

컴퓨터교육학회 논문지 2026년 제29권 제1호
https://doi.org/10.32431/kace.2026.29.1.004



2022 개정 교육과정 ‘데이터 과학’ 과목 수업에 관한 고등학생 인식 분석*

Analysis of High School Students’ Perceptions of the ‘Data Science’ Subject in the 2022 Revised Curriculum

전현희[†] · 김자미^{††}
Hyunhee Jun[†] · Jamee Kim^{††}

요약

2022 개정 정보과 교육과정에 신설된 ‘데이터 과학(Data Science)’ 과목의 교육과정을 기반으로 수업을 설계·운영하고, 고등학교 현장에서의 실행 가능성을 검증하기 위한 기초자료를 마련하고자 하였다. 서울 지역 고등학교 재학생 109명을 대상으로 도입 단계를 달리한 세 가지 프로그램을 운영하여 결과를 분석하였다. 모든 집단에서 ‘데이터 과학’ 과목의 교육적 가치가 높게 인식되었으며, 시각화 도구를 활용한 집단의 이해도가 상대적으로 높게 나타났다. 실습 시간의 배분, 선행 학습 경험과의 연계, 교수 설계 방식이 학습자 인식에 영향을 미치는 요인으로 확인되었다. 본 연구는 ‘데이터 과학’ 과목을 실제 수업에 적용하여 교육과정의 실행 가능성과 교수·학습 설계 방향을 탐색한 점에 의의가 있다.

주제어 컴퓨터 교과교육, 데이터 과학, 2022 개정 교육과정, 학습자 인식 분석, 고등학교 교육과정, 데이터 리터러시

ABSTRACT

Based on the Data Science curriculum newly introduced in the 2022 Revised Informatics Curriculum, this study aimed to design and implement lessons and to provide foundational data for verifying their feasibility in high school settings. Three differentiated programs with varying introductory stages were implemented for 109 high school students in Seoul, and the results were analyzed. All groups perceived the educational value of the Data Science subject to be high, and the group utilizing visualization tools showed a relatively higher level of understanding. The distribution of practice time, connection with prior learning experiences, and instructional design approaches were identified as factors influencing learners’ perceptions. This study is significant in that it applied the newly introduced subject to actual classroom instruction to explore the feasibility.

Keywords Computer Education, Data Science, 2022 Revised Curriculum, Learner Perception Analysis, High School Curriculum, Data Literacy

†정회원 고려대학교 컴퓨터학과 박사과정
††중신회원 고려대학교 교육대학원 컴퓨터교육전공
부교수(교신저자)
논문투고 2025년 07월 15일
심사완료 2025년 11월 22일
게재확정 2025년 12월 18일
발행일자 2026년 01월 30일

* 본 논문은 제1저자의 고려대학교 교육대학원 석사학위
논문 일부를 수정하여 정리한 것임.

1. 서론

디지털 기술의 발전과 더불어 사회 전반에 데이터의 중요성이 강조되고 있다. 2002년 디지털 데이터가 아날로그 데이터를 추월한 이후[1], 데이터는 ‘21세기의 원유’로 불리며 핵심 자원으로 인식되고 있다[2].

산업·행정·교육 등 다양한 영역에서 데이터 기반 의사결정이 확산되면서, 데이터를 이해하고 활용하는 능력인 데이터 리터러시(data literacy)는 현대 시민이 갖추어야 할 필수 역량으로 요구되고 있다.

OECD(2018)는 데이터 리터러시를 모든 학습자에게 요구되는 핵심 자질로 제시했으며[3], UNESCO(2021)는 정보를 비판적으로 해석·활용하는 능력을 시민 기본 소양으로 강조하였다[4]. 데이터 리터러시는 전문가에 국한되지 않고 모든 사회 구성원이 갖추어야 할 보편적 역량으로 인식되고 있다.

교육 분야에서도 데이터 리터러시 함양을 위한 체계적인 교육의 필요성이 제기되고 있으며 컴퓨팅 교육 분야에서는 ‘데이터 과학’을 독립된 교육 영역으로 포함하는 국제 표준이 마련되었다.

2020년 Association for Computing Machinery(ACM)와 Institute of Electrical and Electronics Engineers(IEEE)는 『Computing Curricula 2020』(CC2020)을 통해 ‘데이터 과학’을 컴퓨팅 교육과정의 독립 영역으로 제시하였으며[5], 2021년 발표된 『Computing Competencies for Undergraduate Data Science Curricula』(DS2021)에서는 고등교육 단계의 ‘데이터 과학’ 교육 기준을 구체화하였다[6]. 이 기준은 ‘데이터 과학’ 교육과정의 공통 기반을 제공하고, 각국의 교육 체계 설계에 방향성을 제시하였다.

국내에서는 2022 개정 교육과정을 통해 모든 학생이 갖추어야 할 핵심 역량으로 데이터 활용 능력을 포함한 디지털 기초소양을 제시하였다. 개정 교육과정에 기반하여 중·고등학교 정보 과목은 ‘자료와 정보’영역을 ‘데이터’중심으로 재구성하였으며, 고등학교 진로 선택 과목으로는 ‘데이터 과학’을 신설하였다. 이 과목은 데이터의 역할과 가치를 이해하고 데이터 기반 문제 해결을 통해 합리적 의사 결정 역량을 기를 수 있도록 구성되었다[7].

‘데이터 과학’은 신설 과목으로서 실제 적용 사례가 아직 제한적이다. 2025학년도 적용 초기 단계로, 현장 기반 실증 연구는 축적 과정에 있다. 기존 선행 연구는 다양한 교육적 접근 방식을 통해 교육 효과를 확인하였으나 학습자의 인식을 중요도·필요도·난이도·이해도의 측면에서 종합적으로 분석한 사례는 거의 보고되지 않았다.

본 연구는 2022 개정 교육과정의 ‘데이터 과학’과목 내용 체계를 기반으로 수업을 설계·운영하고, 학습자의 인식을 중요도·필요도·난이도·이해도 측면에서 정량적으로 분석하는 것을 목적으로 하였다. 학습 요소별 인식 분석을 중심으로 수행하였으며, 집단 간 비교 결과는 학습자의 인식 특성을 심화하여 이해하기 위한 보조적 해석 자료로 활용하였다. 이를 통

해 ‘데이터 과학’ 교육의 적용 가능성을 확인하고, 향후 교수·학습 설계를 위한 기초자료를 마련하고자 한다.

2. 관련 연구

2.1 ‘데이터 과학’의 변화

‘데이터 과학’은 전통적 통계학에 기반하여 출발하였으며, 기술의 발전과 학제 간 통합을 거치며 융합적이고 실천 중심의 지식 체계로 발전하였다. 초기에는 통계 분석 중심의 학문으로 인식되었으나, 정보 기술의 급속한 발전과 더불어 다양한 분야의 방법론이 결합하면서 분석 대상과 범위가 확대되었다[6].

Tukey(1962)는 통계학의 한계를 지적하며 데이터 분석을 경험 기반 과학으로 제시하였고[8], Naur(1974)는 ‘데이터 과학’을 데이터가 확정된 이후 이를 다루는 과학으로 정의하였다[9, 10].

1990년대에는 통계학과 데이터 분석이 별도로 발전하여 과학적 진보에 기여하였으며, 동시에 두 분야의 통합을 위한 새로운 접근이 요구되었다.

Wu(1997)는 “Statistics = Data Science”를 제안하며 통합 논의의 기초를 마련하였고[11], Hayashi(1998)는 통계학과 데이터 분석을 포괄하는 개념으로서 ‘데이터 과학’을 제시하였다[12].

Cleveland(2001)는 기존 통계학의 범위를 다학제적 탐구, 모델과 방법, 컴퓨팅, 교육학, 도구 평가, 이론의 여섯 영역으로 확장하고 이를 ‘데이터 과학’이라 명명하였다[13]. 이어 Breiman(2001)은 데이터 분석 문화를 추론 중심과 예측 중심으로 구분하며, 전통적 모형 기반 접근의 한계를 지적하고 블랙박스 모델을 활용한 예측 중심 접근이 ‘데이터 과학’에 적합함을 주장하였다[14].

2000년대 이르러 데이터 처리 기술의 급속한 향상이 이러한 전환을 가속화하였으며, Dhar(2013)은 방대한 데이터로부터 일반화 가능한 지식을 도출하고 예측을 통해 문제를 해결하는 체계적 연구 영역으로서 ‘데이터 과학’을 정의하였다[15].

Donoho(2017)는 빅데이터 환경을 반영한 Greater Data Science 프레임 제안하였으며[16], Brodie(2024)는 ‘데이터 과학’을 존재론·인식론·방법론을 아우르는 총체적 학문으로 재개념화하였다[17].

‘데이터 과학’은 단순한 도구적 기술 활용을 넘어 현실 문제 해결을 지향하는 융합적·실천 중심의 지식 체계로 진화하였다. 이러한 학문적 전개는 교육 영역에도 반영되어, 분석 기법 전달을 넘어 복합적 사고와 맥락 통합을 강조하는 ‘데이터 과학’교육의 필요성을 제기하고 있다.

2.2 ‘데이터 과학’ 교육과정

‘데이터 과학’이 독립된 학문 분야로 정립되면서, 국제적으로 교육과정의 표준화 논의가 본격화되었다. ACM과 IEEE는 『Computing Curricula 2020』(CC2020)을 통

해 ‘데이터 과학’을 컴퓨팅 교육과정의 독립 영역(DS202x)으로 구분하였으며[5], 이후 발표된 『Computing Competencies for Undergraduate Data Science Curricula』(DS2021)에서는 컴퓨터 과학, 통계학, 도메인 지식이 융합된 실전 중심 학문으로 정의하였다[6].

DS2021은 핵심 역량과 성취 기준을 지식·기술·태도의 세 범주로 구조화하고, 데이터 수집·해석·커뮤니케이션·윤리에 이르는 전 과정을 포괄하는 목표를 제시하였다. 이러한 구성은 ‘데이터 과학’학습을 기술 습득을 넘어, 데이터 활용의 전 과정에서 비판적 사고와 윤리적 판단을 함께 기르는 통합적 역량 교육으로 확장하였다는 점에서 의미가 있다.

국제적 논의는 국내 교육과정 개정에도 영향을 미쳤다. 양혜지 외(2021)는 DS202x 체계를 분석하여 국내 교육과정과의 연계 가능성을 검토하였으며, 데이터 적용 역량, 데이터 문해력, 윤리 교육, 컴퓨팅 기초, 유연한 교육과정 설계의 필요성을 제안하였다[18].

이러한 연구와 논의의 흐름은 2022 개정 교육과정에 반영되어, 고등학교 진로 선택 과목으로 ‘데이터 과학’이 신설되었다. 이 과목은 디지털 사회에서 데이터의 가치와 문제 해결의 중요성을 인식하고 합리적 의사 결정과 디지털 시민성을 함양할 수 있도록 융합 역량 중심으로 설계되었다[7]. 학습자는 데이터 분석, 모델링, 비판적 해석 과정을 통해 다양한 분야의 문제를 해결하며 데이터 기반 의사 결정 능력을 함양하도록 구성되었다.

2.3 데이터 교육 관련 연구

국내 고등학생 대상 데이터 활용 교육 연구는 크게 세 가지로 구분된다.

첫째, 교과 수업 적용 사례로서 김현경 외(2022)는 물리 수업에 파이썬(Python) 기반 데이터 분석·시각화를 적용하여 과학적 사고력 향상을 검증하였으며[19], 손미현(2023)은 질문 생성 탐구 수업에 데이터 분석 활동을 결합하여 자기 주도 탐구와 질문 생성이 촉진됨을 확인하였다[20]. 안혜정 외(2024)는 화학 탐구 수업에서 ‘데이터 과학’과정을 반영한 프로그램이 과학적 추론과 의사소통 능력 향상에 효과적임을 제시하였다[21].

둘째, 비교과 및 융합 프로그램 연구에서는 김세민 외(2022)가 대중문화·스포츠 데이터를 활용한 12주 과정에서 학습 동기와 실무 적용 역량이 강화됨을 보고하였고[22], 이은정 외(2024)는 인공지능 기초 교육과정을 재구성한 ‘데이터 과학’수업을 통해 융합적 사고력과 데이터 리터러시가 향상됨을 확인하였다[23].

셋째, 토론 기반 학습 사례로 송유경 외(2021)는 Data-Driven Discussion 모형을 적용한 공공 데이터 기반 토론 활동이 데이터 해석·근거 제시·반박 전략 등 의사소통 역량을 증진한다고 보고하였다[24].

국내 연구들은 교과 핵심 역량, 융합적 사고력, 데이터 리터러시, 자기 주도 탐구 능력 향상 등 다양한 교육 효과를 실증적으로 제시하였으며, 학습 성과나 효과 검증에 초점을

두고 있다. 학습자의 인식을 중요도·필요도·난이도·이해도 의 다차원적 관점에서 정량적으로 분석한 연구는 거의 보고되지 않았다. 이에 본 연구는 2022 개정 교육과정의 ‘데이터 과학’과목 내용 체계를 기반으로 학습자의 인식을 네 가지 측면에서 분석하여, ‘데이터 과학’수업 설계에 활용할 수 있는 기초자료를 마련하고자 한다.

3. 연구 방법

3.1 연구 절차

본 연구는 ‘데이터 과학’수업의 설계부터 효과 분석까지 다섯 단계로 구성하였다. 각 단계의 세부 내용과 흐름은 Fig. 1에 제시하였다.

Stage	Sub-phase	Description
Phase 1	Lesson Design	<ul style="list-style-type: none"> Analysis of the 2022 Revised Curriculum's 'Data Science Content Structure Development of a Data Science Education Program
Phase 2	Tool Development	<ul style="list-style-type: none"> Survey Development <ul style="list-style-type: none"> - Drafting and Refinement of Survey Items (Reflecting Curriculum and Class Design)
Phase 3	Lesson Implementation	[Grouping Based on High School Lesson Design] <ul style="list-style-type: none"> Group A <ul style="list-style-type: none"> - Period: Mar 25 – Jul 15, 2023 (20 hours) - Participants: 60 students Group B <ul style="list-style-type: none"> - Period: Nov 4 – Dec 30, 2023 (28 hours) - Participants: 20 students Group C <ul style="list-style-type: none"> - Period: Mar 23 – Jul 13, 2024 (20 hours) - Participants: 40 students
Phase 4	Survey Administration	<ul style="list-style-type: none"> Survey and Data Collection : Total Participants 109 <ul style="list-style-type: none"> - Group A: 58 - Group B: 14 - Group C: 37
Phase 5	Analysis of Results	<ul style="list-style-type: none"> Data Preprocessing and Filtering Descriptive Statistics and Frequency Analysis Difference Test: Between-Group : <ul style="list-style-type: none"> - One-Way ANOVA (Groups A-C) <ul style="list-style-type: none"> · Welch ANOVA if variance unequal - Paired Samples t-test · Importance vs. Necessity · Difficulty vs. Understanding

Figure 1. Research Procedure

첫째, 2022 개정 교육과정의 ‘데이터 과학’과목에 제시된 내용 체계를 분석하고 이를 기반으로 교육 프로그램을 설계하였다.

둘째, 수업 내용과 연구목적에 적합하도록 설문 문항을 개발한 후 문항의 명확성과 응답 가능성을 검토하여 최종 설문지를 확정하였다.

셋째, 고등학교 수업 운영 환경을 고려하여 A 집단(20차시), B 집단(28차시), C 집단(20차시)으로 구분하고 도입 단계에서 기초 학습 내용을 달리하여 수업을 실시하였다.

넷째, 수업 종료 후 총 120부의 설문지를 회수하였으며 결측치가 포함된 설문을 제외한 A 집단 58명, B 집단 14명, C 집단 37명 등 총 109명의 응답을 최종 분석 대상으로 선정하였다.

다섯째, 응답 자료를 정제된 뒤 기술통계 분석을 실시하였다. 주요 변인의 분포 특성을 확인하기 위해 Shapiro-

Wilk 검정을 적용하였으며, 난이도 변인은 정규성을 충족하였으나 중요도, 필요도, 이해도 변인은 정규성을 충족하지 않았다. 이에 따라 집단 간 차이 검증에서는 t-검정과 일원분산분석(ANOVA)을 실시하였고, 동일 학습 요소 내에서 중요도-필요도, 난이도-이해도의 인식 차이는 동일 응답자의 짝을 이루므로 대응표본 t-검정을 적용하였다.

세 집단 간 평균 차이 비교에서는 일원분산분석(ANOVA)을 활용하였으며, 등분산성이 가정되지 않는 경우 Welch의 분산분석을 병행하여 해석의 신뢰성을 확보하였다. 관심도 수준 문항은 자료 정리의 용이성을 위해 3단계로 재범주화하여 분석에 활용하였다. 자료 분석에는 IBM SPSS Statistics 29.0을 사용하였다.

3.2 연구 대상

본 연구는 2023년부터 2024년까지 두 개 학교에서 세 집단을 대상으로 ‘데이터 과학’수업을 운영하고 설문 조사를 실시하였다. 연구 대상은 서울 소재 S고등학교와 G고등학교 학생으로, A 집단 60명, B 집단 20명, C 집단 40명 등 총 120명으로 구성되었다. A·C 집단은 남학교 학생으로, B 집단은 여학생이 포함된 학교 학생으로 구성되어 성별 분포에 차이가 있었다. 세 집단의 수업은 같은 연구자가 진행하였으며, 연구자는 정보 교과와 데이터 분석 수업을 5년 이상 담당할 경험을 보유하고 있다.

결측치를 제외한 최종 표본은 A 58명, B 14명, C 37명으로 총 109명(N=109)이었다. 응답자의 다수는 남학생(94.5%)과 2학년(89.0%)이었으며, 희망 전공은 공학계열(58명)이 가장 높았다. 또한 대부분은 프로그래밍 경험(94.5%)을 가진 것으로 확인되었고, 절반 이상은 ‘데이터 과학’학습 경험(57.8%)이 있는 것으로 분석되었다. 연구 대상의 특성은 Table 1에 제시하였다.

Table 1. General Characteristics of Participants

Category	Item	Description	Frequency(%)
Basic	Gender	Male	103(94.5)
		Information	6(5.5)
	Grade	10th Grade (1st year)	12(11)
		11th Grade (2nd year)	97(89)
	Preferred Major	Humanities & Social Sci.	6
		Field	34
*		1	
Engineering		58	
Programming Experience	Experience	Yes	103 (94.5)
		No	6 (5.5)
Data Science Experience	Experience	Yes	63 (57.8)
		No	46 (42.2)

* : Multiple responses allowed

3.3 교육 프로그램 구성

본 연구는 2022 개정 교육과정의 ‘데이터 과학’ 과목 내용 체계에 기반하여 교육 프로그램을 설계하였다. 교육과정에서 제시한 네 영역(① ‘데이터 과학’의 이해 ② 데이터 준비와 분석 ③ 데이터 모델링과 평가 ④ ‘데이터 과학’프로젝트)을 지식·이해, 과정·기능, 가치·태도 요소와 통합하여 도입 - 전개(준비·분석·평가) - 프로젝트 세 단계로 구성하였다.

단계별 프로그램 구성은 Table 2에 제시하였다.

Table 2. Educational Programs by Group

Category		A	B	C
Foundation	Program (hours)	Basic Syntax (4)	Visualization (12)	Basic Statistics (4)
	Tools	Python (GoogleColab)	Tableau Public	Statistical Worksheets
Development	Program (hours)	Data Science Curriculum-Based Program (16)		
	Tools	Python (GoogleColab)		
Project*		O		

* : Team-based project conducted outside of class

세 집단의 수업은 도입 단계에서 기초 학습 내용을 달리 하여 설계하였다. A 집단은 파이썬 기초 문법 학습, B 집단은 데이터 시각화와 해석 경험, C 집단은 기초 통계 개념 학습을 중심으로 구성하였다.

전개 단계에서는 세 집단 모두 데이터 분석 전 과정을 동일하게 경험하도록 하였다. 파이썬을 활용해 데이터 수집·정제, 탐색적 분석, 시각화, 결과 해석 및 평가를 진행하였으며 총 16차시로 운영되었다. 학습자는 단계별로 데이터를 준비·분석하고 최종적으로 파이썬 기반 결과를 도출하였다. 세부 운영 프로그램은 Table 3에 제시하였다.

Table 3. Data Science Curriculum-Based Program

Stage	Learning Contents	Tools
Introduction	·Necessity of data literacy ·Exploration of problem-solving cases ·Understanding types and meanings of data	PPT, Worksheets
Preparation	·Data collection - use of public datasets ·Data preprocessing - data type identification, - handling of outliers and missing values	Python (Google Colab)
Analysis	·Exploratory data analysis - descriptive statistics, visualization ·Confirmatory analysis - correlation, hypothesis setting	Python (Google Colab)
Evaluation	·Model application and evaluation - case studies of regression - classification, and clustering	

3단계에서는 팀을 구성하여 프로젝트를 수행하는 활동으로 진행되었으며 수업 외 시간에서 운영되었다.

3.4 연구 도구

2022 개정 교육과정의 ‘데이터 과학’ 과목 내용 체계를 바탕으로 설문지를 개발하였다. 설문은 고등학생의 수업 경험과 인식을 측정하기 위한 도구로, 총 4개 영역(① 개인 배경 ② 프로그래밍·‘데이터 과학’ 경험 ③ 내용 체계 인식 ④ 수업 경험)으로 구성하였다. 각 문항은 Likert 5점 척도로 응답하도록 설계하였으며 총 54문항으로 구성되었다. 설문 문항의 세부 구성은 Table 4에 제시하였다.

Table 4. Structure of the Survey Items

Subcategory Gender Grade	Subcategory	Number of Items			
Data Preparation	Gender	3			
	Grade				
	Intended Major Field*				
Data Analysis	Programming Experience*	4			
	Experience and Interest in Data Analysis				
Data Evaluation	Core Areas of Data Science	Importance	Necessity	Difficulty	Comprehension
	Data Preparation	3	3	3	3
	Data Analysis	6	6	6	6
	Data Evaluation	2	2	2	2
Experience and Perception of Class	Perception of Data Science Education*	2			
	Experience and Perception of Class	1			
Total		54			

* : Multiple responses allowed

설문 문항 중 ‘데이터 과학’ 교육과정 인식 영역은 학습 단계(데이터 준비·데이터 분석·데이터 평가)로 세분화하였다. 동일 문항에 대해 중요도·필요도·난이도·이해도를 측정하도록 설계하였으며, 예시 문항은 Table 5에 제시하였다.

Table 5. Example Survey Items Related to Data Science Learning Stages

Stage	Example Item	Measurement Indicators
Data Preparation	·Ability to collect data without bias	Importance Necessity Difficulty Comprehension
Data Analysis	·Ability to observe and understand collected data from multiple perspectives ·Ability to interpret data numerically through statistical analysis	
Data Evaluation	·Ability to devise creative solutions based on analysis results	

내용 타당도는 2022 개정 교육과정 ‘데이터 과학’ 과목을 개발한 연구진 5인을 대상으로 CVR(Content Validity Ratio)을 통해 내용 타당도를 확보하였다. 설문 도구의 신뢰도는 Cronbach’s α .87로 확인되어 신뢰할 수 있는 수준임이 검증되었다. 또한 ‘필요도’와 ‘중요도’의 개념을 구분하여 설문을 진행하였다. 필요도는 학생들이 자신이 하고 싶은 일을 수행하는 과정에서 ‘데이터 과학’이 실제적으로 도움이 될 수 있는지를 판단하는 측면을 의미하며, 중요도는 특정 의사결정이나 진로 활동과 관련하여 ‘데이터 과학’이 가지는 가치와 영향력의 크기를 평가하는 개념으로 정의하였다.

수업에 참여한 학생들의 데이터는 모두 익명화하였고, 연구에 참여할 것인지에 대해서는 설문 진행 전에 본 연구의 목적 등을 설명하였다. 학부모의 동의가 확인된 학생들의 자료를 연구에 활용하였다. 동의서에는 연구목적, 참여의 자율성, 자료 보관 및 파기 기준을 명시하였으며, 연구 종료 후 일정 기간 보관 뒤 폐기하였다.

4. 연구 결과

‘데이터 과학’ 수업에 대한 학습자의 주요 인식 변수(중요도, 필요도, 난이도, 이해도)의 기술통계 결과는 Table 6과 같다.

Table 6. Descriptive Statistics of Key Perception Variables

Group	Importance	Necessity	Difficulty	Comprehension
	M(SD)	M(SD)	M(SD)	M(SD)
A	4.44(0.39)	4.46(0.39)	3.52(0.66)	3.95(0.58)
B	4.58(0.37)	4.53(0.45)	3.19(0.78)	4.25(0.56)
C	4.61(0.32)	4.61(0.33)	3.39(0.66)	4.17(0.60)
F-value	2.25	1.63	1.53	2.52

* $p < .05$, ** $p < .01$, *** $p < .001$

학생들은 ‘데이터 과학’수업에 대해 전반적으로 중요성과 필요성을 높게 인식하였다. 난이도는 대체로 중간 수준으로 평가되었으며, 이해도는 비교적 긍정적인 반응을 보였다. 네 가지 인식 변인 모두에서 통계적으로 유의한 집단 간 차

이는 나타나지 않았다.

4.1 ‘데이터 과학’수업 단계에 대한 인식 분석 결과

‘데이터 과학’수업은 ‘데이터 준비’, ‘데이터 분석’, ‘데이터 평가’의 세 단계로 구성되어 있다. 본 절에서는 각 단계에 대해 중요도, 필요도, 난이도, 이해도를 기준으로 학습자의 인식을 분석하였다.

4.1.1 수업 단계에 대한 중요도·필요도 인식 분석

‘데이터 준비’와 ‘데이터 평가’ 단계는 세 집단 모두에서 중요도와 필요도 평균이 4.5 이상으로 전반적으로 높은 수준으로 인식되었으며 집단 간 유의미한 차이는 나타나지 않았다. 반면 ‘데이터 분석’ 단계에서는 두 지표 모두에서 통계적으로 유의한 집단 간 차이가 확인되었으며($p < .05$), B 집단은 중요도와 필요도 간 인식 수준에서도 차이를 보였다. 수업 단계별 중요도와 필요도에 관한 인식 결과는 Table 7 과 같다.

Table 7. Importance and Necessity by Lesson Stage

Stage	Group	Importance	Necessity	t-value
		M(SD)	M(SD)	
Data Preparation	A	4.52(0.44)	4.50(0.45)	0.38
	B	4.50(0.43)	4.60(0.35)	1.18
	C	4.58(0.40)	4.57(0.39)	0.42
	F-value	0.28	0.47	-
Data Analysis	A	4.19(0.50)	4.26(0.48)	1.43
	B	4.33(0.44)	4.46(0.42)	2.35*
	C	4.48(0.44)	4.51(0.42)	0.64
	F-value	4.15*	3.97*	-
Data Evaluation	A	4.60(0.50)	4.61(0.48)	0.12
	B	4.75(0.44)	4.54(0.42)	2.12
	C	4.74(0.44)	4.72(0.42)	0.36
	F-value	1.07	0.70	-

* $p < .05$, ** $p < .01$, *** $p < .001$

‘데이터 분석’ 단계에서 수업 구성 방식에 따른 인식 차이가 드러난 것으로 해석된다.

4.1.2 수업 단계에 대한 난이도·이해도 인식 분석

‘데이터 준비’ 단계는 집단 간 차이는 없었으나, A 집단에서만 이해도가 난이도보다 유의하게 높게 나타났다($p < .001$). ‘데이터 분석’ 단계는 이해도 수준에서 집단 간 유의한 차이가 확인되었다. ‘데이터 평가’ 단계는 세 집단 모두 난이도보다 이해도를 유의하게 인식하였으며, 집단 간 차이는 통계적으로 유의하지 않았다. 학습자들은 과제 수행의 난이도보다 개념 이해를 높게 인식하였으며, 수업 설계 방식에 따라 일부 차이가 발생한 것으로 해석된다.

수업 단계별 난이도와 이해도에 대한 인식 분석 결과는 Table 8과 같다.

Table 8. Difficulty and Comprehension by Lesson Stage

Stage	Group	Difficulty	Comprehension	t-value
		M(SD)	M(SD)	
Data Preparation	A	3.26(0.79)	4.05(0.62)	5.57***
	B	3.17(0.81)	4.33(0.58)	1.98
	C	3.24(0.71)	4.19(0.66)	1.09
	F-value	0.08	1.34	-
Data Analysis	A	3.58(0.60)	3.83(0.61)	4.07**
	B	3.23(0.97)	4.17(0.57)	2.07*
	C	3.35(0.65)	4.15(0.63)	2.90*
	F-value	2.30	3.91*	-
Data Evaluation	A	3.73(1.09)	3.96(0.83)	5.85***
	B	3.18(0.99)	4.25(0.67)	5.54***
	C	3.57(0.99)	4.18(0.73)	2.83**
	F-value	1.61	1.32	-

* $p < .05$, ** $p < .01$, *** $p < .001$

4.2 ‘데이터 과학’ 학습 요소에 대한 인식 분석 결과

각 단계에 포함된 세부 학습 요소를 기준으로 학습자의 중요도, 필요도, 난이도, 이해도 인식을 정량적으로 분석하였다. 응답 결과는 수업 집단 간 비교 및 항목별 분석을 통해 인식 특성과 반응 경향을 확인하였다.

4.2.1 데이터 준비 학습 요소에 대한 인식 분석

데이터 준비 단계는 ‘데이터 과학’학습의 출발점으로, 데이터 수집, 신뢰성 검토, 컴퓨팅 도구 활용 등의 기초적 학습 요소로 구성되어 있다. 이 단계에 대한 학습자의 인식 정도를 비교하여 집단 간 차이와 항목별 반응 경향을 분석하였다.

첫째, 중요도 및 필요도 분석 결과 세 학습 요소 모두 평균 4.0 이상으로 전반적으로 높은 수준을 보였으며, 학습자들은 해당 학습요소에 대한 중요도와 필요도를 전반적으로 높게 인식하는 것으로 나타났다.

‘데이터 수집’은 중요도·필요도 간 유의한 차이가 없었으며, 집단 간 차이도 나타나지 않았다.

‘데이터 신뢰성’에서는 A 집단에서 중요도가 필요도보다 유의하게 높았으며($p < .01$), ‘컴퓨팅 활용 능력’에서는 B 집단이 필요도를 상대적으로 높게 평가하였다($p < .05$).

분석 결과는 Table 9에 제시하였다.

Table 9. Group Perception of Importance and Necessity for Data Preparation Elements

Stage	Learning Elements	Group	Importance	Necessity	t-value
			M(SD)	M(SD)	
Data Preparation	Data Collection	A	4.57(0.68)	4.62(0.62)	0.52
		B	4.64(0.50)	4.57(0.51)	0.56
		C	4.57(0.55)	4.57(0.50)	0.00
		F-value	0.09	0.11	-
	Data Reliability	A	4.76(0.51)	4.55(0.71)	2.69**
		B	4.86(0.36)	4.79(0.43)	1.00
		C	4.78(0.48)	4.76(0.49)	0.37
		F-value	0.24	1.66	-
	Computing Tool Utilization	A	4.22(0.77)	4.33(0.76)	1.00
		B	4.00(0.78)	4.43(0.65)	2.48
		C	4.38(0.64)	4.38(0.68)	0.00
		F-value	1.42	0.13	-

* p < .05, ** p < .01, *** p < .001

둘째, 데이터 준비 단계의 세부 학습 내용에 대한 난이도 및 이해도 인식 분석 결과 모든 항목에서 학습자들은 난이도보다 이해도를 더 높게 인식하였다.

‘데이터 수집’과 ‘데이터 신뢰성’ 항목은 세 집단 모두 통계적으로 유의한 차이가 나타났으며, ‘데이터 신뢰성’ 항목에서는 A 집단에서 가장 큰 차이가 확인되었다. ‘컴퓨팅 활용 능력’ 항목에서는 통계적으로 유의한 차이가 나타나지 않았지만, 전체적으로 이해도가 난이도보다 높게 인식되는 경향을 보였다. 분석 결과는 Table 10과 같으며 개념 중심 요소에 비해 도구 활용 요소를 상대적으로 더 어렵게 인식하고 있음을 해석할 수 있다.

Table 10. Group Perception of Difficulty and Comprehension in Data Preparation Elements

Stage	Learning Elements	Group	Difficulty	Comprehension	t-value
			M(SD)	M(SD)	
Data Preparation	Data Collection	A	3.09(1.11)	4.29(0.65)	6.51***
		B	3.29(1.20)	4.36(0.63)	2.79*
		C	3.14(0.89)	4.24(0.72)	5.63***
		F-value	0.20	0.16	-
	Data Reliability	A	2.74(1.07)	4.22(0.82)	7.25***
		B	2.71(0.91)	4.57(0.65)	5.95***
		C	2.81(1.10)	4.41(0.72)	6.55***
		F-value	0.64	1.42	-
	Computing Tool Utilization	A	3.95(0.96)	3.64(0.93)	1.55
		B	3.50(0.85)	4.07(0.73)	1.67
		C	3.78(0.89)	3.92(0.76)	0.68
		F-value	1.41	2.11	-

* p < .05, ** p < .01, *** p < .001

4.2.2 데이터 분석 학습 요소에 대한 인식 분석

데이터 분석 단계는 전처리, 모델링, 시각화, 통계 분석

등 수집된 데이터를 구조화하고 의미화하는 핵심 과정으로 구성된다. 학습자의 응답을 바탕으로 분석 학습 요소에 대한 학습자의 수용 수준과 인지적 부담을 파악하였다.

첫째, 데이터 분석 단계의 세부 학습 요소에 대해 학습자가 인식한 중요도와 필요도 수준을 분석한 결과 모든 항목에서 평균값이 4.0 이상으로 나타나 전반적으로 높게 인식되었다.

‘데이터 관찰 역량’, ‘데이터 시각화’, ‘통계 분석’ 항목은 세 집단 모두에서 중요도와 필요도가 고르게 높았고 집단 간 유의한 차이는 나타나지 않았다. 반면 ‘데이터 전처리’와 ‘데이터 모델 개념’에서는 중요도와 필요도 모두에서 집단 간 차이가 통계적으로 유의하게 나타났으며(p < .05), 수업 도입 단계에서 기초 통계를 진행한 C 집단이 두 항목에서 데이터 구조화 및 모델 개념에 대한 이해 필요성을 상대적으로 높게 인식하였다. 분석 결과는 Table 11과 같다.

Table 11. Group Perception of Importance and Necessity for Data Analysis Elements

Content	Learning Elements	Group	Importance	Necessity	t-value
			M(SD)	M(SD)	
Data Analysis	Data Observation Skills	A	4.33(0.69)	4.38(0.75)	0.57
		B	4.36(0.50)	4.50(0.65)	0.81
		C	4.54(0.65)	4.59(0.60)	0.19
		F-value	1.24	1.13	-
	Data Preprocessing	A	4.16(0.79)	4.21(0.74)	0.55
		B	4.21(0.43)	4.43(0.51)	1.88
		C	4.54(0.56)	4.54(0.51)	0.00
		F-value	3.73*	3.14*	-
	Data Modeling Concepts	A	3.93(0.77)	4.03(0.86)	1.06
		B	4.21(0.80)	4.36(0.50)	1.00
		C	4.43(0.65)	4.43(0.69)	0.00
		F-value	5.35**	3.34*	-
Use of Modeling Techniques	A	4.16(0.83)	4.09(0.88)	0.75	
	B	4.43(0.76)	4.36(0.63)	0.56	
	C	4.27(0.77)	4.46(0.69)	1.64	
	F-value	0.73	2.64	-	
Data Visualization	A	4.17(0.82)	4.34(0.74)	1.80	
	B	4.43(0.65)	4.64(0.63)	1.88	
	C	4.51(0.61)	4.41(0.64)	1.27	
	F-value	2.62	1.04	-	
Statistical Analysis	A	4.41(0.75)	4.48(0.66)	0.78	
	B	4.36(0.74)	4.50(0.65)	1.00	
	C	4.57(0.60)	4.65(0.54)	0.90	
	F-value	0.71	0.85	-	

* p < .05, ** p < .01, *** p < .001

둘째, 데이터 분석 학습 내용에 대한 난이도와 이해도 인식 분석 결과, 학습자들은 난이도보다 이해도를 더 높게 인식하고 있는 것으로 나타났다.

‘데이터 관찰 역량’과 ‘데이터 시각화’는 세 집단 모두에서 통계적으로 유의한 차이가 확인되었으며, 이해도가 난이

도보다 높게 인식되는 것으로 나타났다. 이 중 ‘데이터 관찰 역량’은 난이도에서 집단 간 차이가 유의하게 나타났으나 이 해도에서는 집단 간 차이가 없었다.

‘데이터 전처리’와 ‘데이터 모델 개념’은 C 집단에서 난이도와 이해도의 차이가 통계적으로 유의하게 나타났으며, ‘모델링 활용’과 ‘통계 분석’은 B 집단과 C 집단에서 유의한 차이가 확인되었다. A 집단은 네 항목 모두에서 유의한 차이가 발견되지 않았다.

전반적으로 학습자들은 데이터 분석 활동을 개념적 측면에서 비교적 명확히 이해하고 있는 것으로 나타났으며, 일부 요소에서는 집단 간 인식 양상이 다르게 분포하였다. 비교 분석 결과는 Table 12와 같다.

Table 12. Group Perception of Difficulty and Comprehension for Data Analysis Elements

Stage	Learning Elements	Group	Difficulty	Comprehension	t-value
			M(SD)	M(SD)	
Data Analysis	Data Observation Skills	A	3.59(0.92)	3.93(0.75)	5.18***
		Analysis	2.93(1.14)	4.29(0.83)	5.08***
		C	3.27(0.84)	4.22(0.71)	8.28***
		F-value	3.38*	2.31	-
	Data Preprocessing	A	3.76(0.88)	3.60(0.92)	0.89
		B	3.71(1.07)	4.21(0.80)	1.39
		C	3.59(0.90)	4.08(0.72)	2.70*
		F-value	0.37	5.20*	-
	Data Modeling Concepts	A	3.59(0.97)	3.71(0.82)	0.60
		B	3.36(1.34)	3.86(0.86)	0.98
		C	3.35(0.75)	4.00(0.85)	3.33**
		F-value	0.80	1.41	-
Use of Modeling Techniques	A	3.69(0.90)	3.69(0.84)	0.00	
	B	3.21(1.05)	4.36(0.63)	3.27**	
	C	3.54(0.69)	4.05(0.81)	2.99**	
	F-value	1.79	4.87**	-	
Data Visualization	A	3.21(1.07)	4.14(0.83)	4.66***	
	B	3.14(1.23)	4.14(0.77)	2.39*	
	C	3.11(1.07)	4.30(0.74)	5.13***	
	F-value	0.10	0.49	-	
Statistical Analysis	A	3.67(1.02)	3.90(0.85)	1.16	
	B	3.00(1.18)	4.14(0.77)	2.83*	
	C	3.24(1.09)	4.27(0.69)	4.33***	
	F-value	3.23	2.62	-	

* p < .05, ** p < .01, *** p < .001

4.2.3 데이터 평가 학습 요소에 대한 인식 분석

데이터 평가 단계는 분석된 결과를 기반으로 의미를 도출하고 이를 실제 상황에 적용하는 일련의 해석 및 활용 과정으로 정의된다. 데이터 평가 활동에 대한 이해 수준과 과제 인식 양상을 살펴보았다.

첫째, 데이터 평가 단계의 세부 학습 요소에 대한 중요도와 필요도 분석 결과 두 항목 모두에서 평균이 4.5 이상으로

나타나 전반적으로 높은 수준을 보였다.

‘결과 해석’은 세 집단 모두에서 중요도와 필요도가 높게 인식되었으나, 집단 내·집단 간 비교에서는 통계적으로 유의한 차이가 확인되지 않았다. ‘문제 해결’항목은 B 집단에서만 중요도와 필요도 간 인식 차이가 유의하게 나타났고(p < .05) 집단 간 비교에서는 통계적으로 유의하지 않았다. 데이터 평가 단계의 학습 요소는 세 집단 모두에서 중요하고 필요한 요소로 인식된 것으로 해석된다.

분석 결과는 Table 13에 제시하였다.

Table 13. Group Perception of Importance and Necessity for Data Evaluation Elements

Stage	Learning Elements	Group	Importance	Necessity	t-value
			M(SD)	M(SD)	
Data Evaluation	Interpretation of Results	A	4.72(0.56)	4.74(0.55)	0.27
		B	4.79(0.43)	4.64(0.63)	1.47
		C	4.81(0.46)	4.89(0.39)	1.78
		F-value	0.34	1.54	-
	Analysis-Based Problem Solving	A	4.48(0.73)	4.48(0.71)	0.00
		B	4.71(0.47)	4.43(0.85)	2.28*
	C	4.68(0.63)	4.54(0.73)	1.09	
	F-value	1.27	0.14	-	

* p < .05, ** p < .01, *** p < .001

둘째, 난이도와 이해도 분석 결과 ‘결과 해석’은 B 집단과 C 집단에서 이해도가 난이도보다 유의하게 높게 나타났으며(p < .01), 집단 간 비교에서는 유의하지 않았다.

‘문제 해결’은 세 집단 모두에서 난이도와 이해도의 차이가 통계적으로 유의하지 않았으며, 집단 간 비교 결과 역시 차이가 없었다.

분석 결과는 Table 14에 제시하였다.

Table 14. Group Perception of Difficulty and Understanding for Data Evaluation Elements

Stage	Content	Group	Difficulty	Comprehension	t-value
			M(SD)	M(SD)	
Data Evaluation	Interpretation of Results	A	3.60(1.17)	4.03(0.86)	1.93
		B	2.86(0.95)	4.29(0.73)	3.98**
		C	3.46(1.10)	4.22(0.71)	3.28**
		F-value	2.51	0.90	-
	Analysis-Based Problem Solving	A	3.86(1.19)	3.88(0.94)	0.08
		B	3.50(1.29)	4.21(0.80)	1.51
	C	3.68(1.16)	4.14(0.86)	1.78	
	F-value	0.64	1.35	-	

* p < .05, ** p < .01, *** p < .001

데이터 평가 단계는 일부 항목에서만 난이도와 이해도 차이가 확인되었으며, 전반적으로는 집단 간 유사한 인식 양상이 나타났다. 데이터 평가 활동이 학습자에게 비교적 친숙하게 인식되었고, 과제 수행의 난이도에 비해 이해 수준

이 상대적으로 높게 인식되었다.

4.3 ‘데이터 과학’ 수업 경험에 대한 분석

‘데이터 과학’ 수업을 통해 학습자가 인식한 경험적 효과를 세 가지 측면에서 분석하였다.

첫째 ‘데이터 과학’의 중요성에 대한 인식, 둘째 수업을 통해 향상된 능력에 대한 자기 평가, 셋째 프로젝트 기반 수업 방식의 필요성 인식이다.

각 항목은 객관식 문항의 응답 결과를 정리하여 집단별 차이와 응답 경향을 확인하였다.

4.3.1 데이터 과학의 중요성 인식

전체 109명 중 45명(41.3%)이 ‘데이터 기반으로 의사 결정을 할 수 있다’고 응답하여 가장 높은 빈도를 보였다. 집단별로는 A 집단 25명, B 집단 5명, C 집단 15명으로 나타났다.

그 외 ‘분석 방향에 따라 다른 결과가 도출될 수 있다’(28.4%), ‘홀여져 있던 데이터의 중요함을 알 수 있다’(17.4%)도 비교적 높은 응답을 보였다.

학습자들은 같은 데이터라도 분석 방향에 따라 다양한 결과가 도출될 수 있음을 인식하였으며, 이를 바탕으로 데이터 기반 의사 결정의 중요성을 확인한 것으로 해석된다. 분석 결과는 Table 15와 같다.

Table 15. Perceived Reasons for the Importance of Data Science

Item	A(%)	B(%)	C(%)	total(%)
Recognizes the Importance of Dispersed Data	8(13.8)	1(7.1)	10(27.0)	19(17.4)
Able to Conduct Different Analyses Based on Data Types	9(15.5)	3(21.4)	2(5.4)	14(12.8)
Understands That Different Analytical Approaches Yield Different Results	16(27.6)	5(35.7)	10(27.0)	31(28.4)
Able to Make Data-Driven Decisions	25(43.1)	5(35.7)	15(40.5)	45(41.3)
Others	0	0	0	0
Others	58	14	37	109
total	F-value	0.64	1.35	-

4.3.2 수업을 통해 향상된 능력 인식

A 집단은 ‘프로그래밍 능력’ 응답이 22명으로 가장 높았다. 파이썬 기반 실습 경험이 해당 능력 향상 인식에 영향을 준 것으로 해석된다.

B 집단은 응답 수는 적었으나 ‘데이터의 특성을 파악하고 분석 방법을 적용할 능력’(5명)과 ‘문제 해결 능력’(4명)에서 상대적으로 높은 비율을 보였다. 이는 시각화 도구 중심 수업의 효과로 볼 수 있다.

C 집단은 ‘데이터 리터러시 능력’(11명)과 ‘기초 통계 능력’(10명)에 대한 응답이 높았다. 통계 중심 활동이 능력

향상 인식에 영향을 미친 것으로 해석된다. 응답 결과는 Table 16에 제시하였다.

Table 16. Abilities Improved Through Participation in Data Science Class (Multiple Selections)

Item	A	B	C	total
Programming Ability	22	4	6	32
Data Interpretation Literacy	13	3	11	27
Ability to Identify Data Characteristics and Apply Appropriate Analysis Methods	16	5	10	31
Basic Statistical Ability to Interpret Numerical Data	13	1	10	24
Problem-Solving Ability Using Data	13	4	5	22

4.3.3 프로젝트 기반 수업 방식 필요성 인식

전체 109명 중 102명(93.5%)이 ‘필요하다’고 응답하여 프로젝트 기반 수업 방식의 필요성이 매우 높게 나타났다. 집단별로는 A 집단 91.4%, B 집단 100%, C 집단 94.6%로 모두 높은 수준이었다. 학습자들은 프로젝트 중심 수업이 ‘데이터 과학’ 개념 이해와 실제 적용에 효과적인 학습 방식이라고 인식한 것으로 해석된다. 분석 결과는 Table 17과 같다.

Table 17. Perceived Necessity of Data Science Project-Based Class

	A(%)	B(%)	C(%)	Total(%)
Necessary (%)	53(91.4)	14(100)	35(94.6)	102(93.5)
Neutral (%)	3(5.2)	0	2(5.4)	5(4.6)
Unnecessary (%)	2(3.4)	0	0	2(1.8)
Total (%)	58(100)	14(100)	37(100)	109(100)

‘데이터 과학’수업 경험에 대한 인식은 세 가지 측면에서 공통된 흐름을 보였다. 학습자들은 ‘데이터 과학’의 중요성을 의사 결정과 문제 해결 과정에서 크게 인식하였고, 수업을 통해 체득한 능력은 집단별 수업 특성과 연계된 차이를 보였다. 프로젝트 기반 수업에 대해서는 대부분 긍정적으로 필요성을 인식하였다.

5. 결론 및 제언

본 연구는 2022 개정 교육과정에 새롭게 도입된 ‘데이터 과학’과목을 고등학생 대상으로 설계·운영하고, 학습자 인식을 중요도·필요도·난이도·이해도 관점에서 분석하였다. 이를 통해 교육과정의 실행 가능성과 수업 설계 방향을 탐색하는 데 목적을 두었다.

첫째, 2022 개정 교육과정의 ‘데이터 과학’내용 체계는 학습자에게 중요성과 필요성이 높게 인식되었다. 학습자들은 데이터 준비, 분석, 평가 각 단계와 세부 학습 요소에 전반적으로 긍정적인 반응을 보였으며, 데이터 기반 의사 결

정, 프로그래밍, 데이터 리터러시, 통계 해석 등과 같은 영역이 유용하다고 인식하였다. 교육과정이 학습자에게 의미 있게 연결될 수 있음을 보여준다. 실제 역량 변화 여부는 본 연구의 범위를 넘어서는 부분으로, 후속 연구에서 사전·사후 비교나 종단적 분석을 통해 확인할 필요가 있다.

둘째, ‘데이터 과학’수업의 효과적인 운영을 위해 수업 초기에 학습자가 데이터를 직접 관찰하고 분석하는 탐색 활동을 경험하는 것이 유용할 수 있다. 본 연구에서는 도입 단계에서 시각화 활동이 포함된 경우 개념 이해가 높게 인식되는 경향이 나타났다. 그러나 집단별 수업 설계가 달랐기 때문에 다른 요인이 작용했을 가능성을 배제할 수 없으며, 탐색 및 시각화 활동의 효과는 동일한 조건에서의 후속 검증이 필요하다.

셋째, ‘데이터 과학’수업은 데이터를 직접 다루는 경험을 포함한 수업 설계가 학습자의 인식에 긍정적으로 작용할 수 있다. 이를 위해 학교 교육과정 수준에서 유기적인 연계를 바탕으로 구조적 기반을 마련할 필요가 있다. 탐색 활동과 실습 중심의 도입 수업은 개념 이해를 지원할 수 있으며, ‘정보’, ‘인공지능 기초’ 등 기존 과목과 연계하여 기초 개념 및 실습을 선행하고 이후 ‘데이터 과학’과목으로 확장하는 단계적 구성은 실행 가능성 높은 방안이 될 수 있다[23, 24].

넷째, ‘데이터 과학’수업은 단계별 학습 요소의 난이도와 인지적 특성을 고려한 교수·학습 설계가 요구된다. 데이터 준비 단계의 데이터 수집과 신뢰성은 개념 중심 설명만으로도 충분할 수 있으나, 컴퓨팅 활용 능력은 높은 난이도와 낮은 이해도를 보여 실습 중심 수업이 필요하다. 데이터 분석 단계는 프로그래밍 기반 기술이 포함되어 난이도가 높게 인식되었으며, 반복 실습과 시각화, 피드백을 포함한 교수 전략이 효과적일 수 있다. 평가 단계의 결과 해석은 고차 사고를 요구하는 영역으로, 사례 기반 문제 해결이나 토의 활동과 같은 인지 전략을 활용한 수업 설계가 적절하다. 따라서 학습 요소의 인지 요구 수준에 따라 교수 전략과 활동 유형을 차별적으로 적용해야 한다.

다섯째, 본 연구는 집단 간 인원수 차이에서 제한점이 존재한다. A·C 집단은 30명 이상으로 비교적 안정적인 분석이 가능하였으나, B 집단은 14명으로 소표본에 해당한다. Welch 검정을 통해 집단 간 비교를 보완하였으나, 결과 해석과 일반화에는 한계가 있다. 향후 연구에서는 집단별 표본 수를 균형 있게 확보하여 분석의 신뢰성을 높일 필요가 있다.

본 연구는 학습자의 인식 자료를 바탕으로 ‘데이터 과학’ 교육과정의 현장 적용 가능성과 교수·학습 설계 방향을 탐색했다는 데 의의가 있다. 향후 연구에서는 수업 전후 변화를 비교하는 사전·사후 연구와 학습자의 변화를 장기적으로 추적하는 종단 연구가 필요하다. 이를 통해 ‘데이터 과학’교육의 효과성과 지속 가능성을 보다 명확하게 확인할 수 있을 것이다.

참고문헌

- [1] Jagadish, H. V. (2015). Big data and science: Myths and reality. *Big Data Research*, 2(2), 49–52. <https://doi.org/10.1016/j.bdr.2015.01.005>
- [2] Watson, P. (2021, May 6). *Why data is the new oil*. FutureScot. <https://futurescot.com/why-data-is-the-new-oil/>
- [3] OECD. (2018). *The future of education and skills: Education 2030*. OECD Publishing. <https://doi.org/10.1787/54ac7020-en>
- [4] UNESCO. (2021). *Media and information literacy curriculum for educators and learners: Think critically, click wisely!*. UNESCO Publishing. <https://www.unesco.org/mil4teachers/en/curriculum>
- [5] C2020 Task Force. (2020). *Computing curricula 2020: Paradigms for global computing education*. Association for Computing Machinery & IEEE Computer Society. <https://doi.org/10.1145/3467967>
- [6] ACM Data Science Task Force. (2021). *Computing competencies for undergraduate data science curricula*. Association for Computing Machinery. https://www.acm.org/binaries/content/assets/education/curricula-recommendations/dstf_ccdsc2021.pdf
- [7] Ministry of Education. (2022). *Practical arts (technology & home economics) / information curriculum* (The Notification of MOE 2022-33, The Separate Volume No. 10). Ministry of Education.
- [8] Tukey, J. W. (1962). The future of data analysis. *The Annals of Mathematical Statistics*, 33(1), 1–67. <https://doi.org/10.1214/aoms/1177704711>
- [9] Cao, L. (2017). Data science: A comprehensive overview. *ACM Computing Surveys*, 50(3), 1–42. <https://doi.org/10.1145/3076253>
- [10] Naur, P. (1974). *Concise survey of computer methods*. Studentlitteratur. <https://www.naur.com/Conc.Surv.html>
- [11] Wu, J. (1997). *Statistics = data science*. <http://www2.isye.gatech.edu/~jeffwu/presentations/datascience.pdf>
- [12] Hayashi, C. (1998). *What is data science? Fundamental concepts and a heuristic example*. Springer. https://doi.org/10.1007/978-4-431-65950-1_3
- [13] Cleveland, W. S. (2001). Data science: An action plan for expanding the technical areas of the field of statistics. *International Statistical Review*, 69(1), 21–26. <https://doi.org/10.1111/j.1751-5823.2001.tb00477.x>
- [14] Breiman, L. (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science*, 16(3), 199–231. <https://doi.org/10.1214/ss/1009213726>
- [15] Dhar, V. (2013). Data science and prediction. *Communications of the ACM*, 56(12), 64–73. <https://doi.org/10.1145/2500499>
- [16] Donoho, D. (2017). 50 years of data science. *Journal of Computational & Graphical Statistics*, 26(4), 745–766. <https://doi.org/10.1080/10618600.2017.1384734>
- [17] Brodie, M. L. (2024). *A framework for understanding data*

science. arXiv. <https://arxiv.org/abs/2403.00776>

- [18] Yang, H., Kim, J., & Lee, W. (2021). Analysis of computing competency framework for data science. *The Journal of Korean Association of Computer Education*, 25(1A), 89–92.
- [19] Kim, H., & Cho, J. (2022). Data analysis and visualization training: A study on the effect analysis on physics curriculum competence. *Intelligent Information Convergence and Future Education*, 1(1), 31–44.
- [20] Son, M. (2023). The effect of data analysis process on high school students' question generation. *School Science Journal*, 17(3), 346–360. <https://doi.org/10.15737/ssj.17.3.202308.346>
- [21] Ahn, H., & Kim, H. (2024). Developing and applying a science inquiry program utilizing data science. *Brain, Digital, & Learning*, 14(3), 475–487.
- [22] Kim, S., & Woo, S. (2022). Development of a data science education program for high school students taking the high school credit system. *Journal of Practical Engineering Education (JPEE)*, 14(3), 471–477.
- [23] Lee, E., & Kim, C. (2024). The effect of AI-based data science education on convergence thinking ability and data literacy of high school students. *The Journal of Korean Association of Computer Education*, 27(8), 99–109.
- [24] Song, Y., Song, S., Kim, Y., & Lim, C. (2021). A developmental study of an instructional model and strategies for data-driven debate (DDD) to improve data literacy. *Journal of Educational Technology*, 37(4), 943–982.



전현희

- 2023년 한국방송통신대학교 교육학과(교육학사)
- 2025년 고려대학교 교육대학원 교육정보전공(교육학석사)
- 2025년~현재 고려대학교 대학원 컴퓨터학과 박사 과정

✚ 관심분야 : 정보교육, 정보교육과정, 에듀테크, 교육정책

✉ hsimple@naver.com



김자미

- 1992년 이화여자대학교 교육학과(문학사)
- 1995년 이화여자대학교 교육학과(문학석사)
- 2011년 고려대학교 컴퓨터교육학과(이학박사)
- 2011년~2015년 고려대학교 컴퓨터학과 연구교수
- 2015년~현재 고려대학교 교육대학원 컴퓨터교육 전공 부교수

✚ 관심분야 : 정보교육, 교육과정평가, 에듀테크

✉ celine@korea.ac.kr