

문화적 특성을 고려한 공공도서관 도서 대출수요 분석 : 대구광역시 시립도서관을 사례로

오민기¹, 김경래², 정원웅², 김건욱^{3*}

¹경북대학교 문헌정보학과 학부과정, ²경북대학교 컴퓨터학부 학부과정,
³대구디지털산업진흥원 빅데이터활용센터 센터장

Analysis of Borrows Demand for Books in Public Libraries Considering Cultural Characteristics

Min-Ki Oh¹, Kyung-Rae Kim², Won-Oong Jeong², Keun-Wook Kim^{3*}

¹Undergraduate, Dept. of Library and Information Science, Kyungpook National University

²Undergraduate, Dept. of Computer Science, Kyungpook National University

³Director, Big Data Center, Daegu Digital Industry Promotion Agency

요약 공공도서관은 거주민들의 광범위한 지식과 이념을 학습하는 공간이며, 삶에 직접적으로 연결되어 있는 만큼 관련된 연구들이 다양하게 수행되어 왔다. 대부분의 선행연구에서 인구, 교통, 접근성, 환경 등의 변수가 도서관 이용에 관련성이 높은 것으로 나타났다. 본 연구에서는 도서대출 이력(1,820,407건)과 회원정보(297,222명) 자료를 기반으로 문화적 특성의 변수를 반영하여 도서 대출수요와 관련성을 분석한 점이 선행연구와의 차별성이라 할 수 있다. 분석결과 기술과학 도서 대비 사회과학, 문학 도서 대출이 증가할수록 도서 대출수요가 증가하는 것으로 분석되었으며, 도서관 층상고객들이 도서 대출수요의 주요한 요인으로 확인되었다. 그 외 다양한 기술통계분석을 활용하여 도서관 도서대출 수요 특성을 분석하였으며, 분석결과를 바탕으로 정책적 시사점과 연구의 한계도 제시하였다. 그리고 문화적 특성이 장소, 시점 등에 따라 변화됨을 감안하면 앞으로 관련연구가 지속되어야 할 것으로 판단된다.

주제어 : 공공도서관, 문화적 특성, 머신러닝, 군집분석, 회귀모형

Abstract Public libraries are a space where residents learn a wide range of knowledge and ideologies, and as they are directly connected to life, various related studies have been conducted. In most previous studies, variables such as population, traffic accessibility, and environment were found to be highly relevant to library use. In this study, it can be said that the difference from previous studies is that the book borrow demand and relevance were analyzed by reflecting the variables of cultural characteristics based on the book borrow history (1,820,407 cases) and member information (297,222 persons). As a result of the analysis, it was analyzed that as the increase in borrows for social science and literature books compared to technical science books, the demand for book borrows increased. In addition, various descriptive statistical analyzes were used to analyze the characteristics of library book borrow demand, and policy implications and limitations of the study were also presented based on the analysis results. and considering that cultural characteristics change depending on the location and time of day, it is believed that related research should be continued in the future.

Key Words : Public Library, Cultural characteristics, Machine Learning, Cluster Analysis, Regression Model

*Corresponding Author : Keun-Wook Kim(aut7767@dip.or.kr)

Received December 30, 2020

Accepted March 20, 2021

Revised January 31, 2021

Published March 28, 2021

1. 서론

국내 공공도서관이 설립되기 시작한 것은 1900년대 초반이나 다양한 현대사 문제들로 인해 공공도서관 활성화는 많은 어려움이 있었으며, 본격적으로 1963년 '도서관법' 제정 이후 국내 공공도서관의 양적 발전이 지속되었다. 이후 공공도서관 수는 급격히 증가하여 지난 40년간 약 27배 이상 증가하였으며[1], 이러한 양적 성장 이후에도 정부 주도하에 공공도서관 설립은 지속적으로 이루어지고 있다.

공공도서관은 거주민들의 광범위한 지식과 이념, 작가의 견해에 대해 학습하고 토론하는 장으로써 민주사회를 발전시키고 유지하는 데 중요한 역할을 하는 공공시설물이며, 도서대출 이외에도 지역사회 문화적 중심점, 만남의 장소 등으로 역할을 하며 지역사회에 문화적·사회적 욕구를 충족하는 중요한 기능을 하고 있다[2]. 또한, 사회적인 측면에서도 공공도서관의 입지는 주거인들의 삶의 질과 직접적으로 관련이 있는 만큼 입지선정에 있어 중요성은 증가하고 있다.

하지만 공공도서관 입지선정에 있어 대다수의 지자체에서 이용자 중심의 수요, 편리성, 문화적 특성을 고려하지 않고, 공급자 중심의 부지 매입 용이성, 예산 등의 입지 검토로 인해 많은 문제점 등이 발생하고 있다. 문화체육관광부(2012) 조사에 따르면 과반수의 성인들이 주거지에서 공공도서관 이용에 어려움이 있는 것으로 분석되었으며, 지역과 소득계층 등에 있어서도 편차가 상당수 존재하는 것으로 나타났다[3].

이로 인해 공공도서관 최적 입지 선정 및 도서 이용 활성화를 위해 다양한 연구들이 활발히 진행되어 왔으며, 대다수 선행연구에서 인구, 교통 접근성, 환경 등의 물리적 변수들과의 관련성을 중심으로 연구가 진행되어 왔으며, 문화적 특성을 고려한 연구는 전무하다. 또한, 최근 4차 산업혁명으로 인해 빅데이터, 인공지능 등의 키워드가 중점적으로 부각되면서 도서관 대출 이력 자료를 활용한 빅데이터 분석도 최근 연구가 진행되고 있으나, 기술통계분석을 활용한 연구가 주를 이루고 있다.

이러한 배경 하에, 본 연구에서는 2019년 1월부터 2020년 7월까지의 대구광역시 공공도서관 도서 대출이력 자료(1,820,407건)와 회원 정보(297,222명), 공공데이터 등의 빅데이터를 결합하여 공공도서관 도서 대출수요에 영향을 미치는 다양한 요인을 분석하는 것에 주안점을 두고자 한다. 이를 위해 선행연구에서 공공도서관 이용에 주요 요인으로 분석된 인구, 교통 접근성, 환경 등

의 변수를 다중회귀모형을 활용하여 분석을 실시하였다.

이는 동일한 요인일지라도 대구광역시라는 공간적인 범위와 시점에 따라 공공도서관 도서 대출수요에 영향을 미치는 정도가 상이하고, 타 지자체와 비교가 가능한 점에서 선행연구에서 분석한 변수를 포함하여 연구를 진행하였다. 또한, 본 연구에서는 공공도서관 도서 대출수요에 영향을 미치는 문화적 특성의 변수를 정의하여 추가로 반영하여 연구를 진행하였으며, 이는 다음과 같다.

첫째, 문화적 특성의 차이는 사람들 간의 관계를 맺는 과정에 의해 발생하는 것으로 공간적 범위를 기준으로 그룹 내의 회원들은 동일한 도서를 읽고, 그룹 간에는 이질적인 도서를 읽을 것으로 판단하였다.

둘째, 그룹 내에서 공공도서관 도서대출의 빈도가 높은 회원들이 많을수록 공공도서관 대출수요에 높은 관련성이 존재할 것으로 판단하여, 공간적 문화적 특성 변수에 반영하여 도서 대출수요와의 관련성을 분석하였다.

본 연구의 분석결과는 공공도서관 도서 관리 주체들이 이용활성화를 위한 다양한 정책 수립 시 활용될 수 있을 것으로 보이며, 해당 분야 연구자들은 문화적 특성의 요인들을 정량적으로 분석할 수 있는 기초자료로 활용될 것으로 기대된다.

2. 선행연구 고찰

국내 공공도서관 입지분석 및 이용 활성화와 관련된 연구들은 대부분 사회·경제적, 물리적 변수들 등을 활용하여 접근하는 연구가 대다수 진행되어 왔으며[2-11], 국외 연구들의 경우에도 사회·경제적, 물리적, 환경 변수를 기반으로 연구가 진행되어 왔다[12, 13].

이를 세부적으로 살펴보면 원종준·원건혁(2010)은 공공도서관 이용 활성화에 영향을 미치는 변수들을 파악하기 위해 변수증감법을 활용하여 다중회귀모형을 구축하였으며, 회귀분석 결과 건물면적, 일회성 문화행사, 인구밀도, 지하철 접근성, 대형점포 시설연계, 지역 중심 입지가 유의미한 변수로 나타났다[4].

김현중·이종길·여관현(2015)는 공간적 접근성을 수식을 통해 산출하고, 인구·경제적 특성, 교통·입지 특성, 자연특성 변수를 반영하는 자료와 결합하여 공공도서관 접근성과의 연관성을 공간계량경제모형을 활용하여 분석을 수행하였으며, 분석결과 공시지가, 간선도로와의 거리, 버스 정류장과의 거리, 주거지 집중도, 표고가 공공도서관 접근성과 관련이 높은 것으로 나타났다[6].

Table 1. Prior Research on Location Analysis and Activation of Public Libraries

Author (Year)	Purpose of Study	Method	Period	Variables
J.J.Won (2010)	The Effect of Locational and Facility Characteristics on Public Library Use	Multiple Regression Analysis	2006	Library Characteristics (Facility, Collection, Events) Population(Population & Population Density) Location Characteristics (Accessibility, Connectivity, Center Hierarchy)
Y.Y.Kim (2014)	Location Evaluation and Optimal Location Selection Using a Local Space Search Approach	Minimum Distance & Coverage Approach, Gravity Model	2010	Service Population, Accessibility
S.Heo (2014)	Library User Behaviors Based upon Circulation Data	Descriptive Statistics	2012	User, Collection, Time and etc
H.J.Kim (2015)	Regional Disparities and Determinants of Spatial Accessibility	Spatial Econometric Model	2010	Population & Economy(Resident, University Graduate Rate, Land Value by Public Announcement) Transportation & Location(Distance to Arterial Road and Public Transport, Residential Concentration) Natural Characteristics(Elevation)
B.C.Park (2015)	Service Effect and Propriety Area Analysis	Spatial Analysis	2013	Service Population, Accessibility
J.W.Yoo (2015) J.W.Yoo (2015)	Selection of a Metropolitan Library (Busan) Assessment of Location Environment of a Regional Central Library (Daegu)	Research, Feasibility Study	-	Social Condition(Location, Cognitive, Centrality between Facilities, Similar Facilities Redundancy, Change Responsiveness) Physical Condition(Possibility of Construction, Convenience of Transportation Facilities, Appropriateness of Site Size, Environmental Comfort) Economic Conditions(Budget) Project Implementation Conditions(Population, Ease of Business, Legality of Laws)
I.Kim (2016)	An Analysis of Geographical Accessibility and Related Implications	Survey	2016	Library Form(open/closed), Accessibility, Transportation
C.H.Yu (2017)	A Study on the Appropriate Location of Public Facilities	Survey, Spatial Analysis	2017	Accessibility (Bus Stop, Types and Lanes of Access Roads) Demand & Redundancy(Resident, Other Public Library) Connectivity(Educational Facilities, Rest Facility)
T.Y.Kim (2018)	An Analysis of Library User based on Bigdata Logs	Descriptive Statistics	2014 ~ 2017	User(Age, Gender, Address), Loan Logs, Wireless AP Access
S.J.Yim (2020)	Selecting an Appropriate Location for Infrastructure Considering Service Area and Accessibility	Spatial Analysis	-	Service Population, Accessibility
F.J. Zakrajše (2020)	Agent-based geographical modeling of public library locations	Spatial Analysis	2017	Location and Visitor of Library, Location and Type of Road Network, Population, Age and etc
N.S. Saman (2020)	Spatial modeling of areas suitable for public libraries	Spatial analysis	2019	Population, Accessibility, Air and noise Pollution, Convenience, Safety and etc

유재우·김신영(2015)과 유재우·구본진·장덕현(2015)은 대구광역시와 부산광역시를 대상으로 지역 대표 도서관 입지 선정에 있어 사회경제적 여건, 물리적 여건, 경제적 여건, 사업 시행 여건 등을 고려하여 전문가 집단을 통한 조사분석 및 타당성 검토로 지역별 대표 도서관 후보지를 제안하였다[8, 9].

유창호·이왕무·박성현·신승식(2017)은 광주광역시 도서관 적정입지 선정에 관한 연구를 진행하였으며, 선행연구를 토대로 접근성, 이용수요 및 중복성, 연계성 요인들을 활용하여 1차 평가를 진행하고, 추가로 설문조사와 공간분석을 통해 공공도서관 필요지역 및 최적 입지를 도

출하였다[10].

그 외에도 대부분 선행연구에서 서비스 인구나 공공도서관과의 접근성 요인을 활용하여 공공도서관 입지평가 및 접근성 분석을 수행하였다[2, 3, 7, 11]. 또한, 대부분의 연구가 설문조사, 행정동 기반의 조사 자료를 기반으로 진행됐으며, 최근에는 공공도서관 이용행태 분석과 활성화를 위한 실제적인 분석을 수행하기 위해 대출 이력 자료를 활용한 연구들도 수행되었다[5, 14].

허산·정연경(2014)는 공공도서관 이용자의 이용행태를 분석하기 위해 서울특별시 강서·양천 지역 공공도서관 2개 도서관의 대출 이력 자료 754,198건을 이용하여

기술통계분석(descriptive statistics analysis)을 하였다. 분석을 통해 두 도서관 간 이용자 계층, 방문 횟수, 자료의 주제, 이용시간 등을 확인하였고 관련된 정책적 제언을 하였다[5].

김태영·백지연·오효정(2018)는 국립세종도서관의 빅데이터 로그를 활용하여 대출 이력과 이용자 정보를 함께 활용하여 기술통계분석을 진행하였다. 이용자 그룹별로 연령별, 거주지별 등의 특성 분석과 그 요인을 분석하기 위한 FGI¹⁾ 및 외부데이터와의 연계분석을 통해 반복되는 이용자와 대출 현황의 특이점을 파악하고 분석해 국립세종도서관에 운영 방안에 대해 제안을 하였다[14].

해외 연구 사례로는 Franc J. Zakrajše · Vodeb, Vlasta(2020)가 슬로베니아의 인구자료와 도로망 데이터, 도서관 통계자료 등을 활용하여 도서관 최적 입지를 선정하였다. 행위자 기반모형을 바탕으로 특정 위치에 거주자들이 특정 도서관을 이용할 확률을 계산하여 이동식 도서관의 최적 정류소 위치를 도출하였고[12], Saman Nadizadeh Shorabeh 외(2020)는 인구, 접근성, 대기 및 소음 공해도, 경사 등을 요인으로 공공도서관 건설에 적합한 공간 모델링을 연구하였다[13].

이상의 선행연구들은 공공도서관 적정입지 및 이용 활성화를 위해 사회·경제적, 환경적 측면에서 분석하여 공공도서관 입지 및 이용 활성화에 대한 객관적 논리를 제 공한 장점이 있으나, 공공도서관 이용에 미치는 요인들이 인구, 교통 접근성, 환경 등으로 다소 제한되어 연구가 진행되었다. 이는 자료의 범위가 설문조사, 행정동 기반의 수집 자료로 한정되어 있고, 공공도서관 개인별 대출 이력 자료를 수집하기에는 현실적 어려움이 존재하기 때문으로 판단된다.

허선·정연경(2014)과 김태영 외(2018)의 연구에서는 공공도서관 대출 이력 자료를 활용하여 분석을 수행하였으나, 기술통계분석으로 한정되어 통계모형 기반의 공공도서관 도서 대출수요와의 관계에 관한 연구는 진행되지 않았다[5, 14]. 더욱이 선행연구들은 인구통계학적인 요인들과의 관련성에 중점을 두고 있어서 도서대출 이력 자료를 활용한 공간단위별 문화적 특성을 반영한 연구는 찾아보기 힘들다.

따라서 본 연구에서는 2019년 1월부터 2020년 7월까지의 대구광역시 공공도서관 도서대출 이력자료와 회원정보를 기반으로 문화적 특성의 변수를 추가하여 공공

도서관 도서 대출수요와의 관련성을 분석한 점이 선행연구와의 차별성이라 할 수 있으며, 이러한 공간을 기반으로 하는 문화적 특성은 동일한 요인일지라도 장소와 위치, 시점 등에 따라 미치는 영향이 다를 수 있음을 감안하면 추후에도 공공도서관 도서 대출이력을 활용한 문화적 특성을 파악하는 연구가 활발히 진행되어야 할 것으로 보인다.

3. 공공도서관 도서 대출수요 분석

3.1 자료 수집

본 연구는 공공도서관의 도서 대출수요를 분석하기 위해 2019년, 2020년 7월까지의 대구광역시 9관의 시립도서관 도서대출 이력 자료(1,820,407건)와 회원 정보 자료(297,222명)를 활용하였다. 제공된 데이터의 대출자 주소는 우편번호 단위로 데이터가 구성되어 있어, 분석을 위해 우편번호 기초구역을 기준으로 데이터를 전처리하여 도서관 도서 대출이력 자료를 생성하였다. 그리고 도서 대출에 영향을 미칠 수 있는 다양한 변수들을 결합하기 위해 국토지리원에서 제공하고 있는 인구 데이터와 그 외 공공데이터인 교통수단별 정류장 위치 정보, 공시지가, 건물정보 등을 수집하여 데이터 세트를 구축하였다. 공공도서관 도서대출 이력 자료의 경우 대출 건별로 이용자 번호, 도서 번호, KDC²⁾, 제목, 출판사, 대출일, 반납일, 관리도서관, 미디어 형태 등이 포함되어 있으며, 도서관 회원 정보는 연령, 연체일(누적), 성별, 대출 건수, 이용자 번호, 우편번호 등이 포함되어 있다(Table 2 참조).

Table 2. Components of the Data

	Public Library	Public Data
C o m p o n e n t	- Borrow Log : User Key, Book Key, KDC, Title, Author, Publisher, Loan Date, Return Date, Mange Library, Media Format and etc	- Population : Total Population, Population by Gender and Age, Student and Infant Population
	- User : User key, Date of Birth, Gender, Zipcode and etc	- Location Characteristics : Bus Stop, Subway Station, Building Location by Use, Land Value by Public Announcement
	- National Library Standard Data	

1) Focus Group Interview : 특정 분야 경험자 혹은 전문가들과 함께 인터뷰 방식으로 조사하는 방법

2) KDC(한국십진분류법) : 도서관 도서 분류체계로 1차 10개에서 세분화되어 분류됨

3.2 데이터 전처리 및 활용변수

제공된 데이터는 회원의 주소자료가 우편번호로 구성되어 있으며, 개정 전 5자리 우편번호와 개정 후 6자리 우편번호가 혼재하여, 이를 네이버 우편번호 검색기를 활용하여 주소를 변환하였다. 그리고 주소 정보를 공간분석에 활용하기 위해 Python을 활용하여 위경도 좌표로 변환하여, 데이터 전처리를 진행하였다. 또한, 공간분석을 위해 우편번호 기초구역 단위별 다양한 요인들을 수집 후 결합하여 데이터 세트를 구축하였다.

추가적으로 문화적 특성을 고려한 변수 생성을 위해 Fig 1과 같이 행이 우편번호 기초구역, 열이 KDC로 구분된 도서 대출이력 자료(1,495건)를 생성하였으며, 기초구역간 도서 대출수요의 스케일 보정을 위해 정규화(Normalization) 작업을 수행하였다. 그리고 대표적인 머신러닝 알고리즘 중 하나인 K-Means Clustering을 활용하여 기초구역별 유사한 도서대출 이력을 가지는 군집화 분석을 하였다.

MAIN_CLASS	0.00	100.00	200.00	300.00	400.00	500.00	600.00	700.00	800.00	900.00
41000	0.00	0.01	0.00	0.07	0.00	0.02	0.09	0.00	0.99	0.07
41001	0.00	0.20	0.28	0.02	0.06	0.39	0.14	0.04	0.84	0.06
41002	0.04	0.05	0.03	0.22	0.19	0.17	0.05	0.05	0.89	0.27
41003	0.31	0.00	0.00	0.52	0.52	0.00	0.00	0.00	0.52	0.31
41004	0.00	0.00	0.00	0.00	0.71	0.00	0.71	0.00	0.00	0.00

Fig. 1. Dataset for K-means Clustering(Normalization)

분석한 결과 Fig 2와 같이 나타났으며, 군집의 품질을 평가하는 실루엣 계수(Silhouette Coefficient)가 군집이 4개일 때 가장 높게 나타남을 알 수 있다.

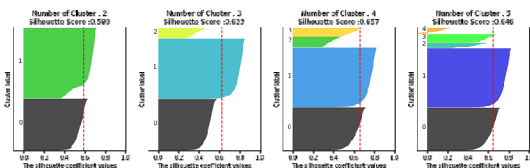


Fig. 2. Silhouette Coefficient by Cluster

4개의 군집으로 분석한 결과 Table 3과 같이 나타나며, 군집 A는 기술과학이 22.4%로 높은 비중을 차지하는 것으로 나타났으며, 군집 B는 문학이 50%, 군집 C는 문학 36.0%, 사회과학 15.7%로 높게 나타났다. 그리고 군집 D는 사회과학 38.4%가 가장 높게 나타나 지역별로 선호하는 도서가 서로 상이함을 확인하였다.

Table 3. Borrow Ratio by Cluster

(unit : %)

No	A (Technology Science)	B (Literature)	C (Literature/ Social Science)	D (Social Science)
General	3.4	2.6	3.4	3.8
Philosophy	8.2	4.1	6.1	6.8
Religion	5.9	1.6	2.2	3.0
Social Science	13.7	10.8	15.7	38.4
Natural Science	9.6	8.6	8.8	4.0
Technology Science	22.4	6.5	8.8	12.0
Art	7.2	3.6	4.8	5.4
Language	4.2	3.9	4.6	3.7
Literature	15.7	50.0	36.0	17.0
History	9.7	8.5	9.8	5.8

마지막으로 기초구역별 도서관 충성고객이 많을수록 도서관 도서 대출수요에 높은 영향이 있을 것으로 판단하여 대출 이력 자료를 활용하여 전체 이용자 중 상위 25% 이상(4사분위)에 해당하는 대출자(7회 이상)를 추출하여 추가 변수를 생성하였다. 그 외에도 선행연구를 참고하여 도서관 도서 대출수요에 영향을 미치는 인구수, 교통수단별 정류장, 공시지가, 건물 등의 자료를 결합하여 분석을 위한 최종 데이터 세트를 구축하였다[4].

3.3 도서관 도서 대출수요 특성 분석

3.3.1 일자별 도서 대출수요 특성

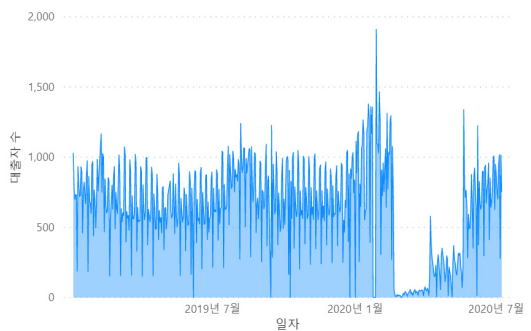


Fig. 3. Time Series Analysis by Date

시립도서관 도서 대출자 수요의 특성을 살펴보면(Fig 3 참조) 평균 대출자 수는 약 1,000명으로 나타났으며, 코로나 19로 인해 2020년 2월 20일부터 대구시 시립도서관 9개 관이 휴관으로 2월 20일부터 대출자 수가 급격

히 감소한 것으로 분석되었으며, 4월 이후 점차 회복세를 보인다. 또한, 선행연구의 분석 결과[14]와 동일하게 통계, 하계 방학 기간에 있어 대출자 수가 증가하는 것으로 분석되었다.

3.3.2 요일별, 시간대별 도서 대출수요 특성

도서관 도서 대출자의 요일별, 시간대별 대출수요 특성을 살펴보면(Fig 4 참조), 월요일의 경우 수요가 가장 낮게 나타났으며, 이는 대부분의 도서관 휴관일이 월요일로 지정된 결과로 판단되며, 월요일을 제외한 평일에서는 유사한 대출수요 분포를 나타내고 있다. 또한, 평일에 비하여 주말의 경우 상대적으로 대출수요가 14~17시에 집중되어 있음을 알 수 있는데, 이는 평일에 비하여 주말의 경우 운영시간이 1시간가량 앞당겨져 17시에 종료되는 것에 따른 것으로 판단된다. 그리고 도서관 운영시간 이외에 집계된 대출자 수는 지하철 역사 내 설치된 무인 대출반납기³⁾와 야간 대출서비스⁴⁾를 활용한 것으로 판단된다.

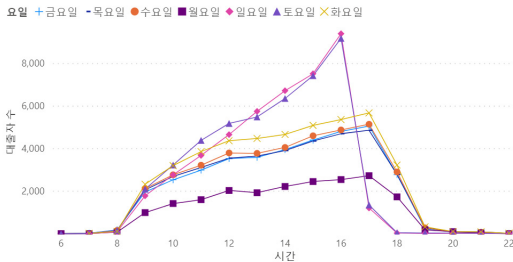


Fig. 4. Time Series Analysis by Day and Time

3.3.3 성별, 연령대별 도서 대출수요 특성

도서관 도서 대출자의 성별, 연령별 특성을 살펴보면 Fig 5와 같이 나타나며, 남성보다 여성의 대출수요가 전반적으로 높게 나타난다. 단, 60대 이상의 연령에서는 남성 대출자 수가 여성보다 높게 나타남을 알 수 있다. 연령대별로는 40대 대출자가 28.9%로 가장 높으며, 다음으로 30대, 20대, 50대 순으로 나타났다. 60세 이상 고령자의 경우 선행연구와 동일하게 대출자 수가 급격히 감소하는 것으로 보인다[5, 14].

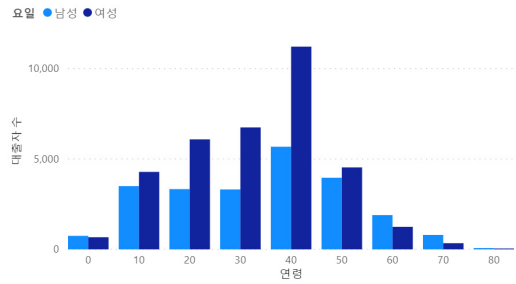


Fig. 5. Demand Analysis by Gender and Age

3.3.4 도서관 접근 거리별 도서 대출수요 특성

도서관 접근 거리별 도서 대출수요 특성을 살펴보면, 도서관과의 거리가 1~2km 이하는 33.4%로 가장 높은 비중을 차지하였으며, 다음으로 1km 내가 31.6% 비율로 나타났다. 또한, 2km 이상의 거리부터는 도서 대출수요가 급격히 감소하였으며, 이는 문화체육부에서 발간하는 공공도서관 건립 운영 매뉴얼 내 시립도서관 봉사권역(2~2.5km)과도 유사한 것으로 확인할 수 있다[15]. 세부적으로 지자체들의 접근거리별 도서 대출수요를 살펴보면 Fig 6과 같이 나타나며, 북구가 상대적으로 봉사권역 내 대출수요가 집중되어 있으며, 달서구의 경우 도서관과 주거지가 상대적으로 이격되어 있음을 추정할 수 있다.

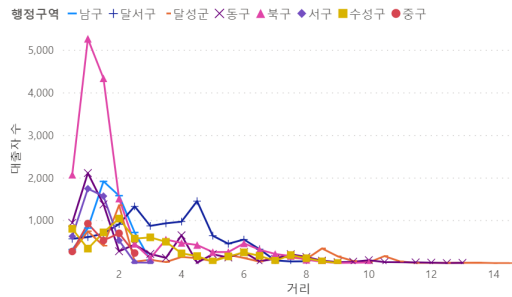


Fig. 6. Demand Analysis by Local Government and Distance

4. 공공도서관 도서 대출수요 모형

4.1 다중회귀 모형 추정

본 연구에서는 도서관 도서 대출자 수에 영향을 미치는 요인을 구체적인 인과관계로 분석하기 위해 기초구역별 인구, 입지특성, 도서대출 이력 기반의 문화적 특성 등

3) 무인 대출반납기 : 지하철 운영시간에 사용 가능한 무인화 대출반납 기기
 4) 야간대출 서비스 : 평일 17시 이전 예약한 자료를 예약 당일 18~22시 대출할 수 있는 제도

의 변수들을 독립변수로 설정하고, 기초구역별 도서관 도서 대출자 수를 종속변수로 설정하여 다중선형회귀모형(multiple linear regression model)을 추정하였다. 다중선형회귀모형의 수학적 표현은 다음과 같다.

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \epsilon_i, i = 1, 2, \dots, n$$

단, y_i = 종속변수

$x_{1i}, x_{2i}, x_{3i}, \dots$ = 독립변수

$\beta_0, \beta_1, \beta_2, \dots$ = 회귀계수

ϵ_i = 오차항

Table 4. Variables in Multiple Regression Model

No	Variable	Description	Format
D e p e n d e n t	Number of Borrowers	The Number of Borrowers of Municipal Library	Continuous
	Total population	Total Population	Continuous
I n d e p e n d e n t	Infant Ratio	Ratio of Infant Population	Continuous
	50s Male Ratio	Ratio of 50s Male Population	Continuous
	50s, 80s Female Ratio	Ratio of 50s, 80s Female Population	Continuous
	Density	Population Density	Continuous
	Area	Area	Continuous
	Land Value	Land Value by Public Announcement	Continuous
	Walking Distance Ratio	Including Library Walking Pass	Continuous
	Bus stop	Number of Bus Stops	Continuous
	Distance	Distance to Nearest Library	Continuous
	Small Library	Number of Small Library	Continuous
	Middle school	Number of Middle School	Continuous
	Cluster B	Cluster based on Library Loan Log(Literature)	Dummy
	Cluster C	Cluster based on Library Loan Log(Literature/Social Science)	Dummy
	Loyal Customer	Number of Users in the Quartile of Loan Logs(Number of Users who Borrowed Books 7 or more times in 1 year and 6 months)	Continuous

다중회귀모형 추정에 활용된 변수는 총 38개이며, 모형의 적합성 및 다중공선성 문제 등을 고려하여 도서관 도서 대출자 수에 유의하지 않는 변수를 변수 증감법(단계적 선택법)을 활용하여 제거 후 모형을 구축하였다. 모형 추정결과 도서관 도서 대출자 수에 영향을 미치는 요인으로는 인구변수에서는 총인구수, 유아 인구비율, 50대 남성 비율, 50대 여성 비율, 80대 여성 비율, 인구밀도가 포함되어 있으며, 입지특성 변수로는 버스 정류장 수, 작은 도서관 수, 최근접 도서관과의 거리, 도보권 비율(도서관 기준 500M), 공시지가, 중학교 수가 포함되었다. 또한, 도서관 도서대출 이력 기반의 문화적 특성은 군집 B(문학), 군집 C(문학, 사회과학), 충성고객 수가 포함되었다.

4.2 다중회귀 모형 분석결과

추정된 모형의 통계적 검정을 하기 위해 R^2 , F-검정, t-검정의 통계적 기법을 활용하였다. 모형의 적합도(goodness of fit)를 살펴보면 R^2 가 0.9574로 분석되어 모형의 설명력은 매우 우수한 것으로 판단된다. 그리고 추정된 다중회귀모형의 전체적인 유의성을 검정하는 F-검정 통계치의 경우 유의확률이 0.000으로 나타나 추정된 모형이 통계적으로 의미가 있는 것으로 나타났다. 본 연구에서 추정된 다중회귀모형을 바탕으로 분석결과를 살펴보면 다음과 같다.

첫째, 인구변수의 경우 총인구수, 유아 인구비율, 여성 50대/80대 인구비율이 통계적으로 유의한 것으로 분석되었으며, 유아 인구비율, 여성 50대/80대 인구비율의 경우 부(-)의 영향을 미치는 것으로 나타났는데, 이는 기초구역별 해당 계층을 제외한 인구비율 증가가 도서관 도서 대출수요 증가에 영향을 미치는 것으로 확인되었다.

둘째, 입지특성 변수의 경우 도보권 비율, 작은 도서관 수, 중학교 수는 통계적으로 유의한 것으로 분석되었으며, 대다수의 선행연구[2-11]와 동일하게 교통 접근성과 학교의 수에 따라 도서관 도서 대출수요는 증가하는 것으로 확인되었다.

셋째, 본 연구의 차별성이라 할 수 있는 문화적 특성의 경우 군집 C(사회과학, 문학), 도서관 충성자 수가 도서 대출수요에 유의한 것으로 분석되었다. 특히 군집 A(기술과학)를 기준으로 군집 C(사회과학, 문학)의 도서가 증가할수록 도서대출 수요가 증가하는 것으로 분석되었으며, 도서관 충성고객의 수도 높은 정(+)의 영향을 미치는 것으로 나타났다.

Table 5. Multiple Regression Model Results

	Variable	Estimate	P-value	VIF
P o p u l a t i o n	Total Population	69.10	0.00	2.82
	Infant Ratio	-9.94	0.00	1.98
	50s Male Ratio	-12.03	0.13	1.20
	50s Female Ratio	-7.13	0.03	1.25
	80s Female Ratio	-12.14	0.05	1.46
	Density	6.24	0.06	2.23
L o c a t i o n	Area	15.67	0.00	1.83
	Land Value	6.58	0.06	1.12
	Walking Distance Ratio	7.45	0.04	1.55
	Bus Stop	-10.89	0.13	1.19
	Distance	10.30	0.13	1.85
	Small Library	-27.78	0.00	1.79
	Middle School	10.38	0.00	1.07
C u l t u r e	Cluster B (Literature)	1.94	0.14	2.51
	Cluster C (Literature/Social Science)	3.37	0.01	2.48
	Loyal Customer	1030.03	0.00	1.45
e t c	Constant	12.96	0.00	-

5. 결론

5.1 분석결과 요약

본 연구에서는 대구광역시 공공도서관 도서 대출수요 분석을 위해 도서관 대출 이력, 회원 정보, 공공데이터 등을 수집하여 우편번호 기초구역 단위로 데이터를 전처리하여 분석하였다. 도서대출 수요 특성을 파악하기 위해 일별, 요일 및 시간대별, 성별 및 연령별, 지자체 및 접근 거리별로 기술통계분석을 진행하였다. 그리고 인구, 교통 접근성, 문화적 특성 등에 따른 도서관 도서대출 수요를 다중회귀모형으로 추정하여 분석하였다. 본 연구의 주요 분석결과를 요약하면 다음과 같다.

첫째, 일자별 도서 대출수요를 살펴보면 선행연구[14]와 동일하게 동계, 하계 방학 기간에 대출자 수가 증가하

는 것으로 보이며, 코로나19로 인한 시립도서관 휴관으로 도서 대출수요 감소 후에 지속적인 대출수요가 회복되는 것으로 나타났다. 그리고 요일별 시간대별 대출수요는 평일에 대비하여 주말 14시에서 17시에 집중되는 것으로 분석되었다.

둘째, 성별, 연령별로 살펴보면 전반적으로 남성에 비해 여성의 도서 대출수요가 높은 것으로 나타났으며, 선행연구[5, 14]와 동일하게 60대 이상의 고령자의 경우 도서 대출자 수가 급감하는 것으로 나타났다. 접근 거리별 대출수요는 대다수가 2km 내 영역에서 도서관 서비스를 이용하는 것으로 보이며, 달서구의 경우 타 지자체에 비해 주거지와 접근 거리가 상대적으로 이격되어 있는 것으로 분석되었다.

셋째, 도서관 도서 대출수요를 다중회귀모형으로 추정 한 결과 유아, 여성 50대/80대 인구를 제외한 인구비율이 증가할 경우 대출수요 증가에 영향을 미치는 것으로 분석되었으며, 대다수 선행연구와 동일하게 접근성과 학교의 수가 도서관 도서 대출수요 증가에 주요 요인임을 확인하였다.

마지막으로, 도서대출 이력자료를 기반으로 기초구역 단위의 문화적 특성을 군집화하여 도서 대출수요와의 관련성을 분석한 결과 사회과학, 문학 분야의 도서 대출과 도서 대출이력이 많은 충성고객이 증가할수록 도서 대출수요가 증가하는 것으로 분석되었다.

5.2 연구결과의 정책적 시사점 및 연구의 한계

본 연구의 분석결과를 바탕으로 정책적 시사점을 제시하면 다음과 같다.

첫째, 도서관 도서 대출수요를 기술통계분석을 활용하여 살펴본 결과 동계, 하계기간에 있어 도서 대출자 수요가 증가하는 것으로 나타났는데, 해당 기간에 특정 계층을 타깃으로 하는 다양한 도서관 문화행사 및 프로그램 개발이 도서관 활성화에 기여할 수 있을 것으로 판단된다.

둘째, 연령별 도서 대출수요가 선행연구와 동일하게 고령자인 60대 이상부터 급격하게 감소하기 시작하는데, 고령화 시대를 맞이하고 있는 시대적 환경을 감안하여, 고령자를 위한 시설물 보완, 프로그램 개발, 마케팅 등이 다양하게 개발되어야 할 것으로 보인다.

셋째, 접근 거리별 대출수요 분석결과 대다수가 2km 범위내에서 도서관 이용이 이루어짐을 알 수 있었으며, 달서구의 경우 상대적으로 시립도서관을 이용에 있어 주거지와 거리가 이격됨을 알 수 있다. 이러한 지자체별 격차를 도서관 설립 시 고려하여 반영한다면 지역 간 도

서관 이용 격차 해소에 도움이 될 수 있을 것으로 보인다.

넷째, 인구특성 변수 중 유아, 여성 50대/80대 인구비율이 증가할 경우 도서 대출수요의 부(-)의 관계가 나타났는데, 이는 역설적으로 해당 계층의 도서관 도서대출 수요가 낮은 것을 의미하는 것으로 해당 계층을 대상으로 하는 다양한 프로그램 개발 및 지원이 필요할 것으로 보인다.

마지막으로 도서대출 이력을 기반으로 문화적 특성 변수와의 관련성을 분석한 결과 기술과학 도서에 비하여 문학, 사회과학 도서의 증가와 충성고객 수 증가가 도서관 도서대출 수요 증가에 관련성이 있는 것으로 분석되었는데, 지역별 도서 대출 이력 특성이 상이함을 감안하여 차별화된 마케팅 접근 및 프로그램 개발로 도서관 이용 활성화에 기여할 수 있을 것으로 판단된다.

본 연구의 학술적 함의는 다음과 같다. 공공도서관 도서 대출수요에 영향을 미치는 다양한 요인들의 분석을 수행하였으며, 선행연구와 동일한 요인이라도 대구광역시라는 공간적인 범위와 특정 기간에서의 미치는 영향정도를 정량화하여 분석하였다. 그리고 문화적 특성의 변수를 반영하여 다중회귀모형으로 도서대출 수요와의 관련성을 확인하였다. 또한, 지역별 충성고객의 수가 공공도서관 도서 대출수요의 주요한 변수로 나타났다. 이러한 연구결과는 문화적 특성이 시간과 장소, 범위에 따라 지속적으로 변화함을 감안하면 관련된 후속연구의 기초자료로 활용될 수 있을 것으로 보인다.

한편, 본 연구의 한계는 다음과 같다.

첫째, 회원 정보의 주소지 자료의 경우 우편번호로 입력되어 있어, 우편번호 기초구역 범위의 중심점을 좌표로 변환하여 분석하였다. 따라서 추후 회원별 정확한 주소지 정보수집이 이루어진다면 공간적인 측면에서 고도화된 분석이 가능할 것으로 판단된다.

둘째, 도서관 도서 대출자 수를 종속변수로 하여 분석하였는데, 도서관을 실제 이용하는 이용자의 경우 대출뿐만 아니라 문화프로그램, 평생교육, 시청각실 대여 등의 다양한 문화생활을 위해 방문하는 이용객이 있음을 고려하면, 분석의 범위가 다소 제한적으로 수행된 것으로 판단된다. 추후 도서관에서 발생하는 다양한 문화프로그램 등의 사용내역 자료와 결합하여 분석한다면 다양한 측면에서 분석이 가능할 것으로 보인다.

셋째, 자료 수집의 한계로 공간적 범위를 시립도서관으로 한정하여 도서관 도서 대출수요 분석하였는데, 작은 도서관 및 구·군립도서관 대출 이력 자료와 회원 정보 등이 추후 수집된다면 종합적인 도서관 도서 대출수요

분석이 가능할 것으로 판단된다.

마지막으로 본 연구에서 추정된 다중회귀모형은 선행 연구에서 진행되지 않은 대출 이력 기반의 문화적 특성을 변수를 포함하여 도서 대출수요에 미치는 관련성을 분석하는 데 초점을 두었으므로 도서관 도서 대출수요를 정확히 예측하는 모형으로는 다소 한계가 있다. 따라서 최근 인공지능 분야에서 각광받고 있는 딥러닝, 머신러닝 기반의 예측모형을 개발하는 연구가 향후 필요할 것으로 판단된다.

REFERENCES

- [1] Y. S. Kim. (2007). A Study on the Effective Enhancement of the Public Library Infrastructure in Korea. *Journal of Korean Library and Information Science Society*, 38(1), 29-48.
DOI : 10.16981/kliss.38.1.200703.29
- [2] I. Kim & J. K. Park. (2016). An Analysis of Geographical Accessibility to Public Libraries in Metropolitan Area and Related Implications: The Case of Busan. *The Korean Journal of Local Government Studies*, 20(3), 191-221.
DOI : 10.20484/klog.20.3.10
- [3] Y. Y. Kim & K. H. Yeo. (2014). Location Evaluation and Optimal Location Selection of Public Library Using a Local Space Search Approach - Focusing on the Case of Seoul City -. *Journal of The Residential Environment Institute of Korea*, 12(4), 55-71.
DOI : 10.20484/klog.20.3.10
- [4] J. J. Won & K. H. Ahn. (2010). The Effect of Locational and Facility Characteristics on Public Library Use - Focused on Public Libraries in Seoul -. *Journal of the Architectural Institute of Korea*, 26(2), 79-86.
- [5] S. Heo & Y. K. Chung. (2014). A Study on Public Libraries' User Behaviors Based upon Circulation Data in Gangseo and Yangcheon Region. *Journal of the Korean Biblia Society for Library and Information Science*, 25(4), 187-207.
DOI : 10.14699/KBIBLIA.2014.25.4.187.
- [6] H. J. Kim, J. G. Lee & K. H. Yeo. (2015). Regional Disparities and Determinants of Spatial Accessibility of Public Libraries in Seoul. *The Seoul Institute*, 16(4), 109-127
- [7] B. C. Park & S. W. Seo. (2015). Service Effect and Propriety Area Analysis of Public Library in Incheon Metropolitan City. *Incheon Development Institute. Urban Research*, (9), 217-241
- [8] J. W. Yoo & S. Y. Kim. (2015). A Study on the Assessment of Location Environment of a Regional Central Library - The Case of Daegu Metropolitan City -. *Journal of Korean Library and Information*

Science Society, 46(4), 427-450.
DOI : kliss.46.201512.427

- [9] J. W. Yoo, B. J. Ko & D. H. Chang. (2015). A Research on the Site Selection of a Metropolitan Library : The Case of City of Busan. *Journal of the Korean Society for Library and Information Science*, 49(3), 291-309.
DOI : 10.4275/KSLIS.2015.49.3.291
- [10] C. H. Yu, W. M. Lee, S. H. Park & S.S.Shin. (2017). A Study on the Appropriate Location of Public Facilities. *Journal of The Korean Cadastre Information Association*, 19(3), 151-162.
DOI : 10.46416/JKCIA.2017.12.19.3.151
- [11] S. J. Yim & C. H. Kim. (2020). Analysis of Selecting an Appropriate Location for Infrastructure Considering the Range of Service Area and Accessibility : Focusing on Public Libraries. *Journal of the Korean Regional Development Association*, 32(3), 45-66.
DOI : 10.46416/JKCIA.2017.12.19.3.151
- [12] F. J. Zakrajsek & V. Vodeb. (2020). Agent-based geographical modeling of public library locations. *Library & Information Science Research*, 42(2).
DOI : 10.1016/j.lisr.2020.101013
- [13] N. S. Saman, V. Ahmadsreza, K. F. Mohammad, N. Fatemeh & N. S. Najme. (2020). Spatial modeling of areas suitable for public libraries construction by integration of GIS and multi-attribute decision making: Case study Tehran, Iran. *Library & Information Science Research*, 42(2).
DOI : 10.1016/j.lisr.2020.101013
- [14] T. Y. Kim, J. Y. Baek & H. J. Oh. (2018). An Analysis of Library User and Circulation Status based on Bigdata Logs A Case Study of National Library of Korea, Sejong. *Journal of Korean Library and Information Science Society*, 49(2), 357-388.
DOI : 10.16981/kliss.49.201806.357
- [15] Ministry of Culture, Sports and Tourism. (2019). *A Manual on Construction and Management for Public Library*.

오 민 기(Min-Ki Oh)

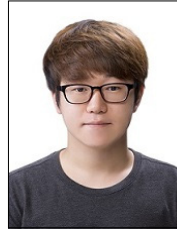
[학생회원]



- 2016년 3월 ~ 현재 : 경북대학교 문헌정보학과
- 2020년 7월 ~ 현재 : 대구디지털산업진흥원 인턴
- 관심분야 : 빅데이터, 머신러닝, 텍스트마이닝, NLP
- E-Mail : dmg04023@gmail.com

김 경 래(Kyung-Rae Kim)

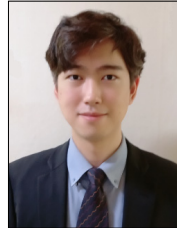
[학생회원]



- 2015년 3월 ~ 현재 : 경북대학교 컴퓨터학부
- 2020년 7월 ~ 현재 : 대구디지털산업진흥원 인턴
- 관심분야 : 빅데이터, 인공지능, 데이터마이닝, 자연어 처리
- E-Mail : frt1145@gmail.com

정 원 웅(Won-Oong Jeong)

[학생회원]



- 2015년 3월 ~ 현재 : 경북대학교 컴퓨터학부
- 2020년 7월 ~ 현재 : 대구디지털산업진흥원 인턴
- 관심분야 : 빅데이터, 인공지능, 데이터마이닝, 입지 분석
- E-Mail : wonoong27@gmail.com

김 건 욱(Keun-Wook Kim)

[정회원]



- 2009년 2월 : 영남대학교 도시공학(공학사)
- 2011년 8월 : 아주대학교 교통공학(공학석사-교통빅데이터)
- 2019년 7월 ~ 현재 : 대구디지털산업진흥원 빅데이터활용센터 센터장
- 관심분야 : 도시데이터분석, 빅데이터, 인공지능, 텍스트마이닝
- E-Mail : aut7767@dip.or.kr