

초등학생을 위한 데이터 분석대회를 활용한 데이터 분석 프로그램 개발

고학능^o, 정재리*, 이영준*

^o가득초등학교,

*한국교원대학교 컴퓨터교육과

e-mail: snddl3@outlook.kr^o, jelly6679@naver.com*, yjlee@knue.ac.kr*

Development of a Data Analysis Program Using a Data Analysis Competition for Primary School Students

HakNeung Go^o, JaeRi Jeong*, Youngjun Lee*

^oGadeuk Primary School,

*Dept. of Computer education, Korea National University of Education

● 요약 ●

본 논문에서는 초등학생을 위한 데이터 분석 대회를 활용한 데이터 분석 프로그램을 개발하였다. 데이터 분석 프로그램은 ADDIE 모형에 개발하였다. 분석 단계에서 G초등학교 학생들의 데이터 분석 도구인 스프레드 시트를 학습한 경험이 적고 배우고자 하는 동기가 없었다. 하지만 교육과정에서 공학도구로 활용하도록 제시되어 있다. 이를 바탕으로 디자인 학생들이 스프레드 시트를 학습할 수 있는 프로그램과 이를 실습할 수 있는 데이터 분석 대회를 디자인 하였다. 개발 단계에서는 LMS를 활용하여 학생들에게 학습을 위한 데이터를 제공하고 학습하며, 데이터 분석 대회에서는 학습한 데이터와 문제만 제공하여 대회에 참여하면서 실습할 수 있는 기회를 제공하였다. 평가 도구로는 데이터 리터러시 평가 도구를 선정하였다.

키워드: 데이터 과학(Data Science), 데이터 분석(Data Analysis), 캐글(Kaggle), 스프레드 시트(Spread sheet), 공공 데이터(Public data), 데이터 리터러시(Data Literacy)

I. Introduction

정보 기반 사회에서는 데이터의 가치는 중요하다. 의사결정, 마케팅, 연구 개발 등 다양한 분야에서 활용할 수 있으며 특히 데이터를 수집하고 이를 처리할 수 있는 기술이 발달함에 따라 데이터를 분석하는 능력이 중요해지고 있다[1]. 교육부는 수학 교육과정에서 공학도구를 활용하여 데이터를 시각화하는 교수학습 방법을 제시하고 있으며 데이터 과학 과목을 개설하여 데이터를 이해하고 데이터를 처리하는 방법을 익혀 문제 해결과 의사 결정에서 통찰력을 얻는 것을 목표로 한다[2][3].

캐글은 데이터 분석 대회로 알려져 있는 데이터 과학자 커뮤니티이다. 기업, 정부 등은 데이터를 제공하여 대회를 개최하면 데이터 과학자는 데이터 분석 대회에 참여하여 해결 방안, 예측 모델 개선 방안 등을 제출한다. 데이터 과학자는 데이터를 분석하고 협력하여 문제를 해결할 수 있는 기회를 제공받고 기업과 정부는 데이터를 분석하여 문제를 해결할 수 있는 기회를 얻을 수 있다[4].

데이터 분석의 중요성은 커지고 있으나 초등학교에서는 이를 반영한 교과가 없으며 연구는 공학 도구를 익히거나 프로젝트를 진행하는

데 초점이 맞춰져 있다. 이에 본 연구에서는 초등학생들이 데이터 분석을 학습하고 이에 실습할 수 있는 경험을 제공할 수 있는 프로그램을 개발하고자 한다.

II. Preliminaries

1. Kaggle

Kaggle은 예측모델 및 분석 대회 플랫폼이다. 기업 및 단체에서 데이터와 과제를 제공하고 상금을 제시하며, 데이터 과학자를 비롯한 사용자들이 이를 해결하기 위한 모델을 개발하고 경쟁하여 등수를 매기고 상금을 배분한다.

Kaggle에서 NH 투자증권, 카카오, 미국 국토 안보부, Santander Bank 등 다양한 기업, 정부, 연구 기관 등이 데이터 분석 대회를 개최했으며 이를 통해 기업의 서비스 개선이 이루어졌다[5].

III. The Proposed Scheme

데이터 분석 대회를 활용한 데이터 분석 프로그램 개발의 학습 목표는 스프레드 시트를 활용하여 실제 데이터를 분석하는 것이다. 학습 내용은 스프레드 시트를 사용하는 방법과 교육과정과 연계한 데이터를 분석하는 것이며 학습 대상은 초등학생이다.

G초등학교 6학년에 재학중인 학생들의 스프레드 시트 학습 경험은 9%로 매우 적었으며 스프레드 시트를 학습하고자 하는 동기는 적었다. 하지만 수학 교육과정에서는 공학도구를 활용하여 데이터 시각화 할 수 있도록 제시하였으며 고등학교에서는 데이터 과학 과목이 개설되었다. 사회적으로 데이터를 처리하고 분석할 수 있는 기술의 발달과 데이터를 수집하는 기술이 발달하여 데이터를 분석할 수 있는 환경이 갖춰졌으며 이를 바탕으로 데이터 분석하여 통찰을 얻고 정보를 얻는 것이 중요해졌다. 이를 바탕으로 학생들이 공학도구를 학습한 경험이 적고 동기가 부족하지만 교육과정, 사회적 요구에 따라 공학 도구를 바탕으로 한 데이터 분석의 필요성을 확인하였다.

교육 프로그램은 스프레드 시트를 학습하는 것과 학습한 내용을 실습하는 것으로 데이터 분석 대회로 구성하였다.

구체적으로 스프레드 시트를 학습하는 내용은 학습자 수준에 맞춰, 셀 주소를 이용한 계산, 자동 계산, 함수, 필터, 그래프 삽입으로 선정하였다. 또한 학생들에게 스프레드 시트의 같은 버전을 제공하고, 학습하기 위한 데이터를 제공하기 위해 마이크로소프트의 Teams나 구글의 Classroom을 활용할 수 있도록 하였다. 데이터 분석 대회는 학습한 내용을 실습할 수 있도록 하였으며 대회에 사용하는 데이터는 공공 데이터를 제공하며 Teams나 Classroom을 활용할 수 있도록 하였다.

적용 후 학생들의 데이터 분석에 대한 이해도를 평가하기 위해 데이터 리터러시를 평가하였으며 송유경(2020)이 개발한 데이터 리터러시 평가 도구를 활용하였다[6].

IV. Conclusions

Kaggle의 데이터 경진대회를 참고하여 데이터 분석 대회를 활용한 데이터 분석 프로그램을 개발하였다. 분석 단계에서 대상이 되는 초등학생의 경우 스프레드 시트를 학습한 경험과 적고, 이를 학습하기 위한 동기가 부족하다. 하지만 수학 교육과정에서는 공학 도구를 사용할 수 있도록 하며 사회적으로 데이터 분석이 필요성이 강조되고 있다. 디자인 단계에서 데이터 분석 학습하는 방법으로 스프레드 시트를 배우는 과정과 이를 실습할 수 있는 데이터 분석 대회 과정으로 구성하였다. 개발 단계에서 스프레드 시트를 학습하기 위해 LMS를 이용하여 학생들에게 학습하기 위한 데이터를 제공하고 필터, 함수, 그래프 삽입, 자동 계산, 셀 계산을 학습하도록 하고, 데이터 분석 대회에서는 공공 데이터를 제공하여 학생들이 학습한 내용을 적용할 수 있도록 하였다. 평가 도구로는 송유경이 개발한 데이터 리터러시 평가 도구를 활용하였다.

스프레드 시트를 학습하고 데이터 분석 대회라는 실습을 기회를 제공하는 방식으로 데이터 분석 프로그램을 개발하였다. 이를 실제 적용하여 효과성을 검증하고자 해야 할 것이다.

REFERENCES

- [1] Fry, B. Visualizing Data: Exploring and Explaining Data with the Processing Environment. O'Reilly Media, Inc. 2007.
- [2] Ministry of Education "Math Curriculum," December 2022
- [3] Ministry of Education "Data Science Curriculum," December 2022
- [4] Sakamoto, "Kaggle Guide", DongYang books, 2020.
- [5] Kaggle Wikipedia, "https://ko.wikipedia.org/wiki/캐글"
- [6] Uksong, "The Data-Driven Debate(DDD) Instructional Model for Improving Data Literacy", Master's Dissertation, Seoul University, 2021.