

데이터 리터러시 향상을 위한 프로젝트 기반 데이터 분석 교육 프로그램 연구

김지현¹ · 김한일^{2*} · 김성백^{2*}¹제주대학교 대학원 과학교육학부 컴퓨터교육전공 석사과정²제주대학교 컴퓨터교육과 교수

Studying a Project-Based Data Analysis Education Program to Enhance Data Literacy

Ji Hyeon Kim¹ · Hanil Kim^{2*} · Seong Baeg Kim^{2*}¹Graduate Student, Computer Education Major, Jeju National University, Jeju 63243, Korea²Professor, Department of Computer Education, Jeju National University, Jeju 63243, Korea

[요약]

빅데이터를 포함한 디지털 기술의 발달로 데이터 분석 교육의 필요성이 높아지고 있다. 고등학교 진로 선택과목으로 인공지능 분야에서 빅데이터를 이해하고, 수집, 분석, 평가, 시각화하는 과정을 학습할 수 있는 <인공지능 기초>, <데이터 과학과 머신러닝>, <데이터 과학>이 신설되었다. 실제 고등학교에서는 구체적인 데이터 분석 연구 사례가 부족한 상황이며, 이론적인 지식뿐만 아니라 실제 상황에서의 문제 해결 능력, 창의성을 발전시킬 수 있는 프로젝트 기반 데이터 분석 교육 자료 개발이 필요한 상황이다. 따라서 일반계 고등학교에서 데이터 분석 교육 관련 과목의 성공적인 운영을 위해 프로젝트 기반 데이터 분석 수업 학습 자료를 개발하고 수업을 통해 효과성을 분석하였다.

[Abstract]

With the rapid advancement of digital technologies, including big data, the importance of data analysis education has continued to grow. In the field of artificial intelligence, courses such as “Introduction to Artificial Intelligence,” “Data Science and Machine Learning,” and “Data Science” have been introduced as elective subjects within high school career paths. However, in actual high school settings, there is a lack of concrete case studies and practical examples related to data analysis education. To address this gap, it is necessary to develop project-based data analysis instructional materials that not only provide theoretical knowledge but also promote problem-solving skills and creativity in real-world contexts. Accordingly, to support the effective implementation of data analysis-related subjects in general high schools, we developed project-based data analysis course materials and examined their effectiveness through classroom-based instruction.

색인어 : 프로젝트, 데이터, 데이터 분석, 데이터 분석 교육, 데이터 리터러시**Keyword** : Project, High School, Data Analysis, Data Analysis Education, Data Literacy<http://dx.doi.org/10.9728/dcs.2026.27.2.463>

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 06 January 2026; Revised 02 February 2026

Accepted 05 February 2026

***Corresponding Author; Hanil Kim, Seong Baeg Kim**

Tel: +82-64-754-3291, +82-64-754-3292

E-mail: hikim@jejunu.ac.kr, sbkim@jejunu.ac.kr

1. 서론

전 세계적으로 빅데이터, 인공지능을 포함한 첨단 디지털 기술의 발달로 다양한 분야의 획기적인 변화를 초래하며 삶의 많은 변화가 일어나고 있다. 데이터가 새로운 가치를 창출해 내고 경제 성장을 이끌면서 세계 각국은 데이터 기반의 디지털 전환을 가속화 하고 있으며, 우리나라도 데이터 생태계의 국가 경쟁력을 높이기 위하여 국가 데이터 정책 추진 방향을 수립하여 데이터 가치와 활용을 강조하였다[1]. 교육부에서는 인공지능 기반 디지털 교육 정책 및 기반을 마련하고 소프트웨어·인공지능 융합 교육 활성화에 대한 계획이 담겨있는 디지털 인재양성 종합방안을 발표하였다[2]. 2015 개정 교육과정에서는 고등학교 진로 선택과목으로 <인공지능 기초>, <데이터 과학과 머신러닝>을 신설하여 시행되고 있으며 [3], 2022 개정 교육과정에서는 <데이터 과학>을 신설하여 2025년도부터 시행될 예정이다[4]. 과목의 일부로 선택과목이 개설 되었지만 실제 고등학교에서는 미리 제시된 학습용 데이터를 통한 데이터 분석으로만 그쳐 실제적 맥락에서 데이터를 분석하고 의미 있는 통찰을 얻어내지 못하였으며, 구체적인 데이터 분석 수업자료와 연구 사례가 부족한 상황이다. 따라서 고등학교에서 정보·인공지능 신설 과목의 내실있는 운영을 위해 데이터 분석 수업 사례와 효과성 논의가 필요하다.

이에 본 연구에서는 고등학교 정보와 연계가 가능하며 이론적인 개념뿐만 아니라 실제 데이터를 다루고 분석하는 경험을 학습할 수 있는 프로젝트 기반 데이터 분석 교육 프로그램을 개발하였다. 일반계 고등학생 대상으로 수업을 적용하여 데이터 리터러시에 대한 사전·사후 조사 분석을 통해 프로그램의 효과성을 분석하여 국내 인공지능 교육의 활성화와 정착을 도모하고자 한다.

II. 선행연구 및 이론적 배경

본 연구에 있어서 기존 선행 연구로 제시된 다양한 데이터 분석 교육 및 프로젝트 기반 수업 연구를 정리하고자 한다.

2-1 데이터 분석 교육 선행 연구

인공지능 교육에서 데이터 교육에 대한 필요성이 높아짐에 따라 데이터 교육에 관한 연구가 이루어 지고 있다. 김민규의 연구에서는 파이썬을 활용한 데이터 시각화 교육 교재를 개발하고 초등학교 6학년 대상으로 36차시 수업을 데이터 시각화 3단계 절차에 따라 교육에 적용하였다[5]. 그런 다음 데이터 프로젝트를 구현하는 과정을 통해 컴퓨터 사고력 신장이 이루어질 수 있는 수업 모형을 제안하였다.

오진희 연구에서는 데이터 리터러시 구성 요소인 지식, 기능, 태도에 따라 파이썬을 활용한 데이터 교육 내용을 20차시로 구성하였다. 초등학교 영재 학생 대상으로 데이터 리터러

시 신장이 이루어질 수 있는 수업 모형을 제시하였다[6].

이은정의 연구에서는 데이터 리터러시 신장을 위한 데이터 분석 도구를 활용한 교육 프로그램을 현장에 적용하고 효과성을 검증하였다. 블록형 ‘Machine Learning for Kids’ 분석 도구, 위젯형 ‘Brightics Education’ 분석 도구를 활용한 데이터 분석 프로그램을 개발하여 이를 일반계 고등학생 대상으로 교육하였으며 데이터 리터러시 향상을 검증하였다[7].

위의 선행연구를 종합하였을 때, 데이터 교육에 관한 연구들이 이루어지고 있음을 알 수 있다. 블록이나 위젯 형태의 데이터 분석 도구를 활용한 연구나, <고등학교 정보>, <인공지능 기초>에서 사용되고 있는 텍스트 프로그래밍 언어인 파이썬 기반 데이터 분석 연구들이 진행되고 있으나, 학생들이 실제로 관심을 가질 수 있는 실생활 데이터를 수집하고 전처리 후 결과를 도출하는 프로젝트 수업 사례에 대한 연구가 많이 필요하다. 또한 초등학생, 초등 영재 대상 중심의 연구들이 주로 이루어지고 있어 일반계 고등학생을 대상으로 파이썬 데이터 분석 수업 사례 연구가 필요하다.

2-2 데이터 리터러시

데이터 리터러시(Data Literacy)는 ‘데이터’와 ‘리터러시’의 의미가 결합 되어 최근 정보의 표현 방법이 다양화되고 있다. 배화순 연구에서는 다양한 형태의 데이터로부터 의미있는 정보를 찾아내고 다양한 실생활 문제를 해결하기 위해 이를 분석하여 결론을 끌어내고 합리적인 의사결정을 할 수 있도록 할 뿐만 아니라 이를 토대로 소통하는 능력을 뜻한다고 정의하였다[8]. 이승철 외는 컴퓨터 교육 분야에서 데이터 리터러시 개념을 탐구하여, 다양한 학사들의 개념을 정리하여 ‘수집, 분석, 표현, 이해, 활용, 윤리, 판단’을 요소로 정리하여 2015 개정 교육과정 정보 교과 내용 체계의 목표, 교과 역량을 바탕으로 정확하고 효과적인 실생활 데이터를 이용하여 문제를 해결할 수 있는 능력으로 정리하였다[9]. 데이터 리터러시는 통계 활용 능력에 초점을 맞춘 관점과 이은정의 연구에서는 데이터 활용이 증가하는 현재 요구되는 역량, 실시간으로 생성되는 데이터의 의미를 파악하여 해석해 내는 능력은 데이터 활용 과정 전반에 필요한 역량으로 정의하고 있다 [7]. 박성식의 연구에서는 데이터 리터러시는 다양한 형태의 데이터로부터 의미 있는 정보를 찾고, 실생활 문제를 해결하기 위해 데이터를 분석하여 결론을 이끌어 합리적인 의사결정을 할 수 있도록 하는 의사소통 능력을 의미한다고 정의하며 [10], 이가원의 연구에서는 의사소통을 바탕으로 실생활의 다양한 문제를 해결할 수 있는 능력으로 정의하였다[11].

2-3 프로젝트 기반 수업

프로젝트 기반 수업은 학습자가 실제 문제를 중심으로 탐구와 협력을 수행하며 문제 해결 역량을 기르는 교수·학습 방법으로, 정보 교과에서 효과적인 수업 방식으로 보고되고 있다.

서주영의 연구에서는 데이터 분석 수업을 문제 정의, 데이

터 준비 및 분석, 결과 적용의 단계로 구조화하고 이를 이론-실습-미션 형태의 학습 모듈로 구성하여 교육 효과를 검증하였다[12]. 또한 파이썬 기반 프로그래밍 교육 연구에서는 학습 역량을 중심으로 단계적 모듈을 설계하고 프로젝트 수행을 통해 학습자의 문제 해결 능력과 개념 이해를 체계적으로 신장시킬 수 있음을 제시하였다[13]. 임미숙의 연구에서는 빅데이터를 활용한 인공지능 교육에서 주차별 학습 목표와 활동을 단계적으로 설계한 경험형 프로젝트 수업이 학습자의 데이터 리터러시와 학습 몰입도를 동시에 향상시키는 데 효과적임을 보고하였다[14].

정리하면 데이터 분석 교육에서 프로젝트 기반 수업은 실제 문제를 중심으로 단계적 차시 구성과 모듈화된 학습 활동을 통해 학습자의 문제 해결 역량과 데이터 리터러시를 효과적으로 신장시킬 수 있음을 시사한다.

III. 교육 프로그램 설계 및 개발

3-1 수업 설계 방향 및 목적

본 연구는 고등학교 데이터 과학 교육 프로그램으로 총 14 차시를 기준으로 구성하였으며, 일반계 고등학생을 대상으로 프로젝트 기반 데이터 분석 수업을 설계하였다. 본 프로그램은 기존에 개설되어 운영 중인 고등학교 <정보과학> 정규 교과 시간에 적용되었으며, 2015 개정 정보과학 교육과정의 프로그래밍 단원 성취기준에 근거하여 파이썬을 활용한 데이터 이해·수집·분석·평가·시각화 활동을 통해 문제 해결 중심의 역량을 기르는 데 목적을 두었다. 정보과학 프로그래밍 단원은 단순한 문법 습득을 넘어 자료를 구조화하고 문제 해결 절차를 설계·구현하는 역량을 강조한다. 이에 본 연구의 데이터 분석 활동에서 데이터를 모델링하고 적합한 자료형을 선택하며, 제어구조를 활용하여 분석 절차를 구현하였다. 이러한 점에서 본 수업은 프로그래밍 단원의 성취기준이 지향하는 문제 해결 중심 학습을 실제 데이터 맥락에서 구현한 사례라 할 수 있다. 차시 계획은 표 1과 같다.

3-2 차시별 수업 내용

1차시 수업은 빅데이터의 정의, 특징, 사례에 대해 설명식 수업으로 진행하며, Padlet 실시간 협업 교육 도구를 활용하여 빅데이터의 사례를 조사하고 기록하여 공유할 수 있도록 한다. 빅데이터 사례 공유를 통해 보건 의료, 기상 정보, 자동 번역, 교통 등 정치·경제·사회·문화 모든 분야에 걸쳐 다양하게 활용되는 빅데이터의 사례를 조사하고 데이터의 가치를 찾을 수 있도록 한다.

2차시 수업은 데이터의 속성, 종류, 데이터 전처리 과정에 대해 알아본다. 스프레드시트 도구를 활용하여 제시된 초등학교 건강검진 데이터에서 결측치, 이상치를 찾고 문제 상황에 따라 필요한 핵심 속성을 추출하는 실습을 진행한다. 인공지

표 1. 데이터 분석 프로젝트 교육 프로그램

Table 1. Data analysis project education program

Lesson	Contents
1	What is Big Data?
2	Data Preprocessing
3	Analyzing Data Without Libraries
4-5	Data Analysis and Visualization Python Library
6	Data Visualization Library
7	Data Analysis Project Practice
8	Project Topic and Planning Establishment
9-12	Project Implementation
13-14	Project Presentation

능 모델을 구현하기 위해서도 핵심 속성을 추출하는 것이 중요하다라는 점을 설명한다.

3차시 수업은 파이썬 기본 문법을 활용한 날씨 데이터 분석하기 수업을 진행한다. 반복문을 활용한 기상 자료 개방 포털의 기존 데이터를 활용하여 ‘기상 관측 이래로 서울의 가장 높은 기온은 얼마이며, 그 날씨는 언제일까?’ 라는 주제로 데이터를 분석한다. 데이터 분석 결과를 도출해낼 수 있으며 라이브러리 없이 조건문, 반복문을 활용하여 데이터를 분석하는 과정을 실습한다.

4-5차시 수업은 데이터를 관리하고 연산을 수행하는 Array(행렬)에 대한 개념을 이해하고 활용할 수 있는 numpy 라이브러리, 데이터 조작 및 분석을 위한 pandas 라이브러리, 간단한 성적 데이터를 사용하여 필요한 라이브러리를 검색하여 찾고 활용할 수 있도록 한다.

6차시 수업은 데이터 시각화 matplotlib 라이브러리를 통해 데이터를 표현할 수 있는 방법을 학습한다. 구체적으로 막대 그래프, 산점도 그래프, 선 그래프를 학습한 후 데이터를 시각화하고 데이터에 적절한 그래프 표현 방식이 무엇인지 생각할 수 있도록 수업 내용을 구성하였다.

7차시 수업은 문제 정의, 데이터 수집, 데이터 전처리, 데이터 시각화, 데이터 분석을 진행할 수 있도록 안전 문제와 관련하여 “인구 대비 CCTV설치 대수가 적절한가?”라는 문제 상황을 해결하기 위해 데이터 분석을 진행한다. 통계 포털 사이트의 읍면동별 세대 및 인구 데이터와 로컬데이터 사이트의 지역의 시군구 CCTV 정보 데이터를 수집하여 데이터 분석을 진행한다. 시각화한 자료를 토대로 CCTV 설치가 필요한 곳을 분석 결과로 찾도록 한다. 학습을 통해 데이터 분석 프로젝트 과정을 경험할 수 있도록 한다.

8차시 수업은 “~을 위한 데이터 분석”을 주제로 실생활에서 해결할 수 있는 문제 상황을 찾아 주제를 선정하도록 한다. 데이터 분석으로 해결할 수 있는 사회적 문제를 찾아 주제를 선정하고, 공공데이터, 민간 데이터를 문제 해결에 필요한 실생활 데이터를 수집할 수 있도록 자료를 제시하여 프로젝트 주제에 가장 적합한 데이터를 찾을 수 있도록 하고, 신뢰성 있는 정보를 수집하고 오류를 최소화하도록 한다.

9-12차시에서는 수집한 데이터를 전처리하고 시각화하여 분석 결과를 도출하도록 하였다. 선정 이유, 데이터 설명, 분석을 위한 알고리즘, 시각화, 결과 분석 보고서를 작성하도록 수업 내용을 구성하였다.

13-14차시 데이터 전처리, 시각화, 분석 과정을 발표 자료로 제작하도록 하여 학생들과 발표를 통해 공유하도록 하도록 한다. 프로젝트 결과 발표 시 분석 결과를 통해 실생활 문제를 해결하기 위한 제언을 할 수 있도록 안내한다. 또한 데이터 수집을 위한 노력, 전처리 과정, 분석 시 어려움을 전달하도록 한다.

3-3 타당도 검사

컴퓨터 교육 및 인공지능 융합 교육 전공 전문가 6인에게 차시 내용 및 주안점에 대한 타당도 검사를 진행하였다. 내용 및 주안점 바탕으로 다음과 같이 프로젝트 기반 데이터 분석 교육 프로그램을 수정 및 보완하였다.

전문가 의견으로는 라이브러리 없이 데이터 분석 수업 내용에서 “자료를 불러올 때 등 라이브러리를 활용하기 때문에 데이터 분석 수업에서 라이브러리를 활용하지 않고 데이터 분석하는 수업보다는 분석 라이브러리를 통해 다양한 데이터를 다루는데 중점을 둘 필요가 있음”, “프로젝트 수업에서 효과적인 프로젝트 진행을 위해 교사의 피드백을 제시하여 올바른 데이터 분석 결과를 얻을 수 있도록 할 필요가 있음”, “효과적인 피드백 제공을 위해 프로그램 참여 인원수의 조정이 있다면 좋겠다는 생각이 듦, ”프로그램에 참여하는 학생들의 출발점에 따라 효과성의 차이가 클 것으로 보임”이라는 의견을 반영하여 3차시의 라이브러리 없이 데이터 분석하기 수업 차시 내용을 생략하고, 8차시부터 시작되는 프로젝트 차시를 7차시부터 진행하여 프로젝트 구현 수업을 진행하여 다양한 데이터를 다룰 수 있도록 수정하였다. 프로젝트 주제 및 계획 수립 시 학생 수준차에 따라 개인 뿐만 아니라 2 ~ 3명의 학생들이 프로젝트에 참여할 수 있도록 프로젝트 참여 인원수를 조정하였다. 또한, 11차시에 프로젝트 진행 중 주제와 데이터에 대한 교사의 피드백을 제시하여 올바른 데이터 분석 결과를 얻을 수 있도록 하고 피드백 사항을 반영할 수 있도록 한다.

이를 종합하여 선정한 학습 지도 계획을 표 2와 같이 정리하였다.

3-4 수정된 차시별 수업 내용

컴퓨터 교육 및 인공지능 융합 교육 전공 전문가 6인에게 차시 내용 및 주안점에 대한 타당도 검사를 진행하였다. 내용 및 주안점 바탕으로 다음과 같이 프로젝트 기반 데이터 분석 교육 프로그램을 수정 및 보완하였다.

7차시 수업은 “~을 위한 데이터 분석”을 주제로 실생활에서 해결할 수 있는 문제 상황을 찾아 주제를 선정하도록 한다. 특히, 데이터 분석으로 해결할 수 있는 사회적 문제를 주제로

표 2. 수정된 데이터 분석 프로젝트 프로그램

Table 2. Revised data analysis project program

Lesson	Contents
1	What is Big Data?
2	Data Preprocessing
3-4	Data Analysis and Visualization Python Library
5	Data Visualization Library
6	Data Analysis Project Practice
7	Project Topic and Planning Establishment
8-12	Project Implementation
13-14	Project Presentation

하여 공공데이터, 민간 데이터를 문제 해결에 필요한 실생활 데이터를 수집하도록 한다. 또한, 프로젝트 주제에 가장 적합한 데이터를 찾을 수 있도록 하고, 신뢰성 있는 정보를 수집하고 오류를 최소화하도록 한다. 프로젝트 주제 및 계획 수립 단계에서 주제, 선정 이유, 수집한 데이터에 대한 설명, 데이터 분석을 위한 과정, 데이터 시각화 그래프, 데이터 분석으로 알 수 있는 결과 보고서를 그림 1과 같이 작성하도록 한다.

Data Analysis Project Report			
Name:		Desired career path	
▶ Topic			
▶ Reason for Choosing the Topic			
▶ Description of Collected Data			
▶ Process for Data Analysis			
▶ Data Visualization Graph			
▶ Findings from Data Analysis			

그림 1. 데이터 분석 프로젝트 보고서 구성

Fig. 1. Data analysis project report structure

8-12차시 수업은 학생이 해결하고자 하는 분야의 수집한 데이터를 불러와 전처리하고 시각화하여 결과를 도출하도록 한다. 실제 학생들이 데이터 분석 프로젝트를 진행하였던 주제를 표 3과 같이 정리하였다.

프로젝트 진행 중 주제와 데이터에 대한 교사의 피드백을 제시하여 올바른 데이터 분석 결과를 얻을 수 있도록 하고 피드백 사항을 반영하도록 하였다. 13-14차시에 완성된 보고서를 활용하여 데이터 분석 프로젝트 결과를 발표하도록 프로그램을 설계하였다. 이를 통해 데이터 분석 전체 과정을 학습하고, 결과를 도출해내도록 하였다.

표 3. 학생 데이터 분석 프로젝트 주제들

Table 3. Student data analysis project topics

Subject	Topic
Science	Trend in COVID-19 Vaccination Numbers by Province
Education	Comparison of Changes in Elementary School Student Population in Provinces Due to Decreasing Birth Rates
Business Management	Changes in Multicultural Household Numbers
Crime	Regional Crime Rates
Society	Distribution Rate of Youth Centers
Computer	Survey on Computer Usage by Age
Energy	Trends in Energy Production by Sector
Education	Student Math Achievement Based on Income
Sports	Correlation Between Pitcher's Average Earned Run Average and Various Records
Engineering	Comparison of Greenhouse Gas Emissions and Electricity Generation

IV. 프로젝트 기반 데이터 분석 교육의 연구 내용 및 방법

4-1 연구 대상

본 연구는 연구자가 재직 중인 00고등학교에서 2023년 1학기에 3학년 학생 총 24명(2학급)을 대상으로 실시하였다. 수업은 <정보과학> 정규 교과 시간에 진행되었다. 학생들은 1학년 정보 교과 시간에 파이썬 프로그래밍 언어를 학습하였다. 하지만 파이썬을 활용한 데이터 분석 경험이 거의 없는 편이다. 연구가 진행된 00고등학교는 특별실에 와이파이(Wi-Fi)가 구축되어 있으며, 노트북을 제공하여 수업을 진행하였다. 본 연구 대상자는 과목을 선택한 남학생 16명, 여학생 8명을 대상으로 선택 교과 수업에서 연구가 진행되다보니 데이터 분석 교육 프로그램을 적용하는 실험 반과의 차이를 비교할 집단을 설정하기 어려워 단일집단의 사전, 사후 검사 결과의 차이를 보고자 대응 표본 t-test를 실시하였다.

4-2 연구 절차

본 프로그램의 효과성을 살펴보기 위해 표 4와 같이 프로그램 적용 전과 후의 변화된 종속 변수를 분석하는 단일집단 사전-사후 분석을 하였다.

표 4. 수업 절차

Table 4. Class procedures

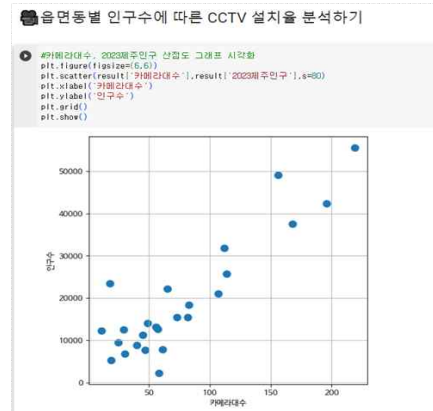
O1	X1	O2
O1, O2: Pre-and post-research on data understanding, data collection, data analysis, data evaluation, and data visualization		
X1: Python-Based Data Analysis Education		

첫째, 데이터 리터러시에 대한 데이터 이해, 수집, 분석, 평가, 시각화에 대한 사전 조사를 하였다. 둘째, 파이썬 기반 데

이터 분석 교육 프로그램을 14차시 진행하였다. 셋째, 데이터 리터러시에 대한 데이터 이해, 수집, 분석, 평가, 시각화에 대한 사후 조사를 하였다.

4-3 프로젝트 기반 데이터 분석 교육의 적용

설계한 데이터 분석 프로젝트 교육 프로그램의 차시별 수업 내용을 적용하였다.



*Since these are the students' original outputs, it is necessary to present them as they are.

그림 2. 6차시 데이터 분석 프로젝트 보고서

Fig. 2. Lesson 6 data analysis project report

그림 2는 6차시의 데이터 분석 프로젝트 실습에서 인구 대비 CCTV설치 대수가 적절한가? 문제를 해결하기 위한 인구와 CCTV 데이터 분석하는 코랩 수업 자료 중 일부이다. 실제 학생들이 스스로 데이터 분석 프로젝트를 진행하기 전에 같은 문제 상황과 데이터로 프로젝트 실습을 해보는 수업을 진행하여 프로젝트의 전반적인 과정을 학습한다.

그림 3은 7~12차시에 진행된 데이터 분석 프로젝트 수업

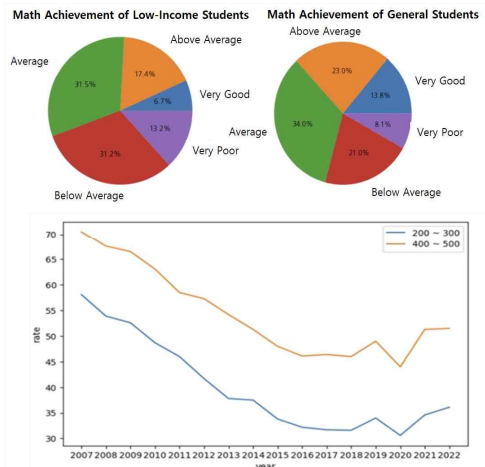


그림 3. 7~12차시 데이터 분석 프로젝트 보고서 일부

Fig. 3. Part of a student report from the data analysis project (lessons 7-12)

을 적용하여 학생이 작성한 보고서의 일부이다. 해당 프로젝트에서는 ‘소득에 따른 학생들의 수학 성취도’를 주제로 설정하고, 가구의 월평균 소득별 사교육 참여율 데이터와 저소득층 및 일반 학생의 수학 성취도 데이터를 수집하였다. 학생은 수집한 데이터를 전처리한 후, 소득 수준별 분포를 원그래프로, 소득별 사교육 참여율의 변화를 선그래프로 시각화하였다. 시각화 결과를 바탕으로 소득 수준이 학업 성취도에 영향을 미치며, 소득에 따른 사교육 참여의 차이가 성취도에 영향을 줄 수 있음을 해석하였다. 이처럼 본 수업에서는 학생들이 실생활 데이터를 수집·전처리하고 분석 및 평가하는 전 과정을 총 14차시에 걸쳐 적용하였다.

4-4 연구 도구

본 연구에서는 프로젝트 기반 데이터 분석 교육 프로그램에 대한 효과성을 분석하기 위해 표 5와 같이 설문 문항을 구성하였다. 데이터 리터러시 측정을 위해 ‘데이터 이해’, ‘데이터 수집’, ‘데이터 분석’, ‘데이터 평가’, ‘데이터 시각화’ 항목은 박성식의 데이터 리터러시 문항 중에서 발췌하여 사용하였다[8].

교육 프로그램의 효과성 분석을 위한 사전-사후 검사 도구는 데이터 이해에 대한 3문항, 데이터 수집에 대한 4문항, 데이터 분석에 대한 3문항, 데이터 평가에 대한 4문항, 데이터 시각화에 대한 4문항 총 18문항으로 구성하였다. 모든 문항은 ‘매우 아니다(1점)’부터 ‘매우 그렇다(5점)’의 5점 Likert 척도

로 측정되었다. 측정 도구의 내적 신뢰도를 검증한 결과, 사전 검사에서 Cronbach’s $\alpha = .893$, 사후검사에서 Cronbach’s $\alpha = .964$ 로 나타났으며, 사전·사후 자료를 통합한 전체 신뢰도는 Cronbach’s $\alpha = .957$ 로 산출되었다. 교육 프로그램의 효과성 분석을 위해 사전-사후 검사는 파이썬 numpy, pandas, scipy 라이브러리를 활용하여 동일 집단의 사전-사후 점수 차이를 비교하기 위한 대응 표본 t검정을 실시하였다.

4-5 프로젝트 기반 데이터 분석 교육 적용 결과 분석

본 프로그램의 데이터 리터러시 효과성 분석 결과는 표 6과 같다. ‘데이터 분석’모든 항목에서 통계적으로 가장 유의미한 차이가 있었다. 즉 데이터의 여러 가지 특성을 찾고, 특성 속에서 상관관계를 이해하고 표현된 자료의 의미를 해석하는 능력이 높아졌음을 알 수 있다.

데이터 수집의 ‘나는 데이터로부터 확인하고 싶은 내용을 가설의 형태로 나타낼 수 있다.’ 문항은 2.33점에서 3.62점으로 평균의 차이가 1.29로 향상되었으며, 통계적으로도 매우 의미 있는 차이를 보이고 있다($p < 0.001$). 이는 연구 이전에는 데이터 분석을 통해 해결할 수 있는 문제를 정의하지 못했지만, 프로그램을 통해 데이터 분석으로 해결하고자 하는 문제를 표현할 수 있다는 점을 알 수 있다.

데이터 이해의 ‘나는 가설을 확인하기 위해 필요한 데이터를 검색하여 접근할 수 있다’ 문항은 2.96에서 3.67으로 평균의 차이가 0.71점이 향상 되었고 통계적으로 유의미한 차이

표 5. 사전-사후 설문 문항들

Table 5. Pre-post survey questions

Item	Questions
Data Understanding	① I understand the definition of data.
	② I understand that there are various types of data.
	③ I understand the meanings of each row and column in collected data.
Data Collection	① I can evaluate and choose the sources of data.
	② I can express what I want to confirm from the data in the form of a hypothesis.
	③ I can search for and access the data necessary to verify the hypothesis.
	④ I can collect data in a suitable format based on the needs.
Data Analysis	① I can identify the necessary characteristics for data analysis among various attributes of the data.
	② I can identify correlations among various attributes of the data.
	③ I can understand the meaning of data presented in tables, charts, graphs, etc., and interpret them.
Data Evaluation	① I can critically assess the reliability of data by considering its source and collection process.
	② I can find errors in my data analysis and suggest ways to fix them right away.
	③ I can logically explain or criticize errors that occurred during someone else's data analysis process.
	④ I can critically evaluate distortions of facts that occurred during my own or someone else's data analysis and interpretation process.
Data Visualization	① I understand the characteristics of various visualization methods such as tables, charts, and graphs.
	② I can select the appropriate visualization method such as tables, charts, and graphs based on the needs and visualize the data accordingly.
	③ I can identify the key features of visualized data and explain them verbally or in writing.
	④ I can derive the analysis result based on the data.

표 6. 사전-사후 설문 결과

Table 6. Pre-post survey results

category		Pre-test		Post-test		t	p
		M	SD	M	SD		
Data Understanding	①	2.96	0.89	3.67	0.85	-3.205	0.004
	②	3.75	1.05	4.04	0.79	-1.372	0.183
	③	3.58	1.04	4.04	0.93	-2.298	0.031
Data Collection	①	3.50	0.76	3.88	0.88	-1.989	0.059
	②	2.33	0.90	3.62	0.90	-5.639	0.000
	③	3.21	0.71	3.92	0.91	-3.331	0.003
	④	3.42	0.95	3.83	0.85	-1.735	0.096
Data Analysis	①	2.96	1.06	3.54	0.91	-2.429	0.023
	②	2.96	0.98	3.62	0.99	-2.892	0.008
	③	3.79	0.82	4.17	0.99	-2.099	0.047
Data Evaluation	①	3.50	1.00	3.92	0.91	-2.198	0.038
	②	3.29	1.02	3.58	0.86	-1.372	0.183
	③	3.33	0.90	3.29	0.98	0.196	0.846
	④	3.08	1.04	3.50	0.87	-2.46	0.022
Data Visualization	①	3.58	0.86	4.12	0.93	-3.186	0.004
	②	3.50	1.00	4.04	0.89	-3.186	0.004
	③	3.46	0.87	3.83	1.11	-1.436	0.164
	④	3.46	1.08	4.04	0.98	-2.933	0.007

를 보인다($p < 0.05$). 이는 데이터 분석 프로젝트를 통해 데이터의 의미를 이해하고 다양한 데이터의 종류와 데이터 의미를 파악하는 능력이 높아진 것을 나타낸다.

반면에 데이터 평가 영역 중 ‘다른 사람이 수행한 데이터 분석 과정에서 발생한 오류를 찾아 논리적으로 설명하거나 비판하는 능력’과 관련된 문항에서는 통계적으로 유의미한 차이가 나타나지 않았다. 이는 본 연구에서 운영한 프로젝트 기반 데이터 분석 수업이 주로 학습자 개인 또는 소집단 중심의 데이터 수집·분석·시각화 활동에 초점을 두었기 때문으로 해석된다.

특히 다른 학생의 분석 결과를 검토하고 오류를 비판적으로 분석하는 활동은 충분한 상호 피드백 시간과 토의 과정이 요구되나, 실제 수업에서는 제한된 차시 운영으로 인해 학생 간 발표 결과에 대해 심층적으로 피드백하고 논리적으로 비판하는 활동이 충분히 이루어지지 못하였다. 이로 인해 해당 문항에서 요구하는 고차적 데이터 평가 능력이 단기간 내에 유의미하게 향상되기에는 한계가 있었던 것으로 판단된다.

이외의 데이터 리터러시 향상을 측정하는 데이터 이해, 수집, 분석, 평가, 시각화 분류 문항에서 대부분의 문항이 평균의 향상을 보였으며 데이터 리터러시 전체 점수의 사전-사후 차이에 대한 효과 크기를 분석한 결과, 대응 표본 t-검정에서 Cohen’s $d = 0.92$ 로 산출되었다.

한편, 정량적 분석 결과를 보완적으로 해석하기 위해 학생들의 서술형 소감을 분석하였다. 분석 결과, 학습자들은 본 프로젝트 기반 데이터 분석 수업을 통해 데이터 전처리 및 분석, 데이터 간 관계 탐색, 시각화 및 해석 능력이 향상되었음을

인식하고 있었다. 예를 들어 한 학습자는 “통계 데이터를 가공하여 시각적인 그래프로 표현하고, 필요한 데이터만 추출하여 활용할 수 있게 되었다”고 응답하였으며(학습자 A), 이는 데이터 분석 및 시각화 영역에서 나타난 평균 점수 향상을 뒷받침하는 결과로 볼 수 있다. 또한 다른 학습자는 “많은 데이터를 한눈에 볼 수 있도록 정리하고, 변화 추세를 바로 파악할 수 있었다”고 응답하여(학습자 B), 시각화된 자료를 통해 데이터 간 관계를 해석하는 경험이 학습에 도움이 되었음을 보여준다. 이러한 학습자 인식은 데이터 분석 영역에서 나타난 통계적으로 유의미한 향상 결과와 일관된 방향을 보인다.

종합적으로, 정량적 분석과 정성적 분석 결과를 함께 고려할 때, 본 프로젝트 기반 데이터 분석 교육은 고등학생의 데이터 리터러시, 특히 데이터 이해·수집·분석 및 시각화 역량 향상에 효과적인 교수·학습 방안을 확인할 수 있다.

V. 결론 및 제언

본 연구에서는 데이터 분석을 위한 프로젝트 교육 프로그램을 고등학교 학생 대상으로 설계하고 적용하여 데이터 리터러시 향상에 대한 효과를 검증하였다.

인공지능 융합 교육에서 데이터 분석 부분의 교육 프로그램을 위한 다양한 수업자료를 제시하기 위해 교육 프로그램을 설계하였다. 연구 대상인 학생들이 데이터에 대해 이해하고 프로젝트 발표에서 실생활 데이터를 수집, 분석, 평가, 시각화하여 데이터 분석 프로젝트를 진행한 결과 작품을 발표

하는 교수학습 활동을 설계하였다. 14차시에 걸친 데이터 분석 프로젝트 교육 프로그램을 개발하였다.

본 연구에서는 설계된 프로그램을 실제 수업에 적용하고, 데이터 분석 프로젝트 학습은 학습자의 디지털 리터러시 능력 향상에 유의미한 영향을 미친 것으로 확인되었다. 이는 2022 교육과정의 데이터 분석 교육에 효과적으로 적용되기 위해 데이터를 이해하고, 수집, 분석, 평가, 시각화 능력 즉 데이터 리터러시 능력이 필요하다는 점과 프로젝트 수업을 위해 다양한 교육 데이터가 필요하다는 점, 수업을 위한 데이터 분석 수업자료 및 교사 연구가 필요하다는 점을 감안할 때 교육의 대안으로서 본 연구가 유용할 수 있다는 점을 시사한다.

다만 본 연구의 실험 집단은 프로그래밍을 경험해본 학생들로 구성되었지만, 데이터 분석을 위해 필요한 프로그래밍 능력을 갖추고 있는 학생 대상으로 수업을 일반화하는데 한계가 있다. 또한 추후 연구에서는 30명 이상의 참여자를 확보하고, 실험 집단과 비교집단을 구성하여 연구 결과에 대한 요인들을 좀 더 체계적으로 분석할 필요가 있다. 향후 본 연구가 디지털 인재 양성을 위한 인공지능 융합 교육 프로그램 개발에 기여할 수 있기를 기대한다.

감사의 글

제1저자의 석사 학위 논문에 바탕으로 본 연구 결과를 도출할 수 있도록 연구에 참여한 학생 및 전문가에게 감사를 드린다.

참고문헌

- [1] 4th Industrial Revolution Commission, 21st Meeting of the 4th Industrial Revolution Commission: National Data Policy Direction, Seoul, 2021.
- [2] B. Ryu, Customization and Choice: Digital-Based Educational Innovation, Korean Educational Development Institute, Jincheon, 2023.
- [3] The Ministry of Education, The National Curriculum for the Primary and Secondary Schools (Ministry of Education Notice 2020-236), 2022.
- [4] The Ministry of Education, The National Curriculum for the Primary and Secondary Schools (Ministry of Education Notice 2022-33), 2020.
- [5] M. G. Kim, Development and Application of Data Visualization Education using Python and Their Effects For Sixth Grade in Elementary School, Master's Thesis, Jeju National University, Jeju, 2020.
- [6] J. H. Oh, Development and Application of the Program for Data Education Using Python, Master's Thesis, Daegu University of Education, Daegu, 2023.
- [7] E. J. Lee, The Effects of AI-Based Data Analysis Education on Convergent Thinking Ability and Data Literacy of General High School Students, Master's Thesis, Kongju National University, Kongju, 2021.
- [8] H. S. Bae, "Educational Implications of Data Literacy in Social Studies," *Theory and Research in Citizenship Education*, Vol. 51, No. 1, pp. 95-120, 2019. <http://dx.doi.org/10.35557/trce.51.1.201903.004>
- [9] S. C. Lee and T. Y. Kim, Exploring the Concept and Components of Data Literacy in Computer Education, in *Proceedings of the Korean Association of Computer Education*, Vol. 23, No. 2, pp. 33-36, 2019.
- [10] S. Park, Development and Application of AI Convergence Education Program to Improve Data Literacy, Master's Thesis, Seoul National University of Education, Seoul, 2023.
- [11] G. Lee, Development of Data Literacy Education Program Based on Cooperative Learning to Improve Computational Thinking and Cooperative Communication Competency, Master's Thesis, Korea National University of Education, Cheongju, 2023.
- [12] J. Y. Seo, "A Case Study on Teaching and Learning Methods of Software Education for Data Analysis Problem Solving," *Journal of Digital Contents Society*, Vol. 20, No. 9, pp. 1953-1960, 2019.
- [13] H. S. Kang, J. M. Lee, and H. Kim, "A Study on Computer Programming Education Model Based on Python," *Journal of Digital Contents Society*, Vol. 21, No. 4, pp. 693-700, 2020.
- [14] M. Lim, "Development of Artificial Intelligence Education Using Big Data," *Journal of Digital Contents Society*, Vol. 26, No. 12, pp. 3591-3597, 2025.



김지현(Ji Hyeon Kim)

2018년 : 제주대학교 사범대학 컴퓨터교육과(학사)
2024년 : 제주대학교 교육대학원 인공지능융합교육전공(석사)

2023년~현 재: 제주대학교사범대학부설고등학교 교사
※관심분야 : 데이터 분석, 인공지능융합교육



김한일(Hanil Kim)

1988년 : 서울대학교 컴퓨터공학과(공학사)
1990년 : 서울대학교 컴퓨터공학과(공학석사)
1995년 : 서울대학교 컴퓨터공학과(공학박사)

1995년~현 재: 제주대학교 컴퓨터교육과 교수
2020년~현 재: 제주대학교 교육대학원 인공지능융합교육전공 교수
※관심분야 : 컴퓨터교육, 운영체제, 시스템 소프트웨어



김성백(Seong Baeg Kim)

1989년 : 서울대학교 컴퓨터공학과(공학사)
1991년 : 서울대학교 컴퓨터공학과(공학석사)
1995년 : 서울대학교 컴퓨터공학과(공학박사)

1996년~현 재: 제주대학교 컴퓨터교육과 교수
2020년~현 재: 제주대학교 교육대학원 인공지능융합교육전공 교수
※관심분야 : 컴퓨팅시스템, 인공지능융합 교육, 컴퓨터교육 등