

## A Study on the Development and Application of an Educational Program to Enhance the Data Utilization Capabilities of Elementary Teachers

Junghee Jo\*

\*Associate Professor, Dept. of Computer Education, Busan National University of Education, Busan, Korea

### [Abstract]

This study designed and implemented an educational program to improve data utilization competencies, applying it to a group of 25 elementary school teachers. In particular, the program emphasized learning databases and SQL to strengthen structural understanding of data, while collecting and analyzing errors that occurred during the process of writing SQL queries. The analysis revealed that the most frequent errors committed by teachers were semantic errors(79.1%), which occurred approximately four times more often than syntax errors. Furthermore, when errors were classified according to SQL clauses, a large proportion were found in the WHERE, GROUP BY/HAVING, and ORDER BY, with the main cause being the omission of the relevant clause. The findings of this study are expected to serve as foundational resources for future database and SQL education targeting elementary school teachers.

▶ **Key words:** Data Science, Elementary School Teacher, Error, Education, Database, SQL

### [요 약]

본 연구에서는 2022 개정 교육과정을 기반으로 교육 프로그램을 설계하고, 25명의 초등 교사를 대상으로 적용하였다. 특히 데이터의 구조적 이해를 강화하기 위해 데이터베이스와 SQL 학습에 중점을 두었으며, SQL 질의 작성 과정에서 발생한 오류를 수집, 분석하였다. 교사들이 가장 빈번하게 범한 오류는 의미 오류(79.1%)였으며, 이는 문법 오류의 발생 빈도보다 약 4배 높았다. 또한 SQL의 구성 요소 중 절(clause)을 기준으로 오류를 분류한 결과, WHERE, GROUP BY/HAVING, ORDER BY에서 오류가 다수 발생하였으며, 주요 원인은 해당 절의 누락이었다. 본 연구의 결과는 향후 초등 교사를 대상으로 하는 데이터베이스 및 SQL 교육의 기초 자료로 활용될 수 있을 것으로 기대된다.

▶ **주제어:** 데이터 과학, 초등 교사, 오류, 교육, 데이터베이스, SQL

- 
- First Author: Junghee Jo, Corresponding Author: Junghee Jo
  - Junghee Jo (dreamer@bnue.ac.kr), Dept. of Computer Education, Busan National University of Education
  - Received: 2025. 09. 11, Revised: 2025. 09. 30, Accepted: 2025. 10. 15.

## I. Introduction

인공지능 연구가 급속도로 진행되고 있는 현 상황과 더불어 데이터 활용 역량은 특정 전공 영역을 뛰어넘어 누구나 함양해야 하는 필수 소양으로 인식되고 있다. 이에, 2022 개정 교육과정에서는 인공지능 소양 함양의 하나로 데이터 과학을 중/고등뿐만 아니라 초등교육에도 시행하도록 권고하고 있다[1]. 따라서, 초등 교사들도 기초 개념의 이해뿐만 아니라 실생활에서 발생하는 다양한 문제 해결을 위해 데이터를 활용할 수 있는 능력이 필요하게 되었다.

이와 관련하여, 교육 현장에는 초등학생을 대상으로 한 다양한 종류의 데이터 분석 도구가 급속도로 개발되어 출시되고 있다. 초등 교사들은 이러한 도구들을 단시간 내에 교육에 활용하기 위해 교원을 대상으로 실시하는 AI·디지털 역량 강화 연수에 참여하거나 대학원에 진학하고 있다. 데이터를 체계적으로 이해하여 활용하기 위해서는 근간이 되는 학문인 컴퓨터 과학(Computer Science)을 학습하여야 하는데 교육용으로 개발된 도구의 사용법을 익히는 것만으로는 데이터에 관한 깊은 이해를 끌어내기에는 한계가 존재한다[2].

일반 대학에서는 컴퓨터 과학 전공자의 양성을 위해 학부 과정에서 반드시 이수해야 할 핵심 교과목의 하나로 데이터베이스(Database) 및 SQL(Structured Query Language)을 교육하고 있다[3]. 데이터베이스는 데이터를 체계적으로 저장하고 관리하는 시스템으로 데이터 과학의 모든 분석은 데이터에 접근하고 추출하는 단계에서 시작되므로 데이터베이스를 제외하고 데이터를 논하기는 어렵다. 따라서, 데이터를 다루는 여러 전공에서는 데이터베이스를 기본적으로 교육하고 있다[4]. SQL은 데이터베이스로부터 데이터를 꺼내오기 위해 사용하는 기본 언어로 파이썬이나 자바와 같은 프로그래밍 언어와 비교해 직관적으로 구성되어 있으므로 비전공자가 습득하기에 수월하다.

미국, 캐나다, 핀란드의 경우, 초등학생을 대상으로 데이터베이스 교육이 다양한 교과와 연계하여 실시되고 있으나 국내는 아직 그렇지 못한 현실이다[5-7]. 그러나, 초등에 이러한 교육을 도입하여야 한다는 목소리는 지속해서 높아지고 있다. 이에, 초등 교사들은 전공자만큼은 아니더라도 데이터에 관한 구조적인 이해를 위해 데이터베이스 및 SQL 관련 지식을 습득하는 것이 필요하다.

기존 연구들은 주로 전공자나 비전공 대학생을 대상으로 이루어졌으므로 비전공자이면서 동시에 초등학생에게 교육 내용을 전달해야 하는 특수한 위치에 있는 초등교사를 고려한 연구는 찾아보기 어렵다.

본 논문에서는 이러한 기존 연구의 제약점을 고려하여 초등 교사를 대상으로 데이터 활용 역량 함양을 위한 단계적 학습 구조 기반의 교육 프로그램을 설계하였다. 이를 위해 블록형 프로그래밍 교육 플랫폼인 엔트리(Entry)의 데이터 분석 블록들을 이용하여 데이터 불러오기, 탐색하기, 분석하기, 시각화를 통해 데이터의 의미를 직관적으로 이해하도록 하였다. 그 후, 컴퓨터 과학에서 정의하는 데이터베이스의 기본 용어 및 개념, SQL 기반의 입출력, 그리고 실생활 데이터의 분석을 통해 데이터에 관한 체계적인 이해를 돕도록 하였다. 그리고 교사들 스스로 엔트리와 데이터베이스의 기능을 비교 분석하여 엔트리에 부가적으로 필요한 데이터 분석 블록을 도출하도록 구성하였다. 개발된 프로그램은 교육 현장에 적용되었고 교사들의 학습 수준을 평가한 결과를 수집, 분석하여 교육적 시사점을 도출하였다.

본 논문의 구성은 다음과 같다. II장에서는 관련 연구로써 엔트리의 데이터 분석 블록, 데이터베이스, SQL, 그리고 초등교사 대상의 데이터 교육에 관한 내용을 설명하였다. III장에서는 연구 대상자 및 연구의 세부 절차, 데이터 수집 과정에 관해 설명하였다. IV장에서는 데이터의 분석 절차 및 결과를 설명하였고 V장에서는 연구의 결과 및 한계점, 그리고 향후 계획에 관해 설명하였다.

## II. Related Works

### 1. Data Analysis Block of Entry

엔트리의 데이터 분석 블록은 초등학생이 데이터를 다루기 쉽도록 직관적으로 경험하도록 만든 블록형 코딩 도구이다[8]. CSV(Comma Separated Values) 파일과 같은 외부 데이터를 엔트리 플랫폼으로 불러올 수 있으며, 특정 행(row)이나 열(column)의 값을 가져올 수 있고 특정 셀의 값도 확인할 수도 있다. 또한, 합계, 평균, 최댓값, 최솟값 등의 기본적인 통계 연산 블록을 제공하고, 조건에 맞는 데이터만 선택하여 계산하거나 원하는 기준으로 데이터를 분류할 수 있다. 간단한 수준의 데이터 시각화 기능도 제공하여 막대, 선, 원그래프 등 기본 그래프를 쉽게 생성할 수도 있다.

### 2. Database and SQL

데이터베이스는 다량의 데이터를 체계적으로 저장하고, 필요할 때 원하는 데이터를 빠르고 정확하게 검색, 수정, 추가할 수 있도록 관리하는 시스템이다. 여러 사용자가 데

이터에 동시에 접근하여 수정할 수 있으며 데이터의 정확성을 보장하도록 설계되어 있다. 일반적으로 표(Table)의 형태로 구성되어 있으며 열(Column)은 데이터의 속성(Attribute)을, 행(Row)은 개별 데이터(Record)를 의미한다. 최근 들어, 데이터 리터러시(Data Literacy) 및 컴퓨팅 사고력(Computational Thinking)을 길러야 한다는 사회적 요구와 맞물려서 데이터 기반의 탐구학습을 위해 초등교육에도 그 중요성이 커지고 있다[5-7].

SQL은 데이터베이스에 저장된 데이터를 불러오고, 추가하고, 수정하고, 삭제하기 위한 표준 언어이다. 주요 기능으로는 데이터 조회(Select), 데이터 추가(Insert), 데이터 수정(Update), 데이터 삭제>Delete) 등이 있다. 기업이나 기관의 데이터는 대부분 데이터베이스에 저장되어 있으므로 분석을 위해서는 SQL이 필수적이며, 프로그래밍 언어 중에서도 문법이 비교적 단순하여 비전공자도 단시간 내에 습득할 수 있다. 비전공자의 SQL 습득은 데이터 리터러시와 분석적 사고를 키우는 데 중요한 출발점으로 인식되고 있으며 대표적인 연구 영역으로는 비전공자들이 발생시키는 다양한 SQL 오류의 분석이 있다[9-14].

### 3. Educational Program for Data Utilization

초등교사 대상의 데이터 활용 교육에 관한 선행 연구를 분석한 결과, 인공지능 분야에 비해 연구가 충분히 축적되지 않은 상태임을 확인할 수 있었다. 허경(2021)은 현직 및 예비 초등교사를 위해 스프레드시트 기반의 데이터 분석 도구를 활용한 교육과정을 개발하고 적용한 후, 만족도 조사를 통하여 프로그램을 평가하였다[17]. 조정희(2023)는 현직 초등교사를 대상으로 데이터를 이해하고 사용하는 능력을 배양하기 위해 파이썬을 활용한 시각화를 중심으로 교육 프로그램을 개발하고 이를 검증하였다[18].

초등학생을 대상으로 한 관련 연구는 초등교사 대상의 연구에 비해 상대적으로 활발히 이루어지고 있는 편이었다. 홍지연(2020)은 초등학생 대상의 데이터과학 교육 프로그램을 개발하고 효과성을 검증하였고[19], 백승인(2024)은 초등학생이 코딩 활동 없이 관심 주제를 선택하여 주도적으로 활동하는 데이터과학 교육 프로그램을 개발하고 적용하여 효과성을 측정하였다[20]. 박남제(2025)는 초등학생을 대상으로 메타데이터를 이용하여 파일을 복구하는 원리를 활용한 디지털 포렌식 교육 프로그램을 개발하고 그 효과를 분석하였다[21].

## III. Research Design

### 1. Research Subject

본 연구는 교원양성대학의 대학원에 재학 중인 초등 교사 25명을 대상으로 진행되었다. 학생들의 세부 전공은 영어교육, 수학교육, 과학교육, 음악교육 등 다양하게 분포되었으며 대부분이 교원연수 또는 동영상 강의를 통해 엔트리 또는 스크래치를 활용한 블록 프로그래밍 경험이 있었다. 그러나 엔트리의 데이터 분석 블록을 사용해 본 경험은 없었으며 데이터베이스 또는 SQL을 접해본 경험 또한 없었다.

### 2. Research Procedure

학생들은 15주 동안 데이터 활용 역량 함양을 위한 교육을 받았다. 세부적인 강의 내용은 Table 1과 같이 데이터 과학의 개요, 엔트리의 데이터 분석 블록, 데이터베이스, SQL, 정리, 그리고 평가로 구성하였다.

Table 1. The Curriculum Designed for the Lecture

Category	Contents
Introduction	Overview of data science
	Trends in elementary education
	The procedure of data science
Data analysis block of Entry	Overview of Entry
	Select, upload, and create data
	Visualize data
	Use data analysis blocks
Database	Analyze small data set
	Overview of database
	Install MySQL
SQL	Use menus of MySQL
	Overview of SQL
	CREATE TABLE, INSERT INTO
	SELECT, INSERT, UPDATE, DELETE
	ORDER BY, LIMIT, DISTINCT
	Aggregate functions(e.g., Max, Min)
	GROUP BY, HAVING
SubQuery	
Wrap up	Exercises
	Comparative analysis of data blocks of Entry and SQL
Evaluation	Formulation of SQL statements

강의 초반에는 데이터 과학의 전반적인 개요, 초등교육에서의 동향, 그리고 데이터 과학의 절차를 설명하였다. 그 후, 학생들에게 익숙한 엔트리의 전반적인 내용을 간략히 언급한 후 데이터 분석 블록의 사용법을 설명하였다. 데이터 블록의 실습 과정에서는 엔트리에서 기본적으로 제공하는 데이터를 먼저 사용해 보도록 하였고, 외부에서 원하는 데이터를 검색하여 엔트리 내부로 불러와서 활용하도록 하였으며, 엔트리 내에서 간단한 데이터를 직접 생

성하여 분석해보도록 하였다.

강의 중반에는 테이블, 속성, 레코드, 기본키(Primary key), 외래키(Foreign key)를 포함한 데이터베이스의 전반적인 개요를 설명한 후, 전 세계적으로 사용되고 있는 데이터베이스 관리 시스템인 MySQL을 설치하도록 하고 사용법을 설명하였다. MySQL은 별도의 라이선스 비용이 없고 표준 SQL을 지원하므로 비전공자들이 기본적인 기능을 학습하기에 비교적 최적화되어 있다.

그 후, 본격적으로 SQL의 개요를 설명하고 테이블 생성 구문부터 시작하여 데이터의 조회, 추가, 수정, 삭제를 위한 기본적인 문법을 설명하였다. 그리고 질의 결과를 정렬하거나 출력할 행의 개수를 제한하는 등 SQL 문장의 동작 방식을 지정하는 다양한 문법 요소들을 설명하였다. 그리고, 비전공자들에게는 난도가 다소 높은 편에 속할 수 있는 서브쿼리(SubQuery)를 설명하였다.

강의 마지막 시간에는 별도의 세션을 만들어서 학생들 스스로 엔트리와 데이터베이스의 기능을 비교 분석하여 엔트리에 부가적으로 필요한 데이터 분석 블록을 도출해 보도록 하였다.

### 3. Data Collection

교사들의 데이터베이스 및 SQL 학습 수준을 평가하기 위하여 본 연구에서는 SQL 구문 작성 시험을 실시하였으며, 해당 결과를 분석 자료로 활용하였다. 시험은 오픈북으로 진행되었고 인터넷 검색은 제한하였다.

Table 2. The database schema used for the questions. Underlined attribute is a primary key for each table

Table name	Attributes
userTable	<u>userID</u> , userName, birthYear, addr, mobile, height, mDate
buyTable	<u>num</u> , userID, prodName, groupName, price, amount

Table 3. Sample questions

Question type	Question
Formulation	Write a SQL statement that prints the names and birthdays of all users whose birthdays are in or before 1970.
Formulation	Write a SQL statement that outputs the name and price of the most expensive item purchased by the user with ID 'LHJ'.
Formulation	Write a SQL statement that outputs the ID and total purchase amount of users whose total purchase amount is 1,000 won or more.

비전공자임을 고려하여 MySQL 사용을 허용하였고 테이블 생성과 데이터 삽입을 위한 기본적인 SQL 구문은 교수가 제공하였다. 시험에 사용된 데이터베이스의 스키마와 샘플 문제는 각각 Table 2와 Table 3과 같다.

## IV. Results

### 1. Classification of Error Types

학생들의 SQL 구문 작성 평가를 통해 수집된 오류를 분석하여 분류한 결과는 Table 4와 같다. SQL 질의를 작성할 때 발생시키는 오류는 다양한 범주로 정리되어왔다 [11-14]. 본 논문에서는 이러한 선행 연구를 기반으로 초등 교사들이 발생시킨 오류들을 의미 오류(Semantic Errors)와 문법 오류(Syntax Errors)로 분류하였다. 그리고, 타입별로 학생들이 자주 발생시키는 오류들을 세부 항목으로 분류하고 식별을 위해 코드 항목을 생성하였다.

Table 4. Type, Code, and Description of Errors

Type	Code	Description
Semantic Errors	SM_1	Insufficient understanding of logical operators(e.g., AND, OR)
	SM_2	Insufficient understanding of the difference between ASC and DESC
	SM_3	Insufficient understanding of the usage of SUM
	SM_4	Insufficient understanding of the usage of LIMIT
	SM_5	Insufficient understanding of the grouping functions such as HAVING and GROUP BY
	SM_6	Insufficient understanding of the usage of ORDER BY
	SM_7	Insufficient understanding of the nested queries
	SM_8	Invalid column or table name
	SM_9	Insufficient understanding of what the problem requires
Syntax Errors	ST_1	Incorrect use of semicolons
	ST_2	Missing part of the quotation marks
	ST_3	Insufficient understanding of the comparison operations(e.g., <, >)
	ST_4	Insufficient understanding of the nested queries
	ST_5	Insufficient understanding of the grouping functions such as HAVING and GROUP BY
	ST_6	Insufficient understanding of the usage of DISTINCT
	ST_7	Misspellings

의미 오류는 실행 결과가 출력은 되지만 주어진 과제를 논리적으로 접근하지 못하여 의도하지 않은 결과를 생성하는 오류이다. 문법 오류는 SQL 구문을 문법 규칙에 맞지 않게 작성하여 실행이 이루어지지 않으며 오류의 내용이 명시적으로 출력된다.

## 2. Description of the Error

데이터베이스 및 SQL의 학습 과정에서 초등 교사들이 발생시킨 오류들을 수집하여 분류한 Table 4의 각 코드 항목에 관한 상세한 내용과 오류 발생 원인에 관한 설명은 아래와 같다.

**논리 연산(SM\_1):** AND, OR와 같은 논리 연산에 관한 이해 부족으로 의도하지 않은 실행 결과가 발생하였다.

```
select  userName, height
from    userTable
where   height=178 and height=180
        and height=182
order  by height;
```

Fig. 1. Example of SM\_1, insufficient understanding of logical operators such as AND and OR

예를 들어, '키가 178, 180, 182 중 하나인 사용자의 이름과 키를 출력하는 SQL 문을 작성하세요. (단, 키를 기준으로 오름차순으로 정렬하여 출력하세요)'의 문항에 대하여 Fig. 1 같이 OR 대신에 AND를 사용함으로써 '키가 178이고 180이고 182'라는 의도하지 않은 조건을 검색하는 오류를 생성하였다.

**정렬(SM\_2):** ASC(오름차순), DESC(내림차순)와 같은 정렬 옵션에 관한 이해 부족으로 인하여 의도하지 않은 방법으로 결과가 정렬되는 오류를 생성하였다.

**SUM(SM\_3):** 집계함수 중의 하나인 SUM의 사용법을 정확히 인지하지 못하여 오류를 생성하였다.

**LIMIT(SM\_4):** 질의 결과로 출력되는 행의 개수를 제한하는 키워드인 LIMIT의 사용법을 인지하지 못하여 오류가 발생하였다. 예를 들어, '회원 테이블(userTable)에서 이름(userName)을 조회하는 SQL을 작성하세요. (단, 가입일(mDate)이 빠른 순서로 정렬하고 세 번째 가입한 사람부터 2명만 출력할 것)'의 문항에 대하여 Fig. 2의 'limit 3,2'는 네 번째 가입한 사람부터 2명만 출력하라는 의미가

므로 '세 번째 가입한 사람부터 2명만 출력'이라는 원래의 의도와는 다른 실행 결과가 발생하였다.

```
select  userName
from    userTable
order  by mDate
limit  3,2;
```

Fig. 2. Example of SM\_4, insufficient understanding of the usage of LIMIT

**그룹화(SM\_5):** HAVING, GROUP BY를 사용한 그룹화 기능을 이해하지 못하여 의도하지 않은 실행 결과가 발생하였다. 오류를 발생시킨 학생들의 대다수(80%)가 HAVING 또는 GROUP BY 구문 자체를 빠뜨렸으며 문법적으로는 오류가 없으므로 실행은 되었으나 주어진 문항의 요구와는 다른 결과가 출력되었다.

**ORDER BY(SM\_6):** 질의 결과를 특정 속성을 기준으로 정렬하는 ORDER BY의 기능을 이해하지 못하여 오류를 발생시켰다.

```
select  prodName, price
from    buytable
where   (prodName, price) in (
select  prodName, max(price)
from    buytable
where   userID = 'LHJ');
```

Fig. 3. Example of SM\_6, insufficient understanding of the usage of ORDER BY

단순한 경우는 ORDER BY 구문 자체를 빠뜨리거나 ORDER BY 뒤에 기재하는 조건에 오류가 있는 구문이었고, 복잡한 경우는 Fig. 3과 같이 ORDER BY를 사용하지 않고 다른 방법으로 해결하고자 하는 과정에서 오류가 발생하였다. Fig. 3은 '아이디가 'LHJ' 인 사용자가 구매한 물건 중에 가장 비싼 물건의 이름과 가격을 출력하는 SQL 문을 작성하세요.'의 문항에 일부 학생들이 제출한 답안이다. '가장 비싼 물건'이라는 조건 검색을 위해 ORDER BY 구문을 사용하면 간단하게 해결할 수 있는데 난도가 있는 중첩 질의를 사용하려고 시도하다가 오류가 발생하였다.

**중첩 질의(SM\_7):** 질의 안에 또 다른 질의가 포함된 형태로서 단순 질의보다 학생들이 어려워하는 구조이다. 복잡한 조건 처리에 사용을 하다 보니 충분한 이해 부족으로 인하여 Fig. 3과 같은 의도하지 않은 검색 결과가 출력되는 오류가 빈번히 발생하였다.

**속성 명, 테이블 명(SM\_8):** 존재하지 않는 테이블이나 속성을 참조할 때 발생하는 오류이다. 이 오류는 SQL 문법에는 문제가 없지만 참조하려는 테이블이나 속성이 존재하지 않을 때 발생한다.

```
select userName, height
from buyTable
where height in(178, 180, 182)
order by height;
```

Fig. 4. Example of SM\_8, invalid column or table name

```
select useName, height
from userTable
where height in(178, 180, 182)
order by height;
```

Fig. 5. Example of SM\_8, invalid column or table name

예를 들어, Fig. 4에서 Table 2에 의하면 buyTable에는 userName과 height 속성이 존재하지 않는데 이러한 경우에 발생하는 오류이다. 또한, Fig. 5와 같이 속성 명인 'useName'을 'userName'으로 잘못 기재한 경우도 이에 해당한다. 이러한 경우, MySQL에서는 Error Code와 함께 오류 메시지를 출력하므로 문법 오류로 여길 수 있으나 SQL 문법적인 측면에서는 문제가 없으므로 의미 오류에 해당한다[14-16].

**문제 해석(SM\_9):** 문제가 요구하는 바를 정확히 이해하지 못하여 전혀 다른 구문을 작성한 경우에 해당된다.

**세미콜론(ST\_1):** 한 구문의 끝을 의미하는 구분자 역할을 하는 세미콜론과 관련된 오류이다. MySQL에서는 구문 끝에 세미콜론이 없어도 편의상 실행이 되지만 Oracle을 포함한 다른 데이터베이스 시스템에서는 오류를 발생시킨다.

```
select userName, birthYear
from userTable;
where birthYear <= 1970;
```

Fig. 6. Example of ST\_1, incorrect use of semicolons

본 연구에서 수집한 데이터에 의하면 Fig. 6과 같이 SQL 문장의 끝뿐만 아니라 중간에도 세미콜론을 표기하여 오류가 발생하였다.

**따옴표(ST\_2):** 데이터와 식별자를 구분하는 역할을 하는 따옴표와 관련된 오류이다. 수집한 데이터에 의하면 시작 따옴표는 있는데 끝 따옴표가 없는 경우로 인하여 오류

가 발생하였다.

**비교 연산(ST\_3):** 속성값과 특정 값, 혹은 두 표현 식을 비교하여 조건의 참, 거짓을 평가하는 역할을 수행하는 비교연산과 관련된 오류이다.

```
select userName
from userTable
where height >= 180 and < 182;
```

Fig. 7. Example of SM\_8, invalid column or table name

예를 들어, '키가 180 이상이고 182 미만인 모든 사용자의 이름을 출력하는 SQL 문을 작성하세요.'라는 문항에 대하여 'height >= 180 and height < 182'의 조건을 Fig. 7과 같이 잘못 표기하여 오류가 발생하였다.

**중첩 질의(ST\_4):** SM\_7의 경우 문법적 오류는 없으므로 실행은 가능하나, 제시된 문항의 요구와는 다른 결과를 산출한다. 반면, ST\_4는 중첩 질의를 구성하는 과정에서 문법 오류가 발생하여 실행 자체가 불가능한 사례에 해당한다.

**그룹화(ST\_5):** HAVING, GROUP BY를 사용한 그룹화의 과정에서 발생하는 오류로써 이미 설명한 SM\_5와 달리, 집계함수 없이 HAVING을 사용하는 등의 문법적 오류로 인하여 실행이 되지 않는 경우이다.

**DISTINCT(ST\_6):** 질의 결과에서 중복된 값을 제거하고 유일한 값만 반환하도록 하는 키워드인 DISTINCT와 관련된 오류로써, 괄호와 결합하여 잘못 사용하는 등의 문법적 오류로 인하여 실행이 되지 않는 경우이다.

**스펠링(ST\_7):** 문법 오류에서 가장 단순한 오류로써 스펠링 오류로 인하여 발생하였다.

```
select userID, sum(price*amount)
from buyTable
group by userID
haveing sum(price*amount) >= 1000
order by userID desc;
```

Fig. 8. Example of ST\_7, misspellings

예를 들어, '총 구매금액이 1000 이상인 모든 사용자의 아이디와 총 구매금액을 출력하는 SQL 문을 작성하세요. (단, 사용자의 아이디를 기준으로 내림차순으로 정렬할 것)'의 문항에 대하여 Fig. 8과 같이 'having'을 'haveing'으로 잘못 표기하여 문법 오류가 발생한 경우에 해당한다.

### 3. Analysis of Error Rate

학생들이 제출한 SQL 구문 작성 시험에서 수집한 오류들을 Table 4에서 제시한 오류 타입별로 분류하고 각 항목에 따른 발생 횟수를 Table 5와 같이 정리하였다.

수집된 전체 오류 데이터 중에서 의미 오류는 79.1%, 문법 오류는 20.9%이었으며, 의미 오류가 문법 오류보다 약 4배 정도 많이 발생하였다. 문법 오류의 경우는 항목별로 1건 또는 2건이 발생하여 항목별 차이가 미미하였지만, 의미 오류는 항목별로 차이가 존재하였다.

의미 오류중에서 가장 많이 발생한 오류는 ORDER BY(SM\_6)와 관련된 것으로서 수집된 전체 의미 오류중에 총 20.6%에 해당하였다. 주로 ORDER BY 뒤에 따르는 조건 기술에서 오류가 발생하였고 일부 오류는 구문 자체가 생략된 경우였다. 다음으로 많이 발생한 의미 오류는 속성명/테이블 명 오류(SM\_8, 17.6%) 및 HAVING, GROUP BY를 사용한 그룹화 구문(SM\_5, 14.7%)과 관련된 오류였다.

Table 5. Frequency of Errors by Type and Code

Type	Code	Frequency
Semantic Errors	SM_1	3(8.8%)
	SM_2	2(5.9%)
	SM_3	3(8.8%)
	SM_4	4(11.8%)
	SM_5	5(14.7%)
	SM_6	7(20.6%)
	SM_7	3(8.8%)
	SM_8	6(17.6%)
Syntax Errors	SM_9	1(2.9%)
	ST_1	2(22.2%)
	ST_2	1(11.1%)
	ST_3	1(11.1%)
	ST_4	2(22.2%)
	ST_5	1(11.1%)
	ST_6	1(11.1%)
	ST_7	1(11.1%)

Table 6. Frequency of Errors Classified based on SQL Clauses

Clause	Mistakes	Frequency
SELECT	Using wrong columns, Using unnecessary aggregate functions(e.g., MAX, MIN)	3(9.7%)
FROM	Missing FROM clause, Invalid table name	2(6.5%)
WHERE	Missing/Wrong condition	8(25.8%)
GROUP BY, HAVING	Missing HAVING/GROUP BY clauses, Wrong condition	7(22.6%)
ORDER BY	Missing ORDER BY clauses, Wrong ORDER BY keywords(e.g., ASC, DESC)	7(22.6%)
LIMIT	Wrong condition	4(12.9%)

Table 6은 SQL 구문을 구성하는 여러 가지 요소 중에 절(Clause)을 기준으로 수집한 데이터를 기반으로 오류가 발생한 횟수를 정리한 것이다. 오류가 가장 자주 발생한 절은 WHERE(25.8%)가 해당하였고, 다음으로 GROUP BY/HAVING(22.6%)과 ORDER BY(22.6%)가 해당하였다. 오류가 발생한 주요 원인은 모두 해당 절의 ‘누락(Omission)’이었다. 이는, 선행연구[12]에서 9년동안 수집한 2300명의 초보 프로그래머 학생의 약 16만 개 SQL 구문을 분석한 결과와도 일치한다. 선행연구에서는 JOIN, 서브쿼리, GROUP BY에서 누락 오류가 빈번함을 확인하였는데, 본 연구의 학생들도 집계 함수 사용 시 GROUP BY 절을 생략하거나, 조건문 작성 과정에서 HAVING 절을 누락하는 오류가 빈번히 발생함을 확인하였다.

Table 7. Frequency of Errors Classified based on Major SQL Clauses: Semantic vs Syntax Errors

	Semantic Errors	Syntax Errors
WHERE	5(62.5%)	3(37.5%)
GROUP BY, HAVING	5(83.3%)	1(16.7%)
ORDER BY	7(100%)	0(0.0%)

Table 7은 수집한 데이터를 기반으로 오류가 가장 자주 발생한 절인 WHERE, GROUP BY/HAVING, ORDER BY에 한하여, 해당 오류가 의미 오류와 문법 오류 중 어디에 해당하는 것인지 분석한 결과이다. 특히, ORDER BY와 관련하여 발생한 오류는 모두 의미 오류(100%)에 해당하였으며 문법 오류로는 발생하지 않았다. 이는, ORDER BY가 WHERE 또는 GROUP BY/HAVING에 비하여 비교적 문법의 난도가 낮기 때문으로 추측된다.

## V. Conclusions

초등 교사들의 데이터에 관한 체계적인 이해 및 활용을 목적으로 설계한 프로그램에서 데이터베이스 및 SQL의 학습 수준의 평가 과정에서 수집한 데이터를 기반으로 도출한 결과는 다음과 같다.

첫째, 초등 교사들이 가장 많이 발생시킨 SQL 오류는 의미 오류였으며 문법 오류보다 약 4배 정도 많이 발생하였다. 문법 오류는 데이터베이스 시스템에서 즉각적으로 오류 메시지를 통해 피드백을 주지만, 의미 오류는 질의가 실행되면서 결과가 출력되므로 이러한 오류를 즉각적으로 인지하기 위해서는 높은 수준의 프로그래밍 역량이 요구된다.

의미 오류가 문법 오류보다 훨씬 더 빈번하게 나타난다는 결과는 단순한 SQL 문법 지식 전달을 넘어서 의미 오류의 유형과 원인을 집중적으로 다루는 교수 전략이 필요함을 시사한다고 할 수 있다. 특히, 본 연구를 통해 오류 발생의 주요 원인으로 파악된 ORDER BY, GROUP BY/HAVING, 속성 및 테이블 명과 관련한 오류 사례를 반복적으로 교정하는 활동을 통해 구문의 개념적 이해를 높이는 데 도움이 될 것으로 보인다.

둘째, SQL을 구성하는 요소 중에 절을 기준으로 초등 교사들이 발생시킨 오류를 분석해보니 WHERE, GROUP BY/HAVING, ORDER BY에서 많이 발생하였고 주요 원인은 해당 절의 ‘누락’이었다. ‘누락’ 오류가 많이 발생한 것은 학습자들이 질의 구조를 올바르게 설계하는 데 어려움을 겪고 있음을 보여준다. 이를 개선하기 위해 특정 절을 생략할 경우 잘못된 결과가 산출되는 사례를 학습자가 직접 경험하게 하여 각 절의 필요성을 인식하도록 하는 교육적 접근이 효과적일 수 있다.

셋째, 비전공자인 교사들의 특성을 고려할 때 반복 학습과 단계적 지원이 필요하다. 초기 학습 단계에서는 모든 절을 포함한 문법 서식을 제공하여 연습하게 하고, 이후 점차 불필요한 절을 생략하도록 유도함으로써 학습 부담을 줄이면서도 구조적 이해를 심화할 수 있다.

마지막으로, 데이터베이스/SQL 교육은 교사들의 엔트리 활용 기반 데이터 분석 역량을 긍정적으로 향상시켰다. 교사들은 엔트리에서 데이터셋을 처리하는 과정이 SQL의 SELECT 구문과 유사함을 인식하였으며, 이를 토대로 새로운 데이터 분석 블록을 구상할 수 있는 통찰력을 형성하였다.

본 연구는 단일 대학에 소속된 25명의 초등 교사들을 대상으로 연구를 수행하였기에 결과를 일반화하기에는 무리가 있다. 또한, 학생들의 전반적인 데이터 활용 역량을 보다 정확하게 파악하기 위해서는 교육 이전과 이후의 변화를 비교할 필요가 있으나 본 연구는 교육 후 평가 결과에 국한하여 분석하였다는 점에서 평가 범위에 한계가 존재할 수 있다. 이에, 향후에는 다양한 환경의 초등 교사들을 대상으로 사전·사후 검사를 병행하여 학생들의 데이터 활용 능력 변화를 체계적으로 분석함으로써, 본 연구의 한계를 보완하고자 한다.

## REFERENCES

- [1] Ministry of Education, 2022 Revised Curriculum, <https://www.moe.go.kr/boardCnts/viewRenew.do?boardID=141&lev=0&status>
- [2] S.H. Kim, S.H. Kim, M.J. Lee, and H.C. Kim, “Review on Artificial Intelligence Education for K-12 Students and Teachers,” *Journal of the Korean Association of Computer Education*, Vol. 23, No. 4, pp. 1-11, 2020.
- [3] S.H. Choi, “A Curriculum to Improve the Lecture of Database SQL,” *Journal of the Korean Institute of Electronic Communication Sciences*, Vol. 18, No. 9, pp. 1005-1010, 2014.
- [4] J.Y. Ahn, K.S. Han, and S.H. Choi, “A Note on Database Education in Statistics Undergraduate Course,” *Journal of Applied Statistics*, Vol. 17, No. 3, pp. 595-603, 2004.
- [5] S.R. Jeon, S.H. Kim, D.S. Nam, and T.W. Lee, “Proposals for Database Education for Elementary School Students,” *Proceedings of the Korean Society for Computer and Information Conference (Summer)*, Vol. 20, No. 2, pp. 179-182, July 2012.
- [6] H.Y. Kim, H.S. Lee, and J.S. Kim, “A Study on Database Education in Korean Elementary and Secondary Schools Focused on Gyeongsangbuk-do,” *Journal of the Korean Society for Industrial Education*, Vol. 32, No. 2, pp. 132-154, 2007.
- [7] S.R. Jeon, S.H. Kim, D.S. Nam, and T.W. Lee, “Development of a Problem-Based Learning (PBL) Based Database Learning Program for Elementary Gifted Students,” *Proceedings of the Korean Society for Computer Education Conference (Summer)*, Vol. 16, No. 2, pp. 225-228, Aug. 2012.
- [8] K. Hur, “A Study on Elementary Education Examples for Data Science using Entry,” *Journal of the Korean Association of Information Education*, Vol. 24, No. 5, pp. 473-481, 2020.
- [9] D. Miedema, E. Aivaloglou, and G. Fletcher, “Identifying SQL Misconceptions of Novices: Findings from a Think-Aloud Study,” *Proceedings of the 17th ACM Conference on International Computing Education Research*. ACM, New York, pp. 355-367, 2021.
- [10] T. Taipalus, M. Siponen, and T. Vartiainen, “Errors and Complications in SQL Query Formulation,” *ACM Transactions on Computing Education*, Vol. 18, No. 3, Article 15, Sep. 2018.
- [11] D. Miedema, G. Fletcher, and E. Aivaloglou, “Expert Perspectives on Student Errors in SQL,” *ACM Transactions on Computing Education*, Vol. 23, No. 1, Article 11, pp. 1-28, Dec. 2022.
- [12] A. Ahadi, J. Prior, V. Behbood, and R. Lister, “Students’ Semantic Mistakes in Writing Seven Different Types of SQL Queries,” *Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education*, New York, USA, pp. 272-277, July 2016.
- [13] A. Ahadi, V. Behbood, A. Vihavainen, J. Prior, and R. Lister, “Students’ Syntactic Mistakes in Writing Seven Different Types of SQL Queries and Its Application to Predicting Students’ Success,” *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, New York, USA, pp. 401-406, Feb. 2016.

- [14] S. Brass and C. Goldberg, "Semantic Errors in SQL Queries: A Quite Complete List," *Journal of Systems and Software*, Vol. 79, No. 5, pp. 630-644, May 2006.
- [15] S. Poulsen, L. Butler, A. Alawini, and G. L. Herman, "Insights from Student Solutions to SQL Homework Problems," *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE '20)*, Norway, pp. 404-410, June 2020.
- [16] K. Presler-Marshall, S. Heckman, and K. T. Stolee, "SQLRepair: Identifying and Repairing Mistakes in Student-Authored SQL Queries," *Proceedings of the 43rd IEEE/ACM International Conference on Software Engineering: Software Engineering Education and Training (ICSE SEET '21)*, pp. 199-210, 2021.
- [17] K. Hur, "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*, Vol. 25, No. 3, pp. 459-469, 2021.
- [18] J.H. Jo, "Development and Application of Python-based Courses for Improving Data Science Competence of Elementary Teachers", *Journal of The Korean Association of Information Education*, Vol. 27, No. 1, pp. 47-55, 2023.
- [19] J.Y. Hong and Y.S. Kim, "Development of AI Data Science Education Program to Foster Data Literacy of Elementary School Student", *Journal of The Korean Association of Information Education*, Vol. 24, No. 6, pp. 633-641, 2020.
- [20] S.I. Baek and Y.B. Lee, "Design and Implementation Effects of Data Science Education Program for Elementary School Students", *Journal of The Korean Association of Information Education*, Vol. 28, No. 5, pp. 633-643, 2024.
- [21] P.R. Kim and N.J. Park, "Development and Effectiveness Analysis of a Metadata-Based Digital Forensics Education Program Utilizing Unplugged Activities", *Journal of The Korean Association of Information Education*, Vol. 29, No. 2, pp. 183-193, 2025.

## Authors



Junghee Jo received the Ph.D. degree in Computer Science from University of Massachusetts Amherst, USA, in 2014. She had worked at LG Electronics and Electronics and Telecommunications Research

Institute (ETRI), respectively. She is currently an associate professor in the Department of Computer Education, Busan National University of Education, Korea. Her research interests include computer education, human factors, and health informatics.