

# 예비교사의 데이터 리터러시 역량 증진을 위한 빅데이터 분석 교양강좌의 개발 및 적용

김슬기 · 김태영  
한국교원대학교

## 요약

최근, 급격히 발전하는 미래 디지털 사회를 살아갈 학생들의 디지털 리터러시와 데이터 리터러시 관련 기초소양 교육이 강조되고 있다. 이에 일반 대학과 교육 대학에서도 기초소양으로서 빅데이터 및 데이터 리터러시 향상을 위한 교육의 수요가 많아지고 있다. 이에 본 연구는 예비교사를 위한 빅데이터 분석 교양강좌를 설계 및 적용하고 데이터 리터러시에 미치는 영향을 분석하였다. 투입 프로그램에 대한 흥미도와 이해도 분석 결과, 예비교사의 수준에 적절한 형태임을 확인했으며, 데이터 리터러시의 '지식', '기능', '가치와 태도'의 모든 영역에서 유의미한 역량의 향상이 있는 것을 확인하였다. 본 연구의 결과가 체계적인 데이터 리터러시 관련 교육 연구에 도움을 주어 학생과 예비교사들의 데이터 리터러시를 증진하는데 이바지할 수 있기를 기대한다.

키워드 : SW/AI교육, 빅데이터, 데이터 리터러시, 기초소양, 예비교사

## The Development and Application of the Big Data Analysis Course for the Improvement of the Data Literacy Competency of Teacher Training College Students

Seulki Kim · Taeyoung Kim  
Korea National University of Education

## Abstract

Recently, basic literacy education related to digital literacy and data literacy has been emphasized for students who will live in a rapidly developing future digital society. Accordingly, demand for education to improve big data and data literacy is also increasing in general universities and universities of education as basic knowledge. Therefore, this study designed and applied big data analysis courses for pre-service teachers and analyzed the impact on data literacy. As a result of analyzing the interest and understanding of the input program, it was confirmed that it was an appropriate form for the level of pre-service teachers, and there was a significant improvement in competencies in all areas of 'knowledge', 'skills', and 'values and attitudes' of data literacy. It is hoped that the results of this study will contribute to enhancing the data literacy of students and pre-served teachers by helping with systematic data literacy educational research.

Keywords : SW/AI Education, Big Data, Data Literacy, Basic Knowledge, Pre-service Teacher

---

교신저자 : 김태영(한국교원대학교 컴퓨터교육과)

논문투고 : 2022-03-22

논문심사 : 2022-04-04

심사완료 : 2022-04-06

## 1. 서론

인공지능 기술을 중심으로 한 디지털 기술이 사회에 미치는 영향이 커짐에 따라 미래 사회를 살아갈 학생들이 길러야 할 핵심역량으로 디지털 기초소양의 중요성이 강조되고 있다. 이에 세계 각국에서는 디지털 기초소양을 길러주는 교육을 위해 정책적으로 많은 노력을 기울이고 있다[13][16][20].

대한민국에서도 ‘인공지능(AI) 국가전략’ 발표를 통해 국가 경쟁력을 높일 수 있는 사회 및 산업, 교육 관련 인공지능 정책의 전반적인 방향을 제시하였다. 또한, ‘전국민 AI·SW교육 확산 방안’ 통해 AI 시대를 살아가기 위한 핵심역량으로 컴퓨팅 사고력을 기반으로 한 데이터 리터러시, 문제해결력, 변화 적응력, 공익적 사고력 등을 제시하였다. 이러한 흐름을 반영하여, 공교육의 방향성을 제시하는 ‘2022 개정 교육과정을 위한 총론 주요 사항’ 발표에서도 디지털 기초소양을 학생들의 필수 역량으로 강조하였다[18][19][25].

이처럼 필수 역량으로 강조되고 있는 학생들의 디지털 소양을 길러주기 위해서는 인공지능을 중심으로 한 디지털 사회를 이해하고 다방면에서 관련 기술을 활용할 수 있는 역량을 길러주어야 한다. 역사적으로 볼 때, 디지털 사회와 인공지능 기술은 데이터와 밀접한 관련이 있으며 필연적으로 많은 수의 데이터를 활용하기 때문에 디지털과 데이터에 대한 기본 소양이 함께 강조되어야 한다고 볼 수 있다. 이는 데이터에 대한 기본 소양인 데이터 리터러시 교육이 필수적으로 함께 이루어져야 함을 의미하며, 많은 교육과정 연구에서 기초 연구로 활용하는 ‘OECD Future of Education and Skills 2030’에서 디지털 리터러시와 데이터 리터러시를 함께 강조하고 있는 것과 같은 맥락이라고 볼 수 있다[27].

이러한 흐름에 발맞춰 AI와 데이터 관련 대학 교양 수업의 수요가 증가하고 있으며, 다양한 전공의 대학교육 대상자를 위한 교육적 접근이 진행되고 있다. 또한, 미래의 초·중등 학생들의 데이터 리터러시와 디지털 리터러시 교육을 제공하게 될, 교사 양성 교육 기관에서도 관련 전공의 심화 교육을 통한 수직적인 교육뿐만 아니라 누구나 알고 활용할 수 있는 보편교육으로서의 AI와 데이터 관련 강좌들의 요구들이 많아지고 있다[10].

교사 양성 교육 기관의 교양강좌의 경우, 교육 대상

자들이 미래의 교사이며, 다양한 교양 지식을 교육적으로 활용하고자 함이 목적이라고 할 수 있다. 특히 예비 교사들의 데이터 리터러시가 함양된다면, 미래의 학생들에게 다양한 관점에서의 데이터 리터러시를 길러 줄 수 있는 교육을 제공할 수 있어 그 파급력이 클 것이다. 이에 본 연구는 예비교사를 위한 빅데이터 분석 교양강좌를 15주에 걸쳐 운영하고 데이터 리터러시에 미치는 영향을 연구하여 시사점을 도출하고자 한다.

## 2. 연구 배경

### 2.1. 데이터 리터러시의 정의

OECD(2019)는 미래 사회를 살아갈 학생들이 갖추어야 할 역량을 증진 시키기 위한 교육의 비전과 학습 프레임을 ‘OECD Learning Compass 2030’을 통해 발표하였다. 그 중, 전체 교육과정을 학습하기 위한 전제 조건으로 ‘핵심기반(Core Foundation)’을 제시하고 그 하위 요소 중 하나인 수 문해력을 수를 활용하는 능력에서 확장하여 데이터 사용능력, 디지털 사용능력까지 포함하는 의미로 해석하고 강조하였다[27].

데이터 리터러시는 관련 연구를 진행한 연구자와 기관에 따라 상세 정의와 범주, 내용 요소가 다양하게 나타난다. 해외의 선행 연구를 먼저 살펴보면 Shield(2005)는 데이터 리터러시를 ‘데이터를 정의하고 통계적으로 활용하는 능력’으로 정의하며, 통계 리터러시와 정보 리터러시, 데이터 리터러시는 서로 상호 보완적이며, 비판적 사고로 확장되어야 한다고 주장하였다[32].

Carlson 외(2011)는 데이터 리터러시를 ‘주어진 데이터로부터 의미 있는 정보를 도출하는 능력’이며, 읽고 쓰고 활용하고, 분석을 통해 해석하여 결론을 도출하는 능력’으로, 기술적 측면과 함께 사회적 측면까지 포함하여 정의하였다[3].

Mandinach 외(2013)는 ‘데이터의 활용적 측면에서 의사 결정에 정보를 제공하기 위해 데이터를 이해하고 효과적으로 사용하는 것’으로 정의하고 교사를 위한 데이터 리터러시를 강조하였다[23].

D’Ignazio 외(2015)는 ‘데이터를 읽고 일하고 분석하며 논증하는 능력’으로 정의하였으며, 이를 확장하여 빅데이터 리터러시를 데이터 리터러시에 ‘식별’, ‘이해’, ‘평

가'라는 능력을 추가하여 정의하였다[4].

데이터 리터러시 관련 국내 연구는 주로 상기 언급된 해외 연구를 활용하여 정의하고 있으며 한상우(2018)는 인문학 기반의 데이터 리터러시를 정의하였다. 그 영역으로 데이터의 이해를 포함하는 공통 영역과 추론 및 소통을 중심으로 한 인문학 영역, 데이터 처리의 기술적 측면을 중심으로 한 디지털 기술 영역으로 구분하고 주요 개념을 제시하였다[5].

배화순(2019)은 '데이터로부터 의미 있는 정보를 추출하고 실생활의 다양한 문제를 해결하기 위해 데이터를 활용하며, 적절한 도구를 활용해 데이터를 분석하고 결론을 도출하는 능력에 타인과 효과적으로 의사소통하는 능력'으로 정의하였다[1].

최근 SW/AI 교육이 강조됨에 따라 컴퓨팅 사고력을 기반으로 한 정보교육 관점에서 데이터 리터러시의 정의를 연구한 결과들이 도출되고 있다. 김슬기 외(2021)는 데이터 리터러시 관련 해외 선행 연구를 통합하고 자연어 처리 딥러닝 방식으로 분석하여 정보교육의 관점에서 인공지능 교육을 위한 데이터 리터러시를 <Table 1>과 같이 정의하였다. 특히 데이터 리터러시의 정의와 함께 '지식', '기능', '가치와 태도'의 영역으로 역량을 나누어 내용 요소를 구체적으로 제시하였다[14]. 이에 본 연구는 데이터 리터러시의 정의와 영역을 김슬기 외(2021)의 연구 결과를 활용하여 영역 수준에서 심층적으로 분석하고자 한다.

<Table 1> Definition and Components of Data Literacy

Definition		
Basic ability of knowledge construction and communication to collect, analyze, and use data and process it as information for problem solving		
Knowledge	Skills	Values and Attitudes
· Concept of Data		
· Role of Data	· Collect	
· Type of Data	· Interpret	· Relationship
· Structure of Data	· Reason	· Integration
· Source of Data	· Summarize	· Continue
· Analysis Method of Data	· Presentation	· Human-Centered
· Data Visualization	· Manage	· Fairness
· Data	· Process	
Pre-processing	· Communicate	

## 2.2. 기초소양 교육으로서의 빅데이터 및 데이터 리터러시 교육

서론에서 언급한 바와 같이 AI와 데이터 관련 대학 교양 수업들의 수요가 증가하고 보편교육을 위한 접근과 함께 기초소양 교육으로서의 빅데이터 및 데이터 리터러시 교육 연구들이 이루어지고 있다.

박윤수 외(2020)는 보편적 빅데이터와 빅데이터 교육의 방향성을 전문가 인식 조사를 통해 조사하고 빅데이터 처리 시스템 자체에 대한 교육보다는 응용 사례와 분석 방법 위주의 교육이 필요하며, 전공자, 비전공자를 구분하지 않고 보편적으로 교육할 수 있는 표준화된 모델이 필요함을 강조하였다[29].

오경선 외(2020)는 보편적인 AI 교육을 위한 빅데이터 교육의 상호 보완적인 측면을 강조하여 빅데이터 기반 AI 기초 교양 수업을 학부생에게 적용하고 자기 주도성 및 자신감의 정의적 태도에 긍정적인 영향을 주는 것을 확인하였다[28].

김혜영(2020)은 비관적 사고를 포함한 21세기 사회가 요구하는 필수 역량은 데이터 리터러시와 매우 밀접한 관련이 있으며, 이를 기르기 위한 교육의 필요성을 강조하였다. 특히, 이공계 및 인문계 전공과 관련 없이 실용적이고 구체적이며 흥미로운 데이터 리터러시에 대한 접근법과 교육이 필요하다고 제안하였다[11].

전수진 외(2021)는 초·중등 학생에게 교육을 제공하게 될 예비교사들의 AI 기초교양 교과목의 필요성을 강조하며 경험 학습 기반으로 교육을 설계하였다. 특히 교육 영역 설정을 위한 전문가 델파이 설문에서 데이터 과학이 가장 높은 중요도를 보였으며, 데이터 과학이 포함된 'AI 이해' 영역의 평균 중요도가 가장 높은 것으로 나타났다. 이를 반영하여 예비교사들의 AI 소양 및 기본 역량과 윤리의식을 높일 수 있으며, 다양한 교과와 연계하여 지도할 수 있는 교육 내용을 설계하였다[10].

예비교사를 위한 기초교양 교육으로서의 데이터 리터러시 관련 해외 연구는 주로 학생들의 데이터 리터러시를 길러주기 위한 교사의 역량 증진 목적보다는 교사가 데이터 기반으로 의사 결정할 수 있는 역량을 기르는데, 초점을 맞추고 있다.

대표적인 연구로 Mandinach 외(2013, 2016)의 연구를 살펴보면, 교사의 데이터 리터러시 구축을 위한 시스템

적 접근법을 제안하면서 데이터 리터러시를 기반으로 한 의사 결정 역량을 길러야 하며, 예비교사 교육부터 교사생활의 전반에 걸쳐 체계적인 데이터 리터러시 교육을 받아야 함을 주장하였다. 특히 예비교사를 위한 데이터 리터러시 교육은 빠르게 도입할수록 좋으며, 기초 소양 독립 교과목으로서 가르쳐야 한다고 주장하였다 [23][24].

### 2.3. 데이터 리터러시 학습자를 위한 도구

데이터 리터러시 혹은 데이터 관련 교육 활동과 이를 위한 도구의 적합성 분석은 주로 초·중등 학생의 교육을 위한 목적에서의 논의가 활발히 이루어졌다. 관련 연구를 분석하면 초등 및 중학교에서는 주로 교육용 프로그래밍 언어인 엔트리와 스크래치를 활용하며, 고등학교 이상 교육에서는 더 심화 된 데이터 분석이 가능한 Orange3를 활용하기도 한다. 이들 도구의 공통점은 그래픽 기반의 프로그래밍 방식을 활용하는 것이며, 학생들의 접근성을 높이기 위한 도구의 선택이라고 볼 수 있다. 대학교 이상의 성인 교육을 위한 도구는 학습자의 수준에 따라 다양하게 나타나며, 데이터 분석 라이브러리를 활용할 수 있는 파이썬 또는, 접근성을 높이기 위해 블록형 프로그래밍 도구를 사용하기도 한다. 선행 연구에서 데이터 분석 교육에 활용한 도구를 대상에 따라 정리하면 <Table 2>와 같다[7][8][9][12][15][17][22]

<Table 2> Data Analysis Education Tools According to Target

Researcher	Target	Educational Tool
Kim, Y. (2018)	Elementary School	Excel, App Inventor
	University	Scratch
Hur, K. (2020)	Elementary School	Entry
Koo, D., Kim, D. (2020)	Elementary School	Excel, Spreadsheet
	Middle School	Python
Hong, J. (2020)	Elementary School	Entry
Kim, J., Son, Y. (2021)	University	Python
Hur, K. (2021)	Adult	Spreadsheet
Lee, Y. et. al (2021)	High School	Orange3

데이터 리터러시 교육 활동 및 도구에 대한 해외 연구 사례를 살펴보면, Bhargava 외(2015)는 기존의 연구들이 학습자에게 적합한 도구에 대해 논의하지 않은 한계를 지적하였다. 이를 바탕으로 현재 교육에 활용되고 있는 데이터 분석 도구들을 학습 난이도와 유연성을 축으로 하여 (Fig. 1)과 같이 특징을 분석하였다[2].



(Fig. 1) Characterization of Data Analysis Tools

D'Ignazio 외(2015)는 학습자에게 효과적인 활동에 대한 연구를 통해, 데이터 리터러시에서 개념이 확장된 빅데이터 리터러시 관련 교육은 알고리즘 시뮬레이션과 시각화하는 활동이 효과적이며, 학습자 중심의 도구가 필요함을 강조하였다. 학습자 중심의 도구는 스스로 탐색하고 운영할 수 있고, 소규모 데이터가 아닌 빅데이터를 활용하여 시각화할 수 있어야 한다고 주장하였다[4].

상기 연구들을 통해 빅데이터 혹은 데이터 리터러시 교육은 최근 대학교육 기관의 기초소양 교육으로서 강조되고 관련 연구들이 진행되고 있음을 알 수 있다. 하지만, 미래의 초·중등 교육을 위한 예비교사 대상의 기초소양 교육 연구와 교육 대상자와 교육 내용을 고려한 데이터 분석 도구를 선정하고 활용하는 교육 효과성 연구가 부족함을 알 수 있다. 이에 본 연구는 대학생 이상의 성인 교양 교육에 많이 활용되며, 유연성과 확장성이 높은 파이썬을 활용하여 빅데이터 분석 교양강좌를 예비교사들에게 적용하고 데이터 리터러시에 미치는 영향을 심층적으로 분석하고자 한다.

3. 연구방법

3.1. 연구대상

본 연구는 교원 양성 대학의 15주, 주별 3시간 총 45시간으로 구성된 교양과목인 ‘컴퓨터의 이해와 활용’ 수강생 44명을 대상으로 진행되었다. ‘컴퓨터의 이해와 활용’ 강좌의 경우 모든 전공과 학년에서 수강 신청 가능하다. 예비교사들의 선행 경험을 조사하기 위해 프로그래밍 경험을 ‘전혀 경험 없음’, ‘블록형 프로그래밍 경험 있음’, ‘텍스트 프로그래밍 경험 있음’으로 나누어 응답할 수 있도록 하였다. 연구 참여자에 대한 상세 사항은 <Table 3>과 같다.

<Table 3> Details of the Research Participants

Factor	Answer	Factor	Answer
Sex	Man 20	Programming experience	No experience 23
	Woman 24		Block programming 16
Grade	1 grade 26	Major Education	Text programming 5
	2 grade 13		Science & Engineering 11
	3 grade 3		Humanities 29
	4 grade 2		Arts & Physical 4

3.2. 연구 설계

Swan 외(2009)는 데이터 리터러시 측정에 있어 학생들의 기술 수준은 다양하고 일관성이 없기 때문에 사전 조사를 통해 출발점을 파악하는 것이 중요하다고 주장하였다[34]. 또한, Hattwig 외(2013)는 일회성의 평가가 아닌 반복 평가가 필요함을 제시하였다[6].

이에 본 연구는 연구 참여자에 대한 투입 프로그램의 적절성을 확인하기 위해 프로그램이 투입되는 매주 주요 내용에 대한 이해도와 흥미도를 측정한다. 또한, 프로그램이 예비교사의 데이터 리터러시에 미치는 영향을 분석하기 위해 <Table 4>와 같이 데이터 리터러시 검사 도구로 사전 검사를 진행하고 독립변수를 투입한 후 사후 검사를 진행함으로써 단일 집단 사전-사후 검사를 설계하였다.

<Table 4> Pretest-Posttest Design

O1	X1	O2
O1 : Pretest(Data Literacy)		
X1 : Big data analysis class using Python		
O2 : Posttest(Data Literacy)		

독립변수인 빅데이터 분석 교양강좌는 ‘파이썬의 기초’와 ‘빅데이터 분석’으로 나누어 구성하였다. ‘파이썬의 기초’는 접근성을 높이기 위해 웹 개발 환경을 활용하여 진행하였으며, 파이썬 터틀(Turtle) 및 소수의 데이터 활용과 처리를 매주 핵심 내용에 맞게 제시하였다. ‘빅데이터 분석’은 공공 데이터를 제공하고 시각화와 데이터 해석 및 분석을 중심으로 구성하였다. 각 주차 별 주요 강의 주제와 내용은 <Table 5>와 같다.

<Table 5> The Main Contents of the Class

No	Topic	Main Contents
1	OT	Lecture OT and curriculum guidance
2	Background Knowledge	The history of computers and Python
3	I/O and Operations	I/O, arithmetic, relation, logic, operations
4	Control Structure	Sequence, repetition, selection structure
5	Variable	Meaning and utilization of variables
	Function	Structure of functions
6	Data Type	Data type and print format.
7	Project Test	Python Turtle Project
8	List	List characteristics and operations
9	Dictionary	Dictionary characteristics and operations
	Tuple	Characteristics of Tuple
10	Input, Save, Output Data	Understanding datasets, file input/output, and data reading
11	Data Manipulation	Save and search, process data
12	Data	Data analysis using Matplotlib and plot(), bar()
13	Visualization / Analysis	Data analysis using hist(), boxplot()
14		Data analysis using pie(), and scatter()
15	Project Test	Big data analysis project

특히, 예비교사의 특수성을 살리기 위해 다양한 교육 전공과 관련된 주제의 데이터를 제시하고, 분석 활동을 통해 교육의 소재로 활용할 수 있도록 하였다. 활용된 데이터셋과 교육의 상세 예시는 <Table 6>과 같다.

<Table 6> Big Data Analysis Lesson Example

Lesson Composition	Contest
Topic	Comparison of students and teachers at public and private schools during the enlightenment period
Related Major	History, Society Education
Data	Number of students and teachers at public and private schools during the enlightenment period ( <a href="http://db.history.go.kr/id/mh_002">http://db.history.go.kr/id/mh_002</a> )
Main Concept	int, list, for, matplotlib bar()
Activity	Search for data and check the data type
	Data save to list and load
	Use bar() to represent in a bar graph Data analysis and finding implications

상기 예시에서는 역사 및 사회와 관련된 데이터를 이용하였다. 학생들이 주제와 관련된 데이터를 찾고 리스트 형태로 저장한 뒤 효율적으로 데이터를 분석할 수 있는 시각화 방법을 논의한다. 이후, 다양한 그래프를 그릴 수 있는 파이썬의 맷플롯립(Matplotlib) 라이브러리를 이용하여 시각화하고 시각화된 데이터를 바탕으로 분석하여, 시사점을 찾아 나누는 시간을 갖도록 한다.

### 3.3. 검사 도구

본 연구는 프로그램이 투입되는 <Table 3>의 각 주차 별 핵심 내용에 대한 이해도와 흥미도 자기 평가를 진행하였다. 자기 평가 결과는 ‘매우 높음’, ‘높음’, ‘보통’, ‘낮음’, ‘매우 낮음’의 5단계 평가 척도로 5점에서 1점으로 제출할 수 있도록 하였으며, 학생들의 긍정적인 자기 평가와 부정적인 자기 평가의 비율 및 프로그램 투입 기간 동안의 변화로 프로그램의 적절성을 분석한다.

또한, 데이터 리터러시를 역량으로 접근하여 세부적인 구성 요소를 제시한, 김슬기 외(2021)의 정의를 중심으로 데이터 리터러시를 ‘지식’, ‘기능’, ‘가치와 태도’의 영역으로 나누어 측정하고자 한다[14]. 연구 참여자의 데이터 리터러시 측정을 위해, 이현숙 외(2019)에 의해

개발된 ‘2018 국가 수준 디지털 리터러시 검사 성취기준’에서 데이터 관련 문항을 도출하고[21], 송유경(2021)의 ‘데이터 리터러시 향상을 위한 데이터 기반 토론 수업 모형 연구’의 검사 도구[33]와 Qlik(2018)의 ‘Data literacy Workforce Assessment’를 재구성하였다[31]. 재구성된 검사 도구에 대한 응답은 ‘매우 그렇다’, ‘그렇다’, ‘보통이다’, ‘그렇지 않다’, ‘전혀 그렇지 않다’의 5단계 평가 척도로 하고 5점에서 1점으로 제출할 수 있도록 하였다.

검사 도구의 19개 문항에 대한 타당성 확보를 위해 교사 양성 교육 기관에서 교육 중인 컴퓨터 공학 전공 교수 2명 및 교육경력 10년 이상의 컴퓨터 교육 전공 박사 4인, 석사 4인의 검토 후, 정량적인 평가를 위해 CVI(Content Validity Index) 값을 활용하였다. CVI 값은 설문에 대한 응답을 4단계로 나누어 1~2는 동의하지 않음, 3~4는 동의함으로 결과를 수합하고 동의한 인원수의 비율로 평가한다[30]. 또한, 검사 문항을 데이터 리터러시의 영역에 따른 핵심 내재 요인으로 그룹화하고자 공통 요인분석을 진행하고 KMO와 Batlett 검증을 이용하여 유의성을 확인하였다. 또한, 영역별로 그룹화된 변수의 내적 일치도 측정을 위해 Cronbach- $\alpha$  계수를 이용하였다[26]. 그 결과는 <Table 7>과 같다.

<Table 7> Validity and Reliability of Data Literacy Measurement Tool

No	Factor			CVI	Cronbach- $\alpha$
	Skills	Value and Attitude	Knowle-dge		
10	0.931	0.057	-0.067	1.00	0.958
14	0.876	0.068	0.027	1.00	
9	0.860	-0.121	0.187	1.00	
7	0.854	-0.076	-0.064	0.80	
8	0.828	-0.008	0.148	1.00	
11	0.803	0.094	0.056	1.00	
12	0.799	0.036	0.150	1.00	
13	0.780	0.142	-0.024	1.00	
16	0.167	0.862	-0.083	1.00	
15	0.121	0.858	0.036	1.00	
18	-0.092	0.842	-0.052	0.80	0.900
19	-0.124	0.765	0.218	1.00	
17	0.339	0.639	-0.025	1.00	
6	-0.097	0.052	0.958	0.90	0.960
2	-0.033	0.060	0.937	1.00	
4	-0.025	0.062	0.936	0.90	

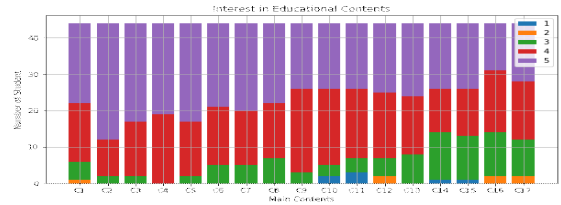
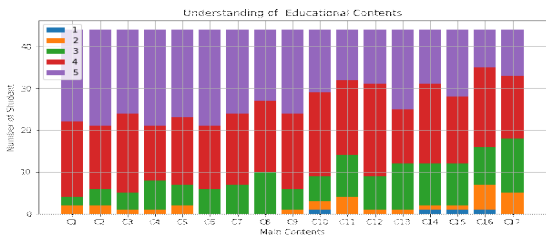
3	0.157	-0.083	0.859	0.80
1	0.097	0.042	0.837	1.00
5	0.188	-0.106	0.797	1.00
KMO(Kaiser-Meyer-Olkin)				0.843
Bartlett's Test of Sphericity				Chi-Square 999.823
				Df 171
				Significance Probability 0.000

CVI 값을 통한 타당도 분석 결과를 먼저 살펴보면, 검토에 참여한 전문가의 수가 10명일 때, CVI 값은 0.78 이상인 경우, 내용 타당도가 확보되었다고 볼 수 있다. 이에, 본 연구를 위해 구성된 검사 도구의 모든 문항에서 타당도가 확보되었다고 해석할 수 있다[30]. 비교적 타당도가 낮은 3번과 7번 문항의 경우 전문가의 의견을 반영하여 명확한 문장으로 서술하였다. 또한, 공통 요인 분석을 진행하고 유의성 검증을 위해 도출한 KMO 값은 0.843으로 사회과학 분야 연구에서 요구하는 0.5 초과 기준을 만족한다. Bartlett 검정을 통한 유의 확률 또한, 기준이 되는 0.05 미만을 만족하여 검사 도구는 데이터 리터러시의 영역에 따른 핵심 내재 요인으로 구분됨을 확인하였으며, ‘지식’, ‘기능’, ‘가치와 태도’로 유목화하였다. 마지막으로 측정하고자 하는 영역으로 유목화된 문항 간 내적 일치도를 확인하기 위한 Cronbach- $\alpha$  계수는 모든 영역에서 0.90 이상으로 도출되어 검사 도구에 대한 신뢰도를 확보하였다고 할 수 있다[26].

#### 4. 연구 결과

##### 4.1. 교육 프로그램에 대한 이해도와 흥미도 분석 결과

교육 프로그램의 주차 별 핵심 내용에 대한 이해도와 흥미도 결과를 시각화하면 (Fig. 2)와 같다.



(Fig. 2) Results of Visualization of Understanding and Interest in Main content

이해도와 흥미도 모두 ‘매우 좋음(5)’과 ‘좋음(4)’의 긍정적인 자기 평가 비율이 ‘매우 낮음(1)’, ‘낮음(2)’의 부정적인 평가 비율에 비해 높게 나타났다. 이를 통해 텍스트 프로그래밍 언어인 파이썬을 도구로 활용한 빅데이터 분석 교육 프로그램의 내용과 수준이 참여자에게 적절했던 것으로 해석할 수 있다.

‘매우 좋음(5)’, ‘좋음(4)’의 응답인 긍정적인 자기 평가를 중심으로 분석하면, 모든 핵심 내용에서 과반수 이상의 학생이 이해도와 흥미도에 대해 긍정적으로 평가한 것을 확인할 수 있다. 또한, 이해도에 비해 흥미도의 긍정적인 자기 평가가 다소 높은 것으로 나타났다.

특히, 이해도의 경우 교육이 진행될수록 프로그램의 핵심 콘텐츠 난이도가 반영되어 긍정적인 자기 평가 비율이 다소 낮아지는 결과를 보였으나, 흥미도는 이해도와 비교해, 수업 전반에서 높은 비율을 비교적 잘 유지하고 있음을 확인할 수 있다. 이는 다양한 교육 전공과 관련된 주제의 데이터 활용과 데이터 분석을 교육적으로 활용할 수 있는 내용 제시를 통해 지속적으로 흥미를 유발하였기 때문으로 해석된다.

##### 4.2. 데이터 리터러시 지식 영역 분석 결과

프로그램 참여자들에 대한 사전 검사를 진행하고 15주에 걸친 독립변수 투입 후 사후 검사를 통한 데이터 리터러시 지식 영역 분석 결과는 <Table 8>과 같다.

<Table 8> Data Literacy Knowledge Area Analysis Results

Factor	M	SD	Std Error	t	P	
Knowledge	Before	2.30	1.03	0.16	-9.02	<.001
	After	3.80	0.69	0.10		

사전-사후 검사 결과, 유의수준은 0.001 이하로 통계적으로 유의미한 결과가 나타났으며, 평균이 1.50 증가해 3가지 영역 중에서 가장 큰 상승을 보였다. 이를 통해 본 연구에서 투입된 빅데이터 분석 교양강좌 프로그램은 예비교사들의 데이터 리터러시 중 지식영역 역량 향상에 긍정적인 영향을 미쳤다고 볼 수 있다. 또한, 표본의 편차가 줄어들었으며, 이를 바탕으로 표본 집단 내의 변화를 확인하기 위해 사전 검사 결과 하위 30%이며, 프로그래밍 경험이 없는 집단과 그 외 집단으로 나누어 사전-사후 검사의 평균 차이를 비교하였다. 그 결과는 <Table 9>와 같다.

<Table 9> Bottom 30% Group and Other Group Analysis

	Pre-Test	Post-Test	Increase Value
Bottom 30% Group	1.21	3.62	2.41
Others	2.81	3.98	1.17
Gap	1.60	0.36	1.24

사전 검사 결과 하위 30% 집단의 평균은 1.21로 그 외 집단에 비해 매우 낮은 것을 확인할 수 있다. 이는 프로그램 참여자가 다양한 교육 전공을 가지고 있으며, 데이터 관련 사전 지식이 상대적으로 부족한 특징이 반영된 것으로 해석된다. 프로그램 투입 이후에는 하위집단의 평균이 그 외 집단에 비해 비교적 높은 상승률을 보였으며 격차가 많이 줄어든 것을 확인할 수 있다. 이를 통해, 투입된 빅데이터 분석 교양강좌 프로그램은 상대적으로 사전 지식이 부족한 예비교사들의 데이터 리터러시 지식 영역의 역량을 높여주었으며, 사전 지식 및 경험으로 인한 예비교사 간 역량의 격차를 줄여주었다고 해석할 수 있다.

**4.3. 데이터 리터러시 기능 영역 분석 결과**

데이터 리터러시의 기능 영역 측면에서의 분석 결과는 <Table 10>과 같다.

<Table 10> Data Literacy Skills Area Analysis Results

Factor	M	SD	Std Error	t	P	
Skills	Before	2.70	0.91	0.14	-7.05	<.001
	After	3.80	0.59	0.11		

기능 측면 또한, 평균이 상승했으며 통계적으로 유의미하게 나타났다. 지식 요소와 비교해 평균의 변화가 크지는 않지만, 1.10의 높은 평균 상승으로 빅데이터를 분석할 수 있는 데이터 리터러시 기능 역량이 향상되었다고 해석할 수 있다. 기능 영역 또한 표본의 편차가 줄어들었으며, 사전 검사 결과 하위 30%이며, 프로그래밍 경험이 없는 집단을 나눠 비교한 결과는 <Table 11>과 같다.

<Table 11> Bottom 30% Group and Other Group Analysis

	Pre-Test	Post-Test	Increase Value
Bottom 30% Group	1.63	3.83	2.20
Others	3.17	3.93	0.76
Gap	1.54	0.10	1.44

프로그래밍 경험이 없는 하위 30%와 그외 그룹의 사전 검사 평균 차이는 1.54이었으며, 사후 검사 차이는 0.10으로 그 격차가 크게 줄어든 것을 확인할 수 있다. 특히, 전반적인 향상 정도가 더 높았던 ‘지식’ 영역에 비해 하위집단과 그 외의 집단 간의 차이가 더 줄어들어 투입된 프로그램이 하위집단의 데이터 리터러시 기능 역량을 길러주는데 효과적이었다고 해석할 수 있다. 또한, 하위집단의 평균 증가와 비교하여 그 외 집단의 평균 증가율이 상대적으로 낮아 다양한 예비교사들의 수준을 만족시킬 수 있는 데이터 분석 활동과 교수 전략이 필요함을 도출할 수 있다.

**4.4. 데이터 리터러시 가치와 태도 영역 분석 결과**

데이터 리터러시의 ‘가치와 태도’ 영역에서의 분석 결과는 <Table 12>와 같다.

<Table 12> Data Literacy Values and Attitudes Analysis Results

Factor	M	SD	Std Error	t	P	
Value and Attitude	Before	3.51	0.78	0.12	-6.03	<.001
	After	4.10	0.60	0.93		

‘가치와 태도’ 영역 또한, 통계적으로 유의미한 결과를 보였으며, 프로그램이 긍정적인 영향을 미쳤다고 해석할 수 있다. 하지만, ‘지식’ 및 ‘기능’ 영역과 비교해 평



준과 표준 편차 모두 변화의 폭이 가장 적게 나타났다. 이는 다른 영역에 비해 사전 검사의 평균값이 높았으며, 사전 검사 결과, 프로그램 참여자는 데이터로 인해 발생할 수 있는 사회 문제에 대한 교육의 필요성에 높은 공감감을 보였고 인간중심, 공정 등의 가치를 중요시하는 태도를 사전에 함양하고 있었던 것으로 분석되었다.

### 5. 결론 및 제언

AI를 중심으로 한 디지털 사회를 살아갈 학생들이 갖추어야 할 기초역량으로 데이터 리터러시가 강조되고 있다. 이에, 본 연구는 학생들의 데이터 리터러시를 길러 줄 수 있는, 예비교사의 역량에 초점을 맞추어 데이터 리터러시 증진 빅데이터 분석 프로그램을 적용하고 그 효과성을 분석하였다.

데이터 리터러시를 ‘지식’, ‘기능’, ‘가치와 태도’의 영역으로 나누어 연구를 진행하였으며, 기초소양 교육과 예비교사 대상이라는 특수성에 맞추어 학습자 중심의 도구 선정하고 강좌 내용을 구성하여 15주, 45시간의 교육 프로그램을 투입하였다. 또한, 학생들의 흥미도와 이해도 자기 평가를 진행하였으며, 재구성된 데이터 리터러시 검사 도구를 통한 사전-사후 검사 결과를 분석하였다. 본 연구의 결과와 시사점은 다음과 같다.

첫째, 15주에 걸친 프로그램에 대한 학생들의 이해도와 흥미도 자기 평가 결과를 통해, 예비교사들의 수준을 고려한 빅데이터 분석 도구인 파이썬의 선정과 내용 구성이 적절했음을 확인하였다. 또한, 다양한 전공을 고려한 데이터 분석 주제가 프로그램 투입의 전반에서 흥미도에 긍정적인 영향을 주었음을 유추할 수 있었다.

둘째, 투입 프로그램이 데이터 리터러시의 영역별 역량에 미치는 영향을 분석하기 위한 사전-사후 검사 결과, 데이터 리터러시의 모든 영역에서 통계적으로 유의미하고, 평균값이 증가하여 학생들의 역량이 향상되었음을 확인하였다. 결과를 상세히 분석하면, ‘지식’ 영역에서 역량이 크게 향상되었으며, ‘지식’과 ‘기능’ 영역은 표본 집단 내의 변화를 확인하기 위해 구분한 하위 30%의 집단 역량이 비교적 많이 향상됨을 확인하였다. 이는 예비교사의 사전 지식과 경험에 따른 격차를 줄이는데, 투입된 프로그램이 긍정적인 영향을 주었다고 해석할 수 있다.

본 연구를 통한 제언은 다음과 같다. 첫째, 데이터 리터러시의 영역과 하위 요소를 면밀하게 분석할 수 있는 표준화된 검사 도구의 개발이 필요하다. 학생들의 데이터 리터러시 역량 증진을 위해서는 ‘지식’, ‘기능’, ‘가치와 태도’의 영역에 대한 분석뿐만 아니라, 각 영역의 하위 요소에 대한 심도 있는 분석이 필요하다. 이를 위해서는 데이터 리터러시의 영역과 구성 요소 모두 측정 가능한 신뢰도 높은 검사 도구의 개발이 선행되어야 한다. 둘째, 교양 수업 내의 다양한 요소 특히, 학생들의 사전 지식 및 수준과 데이터 리터러시 간의 요인분석을 통해서 교양 수업에 참여하는 다양한 수준의 교육 대상자의 데이터 리터러시를 효과적으로 길러 줄 수 있는 단계별 교수학습 전략과 수업 모형의 개발이 필요하다.

학생들이 살아갈 미래 사회의 필수 역량으로 강조되고 있는 데이터 리터러시 관련 교육 연구가 최근 본격적으로 이루어지고 있다. 본 연구의 결과와 제언이 추후 데이터 리터러시 후속 연구에 활용되어 학생과 교사의 미래 역량 증진에 도움을 줄 수 있기를 기대한다.

### 참고문헌

- [1] Bae, H. (2019). Educational Implications of Data Literacy in Social Studies. *Theory and Research in Citizenship Education*, 51(1), 95-120.
- [2] Bhargava, R., & D'Ignazio, C. (2015, June). *Designing tools and activities for data literacy learners*. In Workshop on Data Literacy, Webscience.
- [3] Carlson, J., Fosmire, M., Miller, C. C., & Nelson, M. S. (2011). Determining data information literacy needs: A study of students and research faculty. *portal: Libraries and the Academy*, 11(2), 629-657.
- [4] D'Ignazio, C., & Bhargava, R. (2015, September). Approaches to building big data literacy. *In Proceedings of the Bloomberg data for good exchange conference*.
- [5] Han, S. W. (2018). A study about the concept of data literacy based on digital humanities. *Journal of the Korean Society for information Management*, 35(4), 223-236.

- [6] Hattwig, D., Bussert, K., Medaille, A., Burgess, J. (2013). Visual literacy standards in higher education: New opportunities for libraries and student learning. *Libraries and Academy*, 13(1), 61-89.
- [7] Hong, J., Kim, Y. (2020) Development of AI Data Science Education Program to Foster Data Literacy of Elementary School Students. *Journal of The Korean Association of Information Education*, 24(5), 463 - 471.
- [8] Hur, K. (2020). A Study on Elementary Education Examples for Data Science using Entry. *Journal of The Korean Association of Information Education*, 24(5), 473-481.
- [9] Hur, K. (2021). A Case Study of Basic Data Science Education using Public Big Data Collection and Spreadsheets for Teacher Education. *Journal of The Korean Association of Information Education*, 25(3), 459-469.
- [10] Jun, S., Jeon, Y., Jeong, I. (2021). Design of Liberal Arts Subjects for Artificial Intelligence Education for Pre-Teachers in Elementary and Secondary Schools. *Journal of The Korean Association of Information Education*, 25(5), 859-869.
- [11] Kim, H. (2020). Analysis of Data Literacy in the Core Curriculum to Improve Students' 4C Skills: Communication, Collaboration, Critical Thinking, and Creativity. *Korean Journal of General Education*, 14(6), 147-159.
- [12] Kim, J., Sohn, E. (2021) Development of Data Science Course and Analysis of Computational Thinking Effect for non-Major Students. *Journal of The Korean Association of Information Education* 24(3), 23-31.
- [13] Kim, S., Kim, S., Kim, H. (2019). Analysis of International Educational Trends and Learning Tools for Artificial Intelligence Education. *The Journal of Korean Association of Computer Education*, 23(2), 25-28.
- [14] Kim, S., Kim, T. (2021). A Study of the Definition and Components of Data Literacy for K-12 AI Education. *Journal of The Korean Association of Information Education*, 25(5), 691-704.
- [15] Kim, Y. (2018). *Data Science Education Program to Improve Computational Thinking and Creativity*. Doctor thesis. Jeju. Jeju National University.
- [16] KOFAC. (2019). *2019 Global SW Education Conference Issue Paper: Age of AI, SW Education is reconsidered*. Seoul. KOFAC.
- [17] Koo, D., Kim, D. (2020). Data science education program based on problem solving learning. *The Journal of Korea Elementary Education*, 31, 203-215.
- [18] Korean Government. (2020). *A Plan to Spread AI SW Education for People Across the Country*. <https://www.4th-ir.go.kr/article/download/710>.
- [19] Korean Government. (2020) *National Strategy for Artificial Intelligence*. <https://www.korea.kr/com-mon/download.do?fileId=190114031&tblKey=GMN>.
- [20] Lee, E. (2020). A Comparative Analysis of Contents Related to Artificial Intelligence in National and International K-12 Curriculum. *The Journal of Korean Association of Computer Education*, 23(1), 37-44.
- [21] Lee, H., Kim, S., Kim, H., Lee, W., Lim, S., Park, S. (2019). *A study on measuring the digital literacy level of elementary and middle school students at the national level in 2018*. Daegu, KERIS.
- [22] Lee, Y., Jang, B., Lim, J., Park, J., Seo, M., Lee, D., Kim, H. (2021) *Basic Artificial Intelligence*, Seoul, Cmass.
- [23] Mandinach, E. B., Gummer, E. S. (2013). A systemic view of implementing data literacy in educator preparation. *Educational Researcher*, 42(1), 30-37.
- [24] Mandinach, E. B., Gummer, E. S. (2016). What does it mean for teachers to be data literate: Laying out the skills, knowledge, and dispositions. *Teaching and Teacher Education*, 60, 366-376.

- [25] Ministry of Education. (2021) Key Points of the 2022 Revised Curriculum Introductions, <https://www.moe.go.kr/boardCnts/view.do?boardID=294&boardSeq=89671&lev=0&searchType=null&statusYN=W&page=1&s=moe&m=020402&opType=N>.
- [26] No, K. (2019). *The proper methods of statistical analysis for dissertation SPSS & AMOS*. Seoul. Hanbit.
- [27] OECD. (2018). *The future of education and skills: Education 2030*. OECD Education Working Papers.
- [28] Oh, K., Kim, H. (2020). An Analysis of the Influence big data analysis-based AI education on Affective Attitude towards Artificial Intelligence. *Journal of The Korean Association of Information Education*, 24(5), 463 - 471.
- [29] Park, Y., Lee, S. (2020). Study on the Direction of Universal Big Data and Big Data Education-Based on the Survey of Big Data Experts. *Journal of The Korean Association of Information Education*, 24(2), 201 - 214.
- [30] Polit, D. F., & Beck, C. T. (2006). The content validity index: are you sure you know what's being reported? Critique and recommendations. *Research in nursing & health*, 29(5), 489-497.
- [31] Qilk. (2018). *Data literacy workforce assesment*. <https://thedata literacyproject.org/resources>.
- [32] Shields, M. (2005). Information literacy, statistical literacy, data literacy. *IASSIST quarterly*, 28(2-3), 6-6.
- [33] Song, Y. (2021). *The Data-Driven Debate (DDD) Instructional Model for Improving Data Literacy*. Master Thesis. Seoul. Department of Education The Graduate School Seoul National University
- [34] Swan, K., Vahey, P. J., Rafanan, K., Stanford, T. (2009). *Challenges to cross-disciplinary curricula: Data literacy and divergent disciplinary perspectives*. In Annual Conference of the American Educational Research Association. San Diego, CA.

### 저자소개

#### 김 슬 기



2008년 경인교육대학교 초등교육학과 (교육학학사)  
 2016년 경인교육대학교 융합교육학과 (교육학석사)  
 2020년~현재 한국교원대학교 컴퓨터교육과 박사과정  
 관심분야: 정보(SW·AI) 교육, 컴퓨팅 사고력, 프로그래밍, 빅데이터, 데이터 리터러시  
 e-mail: tmfrlska85@gmail.com

#### 김 태 영



1985년 한양대학교 산업공학과(공학사)  
 1990년 Texas A&M University Computer Science(M.S.)  
 1994년 Texas A&M University Computer Science(Ph.D)  
 1994년~현재 한국교원대학교 컴퓨터교육과 교수  
 관심분야: 컴퓨터교육, 데이터베이스, 프로그래밍  
 e-mail: tykim@knue.ac.kr