

A Meta-Analysis of Represented Types of Data Visualization in Youth: Awards in the National Student Statistics Utility Competition from 2019 to 2021

MinYoung Baek¹, EunRyung Hyun^{2*}

¹Department of Learning Science, Major in Cognitive Information Design, Graduate School, Student of Doctor Course, Hanyang University, Seoul, Korea

²Department of Applied Art Education, College of Education & Department of Learning Science, Major in Cognitive Information Design, Graduate School, Associate professor, Hanyang University, Seoul, Korea

Abstract

Background The purpose of the study is to derive characteristics and concerns through an empirical case study utilizing data visualization charts of South Korean youth and to determine the necessity and importance of educating data visualization in design for youth.

Methods The theoretical research method contemplated the concepts, processes, and types of charts of data visualization based on preceding research that related to data visualization. The analysis from 1,387 charts that scored 117 by middle school students in the National Student statistics utility competition by Statistics Korea from 2019 to 2021, was classified and analyzed into six categories: comparison, trend, distribution, relation, location data, and other to identify the characteristics and concerns.

Results The data was unequally distributed to 'distribution' and 'comparison' visualized charts regardless of the characteristics of the data and the purpose of the data collection. There were also errors in the chosen chart that was inadequate to the purpose of types of charts, errors of visual expression that was difficult to analyze due to many unnecessary visual expressions without a classified scale, and errors of a proportion expression that showed the data value was either exaggerated or diminished.

Conclusions Youth must experience appropriate visualization images to express accurate data visualization by comprehending contents and purposes of data. The data has already drawn attention in each field as an important resource for future society. For this reason, it is time to link design researchers with an educational program in various data visualizations to promote data literacy for youth.

Keywords Data Visualization, Data Visualization Design, Data Visualization Education

*Corresponding author: EunRyung Hyun (pariosa@hanyang.ac.kr)

Citation: Baek, M. Y., & Hyun, E. R. (2023). A Meta-Analysis of Represented Types of Data Visualization in Youth: Awards in the National Student Statistics Utility Competition from 2019 to 2021. *Archives of Design Research*, 36(2), 285-301.

<http://dx.doi.org/10.15187/adr.2023.05.36.2.285>

Received : Aug. 04. 2022 ; **Reviewed :** Jan. 11. 2023 ; **Accepted :** Jan. 11. 2023

pISSN 1226-8046 **eISSN** 2288-2987

Copyright : 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 educational and non-commercial use, provided the original work is properly cited.

1. 서론

1. 1. 연구의 배경 및 목적

본 연구는 빠르게 변화하는 사회 환경과 IT 기술의 융합으로 그 중요도와 활용도가 급격히 확장된 데이터 시각화의 체계적인 교육 방법을 제안하기 위하여 시작되었다. 김현정(Kim, 2018)은 막대한 정보를 담고 있는 빅 데이터는 그 속성으로 '크기, 속도, 다양성'을 포함한다고 보았다. 따라서 이러한 데이터는 단순히 숫자나 글자를 색이나 이미지를 통해 표현하는 것이 아니라, 각각의 정보를 효과적으로 시각화할 수 있는 기획과 분석을 통한 디자인 표현 과정이 포함되어야 한다. 김명신(Kim, 2022)에 의하면 데이터 시각화란 다양하고 방대한 데이터 자료 중 특정한 데이터를 택하여 조직화하고 시각적으로 표현함으로써 의미 있는 정보로 전달하는 연구 영역이라고 하였다. 따라서 데이터 시각화 과정에서는 정보 수용자가 데이터 분석 결과를 쉽게 이해할 수 있도록 핵심 개념 및 아이디어를 효과적으로 시각화할 수 있는 전략이 필요하다.

최근 데이터를 활용한 사회, 경제, 문화 등의 부가가치 창출이 늘면서 데이터 시각화 또한 경영, 언론, 산업계, 과학계 등 다양한 분야에서 주목 받고 있다(김현정, 2018). 이에 따라 대학을 비롯한 공공기관에서도 데이터 시각화에 대한 교육과정 및 공모전을 개설하는 추세가 있다. 서울의 H대학의 경우 2021년부터 대학원 과정에서 '인지정보 데이터 시각화' 과목을 운영하고 있으며, 서울디지털재단은 2019년부터 '도시문제 분석을 위한 데이터 시각화 및 탐색 과정'을 개설하여 운영 중이다. 한국디자인진흥원(KIDP)은 2020년 '데이터비즈 인포그래픽 공모전'을 통해 정부와 공공기관의 공공데이터를 비즈니스에 활용할 수 있는 데이터 시각화 공모전을 시행한 바 있다. 이러한 분위기를 반영하여 2017년 통계청은 데이터 시각화 도구인 '통그라미(사이트)'를 제공하고 있다. 통그라미는 K-plot의 새로운 버전으로 일반 통계 패키지에서는 지원되지 않는 초·중등 교과서 수록 그림이 추가되어 실 수업과도 연계하여 데이터를 시각화할 수 있는 장점이 있지만(이정진 외, 2014), 사용자의 의도에 맞는 적절한 차트 선택 작업은 사용자의 기반 지식에 의존할 수밖에 없는 상황(변정운 & 박용범, 2016)이므로 데이터를 다양한 관점에서 분석하고 시각화하는 데는 다소 부족함이 있다. 야우(Yau, 2013)는 양질의 데이터 시각화 과정은 통계 지식뿐만 아니라 디자인 지식이 동시에 동반될 때 그 효과가 극대화된다고 보았다. 그는 통계 지식에 기반 하지 않은 시각화 산출물은 단지 미적 목표 달성을 위한 습작이 될 뿐이고, 디자인 지식이 기반되지 않으면 그저 분석 산출물에 불과하다고 주장하면서, 포편적인 데이터 시각화 도구로 사용되고 있는 엑셀(Excel) 프로그램 사용 시에는 별도의 차트 환경 설정 조정이 필요하다고 보았다. 변정운(Byun, 2016) 역시, 다양한 데이터 시각화 툴에서 제공하는 차트 추천 방식은 사용자의 의도를 고려하지 않은 상태로 차트를 추천한다고 하였다. 따라서 자신의 의도에 맞게 효과적인 데이터 시각화를 표현하기 위해서는 차트를 생성하는 시각화에 대한 지식이 필요하다(Mackinlay et al., 2007). 하지만 지금까지의 데이터 시각화 교육 관련 선행 연구들을 살펴보면, 연구 분석의 대상이 초등학교에 편중되어 있었으며(김민범, 2020; 김민규, 2020; 이진영, 2021), 청소년을 대상으로 하는 데이터 시각화 연구와 교육과정은 드문 상황이다. 따라서 본 연구에서는 우리나라 청소년을 위한 데이터 시각화 교육과정 개발에 대한 첫 단계로서, '전국학생통계활용대회(통계청)'의 최근 3년(2019~2021)간의 수상작을 중심으로 우리나라 청소년의 데이터 시각화 표현 유형에 대한 메타 분석을 실시하고자 한다.

1. 2. 연구의 방법 및 내용

본 연구는 이론적 연구와 실증 분석 연구로 진행된다. 이론적 연구는 데이터 시각화에 관련된 선행 연구 자료를 바탕으로 데이터 시각화의 개념, 과정, 데이터 시각화 차트 유형을 중심으로 탐색한다. 실증 분석으로는 최근 3년(2019~2021)간 통계청에서 주관한 '전국학생통계활용대회'의 중학생부 수상작에 나타난 데이터 시각화 표현 유형을 분석하였다. 분석 자료는 통계청활용대회 웹사이트(<http://통계활용대회.kr>)에 수록된 중학생부 수상작 117점이다.

분석 도구로는 아벨라(Abela, 2011)의 목적에 따른 차트 선택 방법과 강원양 외(Kang et al., 2020)가 제시한 '5가지(비교, 추이·트렌드, 구성 비중·분포, 관계, 위치 데이터) 데이터 시각화 유형 분류 기준'을

수정·보완하여 ‘6가지(비교, 추이·트렌드, 구성 비중·분포, 관계, 위치 데이터, 기타)’ 데이터 시각화 유형 분류 기준으로 재구성하여 활용하였다.

2. 이론적 배경

2. 1. 데이터의 시각화

(1) 데이터 시각화의 개념

데이터 시각화는 ‘단순한 기호, 문자, 숫자들을 조직화하여 그림이나 그래프 등을 통해 보다 쉽게 이해할 수 있는 형태로 변환하는 과정’을 의미하는데, 이때 데이터는 ‘시각화’를 통해 정보로서의 가치를 가지고, 지식으로서 의사 결정이나 지적 발전의 역할을 확대하여 수행하는데 도움을 주는 전략으로 보았다(계보경 & 박연정, 2017). 카이로(Cairo, 2019)에 의하면 데이터 시각화란 ‘데이터의 분석, 탐색, 발견을 가능하게 하기 위해 디자인된 데이터 표현’이다. 그는 데이터 시각화의 주된 목적은 정보 수용자로 하여금 어떤 결론을 직접 도출할 수 있도록 하는 도구로 인식되는 경우가 더 많으며, 단편적인 디자이너의 메시지 전달이 주목적이지 아니라고 하였다. 카드 외(Card et al., 1999) 역시 시각화의 주된 목적은 시각화 그 자체가 아니라 시각화를 통한 새로운 발견 또는 의사 결정을 위한 통찰력에 있다고 보았으며, 정보 시각화란 추상 데이터에 대한 인간의 인지능력을 최대화하기 위한 컴퓨터와 상호 작용성을 기반으로 하는 시각적 재현이라 하였다. 많은 연구들이 데이터 시각화, 인포그래픽, 정보 시각화와 같은 용어에 대해서 새로운 해석과 범주화를 시도하고 있으나 모호한 개념 차이를 보인다고 하였다(곽민구, 2018).

이처럼 데이터 시각화는 현재로서는 연구자간 모호한 개념의 차이가 존재하고 광범위하게 해석될 수 있으나, 본 연구에서는 이러한 관점 중 비교적 데이터 시각화의 구분과 내용이 구체적인 후지 토시쿠니, 와타나베 료이치(Fuji & Watanabe, 2020)의 데이터 시각화 개념을 데이터 시각화라 규정하고 그 기준을 통해 연구 대상인 데이터 시각화 수상작을 분류하였다. 그들에 의하면 데이터 시각화는 차트를 이용하여 데이터를 알기 쉽게 시각화하는 것으로 사실과 주장에 의해 [Figure 1]과 같이 인포메이션 디자인과 데이터 아트 디자인으로 나눌 수 있다.

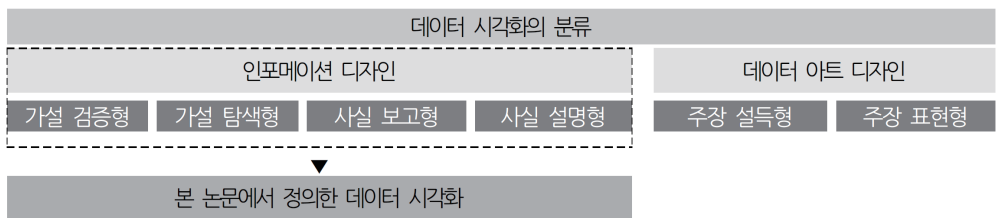


Figure 1 Classification of Data Visualization (Fuji Toshikuni & Watanabe Ryoichi, 2020)

[Figure 1]에서 데이터를 시각화할 때 자신을 포함한 조직과 상대방의 과제해결이 목적이면 ‘인포메이션 디자인’, 자신의 주장 및 표현을 위한 데이터의 시각화가 목적이면 ‘데이터 아트 디자인’으로 분류된다. ‘인포메이션 디자인’은 사실 검증이 필요한 ‘가설 검증형’과 ‘가설 탐색형’, 사실을 다루는 ‘사실 보고형’과 ‘사실 설명형’으로 나눌 수 있다. ‘데이터 아트 디자인’은 자신의 주장 및 표현을 독자에게 이해, 납득시키기 위한 ‘주장 설득형’과 특정한 대상 없이 자기표현에 중점을 두는 ‘주장 표현형’으로 구분된다(Fuji Toshikuni & Watanabe Ryoichi, 2020).

통계 교육원에 따르면 전국학생통계활용대회의 수상작은 여러 관점에서 문제에 접근하고, 그 과정을 통해 해답을 찾아 분석한 자료를 시각적으로 표현하는 것을 말한다. 따라서 전국학생통계활용대회의 수상작은 문제해결을 위한 데이터 시각화의 유형으로 분류되며 본 연구에서는 이를 ‘인포메이션 디자인 영역’의 통계로 분류하고 연구를 진행하고자 한다.

(2) 데이터 시각화의 과정

데이터 시각화의 과정은 정보화되기 이전의 상태, 즉 자료 형태의 다양한 데이터 중 어떤 종류를 모을 것인지, 시각화 이후 어떤 결과를 유도, 기대할 것인지를 의미 있게 설계·디자인하는 작업이며 결과를 보여주는 것이 아니라 생각을 유도해 나가는 것이다(민세희, 2013; 김현정, 2018 재인용). 또한 야우(Yau, 2013)에 의하면 일반적으로 시각적인 데이터를 탐색할 때는 [Figure 2]와 같이 네 단계를 거친다. 즉 반복적 데이터 탐색 과정의 중요성을 강조한 것이다.

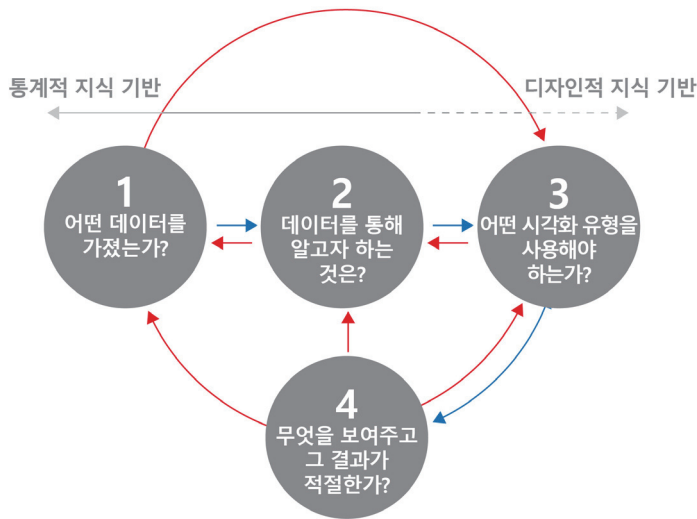


Figure 2 Reconstructing the Iterative Data Exploration Process through Prior Research (Nathan Yau, 2013)

[Figure 2]에서와 같이 데이터 시각화의 1~2단계는 통계적 지식을 기반으로 데이터를 탐색하며, 3단계는 디자인적 지식을 기반으로 데이터의 기하 형태와 색상에 매핑시키는 시각화 작업이 중심이 된다. 위치, 길이, 각도, 방향, 면적과 부피, 색상 등에 의해 적절한 시각적 요소의 선택은 중요하게 다루어지며, 시각 태스크들 또한 부호화하여 표현할 수 있다. 4단계는 통계적 지식과 디자인적 지식이 동반되어 결과가 이치에 맞는지와 그 결과가 적절한지 등을 결과로서 종합하는 단계이다(Yau, 2013). 이 과정에서 오류나 부족한 점이 발견되면 순환구조를 통해 전 단계로 이전하여 데이터 시각화 과정을 재구성할 수 있다.

결과적으로 데이터 시각화는 데이터 시각화 디자인 요소, 즉 위치, 길이, 각도, 방향, 면적과 부피, 색상 등 시각적 속성의 이해가 수반될 때 더욱 더 효과적으로 제안될 수 있다. 이는 복잡하고 방대한 데이터를 목적에 맞는 내용으로 강력하게 보여주기 때문이다.

(3) 데이터 시각화의 차트 유형

데이터 시각화를 디자인할 때 정보 또는 메시지를 정확하게 전달할 수 있는 적합한 차트의 유형 선택이 매우 중요하다. 서미란 외(Seo et al., 2018)는 같은 데이터라도 어떤 시각화 차트 유형을 선택하느냐에 따라 시각화의 결과와 해석이 달라질 수 있음을 제기하며, 시각화 차트 유형 선택의 중요성을 강조하였다. 이러한 배경에 따라 본 연구에서는 미국의 프리젠테이션 전문가 아벨라(Abela, 2011)가 사용 목적에 따라 적합한

차트를 선택할 수 있도록 분류한 ‘비교, 구성, 분포, 관계’에 따른 시각화 차트 유형과 강원양 외(Kang et al., 2020)가 제시한 ‘5가지(비교, 추이·트렌드, 구성 비중·분포, 관계, 위치 데이터) 데이터 시각화 유형 분류 기준’을 수정·보완하여 ‘6가지(비교, 추이·트렌드, 구성 비중·분포, 관계, 위치 데이터, 기타)’ 데이터 시각화 유형 분류 기준을 [Figure 3]으로 재구성하였다.

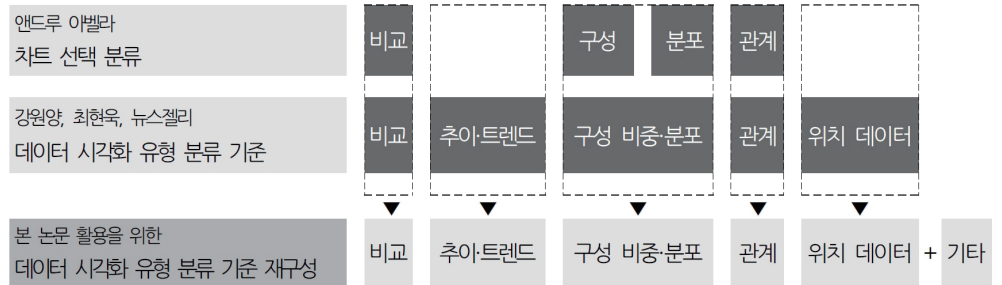


Figure 3 Reorganization of the Criteria for Data Visualization Classification

강원양 외(Kang et al., 2020)에 의하면 시각화 유형에 대한 정확한 이해는 상황 및 목적에 맞는 데이터 시각화 차트 유형을 선택하는 데 도움을 줄 수 있다고 하였다. 이러한 연구 배경을 바탕으로 세부 차트들을 구분하였으며, 6가지 데이터 시각화 유형 분류 기준의 세부 내용은 다음과 같다.

① 비교 시각화 차트

비교를 위한 기본적인 시각화 차트 유형은 막대 차트, 그룹 막대 차트가 있으며 길이, 위치 등의 속성으로 비교한다. 다음의 [Table 1]은 이장미(Lee, 2020), 강원양 외(Kang et al., 2020), 후지 토시쿠니, 와타나베 료이치(Fuji & Watanabe, 2020), 나가타 유카리(Nagata, 2021)의 선행 연구 및 문헌에 정의된 ‘비교 시각화’의 대표적인 차트 유형들의 특징을 분석한 표이다.

Table 1 Characteristics of Various ‘Comparative Visualization Chart’ according to Prior Research

비교 시각화 차트 (18)	비교 시각화 차트의 특징
막대	데이터 크기에 따라 막대 길이를 달리함
그룹 막대	막대 차트를 그룹별로 그림
누적 막대	그룹 막대 차트에서 세부 항목별 막대를 하나로 쌓은 것
양방향 막대	2개의 세부 항목으로 구성된 그룹별 막대 차트이나, 세부 항목 막대 방향이 반대되게 그림
피라미드	양방향 막대 차트와 형태는 동일하나 인구수, 연령별, 성별 데이터를 시각화할 때 활용
양방향 누적 막대	양방향 막대 차트의 막대를 세부 항목별로 나누어 표현한 차트
범위	데이터의 최소값과 최대값 사이를 막대 길이로 표현
점	원의 위치를 달리하여 데이터를 표현
막대 사탕	막대 끝에 원을 표시하여 데이터를 표현
연결된 점	덤벨 차트라고도 하며 2개의 수치 데이터를 원으로 표시, 그 사이를 선으로 연결한 차트
콤보	막대 차트와 선 차트를 동시에 표현
픽토그램	도형, 아이콘을 나열하여 수치형 데이터 크기를 시각화
비례도형도	범주형 변수를 도형으로 그리고 도형 크기로 수치형 변수 데이터의 크기를 시각화
버블	비례도형도인데 원을 도형으로 활용한 차트
워드 클라우드	텍스트 형태로 범주형 변수 값을 시각화, 수치형 변수의 크기를 기준으로 데이터 크기 비교
XY 히트 맵	2개의 범주형 변수의 항목을 x축과 y축에 배치하여 격자무늬 형태로 시각화
레이더	다수의 평가 항목에 대한 수치를 한 번에 볼 수 있는 시각화
플라	막대 차트를 원으로 말아 놓은 형태로 막대를 파이 차트처럼 표현하여 한눈에 비교 할 수 있는 차트

비교를 위한 차트 유형은 총 18개의 세부 차트로 분류되며, 이러한 ‘비교 시각화 차트 유형’은 흔히 지표를 속성별로 나누어 많고 적음, 좋고 나쁨과 같은 판단을 할 수 있는 통찰력을 제공한다.

② 추이·트렌드 시각화 차트

‘추이·트렌드 시각화’는 시간 흐름에 따라 변화하는 데이터의 경향을 보는 데 유용한 차트로서 선 차트, 영역 차트, 타임라인 차트가 대표적이며, 방향의 시각적 단서가 중요한 정보가 된다(이장미, 2020). 다음의 [Table 2]는 이장미(Lee, 2020), 강원양 외(Kang et al., 2020), 후지 토시쿠니, 와타나베 료이치(Fuji & Watanabe, 2020), 나가타 유카리(Nagata, 2021)의 선행 연구 및 문헌에 정의된 ‘추이·트렌드’의 대표적인 차트 유형들의 특징을 분석한 표이다.

Table 2 Characteristics of Various ‘Trend Visualization Charts’ according to Prior Research

추이·트렌드 시각화 차트 (12)	추이·트렌드 시각화 차트의 특징
선	연결선의 높낮이 변화를 통해 데이터 변화 경향을 파악할 수 있도록 시각화
영역	선 차트 하단 영역을 동일 색 혹은 동일 계열 색으로 채운 형태
누적 영역	시간 흐름에 따른 항목별 누적 데이터 변화를 면적 크기로 시각화
팬	시계열 기준 데이터의 예측 값까지 시각화할 때 활용
범프	시간에 따라 변하는 항목별 데이터의 순위 변동을 표현
경사	두 시점의 데이터 추이를 비교할 때 활용(직관적 표현)
추가	시간의 흐름에 따라 가격 변동 정도를 시각화할 때 활용
폭포	월별 현금 추이, 분기별 예산 변동 등 시계열 데이터의 변화와 총합을 한눈에 볼 때 활용
타임라인	특정 이벤트에 대한 기간을 시각화할 때 활용
연결된 산점도	산점도에 표현된 2개의 수치형 데이터를 시간 순서에 따라 선으로 연결하여 시각화
방사형 선	선 차트를 원의 형태로 시각화
캘린더	‘주’와 ‘요일’ 2개의 범주형 항목을 나열해 색이나 진하기로 데이터 값을 표현하는 시각화

추이·트렌드를 위한 차트 유형은 총 12개의 세부 차트로 분류되며, 이러한 ‘추이·트렌드 시각화’ 차트 유형은 시간이 경과함에 따라 일정한 간격을 기준으로 측정되는 데이터를 평가하고 해석할 수 있는 차트이다(서미란 외, 2018).

③ 구성 비중·분포 시각화 차트

데이터가 어떤 항목으로 이루어져 있는지 전체 데이터를 구성하는 세부 항목별 구성 비중과 분포를 표현하기 위해 대표적으로 파이 차트, 도넛 차트, 트리맵 등이 활용된다. [Table 3]은 이장미(Lee, 2020), 강원양 외(Kang et al., 2020), 후지 토시쿠니, 와타나베 료이치(Fuji & Watanabe, 2020), 나가타 유카리(Nagata, 2021)의 선행 연구 및 문헌에 의해 정의된 ‘구성 비중·분포 시각화’의 대표적인 차트 유형들의 특징을 분석한 표이다.

Table 3 Characteristics of Various ‘Composition · Distribution Charts’ according to Prior Research

구성 비중·분포 시각화 차트 (12)	구성 비중·분포 시각화 차트의 특징
파이	전체 데이터의 원에서 차지하는 구성 비중만큼 조각별 면적 크기를 나눔
도넛	전체 데이터의 도넛 형태에서 구성 비중만큼 조각별 면적 크기를 나눔
100% 누적 막대	막대 하나가 전체 데이터 100%를 의미하며, 구성 비중만큼 조각별 면적 크기를 나눔
게이지	자동차 계기판의 형태이며 도형을 나눈 조각의 크기를 기준으로 전체 중 구성비를 파악
와플	전체 데이터를 격자무늬로 표현한 뒤 해당 항목 데이터 비율을 색으로 채워 비중 표현
트리맵	하나의 사각형을 전체 데이터로 보고 데이터의 크기에 따라 네모 조각으로 나누어 표현
계층형 트리맵	여러 개의 트리맵을 합쳐서 하나의 큰 사각형으로 표현
서클 패킹	계층형 구조의 데이터를 원으로 표현, 원의 크기로 특정 항목이 전체 데이터에서 차지하는 비중을 표현
선버스트	원 안쪽은 가장 큰 범주 계층, 항목에 따라 비중 조각을 나눔 뒤 하위 계층 항목은 원 밖에 조각으로 표현

히스토그램	통계적 분포를 표현하는 차트로 가로축에 구간, 세로축에 정도를 표현
점 나열	점을 찍어 데이터 값을 표현하며 점의 간격, 점의 밀집도, 위치 등을 통해 데이터를 해석하는 차트
상자 수염 그림	여러 항목 사이의 범위나 분포를 간단하게 표현하고 최솟값, 최댓값, 사분위수의 정보를 포함

구성 비중·분포를 위한 차트 유형은 총 12개의 세부 차트로 분류되며, 이러한 ‘구성 비중·분포 시각화’ 차트 유형은 서로의 상대적인 값을 표현하는 차트로 전체 대비, 그 비율의 정도를 효과적으로 나타낼 수 있다.

④ 관계 시각화 차트

‘관계 시각화’ 차트는 데이터 간의 관계를 확인하는 데 유용한 차트로 대개 둘 이상의 데이터를 하나의 시각화로 표현해서 데이터 간의 관계를 파악할 수 있게 한다. 가장 기본적인 산점도를 비롯해 버블차트, 네트워크 시각화가 포함된다. 다음의 [Table 4]는 이장미(Lee, 2020), 강원양 외(Kang et al., 2020), 후지 토시쿠니, 와타나베 료이치(Fuji & Watanabe, 2020), 나가타 유카리(Nagata, 2021)의 선행 연구 및 문헌에 의해 정의된 ‘관계 시각화’의 대표적인 차트 유형들의 특징을 분석한 표이다.

Table 4 Characteristics of Various ‘Relationship Visualization Charts’ according to Prior Research

관계 시각화 차트 (7)	관계 시각화 차트의 특징
산점도	차트 영역에 분포된 형태를 바탕으로 데이터 간의 관계를 파악하는 차트
버블차트	3개의 변수를 다루며 1개 변수의 데이터 크기를 원의 크기로 표현
평행 좌표	동일한 범주형 수치 항목의 데이터끼리 선으로 연결하는 차트
생키 다이어그램	각 항목별 데이터 이동 경로를 곡선으로 표현, 선의 경로와 두께를 기준으로 데이터를 해석하는 차트
패럴렐 셋	같은 계열 데이터를 영역으로 연결, 영역 너비로 데이터를 해석하는 차트
코드 다이어그램	호의 길이로 데이터 비중을 표현해서 하나의 원을 만들고 데이터 간의 관계를 곡선으로 연결하는 차트
네트워크 시각화	관계형 데이터를 나타내는 데 활용되는 시각화 차트로 선을 이어 데이터 관계를 표현하는 차트

관계 시각화를 위한 차트 유형은 총 7개의 세부 차트로 분류되며, 이러한 ‘관계 시각화 차트’ 유형은 차트에 산포되어 있는 데이터를 통해 서로의 상관관계, 군집을 찾아내고 그로 인해 인사이트를 발견할 수 있다.

⑤ 위치 데이터 시각화 차트

‘위치 데이터 시각화’ 차트는 데이터가 위치 정보를 포함하고 있을 때 차트 배경을 지도로 활용하는 지도 시각화 유형이다. 데이터 크기를 지역마다 색으로 표현한 단계 구분도, 도형의 크기로 데이터를 표현하는 도형 표현도가 있다. 다음의 [Table 5]는 이장미(Lee, 2020), 강원양 외(Kang et al., 2020), 후지 토시쿠니, 와타나베 료이치(Fuji & Watanabe, 2020), 나가타 유카리(Nagata, 2021)의 선행 연구 및 문헌에 의해 정의된 ‘위치 데이터 시각화’의 대표적인 차트 유형들의 특징을 분석한 표이다.

Table 5 Characteristics of Various ‘Location Data Visualization Charts’ according to Prior Research

위치 데이터 시각화 차트(11)	위치 데이터 시각화 차트 특징
점 밀집도	모든 데이터의 위치를 지도상에 점으로 나타내며, 각 지역 점의 분포를 통해 데이터를 해석하는 차트
도형표현도	각 지역별 위치에 표현할 데이터를 도형으로 나타내고, 도형의 크기로 데이터를 표현
단계 구분도	색을 통해 수치형 데이터의 크기를 지도상에 표현하며 색의 패턴으로 데이터를 해석하는 차트
히트 맵	지도 위 지점마다 색으로 데이터를 표현하고 데이터의 밀집도를 나타내는데 효과적인 차트
등고선 지도	동일 수치의 데이터 값을 갖는 지점을 선으로 연결하여 영역을 구분하고 색을 통해 표현하는 차트
연결 지도	지도 위 특정 지점 사이를 직선 또는 곡선으로 그려 데이터를 표현하는 시각화 차트
이동 경로 지도	여러 지점 간의 관계를 직선 또는 곡선을 그려 표현
흐름 지도	지도 위 특정 지역 사이를 직선 또는 곡선으로 연결하여 표현하며 데이터에 따라 선의 두께를 달리함
카토그램	특정한 데이터 수치 값에 따라 지역별 면적을 왜곡해서 표현하는 시각화 지도 차트
돌리 카토그램	카토그램 변형 형태로 데이터 수치 값에 따라 지역의 영역을 원의 크기로 표현하는 시각화 차트
타일 격자 지도	모든 행정구역상의 지역 크기를 정사각형 타일 형태로 고정하고, 대신 데이터 수치 값에 따라 색을 달리 표현하는 시각화 차트

‘위치 데이터 시각화’를 위한 차트 유형은 총 11개의 세부 차트로 분류되며, 이러한 위치 데이터 시각화 유형은 지도와의 결합을 통해 위치 데이터를 쉽게 이해하고 활용할 수 있다. 이외에 학생의 개인적인 아이디어에 의해 표현된 차트, 여러 개념이 혼재되어 있는 차트, 의미 파악이 어려운 차트는 기타 시각화로 분류하였다.

이상과 같이 차트 유형은 비교 시각화 차트 18개, 추이·트렌드 차트 12개, 구성 비중·분포 시각화 차트 12개, 관계 시각화 차트 7개, 위치 데이터 시각화 차트 11개, 기타 포함 총 61개로 정리할 수 있었다. 이처럼 데이터를 시각화하는 차트의 유형은 매우 다양하다. 나가타 유카리(Nagata, 2021)에 의하면 데이터 시각화를 위한 ‘레퍼토리’ 즉 다양한 시각화에 대한 이해에 관해서는 목적에 적합한 데이터 차트 유형이 무엇인지에 대한 선택의 지침을 세울 수 있다고 보았다. 이를 토대로 다음 장에서는 청소년의 데이터 시각화 표현 유형을 분류해보고자 한다.

3. 청소년의 데이터 시각화 표현 유형 메타 분석

3.1. 분석 대상

본 연구에서는 청소년의 데이터 시각화 표현 유형 메타 분석을 위해 분석 대상의 범위를 다음과 같이 한정한다. 일반 청소년에 비해 데이터 시각화에 대한 관심을 가지고 적극적으로 활용하고 있다고 생각되는 통계청 주관 ‘전국학생통계활용대회’의 중학생부 수상작 2019년 39작품, 2020년 39작품, 2021년 39작품 총 117점을 수집하였다. 한 수상작당 최소 1개에서부터 최대 32개의 차트가 활용되었으며, 각각의 시각화 표현 차트 유형 결과에 대한 분석을 실시하였다.

전국학생통계활용대회는 통계적 지식 증진을 목표로 1998년 초등학교를 대상으로 시작되었으나 현재 초·중·고등학교를 대상으로 그 범위가 확대되었으며, 시험 형식으로 진행되던 경진대회는 2012년부터 통계포스터 경진 대회로 변경되어 통계포스터 작성을 통해 문제해결능력과 통계적 사고력 증진을 목표로 한다. 전국학생통계활용대회의 수상작은 학생들이 직접 작성한 통계포스터로 주제 정하기에서부터 자료수집 및 분석의 과정을 거쳐 이를 시각화하여 결론을 정리하는 일련의 과정을 통해 만들어지는 데이터 시각화의 결과물이다. 전국학생통계활용대회의 통계포스터 작성 방법에 따르면 자료의 분석과 시각화 단계에서 컴퓨터를 활용하여 분석하고 출력하는 방법과 손으로 직접 그리거나 만드는 방법을 제시하고 있으며, 틀 사용에 대한 특별한 제시는 없었다.

중학생으로 그 대상을 선정한 이유는 초등학생의 경우 데이터 내용의 다양성을 확보하기 어려운 점이 있고, 고등학생 경우는 대회 참여 대상으로 포함될 지 오래되지 않아 데이터 빈도수가 작기 때문이다. 무엇보다 중학생은 대회 초기부터 참여하여 타 학교 급에 비해 데이터 주제 다양성이 확보되었기에 연구 분석 대상으로 선정하였다. 단, 본 연구에서 분석한 수상작은 학생들의 실명이 공개된 자료이므로 실제 학생 작품을 예시로 사용하기에는 연구 윤리에 어긋나는 부분이 있어 본 연구자가 그 사례를 재구성하여 활용하였다.

3.2. 분석 방법 및 도구

메타분석은 다양한 연구의 결과물을 분석하는 연구 방법이다(신태범 & 권상희, 2013). 본 연구에서는 분석 도구로 아벨라(Abela, 2011)의 목적에 따른 차트 선택 방법과 강원양 외(Kang et al., 2020)가 제시한 ‘5가지(비교, 추이·트렌드, 구성 비중·분포, 관계, 위치 데이터) 데이터 시각화 유형 분류 기준’을 수정·보완하여 ‘6가지(비교, 추이·트렌드, 구성 비중·분포, 관계, 위치 데이터, 기타) 데이터 시각화 유형 분류 기준’으로 재구성하고 각 유형별로 세부 차트를 나누어 [Table 6]과 같이 정리하였다.

Table 6 The 6 Criteria to Classify Data Visualization Types

비교 시각화	추이·트렌드 시각화	구성 비중·분포 시각화	관계 시각화	위치 시각화	기타
막대 그룹 막대 누적 막대 양방향 막대 피라미드 양방향 누적 막대 범위 점 막대 사탕 연결된 점 콤보 픽토그램 비례도형도 버블 워드 클라우드 XY 히트 맵 레이더 플라	선 영역 누적 영역 팬 범프 경사 추가 폭포 타임라인 연결된 산점도 방사형 선 캘린더	파이 도넛 100% 누적 막대 게이지 와플 트리맵 계층형 트리맵 서클 패킹 선버스트 히스토그램 점 나열 상자 수염 그림	산점도 버블차트 평행 좌표 생키 다이어그램 패럴렐 셋 코드 다이어그램 네트워크 시각화	점 밀집도 도형 표현도 단계 구분도 히트 맵 등고선 지도 연결 지도 이동 경로 지도 흐름 지도 카토그램 돌링 카토그램 타일 격자 지도	학생의 개인적 아이디어에 의해 표현된 차트 여러 개념이 혼재되어 있는 차트 의미 파악이 어려운 차트 등

합계(개), 비율(%)

첫째, 비교 시각화 차트 유형으로 세부 차트는 18개, 둘째, 추이·트렌드 시각화 유형으로 세부 차트는 12개이다. 셋째, 구성 비중·분포 시각화 차트 유형으로 세부 차트는 12개로 나뉘며 넷째, 관계 시각화 유형의 세부 차트는 7개이다. 또한 위치 데이터 시각화 유형의 세부 차트는 11개로 분류되며 마지막으로 기타 시각화는 학생의 개인적 아이디어에 의해 표현된 차트, 여러 개념이 혼재되어 있는 차트, 의미 파악이 어려운 차트 등을 포함한다.

3. 3. 분석 결과

전국학생통계활용대회 중학생부 수상작 117점에서 ‘6가지(비교, 추이·트렌드, 구성 비중·분포, 관계, 위치 데이터, 기타)’ 분류 기준을 사용하여 1,387개의 차트를 분석하였고, 그 결과를 [Table 7]로 정리하였다.

Table 7 Analysis Results Using the 6 Classification Criteria

차트 유형	연도			합계(개)	비율(%)
	'19	'20	'21		
비교 시각화	131	221	199	551	39.73
추이·트렌드 시각화	35	24	38	97	6.99
구성 비중·분포 시각화	176	175	255	606	43.69
관계 시각화	3	0	5	8	0.58
위치 데이터 시각화	0	0	0	0	0.00
기타 시각화	34	74	17	125	9.01
합계	379	494	514	1,387	100.00

최근 3년간의 전국학생통계활용대회 중학생부 수상작에서 ‘구성 비중·분포’의 시각화 유형이 43.69%로 활용 비율이 가장 높게 나타났다. ‘비교 시각화’가 39.73%로 그 뒤를 이었으며, 다음 순으로는 ‘추이·트렌드’ 시각화가 6.99%로 나타났다. 반면 ‘관계’ 시각화는 0.58%로 현저히 낮은 비율로 나타났으며, ‘위치’ 시각화(0.00%)는 전혀 활용되지 않았음을 알 수 있다. 이를 종합해보면, 우리나라 중학생의 데이터 시각화 차트 유형의 활용은 ‘구성 비중·분포’의 시각화 차트 유형과 ‘비교’ 시각화 차트에만 지나치게 편중되어 있음을 확인할 수 있었다.

다음 [Table 8]은 전국학생통계활용대회 중학생부 수상작에 나타난 통계포스터를 6가지 차트 유형 중 구체적으로 어떤 차트로 시각화하였는지 세부 분석한 내용이다.

Table 8 Detailed Analysis Results according to 6 Data Visualization Chart Type

차트 유형(개)	차트	년도			합계 (개)	비율 (%)
		'19	'20	'21		
비교 시각화 (18)	막대	65	142	113	320	58.08
	그룹 막대	23	39	37	99	17.97
	누적 막대	2	18	10	30	5.44
	양방향 막대	3	1	1	5	0.91
	피라미드	-	-	-	-	-
	양방향 누적 막대	-	-	-	-	-
	범위	-	-	-	-	-
	점	-	-	2	2	0.36
	막대 사탕	-	-	-	-	-
	연결된 점	-	-	-	-	-
	콤보	3	4	7	14	2.54
	픽토그램	4	7	13	24	4.36
	비례도형도	-	-	-	-	-
	버블	2	2	6	10	1.81
	워드 클라우드	1	-	-	1	0.18
	XY 히트 맵	10	-	-	10	1.81
	레이더	18	8	8	34	6.17
	플라	-	-	2	2	0.36
비교 시각화 합계		131	221	199	551	100.00
추이·트렌드 시각화 (12)	선	29	16	27	72	74.23
	영역	4	-	7	11	11.34
	누적 영역	-	5	-	5	5.15
	팬	-	-	-	-	-
	범프	-	-	-	-	-
	경사	2	3	4	9	9.28
	추가	-	-	-	-	-
	폭포	-	-	-	-	-
	타임라인	-	-	-	-	-
	연결된 산점도	-	-	-	-	-
	방사형 선	-	-	-	-	-
	캘린더	-	-	-	-	-
추이·트렌드 시각화 합계		35	24	38	97	100.00
구성 비중·분포 시각화 (12)	파이	101	82	112	295	48.68
	도넛	28	49	77	154	25.41
	100% 누적 막대	26	28	48	102	16.83
	게이지	6	8	17	31	5.12
	외플	-	-	-	-	-
	트리맵	3	6	-	9	1.49
	계층형 트리맵	-	-	-	-	-
	서클 패킹	-	-	-	-	-
	선버스트	-	-	-	-	-
	히스토그램	11	2	1	14	2.31
	점 나열	-	-	-	-	-
	상자 수염 그림	1	0	0	1	0.17
구성 비중·분포 시각화 합계		176	175	255	606	100.00

관계 시각화 (7)	산점도	3	-	5	8	100.00
	버블차트	-	-	-	-	-
	평행 좌표	-	-	-	-	-
	생키 다이어그램	-	-	-	-	-
	패럴렐 셋	-	-	-	-	-
	코드 다이어그램	-	-	-	-	-
	네트워크 시각화	-	-	-	-	-
	관계 시각화 합계	3	0	5	8	100.00
위치 데이터 시각화 (11)	점 밀집도	-	-	-	-	-
	도형 표현도	-	-	-	-	-
	단계 구분도	-	-	-	-	-
	히트 맵	-	-	-	-	-
	등고선 지도	-	-	-	-	-
	연결 지도	-	-	-	-	-
	이동 경로 지도	-	-	-	-	-
	흐름 지도	-	-	-	-	-
	카토그램	-	-	-	-	-
	돌링 카토그램	-	-	-	-	-
	타일 격자 지도	-	-	-	-	-
위치 데이터 시각화 합계	-	-	-	-	-	
기타	학생의 개인적 아이디어에 의해 표현된 차트, 여러 개념 혼재되어 있는 차트, 의미 파악이 어려운 차트 등	34	74	17	125	100.00
	기타 합계	34	74	17	125	100.00
전체 시각화 차트		379	494	514	1,387	100.00

데이터 시각화 차트 유형에 있어 활용된 차트의 세부 분석 내용은 다음과 같다. ‘비교’ 시각화 차트에서는 ‘막대 차트’가 58.08%의 비율로 가장 높게 나타났고, ‘추이·트렌드’ 시각화에서는 ‘선 차트’가 74.23%로 현저히 높은 비율로 나타났다. ‘구성 비중·분포’ 시각화에서는 ‘파이 차트’가 48.68%의 비율로 가장 높게 나타났으며, ‘도넛 차트’ 또한 25.41%의 높은 비율로 나타났다. 특이점으로는 ‘관계’ 시각화 차트는 ‘산점도’만 활용되어 그 비율이 100.00%로 나타난 반면 ‘위치’ 시각화는 전혀 활용되지 않아 그 비율을 확인할 수 없었다. 전국학생통계활용대회 중학생부 수상작 117점의 분석 결과를 종합해보면 데이터 시각화의 세부적 차트 활용에 있어서 데이터가 가지는 주제 및 특징과 무관하게 ‘막대, 콤보, 파이, 도넛, 산점도’ 등 익숙하거나 전통적인 차트가 주로 활용된 것을 알 수 있었다.

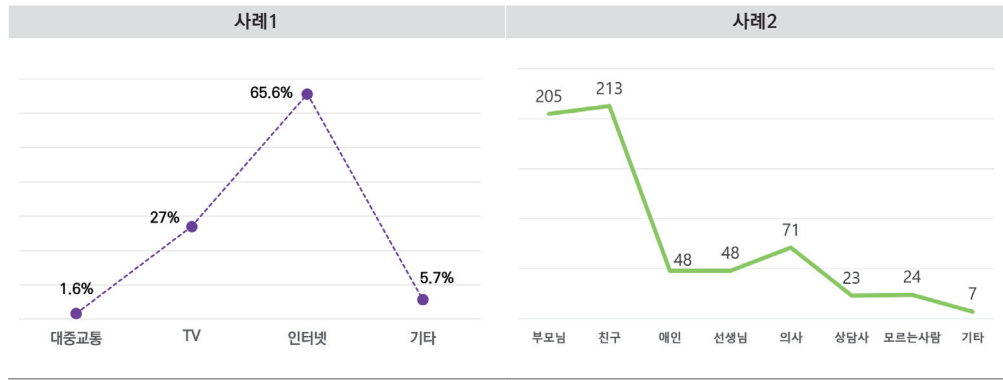
3. 4. 분석 결과에 따른 데이터 시각화의 오류

전국학생통계활용대회 중학생부 수상작의 분석 결과, 활용 빈도가 높은 차트에서 나타난 대표적인 오류 사례를 살펴보고자 한다. 단, 비교 시각화에서 막대를 활용한 차트들은 오류의 비율이 현저히 낮았으므로 데이터 시각화의 오류 사례에서는 제외되었다. 데이터 시각화의 오류 분석의 대상은 추이·트렌드 시각화의 ‘선 차트’, 구성·비중·분포시각화의 ‘파이 차트’, 기타 시각화의 ‘학생의 개인적 아이디어에 의해 표현된 차트’에 대한 오류 사례로 3가지 유형으로 정리할 수 있다.

첫 번째 목적에 맞지 않는 차트 유형을 선택한 ‘차트 선택의 오류’, 두 번째 정보를 포함하지 않은 불필요한 시각적 표현 또는 지나치게 복잡하거나 단순화되어 데이터의 해석에 어려움이 있는 ‘시각적 표현의 오류’, 세 번째 비율에 대한 고려 없이 데이터의 값이 과장 또는 축소되어 표현된 ‘비율 표현의 오류’이다.

[Table 9]는 ‘차트 선택의 오류’ 사례로 <사례 1>, <사례 2> 모두 비교의 목적에 맞지 않는 ‘선 차트’를 선택하여 차트 유형 선택의 오류를 범하였다.

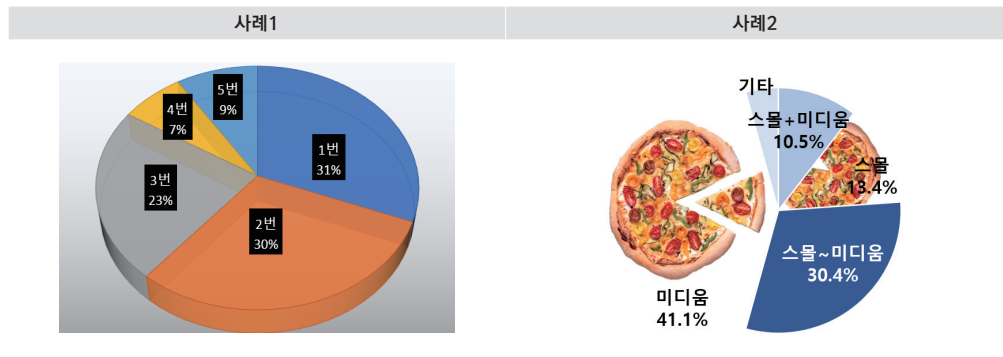
Table 9 An Error of the Chosen Chart



특히 오류가 두드러지게 나타난 사례는 ‘추이·트렌드’ 시각화의 ‘선 차트’이다. ‘선 차트’는 74.23%로 그 활용이 매우 높게 나타났음에도 불구하고 목적에 맞게 활용한 차트의 사례는 극히 드물었다. ‘선 차트’는 시간 흐름에 따라 변화하는 데이터를 표현하는 것이 그 목적이지만 수상작에 나타난 선 차트들은 단순 비교에 초점에 맞추어 표현된 것을 확인할 수 있다.

[Table 10]은 ‘시각적 표현의 오류’ 사례이며 <사례 1>은 ‘구성 비중·분포’ 시각화의 ‘파이 차트’를 3차원으로 표현한 시각화로 ‘엑셀’이나 ‘통그라미’ 시각화 도구를 활용하고 틀 내에서 지원 가능한 차트 중 추천 차트 또는 기본 설정 된 차트 디자인의 형태를 활용한 사례로 볼 수 있으며, 3D로 표현된 데이터의 시각화는 해석의 오류를 야기할 수 있다. <사례 2>는 파이 차트로 파이 조각을 통해 직관적으로 데이터를 이해할 수 있어야 하나 불필요한 이미지를 추가하여 오히려 정보 수용자에게 해석의 어려움을 초래할 수 있다.

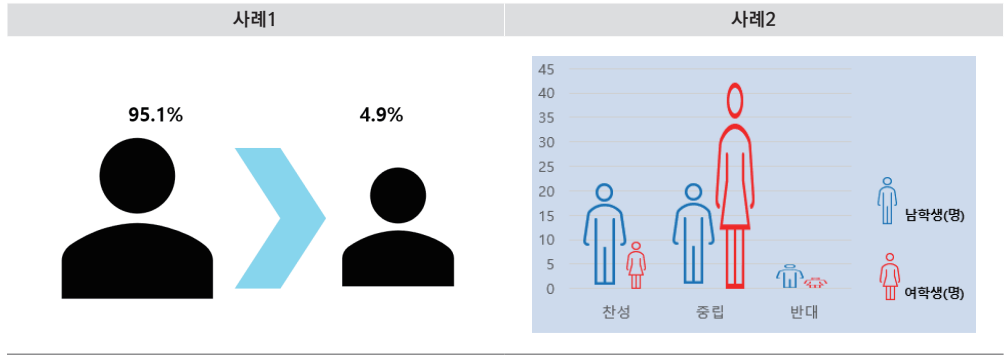
Table 10 An Error of Visual Expression



특히 3D 차트는 많은 시각화 전문가들이 지양해야 한다고 주장하는 시각화 방법임에도 불구하고 수상작들에서는 ‘파이 차트’뿐만 아니라 ‘막대 차트’ 등 3D를 활용한 사례를 다수 발견할 수 있었다. 특히 <사례 1>은 파란색 비율 31%, 주황색 비율 30%로 파란색 부분의 비율이 높으나 3D 효과로 인해 주황색 비율이 높은 것 같은 시각적 오류를 범하고 있다. 파이 차트를 인식하는 가장 큰 시각적 요소는 ‘면적’임에도 불구하고 <사례 2>는 면적에 대한 시각적 요소 대신 이미지를 삽입하고 파이 차트의 형태를 임의적으로 변경하여 해석에 어려움을 준다.

[Table 11]은 비율 표현의 오류로 <사례 1>, <사례 2> 모두 실질적 데이터 값에 근거하지 않고 비율을 표현하여 오류를 범하였다.

Table 11 An Error of a Proportion Expression



대표적인 비율 표현의 오류 사례는 ‘기타 시각화’의 학생의 개인적 아이디어에 의해 표현된 차트에서 많이 나타났다. <사례 1>은 95.1%와 4.9%의 실질적 데이터 값에 대한 정확한 비율로 표현되지 않고 4.9%의 데이터 값이 과장되어 표현되었다. 95.1%와 4.9%의 차이가 (78x58)pixel:(57x62)pixel, 다시 말해 1:19.4 차이가 1:1.28차이로 표현되었음을 알 수 있다. <사례 2>는 xy 좌표 위에 데이터 수치를 면적으로 표현하려 했으나 실질적 데이터의 면적과 관계없이 이미지의 비율이 조절되어 왜곡된 형태로 표현되었다.

이 밖에도 비증의 합에 대한 오류, 범례의 생략, 그리드의 생략, 축의 생략 등 다양한 오류가 발견되었다.

이를 종합해보면 데이터 시각화 차트의 활용 빈도가 높아도 올바른 데이터의 시각화를 표현하는 것에 어려움이 있을 수 있을 수 있다는 것을 발견하였다.

4. 결론 및 제언

본 연구는 최근 정보화 사회에서 매우 중요한 무형 자원으로 활용되고 있는 데이터를 우리나라 청소년들이 적극적으로 활용할 수 있도록 하기 위한 첫 단계로서, 우리나라 청소년들의 데이터 시각화 활용에 대한 동향을 실증적으로 규명하는 데 그 목적을 두었다. 이를 위해 이론적 배경으로 데이터 시각화의 개념과 과정, 유형을 고찰한 뒤, 전국학생통계활용대회의 중학생부 수상작 총 117점의 수상작을 ‘6가지(비교, 추이·트렌드, 구성비중·분포, 관계, 위치 데이터, 기타)’ 분류 기준을 사용해 1,387개의 차트 유형을 분석하여 문제점과 제언을 도출하였다. 이러한 과정을 통한 본 연구의 결론은 다음과 같다.

첫째, 데이터 시각화의 목적에 따른 차트 유형 교육을 필요로 한다.

차트 유형의 선택 오류가 다수 발견되는 것은 통계적 기반에 초점이 맞추어져 있는 현행 교육과 더불어 데이터를 목적에 맞게 표현할 수 있는 디자인적 방법에 대한 학습 부족을 그 원인으로 해석할 수 있다. 서미란 외(Seo et al., 2018)에 의하면 데이터 시각화는 원 데이터를 알맞게 시각화하여 올바른 정보를 전달하는 데 큰 목적이 있기 때문에 시각화 과정에서의 유형 선택을 매우 중요하다고 보았다. 따라서 목적에 맞는 차트 유형을 선택할 수 있도록 데이터 시각화에 대한 기초 지식 교육 및 활용 교육이 필요하다.

둘째, 데이터 시각화의 틀을 사용하기에 앞서 시각 속성의 원칙들에 대한 교육이 선행되어야 할 것이다.

후지 토시쿠니, 와타나베 료이치(Fuji & Watanabe, 2020)에 따르면 데이터 시각화의 주요 시각 속성은 위치, 길이, 방향(각도), 굵기(폭), 크기(면적), 색(채도), 색(색상), 형태 총 8가지로 분류되며, 똑같은 양의 정보라 할지라도 정확도가 더 높은 시각 속성으로 구성된 그래픽을 사용한다면 데이터의 조직력과 패턴 행동이 명확하게 감지될 수 있다(Cairo, 2013).

활용 빈도가 높은 파이 차트와 학생들의 창작품에서 오류의 비율이 높게 나타난 원인은 데이터 시각화를

표현하는 시각 속성의 요소인 면적, 비율 등의 원칙들에 대한 이해의 부족으로 사료된다. 따라서 각 차트에서 중요시되는 시각 속성의 표현 및 차트를 읽어내는 과정에서의 원칙들에 대한 교육을 통해 보다 정확하고 효과적인 시각화를 표현할 수 있을 것이다.

셋째, 데이터 시각화를 응용하고, 인지력을 높일 수 있는 방법들에 대한 기회를 제공해야 할 것이다. 코스린(Kosslyn, 2013)의 색, 채움, 선택적 구성요소 등 차트 디자인에 응용할 수 있는 다양한 교육의 활성화를 통해 올바르게 다양한 차트를 표현할 수 있는 능력을 함양할 수 있기를 기대해 본다.

데이터는 무형 자원이지만 미래사회의 기술과 산업사회를 이끄는 모든 영역의 원 데이터로서 누가, 어떻게 그 잠재적 가치를 발견하고 활용하느냐에 따라 무한한 부가가치를 선점할 수 있다. 따라서 점차 경쟁이 치열해지는 글로벌 경쟁 시대에 데이터 시각화 역량에는 한 개인의 생존뿐 아니라 국가의 생존이 걸려 있다고 해도 과언이 아니다. 본 연구가 데이터 시각화 표현의 함양을 위한 디자인 교육 개발의 기초 자료로 활용되길 기대하며, 후속 과제로 데이터 시각화 교육 방안에 대한 구체적인 디자인 교육 프로그램을 제안하고자 한다.

References

1. Abela, A. (2011). *익스트림 프레젠테이션 [Advanced Presentations by Design]*. Seoul:communicationbooks.
2. Byun, J., & Park, Y. (2016). 사용자 의도 기반 정량적 빅데이터 시각화 가이드라인 툴 [A Guiding System of Visualization for Quantitative Bigdata Based on User]. *KIPS Transactions on Software and Data Engineering*, 5(6), 261-266.
3. Byun, J. (2017). *사용자 의도 기반 정량적 빅데이터 시각화 가이드라인 시스템 [An User Intention Guiding System for a Quantitative Bigdata Visualization]* (Master's thesis). Dankook University, Gyeonggi-do, Korea.
4. Cairo, A. (2013). *The functional art : 인포그래픽과 데이터 시각화 기법을 활용한 스토리텔링 [The Functional Art: An introduction to information graphics and visualization]*. Paju: WIKIBOOKS.
5. Cairo, A. (2019). *The Truthful Art: 진실을 드러내는 데이터 시각화의 과학과 예술 [The Truthful Art: Data, Charts, and Maps for Communication]*. Seoul: insight.
6. Card, S., Mackinlay, J., & Schneiderman, B. (1999). *Readings in Information Visualization: Using Vision to Think*. San Francisco: MO:Morgan Kaufman Publishers.
7. Fuji, T., & Watanabe, R. (2020). *데이터 시각화 입문 [Data Visualiztion]*. Seoul: Roadbook.
8. Healy, K. (2020). *데이터 시각화 기본기 다지기 [Data Visualization: A Practical Introduction]*. Seoul: Acorn.
9. Kang, W., Choi, H., & Newsjelly. (2020). *데이터가 한눈에 보이는 시각화 [Visualization of data at a glance]*. Paju: WIKIBOOKS.
10. Kim, E. (2014). *정보미학공연: 정보시각화의 미학적 확장에 관한 연구 [Information Aesthetic Performance: A Study on the Aesthetic Expansion of Information Visualization]* (Master's thesis). Soongsil University, Seoul, Korea.
11. Kim, H. (2018). *고등학교 미술과 정보디자인 교육으로서 데이터 시각화 수업 실행에 관한 질적 연구 [A Qualitative Study on Implementing Data Visualization Class in High School Art and Information Design Education]* (Doctoral dissertation). Seoul National University, Seoul, Korea.
12. Kim, M. (2021). *초등학생을 위한 데이터 시각화 리터러시 교육 방안 연구 [A Study on the Educational Method of Data Visualization Literacy for Elementary School Students]*. *Korea Institute of Design Research Society*, 6(1), 294-305.
13. Kim, M. (2022). *데이터 시각화를 활용한 창의융합수업 설계원리 개발 -귀추적 사고를 중심으로- [Creative Convergence Class Using Data Visualization-Focusing on abductive thinking-]* (Master's thesis). Seoul National University, Seoul, Korea.
14. Kim, M. B. (2020). *구글 스프레드시트를 활용한 데이터 시각화 교육이 창의성 향상에 미치는 효과 [The effect of education data visualiztion [i.e. visualization] using Google spreadsheet program on improvement of creativity : for forth and fifth grade students]* (Master's thesis). Jeju National University, Jeju, Korea.

15. Kim, M. G. (2020). 파이썬을 활용한 데이터 시각화 교육 교재 개발 및 적용과 그 효과 : 초등학교 6학년을 대상으로 [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, Korea.
16. Kosslyn, S. M. (2013). 눈과 마음을 사로잡는 그래프 디자인 [Graph Design for Eye and Mind]. Seoul:communicationbooks.
17. Kwak, M. (2018). 인터랙티브 데이터 시각화 디자인 사례 연구 : 디지털 뉴스 미디어 블룸버그, 뉴욕타임즈를 중심으로 [A case study on interactive data visualization : Focused on digital news media, Bloomberg and The New York Times] (Doctoral dissertation). Sungkyunkwan University, Seoul, Korea.
18. Kye, B., & Park, Y. (2017). EdTech Trend #3. 교육에서의 데이터 시각화 동향: 사례 및 이슈를 중심으로 [Data Visualization Trends in Education: Focusing on Cases and Issues]. *Korea Education & Research Information Center*.
19. Lee, J. (2020). 데이터 시각화를 위한 대시보드 융합 디자인 연구 [A Study on Dashboard Convergence Design for Data Visualization]. *The Korean Society of Science & Art*, 38(5), 423-435.
20. Lee, J., Lee, T., Kang, G., Kim, S., Park, H., Lee, Y., & Sim, S. (2014). 초중등학생 교육용 통계패키지 통그라미 개발 [A Statistics Education Package Tong-Gramy for 5-8 Graders]. *The Korean Journal of Applied Statistics*, 27(3), 487-500.
21. Lee, J., & Yun, J. (2018). 데이터 시각화의 시각적 형태유형과 은유에 관한 연구 [Metaphor and Visual Form in Data Visualization]. *Proceedings of HCI Korea, 2018*(1), 992-996.
22. Lee, J. Y. (2021). 파이썬을 활용한 데이터 시각화 교육이 초등정보영재 학생의 컴퓨팅 사고력 증진에 미치는 영향 [Effects of Data Visualization Education Using Python on Improvement of Computational Thinking Ability in Information Gifted Students of Elementary school] (Master's thesis). Korea National University of Education. Chung-Buk, Korea.
23. Mackinlay, J., Hanrahan, P., & Stolte, C. (2007). Show me : Automatic presentation for visual analysis. *IEEE TransactionsonVisualization and ComputerGraphics*, 13(6), 1137-1144.
24. Nagata, Y. (2021). 데이터 시각화 디자인 [Data Visualization Design]. Paju: WIKIBOOKS.
25. Yau, N. (2013). 데이터 포인트 : 데이터 시각화 + 인포그래픽을 깨우치다 [Data points: Visualization that means something]. Seoul: bjpublic.
26. Seo, M., Kim, H., Choi, E., Choi, Y., & Suh, J. (2018). 데이터 시각화 유형 및 툴의 기능 분석 [Analysis of Data visualization types and tools]. *Proceedings of the Korea Information Processing Society Conference*, 25(2), 449-452.
27. Shin, T. B., & Kweon, S. H. (2013). 국내 온라인저널리즘 연구에 대한 메타분석: 분석 대상과 주제, 이론 및 방법론적 특성을 중심으로 [A Meta Analysis for Online Journalism Researches: Focused on the characteristic of analysis object, theme, theory and methodology]. *Korean Society for Journalism & Communication Studies*, 9(2), 200-249.

청소년의 데이터 시각화 표현 유형 메타 분석: 2019~2021 '전국학생통계활용대회' 3년간의 중학생부 수상작 분석을 중심으로

백민영¹, 현은령^{2*}

¹한양대학교 대학원 러닝사이언스학과 인지정보디자인전공, 박사과정 학생, 서울, 대한민국

²한양대학교 사범대학 응용미술교육과&대학원 러닝사이언스학과 인지정보디자인전공, 부교수, 서울, 대한민국

초록

연구배경 본 연구는 우리나라 청소년들의 데이터 시각화 차트 활용 사례 실증 분석을 통해, 그 특성과 문제점을 도출하고자 한다. 이를 통해 청소년 데이터 시각화 디자인 교육의 필요성과 중요성을 확인하고자 한다.

연구방법 이론적 연구에 대해서는 데이터 시각화에 관련된 선행 연구 자료를 바탕으로 데이터 시각화의 '개념, 과정, 차트 유형'을 고찰하였다. 실증 분석은 최근 3년(2019~2021)간 통계청에서 주관한 '전국학생통계활용대회'의 중학생 부분 수상작 117점에서 나타난 1,387개의 차트를 '6가지(비교, 추이·트렌드, 구성 비중·분포, 관계, 위치 데이터, 기타)' 분류 기준을 사용하여 분석하고 그 특징과 문제점을 도출하는 과정으로 이루어졌다.

연구결과 데이터 특성과 수집 목적에 상관없이 '구성 비중·분포'와 '비교' 시각화 차트에 지나치게 편중되어 있음을 확인할 수 있었다. 또한 목적에 맞지 않는 차트 유형을 선택한 '차트 선택의 오류', 정보가 포함되지 않은 불필요한 시각적 표현이 많아 해석이 어려운 '시각 표현의 오류', 데이터 값이 과장 또는 축소되어 표현된 '비율 표현의 오류'가 많이 발견되었다.

결론 청소년들의 올바른 데이터 시각화 표현을 위해서는 데이터가 가지는 내용과 목적을 이해하고, 이에 적합한 시각 이미지로 표현하는 경험이 필요하다. 데이터는 미래사회의 중요한 자원으로 이미 각 분야에서 주목 받고 있다. 따라서 우리나라 청소년들의 데이터 리터러시 함양을 위한 다양한 데이터 시각화 교육 프로그램에 디자인 연구자들도 관심을 가져야 할 때이다.

주제어 데이터 시각화, 데이터 시각화 디자인, 데이터 시각화 교육
