

초등교사 대상의 기초 데이터 과학 교육의 사례 연구

조정희

부산교육대학교 컴퓨터교육과

요약

데이터 과학은 통계학, 컴퓨터 과학, 정보기술, 도메인 지식 등 여러 분야의 융합 학문으로써 다양한 학문에서 제공하는 복합적인 기술을 이용하여 데이터를 분석하고 의미 있는 결과를 도출한다. 데이터 과학은 인공지능과 함께 4차산업혁명의 핵심기술로써, 고도의 전문성을 요하는 데이터 과학자의 양성을 위해 세계의 대학과 기업에서는 다양한 프로그램들을 활발히 개발하고 있다. 이러한 사회적 흐름에 맞추어, 초등 교육 현장에서도 데이터 과학 교육의 중요성을 인식하고 학생들이 데이터를 이해하고 활용하도록 관련 콘텐츠를 개발하고자 연구가 진행되고 있다. 본 논문에서는 컴퓨터 분야의 비전공자가 대다수인 현직 초등교사들의 데이터 과학 교육을 목적으로 강의 콘텐츠를 제안하고, 인공지능융합대학원에 재학 중인 현직 초등교사 집단을 대상으로 15차시 교육 과정을 통해 적용하였다. 그리고, 본 논문에서 제안된 데이터 과학 교육 사례의 효과성을 분석하기 위해 학습자들로부터 수집한 설문을 바탕으로 만족도 분석을 실시하였다.

키워드 : 소프트웨어 교육, 데이터과학, 교육과정, 초등교육

A Case Study of the Curriculum of Data Science for Elementary School Teachers

Junghee Jo

Busan National University of Education

Abstract

Data science is a discipline comprised of the academic fields of statistics, computer science, information technology, and domain knowledge. It analyzes data and derives meaningful results using complex technologies. Data science, along with artificial intelligence, is a core technology of the 4th industrial revolution; consequently, universities and companies worldwide are actively developing programs to develop data scientists who require high levels of expertise. In line with this undertaking, the field of elementary education has recognized the importance of data science education and so various studies have been conducted to develop curricula designed to help students understand how to use data. This paper proposes a curriculum for the purpose of educating elementary school teachers who are mostly non-majors in the computer field about data science. Satisfaction analysis was conducted based on questionnaires collected from students to analyze the effectiveness of the data science education proposed in this paper.

Keywords : Software Education, Data Science, Education Curriculum, Elementary Education

본 연구는 2021년도 부산교육대학교 학술연구과제로 지원을 받아 수행되었음.

논문투고 : 2021-09-27

논문심사 : 2021-11-01

심사완료 : 2021-11-07

1. 서론

세계는 데이터를 미래 경제를 이끌어 갈 핵심 기술로 판단하고 데이터 시장을 주도하기 위해 사활을 걸고 있다. 미국, 중국, 그리고 EU 등의 주요 국가는 정부 차원에서 데이터를 활용해 시장 경제의 확대를 도모함과 동시에 데이터 보안을 통해 국민을 보호할 수 있는 대응 정책들을 빠른 속도로 개발하고 있다. 국내의 경우, 2020년 데이터 산업의 시장규모는 2019년 대비 14.3% 증가한 19조 2,736억원, 데이터 직무 인력은 14.5% 증가한 101,967명으로 각각 추정되었다. 데이터 직무 인력이란, 데이터 아키텍트, 데이터베이스 관리자, 그리고 데이터 과학자 등 데이터와 관련된 업무를 수행하는 인력을 의미한다. 향후 5년 이내에 기업에서 추가로 필요로 하는 인력은 12,114명으로 부족률이 약 10.6%로 파악되었으며, 고도의 전문성을 요하는 ‘데이터 과학자’의 부족률은 31.4%로 가장 높은 것으로 조사되었다[1].

데이터 과학자는 원활한 업무 수행을 위해 컴퓨터 사이언스 및 통계학을 비롯하여 분석하고자 데이터에 관한 배경 지식 또한 충분히 갖추었다고 가정한다. 조완섭[2]은 방대한 데이터로부터 의미를 발굴하고 그것을 비즈니스 가치로 연결하는 사람들로 데이터 과학자를 정의하였고, 송현민[3]은 데이터 과학자의 업무를 ‘컴퓨터 공학, 데이터 공학, 경영학, 통계학적인 지식과 기술을 활용, 데이터 기반의 정책 제도, 기획, 처리, 분석, 시각화, 운영관리 업무를 수행하고 빅데이터를 수집, 저장, 분석, 표현하는 체계를 만드는 지식과 기술을 습득하여 프로세스 혁신 및 신제품 개발, 마케팅 전략 결정 등의 과학적 의사결정을 끌어내는 직무’로 정의하였다. 미국을 선두로 한 세계의 우수 대학에서는 데이터 과학자 양성을 위한 다양한 프로그램들을 개발하고 있으며[4], 국내의 관련 정부 부처도 적극적으로 대응하고 있다.

이러한 사회적 흐름에 맞추어, 초등 교육 현장에서도 학생들의 컴퓨팅 사고력과 창의성 향상을 목적으로 데이터 과학 교육을 도입하고자 다양한 연구가 진행되고 있다. 구덕희(2020)는 데이터 과학 교과 내용의 특성을 반영한 교육 프로그램을 개발하였다. 즉, 교수 학습의 단계를 문제 정의, 데이터 수집, 데이터 분석, 데이터 표현, 평가 및 일반화로 설계하였고 이를 적용하기 위해 차시별 세부 지도계획을 수립하고 예시 교재를 개발하

였다[5]. 유사한 사례로, 홍지연(2020)은 초등 학생들을 대상으로 데이터에 대한 이해, 수집, 분석, 표현 능력의 향상을 위한 교육 프로그램을 개발하였다[6]. 허경(2020)은 도구를 활용한 데이터 과학 교육에 관한 연구의 일환으로 엔트리에서 제공하는 데이터 과학 관련 기능을 활용한 초등 교육용 수업 콘텐츠를 제안하였으며[7], 김용민(2020)은 ADDIE(Analysis, Design, Development, Implementation, Evaluation) 모형의 절차에 따라 PBL(Project Based Learning) 기반의 데이터 과학 교육 수업을 엠 인벤터를 활용하여 설계하고, 이를 적용하였을 때 초등학생들의 컴퓨팅 사고력 향상의 유무를 검증하였다[8].

이에, 본 논문에서는 현직 초등교사를 대상으로 개발하여 교육 현장에서 활용하고 있는 데이터 과학 수업의 구성을 설명하고, 학생들의 데이터 과학과 관련한 사전 지식의 정도, 특별히 선호했던 내용 및 이해에 어려움을 느꼈던 부분을 분석하고, 이 과정에서 얻은 경험을 반영하여 향후 수업에의 적용 방안을 도출하였다.

2. 관련 연구

2.1. 데이터 과학

데이터란, 모든 행위를 숫자나 문자의 기록으로 남긴 것을 의미한다. 사물 인터넷 기술의 발달로 다양한 기능과 종류의 센서가 발명되었고 이러한 센서들로부터 여러 종류의 데이터가 빠른 속도로 수집되고 있다. 노르웨이의 연구기관인 SINTEF는 현재 시점에서 수집된 데이터의 90%가 지난 2년간 만들어진 최신 데이터임을 연구 결과로 증명하였다[9]. 다른 연구에 의하면, ‘1분’이라는 상당히 짧은 시간동안 유튜브 시청은 433만건, 기상예측접속은 1,805만건, 넷플릭스 시청은 9만7천건, 문자발송은 1,298만건, 구글검색은 387만건등의 수천만건의 데이터가 동시 다발적으로 발생한다고 한다[10]. 이러한 데이터로부터 의미 있는 지식을 추출하는 학문을 데이터 과학이라고 한다. 데이터는 그 자체로는 가치가 미약하지만 내재된 지식을 찾아내면 비로소 가치를 갖게 된다.

데이터 과학은 통계학, 컴퓨터 과학, 정보기술, 도메인 특화 영역 등 여러 분야의 융합 학문으로써[11-13],

다양한 학문에서 제공하는 복합적인 기술을 이용하여 데이터에 담긴 패턴이나 예측에 도움이 되는 결과를 도출한다. 데이터 과학의 절차로는, 해답을 얻고자 하는 문제의 정의, 문제를 해결하기 위한 데이터의 수집, 전처리, 그리고 분석 과정 등이 포함된다. 특히, 데이터의 분석은 데이터간 연관성을 찾아내거나 데이터를 분류하거나 군집화 하는 등 해결하고자 하는 문제와 주어진 데이터의 특징을 바탕으로 의미 있는 결과를 얻기 위한 기술을 의미한다[14]. 이러한 분석 결과를 기반으로 데이터를 가장 잘 설명할 수 있는 모델을 찾고, 최종적으로 이러한 결과들을 효율적으로 전달하기 위한 자료를 작성하는 과정을 거친다.

2.2. 데이터 과학의 연구 동향

포브스(Forbes)는 2020년을 선도할 7가지 주요 기술들을 발표하였는데 그 중의 하나가 ‘빅데이터’이다[15]. 빅데이터는 비정형의 데이터까지도 포함하는 대용량의 데이터로써 이러한 정보를 처리하기 위해서는 고난이도의 기술이 필요하다. 예를 들어, 빅데이터의 상당한 부분을 차지하고 있는 ‘공간 빅데이터’의 경우는 빅데이터의 일반적인 특징과 더불어 공간정보의 특징도 보유하고 있으므로 다른 종류의 빅데이터에 비하여 다양성과 복잡도가 높은 특징을 가지고 있다[16]. 이러한 공간 빅데이터를 분석하기에는 기존의 시스템들은 성능의 한계가 발생하므로 새로운 시스템에 관한 연구가 활발히 이루어지고 있다. 대표적인 연구 결과로는, 아파치의 하둡(Apache Hadoop), 아파치의 스파크(Apache Spark), 미국의 미네소타 대학의 공간 하둡(SpatialHadoop)등이 있다. 하둡은 다수의 컴퓨터를 하나의 컴퓨터인 것처럼 묶는 기술을 통해 시스템의 저장 공간과 계산 능력을 향상시킴으로써 빅데이터 처리가 가능한 공개 소프트웨어 프레임워크이다[17]. 스파크는 하둡의 한계를 보완하기 위한 인메모리 기반의 분산 데이터 분석 시스템으로써, 데이터를 메모리에서 처리하므로 디스크 기반의 하둡에 비하여 속도가 개선된 장점이 있다[18]. 공간 하둡은 기존의 하둡에 공간 정보를 처리하는 기능을 추가하여, 영역 질의, 최단 경로 질의 및 공간 조인과 같은 공간 분석에 특화된 기능을 기반으로 공간 빅데이터 분석을 수행할 수 있도록 하였다[19].

2.3. 데이터 과학의 교육 동향

교육부는 2018년부터 초, 중, 고등학교의 교육과정에 소프트웨어 교육을 필수적으로 적용하였다. 데이터 과학이 초, 중, 고등학교의 정보 교과에 적용된 실태를 조사한 결과에 의하면, 정부에서 승인한 정보 교과서 중에 금성출판사의 경우에는 데이터와 가까운 주제인 ‘자료와 정보’가 전체 정보 교과 내용의 15%를 차지하고 있으며, 교학사의 정보 교과서의 경우에는 17%를, 교문사의 정보 교과서의 경우에는 22%를, 이오북스의 정보 교과서의 경우에는 21%를 차지하고 있다. 즉, 정보교과서에서는 데이터와 관련된 내용을 약 15%-22%의 분량으로 다루고 있다. 이들 교과서는 데이터와 관련한 실질적인 예시와 구체적인 내용보다는 간단한 정도의 이론 위주로 교육을 하도록 구성된 것으로 조사되었다[20].

초등교육에 데이터 과학의 개념을 적용하기 위한 다양한 연구도 활발히 진행되고 있다[21-23]. 이명호(2016)는 국내외에 증가하고 있는 데이터 과학 프로그램에 대한 교과과정을 조사하고, 국내 및 국외의 데이터 과학 교과과정을 국내와 비교하여 향후 개선 방향을 제시하였다[24]. 이해원(2020)은 국내에 개설된 데이터 과학과 관련한 교육 과정과 교과목들을 대상으로 학문 영역 특징 기반 분석, 데이터 전문가 역량 기반 분석, 그리고 교과목명 내용 분석을 시행하였다[25]. 김재경(2021)은 컴퓨터 비전공자의 컴퓨팅 사고력을 향상시키는 목적으로 ADDIE 모델을 기준으로 데이터 과학 교육을 위한 수업안을 개발하였다. 그리고 학부생을 대상으로 15주간 교육하고 사전/사후 설문을 시행하여 결과를 분석하고 교육의 효과가 있음을 입증하였다[26].

3. 데이터 과학 수업 설계

3.1. 이론적 내용

본 연구를 통해 개발된 데이터 과학 수업 콘텐츠는 현직 초등교사를 대상으로 하였다. 수업의 도입부에는 ‘데이터’의 기본적인 정의를 설명하였고 빅데이터의 생성 배경을 사물 인터넷 및 인공지능과 연결시켜 제시하였다. 이 과정에서, 인공지능의 발전에 크게 이바지한 대표적인 데이터 과학자들에 관한 소개를 포함시켰다.

예를 들어, 대용량 이미지 데이터베이스인 이미지넷을 구축한 스탠포드 대학의 페이페이 리 교수의 업적을 인공지능과 빅데이터를 연결지어 설명함으로써 학습자들이 두 기술 간의 연관성을 이해하도록 유도하였다. 이어서, ‘데이터 과학’의 정의 및 배경지식, 그리고 요소 기술들을 차례로 소개하였고, 데이터 과학의 개념을 전반적으로 정리하는 동영상상을 시청하도록 함으로써 개요에 대한 설명을 마무리하였다.

다음으로, ‘데이터 과학자’의 정의를 설명하고 갖추어야 할 역량들을 (Fig. 1)의 자료를 토대로 순차적으로 설명하였다. 그리고, 일반인의 인지도가 높은 대기업들의 인재 영입 사이트에 접속해서 데이터 분석과 관련된 다양한 인력 모집 공고들을 학습자들에게 보여줌으로써 산업계에서 필요로 하는 데이터 과학자의 급진적인 수요에 대해 본인들이 직접 느끼도록 유도하였다. 국외의 사례 또한 설명하기 위해, 컴퓨터 분야의 세계적인 선도 기업에서 데이터 과학자로 근무중인 한국인의 인터뷰 동영상상을 시청하도록 하여 다양한 경험을 간접적으로나마 체험할 수 있는 기회를 제공하였다.



(Fig. 1) Skills a data scientist should have (source: <https://www.proschoolonline.com/blog/data-science-skills>)

이후에는 데이터 과학의 절차에 대한 설명이 순차적으로 이어지고, 교수자가 직접 경험한 데이터 과학 관련 프로젝트들의 상세한 내용을 학습자들과 공유하고 이를 기반으로 토론하는 시간을 가짐으로써 이론으로 배운

데이터 과학과 현업에서 이루어지는 데이터 과학의 내용을 스스로 고찰해보고 연결해볼 수 있는 기회를 제공하였다.

3.2. 실습적 내용

대용량 데이터를 대상으로 파이썬과 같은 데이터 분석용 언어를 활용한 실습 수업을 위해서는 텍스트 프로그래밍 언어의 이해가 필수적으로 수반된다. 따라서, 프로그래밍 언어의 경험이 전혀 없는 학습자들에게는 누구나 쉽게 배워서 활용할 수 있는 블록 코딩 방식의 언어가 유용하다. 엔트리는 블록 코딩 방식의 소프트웨어 교육 플랫폼으로 컴퓨터 비전공자에게도 친숙하게 사용될 수 있는 도구이다. 엔트리는 꾸준히 업그레이드되고 있으며 최근에는 데이터를 수집하고, 시각화하고, 프로그래밍과 접목하여 활용하는 기능이 추가됨으로써 초등 교육 현장에서도 데이터 과학의 교육이 가능하게 되었다. (Fig. 2)는 엔트리의 데이터 과학 블록을 사용하여 ‘성별 학생수’ 데이터를 분석해보는 간단한 프로그램을 구현한 예제이다.



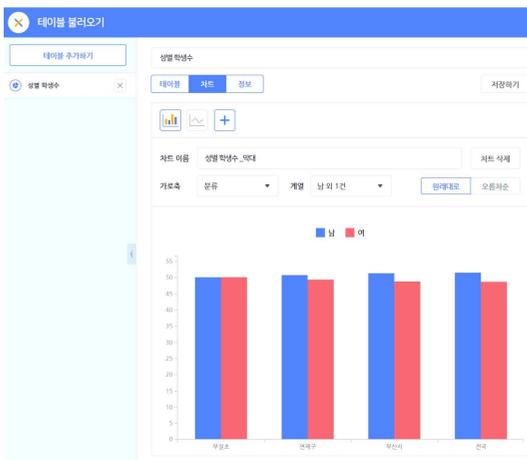
(Fig. 2) An example of data science blocks in Entry

엔트리에서 제공하는 데이터 분석 기능을 활용하기 위해서는 기본적으로 데이터가 ‘테이블’의 형태로 제공되어야 하므로, 실습의 초반에 테이블의 기본적인 정의를 설명하였고 ‘행’, ‘열’, ‘레코드’, ‘속성’ 등의 테이블 관련 용어들에 관해 설명을 하였다. 엔트리에 데이터를 입력하는 방법인, ‘새로 만들기’, ‘테이블 선택’, ‘파일 올리기’의 기능은 난이도의 순서대로 예시와 함께 차례로

설명하였다. ‘새로 만들기’의 경우는 사용자가 엔트리에 직접 빈 테이블을 생성하여 행과 열을 추가하고 데이터를 입력하는 방법인데 ‘학교알리미’ 사이트에서 제공하는 성별 학생 수에 관한 공공 데이터를 참조하여 (Fig. 3)과 (Fig. 4)와 같이 엔트리에 직접 데이터를 입력하고 간단한 차트를 생성하여 시각화해보도록 하였다.



(Fig. 3) An example of creating a data table in Entry



(Fig. 4) An example of creating visualizations in Entry

이어서, ‘테이블 선택’의 경우는 ‘네이버 블로그 평균 사용 시간’, ‘연도별 배추 생산량’ 등과 같이 엔트리에 기본적으로 제공하는 데이터를 활용하여 간단한 엔트리 프로그램을 스스로 작성해보도록 하였다. 마지막으로, ‘파일 올리기’ 기능을 활용하여 외부에서 가져온 공공 데이터를 엔트리로 불러와서 분석해보는 실습을 수행하였다. 이후부터는, 다양한 종류의 데이터를 수집, 분석하

고 변수 활용 및 신호 보내기와 같은 기능을 활용하여 데이터 분석용 엔트리 프로그램을 작성해보도록 하였다.

4. 적용 결과

본 연구를 통해 개발된 데이터 과학 수업 콘텐츠는 인공지능융합대학원에 재학 중인 초등교사 22명을 대상으로 15차시의 수업에 사용되었으며, 종강 후에 학습자들로부터 수집한 설문을 바탕으로 결과를 도출하였다.

<Table 1> Survey results of block programming experience

Level	Num. of students	Percentage
Good	1	5%
Fair	12	54%
Poor	7	32%
No experience	2	9%
Sum	22	100%

<Table 1>은 학습자들이 대학원에 입학하기 전, 엔트리와 같은 블록 프로그래밍에 대한 경험의 수준에 대한 응답 결과이다. 학습자는 다음의 4단계 중에 한 가지를 선택하도록 하였다. ‘상’은 원하는 응용 프로그램 개발이 자유자재로 가능한 단계이고, ‘중’은 블록 프로그래밍의 이론은 충분히 알고 있으나 응용 프로그램 개발은 미흡한 단계이며, ‘하’는 블록 프로그래밍의 이론도 미숙하여 응용 프로그램 개발이 힘든 단계이다. 그리고, ‘없음’은 블록 프로그래밍을 경험해 보지 못한 단계를 의미한다. 가장 많은 응답을 얻은 단계는 ‘중’으로 총 22명 중에 12명(54%)의 학습자들이 본인이 이 단계에 해당한다고 응답하였다. 이어서, ‘하’ (7명, 32%), ‘없음’ (2명, 9%), ‘상’ (1명, 5%) 순으로 수집되었다.

<Table 2> Survey results of text programming experience

Level	Num. of students	Percentage
Good	1	4%
Fair	3	14%
Poor	5	23%
No experience	13	59%
Sum	22	100%

<Table 2>는 학습자들이 대학원에 입학하기 전, 파이썬과 같은 텍스트 프로그래밍에 대한 경험치의 수준에 대한 응답 결과이다. 학습자는 앞에서 설명한 4단계 중에 한 가지를 선택하도록 하였다. 가장 많은 응답을 얻은 단계는 ‘없음’으로 총 22명 중에 13명(59%)의 학습자들이 본인이 이 단계에 해당한다고 응답하였다. 그 뒤로 ‘하’ (5명, 23%), ‘중’ (3명, 14%), ‘상’ (1명, 4%) 순으로 수집되었다. 특히, ‘상’ 이라고 응답한 1명의 학습자는 교육대학교에 입학하기 전에 일반대의 기계공학전공 과정을 수강하면서 포트란, C++, 비주얼 베이직 등의 다양한 종류의 텍스트 프로그래밍 언어를 경험해 보았던 특수한 경우였다.

<Table 3> Survey results of contents that learners had difficulty understanding

Category	Num. of questions	Num. of students	Avg. of students
Theoretical Foundations of Data Science	3	6	2(7%)
Theoretical Foundations of Entry	1	1	1(3%)
Practice of Data Science using Entry	6	23	4(14%)
Theoretical Foundations of Python	2	13	7(24%)
Practice of Data Science using Python	1	15	15(52%)
Sum	13	58	29(100%)

<Table 3>은 강의의 전반적인 콘텐츠 중에서 학습자들이 이해하기에 어려움을 느낀 부분들에 관하여 묻는 질문에 대한 응답 결과이다. 이 질문을 위해, 전체 강의를 총 13가지의 내용으로 분류하였고, 학습자들은 복수개의 내용을 선택하도록 허용하였다. 13가지 내용에 대한 학습자들의 선택 결과는 ‘데이터 과학 이론’, ‘엔트리 이론’, ‘엔트리 기반의 데이터 과학 실습’, ‘파이썬 이론’, ‘파이썬 기반의 데이터 과학 실습’의 5가지 카테고리로 합산하였으며, 각 카테고리에 응답한 평균 학습자들의 수를 계산하여 결과를 도출하였다. 가장 많은 응답을 얻은 강의 내용은 ‘파이썬 기반의 데이터 과학 실습’ (15명, 52%) 이었고, 이어서 ‘파이썬 이론’ (7명, 24%), ‘엔

트리 기반의 데이터 과학 실습’ (4명, 14%), ‘데이터 과학 이론’ (2명, 7%), ‘엔트리 이론’ (1명, 13%) 순으로 수집되었다.

<Table 4>는 향후에 추가 또는 보완이 필요한 강의 내용을 묻는 질문에 대한 주관식 형태의 응답 결과이며, 총 22명의 학습자 중에서 구체적으로 응답한 8명(36%)을 대상으로 결과를 도출하였다.

<Table 4> Categories of contents needing improvement

Category	Num. of students	Percentage
Overseas AI Education	1	13%
Text Programming	4	50%
Theoretical Foundations of Data Science	1	13%
Practice of Data Science	1	13%
Practice of Artificial Intelligence	1	13%
Sum	8	100%

가장 많은 수의 학습자들이 텍스트 프로그래밍 언어라고 응답하였으며 (4명, 50%), 그 외에 해외 인공지능 초등 교육 사례 (1명, 13%), 데이터 과학 이론 (1명, 13%), 데이터 과학 실습 (1명, 13%), 인공지능 교육 도구 실습 (1명, 13%)의 응답들이 수집되었다.

5. 결론

본 논문의 사례 연구를 통해 도출된 내용은 다음과 같다. 첫째, 현직에서 활동 중인 초등교사들은 학부 재학시 관련 교과를 수강을 통하여, 또는 교육 현장에서의 경험으로 인하여 블록 프로그래밍에 대한 경험이 풍부할 것이라고 예상하였다. 그러나 설문 결과에 의하면, 학부 재학 시 관련 교과를 수강하지 못했거나 수강하였다더라도 시간이 흘러 기억을 못하는 등의 이유로 블록 프로그래밍이 힘든 학습자들이 상당수 존재하였다. 본 논문에서 제안한 수업 콘텐츠는 학습자들이 중급 이상의 블록 프로그래밍의 경험이 있을 것으로 가정하고 제작하였기에 일부 학습자들은 블록프로그래밍의 기초를

혼자 학습하여 따라 가야하는 어려움이 발생하였다. 따라서, 향후에는 이를 보완하여 블록 프로그래밍의 개념에 관한 내용을 콘텐츠에 추가하여야 할 필요가 있음을 파악하였다. 둘째, 데이터 과학의 실습을 위해 엔트리의 데이터 블록을 활용하는 방법 이외에도, 텍스트 프로그래밍을 활용하면 고도의 데이터 분석이 가능하다는 점을 전달하기 위해, 간단한 파이썬 문법을 활용한 데이터의 시각화 생성에 관한 내용을 수업 후반부의 2차시에 걸쳐서 실험적으로 실시하였다. 코로나로 인한 비대면 강의의 특성상, 학습자들의 이해도를 즉시 파악할 수 없었으므로 수업 종료후에 텍스트 프로그래밍에 대한 경험치를 묻는 설문을 시행하였고 예상대로 절반 이상의 학생들이 해당 경험이 전혀 없었다. 이는, 파이썬의 기초 이론과 시각화 실습이 비교적 단순한 내용이었음에도 불구하고 상당수의 학습자가 이해에 어려움을 느꼈던 원인으로 파악하였다. 마지막으로, 파이썬을 이용한 간단한 데이터의 시각화에 어려움을 느낀 학습자들이 상당수임에도 불구하고, 향후에 보완을 희망하는 강의의 내용으로 많은 수의 학습자가 ‘텍스트 프로그래밍’으로 응답하였다. 비록, 생소한 언어를 처음 사용함으로써 어려움은 느꼈으나 교육 현장에서 매우 익숙한 엔트리보다는 새로운 분야인 텍스트 프로그래밍을 학습해보고자 하는 학생들이 상당수 존재함을 파악하였다. 본 사례 연구를 통해 도출된 내용은 학술적으로 일반화하기에는 한계가 있다. 하지만, 향후 초등 데이터 과학 교육의 깊이와 질을 고도화하는 과정에 기여할 수 있을 것으로 기대한다. 향후 연구 과제로는, 본 연구를 통해서 파악된 내용들을 반영하여 수업 콘텐츠를 개선하고 이를 교육 현장에 활용하여 검증해 보는 것이다.

참고문헌

- [1] Korea Data Agency(2020). 2020 Data Industry Survey, Korea Data Agency.
- [2] Jo, W.S.(2014). Big Data Utilization and Data Scientist. *Communications of the Korean Institute of Information Scientists and Engineers*, 32(1), 59-65.
- [3] Song, H.M.(2015). A Study of Development for job Competency in the bigdata using DACUM method, Korea University master's thesis.
- [4] Rappa, M.(2011). Master of Science in Analytics: Goals, Learning, and Outcomes, NC State University, Internal Report.
- [5] Koo, D.H., Kim D.J.(2020). Data science education program based on problem solving learning. *Korean Journal of Elementary Education*, 31, 203-215.
- [6] Hong, J.H., Kim, Y.S.(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(6), 633-641.
- [7] 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.
- [8] Kim, Y.M.(2020) The Effects of PBL-based Data Science Education classes using App Inventor on elementary student Computational Thinking and Creativity improvement. *Journal of The Korean Association of Information Education*, 24(6), 551-562.
- [9] SINTEF(2013). Big Data, for better or worse: 90% of world's data generated over last two years. Retrieved from <https://www.sciencedaily.com/releases/2013/05/130522085217.htm>.
- [10] Data Never Sleeps 6.0, Retrieved from <https://www.domo.com/learn/information-graphic/data-never-sleeps-6>
- [11] Shi, Y., Philip, S.Y., Zhu, Y., Tian, Y.(2014). Explore New Field of Data Science Under Big Data Era: Preface for ICDS 2014. *Procedia Computer Science*, (30), 1-3.
- [12] Bruce, P., Bruce, A.(2017). Practical statistics for data scientists: 50 essential concepts. O'Reilly Media, Inc.
- [13] O'Neil, C., Schutt, R.(2013). Doing data science: Straight talk from the frontline. O'Reilly Media, Inc.

- [14] Rajaraman, A., Ullman, J.D.(2011). Mining of massive datasets. Cambridge University Press.
- [15] Forbes(2019). The 7 Biggest Technology Trends That Will Transform Telecoms In 2020. Retrieved from <https://www.forbes.com>.
- [16] Jo, J.H., Lee, K.W.(2018). Marmot: A Hadoop-based High Performance Data Storage Management System for Processing Geospatial or Geo-Spatial Big Data. *Journal of Korean Society for Geospatial Information Science*, 28(1), 3-10.
- [17] White, T.(2012). Hadoop: The Definitive Guide, 3rd ed.; O'Reilly Media, Inc. Sebastopol, CA, USA: 2012; ISBN 1449338771.
- [18] Zaharia, M., Chowdhury, M., Franklin, M.J., Shenker, S., Stoica, I.(2010). Spark: Cluster Computing with Working Sets. *Proceedings of the 2nd USENIX conference on Hot topics in cloud computing*, 10, 10-10.
- [19] Eldawy, A.(2014). SpatialHadoop: Towards flexible and scalable spatial processing using MapReduce. *Proceedings of the SIGMOD PhD symposium 2014*, Snowbird, UT, USA, 22 June 2014, 46 - 50
- [20] Park, H.S.(2020). Analyzing the current status of how education is conducted for data science, suggesting directionality, and present the data science integration class plan. Incheon University master's thesis.
- [21] Kim, B.C., Kim, J.J., Moon, W.J., Seo, Y.H., Kim, J.A., Oh, J.C., Kim, Y.M., Kim, J.H.(2021). The Effect of Data Science Education on Elementary School Students' Computational Thinking: Focusing on Micro:bit's Sensor Function. *Journal of The Korean Association of Information Education*, 25(2), 337-346.
- [22] 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.
- [23] Jung, S.Y., Ma, Y.G., Koo, D.H.(2021). Development of a data analysis system for preventing school violence based on AI unsupervised learning, *Journal of The Korean Association of Information Education*, 25(5), 741-750.
- [24] Lee, M.H.(2016). A Study on the Curriculums of Data Science. *Journal of the Korean Biblia Society for Library and Information Science*, 27(1), 263-290.
- [25] Lee, H.W., Han, S.H.(2020). An Analysis of Data Science Curriculum in Korea. *Journal of the Korean Society for Library and Information Science*, 54(1), 365-385.
- [26] Kim, J.K, Sohn, E.S.(2021). Development of Data Science Course and Analysis of Computational Thinking Effect for non-Major Students. *The Journal of Korean Association of Computer Education*, 24(3), 23-31.

저자소개



조 정 희

2014년 University of Massachusetts Amherst, Computer Science, 공학박사
 2005~2019 한국전자통신연구원 (ETRI), 책임연구원
 2019~현재 부산교육대학교 컴퓨터교육과 조교수
 관심분야: 컴퓨터교육, 인공지능, 인간공학, 헬스케어
 e-mail: dreamer@bnue.ac.kr