활용하고자 국립중앙도서관
ISBN 서지정보 API를 활용하여 ISBN 부가기
호(국립중앙도서관, 2018)를 수집하여 ‘독자대
상기호’와 ‘내용분류기호’를 베스트셀러 도서
정보에 추가하였다.

3.2 데이터 개요

데이터 수집 및 가공과정을 통해 생성한 데이
터 세트는 공공도서관 주간별 대출 데이터와 베
스트셀러 도서 주간별 순위 데이터이다. 이들
데이터 세트의 세부 요소는 다음 <표 2>와 같다.
먼저, 공공도서관 대출 데이터는 2018~2019년
까지 전국 659개 공공도서관의 172,818,097건의
대출 이력이 존재했으며, 이 가운데 ISBN을 기
준으로 고유한 도서의 수는 1,125,935권으로 나
타났다. 이를 바탕으로 2018년 1월 1일부터 2019
년 12월 29일까지 총 104주 동안 전국 공공도서
관 도서의 주간별 대출 건수 데이터 세트를 생
성하였다.

베스트셀러 도서 순위 데이터는 YES24 주간

<표 2> 데이터 세트 개요

데이터 세트	세부 요소
공공도서관 (주간별) 대출 데이터	ISBN, 도서 제어번호, 공공도서관 코드, 주간
베스트셀러 도서 (주간별) 순위 데이터	ISBN, 순위, 주간, 도서 제목, 저자, 출판사, 독자대상기호 [†] , 내용분류기호 [‡]

† 독자대상기호는 0: 교양, 1: 실용, 2: 여성, 3: 예비번호, 4: 청소년, 5: 중·고교용 학습참고서, 6: 초등용 학습참고서, 7: 아동, 8: 예비번호, 9: 전문·학술로 구성됨

‡ 내용분류기호는 0: 총류·컴퓨터과학, 1: 철학·심리학·윤리학, 2: 종교, 3: 사회과학, 4: 자연과학, 5: 기술과학, 6: 예술, 7: 어학, 8: 문학, 9: 역사·지리·관광으로 구성됨

베스트셀러 목록을 활용하였으며, 2018년 1월 1일부터 2019년 12월 29일까지 총 104주 동안 매주 1,000위까지의 도서 목록을 수집하였다. 총 103,050건의 도서 목록이 취합되었으며, 이 가운데 ISBN을 기준으로 식별한 고유한 도서의 수는 12,592권이였다. 이를 바탕으로 도서의 주간별 순위 데이터 세트를 생성하였다. 도서의 주간별 순위가 1,000위를 벗어나는 경우 순위를 파악할 수 없으므로 분석에서 제외하였다.

베스트셀러 순위와 대출 건수 간의 관계를 분석하기 위하여 104주의 기간 동안 최소 1주라도 100위 안에 진입한 이력이 있고, 공공도서관 대출 이력이 존재하는 도서를 선별하였다. 그러나 분석 대상 도서 선별과정 중 베스트셀러 도서 가운데는 일시적으로 100위권에 진입한 후, 1,000위 밖으로 빠르게 벗어나는 경우가 상당수 존재하였다. 데이터 분석에 있어 결측된 순위 정보는 분석 결과를 편향시킬 가능성이 존재하여 이를 최소화하기 위해 다수의 결측이 포함된 도서들을 제외하는 과정이 필요하였다. 순위 정보에 있어 결측된 횟수가 적은 상위 10% 도서는 104주 중 47주 미만으로 1,000위권 밖으로 벗어난 것으로 나타났고, 해당 도서들을 분석과정에 포함하였다. 결과적으로 베스트셀러 도서 179권이 분석 대상으로 선정되었고, 이들 도서의 순위와 공공도서관에서의 대출 건수 간의 관계를 분석하였다.

3.3 데이터 분석

데이터 분석의 첫 번째 단계로는 순위 결측값이 전혀 없는 도서의 평균 순위 변화와 공공도서관 평균 대출 건수 변화를 살펴보고, 분석

대상 도서 179권의 내용분류에 따른 기술 통계치(평균, 표준편차, 최소값, 최대값)를 살펴보았다. 그리고 시간의 흐름에 따른 내용분류별 도서의 평균 순위 변화와 평균 대출 건수 변화를 시각화하여 각각의 특성을 파악하였다.

다음 단계로 패널분석을 수행하였는데, 패널 분석 방법은 예측변수가 오차항과 어떤 방식으로 상관관계를 가지는가에 따라 다양한 방법론이 존재한다. 패널분석 방법 중 가장 일반적인 선형회귀모형은 모형에 포함된 오차항과 예측변수 간 독립을 가정한다. 그러나 자료의 특성에 따라 오차항과 예측변수가 서로 상관을 가지는 내생성의 문제가 발생할 수 있다. 이러한 내생성의 문제에 따른 추정의 편의(bias)를 최소화하는 방법으로 고정효과 모형이 있다.

선형회귀모형은 패널 자료의 시간에 따른 차이뿐 아니라 한 시점의 관찰치 간 차이를 모두 사용하여 독립변수의 종속변수에 대한 효과를 추정하는 방법이다. 만일 선형회귀모형을 바탕으로 추정된 독립변수의 효과가 유의하다면, 이는 시간의 흐름에 따른 관찰치의 변화 때문일 수도 있으며, 관찰치 간 차이와 관련된 다른 관찰되지 않은 이질성에 의한 차이 때문일 수도 있다. 반면 고정효과 모형은 한 시점의 관찰치 간 차이는 고려하지 않고 동일한 관찰치의 시간에 따른 독립변수의 변화만을 고려하여 종속변수에 미치는 효과를 파악한다. 이 방법의 특징은 관찰치 간 차이 때문에 발생하는 효과를 비관측된 오차로 흡수시켜 통제된 상황에서 독립변수의 변화가 종속변수에 미치는 영향 자체를 파악하는 것이다. 구체적으로 패널자료 분석은 다음 수식(1)으로 표현할 수 있다.

$$y_{ti} = \alpha + \beta X_{ti} + u_i + e_{ti} \quad (1)$$

모형이 포함하고 있는 두 개의 오차항 중 (u_i)는 시간에 따라 변화하지 않는 관찰치의 이질적 특성을 나타내며, (e_{ti})는 관찰치 간에도 변화하며, 시간에 따라서도 변화하는 오차를 나타낸다. 모수의 추정과정에서 관찰치의 이질적 특성(u_i)을 추정해야 할 고정된 모수로 간주하면 고정효과모형을 적용하고, 특정 분포에 대한 가정을 바탕으로 한 확률변수로 간주하면 확률효과모형을 적용한다.

본 연구에서는 공공도서관(주간별) 대출과 베스트셀러 도서 순위 간 정확한 관계를 확인하기 위해 패널데이터의 특성을 활용하여 다른 추정방법을 사용하는 세 개의 회귀분석 방식을 비교하였다. 분석과정에서 도서의 종류에 따른 베스트셀러 도서의 순위 효과의 차이를 살펴보기 위해 독자대상기호와 내용분류기호를 통한 세부 분류를 바탕으로 자료를 나누어 분석하였다. 그러나 분석과정에서 내용분류의 효과만이 대출량에 유의한 효과를 갖는 것으로 나타나 독자대상에 따른 분석은 제외하였다. 분석은 내용분류에 따른 베스트셀러 도서 순위가 공공도서관 주간별 대출에 영향을 미치는지를 확인하기 위해 다중회귀분석을 실시한 후(모형1), 다음으로 동일한 변수들을 이용해 일련의 패널 회귀분석을 실시하였다.

패널 회귀분석에서는 고정효과모형(모형2)과 확률효과모형(모형3) 분석을 실시하여 결과를 비교하였다. 구체적으로 고정효과모형과 확률효과모형의 적합성을 확인하기 위해서 개별 도서의 이질적 특성과 독립변수 간에 상관관계가 없다는 영가설에 대한 하우스만 검정

(Hausman, 1978)을 통해 두 모형 간 적합성을 비교하였다. 모든 분석은 'R' 4.1.2 프로그램을 이용하여 수행하였다.

4. 연구결과

4.1 기술통계 분석

본 연구의 분석 대상 도서는 104주 동안 최소 1주 이상 100위권에 들어간 적이 있으며, 최소 57주 이상 1,000위권에 들어가 있었던 도서 179권이다. 이들 도서 가운데 104주 전체기간 동안 1,000위권에 속한 도서 43권(순위 결측치가 없는 도서)의 순위 변화를 살펴보면 <표 3>과 같다. 실제 데이터 분석에서는 주간 순위 데이터가 활용되었으나 아래에서는 가독성을 높이기 위해 13주간 평균 순위를 제시하였다.

같은 기간 동안 이들 도서(순위 결측치가 없는 도서)에 대한 공공도서관 대출 건수는 다음과 같다. 실제 데이터 분석에서는 주간 대출 건수 데이터를 활용하였으나, <표 4>에서는 가독성을 높이고자 13주간 공공도서관 대출 건수 합계에 대한 평균값을 제시하였다.

분석 대상인 베스트셀러 도서 179권의 내용분류에 따른 기술 통계치 결과는 <표 5>와 같다. 내용분류에 따른 베스트셀러에서의 평균 순위는 어학 분야 도서가 해당 기간 내 평균 323.0위(SD=293.3)로 가장 높았으며, 다음으로 역사·지리·관광 분야 357.5위(SD=350.5)등의 순위였으며, 총류·컴퓨터과학 분야 도서의 순위가 723.9위(SD=289.4)로 가장 낮았다. 공공도서관 평균 대출 건수를 살펴보면 문학 분야 도서

〈표 3〉 순위 결측치가 없는 도서의 13주간 평균 순위 변화

도서명	2018년				2019년			
	1주-13주	14주-26주	27주-39주	40주-52주	1주-13주	14주-26주	27주-39주	40주-52주
82년생 김지영	4.5	8.5	33.1	30.6	65.1	126.0	213.0	17.5
ETS TOEIC 토익 정기시험 기출문제집 RC 리딩	29.6	56.6	44.5	88.2	284.1	614.1	304.2	569.2
개인주의자 선언	52.6	67.5	28.5	94.0	134.8	200.4	239.8	345.2
꽃을 보듯 너를 본다	254.2	170.3	243.8	168.6	10.3	32.5	60.5	49.3
나는 나로 살기로 했다	14.8	13.0	17.2	12.3	12.3	19.6	27.7	28.9
나미야 잡화점의 기적	8.2	17.8	16.8	34.4	47.1	74.5	170.1	195.6
데미안	242.8	281.7	300.8	291.8	229.8	277.7	244.8	263.8
라틴어 수업 리커버	36.5	92.4	191.8	219.1	357.6	275.1	505.8	331.8
말 그릇	13.3	35.1	118.8	215.8	289.5	315.6	449.1	508.0
말의 품격	26.5	44.2	65.7	106.9	134.5	208.8	316.3	404.5
못 잡는 아이 육하는 부모	177.1	169.5	191.7	204.8	135.2	113.4	188.1	239.6
미움받을 용기	114.8	133.3	162.2	155.5	148.5	147.6	199.5	231.5
부의 추월차선	79.8	99.5	119.0	78.5	134.0	232.9	240.2	241.5
사피엔스	81.2	64.1	28.3	31.4	37.4	70.2	61.1	16.8
설민석의 조선왕조실록	94.2	86.7	63.7	48.8	82.6	106.0	72.0	39.5
설민석의 한국사 대모험 1	94.2	67.0	65.5	64.3	100.8	63.8	122.9	127.5
설민석의 한국사 대모험 2	103.0	82.9	77.9	76.4	122.2	81.4	152.5	164.3
설민석의 한국사 대모험 3	101.2	96.5	95.5	89.9	156.9	93.4	187.5	208.5
설민석의 한국사 대모험 4	83.4	91.2	97.4	93.3	164.0	102.8	222.4	280.6
설민석의 한국사 대모험 5	36.6	61.7	86.1	91.2	155.2	105.6	215.5	268.6
세상에서 제일 쉬운 엄마표 생활영어	182.2	381.4	435.5	411.2	173.2	390.5	509.8	543.9
소년이 온다	299.5	123.7	164.5	250.5	330.0	186.1	280.5	403.0
신경 끄기의 기술	3.7	40.1	132.6	161.2	295.4	434.5	602.9	416.8
이문드	183.2	91.2	63.5	103.8	99.2	63.4	47.6	68.1
아홉 살 마음 사전	177.9	97.4	191.7	120.6	100.3	108.4	233.1	161.2
알사탕	229.4	90.4	211.2	190.6	385.5	108.6	165.0	121.2
어떻게 살 것인가	28.9	41.9	38.8	68.8	153.4	189.4	254.5	492.2
언어의 온도	7.5	13.0	9.2	19.7	34.9	78.8	81.2	75.3
여행영어 100일의 기적	136.2	144.3	144.5	214.2	230.2	361.5	393.1	429.2
영어회화 100일의 기적	34.6	46.2	55.2	76.8	68.4	104.5	100.5	136.2
왕초보 영어회화 100일의 기적	296.0	401.8	328.9	391.8	156.5	300.2	363.4	564.6
유시민의 글쓰기 특강	118.2	160.1	152.8	178.3	334.5	427.5	402.1	547.4
일빵빡 입에 달고 사는 기초영어	98.1	106.3	139.4	209.0	207.2	274.1	322.2	494.2
자존감 수업	42.9	42.5	72.2	92.2	79.0	110.9	130.5	162.6
지적 대화를 위한 넓고 얇은 지식: 현실너머 편	123.1	161.8	197.5	169.4	171.4	261.4	234.1	239.7
충, 균, 쇠	372.2	399.5	301.4	208.1	204.5	227.1	204.7	91.1
추리 천재 엉덩이 탐정 1	349.2	201.3	180.6	253.5	273.6	160.4	93.3	225.1
추리 천재 엉덩이 탐정 2	405.4	236.3	231.8	283.2	374.2	196.9	118.5	285.3
추리 천재 엉덩이 탐정 3	436.2	304.5	215.2	318.5	452.2	227.9	154.4	382.8
코스모스	94.0	95.0	118.6	87.6	59.8	48.5	52.4	56.8
타이탄의 도구들	177.5	155.2	366.4	177.3	327.8	412.5	140.5	186.0
호모 테우스	118.7	160.9	104.2	91.4	169.6	321.4	290.0	258.2
환자 혁명	338.0	286.1	246.8	142.1	188.7	180.3	258.2	287.0

비고, 도서 목록은 자모순 배열임

〈표 4〉 순위 결측치가 없는 도서의 공공도서관 대출 건수 합계 평균

도서명	2018년				2019년			
	1주-13주	14주-26주	27주-39주	40주-52주	1주-13주	14주-26주	27주-39주	40주-52주
82년생 김지영	605.4	656.5	681.6	666.6	655.5	575.2	542.9	621.9
ETS TOEIC 토익 정기시험 기출문제집 RC 리딩	0.5	0.6	0.5	0.8	1.0	0.8	1.2	0.8
개인주의자 선언	204.9	230.1	281.8	249.5	271.3	261.5	258.2	197.5
꽃을 보듯 너를 본다	82.4	83.5	92.5	118.8	203.7	184.9	157.5	153.2
나는 나로 살기로 했다	304.6	342.7	346.5	334.8	356.6	351.9	346.3	306.6
나미야 잡화점의 기적(봄날 에디션)	550.3	573.4	589.1	568.4	577.1	504.5	468.8	365.1
데미안	132.7	117.4	137.5	122.8	140.8	112.9	134.3	130.5
라틴어 수업 리커버	248.8	246.5	221.7	183.2	185.3	170.5	138.8	130.6
말 그릇(벚꽃 에디션)	184.5	261.6	243.7	194.4	201.2	175.0	163.8	125.1
말의 품격	329.5	341.2	339.1	294.8	287.1	222.2	217.5	175.5
못 잡는 아이 육하는 부모	181.8	178.3	172.3	145.9	159.3	156.8	145.1	113.8
미움받을 용기	371.9	348.5	349.2	323.6	351.0	299.4	288.4	240.9
부의 추월차선	143.2	139.2	131.9	138.2	154.3	131.6	130.3	125.5
사피엔스	283.2	261.5	318.1	313.6	320.0	288.5	297.8	312.2
설민석의 조선왕조실록	307.2	262.5	308.7	266.2	286.1	223.5	264.4	257.5
설민석의 한국사 대모험 1	230.4	232.1	269.9	256.3	281.3	261.4	280.3	261.5
설민석의 한국사 대모험 2	232.4	223.4	266.0	246.3	274.5	252.8	273.0	249.3
설민석의 한국사 대모험 3	233.4	222.5	265.3	241.3	275.2	249.2	267.2	251.3
설민석의 한국사 대모험 4	216.8	226.5	269.2	255.4	284.2	260.4	291.5	275.3
설민석의 한국사 대모험 5	117.9	221.7	271.5	254.1	283.3	254.0	273.9	247.1
세상에서 제일 쉬운 엄마표 생활영어	28.6	59.4	52.6	45.5	58.5	43.8	36.4	31.8
소년이 온다	301.4	310.6	345.2	275.3	291.9	275.7	274.8	228.0
신경 끄기의 기술	292.2	347.5	336.8	290.8	267.2	195.7	188.2	162.8
이문드	255.2	305.2	331.6	313.4	333.5	302.0	331.8	297.5
아홉 살 마음 사전	173.8	182.9	200.9	166.4	185.9	205.9	211.3	179.2
알사탕	245.2	282.8	289.1	260.3	263.0	307.8	302.3	266.8
어떻게 살 것인가	285.5	283.8	316.6	286.2	267.8	229.7	214.3	154.2
언어의 온도	379.3	426.2	460.4	449.0	450.3	380.3	376.3	318.9
여행영어 100일의 기적	60.5	60.5	63.0	56.2	63.5	54.9	51.2	43.7
영어회화 100일의 기적	103.2	93.1	96.1	81.5	91.8	73.7	71.0	66.2
왕초보 영어회화 100일의 기적	61.8	55.8	55.3	47.2	61.4	46.5	49.2	36.5
유시민의 글쓰기 특강	199.2	177.5	193.5	165.8	164.0	133.7	138.2	99.9
일빵빡 입에 달고 사는 기초영어	70.2	62.8	62.5	54.3	63.2	50.7	48.8	40.5
자존감 수업	337.6	329.8	321.0	282.3	289.0	255.8	239.2	189.2
지적 대화를 위한 넓고 얇은 지식: 현실너머 편	262.8	230.8	245.1	224.6	236.6	180.7	192.0	171.5
총, 균, 쇠	147.1	132.1	149.1	144.1	155.8	130.9	140.3	175.3
추리 천재 엉덩이 탐정 1	283.4	301.6	332.4	308.2	357.8	351.8	413.6	361.5
추리 천재 엉덩이 탐정 2	289.2	293.8	344.5	316.2	356.2	348.9	411.0	370.2
추리 천재 엉덩이 탐정 3	281.5	296.0	358.1	331.2	368.1	363.7	422.7	382.5
코스모스	124.2	106.8	118.2	116.7	136.4	119.8	136.8	126.1
타이탄의 도구들(리커버 에디션)	140.7	137.0	131.4	144.4	137.5	119.4	153.5	134.1
호모 테우스	206.5	173.7	207.4	203.3	198.8	159.3	166.5	145.5
환자 혁명	66.2	67.3	73.7	92.2	106.2	103.1	92.4	86.8

비고, 도서 목록은 자모순 배열임

〈표 5〉 내용분류에 따른 기술 통계치

분류	표본수	2018-2019 베스트셀러 순위				2018-2019 공공도서관 대출 건수			
		평균	표준편차	최소값	최대값	평균	표준편차	최소값	최대값
총류·컴퓨터과학	1	723.9	289.4	25	1000	9.7	9.4	0	28
철학·심리학·윤리학	17	400.0	353.2	1	1000	156.3	106.8	0	435
사회과학	30	532.9	354.7	1	1000	95.6	71.4	0	375
자연과학	5	429.9	407.0	2	1000	96.1	93.1	0	373
기술과학	8	404.9	315.0	5	1000	92.6	63.2	0	245
예술	5	431.2	337.3	11	1000	45.0	58.9	0	237
어학	23	323.0	293.3	1	1000	38.5	50.0	0	275
문학	64	411.1	343.2	1	1000	204.3	130.9	0	755
역사·지리·관광	26	357.5	350.5	1	1000	141.2	112.7	0	449
합계	179	446.1	338.2	1	1000	97.7	77.4	0	755

가 기간 내 평균 204.3건(SD=130.9)으로 가장 높았으며, 다음으로 철학·심리학·윤리학 관련 도서가 156.3건(SD=106.8), 역사·지리·관광 관련 도서가 141.2건(SD=112.7) 등의 순이었다. 총류·컴퓨터과학 관련 서적의 대출량은 평균 9.7건(SD=9.4)으로 가장 낮았다. 특히, 어학 분야 도서의 경우 평균 베스트셀러 순위는 가장 높으나, 공공도서관 평균 대출 건수는 매우 낮았는데, 이는 어학 관련 수험서나 학습참고서 등이 서점에서는 높은 판매량을 보이나 공공도서관에서는 이러한 개인학습 위주의 도서를 원칙적으로 배제(도서관연구소, 2010)하였기 때문으로 볼 수 있다.

다음으로 시간의 흐름에 따른 베스트셀러 순위와 대출 간의 변화를 기술적으로 살펴보았다. 먼저, 시간의 흐름에 따른 내용분류별 베스트셀러 도서의 평균 순위 변화를 그림으로 나타낸 것은 〈그림 2〉와 같다. 그 결과, 시간이 지날수록 순위가 상승하다가 하락하는 'U자' 형태를 띠는 경우가 많았다. 이는 사회문화적 상황과 더불어 전처리 과정에서 순위 결측치가 많은 도서를 제외한 것에 영향을 받은 것으로 추

측된다. 구체적으로 전처리 과정에서 2018년 1월부터 2019년 12월까지의 기간에서 시기적으로 중간 시점에 베스트셀러에 오른 도서는 순위 정보가 보전되어 분석에 포함될 가능성이 높으나 시기적으로 2018년 초와 2019년 말에 베스트셀러 순위가 하락하고 있었던 도서들은 분석에 제외되었을 가능성이 있다. 평균 순위의 변화 폭이 큰 분야의 경우, 출판 외적인 이유로 해당 분야에 속한 도서의 판매량 변화가 컸을 가능성이 있다. 다만, 총류·컴퓨터과학이나 자연과학의 경우, 표본 수가 적어 변화 폭이 컸을 것으로 추측된다. 문학이나 사회과학, 역사·지리·관광 분야의 경우, 변화 폭이 상대적으로 크지 않았는데, 이러한 분야에 속한 도서 가운데는 꾸준한 판매량을 갖는 도서가 많은 것으로 볼 수 있다.

개별 도서의 공공도서관에서의 대출 건수를 합산한 후, 내용분류별로 평균을 구하였으며, 이를 시간의 변화에 따라 시각화한 것은 〈그림 3〉과 같다. 그림에서 살펴보면, 시간이 지날수록 도서의 대출량은 증가하다가 다시 감소하는 '뒤집어진-U자' 경향을 보이는 경우가 많았다.



〈그림 2〉 내용분류별 베스트셀러 도서 평균 순위 변화

이러한 경향은 앞서 베스트셀러 도서의 평균 순위 변화가 'U자' 형태를 보이는 것과 상반되는 것을 알 수 있다. 종합하면, 도서의 베스트셀러 순위의 변화('U자 형태')와 대출 건수 합계 평균의 변화('뒤집어진-U자 형태')를 시각적으로

확인한 결과, 이들 간에는 선행연구(윤휘영, 2021; 이종엽, 2012)에서 확인한 바와 같이 어느 정도의 상관 또는 영향 관계가 존재할 것으로 예상된다.



〈그림 3〉 내용분류별 베스트셀러 도서 대출 건수 합계 평균의 변화

4.2 베스트셀러 순위와 대출 건수 관계 분석

베스트셀러 순위가 공공도서관 대출 건수에 미치는 영향을 파악하기 위해 선형회귀모형, 고

정효과모형, 확률효과모형 분석을 실시하였다. 분석과정에서 세부 도서 유형별 차이를 확인하기 위해 활용한 독자대상분류와 내용분류의 경우, 내용분류의 효과만이 대출량에 미치는 효과가 유의하여 독자대상분류는 분석에 제외하

였다.

본 연구의 최종모형을 선정하기 위해 가장 먼저 선형회귀모형의 오차항을 분석한 결과, 자기상관이 통계적으로 유의한 것으로 나타났다 ($\chi^2=12811$, $df = 105$, $p<0.01$). 이는 오차항의 독립성 가정이 위배되었다는 의미이며, 이 경우 선형회귀모형의 추정치는 예측변수의 효과를 과대 혹은 과소추정하는 편향의 가능성이 있다. 추가로 선형회귀 모형과 고정효과 모형의 적합성을 비교한 결과, 선형회귀 모형보다 고정효과 모형이 더 적합한 것으로 나타났다($F(170, 18615) = 217.5$, $p<0.01$).

다음으로 하우스만(Hausman) 검정을 시행한 결과, 개체 간 이질적 특성과 독립변수 간에 상관관계가 없다는 영가설이 기각되었으며($\chi^2 = 110.23$, $df = 1$, $p<0.01$), 이는 고정효과 모형이 확률효과 모형보다 자료를 설명하는데 더욱 적합하다는 것을 의미한다. 또한 이론적으로도 도서 대출 건수와 같이 유동적이고 복합적인 특

성을 지닌 변수를 설명하기 위해서는, 연구모형에서 관찰되지 않는 개별 도서 간의 보이지 않는 특성이 독립변수에 영향을 미치는 내생성을 통제할 수 있는 고정효과모형을 적용하는 것이 바람직할 것으로 판단된다.

결과적으로 선형회귀 모형은 자기상관이 존재하였고, 하우스만 검정 결과에 따르면 확률효과 모형은 고정효과 모형에 비해 적합하지 않은 것으로 나타났으므로 본 연구에서는 고정효과 모형을 결과의 해석을 위한 최종모형으로 선택하였다. 베스트셀러 순위가 공공도서관 도서 평균 대출 건수 변화에 미치는 영향을 선형회귀분석 모형, 고정효과모형, 확률효과모형을 통해 분석한 결과는 다음 <표 6>과 같다.

다음 결과에 따르면 고정효과모형에서 결정계수가 선형회귀모형에 비해 더 높게 나타났다. 이러한 결과는 선형회귀분석에서 고려하지 못한 관찰되지 않은 개별 도서 간 차이를 고정효과모형이 추가로 설명하였기 때문이다. 또한

<표 6> 패널 회귀분석 결과

분류	분석 방법					
	선형회귀(모형1)		고정효과(모형2)		확률효과(모형3)	
	회귀계수	SE	회귀계수	SE	회귀계수	SE
총류·컴퓨터과학	-0.016**	0.003	추정불가	추정불가	추정불가	추정불가
철학·심리학·윤리학	-0.212**	0.005	-0.150**	0.004	-0.150**	0.004
사회과학	-0.102**	0.003	-0.107**	0.002	-0.095**	0.002
자연과학	-0.146**	0.008	-0.084**	0.009	-0.137**	0.007
기술과학	-0.128**	0.005	-0.086**	0.005	-0.098**	0.005
예술	-0.089**	0.007	-0.071**	0.004	-0.077**	0.004
어학	-0.009**	0.003	-0.012**	0.001	-0.018**	0.001
문학	-0.204**	0.004	-0.128**	0.002	-0.137**	0.003
역사·지리·관광	-0.189**	0.005	-0.090**	0.004	-0.099**	0.004
전체	-0.151**	0.002	-0.108**	0.002	-0.108**	0.002
	R ²	0.192	R ²	0.271	R ²	0.274

**p<.001.

고정효과모형과 확률효과모형 간 회귀계수와 결정계수는 거의 차이가 없었다.

시간에 따라 불변하는 개체 간 차이를 고정된 모수로 추정시켜 통제된 상태에서 시간에 다른 개체 내 변동의 효과만을 고려하였을 때, 시간의 변화에 따라 베스트셀러 순위는 대출량에 부적인 영향을 보였다($b = -0.108, p < .001$). 구체적으로 어떤 도서의 베스트셀러 순위가 한 단계 내려가면 공공도서관에서의 해당 도서 평균 대출량은 0.108권 정도가 유의미한 수준에서 감소한다는 것이다.

도서의 내용분류에 따른 베스트셀러 순위가 도서 평균 대출량에 미치는 효과의 상대적인 크기를 살펴보면, 철학·심리학·윤리학($b = -0.150, p < .001$)과 문학($b = -0.128, p < .001$) 분야의 순위가 가지는 효과는 다른 분야에 비해 상대적으로 크게 나타났으며, 어학($b = -0.012, p < .001$)과 예술($b = -0.071, p < .001$) 분야의 효과는 적은 것으로 나타났다. 이는 철학·심리학·윤리학 및 문학 분야의 도서의 경우, 높은 베스트셀러 순위가 오르는 것이 도서 평균 대출량에 좀 더 크게 영향을 미치는 것을 의미하며, 어학과 예술 분야의 도서는 해당 영향이 상대적으로 적음을 의미한다.

이러한 차이가 발생한 이유를 살펴보기 위해 분석 대상인 된 어학과 예술 분야 도서를 살펴본 결과, 어학 분야에는 주로 토익 시험 준비를 위한 도서나 외국어 회화 관련 도서가 많았고, 초등학생을 위한 국어사전이 속해 있었고, 예술 분야에는 종이접기 관련 도서나 악기 교본 또는 미술 작품 감상 관련 도서가 포함되어 있었다. 즉, 이들 도서들은 꾸준한 수요가 존재하는 것으로 베스트셀러 순위에 영향을 덜 받은

것으로 추측된다. 반면, 철학·심리학·윤리학 및 문학 분야의 도서에는 시간이 흐름에 따라 베스트셀러 목록에 새롭게 진입하는 도서의 비율이 높아 도서 평균 대출량에 더 큰 영향을 미칠 수 있을 것으로 판단된다.

추정된 도서의 내용분류별 회귀계수를 모형별로 비교한 결과, 선형회귀분석에 비해 고정효과 모형의 계수 추정치는 전반적으로 낮은 수준을 보였으며, 확률효과모형의 회귀계수 추정치와는 규칙적인 관련성이 발견되지 않았다. 추가로 시간고정효과(time-fixed effect)를 고정효과 모형에 추가하여 시간에 따라 불변하는 개체 간 차이뿐만 아니라 시간에 따라 변하지만, 개체 간에는 일정한 차이도 동시에 고정된 모수로 추정한 결과, 베스트셀러 순위의 추정된 계수는 -0.108로 고정효과모형의 계수와 거의 유사한 결과를 보였다($p < .001$). 이는 개체 내 변동의 효과와 개체 간 차이를 모두 통제할 경우, 베스트셀러 순위가 한 단계 내려가면 도서 평균 대출량은 0.108권 줄어든다는 의미이다.

5. 논의 및 결론

이용자의 공공도서관 대출에 영향을 미치는 요인을 식별하는 것은 이들의 정보요구 및 이용에 대한 이해를 도모하여 도서관 차원에서 이용자 요구에 부합하는 장서를 구성하고, 서비스나 프로그램을 계획하는 데 활용할 수 있다. 이에 따라 기존의 여러 연구들이 공공도서관 대출 데이터를 활용하여 이용자 개개인의 특성을 비롯한 도서관, 지역 및 사회문화적 특성과 결부시키고 있는 상황이다. 그렇지만 사회문화적 특성

과 공공도서관 대출 간의 관계에 대한 연구는 상대적으로 부족한 편이다. 따라서 이 연구에서는 사회문화적 특성을 반영하는 서점의 베스트셀러를 활용하여 우리나라 공공도서관에 소장된 베스트셀러 도서에 대한 대출에 있어 베스트셀러 순위의 영향을 실증적으로 파악하였다. 그 결과, 베스트셀러 목록에 포함된 모든 주제 도서의 평균 순위가 한 단계 하락하면 공공도서관에서의 해당 도서 평균 대출량이 약 0.1권 통계적으로 유의미하게 감소하는 것을 확인할 수 있었다. 이는 베스트셀러 목록 자체에 대한 긍정적 혹은 부정적 측면과 상관없이 베스트셀러가 공공도서관 이용자 대출행태에 부분적으로 영향을 미칠 수 있음을 실제 자료 분석을 통해 확인한 것이다.

본 연구의 결과는 베스트셀러 순위와 대출 간의 관련성을 밝혀낸 이종엽(2012)의 연구 결과와 베스트셀러 순위가 이용자 요구를 예측하는 데 활용될 수 있음을 제시한 윤휘영(2021)의 연구 결과를 뒷받침한다고 볼 수 있다. 한편 이호신(2018)의 연구에서 도서관의 인기대출 도서 목록과 서점의 베스트셀러 목록 간의 중복 비율이 15.5~19.0%에 불과하고, 주제 분포나 발행연도에서도 차이가 존재한다는 결과를 토대로 도서관에서의 도서 대출이 서점에서의 판매를 감소시킬 수 있다는 주장은 바람직하지 않음을 제시한 바 있으나, 본 연구에서는 서점에서의 도서 판매량이 도서관에서의 대출에 영향을 미칠 수 있음을 밝혀내었다. 다만, 본 연구에서 어학 분야의 도서는 평균적인 베스트셀러 순위는 가장 높았지만, 공공도서관에서의 평균 대출 건수는 매우 낮은 것으로 드러난 바 있는데, 이러한 사실은 서점에서의 도서 판매와 도

서관 대출 간에는 어느 정도의 관련성이 존재하지만, 차별성도 보인다는 것을 의미한다.

이 연구에서는 베스트셀러가 사람들의 독서량이나 다양성 또는 도서관 전체 소장도서에 대한 대출에 미치는 영향을 분석한 것은 아니므로 독서나 도서관 대출에 대한 베스트셀러의 기여도를 논하기는 어렵다. 그렇지만 이용자의 도서 대출행태 즉, 정보요구에 영향을 미칠 수 있는 다양한 요소 가운데 베스트셀러 순위도 부분적 영향을 미칠 수 있음을 감안하면, 도서관 차원에서의 장서 개발이나 정보서비스에 이에 대한 적절한 참고가 필요함을 알 수 있다. 이미 공공도서관에서는 도서 선정과정에서 사서가 추천도서 목록과 서평, 독자적 판단뿐만 아니라 베스트셀러 목록을 활용하고 있음이 알려져 있다(박옥화, 2011). 그렇지만 본 연구가 시사하는 것은 사서가 단순히 베스트셀러 목록 자체만을 활용하는 것이 아닌, 이러한 베스트셀러를 형성하는 사회문화적 맥락에 대한 관심과 이해를 높여 이를 장서 구성이나 서비스 및 프로그램 기획에 활용하여 이용자의 정보요구에 선제적 대응이 필요하다는 것이다.

부가적으로 도서의 내용분류에 따라 베스트셀러 순위가 도서 평균 대출량에 미치는 상대적 효과의 차이가 존재하는 것으로 나타났는데, 특히 철학·심리학·윤리학 및 문학 분야의 도서는 베스트셀러 순위가 평균 대출량에 더욱 많은 영향을 받고, 어학과 예술 분야의 도서는 해당 영향이 상대적으로 적은 것으로 나타났다. 이는 도서관 장서 개발에 있어 베스트셀러의 활용도 또는 사회문화적 환경의 영향력이 도서의 주제 분야에 따라 다를 수 있음을 의미한다.

앞서 언급하였듯이, 공공도서관 대출에 영향을 미칠 수 있는 요인은 매우 다양하나 이 연구에서는 이러한 요인 가운데 사회문화적 특성으로 베스트셀러의 영향을 중점적으로 조사하였다. 후속 연구에서는 도서 대출에 영향을 미칠 수 있는 요인을 추가적으로 조사하여, 궁극적으로는 이를 종합적으로 고려하여 이용자의 정보요구와 이용을 예측하여 공공도서관 서비스 정책 수립에 고려할 필요가 있을 것이다. 본 연구는 주요 서점의 모든 베스트셀러 목록을 활용하지 않고 하나의 서점의 목록만 활용하였다

는 점과 베스트셀러 순위 데이터가 1,000위까지만 존재했던 점, 그리고 시간의 변화에 따른 도서 대출 건수 간의 자기상관 효과를 고려하지 못한 점, 베스트셀러에 오르는 시점과 공공도서관 대출이 이루어지는 시점 간의 시간 간격을 고려하지 못한 점 등의 한계가 있다. 이에 후속 연구에서는 이러한 한계점을 보완하여 베스트셀러 목록의 범위를 확장하고, 순위 구간별 도서관 대출 건수에 대한 영향을 분석하며, 나아가 베스트셀러 순위와 도서 대출 간의 시간차 효과도 고려할 예정이다.

참 고 문 헌

- 국립중앙도서관 (2018). 한국문헌번호편람. 서울: 국립중앙도서관.
- 권나현, 송경진 (2014). 한국 성인의 공공도서관 이용에 영향을 미치는 주요 요인 분석: 개인, 도서관의 특성 및 자치단체의 도서관 투자를 중심으로. 한국비블리아학회지, 25(4), 291-312.
<http://doi.org/10.14699/kbiblia.2014.25.4.291>
- 김기태 (1997). 베스트셀러가 출판문화 발전에 미치는 영향. 출판잡지연구, 5(1), 160-176.
- 김기태 (2015). 국내 문학도서의 베스트셀러 요인 분석 연구. 한국문예창작, 14(2), 225-254.
- 김선애 (2016). 도서관 예산과 도서관 이용과의 관계에 관한 연구: 부산지역 공공도서관을 중심으로. 한국비블리아학회지, 27(2), 193-212. <http://doi.org/10.14699/kbiblia.2016.27.2.193>
- 김영조 (2006). 한국의 실용서 베스트셀러 요인에 관한 연구: 최근 10년간 베스트셀러를 중심으로. 석사학위논문, 중앙대학교 신문방송대학원 신문방송학과 정보미디어전공.
- 김완중 (2015). 유행성 질병이 공공도서관의 대출책수에 미치는 영향: 메르스 사태를 중심으로. 정보관리학회지, 32(4), 273-287. <http://doi.org/10.3743/KOSIM.2015.32.4.273>
- 김태영, 백지연, 오효정 (2018). 빅데이터 로그 기반 도서관 이용자 및 대출 현황 분석. 한국도서관·정보학회지, 49(2), 357-388. <http://doi.org/10.16981/kliss.49.2.201806.357>
- 김형욱 (2004). 베스트셀러가 한국출판계에 미친 영향 연구: 독자의 수용형태에 대한 분석을 중심으로. 석사학위논문, 성균관대학교 언론정보대학원 출판잡지학과.
- 대한출판문화협회 (2020). 2019년 출판시장 통계.

- 도서관연구소 (2010). 공공도서관 장서관리매뉴얼. 서울: 국립중앙도서관.
- 문화체육관광부. 국가도서관통계시스템. 출처: <https://www.libsta.go.kr/>
- 박옥화 (2011). 공공도서관 사서들의 베스트셀러에 관한 인식 연구. 사회과학연구, 22(2), 3-22.
- 박태연, 오효정 (2021). 뉴노멀 시대의 도서관 이용자 요구 분석: 소셜 미디어를 중심으로. 한국도서관·정보학회지, 52(2), 303-330. <http://doi.org/10.16981/kliss.52.2.202106.303>
- 심지영 (2021). 도서대출 데이터를 이용한 공공도서관 이용자 정보요구 분석: 서울시 공공도서관을 중심으로. 정보관리학회지, 38(2), 173-199. <http://doi.org/10.3743/KOSIM.2021.38.2.173>
- 안선희 (2009). 한국 출판의 베스트셀러 다양성에 관한 연구: IMF 이후 10년간 베스트셀러를 중심으로. 석사학위논문, 서강대학교 언론대학원 출판 전공.
- 유경중, 박일중 (2009). 도시지역과 군지역에 위치한 공공도서관의 자료이용 특성에 관한 비교연구: 경남지역 4개 공공도서관의 대출기록을 중심으로. 한국비블리아학회지, 20(1), 39-57.
- 윤휘영 (2021). Gradient Boosting 기반 알고리즘을 활용한 도서 대출 빈도 예측 모델에 관한 연구. 석사학위논문, 중앙대학교 대학원 문헌정보학과 정보학전공.
- 이영희 (1998). 한국의 베스트셀러 유형 연구: 1948년부터 1997년까지 50년 간을 중심으로. 석사학위논문, 이화여자대학교 정보과학대학원 언론정보학과.
- 이정연, 김현애 (2021). 비대면 시대의 지식정보취약계층 일상생활 정보요구와 도서관 이용 경험에 관한 연구. 한국비블리아학회지, 32(1), 223-246.
<http://doi.org/10.14699/kbiblia.2021.32.1.223>
- 이종엽 (2012). 베스트셀러와 도서관 대출과의 상관관계 분석: 공공도서관과 대학도서관을 대상으로. 박사학위논문, 경기대학교 대학원 문헌정보학과.
- 이종욱, 강우진, 이명 (2021). 지역사회 사회경제적 박탈이 공공도서관 대출 책수에 미치는 영향. 한국문헌정보학회지, 55(4), 219-243. <http://doi.org/10.4275/KSLIS.2021.55.4.219>
- 이호신 (2018). 공공대출보상권 도입의 타당성에 대한 실증적 검토를 위한 기초 연구: 인기대출도서와 베스트셀러의 비교를 중심으로. 한국문헌정보학회지, 52(1), 179-202.
<http://doi.org/10.4275/KSLIS.2018.52.1.179>
- 이희수, 김기영 (2014). 지역사회 특성에 따른 지역주민의 도서관 요구에 관한 연구: 3개의 지역유형을 중심으로. 정보관리학회지, 31(1), 207-230. <http://doi.org/10.3743/KOSIM.2014.31.1.207>
- 정옥년 (2005). 베스트셀러와 독서교육. 독서연구, 13, 101-144.
- 최지영 (2006). Best seller 분석을 통한 사회동향 파악에 대한 연구. 독서문화연구, 5, 33-63.
- 허선, 정연경 (2014). 대출기록을 통해 본 공공도서관 이용자 연구: 강서·양천지역을 중심으로. 한국비블리아학회지, 25(4), 187-207.
- 허연 (2006). 단행본도서의 베스트셀러 유발요인에 관한 연구: 2000년 이후 출판물을 중심으로. 석사학위논문, 연세대학교 언론홍보대학원 저널리즘전공.

- Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46(6), 1251-1271.
- Sorensen, A. T. (2007). Bestseller lists and product variety. *The Journal of Industrial Economics*, 55(4), 715-738.

• 국문 참고문헌에 대한 영문 표기
(English translation of references written in Korean)

- Ahn, Seon-Hee (2009). A Study on the Diversity of Bestseller Books in Korean Publishing: Focus on Bestseller Books in the 10 years following IMF Management Crisis. Master's thesis, Graduate School of Mass Communication Sogang University, Major in Publications.
- Choi, Jiyoung (2006). The study of social trend by analysing the best sellers. *Reading Culture Research Institute Journal*, 5, 33-63.
- Heo, Sun & Chung, Yeon-Kyoung (2014). A study on public libraries' user behaviors based upon circulation data in gangseo and yangcheon region. *Journal of the Korean Bibliol Society for Library and Information Science*, 25(4), 187-207.
- Heo, Yeon (2006). A study on Best-Selling Factors for the Independent books: Focused on the Best-Sellers since 2000. Master's thesis, Graduate School of Journalism & Mass Communication Yonsei University, Major in Journalism Journalism.
- Jung, Ok-Ryun (2005). Bestseller books as reading instructional resources. *Journal of Reading Research*, 13, 101-144.
- Kim, Hyeong-Ok (2006). A Study on the Effect pf 'Best Sellers' on the Korean Publishing Market: Concentrating on the Analysis of the Types of Readers' Acceptability. Master's thesis, Graduate school of Journalism, Communication and Information, Sungkyunkwan University.
- Kim, Kitae (1997). The effects of best sellers on the development of publishing industries. *Korean Journal of Publishing and Periodical Studies*, 5(1), 160-176.
- Kim, Kitae (2015). A study on factorial analysis of domestic bestselling literary Books. *The Journal of Korean Literary Creative Writng*, 14(2), 225-254.
- Kim, Sun-Ae (2016). A study on the relationship between library budget and library usage. *Journal of the Korean Biblia Society for Library and Information Science*, 27(2), 193-212. <http://doi.org/10.14699/kbiblia.2016.27.2.193>
- Kim, Tae-Young, Baek, Ji-Yeon, & Oh, Hyojung (2018). An analysis of library user and circulation status based on bigdata logs. *Journal of Korean Library and Information Science Society*,

- 49(2), 357-388. <http://doi.org/10.16981/kliss.49.2.201806.357>
- Kim, Wan-Jong (2015). A study on the impact of the epidemic disease on the number of books checked out of the public libraries: based on the middle east respiratory syndrome coronavirus. *Journal of the Korean Society for Information Management*, 32(4), 273-287. <http://doi.org/10.3743/KOSIM.2015.32.4.273>
- Kim, Young-Jo (2006). A Study on the Factors for the Best Seller Among Practical Books in Korea: With Focus on the Best Sellers for the Latest 10 Years. Master's thesis, Graduate School of Mass Communication Chung-Ang University, Major in Information & Media of Publishing.
- Korea Research Institute for Library and Information (2010). *Public Library Collection Management Manual*. Seoul: National Library of Korea.
- Korean Publishers Association (2020). 2019 Publishing Market Statistics.
- Kwon, Nahyun & Song, Kyeong-Jin (2014). A national study explaining the public library use among korean adults: examining the influence of individual characteristics, local library inputs, and local government investments. *Journal of the Korean Biblia Society for Library and Information Science*, 25(4), 291-312. <http://doi.org/10.14699/kbiblia.2014.25.4.291>
- Lee, Heuisoo & Kim, Giyeong (2014). Study on the needs of local residents for library services based on community characteristics: focusing on three types of local communities. *Journal of the Korean Society for Information Management*, 31(1), 207-230. <http://doi.org/10.3743/KOSIM.2014.31.1.207>
- Lee, Hosin (2018). A preliminary research for empirical analysis on the validity of public lending right: focus on the comparison between popular borrowing books and best sellers. *Journal of the Korean Society for Library and Information Science*, 52(1), 179-202. <http://doi.org/10.4275/KSLIS.2018.52.1.179>
- Lee, Jongwook, Kang, Woojin, & Lee, Myeong (2021). The effects of socioeconomic deprivation on public library book circulation: a community-level study. *Journal of the Korean Library and Information Science*, 55(4), 219-243. <http://doi.org/10.4275/KSLIS.2021.55.4.219>
- Lee, Jong-Yup (2012). The Correlation Analysis of the Bestseller and the Most Frequently Circulated Books in Libraries: Focusing on a Public Library and an Academic Library. Doctoral dissertation, The Graduate School, Kyonggi University.
- Lee, Jungyeoun & Kim, Hyunae (2021). A study on everyday life information seeking and user experience of public library in contactless society. *Journal of the Korean Biblia Society for Library and Information Science*, 32(1), 223-246.

<http://doi.org/10.14699/kbiblia.2021.32.1.223>

- Lee, Young-Hee (1998). The Research about the Type of Bestsellers in Korea: From 1948 to 1997. Master's thesis, Graduate School of Information Science, Ewha Womans University. Ministry of Culture, Sports and Tourism, National Library Statistics System. Available: <https://www.libsta.go.kr/>
- National Library of Korea (2018). A Manual for Korean Document Numbering System.
- Park, Ok-Wha (2011). A study on the perception of public librarians toward the best selling books. *Journal of Institute for Social Sciences*, 22(2), 3-22.
- Park, Tae-yeon & Oh, Hyo-Jung (2021). Analysis of library user needs in the new normal era: focusing on social media. *Journal of Korean Library and Information Science Society*, 52(2), 303-330. <http://doi.org/10.16981/kliss.52.2.202106.303>
- Shim, Jiyoung (2021). Identifying information needs of public library users based on circulation data: focusing on public libraries in Seoul. *Journal of the Korean Society for Information Management*, 38(2), 173-199. <http://doi.org/10.3743/KOSIM.2021.38.2.173>
- Yoo, Kyeong-Jong & Park, Il-Jong (2009). A comparative study of the use characteristic of public library collection in urban and rural areas: focused on the circulation data of four libraries in the Gyung-sangnam-do province. *Journal of the Korean Biblia Society for Library and Information Science*, 20(1), 39-57.
- Yoon, Hwui-Young (2021). Library Loan Prediction using Gradient Boosting based Algorithms. Master's thesis, The Graduate School, Chung-Ang University.