



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

박사학위논문

컴퓨팅 사고력과 창의성 향상을 위한
데이터 과학 교육 프로그램

제주대학교 대학원

과학교육학부 컴퓨터교육전공

김용민

2018년 8월



컴퓨팅 사고력과 창의성 향상을 위한 데이터 과학 교육 프로그램






指導教授 金 鍾 勳

金 瑢 珉

이 論文을 教育學 博士學位 論文으로 提出함

2018年 6月

金瑢珉의 教育學 博士學位 論文을 認准함

審査委員長	김 성 백	
委 員	박 찬 정	
委 員	이 재 호	
委 員	전 흥 식	
委 員	김 종 훈	

濟州大學校 大學院

2018年 6月

Data Science Education Program to Improve Computational Thinking and Creativity

Yongmin Kim

(Supervised by professor Jonghoon Kim)

A thesis submitted in partial fulfillment of the requirement for the
degree of Doctor of Philosophy in Education

2018. 6.

This thesis has been examined and approved.

.....
Thesis director, Jonghoon Kim, Prof. Department of Computer Education

.....
.....
.....
.....
.....
(Name and signature)

.....
Date

Major in Computer Education
Faculty of Science Education
GRADUATE SCHOOL
JEJU NATIONAL UNIVERSITY

목 차

표 목 차	iv
그림목차	vii
국문초록	ix
I. 서론	1
1. 연구의 필요성	1
2. 연구의 내용	3
3. 연구의 기대 효과	4
4. 연구의 제한점	5
II. 이론적 배경	6
1. 컴퓨팅 사고력	6
1) 컴퓨팅 사고력의 정의	6
2) 컴퓨팅 사고력의 구성 요인	8
3) 컴퓨팅 사고력의 학습단계와 주요 개념 요소	12
4) 컴퓨팅 사고력 검사 도구	13
2. 창의성	15
1) 창의성의 정의	15
2) 창의성의 구성 요인	16
3) 창의성 검사 도구	17
3. 데이터 과학 교육	17
1) 데이터 과학 교육의 정의	18
2) 데이터 과학 교육의 동향	22
4. 프로그램 개발 모형	23
1) ADDIE 모형	23
2) Dick & Carey 모형	25

III. 데이터 과학 교육 프로그램	27
1. 데이터 과학 교육 프로그램을 위한 요구 분석	27
1) 요구 분석 목적 결정	28
2) 출처 확인	28
3) 도구 선택	29
4) 요구 분석 실시	29
5) 요구 분석 의사 결정	33
2. 데이터 과학 교육 프로그램의 설계	34
1) 성취 목표 진술	34
3. 데이터 과학 교육 프로그램의 개발	45
1) 교육 프로그램 개발 원리	45
2) 교수 매체 선정 및 개발	46
4. 데이터 과학 교육 프로그램의 적용	48
1) 엑셀을 활용한 초등학생 대상 데이터 과학 교육 프로그램	49
2) 앱인벤터를 활용한 초등학생 대상 데이터 과학 교육 프로그램	49
3) 스크래치를 활용한 대학생 및 예비 코딩 강사 대상 데이터 과학 교육 프로그램	50
5. 데이터 과학 교육 프로그램의 평가	51
1) 양적 연구 결과	51
2) 질적 연구 결과	52
IV. 데이터 과학 교육 프로그램 개발의 실제 및 적용	54
1. 엑셀을 활용한 초등학생 대상 데이터 과학 교육 프로그램	54
1) 요구 분석	54
2) 설계 및 개발	54
3) 적용	57
4) 양적 연구 결과 분석	63
5) 질적 연구 결과 분석	74
2. 앱인벤터를 활용한 초등학생 대상 데이터 과학 교육 프로그램	76
1) 요구 분석	76
2) 설계 및 개발	76

3) 적용	79
4) 양적 연구 결과 분석	86
5) 질적 연구 결과 분석	90
3. 스크래치를 활용한 대학생 및 예비 코딩 강사 대상 데이터 과학 교육 프로그램 ...	92
1) 요구 분석	92
2) 설계 및 개발	92
3) 적용	95
4) 양적 연구 결과 분석	102
5) 질적 연구 결과 분석	105
V. 결론 및 제언	107
참고문헌	111
Abstract	120
부록	124

표 목 차

<표 II-1> 컴퓨팅 사고력 관련 국내·외 연구	7
<표 II-2> 컴퓨팅 사고력 관련 연구에서의 구성 요인 분석	9
<표 II-3> 사고력의 7범주	10
<표 II-4> 컴퓨팅 사고력의 학습단계와 주요 개념 요소	13
<표 II-5> 컴퓨팅 사고력 측정 도구 관련 연구	14
<표 II-6> 창의성 관련 연구들의 창의성 구성 요인	16
<표 II-7> 창의성 측정 검사 도구 관련 연구	17
<표 II-8> 데이터 과학 관련 국내·외 관련 연구	21
<표 II-9> ADDIE 모형의 과정과 산출물	24
<표 II-10> ADDIE 모형과 Dick & Carey 모형 비교	26
<표 III-1> 정보원과 추출 가능한 정보	28
<표 III-2> 초등학생	29
<표 III-3> 교사-컴퓨터 관련 소지 학위	29
<표 III-4> 소프트웨어교육 참여 경험	30
<표 III-5> 교육용 프로그래밍 언어에 대한 관심도	31
<표 III-6> 언플러그드활동에 대한 관심도	31
<표 III-7> 피지컬컴퓨팅에 대한 관심도	31
<표 III-8> 데이터 과학 교육을 통한 능력 향상	32
<표 III-9> 데이터 과학 교육 활용 데이터	32
<표 III-10> 데이터 과학 교육 학습 방법	33
<표 III-11> 스크래치 관련 연구	36
<표 III-12> 앱인벤터 관련 연구	39
<표 III-13> 엑셀 관련 연구	40
<표 III-14> 성취 목표	44
<표 III-15> 교수 전략 설계	45
<표 III-16> 교육 프로그램 주제(엑셀)	47

<표 III-17> 교육 프로그램 주제(앱인벤터)	48
<표 III-18> 교육 프로그램 주제(스크래치)	48
<표 IV-1> 연구 대상	55
<표 IV-2> 프로그램 일정	56
<표 IV-3> 교육 프로그램 주제(엑셀)	56
<표 IV-4> 실험 설계	57
<표 IV-5> 컴퓨팅 사고력 검사 정규성 검정	63
<표 IV-6> 컴퓨팅 사고력 사전 검사 결과 비교(독립표본 t 검정)	64
<표 IV-7> 컴퓨팅 사고력 사후 검사 결과 비교(독립표본 t 검정)	65
<표 IV-8> 사전·사후 검사 결과 집단 내 비교(대응표본 t 검정)	65
<표 IV-9> 개인 데이터반 창의성 검사 정규성 검정	66
<표 IV-10> 공공 데이터반 창의성 검사 정규성 검정	67
<표 IV-11> 창의성 사전 검사 결과(독립표본 t 검정)	68
<표 IV-12> 창의성 사전 검사 결과(Mann-Whitney U 검정)	68
<표 IV-13> 창의성 사후 검사 결과(독립표본 t 검정)	69
<표 IV-14> 개인 데이터반 사전·사후 검사 결과 집단 내 비교(대응표본 t 검정) ..	70
<표 IV-15> 개인 데이터반 사전·사후 검사 결과 집단 내 비교(Wilcoxon 부호 순위 검정) ..	71
<표 IV-16> 공공 데이터반 사전·사후 검사 결과 집단 내 비교(대응표본 t 검정) ..	72
<표 IV-17> 공공 데이터반 사전·사후 검사 결과 집단 내 비교(Wilcoxon 부호 순위 검정) ..	72
<표 IV-18> 연구 대상	77
<표 IV-19> 프로그램 일정	78
<표 IV-20> 교육 프로그램 주제(앱인벤터)	78
<표 IV-21> 실험 설계	79
<표 IV-22> 컴퓨팅 사고력 검사 정규성 검정	86
<표 IV-23> 사전·사후 검사 결과 집단 내 비교(대응표본 t 검정)	86
<표 IV-24> 앱인벤터반 창의성 검사 정규성 검정	87
<표 IV-25> 앱인벤터반 사전·사후 검사 결과 집단 내 비교(대응표본 t 검정) ..	88
<표 IV-26> 앱인벤터반 사전·사후 검사 결과 집단 내 비교(Wilcoxon 부호 순위 검정) ..	88
<표 IV-27> 연구 대상	93

<표 IV-28> 프로그램 일정	94
<표 IV-29> 교육 프로그램 주제(스크래치)	94
<표 IV-30> 실험 설계	95
<표 IV-31> 대학생 및 예비 코딩 강사 창의성 검사 정규성 검정	103
<표 IV-32> 사전·사후 검사 결과 집단 내 비교(대응표본 t 검정)	104
<표 IV-33> 사전·사후 검사 결과 집단 내 비교(Wilcoxon부호 순위 검정) ..	104

그림 목 차

[그림 II-1] Conway 데이터 과학 밴다이어그램	19
[그림 II-2] 데이터 과학의 학제적 연계	20
[그림 II-3] 시간 흐름에 따른 관심도 변화	22
[그림 II-4] 지역별 관심도	22
[그림 II-5] ADDIE 모형의 요소 및 개념화	23
[그림 II-6] Dick & Carey 모형의 요소 및 개념화	25
[그림 III-1] Rossett 모형의 절차	27
[그림 III-2] 데이터 크기에 따른 문제 특성별 사용 도구	34
[그림 III-3] 데이터 과학 설문 결과: 가장 널리 사용되는 데이터 과학의 도구들 ..	41
[그림 III-4] 데이터 과학 교육 문제 해결 단계	43
[그림 III-5] 데이터 과학 교육 보고서 주요 내용	53
[그림 IV-1] 주제 및 주제 선정 이유(동기) 및 내용(이○○)	58
[그림 IV-2] 데이터 문제 정의하기(이○○)	58
[그림 IV-3] 데이터 수집하기(이○○)	60
[그림 IV-4] 현상 이해하기: 탐색적 데이터 분석(이○○)	61
[그림 IV-5] 현상 일반화하기(이○○)	62
[그림 IV-6] 현상 예측하기(이○○)	62
[그림 IV-7] 데이터 스토리텔링 장면	63
[그림 IV-8] 속성 간의 관계 분석(상관관계 분석) 후 현상 일반화하기(김○○)	75
[그림 IV-9] 주제 선정 이유(동기) 및 내용(석○○)	80
[그림 IV-10] 데이터 문제 정의하기(석○○)	80
[그림 IV-11] 데이터 수집하기(석○○)	81
[그림 IV-12] 현상 이해하기: 탐색적 데이터 분석(석○○)	84
[그림 IV-13] 현상 일반화하기(석○○)	85
[그림 IV-14] 데이터 스토리텔링 장면	85
[그림 IV-15] 속성 간의 관계 분석(상관관계 분석) 후 현상 일반화하기(김○○)	91

[그림 IV-16] 주제와 주제 선정 이유(동기) 및 내용(양○○)	96
[그림 IV-17] 데이터 문제 정의하기(양○○)	96
[그림 IV-18] 데이터 수집하기(양○○)	97
[그림 IV-19] 현상 이해하기: 탐색적 데이터 분석(양○○)	101
[그림 IV-20] 현상 일반화하기(양○○)	102

<국문초록>

컴퓨팅 사고력과 창의성 향상을 위한 데이터 과학 교육 프로그램

김 용 민

제주대학교 대학원 과학교육학부 컴퓨터교육전공

지도교수 김 중 훈

이 연구의 목적은 컴퓨팅 사고력과 창의성 향상을 위한 데이터 과학 교육 프로그램을 개발하고 적용하여 교육적 효과를 검증하는 것이다.

18세기 이후 인류가 경험한 3차례의 산업혁명과 달리, ‘4차 산업혁명’의 물결은 알파고의 충격을 시작으로 기존 삶의 방식을 근본적으로 바꿔놓으며, 빅데이터로 대표되는 측정조차 불가능하고 나날이 급격하게 증가하는 많은 양의 디지털 데이터의 홍수 속에서 빠르게 우리 생활 속에 파고들고 있다.

따라서, ‘4차 산업혁명’ 시대의 교육은 데이터를 수집, 분석, 처리하여 데이터에 근거한 결정을 하고 비판적으로 평가할 수 있는 창의적인 인재를 양성하기 위한 데이터 과학 교육과 ‘컴퓨터’를 기반으로 다양한 문제에 대하여 해결방법을 찾기 위해 자료의 수집, 분석 및 효율적 문제 해결 과정 등에 중점을 두는 소프트웨어교육을 통해 컴퓨팅 사고력과 창의성을 기르도록 해야 한다. 이에 본 연구에서는 학생들의 컴퓨팅 사고력과 창의성을 향상시키기 위한 데이터 과학 교육 프로그램을 개발하였다.

본 연구에서는 데이터 과학 교육 프로그램의 개발을 위해 ADDIE 모형과 Dick & Carey 모형의 단계를 적용하여 연구를 진행하였다.

요구 분석 단계에서는 컴퓨팅 사고력, 창의성, 데이터 과학 교육에 대한 문헌

연구 및 선행연구 분석과 Rossett의 요구 분석 모형에 따라 초등학생, 대학생, 교사를 대상으로 학습자의 요구 분석을 진행하였다.

설계 단계에서는 Dick & Carey 모형의 단계에 따라 성취 목표 진술, 교수 전략을 개발하였다.

개발 단계에서는 설계 단계의 내용을 바탕으로 Dick & Carey 모형의 단계에 따라 ‘엑셀을 활용한 초등학생 대상 데이터 과학 교육 프로그램(개인 데이터)’, ‘엑셀을 활용한 초등학생 대상 데이터 과학 교육 프로그램(공공 데이터)’, ‘앱인벤터를 활용한 초등학생 대상 데이터 과학 교육 프로그램’, ‘스크래치를 활용한 대학생 및 예비 코딩 강사 대상 데이터 과학 교육 프로그램’ 등 4가지의 데이터 과학 교육 프로그램에 대한 교수 매체 선정 및 개발을 하였다.

적용 단계에서는 데이터 과학 교육 도구 및 교육 대상에 따라 개발된 4가지의 데이터 과학 교육 프로그램을 총 5회에 걸쳐 초등학생, 대학생, 예비 코딩 강사에게 적용하여 프로그램의 교육적 효과를 검증하였다.

평가 단계에서는 컴퓨팅 사고력과 창의성을 측정하기 위하여 컴퓨팅 사고력 검사와 창의성 검사를 실시한 후 검사 결과를 분석하였다.

평가 결과, 본 연구에서 개발한 ‘엑셀을 활용한 초등학생 대상 데이터 과학 교육 프로그램(개인 데이터)’, ‘엑셀을 활용한 초등학생 대상 데이터 과학 교육 프로그램(공공 데이터)’, ‘앱인벤터를 활용한 초등학생 대상 데이터 과학 교육 프로그램’은 모두 초등학생의 컴퓨팅 사고력과 창의성 향상에 효과적인 것으로 나타났다. 또한, 본 연구에서 개발한 ‘스크래치를 활용한 대학생 및 예비 코딩 강사 대상 데이터 과학 교육 프로그램’은 대학생과 예비 코딩 강사의 창의성 향상에 효과적인 것으로 나타났다.

본 연구의 시사점은 다음과 같다.

첫째, 데이터 과학 교육 프로그램은 컴퓨팅 사고력과 창의성 향상에 효과적이다. 데이터 문제 정의하기(문제 탐색하기), 데이터 수집하기, 현상 이해하기(탐색적 데이터 분석), 현상 일반화하기, 현상 예측하기, 데이터 스토리텔링 등 직접적인 활동 중심의 데이터 과학 교육 문제 해결 과정을 통해 컴퓨팅 사고력과 창의성이 향상되었다.

둘째, 개인 데이터와 공공 데이터를 모두 데이터 과학 교육에 활용할 수 있

음을 입증하였다.

셋째, 대표적인 통계용 프로그램인 ‘엑셀’ 뿐만 아니라 ‘엠펙터’, ‘스크래치’와 같은 교육용 프로그래밍 언어를 데이터 과학 교육 프로그램의 도구로써 활용할 수 있음을 제시하였다.

넷째, 데이터 과학 교육 프로그램은 초등학생, 대학생, 예비 코딩 강사를 대상으로 그 효과가 입증되어, 앞으로 교육현장에서 다양한 연령층을 대상으로 활용할 수 있게 되었다.

향후 후속 연구를 통해, 본 연구에서 개발한 데이터 과학 교육 프로그램을 지속적으로 개선하고, 컴퓨팅 사고력 향상 검증을 위한 타당도와 신뢰도를 갖춘 검사 도구가 개발된다면, 데이터 과학 교육을 통한 컴퓨팅 사고력 향상 및 연구 결과에 대한 각 요인들 간의 상관관계를 체계적으로 분석할 수 있을 것이다.

주요어: 컴퓨팅 사고력, 창의성, 데이터 과학 교육, ADDIE 모형, 요구 분석

I. 서론

1. 연구의 필요성

“4차 산업혁명이 우리에게 쓰나미처럼 밀려올 것입니다. 그것이 모든 시스템을 바꿀 것입니다.” 2016년 WEF(World Economic Forum)에서 Claus Schwab 회장이 언급한 내용이다. ‘4차 산업혁명’은 정보통신기술(ICT)이 제조업 등 다양한 산업들과 결합하며 지금까지는 볼 수 없던 새로운 형태의 제품과 서비스, 비즈니스를 창출하는 것을 말한다(김성범 & 강성현, 2016). ‘4차 산업혁명’의 물결은 18세기 이후 인류가 경험한 3차례의 산업혁명과 달리, 알파고의 충격을 시작으로 기존 삶의 방식을 근본적으로 변화시키며 빠르게 우리 생활 속으로 파고들고 있다. 따라서, 인공지능, 사물인터넷(IoT), 자율주행자동차, 3D프린터, 나노기술 등으로 대표되는 ‘4차 산업혁명’의 핵심기술을 어떻게 교육과 융합시켜 나갈 수 있을지에 대한 고민이 무엇보다 필요한 시점이다(김용민 & 김중훈, 2017a).

또한, 2012년 디지털 데이터의 양이 아날로그 데이터의 양을 앞지르는 것을 시작으로 이제는 측정조차 불가능하고 나날이 급격하게 증가하는 많은 양의 디지털 데이터의 홍수 속에서 살아가게 되었다(이명호, 2016). 이와 같이 디지털 정보의 양이 방대해지고 그 불확실성이 점차 증대되는 시대에 교육에서는 필요한 디지털 데이터를 수집·분석하여 유용한 정보를 추출할 수 있는 능력을 길러주어야 할 필요가 있다(이은희 & 김원경, 2015).

2016년 10월, Claus Schwab 회장은 국내 한 방송에서 ‘4차 산업혁명이 미래다’라는 주제로 강의하며 우리나라 교육이 가야 할 길을 제시하기도 했는데, 그것은 바로 미래 인재를 양성해야 한다는 것이었다. 이는 미래 사회를 이끌어 갈 인재 양성을 위해 세계 각국에서 실시하고 있는 소프트웨어교육과 빅데이터의 시대에 데이터를 수집, 분석, 처리하여 데이터에 근거한 결정을 하고 비

관적으로 평가할 수 있는 지적인 인재를 양성하기 위한 데이터 과학(data science) 교육(강현영, 2012)에서 그 힌트를 찾을 수 있다(김용민 & 김종훈, 2017a).

소프트웨어교육이란, 한글, 엑셀 등 컴퓨터를 단순히 활용하는 방법을 익히는 교육이 아니라 다양한 문제의 해결방법을 찾기 위해 ‘컴퓨터’를 기반으로 자료를 수집, 분석하고, 문제의 효율적 해결 과정 등을 창조하는 일련의 사고력 교육이다(김유리 & 문용은, 2015).

데이터 과학 교육은 수학, 통계학, 경영학, 컴퓨터과학, 문헌정보학 등 다양한 학문의 융합이라고 정의되고 있으며(Shi et al, 2014), 2015 개정 수학과 교육과정의 개정 방향은 ‘실생활 중심의 통계 내용 재구성’, ‘공학적 도구의 활용 강조’ 등을 통해 통계적 문제해결 과정을 강조하고 있다(배혜진 & 이동환, 2016).

ASA(American Statistical Association)의 GAISE(Guidelines for Assessment and Instruction in Statistics Education) 프로젝트(Franklin et al, 2007)에서는 문제 설정, 자료 수집, 자료 분석, 결과 해석의 과정으로 통계적 문제해결 과정을 경험하도록 교육과정을 구성하였다(배혜진 & 이동환, 2016).

우리나라와 미국의 예를 통해 살펴보면 데이터 과학 교육과 소프트웨어교육은 컴퓨터와 같은 공학적 도구를 활용하여 문제를 수집·분석하여 효율적으로 문제를 해결해 나가는 절차와 과정을 강조하고 있다는 공통점이 있다. 따라서, 소프트웨어교육에서 다양한 문제를 해결하기 위해 문제의 특성을 분석하고, 문제의 각 요소들 간의 논리적 상관관계를 파악하며, 해결방법을 설계하는 과정(김유리 외, 2015)을 통해 컴퓨팅 사고력을 향상 시킬 수 있다는 국내외의 여러 연구 결과들처럼 데이터 과학 교육을 통해서도 컴퓨팅 사고력과 창의성을 향상시킬 수 있을 것으로(김용민 & 김종훈, 2017a) 생각된다.

컴퓨팅 사고력(computational thinking)은 Seymour Paper가 1996년 최초로 사용한 용어이며, 2006년 Wing에 의해 처음 제시된 이후, 지난 몇 년간 교육의 여러 분야에서 광범위하게 논의되어 오고 있으며, 프로그래머나 컴퓨터과학자 뿐 아니라 21세기의 모든 사람이 필수적으로 갖추어야 할 새로운 역량으로 간주되고 있다(이은경, 2013). 창의성 또한 컴퓨팅 사고력과 함께 소프트웨어교육을 통해 향상시키고자 하는 인간의 가장 높은 수준의 수행과 성취이다. 따라

서, 데이터 과학 교육을 통한 컴퓨팅 사고력과 창의성 향상은 ‘4차 산업혁명’으로 인한 빅데이터 시대의 미래 인재 양성 교육에 대한 궁극적인 해답이라고 할 수 있다(김용민 & 김종훈, 2017b).

요약하면, ‘4차 산업혁명’으로 인한 빅데이터 시대에 필요한 미래 인재 양성 교육은 컴퓨팅 사고력과 창의성을 함양하는 것이 무엇보다 중요하며, 이에 본 연구에서는 데이터 과학 교육을 통해 학습자의 컴퓨팅 사고력과 창의성을 기를 수 있는 교육 프로그램을 개발하였다.

2. 연구의 내용

본 연구의 내용을 요약하여 제시하면 다음과 같다.

첫째, 컴퓨팅 사고력과 창의성의 향상을 위한 데이터 과학 교육 프로그램을 개발하기 위해 컴퓨팅 사고력, 창의성, 데이터 과학 교육을 대상으로 한 연구들에 대해 조사 및 특징을 비교, 분석한 후 제시한다.

둘째, 컴퓨팅 사고력과 창의성을 향상시키기 위한 데이터 과학 교육 프로그램을 설계, 개발한다.

셋째, 개발한 교육 프로그램을 학습자에게 투입하고 컴퓨팅 사고력, 창의성 검사 도구를 활용하여 본 연구를 통해 개발한 교육 프로그램의 연구 결과를 도출·분석하여 효과를 검증한다.

ISD(Instructional System Development)모형은 기본적으로 ADDIE 모형을 적용하였는데 ADDIE 모형의 경우 준비단계에 해당하는 Analysis(분석)와 Design(설계)의 단계가 복잡하고 교사가 사용하기에 복잡하다는 단점이 있어(김선태, 2011), ADDIE 모형의 단계를 따르면서 효과적인 수업설계 과정을 안내하는 체제적, 단계별 교수설계 모형중의 하나인 Dick & Carey 모형을 함께 적용하였다. 또한, Analysis(분석)의 학습자 요구 분석은 실제 요구 분석 실행자들이 적용하기 쉬운 안내를 제공하는(이재무, 2014) Rossett의 요구 분석 모형을 사용하였다. ADDIE 모형 및 Dick & Carey 모형의 단계에 따라 연구 내

용 및 연구방법을 구체적으로 살펴보면 다음과 같다.

- 분석(Analysis)
 - 컴퓨팅 사고력, 창의성에 대한 연구(문헌 연구)
 - 선행 연구 분석(사례 연구)
 - 학습자의 요구 분석(조사 연구)
- 설계(Design)
 - 성취 목표 진술(조사 연구/문헌 연구)
 - 평가에 활용될 도구 개발(문헌 연구)
 - 교수에 필요한 전략 개발(조사 연구/문헌 연구)
- 개발(Development)
 - 교수 매체 선정 및 개발(개발 연구)
 - 형성평가 설계 및 수행(조사 연구/문헌 연구)
- 실행(Implementation)
 - 데이터 과학 교육 프로그램 적용(실험 연구)
- 평가(Evaluation)
 - 교육 프로그램 적용 결과 분석(실험 연구/조사 연구)

3. 연구의 기대 효과

본 연구의 기대 효과는 다음과 같다.

첫째, 스스로 자료를 수집, 조직, 표현하고 해석하는 직접적인 활동 중심의 데이터 과학 교육을 통해 컴퓨팅 사고력과 창의성이 향상될 것이다.

둘째, 실생활 문제해결을 위해 수집한 자료를 교육용 프로그래밍 언어를 활용하여 데이터 문제 해결 과정의 경험을 통해, 컴퓨터과학과 통계학이 지니고 있는 한계를 극복하고 시너지 효과를 넘으로써 소프트웨어교육의 또 다른 모델을 제시할 수 있을 것이다.

셋째, 데이터 과학 교육 프로그램은 초등학생, 대학생, 예비 코딩 강사를 대

상으로 개발 및 적용되어, 앞으로 교육현장에서 다양한 연령층을 대상으로 하는 소프트웨어교육 방안 중 하나로 활용될 수 있을 것이다.

4. 연구의 제한점

본 연구에서의 제한점은 다음과 같다

첫째, 본 연구의 교육 프로그램의 평가는 초등학생, 대학생, 예비 코딩 강사를 실험대상으로 선정하여 효과를 분석하였다. 하지만 일부 집단의 구성원 수가 많지 않아 정규성 확보에 어려움이 있어, 중심극한정리에 근거한 모수적 방법과 표본 재추출법을 사용하는 비모수적 방법을 병행하여 실시하였다. 따라서, 본 연구 결과를 모든 학습자에게 일반화하기에는 어려움이 있다.

둘째, 본 연구의 목적은 데이터 과학 교육 프로그램을 통해 학습자의 컴퓨팅 사고력과 창의성을 향상시키는 것이다. 하지만, 아직까지는 컴퓨팅 사고력을 객관적으로 평가해 낼 수 있는 공신력 있는 검사 도구가 없기 때문에 본 연구의 교육 프로그램이 학습자의 컴퓨팅 사고력을 향상시켰는지에 대한 직접적인 검증이 불가하였다.

II. 이론적 배경

1. 컴퓨팅 사고력

컴퓨팅 사고력(computational thinking)은 1980년대에 Seymour Papert가 아동의 절차적 사고 연구를 통해 사고 과정으로서 컴퓨팅을 체계화하였고, 이를 통해 컴퓨팅 기기를 문제 해결의 창조적인 도구로 활용하는 컴퓨팅 사고력이라는 용어에 관심을 갖게 되었다(류미영 & 한동관, 2015).

Wing은 모든 사람이 3R(읽기, 쓰기, 셈하기)과 더불어 모든 학습자가 컴퓨팅 사고력을 배우고 학습해야 한다고 주장하였으며, 추상화(abstraction)와 자동화(automation)를 통한 문제해결능력이라고 하였다(Wing, 2008). 추상화는 문제 해결을 위해 문제를 분해하거나 중요한 부분을 끌어내는 것 등을 통하여 해결해야 할 문제의 복잡성을 효과적으로 해소시켜 나갈 수 있도록 하는 것이며, 자동화는 추상 개념이 수행해야 할 일들을 컴퓨팅 기기가 수행할 수 있도록 해결과정을 알고리즘화하는 과정을 의미한다(Wing, 2008).

1) 컴퓨팅 사고력의 정의

컴퓨팅 사고력은 Wing에 의해 ‘미래 지식 정보화시대를 살아갈 모든 사람이 갖추어야 할 기본적인 사고 능력으로서 컴퓨터 과학의 기본 개념 및 컴퓨터 과학 원리에 따른 문제 해결 방식과 시스템 설계 방법, 인간 행동의 이해까지 포함되는 추상적인 사고 능력’(Wing, 2006)이라고 정의되었다.

CSTA(Computer Science Teachers Association)와 ISTE(International Society for Technology in Education)는 컴퓨팅 사고력을 문제를 해결하는 과정으로 정의하고, 문제해결을 돕기 위해 컴퓨터와 다른 도구를 이용할 수 있도록 문제를 정립하고 문제해결을 위한 자료를 논리적으로 구성하고 분석하며, 모델과 시뮬레이션과 같은 추상화를 통해 자료를 재표현하고 절차적 단계와

알고리즘적 사고를 바탕으로 해결책을 자동화하며, 또한 가장 효율적이고 효과적인 단계와 자원의 조합을 얻기 위해 가능한 해결 방법들을 찾아 분석하고, 구현하며 다양한 문제들의 해결과정을 일반화 및 전환시키는 특성을 지닌다고 하였다(CSTA & ISTE, 2011a).

컴퓨팅 사고력은 컴퓨터과학의 중심개념이라고 할 수 있으며, 미국과 영국을 중심으로 컴퓨팅 사고력의 중요성을 인식하고 그에 대한 교육을 정규 교육과정에 포함시키기 위한 연구들이 지속적으로 수행되고 있으며, 우리나라에서도 컴퓨팅 사고력의 개념에 대한 다양한 정의를 합리적으로 통합하여 제시하려는 노력과 함께 활발한 학문적 논의가 지속되고 있다(함성진 외, 2014).

컴퓨팅 사고력에 대해 현재까지 논의가 이루어져 온 관련 연구를 <표 II-1>에 제시하였다.

<표 II-1> 컴퓨팅 사고력 관련 국내·외 연구

제목	내용
Computational thinking(Wing, 2008)	컴퓨팅 사고력은 컴퓨팅의 기본 개념과 원리에 기반하여 문제 해결, 시스템 설계, 인간의 행동 양식 이해에 대한 하는 접근 방식
CSTA K-12 Computer Science Standards (CSTA & ISTE, 2011a)	컴퓨팅 사고력은 컴퓨터를 활용하여 구현하는 방법을 통해 문제 해결에 접근하는 방법
Report of a Workshop on the Scope and Nature of Computational Thinking (National Research Council, 2010)	<ul style="list-style-type: none"> • Lee, P. - 컴퓨팅 사고력은 인간 지능의 확대를 통해 실제적으로 적용할 수 있는 인간 지능 매커니즘에 대한 연구이고, 다시 말하면 인간의 정신 능력의 복잡도의 관리, 자동적인 일처리 추상화 도구를 통해 확장하는 것임 • Wulf, B. - 과학은 물리적인 대상과 관련되고, 컴퓨팅 사고력은 문제 해결 과정과 진행을 가능하도록 하는 추상적인 현상에 초점을 맞춤 • Sussman, G. - 컴퓨팅 사고력은 일 처리에 대한 정확한 방법을 공식화한다. 다시 말해 특정한 문제의 효율적인 처리를 위해 문제의 분석 및 해결을 위해 엄밀한 절차를 만들어가는 과정이 컴퓨팅 사고력임 • Fox, E. - 인간이 세상에 대한 접근 및 과정을 생각하며, 디지털로 표현된 것을 통해 하는 일들로 표현함

<p>Report of a Workshop on the Pedagogical Aspects of Computational Thinking (National Research Council, 2011)</p>	<p>컴퓨팅 사고력의 교육학적 관점에 대한 논의 • Henderson, P. -Computational thinking as generalized problem solving with constraints.(제약을 가진 일반화된 문제 해결로서의 컴퓨팅 사고)</p>
<p>Computational Thinking 능력 향상을 위한 로봇 프로그래밍 교수 학습 모형(이은경, 2009)</p>	<p>21세기를 살아가는 모든 사람이 갖추어야 할 기본적인 사고 능력으로 컴퓨터 과학의 기본 개념과 원리에 따른 문제 해결, 시스템 설계, 인간 행동의 이해를 포함하는 추상적 사고 능력. 이는 문제 해결을 위해 적합한 추상적 개념을 선택하고 구성하기 위한 추상화 능력과 추상적 개념을 자동화하기 위해 적합한 컴퓨팅 도구를 선택하고 사용하기 위한 능력을 포함</p>
<p>컴퓨터과학 교육용 계산 원리 학습도구의 기능요소 고찰(김형철, 2011)</p>	<p>컴퓨팅 사고는, 좁은 의미로는 계산 시스템(computational system)을 활용해 효과적으로 작업하기 위해 습득해야 할 사고방식이나 태도이며, 넓은 의미로는 세상을 이해하는 양식(단순한 방법을 초월한 양식, 광범위한 인간노력에 두루 접목 가능한 양식)</p>
<p>Algorithmic brick based tangible robot and hybrid programming environments for enhancing computational thinking(권대용, 2011)</p>	<p>컴퓨팅 사고는 21세기 중반까지 모든 사람들이 사용하게 될 필수적 기술(skill)으로써, 컴퓨터 과학의 기본 개념과 원리에 따른 문제 해결, 시스템 설계, 인간 행동의 이해를 포함하며 컴퓨터 과학의 다양한 영역을 반영하는 정신적 도구의 범위를 포함</p>
<p>계산적 사고력 신장을 위한 PPS 기반 프로그래밍 교육 프로그램(김병수, 2014)</p>	<p>실세계에서부터 디지털세계를 망라하여 자연적으로 또는 인간 사회에 의해 존재하는 다양한 현상 속의 계산을 계산 대행자(computational agent)를 이용하여 발견하거나 새롭게 창조하기 위해 습득해야 할 인간의 사고 양식과 태도</p>
<p>초중등 단계 Computational Thinking 도입을 위한 기초 연구(이영준 외, 2014)</p>	<p>컴퓨팅 사고력은 컴퓨팅 시스템의 역량을 활용하여 해결하고자 하는 문제를 효과적이고 효율적으로 해결할 수 있는 절차적 사고 능력</p>
<p>컴퓨팅 사고력 신장을 위한 프로그래밍 중심 STEAM 교육 프로그램(김태훈, 2015)</p>	<p>컴퓨팅 사고력을 구성하는 요소를 인지적인 측면과 창의적인 측면으로 구분하여 논리적 사고력과 창의력 검사 도구로 효과를 검증</p>

2) 컴퓨팅 사고력의 구성 요인

컴퓨팅 사고력의 정의와 구성 요인에 대한 접근은 각 연구마다 다양하다. 컴퓨팅 사고력 관련 국내·외의 연구들을 대상으로 컴퓨팅 사고력의 구성 요인

에 대한 연구들을 정리하면 <표 II-2>와 같고, 창의성과의 관계는 <표 II-3>과 같다. 국내·외 관련 연구들에서 제시한 컴퓨팅 사고력 구성 요인들 중 공통적인 것들은 ‘재귀적 사고, 추상적 사고, 비판적 사고, 논리적 사고, 알고리즘적 사고’ 라고 할 수 있다.

<표 II-2> 컴퓨팅 사고력 관련 연구에서의 구성 요인 분석(김병수, 2014)

구 분	Wing (2006)	유중현 & 김중혜 (2010)	이은경 (2009)	Den- ning (2010)	송정미 (2011)	권대용 (2011)	Selby (2012)	Werner et al (2012)	Walden et al (2013)
재귀적 사고	○	○	○						
추상적 사고	○		○				○	○	
분해적 사고	○						○		
경험적 추론	○								
예방/보호/복구 적사고	○								
비판적 사고		○	○						○
논리적 사고		○	○		○				
동시적 사고			○						
절차적 사고			○						
전략적 사고			○						
선행적 사고			○						
알고리즘적 사고		○		○	○	○	○	○	
혁신적 사고					○				
일반화							○		
과학적 사고					○				
효율적인 솔루션					○				
문제해결력									○
모델링 능력							○	○	

<표 II-3> 사고력의 7범주 (김영정, 2005)

← 수리성 방향	비판적 사고					⇒ 예술성 방향
Formal Symbolic Thinking 기호적 사고	Analytical Thinking 분석적 사고	Inferential Thinking 추론적 사고	Synthetical Thinking 종합적 사고	Alternative Thinking 대안적 사고	Divergent Thinking 발산적 사고	Material Symbolic Thinking 상징적 사고
	개념적 분석 텍스트 분석	분석적 추론: 연역 종합적 추론: 귀납귀추	논리퍼증 의사 결정 상황 추리 민감성	관점/발상전환 대안 창안, 시야/시계확장 시각/지평전환 재정의	유창성 융통성 독창성 진교성	
논리적 사고				창의적 사고		
광의의 논리적 사고					협의의 창의적 사고	
협의의 논리적 사고			광의의 창의적 사고			

(1) 재귀적 사고

인지공학의 관점에서, 재귀적 사고(recursive thinking)란 ‘자기 자신 또는 타인의 사고에 관한 사고’를 의미한다(Kurdek, 1977). 유중현과 김종혜는 ‘재귀적 사고란 인간이 영아기 때 사용하는 원초적인 사고과정에서 시작하여 인간이 커가면서 논리적 사고가 접목되어 논리를 바탕으로 하여 조건이 만들어지고 조건을 변경하면서 문제가 해결될 때까지 반복적으로 사고하는 과정’이라 정의하였다. Wing(2006)은 재귀적 사고란 병렬적 처리를 의미한다고 했는데, 코드를 데이터로, 데이터를 코드로 해석하는 것, 하나의 문제의 장·단점을 인식하는 것, 어떤 이에게 또는 사물에게 하나 이상의 이름을 부여하는 것이 그 예가 될 수 있다고 하였으며, 프로그램의 정확성과 효율성만을 평가하는 것이 아니라 간결함과 단순성을 고려한 심미적 측면에서의 시스템 설계도 재귀적 사고라 하였다(김병수, 2014).

(2) 추상적 사고

Trope과 Liberman(2003)은 ‘추상화의 차원(dimension)이라는 것은 스펙트럼의 구체적 부분에서의 낮은 수준의 지역적 사고에서 추상적 부분에서의 높은

수준의 전체적 사고에 의해 특징지어진 연속체'라고 정의하였다. 예를 들어, 추상적 사고(abstract thinking)는 '숲'에 대한 생각인 반면, 구체적 사고는 '나무'에 관한 생각이라고 제시하였다(김병수, 2014). Smith와 Trope(2006)는 '추상적 사고란 자극을 그보다 높은 차원으로 분류할 수 있을 뿐 아니라, 자극의 주요 측면에 초점을 두고 패턴과 핵심을 추출하기 위한 구조와 패턴을 감지하는 사고'라고 정의하였다.

(3) 비판적 사고

비판적 사고(critical thinking)에 대한 개념과 하위 요소들에 정의는 매우 다양하다(이강범, 2008). 비판적 사고와 관련한 연구 중에서 김영정(2004)의 비판적 사고론은 연구의 성과와 그 영향성 모두 인정을 받고 있다(최훈, 2010). 김영정(2004)은 단순 기억 능력이나 상기(recall) 능력과 같은 저차원적 사고 능력을 제외한 고차적 사고 능력을 7범주로 구별하여 제시하였다

(4) 논리적 사고

논리적 사고(logical thinking)란 대상이나 사건들의 인과적 관계에 의하여 합리적인 결론을 추론하는 능력이다(손은경, 1995). 일반적으로 비판적 사고와 논리적 사고는 동일한 개념을 지칭하고 혼용되어 사용되기도 하지만 논리적 사고는 주로 형식 논리/수리 논리적인 뉘앙스가 강한 반면, 비판적 사고는 비형식 논리학을 기반으로 한 일상 언어/비형식 논리적인 뉘앙스가 강한 용어이다(김병수, 2014). 즉, 논리적 사고는 타당성(validity)을 그 준거로 가지는 사고인 반면, 비판적 사고는 건전성(soundness)을 그 준거로 가지는 사고라고 구분될 수도 있을 것이며 이것은 논리적 사고는 구체적인 상황이나 맥락을 가급적 배제하는 추상화의 길을 지향해온 반면, 비판적 사고는 판단이나 추리에 있어 상황이나 맥락의 중요성을 강조하고 상황이나 맥락을 적극적으로 고려하는 구체화의 길을 지향해온 것과는 통한다(김병수, 2014).

(5) 알고리즘적 사고

알고리즘적 사고(algorithmic thinking)는 어떠한 문제를 해결하기 위한 문제

해결 절차를 만들어내는 능력이다(유중현 & 김종혜, 2010). 유중현과 김종혜(2010)는 문제를 해결하기 위해서는 직관적 사고가 필요하며, 직관적 사고에 의해 나온 해결책들을 통해 결과를 예측해 보는 추론적·논리적 사고가 필요하며 각각의 해결책들의 장단점을 파악하는 비판적 사고가 이미 포함되어 있다고 말하고 있다(유중현 & 김종혜, 2010). 알고리즘적 사고에 의해 도출된 결과가 항상 논리적으로 옳고 최적의 알고리즘이라는 것은 아니며 문제를 한 번에 해결하는 ‘통찰’을 의미하는 것도 아니고 하나의 문제를 해결하는데 필요한 최적의 계산을 논리적 사고의 기반 위에서 하려는 사고 과정이라는 의미이다(김병수, 2014).

3) 컴퓨팅 사고력의 학습단계와 주요 개념 요소

CSTA와 ISTE에서는 Barr와 Harrison, Conery의 연구 결과(Bar, Harrison, & Conery, 2011)를 바탕으로 컴퓨팅 사고력의 핵심 개념과 능력으로 자료수집, 자료분석, 자료표현, 문제 분해, 추상화, 알고리즘과 절차, 자동화, 시뮬레이션, 병렬화 의 9가지 개념으로 구분하여 K-12의 컴퓨터과학 표준 교육과정으로 제시하였다(류미영 & 한선관, 2015).

Philip(2010)은 컴퓨팅 사고력을 추상화, 자동화, 적용의 3단계로 구분하였으며(Guo, P., 2010), CSTA와 ISTE에서 제시한 컴퓨팅 사고력의 주요 개념 요소(CSTA & ISTE, 2011b)와 연계하여 본 연구에 적용되는 컴퓨팅 사고력의 학습단계와 주요 개념요소를 <표 II-4>에 제시하였다(Kim, Y. & Kim, J, 2016).

<표 II-4> 컴퓨팅 사고력의 학습단계와 주요 개념 요소

단계	개념	내용
추상화 (abstraction)	자료수집 (data collection)	문제의 이해와 분석을 토대로 문제를 해결하기 위한 자료 수집
	자료분석 (data analysis)	수집된 자료와 문제에 주어진 자료를 세심히 분류하고 분석
	자료표현 (data representation)	문제의 자료를 단어, 차트, 그래프, 이미지 등으로 표현
	문제 분해 (problem decomposition)	문제를 해결하기 위해 문제를 다루기 쉽도록 작은 부분으로 나눠 분석
	추상화 (abstraction)	문제의 복잡한 정도를 낮추기 위해 주요 개념이나 정의 설정
자동화 (automation)	알고리즘과 절차 (algorithms & procedures)	문제를 해결하기 위한 과정을 순서적 단계로 표현
	자동화 (automation)	순서대로 나열 및 표현한 내용을 컴퓨팅 기기를 통해 최선의 해결책을 선택
적용 (application)	시뮬레이션 (simulation)	복잡하고 어렵거나 현실적으로 불가능한 해결책을 선택하기 위한 모의실험 수행
	병렬화 (parallelization)	문제 해결을 위한 공동 목표 달성 작업 수행

4) 컴퓨팅 사고력 검사 도구

컴퓨팅 사고력과 관련된 많은 연구가 이루어져 왔지만 컴퓨팅 사고력을 측정하기 위한 공식적인 검사 도구는 현재까지 개발되어 있지 않다. 다만, 컴퓨팅 사고력의 측정과 평가에 대해 현재까지 이루어져 온 국내외의 몇몇 연구들에 대해 살펴보면 <표 II-5>와 같다.

<표 II-5> 컴퓨팅 사고력 측정 도구 관련 연구

제목	내용
정보과학적 사고 기반의 문제 해결 능력 향상을 위한 중등 교육 프로그램(김중혜, 2009)	정보과학영재교육원 및 PISA의 ‘문제 해결’ 영역 문항들을 활용하여 검사 도구 개발
Computational Thinking 능력 향상을 위한 로봇 프로그래밍 교수 학습 모형(이은경, 2009)	OECD/PISA의 문제 해결 영역 평가 틀을 이용하여 검사 도구 개발
계산적 사고 향상을 위한 창의적 스크래치 프로그래밍 학습(이은경, 2013)	Brennan과 Resnick(2012)이 제시한 계산적 사고 능력 평가 틀을 이용
First year student performance in a test for Computational Thinking(Gouws et al, 2013)	computer olympiad의 Talent Search Test
계산적 사고력 신장을 위한 PPS 기반 프로그래밍 교육 프로그램(김병수, 2014)	컴퓨팅 사고력 구성요소를 계산적 인지와 계산적 창의성으로 구분. 계산적 인지도 검사 도구 개발
컴퓨팅 사고력 신장을 위한 프로그래밍 중심 STEAM교육 프로그램(김태훈, 2015)	컴퓨팅 사고력 구성요소를 인지적인 측면과 창의적인 측면으로 구분, 논리적 사고력 검사(GALT)와 창의력 검사지(TTCT) 활용

김중혜(2009)의 연구에서는 정보과학영재 교육원과 PISA의 ‘문제 해결’ 영역에 해당하는 문항을 선별하여 검사 도구를 개발하였다. 하지만 선정된 평가 문항들이 컴퓨터 과학의 핵심 주제와 어떠한 관계가 있는지 명확한 근거를 제시하지는 못하였다.

이은경(2009)의 연구에서는 PISA의 문제를 바탕으로 검사 도구를 개발하였지만 검사 도구의 각 영역과 컴퓨팅 사고력의 하위 요인 간에 어떠한 관련을 맺고 있는지 명확한 근거를 제시하지는 못하였다.

김병수(2014)의 연구에서는 컴퓨팅 사고력의 구성요소를 계산적 인지와 계산적 창의성으로 구분하여 컴퓨팅 사고력 검사 도구를 개발하였다. 하지만 컴퓨팅 사고력의 인지적인 측면에 대해서만 검사지를 통해 다루고 있는 제한점이 있다.

김태훈(2015)의 연구에서는 선행 연구를 통하여 인지적인 측면과 창의적인 측

면으로 컴퓨팅 사고력 구성요소를 구분하여 기존에 공신력을 확보한 검사지 중 논리적 사고력과 창의력 검사를 실시하였다. 하지만, 김병수(2014)의 연구에서처럼 인지적인 측면은 비판적 사고, 추상적 사고, 논리적 사고, 알고리즘적 사고, 재귀적 사고 등을 포함하는 측면이 강하기 때문에 학자마다 견해가 다양하기는 하지만 단지, 논리적 사고에만 국한시킬 수는 없다는 제한점이 있다.

지금까지 다양한 논의가 진행되어 왔지만, 아직까지는 컴퓨팅 사고력을 측정하는 데에는 한계가 있는 것이 사실이다. 따라서, 본 연구에서는 앞서 살펴본 선행연구들 중 가장 최근에 개발된 계산적 인지력 검사(Computational Cognition Test) A, B형(김병수, 2014)을 활용하여 컴퓨팅 사고력을 측정하였다. 또한, 선행 연구들 중 창의성을 컴퓨팅 사고력의 구성요인으로 제시한 연구도 있지만 창의성이 컴퓨팅 사고력에 포함된다는 명확한 연구 결과가 없기 때문에 본 연구에서는 컴퓨팅 사고력과 창의성을 구분하여 데이터 과학 교육 프로그램의 효과를 검증하였다.

2. 창의성

‘창의성(creativity)’이라는 용어는 다양한 의미를 가지고 있으며 심리학에서 논쟁적으로 토론이 이루어져 왔다. 그 동안의 창의성에 대한 연구 대부분은 창의성의 정의, 평가, 증진에 대한 다양한 관점에 초점을 맞추어 왔다. 또한, 창의적 인간, 창의적 과정, 환경 요소의 영향, 창의적 생산물이 무엇인가에 대한 탐구가 관련 연구에서 계속되고 있다(Runco, 2001).

1) 창의성의 정의

창의성의 개념은 어렵고 복잡하며 다면적인 성격을 띠고 있어서 그 개념을 한 마디로 정의내리기가 어렵다. 그것은 창의성이 인간의 가장 높은 수준의 수행과 성취이기 때문이다(Barnes, 2002). 창의성의 개념은 연구자의 수만큼이나 다양하고 포괄적이며 연구자와 연구 분야에 따라 창의성의 정의에 대한 견해

차이가 있다(박정호 & 김철, 2011).

Guilford(1967)는 창의성을 지적인 능력으로 보고, ‘아이디어의 생성’이라고 정의했으며, 창의성을 기존의 지능에서 강조하는 수렴적 사고와는 달리 확산적인 사고를 필요로 하는 사고과정이라고 하였고, Torrance(1982)는 ‘창의적인 사고능력이란 창의적인 성취를 할 때 작용한다고 생각하는 일반화된 정신능력의 집합’이라고 정의하였으며, 창의적 사고를 구성하는 하위요인으로는 유창성, 융통성, 독창성, 정교성을 들고 있다(김은경 외, 2010).

2) 창의성의 구성 요인

창의성을 분석적으로 접근하기 위해, 이러한 창의성의 구성 요인들에 대한 관련 연구를 조사하여 인지적 영역에서의 각 학자마다 주장하는 창의성의 구성 요인에 대해 <표 II-6>에 정리하였다(김병수, 2014).

<표 II-6> 창의성 관련 연구들의 창의성 구성 요인

관련연구 구성요인	Guilford (1956)	Torrance (1974)	Gold (1981)	Feldhusen (1983)	허경철 외 (1991)	전경원 (2000) K-CCTYC
유창성	○	○	○	○	○	○
독창성	○	○	○	○	○	○
융통성	○	○	○	○	○	○
정교성	○	○		○	○	
민감성		○			○	
조직성						
재정의 및 재구성력		○				
성격적 요인						
지각적 개방성						
상상력						○

3) 창의성 검사 도구

<표 II-7>은 최근 컴퓨터과학 분야에서 창의성과 관련된 논문들에서 사용되는 창의성 측정 검사 도구를 정리한 것인데(김병수, 2014), 창의성 검사 도구로 Torrance의 창의성 측정을 위한 TTCT(Torrance Tests of Creative Thinking) 검사지가 많이 활용되고 있음을 알 수 있다.

<표 II-7> 창의성 측정 검사 도구 관련 연구

제목	내용
마이크로 로봇 교육을 통한 초등학교 창의성 개발에 대한 연구(이태욱, 2006)	TTCT (도형) 검사
LOGO프로그래밍 언어가 초등학생의 창의성 발달에 미치는 영향(이점순, 2008)	TTCT (도형) 검사
두리틀을 이용한 프로그래밍 수업이 창의성, 문제해결력, 프로그래밍 흥미도 향상에 미치는 영향(이민희, 2009)	TTCT (도형) 검사
초등정보영재의 창의성 신장을 위한 교과 통합 로봇 프로그래밍 수업 모형(서영민 & 이영준, 2010)	강충열(2001)의 ‘창의적 성향 검사’와 자체개발한 ‘창의적 인지 능력 검사’
초등학생의 창의성 신장을 위한 스크래치 프로그래밍 교재 개발 연구(김성훈, 2010)	TTCT (도형) 검사
EPL과 로봇 프로그램 교육의 창의성 신장 효과 분석(박경재, 2010)	이경화(2003)의 ‘객관형 초등 창의성 검사’
EPL을 이용한 창의성 증진 교육 프로그램 개발 및 적용에 관한 연구-로고와 스크래치를 중심으로(김종진, 2011)	TTCT (도형) 검사

3. 데이터 과학 교육

데이터를 향한 사람들의 인식이 변화하기 시작했다. 과거에는 데이터란 수집된 자료에 불과했으며, 더 효율적으로 저장하고 더 빨리 열람하는 것이 데이터

를 대하는 주된 관심사였다. 그러나 최근, 데이터는 과학적 의사결정을 위한 필수 재료로 인식되고 있다. 데이터를 활용해 의미 있는 정보를 추출하고, 숨어있던 규칙들을 찾아낸다. 다양한 수학적 모델링 기법을 적용하여 과거의 데이터로부터 미래를 예측하기도 한다(김동일, 2016).

데이터 과학이 경영, 학술, 교육 등의 전 영역에 확산되면서 변화를 이끌어내고 있다. 빅데이터가 대중적인 키워드로서 데이터의 중요성을 상징하는 용어로 대표되고 있다면, 학술적으로 데이터의 중요성을 부각시키는 현상은 데이터 과학의 대두이다(이재윤, 2015).

2009년, Google의 수석 경제학자 Hal Varian은 “10년 뒤 가장 섹시한 직업은 data scientist 일 것이다”라고 말했다. 즉, 정보에 근거한 결정을 하고 비판적으로 평가할 수 있는 지적인 사람으로 학생들을 지도하고자 한다면 통계적 지식은 필수적이라 할 수 있다. 따라서 일상생활에서 만나는 많은 통계적 정보와 주장들을 판단, 평가하는 데에 통계적 추론과 통계적 소양, 경험 등이 필요하다(강현영, 2012). NCTM(National Council of Teachers of Mathematics)은 모든 학생들이 교양을 갖춘 시민과 지혜로운 소비자가 되기 위해서는 자료를 수집하여 조직하고, 표나 그래프로 나타내어 분석하고, 자료에 근거한 추론과 예측을 하고, 평가를 할 수 있도록 일상적인 경험과 연결되는 학습 기회를 제공해야 한다고 하였다(이은희 & 김원경, 2015).

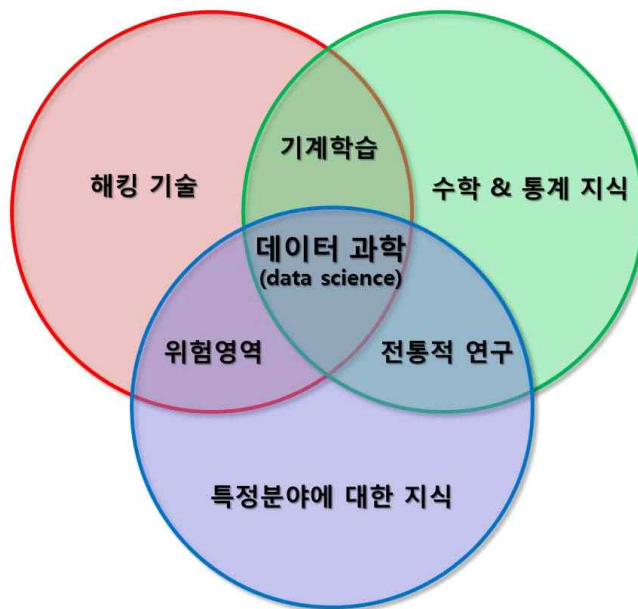
1) 데이터 과학 교육의 정의

데이터 과학이라는 용어는 2001년도에 Cleveland에 의해서 처음 사용이 되었으며(Cleveland, 2001), NCDS(National Consortium for Data Science)는 데이터 과학을 디지털 데이터에 대한 과학적 관찰, 이론 개발, 시스템적 분석, 가설 실험, 검증을 하는 분야라고 정의하고 있다(Jagadish, 2015). Cleveland(2001)의 연구에서는 컴퓨터과학과 통계학이 지니고 있는 각각의 한계를 극복하고 서로 시너지 효과를 낼 수 있도록 하자는 의미에서 데이터 과학을 제안했다(이명호, 2016).

WEF(World Economic Forum)는 2012년 떠오르는 10대 기술 중 하나로 빅데이터 기술을 선정하였으며 국내에서도 지식경제부에서 2012년도에 10대 핵

심기술 하나로 빅데이터를 선정하였다(최윤식, 2012). 최근에 만들어진 많은 양의 다양한 데이터는 수집되고 분석이 되어져야 할 필요성을 인식하지만 기존의 방법으로는 데이터의 처리가 어려워지고 있다(Shi et al, 2014). 이러한 데이터를 수집, 분석, 처리하는 학문 분야 중 하나가 데이터 과학이며, 이는 수학, 통계학, 경영학, 컴퓨터과학, 문헌정보학 등 다양한 학문의 융합이라고 정의되고 있다(Shi et al, 2014).

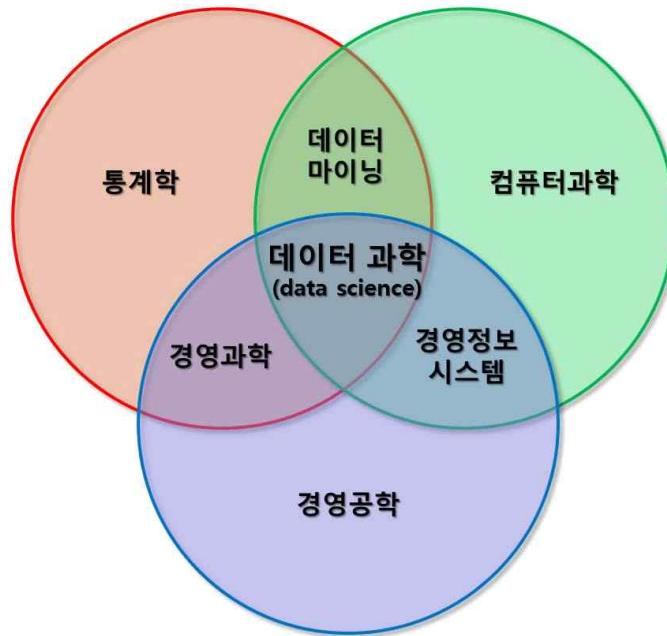
통계학, 컴퓨터학, 그리고 경영공학에 기반을 둔 데이터 과학은 데이터마이닝, 경영과학, 경영정보학을 바탕으로 데이터 구조(structure) 및 수집(collection)에서부터 분석(analysis)과 저장(storage)에 이르는 전체 과정에 대하여 다룬다. 데이터 과학은 바이오정보학(bioinformatics), 문헌정보학(library and information science), 기술경영학(management of technology) 등 다양한 분야에서 발생하는 데이터의 관리에 중요한 역할을 담당하고 있다(Stanton, 2013).



[그림 II-1] Conway 데이터 과학 밴다이어그램

데이터 과학의 정의에 대한 연구는 Conway(2010)의 밴다이어그램이 많이 알려져 있다. Conway는 [그림 II-1]과 같이 각 영역의 교집합 부분을 데이터

과학이라고 정의하고 있는데, 각 영역은 수학 및 통계 지식, 해킹 기술, 특정분야에 대한 지식이다. 세 영역이 겹치는 부분을 기계학습(machine learning), 전통적 연구(traditional research), 그리고 위험영역(danger zone)이라고 추가적인 정의를 하고 있다(이명호, 2016).



[그림 II-2] 데이터 과학의 학제적 연계

또한, 데이터 과학은 통계학, 컴퓨터과학, 그리고 경영공학 등 다양한 학문분야가 학제적(interdisciplinary) 연계를 통하여 이루어진다. [그림 II-2]는 이와 같은 데이터 과학의 학제적 연계에 대하여 나타내고 있다(전성해 & 최준혁, 2015).

데이터 과학에 대해 현재까지 논의가 이루어져 온 국내·외의 관련 연구 내용을 <표 II-8>에 제시하였다.

<표 II-8> 데이터 과학 관련 국내·외 관련 연구

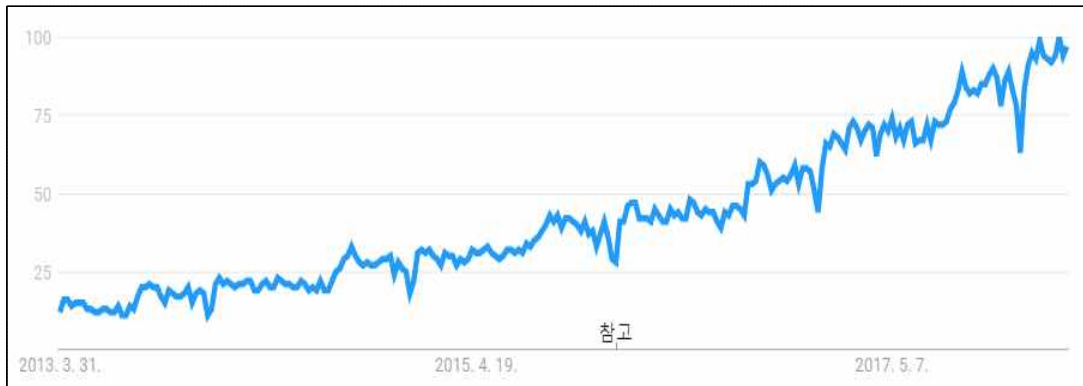
제목	내용
통계적 소양의 교육적 의미 고찰(강현영, 2012)	통계적 소양의 정의, 구성요소를 살펴보고 통계 교육을 위한 시사점을 제시
국내외 통계교육 연구동향 비교 분석(이은희 & 김원경, 2015)	국내외 통계교육 연구 동향을 분석함으로써 국내 통계교육 연구에 도움이 되는 시사점을 찾아내고 통계교육 연구의 기초 자료를 제공하는데 그 목적이 있음. 국내외에서 모두 교수·학습, 통계적 개념의 이해 및 추론의 순서로 연구가 많이 이루어졌음. 특히, 교수·학습에서 국내는 교수자료에 대한 연구가 많았으나 해외는 교수·학습 전략에 대한 연구가 많았음. 국내에서는 실험 연구 대상이 주로 중등 학생인데 비해 해외에서는 주로 대학 이상의 학생임
데이터 사이언스와 데이터 리터러시(이재윤, 2015)	데이터 사이언스가 경영, 학술, 교육 등의 전 영역에 확산되면서 변화를 이끌어내고 있음. 데이터 리터러시는 데이터와 관련된 기술과 역량을 강조하는 용어로서 문헌정보학과 도서관 현장에서 데이터 사이언스를 주도적인 입장에서 수용하는 브랜드로 육성시킬 필요가 있음
특허와 데이터사이언스, 그 연결은?(전성해 & 최준혁, 2015)	통계학, 컴퓨터학, 그리고 경영공학은 데이터사이언스를 이루는 학제적 연계이며 특허데이터사이언스는 특허데이터의 수집에서 분석, 저장에 이르는 전체 과정에 대한 역할을 담당함 이를 통하여 기술경영 분야의 기술 분석 및 R&D 기획에 필요한 정보를 제공하게 됨
Big Data and Science: Myths and Reality (Jagdish, 2015)	빅데이터는 사회의 모든 분야에서 불가피하게 주목을 받고 있기 때문에 많은 부분에서 어려움을 겪고 있음. 부정확한 특성화 등 문제점에 대해 살펴봄
스마트 제조를 위한 데이터사이언스-기법, 응용 및 이슈(김동일, 2016)	제조업은 과거의 열악함에서 벗어나 스마트 제조를 위해 나아가고 있으며, 스마트 공장이라는 공통된 목표를 달성하기 위해 많은 연구와 지원이 활발하게 이루어지고 있음. 데이터사이언스 역시 이 과정에서 매우 핵심적인 역할을 담당하고 있으며, 제조업은 데이터사이언스가 성공적으로 그리고 지속적으로 적용되기 위한 유리한 점들을 보유하고 있음. 아직 해결해야 할 문제나 현장의 난제들이 존재하기는 하지만, 그만큼 연구의 가능성도 매우 높은 분야라고 볼 수 있고, 산업공학의 다양한 분야와의 공동 연구 가능성 또한 매우 높다고 볼 수 있음

데이터 과학에 대해 현재까지 논의가 이루어져 온 국내·외의 관련 연구들 중 데이터 과학 교육에 대한 연구는 전혀 찾아볼 수 없었으며, 컴퓨팅 사고력, 창의성과 관련된 연구들 중에서도 데이터 과학 교육에 대한 연구는 찾아볼 수가 없었다.

2) 데이터 과학 교육의 동향

과거에는 기업이나 정부에서 다양한 의사결정을 할 때 경영자의 직관, 작업자의 경험, 혹은 장인의 노하우 등 개인의 능력과 경험에 의한 의사결정이 큰 부분을 차지했다. 데이터는 이러한 경험적 요소를 뒷받침하는 숫자 정도의 역할을 담당했다. 그러나, 최근 다양한 변수를 고려한 고도화된 의사결정이 요구됨에 따라 개인 능력에 의한 의사결정은 신뢰성과 속도 측면에서 한계를 맞게 되었으며, 데이터 기반 과학적 의사결정이 주목받기 시작했다. 데이터마이닝, 머신러닝, 패턴인식, 통계학 등 다양한 분야의 분석 모델링 기법들과, 데이터를 처리하고 저장하는 기술들, 그리고 시각화 등과 같은 사용자 경험(user experience) 기술들이 모여, 데이터 과학이라는 융합 분야를 이루게 되었다(김동일, 2016).

[그림 II-3], [그림 II-4]에서 보듯이 Google Trends를 통해 살펴본 데이터 과학이라는 용어에 대한 전 세계적인 관심도는 급격하게 증가하고 있다.



[그림 II-3] 시간 흐름에 따른 관심도 변화(최근 5년)



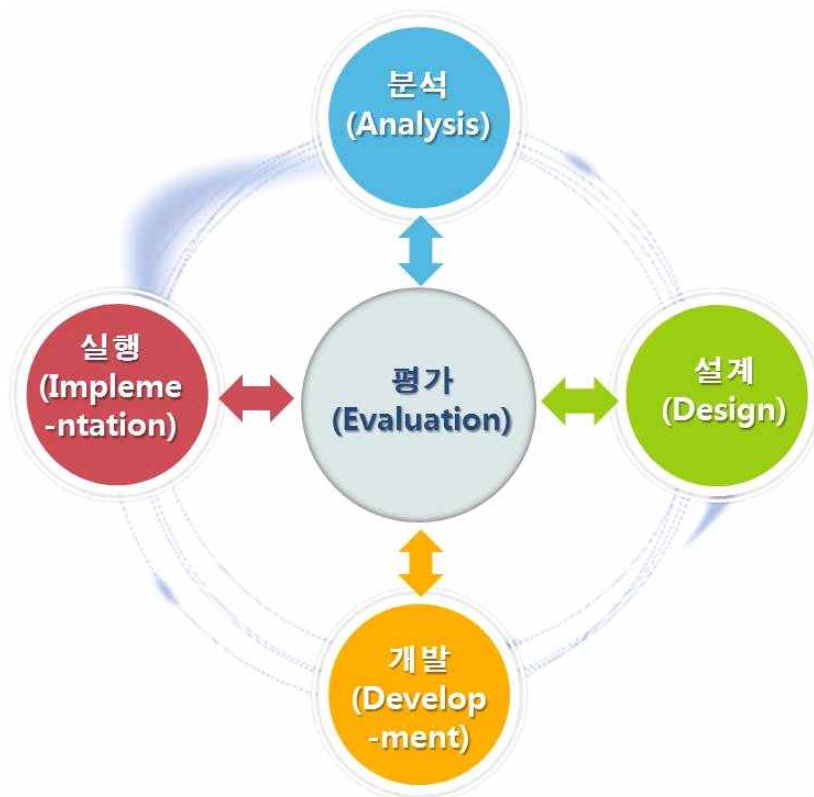
[그림 II-4] 지역별 관심도(최근 5년)

4. 프로그램 개발 모형

프로그램 개발 모형은 기본적으로 ADDIE 모형을 적용하였는데 ADDIE 모형의 경우 준비단계에 해당하는 Analysis(분석)와 Design(설계)의 단계가 복잡하고 교사가 사용하기에 복잡하다는 단점이 있어(김선태, 2011), ADDIE 모형의 단계를 따르면서 효과적인 수업설계 과정을 안내하는 체제적, 단계별 교수설계 모형 중의 하나인 Dick & Carey 모형을 함께 적용하였다.

1) ADDIE 모형

ADDIE 모형은 과정에 초점을 맞추고 있는 교수설계의 대표 모형으로 교수설계의 과정을 크게 다섯 단계, 즉 분석, 설계, 개발, 실행, 평가로 나누고, 각 단계에서 어떠한 작업들이 어떤 방법으로 이루어져야 하는지를 보여준다(Booth, 2011).



[그림 II-5] ADDIE 모형의 요소 및 개념화

[그림 II-5]는 ADDIE 모형의 요소와 개념을 도식화한 것이다. ADDIE 모형의 각 단계는 교수설계 주기에서 한 부분씩을 차지하지만, 분석, 설계, 개발, 실행은 선형의 과정이 되는 반면, 평가는 교수설계의 전체 과정에 영향을 미치는 과정이 된다(Booth, 2011).

분석(Analysis) 단계는 수업설계의 초기 단계로서 설계 상황에서 절대적인 중요성을 지닌 요소들을 분석하는 단계이다. 즉, 현재의 학습자에게 어떤 문제가 있고, 어떤 점이 부족한지를 밝혀서 전체 과정을 수행하는 타당한 근거를 확보하고, 그 문제가 해결된 상태의 진술인 일반 수준의 목표를 도출하게 된다(최재황,2016). 설계(Design) 단계는 분석 단계에서 나온 산출물을 창의적으로 종합하는 단계이며 수행목표의 명세화, 평가도구의 설계, 프로그램의 구조화 및 계열화, 교수 전략과 매체의 선정이 포함될 뿐만 아니라 수업의 전체 모습이 그려진다(최재황,2016). 개발(Development) 단계는 수업에 사용될 교수-학습 자료를 실제로 개발하고 제작하는 단계이며 교수 자료의 초안이 개발되고, 이 초안들은 몇 차례의 수정·보완의 과정을 거치게 된다(최재황,2016). 실행(Implementation) 단계는 개발 단계에서 완성된 교수-학습 자료와 다양한 매체를 실제 현장에서 활용해 보는 단계이며 평가(Evaluation) 단계는 최종 산출물이 의도한 목적을 충실히 달성하였는지를 판단하는 과정이다(최재황,2016).

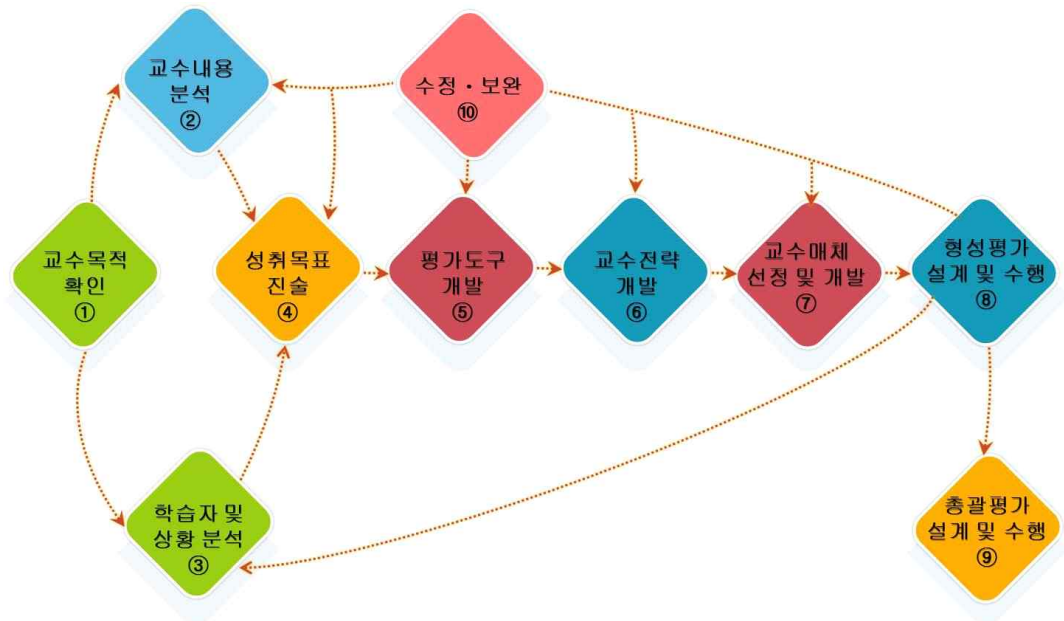
ADDIE 모형의 각 단계는 유기적으로 연관되어 있으며, 각 과정에 따른 역할, 세부 활동 및 그 산출물의 관계는 <표 II-9>와 같다(정재삼, 1998).

<표 II-9> ADDIE 모형의 과정과 산출물

과정	역할(기능)	산출물
분석	학습내용 정의	요구, 목적, 제한점, 과제
설계	교수방법 구체화	성취행동 목표, 교수 전략 등의 설계명세서
개발	교수자료 제작	완성 프로그램
실행	교수자료 실제 상황에 적용	실행 프로그램
평가	교수자료의 효과성, 효율성 결정	프로그램 평가 보고서

2) Dick & Carey 모형

1970년대 후반에 소개된 이후 지금까지 수정, 보완의 변경을 거치면서 계속하여 발전되어 온 Dick & Carey(Dick, W., Carey, L., & Carey, J.)의 모형은 효과적인 수업설계 과정을 안내하는 대표적인 체제적, 단계별 교수설계 모형 중의 하나이다. 이 모형은 한 교수체제에 대한 설계를 완성해가는 과정모형으로 기본적으로 ADDIE 모형의 단계를 따르고 있다(최재황, 2016). [그림 II-6]은 Dick & Carey 모형을 나타낸 것이고 <표 II-10>은 ADDIE 모형과 Dick & Carey 모형을 비교한 것이다(강이철, 2011).



[그림 II-6] Dick & Carey 모형의 요소 및 개념화

<표 II-10> ADDIE 모형과 Dick & Carey 모형 비교

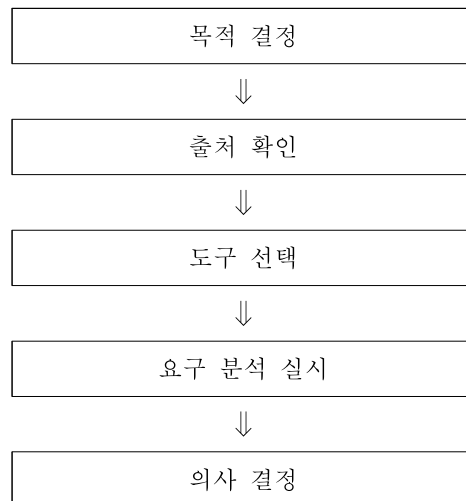
ADDIE 모형	Dick & Carey 모형
분석(A)	① 교수 목적 확인
	② 교수 내용 분석
	③ 학습자 및 상황 분석
설계(D)	④ 성취 목표 진술
	⑤ 평가에 활용될 도구 개발
	⑥ 교수에 필요한 전략 개발
개발(D)	⑦ 교수 매체 선정 및 개발
	⑧ 형성평가 설계 및 수행
실행(I)	—
평가(E)	⑨ 총괄평가 설계 및 수행

III. 데이터 과학 교육 프로그램

1. 데이터 과학 교육 프로그램을 위한 요구 분석

요구 분석은 교수·학습 목표를 추출하고, 교육 프로그램의 내용을 선정하는데 있어서 반드시 필요한 활동이다. 교수·학습자의 요구 상황을 제대로 반영하게 되면 교수·학습자에게 유용하게 활용될 수 있기 때문에 본 연구에서는 학습자뿐만 아니라 컴퓨터 교육 전공 현직 초등교사들을 대상으로도 요구 분석을 실시하였다.

본 연구에서는 Rossett의 요구 분석 모형(Rossett, 1987)을 사용하였다. Rossett 모형을 적용한 이유는 기업 교육에서 널리 활용되고 있는 대표적인 교육 요구 분석 모형으로 요구 분석의 실행과정에 초점을 둬으로써, 실제 요구 분석 실행자들이 적용하기 쉬운 안내를 제공하기 때문이다(이재무, 2014). Rossett 모형은 유발된 문제에 관한 요구 분석 목적 결정에서부터 문제 해결을 위한 의사결정까지 단계별로 제시하고 있으며 Rossett 모형의 절차는 [그림 III-1]과 같다(이재무, 2014).



[그림 III-1] Rossett 모형의 절차

Rossett 모형은 유발된 문제에 관한 요구 분석 목적 결정에서부터 문제 해결을 위한 의사결정까지 단계별로 제시하고 있다. Rossett 모형에 따른 본 연구의 요구 분석 과정(절차)은 다음과 같다.

첫째, 요구 분석 목적 선택으로, 본 연구의 요구 분석은 컴퓨팅 사고력과 창의성 향상을 위한 데이터 과학 교육 프로그램을 개발하기 위하여 학습자들의 흥미, 사전 배경 지식 등을 알아보는 데 있다.

둘째, 요구 분석 출처 확인으로, 문제 유발 상황 확인 및 요구 분석 정보 획득 출처를 결정한다.

셋째, 요구 분석 도구 선택으로, 설문조사 방법 등으로 요구 분석을 실시한다.

넷째, 요구 분석을 실시한다.

다섯째, 요구 분석 결과 분석을 통해 데이터 과학 교육 프로그램 개발에 반영할 요소를 결정한다.

1) 요구 분석 목적 결정

본 연구의 요구 분석 과정을 통해 알아보려고 하는 정보를 결정하는데 본 연구에서는 데이터 과학 교육 프로그램을 개발하기 위한 요구 분석을 실시한다.

2) 출처 확인

출처 확인은 정보의 출처를 확인하는 단계인데, 본 연구에서는 데이터 과학 교육 프로그램의 교수·학습자들을 중심으로 요구 정보를 추출하였고, 정보 출처와 관련된 대상과 내용에 대해 <표 III-1>에 제시하였다.

<표 III-1> 정보원과 추출 가능한 정보

정보원	추출 가능한 정보
학습자	데이터 과학 교육 프로그램 설계·개발에 따른 학습자의 실태 및 요구되는 정보
교수자	데이터 과학 교육 프로그램 설계 및 개발에 따른 학습 환경 및 교수 전략 등에 대한 지식·정보

3) 도구 선택

본 연구의 요구 분석 도구는 연구자가 직접 개발한 설문지를 사용하였고, 설문 내용은 데이터 과학 교육 프로그램 설계 및 개발을 위한 학습자의 실태에 대해 묻는 내용이 중심이다.

설문 형태는 폐쇄형 질문을 중심으로 학습자들의 생각을 효율적으로 알아볼 수 있도록 개방형 질문과 혼합되어 있고, 폐쇄형 질문의 구성은 5단계 리커트 척도(likert scale)이다.

4) 요구 분석 실시

데이터 과학 교육 프로그램에 대하여 다음과 같은 내용의 요구 분석을 실시하였다.

- 소프트웨어교육 참여 경험
- 교육용 프로그래밍 언어, 언플러그드 활동, 피지컬컴퓨팅에 대한 관심도
- 데이터 과학 교육을 통한 능력 향상
- 데이터 과학 교육 활용 데이터
- 데이터 과학 교육 학습 방법

(1) 요구 분석 대상

<표 III-2> 초등학생

4학년	5학년	6학년
84(41%)	39(19%)	82(40%)

<표 III-3> 교사-컴퓨터 관련 소지 학위

학사	석사과정	석사	박사과정	박사수료	박사
3(15%)	3(15%)	6(30%)	3(15%)	2(10%)	3(15%)

<표 III-2>, <표 III-3>과 같이 제주대학교에서 실시하는 ‘창의컴퓨터교실(교육기부 프로그램)’에 지원한 초등학생을 포함하여, 제주특별자치도내 초등학교

4~6학년 205명의 학생들과 제주대학교 1학년 학생 19명, 컴퓨터교육 전공 초등 현직교사 20명을 대상으로 요구 분석을 실시하였다.

(2) 소프트웨어교육 참여 경험

소프트웨어교육 참여 경험에 대하여 설문 실시 결과는 <표 III-4>에 제시하였다.

<표 III-4> 소프트웨어교육 참여 경험

내용	구분	있다	없다.
교내외에서 소프트웨어교육에 참여해 본 경험이 있나요?	초등학생	152(74.1%)	53(25.9%)
	대학생	4(21.1%)	15(78.9%)
	교사	20(100%)	0
‘예’ 라고 대답한 사람만 어떤 교육에 참여해 본 경험이 있나요?	초등학생	<ul style="list-style-type: none"> 교육용 프로그래밍 언어 107(70.4%) 언플러그드활동 23(15.1%) 피지컬컴퓨팅 22(14.5%) 	
	대학생	<ul style="list-style-type: none"> 기타-데이터 과학(엑셀 등) 7(46.7%) 교육용 프로그래밍 언어 5(33.4%) 피지컬컴퓨팅 2(13.3%) 언플러그드활동 1(6.6%) 	
	교사	<ul style="list-style-type: none"> 교육용 프로그래밍 언어 12(60%) 언플러그드활동 5(25%) 기타 2(10%) 피지컬컴퓨팅 1(5%) 	

대학생을 제외하고 초등학생, 교사는 대부분 소프트웨어교육 경험을 갖고 있는 것으로 나타났다. 따라서, 대학생과 예비 코딩 강사를 대상으로 하는 교육 프로그램의 경우에는 초보자들도 다루기 쉬운 교육용 프로그래밍 언어를 활용할 필요가 있다.

(3) 교육용 프로그래밍 언어, 언플러그드 활동, 피지컬 컴퓨팅에 대한 관심도
교육용 프로그래밍 언어, 언플러그드 활동, 피지컬 컴퓨팅에 대한 관심도를

알아보기 위하여 설문을 실시하였는데, 결과는 <표 III-5>, <표 III-6>, <표 III-7>에 제시하였다.

<표 III-5> 교육용 프로그래밍 언어에 대한 관심도

내용	구분	평균	전혀 들어본 적 없다 (1점)	들어본 적 없다 (2점)	보통이다 (3점)	관심이 있다 (4점)	매우 관심이 많다 (5점)
‘교육용 프로그래밍 언어’에 대한 관심도는 얼마나 되나요?	초등학생	3.3	26(12.7%)	23(11.2%)	58(28.3%)	57(27.8%)	41(20%)
	대학생	3	1(5.3%)	7(36.8%)	5(26.3%)	3(15.8%)	3(15.8%)
	교사	4.85	0	0	1(5%)	1(5%)	18(90%)

<표 III-6> 언플러그드활동에 대한 관심도

내용	구분	평균	전혀 들어본 적 없다 (1점)	들어본 적 없다 (2점)	보통이다 (3점)	관심이 있다 (4점)	매우 관심이 많다 (5점)
‘언플러그드활동’ 에 대한 관심도는 얼마나 되나요?	초등학생	3.02	35(17.1%)	26(12.7%)	69(33.6%)	49(23.9%)	26(12.7%)
	대학생	2.79	3(15.8%)	5(26.3%)	5(26.3%)	5(26.3%)	1(5.3%)
	교사	4.8	0	0	1(5%)	2(10%)	17(85%)

<표 III-7> 피지컬컴퓨팅에 대한 관심도

내용	구분	평균	전혀 들어본 적 없다 (1점)	들어본 적 없다 (2점)	보통이다 (3점)	관심이 있다 (4점)	매우 관심이 많다 (5점)
‘피지컬컴퓨팅’에 대한 관심도는 얼마나 되나요?	초등학생	3.02	38(18.5%)	26(12.7%)	67(32.7%)	42(20.5%)	32(15.6%)
	대학생	2.84	5(26.3%)	4(21.1%)	7(36.8%)	0	3(15.8%)
	교사	4.85	0	0	0	2(10%)	18(90%)

초등학생, 대학생 모두 교육용 프로그래밍 언어에 대한 관심도가 가장 높게 나타나, 데이터 과학 교육 프로그램 교수 전략으로 교육용 프로그래밍 언어를 사용하는 것이 효과가 높을 것으로 생각된다.

(4) 데이터 과학 교육을 통한 능력 향상

<표 III-8> 데이터 과학 교육을 통한 능력 향상

내용	구분	컴퓨팅 사고력	창의성	논리적 사고력	문제해결력	정보활용 능력
‘데이터 과학’ 교육에 참여한다면 어떤 능력 향상에 도움이 된다고 생각하는지 2가지를 고르면?	초등학생	63(30.7%)	57(27.8%)	33(16.1%)	21(10.3%)	31(15.1%)
	대학생	15(39.5%)	4(10.5%)	3(7.9%)	4(10.5%)	12(31.6%)
	교사	14(70%)	1(5%)	6(30%)	7(35%)	6(30%)

데이터 과학 교육 프로그램을 통하여 기대되는 효과로는 컴퓨팅 사고력과 창의성 순으로 나타났다. 따라서, 본 연구는 컴퓨팅 사고력과 창의성 등을 높여줄 것으로 여겨진다.

(5) 데이터 과학 교육 활용 데이터

<표 III-9> 데이터 과학 교육 활용 데이터

내용	구분	개인 데이터	공공 데이터	개인 데이터+ 공공 데이터
‘데이터 과학’ 교육에 참여할 기회가 주어진다면 어떤 데이터를 대상으로 학습하고(가르치고) 싶나요?	초등학생	145(70.7%)	60(29.3%)	
	대학생	11(57.9%)	8(42.1%)	
	교사	1(5%)	3(15%)	16(80%)

초등학생의 경우, 개인 데이터와 공공 데이터 사이에 선호도 차이가 뚜렷이 나타났고, 대학생의 경우는 비슷하게 나타났으며, 교사의 경우, 개인 데이터와 공공 데이터를 함께 활용하는 것을 선호하는 것으로 나타났다. 따라서, 교육 프로그램 설계에 이를 고려할 필요가 있다.

(6) 데이터 과학 교육 학습 방법

<표 III-10> 데이터 과학 교육 학습 방법

내용	구분	강의/실습	팀별 프로젝트 학습	개별 프로젝트 학습	개별+팀별 프로젝트 학습
‘데이터 과학’ 교육에 참여할 기회가 주어진 다면 어떤 방법으로 학 습하고(가르치고) 싶나 요?	초등학생	39(19%)	124(60.5%)	42(20.5%)	
	대학생	11(57.9%)	3(15.8%)	5(26.3%)	
	교사	0	6(30%)	0	14(70%)

초등학생, 대학생, 교사 모두 선호하는 학습 방법이 제각각 다른 것으로 나타나 연구 대상에 따라 선호하는 학습 방법을 적합하게 적용한다면 교육적 효과가 높을 것으로 기대된다.

5) 요구 분석 의사 결정

요구 분석 결과는 다음과 같다.

첫째, 초등학생, 대학생, 교사 모두 교육용 프로그래밍 언어에 대한 관심도가 가장 높게 나타나, 데이터 과학 교육 프로그램 교수 전략으로 교육용 프로그래밍 언어를 사용하는 것이 효과적일 것으로 보인다. 한편, 대학생이 상대적으로 소프트웨어교육 경험이 적은 것으로 나타나 대학생과 예비 코딩 강사를 대상으로 하는 교육 프로그램의 경우 초보자들이 다루기에 가장 쉬운 교육용 프로그래밍 언어를 활용할 필요가 있는 것으로 생각된다.

둘째, 초등학생, 대학생, 교사 모두 데이터 과학 교육 프로그램을 통하여 컴퓨팅 사고력과 창의성이 향상될 것으로 기대하고 있다. 따라서, 본 연구에서 개발된 교육 프로그램의 적용을 통해 컴퓨팅 사고력과 창의성이 향상되었는지에 대한 검증이 필요하다.

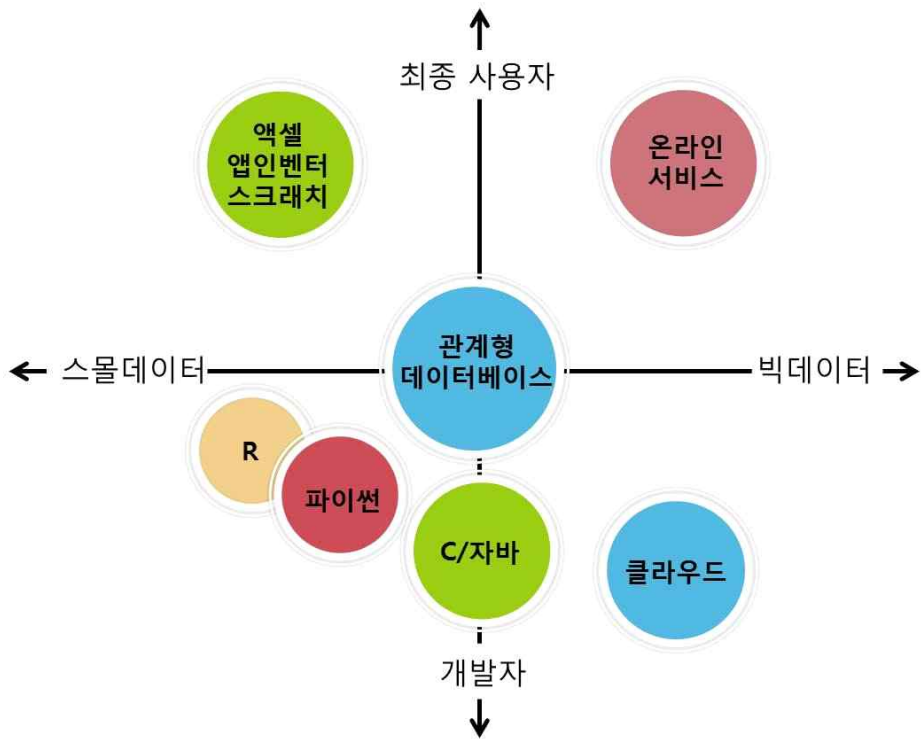
셋째, 데이터 과학 교육 프로그램에 활용할 데이터에 대해 초등학생, 대학생, 교사에 따라 개인 데이터와 공공 데이터 사이에 선호도 차이가 뚜렷이 나타나 설계 단계에서 이를 고려할 필요가 있다.

2. 데이터 과학 교육 프로그램의 설계

1) 성취 목표 진술

(1) 데이터 과학 교육 활용 도구

사용성(usability)이라는 개념의 주창자로 유명한 도널드 노먼은 보조 장치 없는 인간의 지적 능력은 미약하기 짝이 없으며 종이와 펜, 컴퓨터와 같은 인지보조장치의 사용으로 인간의 기억과 사고력은 극적으로 향상된다고 하며 ‘우리를 똑똑하게 만드는 것들(1994)’에서 인간의 미약한 기억과 판단력을 보완해주는 인지 보조 장치의 중요성에 대해 언급하고 있는데, 데이터 과학에 있어서도 도구의 중요성은 예외가 아니다(김진영, 2016).



[그림 III-2] 데이터 크기에 따른 문제 특성별 사용 도구

문제 정의와 데이터 수집, 데이터 준비, 데이터 분석, 결과 구현, 결과 소통 등 데이터 과학의 모든 프로세스를 완벽하게 지원해주는 도구를 찾을 수는 없

다. 하지만 널리 쓰이는 도구 중에서 데이터의 크기, 분석의 성격, 사용 가능한 도구를 파악하는 것이 출발점이 된다고 할 수 있다.

[그림 III-2]는 김진영(2016)이 제시한 도구들에 데이터의 크기 및 인터페이스의 성격에 따라 본 연구의 문제 해결 과정에서 활용한 교육용 프로그래밍 언어인 ‘스크래치’와 ‘앱인벤터’를 넣어 재구성하였다. 데이터 분석을 위해 널리 사용되는 R과 파이썬은 비교적 편리한 사용 환경을 제공하고, C나 자바와 같은 범용 프로그래밍언어는 좀 더 사용법이 복잡하지만 성능과 유연성이 중요한 환경에서는 최선의 선택이며, 데이터가 작고 간단한 분석 작업에는 엑셀과 같은 스프레드시트를 사용한다(김진영, 2016).

도구 선택은 작업 환경과 데이터에 따라 달라지는 부분이지만 초등학생 및 초보자들을 대상으로 한 도구는 [그림 III-2]에 의하면, 편리한 사용 환경과 사용법이 간단하고 비교적 적은 프로그래밍 경험으로도 사용할 수 있어야 한다. [그림 III-2]에 김진영(2016)이 제시한 엑셀과 함께 소프트웨어교육에서 널리 사용되고 있는 스크래치와 앱인벤터가 그 조건에 해당한다고 할 수 있다.

① 스크래치

스크래치(Scratch)란 최근 각광받는 EPL(Educational Programming Language) 중 하나로 2007년 5월 MIT 미디어랩과 UCLA대학의 연구팀이 미국국립과학재단, 마이크로소프트, 인텔, 노키아, MIT 미디어 연구소로부터 재정지원을 받아 개발하였다(길순희, 2011). 스프라이트(sprite)라는 오브젝트에 다양한 명령어를 갖는 블록(block)들을 쌓는 간단하면서도 움직임이나 소리 등의 효과를 즉시 확인할 수 있는 직관적인 조작으로 프로그래밍을 할 수 있으며, 애니메이션, 게임, 음악 등과 같은 다양한 미디어 효과 또한 구현이 가능해 흥미롭고 다채로운 결과물을 만들어낼 수 있다(구덕훈, 2007).

스크래치는 프로그래밍을 처음 접하는 학생들을 위한 교육용 프로그래밍 언어로 다음과 같은 특성을 지닌다(Resnick et al, 2009).

첫째, 누구나 쉽게 사용할 수 있다.

둘째, 다양한 유형, 개인적인 관심·경험 등에 대한 프로젝트 제작이 가능하다.

셋째, 사회적 상호작용 활동이 가능하므로, 학습자들은 웹을 통해 프로젝트

를 공유하고 비평가 의견을 주고받을 수 있다.

스크래치에 대해 현재까지 논의가 이루어져 온 국내·외의 관련 연구를 <표 III-11>에 제시하였다.

<표 III-11> 스크래치 관련 연구

제목	내용
Applying creativity in CS high school education: criteria, teaching example and evaluation(Romeike, 2007)	고등학교에서의 컴퓨터과학 학습에서 창의성에 대한 평가 준거를 설정한 프로그래밍 입문 학습에서 학습자들의 동기유발과 흥미, 학습 성취도가 증진됨
스크래치 프로그래밍 학습이 학습자의 동기와 문제해결력에 미치는 영향(송정범 외, 2008)	초등학교 6학년 학생들의 재량활동 시간을 활용하여 스크래치 프로그래밍 학습한 결과 학습자의 내재적 동기와 문제해결력 향상에 효과가 있는 것으로 나타남
CPS에 기반한 스크래치 EPL이 문제해결력과 프로그래밍 태도에 미치는 효과(조성환 외, 2008)	중학교 남녀 학습자에게 CPS 모형 기반의 스크래치 프로그래밍 학습이 문제해결력 향상과 프로그래밍 교육에 대한 태도에 긍정적인 영향을 미침
교육용 프로그래밍 언어인 로고와 스크래치 교재 개발 및 비교 실험(김종진 외, 2010)	초등학생을 대상으로 로고와 스크래치 프로그래밍 학습 후, 창의성 검사 결과를 교차 분석하여 비교한 결과, 두 언어 모두 창의성에 긍정적 영향을 미침. 특히 로고는 창의성 영역의 유창성, 스크래치는 추상성과 저항 영역에 긍정적인 영향
스크래치 프로그래밍을 활용한 게임중독 치료 프로그램의 개발(한선관 외, 2010)	초등학교 게임중독 고위험군을 대상으로 게임중독 치료 대안활동과 함께 스크래치 교육 프로그램을 적용한 결과, 게임중독 치료에 긍정적인 답변을 얻고, 휴식 시간 게임 접속 시간이 감소된 것으로 분석되었음. 관찰과 면접을 통합 분석 결과 자아존중감과 창작의 기쁨, 자아 제어 능력에도 긍정적인 영향
스크래치 프로그래밍 교육이 초등학생의 학습 몰입과 프로그래밍 능력에 미치는 효과(안경미 외, 2011)	초등학교 고학년을 대상으로 스크래치 프로그래밍 학습 집단과 순서도 학습 집단을 비교한 결과, 스크래치 프로그래밍 학습이 상대적으로 학습몰입 수준과 프로그래밍 능력에 긍정적인 영향
스크래치 프로그래밍 교육이 초등학생의 자기 주도적 학습 능력에 미치는 효과(박용철 &	초등학교 6학년 학생들에게 스크래치 프로그래밍을 실시한 결과, ICT 소양 프로그래밍 수업에 비해 자기주도적 학습 능력의 하위요소인 개방성, 내재적 동기, 자

제목	내용
이수정, 2011)	울성 영역의 신장에 유의한 차이를 보임. 특히 ICT 활용 능력이 우수한 학생들에게 그러한 신장이 크게 나타남
스크래치를 이용한 초등학교 컴퓨터 교육과정 설계(함성진 & 양창모, 2011)	프로그래밍 교육을 위한 교육과정을 설계하여 전문가 집단으로부터 학습 대상, 교육 요소, 학습 내용, 학습량, 교수학습지도안의 적정성 등의 거의 모든 영역에서 적절하다는 평가를 받음
스크래치를 활용한 정보과학적 사고 기반 퍼즐 교육(송정미, 2011)	문제해결과정으로서 퍼즐을 도입하여 스크래치 교육과 연계. 학생들이 펜토미노 구체물이 아닌 스크래치를 이용하여 문제를 해결하고 이 과정에서 도형의 개념과 퍼즐의 원리를 습득
Scratch가 초등 영재들의 창의적 문제해결력에 미치는 효과(류충구, 2012)	초등학교 영재학생들을 대상으로 스크래치 프로그래밍 학습이 비주얼 베이직 프로그래밍 수업보다 창의적 문제해결력에 긍정적인 효과
기술교육에서 중학생의 논리적 사고력 함양을 위한 스크래치 학습 프로그램 개발(조준필, 2012)	중학생들의 기술교육에 있어서 논리적 사고력을 함양시키기 위한 스크래치 학습 프로그램을 전문가의 검증을 통해 개발
프로그래밍 입문 수업에서 스크래치 활용 효과분석(박정신 외, 2012)	전문대학의 컴퓨터 전공 학생들이 프로그래밍언어 수업에서 스크래치를 활용하여 비활용반과의 비교에서 중간고사와 기말고사의 점수를 비교하여 알고리즘 구상과 프로그래밍 작성에 효과
CPS 모형 기반 스크래치 프로그래밍 학습이 언어 창의성에 미치는 영향(김병수 & 김종훈, 2013)	창의적 문제 해결(CPS) 모형 기반의 스크래치 교육 프로그램이 학습자의 창의성에 미치는 영향을 분석. 기존 연구들과는 달리 TTCT 도형 검사지를 사용하지 않고 스크래치 블록 프로그래밍 활동을 언어 창의성과 연결시켜 언어 창의성 신장에 효과

<표 III-11>에서 보듯이 스크래치와 관련된 연구들 중 본 연구에서 개발하고자 하는 컴퓨팅 사고력과 창의성 향상을 위한 데이터 과학 교육 프로그램에 관련된 연구는 전혀 찾아볼 수 없다. 본 연구는 이러한 선행 연구의 제한점을 넘어 컴퓨팅 사고력과 창의성 향상을 위한 데이터 과학 교육 프로그램을 개발하고 학습자에게 투입한 후 효과를 검증한다는 점에서 차이가 있다.

② 앱인벤터

앱인벤터(App Inventor)는 프로그래밍 지식이 없는 초보자도 쉽게 드래그 앤 드롭으로 블록을 조합하여 원하는 앱을 제작할 수 있는 무료 웹서비스이다. 2010년 Google에서 MIT 앱인벤터로 바뀌었고, 별도의 고성능 하드웨어가 필요하지 않아 운영체제의 영향을 받지 않고 웹브라우저에서 사용이 가능하여 교육용으로 탁월하다(CSTA & ISTE, 2011b). 앱인벤터의 개발환경은 Mac OS, GNU/Linux, Windows OS와 안드로이드 환경의 스마트 기기를 지원하며 앱을 제작하기 위한 환경은 자바와 크롬 브라우저에 최적화되어 있다(임화경, 2013).

앱인벤터는 객체지향언어로 교육용 언어로 알맞으며 학생들이 스스로 문제 해결에 집중할 수 있도록 해 준다(성은지, 2014). 뿐만 아니라 학생들의 삶과 밀접한 관련이 있기 때문에 동기 부여가 잘되고, Google App Inventor Group 이 프로그래밍 교육을 하는데 있어 많은 도움이 된다고 한다(Morelli, R., 2010). 또한, 다른 프로그래밍 언어와 달리 누구나 쉽게 접근할 수 있고 실제적인 경험과 아울러 문법을 익히는 데 시간과 노력이 적게 들기 때문에 심리적 부담을 덜어줄 수 있다(김용민 & 김종훈, 2015).

앱인벤터 시스템은 앱 설계기(App Inventor designer), 블록 편집기(block editor)두 개의 화면으로 구성되는데, 앱 설계기는 버튼, 이미지, 소리, 동영상 등 앱에서 사용할 요소들을 배치하고 설계하는 도구로써 별도의 프로그램이 아닌 앱인벤터 홈페이지에서 실행되며, 팔레트 영역, 뷰어 영역, 콤포넌트 영역, 속성 영역의 세로로 구분된 4개의 영역으로 구분된다(성은지, 2014). 블록 편집기는 그래픽 코드 블록들을 사용하여 앱 요소들의 동작을 작성하는 도구로써 일반 안드로이드 SDK에서의 자바 프로그래밍에 해당되는데, 스마트폰 앱의 동작방식은 이벤트 기반이며 주변 상황의 어떤 변화, 즉 발생 가능한 각각의 이벤트에 원하는 동작을 연결하는 방식으로 이 작업은 블록 편집기에서 이루어진다(성은지, 2014).

앱인벤터에 대해 현재까지 논의가 이루어져 온 국내·외의 관련 연구를 <표 III-12>에 제시하였다.

<표 III-12> 앱인벤터 관련 연구

제목	내용
Using a Discourse-Intensive Pedagogy and Android's App Inventor for Introducing Computational Concepts to Middle School Students(Grover & Pea, 2013)	앱인벤터를 활용하여 모바일 앱 프로그래밍을 함으로써 컴퓨터 과학의 개념을 이해하는데 도움이 될 수 있는 교수학습방법
Mobile App Design for Teaching and Learning(Hsu & Ching, 2013)	앱인벤터를 활용하여 교수자들을 위한 모바일 앱 디자인 온라인 수업을 진행하였으며, 학습경험과 인식에 긍정적인 효과
Can android app inventor bring computational thinking to k-12(Morelli et al, 2011)	앱인벤터는 쉽게 접근가능하고 강력한 플랫폼을 갖고 있어서 복잡한 모바일 앱도 자신의 아이디어에 따라 쉽게 만들 수 있음
앱인벤터를 이용한 실시간/비실시간 질의-응답 앱 설계 및 개발(박찬정 외, 2012)	앱인벤터를 활용하여 물리적 공간에 제한받지 않는 실시간/비실시간 질의-응답 앱을 설계하고 구현
앱인벤터를 이용한 초등정보영재 앱 제작 교육프로그램 개발(황성진 외, 2013)	앱인벤터를 활용하여 컴퓨팅 사고력 교육 과정에 기반한 교육프로그램을 개발하여 초등정보영재학생들의 학습에 대한 지속적인 흥미와 몰입이 향상된 것으로 나타남
초등학생을 대상으로 앱인벤터를 활용한 안드로이드용 앱 제작교육(임화경, 2013)	초등학생을 대상을 앱인벤터를 활용하여 앱 제작학습을 실시하였으며, 학습자의 사고의 변화와 동료평가를 통하여 학습자 반응을 관찰. 학습자들은 강한 의욕과 성취감을 나타냈으며 아이디어를 앱으로 실제로 구현할 수 있음에 긍정적인 변화가 나타남
위터폴 모델을 적용한 앱인벤터 프로그래밍 교재개발 연구(설문규 & 손창익, 2013)	소프트웨어 개발에 효과적인 위터폴 모델을 활용하여 실생활과 관련된 상황중심 내용 전개를 토대로 앱인벤터를 활용한 프로그래밍 교육 방안을 제시
안드로이드 앱인벤터를 활용한 컴퓨터 프로그래밍 교육(김병호, 2013)	앱인벤터를 활용하여 안드로이드 스마트폰 앱 개발을 위한 컴퓨터 프로그래밍 교육을 실시하고 교육효과를 분석
CPS에 기반한 스크래치 EPL이 문제해결력과 프로그래밍 태도에 미치는 효과(조성환 외, 2008)	앱인벤터는 실제적인 경험을 주고, 학습자가 쉽게 접근할 수 있는 장점이 있음

<표 III-12>에서 보듯이 국내·외의 앱인벤티 관련 연구는 소프트웨어교육을 처음 접하는 학습자를 대상으로 하는 연구가 대부분이라는 점이 특징적이다. 하지만, 대부분의 해외 연구를 살펴보면, 앱인벤티는 대학생 및 교사를 대상으로 한 프로그래밍 학습 활동에 대부분 활용되고 있으며, 국내의 앱인벤티 관련 여러 가지 연구들 또한 초등학교 및 대학에서 이루어진 프로그래밍 사례 연구로 본 연구에서 개발하고자 하는 초등학생을 대상으로 하는 컴퓨팅 사고력과 창의성 향상을 위한 데이터 과학 교육 프로그램에 관련된 연구는 전혀 찾아볼 수 없었다. 본 연구는 이러한 선행 연구의 제한점을 넘어 초등학생을 대상으로 하는 컴퓨팅 사고력과 창의성 향상을 위한 데이터 과학 교육 프로그램을 개발하고 수업에 투입한 후 효과를 검증한다는 점에서 차이가 있다.

③ 엑셀

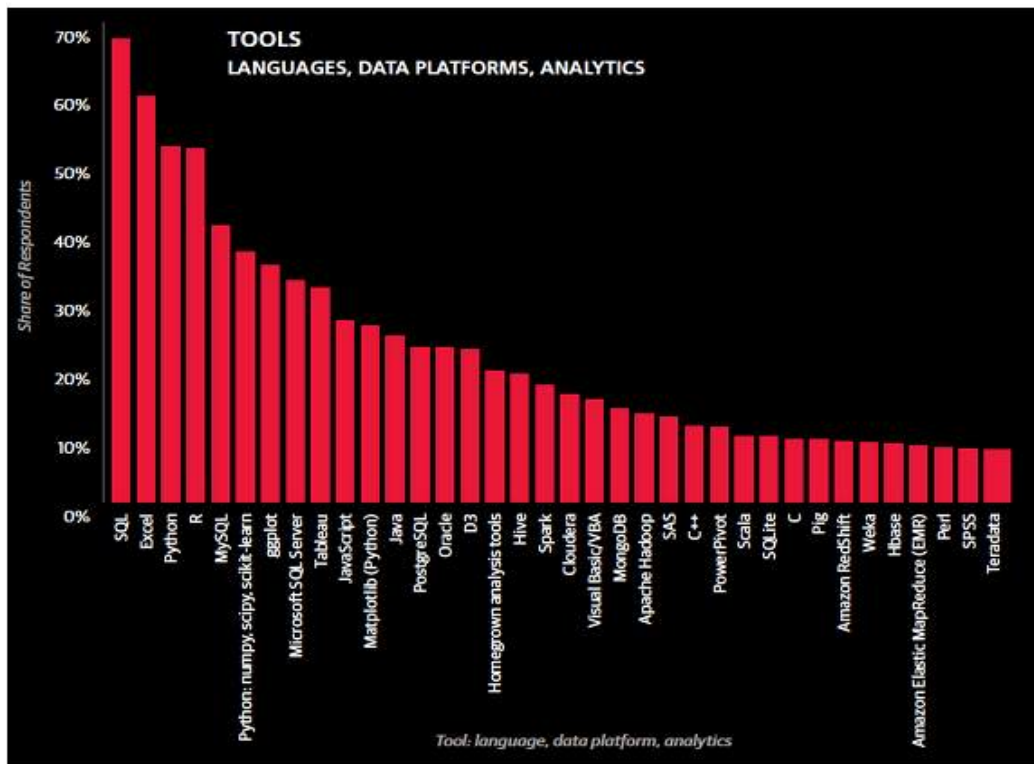
엑셀(Excel)에 대해 현재까지 논의가 이루어져 온 국내·외의 관련 연구를 <표 III-13>에 제시하였다.

<표 III-13> 엑셀 관련 연구

제목	내용
Spreadsheet use in elementary statistics course(Piotr, 2001)	수학적 개념 구성 및 긍정적 태도를 함양할 수 있음
엑셀을 활용한 중학교 수학통계교육에 관한 연구(안기완, 2003)	중학교통계교육에 있어서 엑셀을 활용하여 학습자들이 통계를 쉽고 재미있는 수학의 한 부분으로 인식 할 수 있음
엑셀을 활용한 변수 개념 지도에 관한 연구(양혜진, 2003)	엑셀이라는 소프트웨어가 학생들이 변수 개념을 이해하는데, 초등 수학에서의 구체적인 활동과 중등 수학에서의 형식적인 변수의 의미를 연결시킬 수 있는 도구가 될 수 있음을 알 수 있음
Excel과 SPSS를 이용한 통계지도 자료 개발(김은정, 2004)	통계 프로그램을 이용하여 직접 통계개념을 이해하는 과정을 통해 올바른 판단력과 분석력을 기를 수 있음
엑셀을 활용한 '확률과 통계'의 효율적인 교수·학습 과정안 연구(김윤석, 2005)	Excel을 활용하여 확률과 통계를 가르쳤을 때 문제해결력 향상에 영향을 줌
엑셀을 통한 일차함수의 과정-대상 관점 형성(이광상, 2005)	스프레드시트를 활용한 수업에서 학생들이 귀납적인 관찰을 통해 일반적인 규칙을 발견하려는 성향을 보였으며, 스프레드시트를 활용한 활동은 지필 수업을 보완하는 탐구 학습 형태로 진행할 수 있음

<표 III-13>에서 보듯이 국내·외의 엑셀을 대상으로 한 연구 중에 본 연구에서 개발하고자 하는 데이터 과학 교육 프로그램이나 컴퓨팅 사고력, 창의성 향상에 대한 연구는 전혀 찾아볼 수 없었다. 특히, 2005년 이후에는 엑셀과 관련된 교육 분야의 연구가 전무하였다.

“엑셀? 엑셀로 뭘 할 수 있는데?”라고 생각할지도 모른다. 하지만 2014년 오라일리(<http://oreil.ly>)에서 데이터 과학에 종사하는 800여 명의 현업 실무자를 대상으로 실시한 여론조사에 따르면 [그림 III-3]¹⁾과 같이 가장 널리 사용되는 데이터 과학의 도구는 SQL, 엑셀, 파이썬(python), R 순으로 나타났다.



[그림 III-3] 데이터 과학 설문 결과: 가장 널리 사용되는 데이터 과학의 도구들

엑셀은 행과 열로 짜여진 전자 작업지 위에서 수량적 자료 처리와 그래프 그리기가 가능한 스프레드시트 프로그램의 일종으로, 우리 주위에서 가장 쉽게 접할 수 있고 가장 많이 사용되는 프로그램이며, 컴퓨터의 강력한 계산 기능은

1) 2015 Data Science Salary Surbay, <http://oreil.ly/1KENyAf>

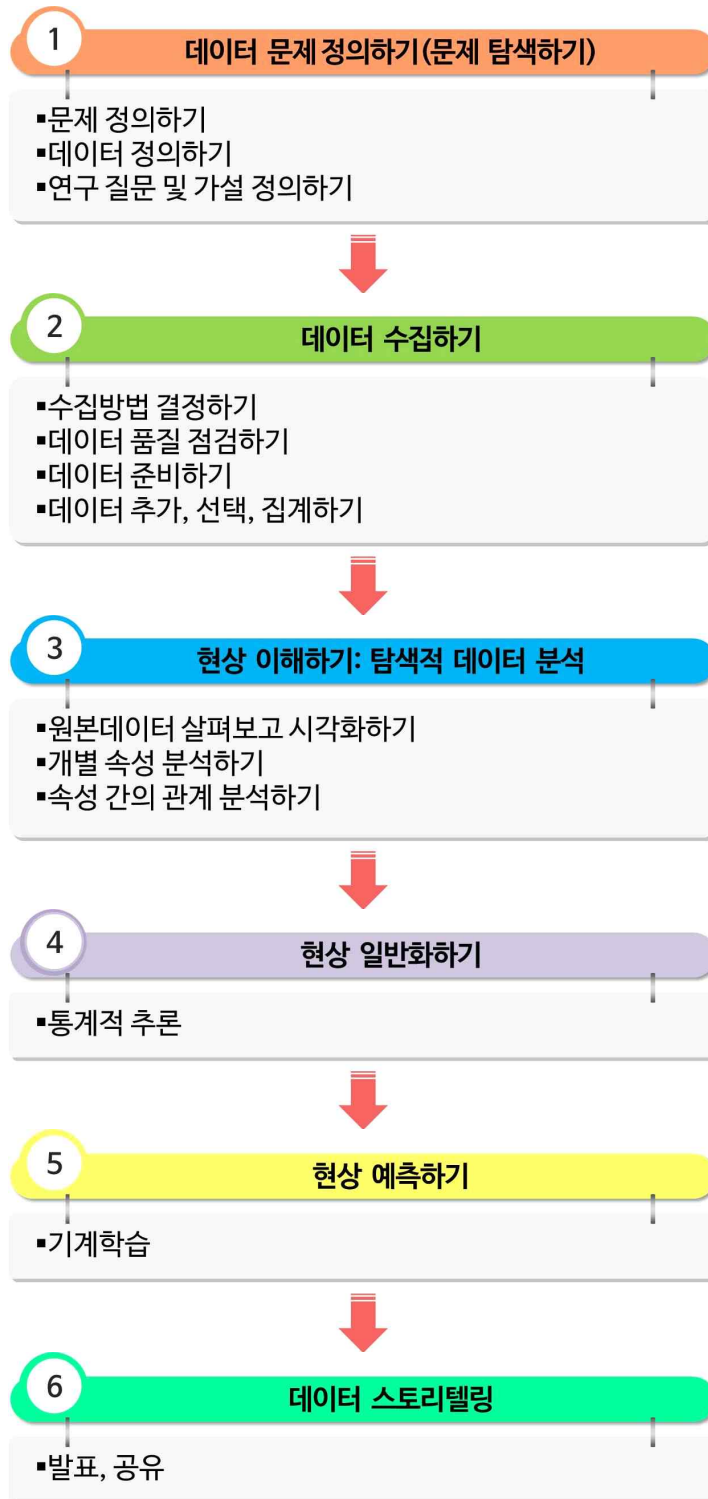
물론 시각화의 기능을 가장 직접적이고 효과적으로 사용할 수 있다(강문봉, 2009). 그래서 자료를 입력하여 행 또는 열에 따라 자료를 정렬하고, 함수식을 사용하여 계산 결과를 출력하는 경우에, 또는 함수 그래프를 그리고 통계 그래프를 그리며 탐구 활동을 하는 데 사용할 수 있다(강문봉, 2009).

데이터 과학을 위한 도구로써 엑셀이 가진 장점은 다음과 같다. 첫째, 프로그래밍에 익숙하지 않은 사람에게 새로운 언어를 공부한다는 것은 분명 큰 부담이 될 수 있는데 반해 엑셀은 코드가 아닌 데이터가 중심이다. 둘째, 데이터 과학의 전 과정을 엑셀로 해볼 수 있기 때문에 올인원 솔루션이다. 셋째, 결과물을 쉽게 공유할 수 있다(김진영, 2016).

앞에서 살펴본 것처럼 스크래치와 애플벤터는 엑셀이 가진 장점을 충족하는 교육용 프로그래밍 언어이기 때문에 본 연구에서는 엑셀과 함께 스크래치, 애플벤터를 데이터 과학 교육 프로그램 개발 도구로 활용하였다.

(2) 데이터 과학 교육 문제 해결 단계

문제 해결의 세부적인 단계는 문제의 특성에 따라 달라지겠지만, 데이터에 기반한 문제 해결의 일반적인 단계(김진영, 2016)는 [그림 III-4]와 같이 정의할 수 있다.



[그림 III-4] 데이터 과학 교육 문제 해결 단계²⁾

2) '김진영(2016). 헬로 데이터과학, 한빛미디어'를 연구자가 재구성

(3) 개인 데이터와 공공 데이터

옛날부터 사람들은 자신의 삶과 주변에서 일어나는 여러 가지 일을 기록하고 이를 통해 배우고 무언가를 향상하기 위해 노력해왔다. 이런 기록은 대부분 산문(일기) 형태를 띠지만 더 계량화된 접근을 취한 경우도 종종 발견됐다. 디지털 기기를 이용하지는 않았지만 이런 기록들이 개인 데이터 활용의 원형이라고 할 수 있다(김진영, 2016). 최근 들어 스마트 워치를 비롯한 웨어러블 기기의 확산에 따라 개인 데이터 수집과 활용에 대한 관심이 늘어나고 있다.

공개 데이터 중에서 가장 잘 알려진 것이 정부 등 각종 기관에서 공개하는 공공데이터일 것이다. 우리가 관심을 가질 만한 대부분의 영역에 공공기관이 존재하기 때문에, 공공 데이터를 잘 활용하면 다양한 데이터를 손쉽게 구할 수 있으며, 접근하기에 편리하고 다양한 통계자료를 손쉽게 얻을 수 있는 곳으로 통계청 누리집이 있다.

(4) 성취 목표

요구 분석 결과 및 데이터 과학 교육 문제 해결 단계를 바탕으로 <표 III-14>와 같이 성취 목표를 선정하였다.

<표 III-14> 성취 목표

연번	성취 목표
1	데이터 과학의 의미를 이해하고 이를 위한 데이터 과학 도구의 기본 기능을 익힐 수 있다.
2	데이터를 수집하고 데이터 과학 도구를 활용하여 데이터 시각화를 할 수 있다.
3	데이터 과학 교육 문제 해결 단계에 따라 문제를 해결하고 데이터 스토리텔링을 할 수 있다.

(5) 교수 전략 개발

요구 분석 결과 및 데이터 과학 교육 도구의 특징, 데이터 과학 교육 문제 해결 단계, 데이터의 종류를 바탕으로 <표 III-15>와 같이 4가지의 교육 프로

그램 교수 전략을 설계하였다.

<표 III-15> 교수 전략 설계

연번	대상	표본수 (명)	데이터 종류	교육 도구 (교수 매체)	학습 방법	수업시수 (차시)
1	초등학생	20	개인 데이터	엑셀	팀별 프로젝트학습	42
2		20	공공 데이터			42
3	초등학생	26	개인 데이터+ 공공 데이터	앱인벤터	팀별 프로젝트학습	42
4	대학생	19	개인 데이터+ 공공 데이터	스크래치	개별 프로젝트학습	21
	예비 코딩 강사	13				21

3. 데이터 과학 교육 프로그램의 개발

1) 교육 프로그램 개발 원리

본 연구의 데이터 과학 교육 프로그램은 다음과 같은 원리에 따라 개발하였다.

첫째, 문제 정의와 데이터 수집, 데이터 준비, 데이터 분석, 결과 구현, 결과 소통 등 데이터 과학 교육의 모든 프로세스를 완벽하게 지원해주는 도구를 찾을 수는 없지만 소프트웨어교육에서 널리 쓰이는 도구 중에서 데이터의 크기, 분석의 성격, 초등학생이나 초보자들을 대상으로 한 편리한 사용 환경과 사용법이 간단한 도구로 엑셀, 앱인벤터, 스크래치를 활용하였다.

둘째, 데이터 과학 교육 프로그램에 활용할 데이터는 개인 데이터와 공공 데이터로 구분하여 개인 데이터는 구글 설문지를 이용하여 해당 데이터 과학 교육 프로그램에 참여한 학습자들을 대상으로 필요할 때마다 실시간으로 수집하여 활용하고, 공공 데이터는 통계청 누리집에 있는 데이터를 다운 받아 활용하였다.

셋째, 요구 분석 결과를 고려하여 학습자들의 특징에 알맞게 개별 프로젝트 학습과 팀별 프로젝트 학습을 적절하게 활용할 수 있도록 데이터 과학 교육 프로그램을 개발하였다.

넷째, 프로그래밍에 익숙하지 않은 학습자가 새로운 언어를 익힌다는 것은 큰 부담이 될 수 있으므로 프로그래밍 언어의 습득과 같은 기능적 측면보다는 데이터 문제 해결 과정에 중점을 두어 데이터 과학 교육 프로그램을 개발하였다.

다섯째, 데이터 문제 해결 과정 중 <속성 간의 관계 분석하기>에서 초등학생에게 상관관계(correlation)분석은 어렵기도 하지만 평소 인과 관계(causal relation) 위주로 문제 해결을 했던 경험 때문에 상관관계 분석을 충실히 하지 않고 검증되지 않은 결론을 내려버리는 경우가 있을 수 있다. 하지만, 초등학생을 대상으로 한 교육 프로그램에서는 인지적 부담감을 줄이면서 학습 동기 향상을 통해 창의성을 향상시킬 수 있도록(김용민 외, 2015) 맞고 틀리고 보다는 스스로 데이터를 수집하고 문제를 해결해 나가는 과정에 더 큰 의미를 부여하여 데이터 과학 교육 프로그램을 개발하였다.

2) 교수 매체 선정 및 개발

엑셀은 우리 주위에서 가장 쉽게 접할 수 있고 가장 많이 사용되는 스프레드시트 프로그램이며, 컴퓨터의 강력한 계산 기능은 물론 시각화의 기능을 가장 직접적이고 효과적으로 사용할 수 있다(강문봉, 2009). 하지만, 엑셀에서 활용하기 위해서는 테이블 형태로 데이터를 재가공 하는 과정이 필요하다. 따라서, 엑셀을 활용한 데이터 과학 교육 프로그램에 활용한 개인 데이터는 체크리스트를 통해 스스로 본인의 데이터를 수집할 수 있도록 과제를 통해 학습자들에게 사전 제시하거나 강의실 내 학습자들을 대상으로 서로 실시간으로 수집한 ‘몸무게와 체지방’, ‘학교에서 발생하는 안전사고’ 등의 데이터와 통계청 누리집에서 다운 받은 다양한 공공 데이터를 테이블 형태로 재가공하여 활용해야 하기 때문에 데이터 과학 교육 문제 해결 단계(김진영, 2016) 중 <데이터 품질 점검하기 단계>가 매우 중요하다고 할 수 있다.

앱인벤터는 프로그래밍 지식이 없는 초보자도 쉽게 블록을 조합하여 원하는

앱을 제작할 수 있다(CSTA & ISTE, 2011b). 하지만, 엑셀에서는 기본으로 제공되는 함수식을 앱인벤터에서는 블록을 조합하여 직접 만들어야 하므로 다양하게 데이터 시각화를 하는 것은 초등학생에게는 어려울 수가 있다. 따라서, 활용되는 데이터의 크기를 너무 크지 않도록 하여 간단하게 막대그래프 형태로 데이터 시각화를 하도록 데이터 과학 교육 프로그램을 개발하였다.

스크래치도 앱인벤터와 마찬가지로 엑셀에서는 기본으로 제공되는 함수식을 블록을 조합하여 직접 만들어야 하므로 다양하게 데이터 시각화를 하는 것은 초보자들에게는 어려울 수가 있다. 따라서, 활용되는 데이터의 크기를 너무 크지 않도록 하여 간단하게 띠그래프, 막대그래프, 원그래프 형태로 데이터 시각화를 하도록 데이터 과학 교육 프로그램을 개발하였다.

요구 분석 결과 및 데이터 과학 교육 도구의 특징에 따라 <표 III-16>, <표 III-17>, <표 III-18>과 같이 교수 매체를 선정하여 데이터 과학 교육 프로그램을 개발하였다.

<표 III-16> 교육 프로그램 주제(엑셀)

차시	학습 주제	
	개인 데이터 중심	공공 데이터 중심
1~7	데이터 과학의 의미를 이해하고 이를 위한 엑셀의 기본 기능 익히기	데이터 과학의 의미를 이해하고 이를 위한 엑셀의 기본 기능 익히기
8~14	생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 보기 1	데이터를 수집하고 엑셀 프로그램을 활용하여 데이터를 가공하고 문제를 해결하기 1
15~21	생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 보기 2	데이터를 수집하고 엑셀 프로그램을 활용하여 데이터를 가공하고 문제를 해결하기 2
22~28	생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 보기 3	데이터를 수집하고 엑셀 프로그램을 활용하여 데이터를 가공하고 문제를 해결하기 3
29~25	평소에 내 자신 또는 내 주변에서 고쳤으면 하는 점이나 궁금한 점을 생각하여 주제를 정하고 스스로 데이터를 수집하여 이를 해결해 보기	평소에 내 자신 또는 내 주변에서 고쳤으면 하는 점이나 궁금한 점을 생각하여 주제를 정하고 데이터를 수집하여 이를 해결해 보기
35~42	데이터 스토리텔링하기	데이터 스토리텔링하기

<표 III-17> 교육 프로그램 주제(앱인벤터)

차시	학습 주제
1~7	데이터 과학의 의미를 이해하고 이를 위한 앱인벤터의 기본 기능 익히기
8~14	생활 속에서 궁금한 내용을 공공 데이터를 수집하여 해결해 보기 1
15~21	생활 속에서 궁금한 내용을 공공 데이터를 수집하여 해결해 보기 2
22~28	생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 보기 1
29~25	생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 보기 2
35~42	데이터 스토리텔링하기

<표 III-18> 교육 프로그램 주제(스크래치)

차시	학습 주제
1~3	데이터 과학의 의미를 이해하고 이를 위한 스크래치의 기본 기능 익히기
4~6	반복 구조, 조건 구조, 변수와 리스트, 펜 기능의 이해, 블록 정의하기(재귀)
7~9	생활 속에서 궁금한 내용을 공공 데이터를 수집하여 해결해 보기 1
10~12	생활 속에서 궁금한 내용을 공공 데이터를 수집하여 해결해 보기 2
13~15	생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 보기 1
16~18	생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 보기 2
19~21	데이터 스토리텔링하기

4. 데이터 과학 교육 프로그램의 적용

컴퓨팅 사고력과 창의성 향상을 위해 설계하고 개발한 데이터 과학 교육 프로그램을 초등학생, 대학생, 예비 코딩 강사에게 적용하여 프로그램의 교육적 효과를 검증하였다. 본 연구에서는 데이터 과학 교육 도구 및 교육 대상에 따라 4가지의 데이터 과학 교육 프로그램을 개발하였다.

1) 엑셀을 활용한 초등학생 대상 데이터 과학 교육 프로그램

엑셀을 활용한 데이터 과학 교육 프로그램은 제주대학교에서 실시하는 ‘창의 컴퓨터교실’에 지원한 지원자 표본(volunteer sample) 40명의 학생을 20명씩 무작위로 두 그룹으로 나누어 2개의 데이터 과학 교육 프로그램을 적용하였다.

첫 번째 그룹은 평소에 내 자신이나 또는 내 주변에서 고쳤으면 하는 점이나 궁금한 점을 생각하여 주제를 정하고 스스로 데이터를 수집하여 이를 해결해 보기 등 개인데이터를 활용하여 초등학생들이 가장 선호한 팀별 프로젝트 학습 중심으로 교육을 진행한 후 컴퓨팅 사고력에 어떠한 변화가 있는지를 분석하였다.

두 번째 그룹은 ‘제주의 자연과 온실 가스 배출량’, ‘패스트푸드와 비만을 상관관계 알아보기’ 등 제주 지역과 관련된 공공데이터를 활용하여 팀별 프로젝트 학습 중심으로 교육을 진행한 후 컴퓨팅 사고력에 어떠한 변화가 있는지를 분석하였다.

또한, 개인 데이터를 활용한 그룹과 공공 데이터를 활용한 그룹의 컴퓨팅 사고력과 창의성 향상에 어떤 차이가 있는지 프로그램 적용 후 집단 간 비교 검증을 실시하였다.

초등학생을 대상으로 한 요구 분석 결과 소프트웨어교육에 대한 경험은 많지만 데이터 과학 교육을 처음 접하는 학생들이 대부분이기 때문에 김진영(2016)의 연구에서 제시한 것처럼 데이터가 작고 간단한 분석 작업에 초보자들도 쉽게 활용할 수 있는 엑셀을 활용하였다.

2) 앱인벤터를 활용한 초등학생 대상 데이터 과학 교육 프로그램

앱인벤터를 활용한 데이터 과학 교육 프로그램은 제주대학교에서 실시하는 ‘창의 컴퓨터교실’에 지원한 지원자 표본 26명의 초등학생을 대상으로 데이터 과학 교육 프로그램을 적용하였다.

초등학생을 대상으로 한 요구 분석 결과 교육용 프로그래밍 언어, 언플러그드 활동, 피지컬 컴퓨팅 중 교육용 프로그래밍 언어에 대한 관심도가 가장 높게 나타나, 데이터 과학 교육 프로그램 교수 전략으로 교육용 프로그래밍 언어를 활용하였다. 또한, 초등학생들 대부분이 소프트웨어교육 경험이 많은 것으

로 나타나 다양한 교육용 프로그래밍 언어 중 ‘미래부 지정 소프트웨어교육 선도학교’, ‘제주특별자치도교육청 지정 소프트웨어교육 중점학교’를 포함한 일반 학교에서 많이 다루고 있는 스크래치를 제외하고 앱인벤터를 활용하였다.

교사를 대상으로 한 요구 분석 결과 개인 데이터와 공공 데이터를 함께 활용하는 것이 컴퓨팅 사고력과 창의성 향상에 도움을 줄 것으로 기대하고 있을 뿐만 아니라, ‘엑셀을 활용한 초등학생 대상 데이터 과학 교육 프로그램’ 적용 결과 개인 데이터와 공공 데이터의 활용에 유의미한 차이가 나타나지 않아 앱인벤터를 활용한 초등학생 대상 데이터 과학 교육 프로그램에서는 개인 데이터와 공공 데이터를 함께 활용하였다. 또한, 초등학생들이 가장 선호한 팀별 프로젝트 학습 중심으로 교육을 진행한 후 컴퓨팅 사고력과 창의성에 어떠한 변화가 있는지를 분석하였다.

3) 스크래치를 활용한 대학생 및 예비 코딩 강사 대상 데이터 과학 교육 프로그램

스크래치를 활용한 데이터 과학 교육 프로그램은 제주대학교 1학년 대상 전공필수 ‘초등컴퓨터교육 I’을 수강하는 학생들 중 19명과 제주창조경제혁신센터에서 주관하고 제주대학교에서 진행한 ‘청년 코딩 사업’에 지원한 지원자 표본 13명의 예비 코딩 강사를 대상으로 적용하였다.

대학생을 대상으로 한 요구 분석 결과 교육용 프로그래밍 언어, 언플러그드 활동, 피지컬 컴퓨팅 중 교육용 프로그래밍 언어에 대한 관심도가 가장 높게 나타나, 데이터 과학 교육 프로그램 교수 전략으로 ‘교육용 프로그래밍 언어’를 활용하였다. 대학생은 상대적으로 소프트웨어교육 경험이 적은 것으로 나타나 초보자들이 다루기에 가장 쉬운 교육용 프로그래밍 언어인 스크래치를 활용하였다.

또한, 요구 분석 결과 개인 데이터와 공공 데이터에 대한 선호도 차이가 거의 없어 개인 데이터와 공공 데이터를 함께 활용하였고 대학생들이 가장 선호한 개별 프로젝트 학습 중심으로 교육을 진행한 후 컴퓨팅 사고력과 창의성에 어떠한 변화가 있는지를 분석하였다.

5. 데이터 과학 교육 프로그램의 평가

본 연구에서는 데이터 과학 교육 프로그램의 효과를 알아보기 위하여 데이터 과학 교육 프로그램을 적용한 후 학습자의 변화를 비교·분석하여 효과를 검증하였다.

1) 양적 연구 결과

본 연구에서 개발한 데이터 과학 교육 프로그램은 학습자의 컴퓨팅 사고력과 창의성의 향상을 목표로 하였으며, 컴퓨팅 사고력과 창의성의 변화를 측정하였다.

(1) 검사 도구

① 컴퓨팅 사고력 검사 도구

컴퓨팅 사고력 측정은 사전-사후검사 통제집단 설계(pretest-posttest control group design)를 사용하였고, 검사 도구는 계산적 인지력 검사(Computational Cognition Test) A, B형(김병수, 2014)을 사용하였다.

컴퓨팅 사고력 검사에서 계산적 인지력의 측정은 컴퓨팅 사고력의 구성 요인 중 창의성을 제외한 추상적 사고, 비판적 사고, 논리적 사고, 재귀적 사고, 알고리즘적 사고를 통합적으로 측정하는 것을 의미한다. 계산적 인지력 검사는 평가 문항의 제작, 타당도, 변별도, 난이도, 신뢰도 검증을 통해 문항을 수정·보완하여 개발되었고 개발된 검사 도구는 논리적 사고력 검사인 GALT와 높은 상관도를 보였다. 컴퓨팅 사고력에 논리적 사고가 포함되었기 때문에 이들의 높은 상관관계는 개발된 컴퓨팅 사고력 검사 도구의 타당성을 더욱 높인다고 볼 수 있으며, 컴퓨팅 사고력 검사 도구인 A형, B형 검사지는 동형 검사 신뢰도가 검증되었다(김병수, 2014).

② 창의성 검사 도구

창의성 검사 도구는 TTCT(Torrance Tests of Creative Thinking) 검사지 도형 A, B형을 사용하였다.

TTCT 검사는 그림 구성하기, 그림 완성하기, 쌍의 두 직선-선 그리기의 세 가지 활동으로 이루어져 있으며 창의력의 하위요소를 유창성, 독창성, 제목의 추상성, 정교성, 성급한 종결에 대한 저항으로 구분하였다.

유창성은 주어진 자극을 유의미하게 사용하여 해석 가능한 반응으로 표현해 내는 능력을 말하고, 독창성은 수검자의 반응이 얼마나 드물게 일어나며 특별한 것인지를 판단하며, 제목의 추상성은 수검자가 종합하고 조직화할 줄 아는 사고 과정과 관련되어 있는데, 가장 높은 수준에서는 관련 정보들의 핵심을 포착해 내고, 무엇이 중요한 것인지를 알며, 그림을 보다 깊게 풍부하게 볼 줄 아는 능력이 작용한다. 정교성은 자극도형에 대한 최소의 일차적인 반응은 하나의 단일 반응이고 내용을 자세하게 상상하고 설명할 수 있는 것은 정교성이라는 창의력의 함수라고 보는 것이며, 성급한 종결에 대한 저항은 독창적인 아이디어를 가능하게 하는 정신적 비약을 할 수 있을 만큼 충분히 긴 시간 동안 마음을 열고 있으며 그래서 성급하게 반응을 폐쇄하고 종결시키는 것을 지연시킬 줄 아는 능력이다. 창의력의 5가지 하위 요소들은 모두가 기준 자료를 이용하여 해석하는 기준 관련 척도인데, 창의적 강점 체크리스트는 절대 기준을 사용하는 준거관련 척도들로 정서적 표현, 이야기의 명료성, 운동 또는 행위, 제목의 표현성, 불완전 도형들의 종합, 선들의 종합, 독특한 시각화, 내적인 시각화, 경계의 확대 또는 파괴, 유머, 심상의 풍부함, 심상의 다채로움, 환상의 13가지 항목으로 구성되어 있으며, 이러한 창의적 강점을 교육과정과 수업방법을 개발하는데 사용할 수 있다(김영채, 2010a, 2010b).

창의력 점수는 유창성, 독창성, 제목의 추상성, 정교성, 성급한 종결에 대한 저항과 각 하위요소의 평균점수, 창의적 강점이 포함된 창의성 지수로 구분되며 각각 표준점수와 백분위 점수를 사용할 수 있는데 본 연구에서는 표준점수를 사용하였다.

2) 질적 연구 결과

본 연구는 최종적으로 데이터 과학 교육 프로그램을 통한 학습자의 컴퓨팅 사고력과 창의성의 향상을 목표로 질적 연구를 병행하였다.

(1) 연구 도구

본 연구의 자료 수집 및 분석은 질적 사례 연구에서 대표적으로 사용되는 방법인 내용분석을 주로 활용하였다. 내용분석에는 데이터에 기반한 문제 해결의 일반적인 단계(김진영, 2016)에 따라 연구자가 직접 구안하고 학습자들이 데이터 문제 해결 단계에 따라 작성한 데이터 과학 교육 보고서 자료를 활용하였다. 엑셀을 활용한 초등학교 대상 데이터 과학 교육 보고서의 주요 내용은 [그림 III-5]와 같다.

내가 만드는 데이터 과학 보고서				
주제				제출일
학교		번호		이름
<주제 선정 이유(동기) 및 내용>				
<단계별 데이터 과학 분석>				
<p>◆ 1단계-문제 정의 ◆</p> <p>1. 문제:</p> <p>2. 목표:</p> <p>3. 가설:</p> <p>◆ 2단계-데이터 수집 방법 ◆</p> <p>1. 수집 방법</p> <p>- 대상자:</p> <p>- 어떤 방법:</p> <p>2. 설문지 내용 (써도 좋고, 구글 설문지 그림을 붙여도 됨)</p> <p>◆ 3단계-탐색적 데이터 분석 ◆</p> <p>1. 원본데이터 살펴보고 시각화하기 (조건부 서식, 차트 그림 넣고 설명)</p> <p>2. 개별 속성 분석하기(피벗 테이블, 피벗 차트 넣고 설명)</p> <p>3. 속성간의 관계 분석하기(상관도 그림 넣고 설명, 할 수 있는 사람만 해도 됨)</p> <p>◆ 4단계-현상 일반화하기, 통계적 추론 ◆</p> <p>1. 결론:</p> <p>2. 통계적 추론:</p>				
<소감>				

[그림 III-5] 데이터 과학 교육 보고서 주요 내용

IV. 데이터 과학 교육 프로그램 개발의 실제 및 적용

1. 엑셀을 활용한 초등학생 대상 데이터 과학 교육 프로그램

1) 요구 분석

제주대학교에서 실시하는 ‘창의컴퓨터교실’에 지원한 지원자 표본(volunteer sample) 40명의 학생을 대상으로 아래와 같은 내용에 대하여 요구 분석을 실시하였다.

- 소프트웨어교육 참여 경험
- 교육용 프로그래밍 언어, 언플러그드 활동, 피지컬 컴퓨팅에 대한 관심도
- 데이터 과학 교육을 통한 능력 향상
- 데이터 과학 교육 활용 데이터
- 데이터 과학 교육 학습 방법

요구 분석 결과를 종합하여 초등학생들이 가장 선호한 팀별 프로젝트 학습 중심으로 엑셀을 활용하여 개인 데이터와 공공 데이터를 대상으로 2가지의 데이터 과학 교육 프로그램을 설계 및 개발하여 두 집단에 투입한 후 그 결과를 분석하였다.

2) 설계 및 개발

(1) 연구 가설

- ① 개인 데이터와 공공 데이터를 활용한 두 집단 모두 학습자의 컴퓨팅 사고력이 향상될 것이다.
- ② 개인 데이터와 공공 데이터를 활용한 두 집단 모두 학습자의 창의성이 향상될 것이다.
- ③ 엑셀을 활용한 데이터 과학 교육 프로그램은 학습자의 컴퓨팅 사고력을 향상시킬 수 있다.

- ④ 엑셀을 활용한 데이터 과학 교육 프로그램은 학습자의 창의성을 향상시킬 수 있다.

(2) 연구 대상

본 연구에서 적용한 프로그램의 교육적 효과를 확인하기 위해 <표 IV-1>과 같이 제주대학교에서 실시한 교육기부 프로그램의 지원자 표집에 의한 지원자 표본 40명의 학생을 무작위로 20명씩 두 집단으로 나누었다.

지원자인 초등학교 4, 5, 6학년 학생들에게 오리엔테이션, 사전·사후 컴퓨팅 사고력 검사·창의성 검사를 포함하여 총 6일 동안 42차시 수업으로 진행되었다.

개인 데이터반과 공공 데이터반으로 나누어 각각 주 강사 1인이 학습을 진행하였고 보조 강사 2인은 학습자 출석 관리 및 여러 가지 문제해결 활동에 보조 강사로서 도움을 주었다. 모든 학생들은 ‘데스크 탑 PC’와 ‘Microsoft Excel 2010’버전을 사용하였으며, 연구 대상에 대한 구체적인 내용은 <표 IV-1>과 같다.

<표 IV-1> 연구 대상

구분	개인 데이터반(명)			공공 데이터반(명)		
	남학생	여학생	소계	남학생	여학생	소계
4학년	5	3	8	6	3	9
5학년	4	2	6	3	2	5
6학년	2	4	6	3	3	6
소계	11	9	20	12	8	20

(3) 연구 설계

본 연구에서 개발한 교육 프로그램은 집중이수제 형식으로 6일간 총 42차시의 집합 강의 및 실습으로 이루어졌다. 첫 날에는 사전 검사와 오리엔테이션 중심으로, 마지막 날에는 사후 검사 및 학습자들이 만든 프로젝트를 학생들과 학부모를 대상으로 최종 발표(데이터 스토리텔링)하는 시간으로 이루어졌다.

매일 오전 9시부터 오후 15시까지 40분 단위 수업 7차시와 휴식 시간, 점심 시간으로 운영되었고, 학습의 내면화를 위하여 배운 내용을 가정에서 복습 및 연습을 할 수 있도록 매일 과제를 제시하였다.

<표 IV-2> 프로그램 일정

일차	1	2~5	6
일정	사전 검사, 오리엔테이션, 엑셀 기초	데이터 과학 교육	사후 검사 데이터 스토리텔링

(4) 교육 내용의 선정

요구 분석 결과 및 데이터 과학 교육 도구의 특징에 따라 <표 IV-3>과 같이 엑셀을 활용한 두 집단의 교육 내용을 선정하였다. 두 집단은 개인 데이터를 활용하는 개인 데이터반과 공공 데이터를 활용하는 공공 데이터반으로 구분하였다.

<표 IV-3> 교육 프로그램 주제(엑셀)

차시	학습 주제	
	개인 데이터	공공 데이터
1~7	데이터 과학의 의미를 이해하고 이를 위한 엑셀의 기본 기능 익히기	데이터 과학의 의미를 이해하고 이를 위한 엑셀의 기본 기능 익히기
8~14	스스로 데이터를 수집하여 나에게 안전사고가 자주 일어나는 이유에 알아보고, 안전한 생활을 위해 노력해야 할 점을 실천하기	데이터를 수집하고 엑셀 프로그램을 활용하여 데이터를 가공하고 문제를 해결하기
15~21	스스로 데이터를 수집하여 학원은 학업성적 향상에 도움이 되는지 알아보고, 바람직한 학습 방법에 대해 생각해보기	우리나라의 인구 변화 데이터를 보고 대응하는 방안을 마련해보기
22~28	스스로 데이터를 수집하여 독서가 창의력 향상에 도움을 주는지 알아보기	청소년들의 안전을 위한 사망 원인을 분석하고 대안 마련해보기
29~35	평소에 내 자신이나 또는 내 주변에서 고쳤으면 하는 점이나 궁금한 점을 생각하여 주제를 정하고 스스로 데이터를 수집하여 이를 해결해 보기	미래의 진로에 관한 다양한 통계를 접해보고, 나의 진로를 탐색해보기
36~42	데이터 스토리텔링하기	데이터 스토리텔링하기

(5) 검사 도구

① 양적 연구

<표 IV-4> 실험 설계

대상	사전 검사	처치	사후 검사
개인 데이터반	O ₁	X ₁	O ₂
공공 데이터반	O ₃	X ₂	O ₄

- X₁ : 개인 데이터 활용 데이터 과학 교육 프로그램
- X₂ : 공공 데이터 활용 데이터 과학 교육 프로그램
- O₁, O₂ : 사전·사후 검사(컴퓨팅 사고력, 창의성 검사)
⇒ 정규성 검정 후 대응 표본 t 검정 또는 Wilcoxon 부호 순위 검정
- O₃, O₄ : 사전·사후 검사(컴퓨팅 사고력, 창의성 검사)
⇒ 정규성 검정 후 대응 표본 t 검정 또는 Wilcoxon 부호 순위 검정
- O₁, O₃ : 사전 검사(컴퓨팅 사고력, 창의성 검사)
⇒ 각 집단별 정규성 검정 결과에 따라 동질 집단 확인을 위한 독립 표본 t 검정 또는 Mann-Whitney U 검정
- O₂, O₄ : 사후 검사(컴퓨팅 사고력, 창의성 검사)
⇒ 동질 집단 사전 검정 결과에 따른 독립 표본 t 검정 또는 Mann-Whitney U 검정

컴퓨팅 사고력과 창의성 측정은 사전-사후검사 통제집단 설계를 활용하였다. 컴퓨팅 사고력 검사 도구는 계산적 인지력 검사 A, B형을 활용하였고, 창의성 검사 도구는 TTCT 검사지 도형 A, B형을 사용하였다.

② 질적 연구

총 40명의 학습자들이 6일간 총 42차시의 집합 강의 및 실습을 통해 작성한 최종 데이터 과학 교육 보고서 각 1편씩을 대상으로 내용분석을 하였다.

3) 적용

앞서 제시한 6단계의 데이터 과학 교육 문제 해결 단계에 따른 실제 학습 과정은 다음과 같이 진행되었다. 개인 데이터반의 5일차 학습 주제인 ‘평소에 내 자신이나 또는 내 주변에서 고쳤으면 하는 점이나 궁금한 점을 생각하여 주제를 정하고 스스로 데이터를 수집하여 이를 해결해 보기’에 따라, ‘이○○’

학생이 작성한 ‘주제와 주제 선정 이유(동기) 및 내용’은 [그림 IV-1]과 같다.

주제: 사용방법에 따라 휴대폰 배터리 소모량이 다를까?
주제 선정 이유 및 내용: 나는 가끔 우리 엄마 휴대폰을 쓴다. 그런데 이상하게 엄마가 쓸 때는 배터리가 별로 안 줄고 내가 쓸 때는 콕콕 줄어든다. 왜 같은 휴대폰을 같은 시간 동안 사용하는데 내가 쓸 때만 배터리가 빨리 나가는지 알고 싶어 이 주제를 선택하게 되었다.

[그림 IV-1] 주제 및 주제 선정 이유(동기) 및 내용(이○○)

(1) 데이터 문제 정의하기(문제 탐색하기)

이 단계에서는 문제(문제점이나 궁금한 점이 무엇인가?), 목표(문제점이나 궁금한 점을 해결하여 얻고 싶은 점은 무엇인가?), 가설(문제나 궁금한 점에 대해 내가 생각한 이유는 무엇인가?)을 [그림 IV-2]와 같이 정해 보도록 하였다.

문제: 하나의 휴대폰을 같은 시간 동안 사용하는데 내가 쓸 때만 휴대폰 배터리가 빨리 줄어든다.
목표: 휴대폰 배터리가 빨리 줄어드는 이유를 알고, 배터리 소모량을 줄이는 법을 익혀 실천할 수 있다.
가설: 용량이 큰 앱을 사용할 경우 배터리 소모량이 클 것이기 때문에 용량이 작은 앱을 사용하면 배터리 소모량이 감소할 것이다.

[그림 IV-2] 데이터 문제 정의하기(이○○)

(2) 데이터 수집하기

이 단계에서는 필요한 데이터를 스스로 만들어보는 단계로 먼저, ‘체크리스트 등을 사용하여 내 자신의 누적 자료로 수집할 것인가?, 구글 설문지 등을 사용하여 내 주변에서 자료를 수집할 것인가?’ 등의 데이터 수집 방법을 정해 보도록 하였다. 다음으로는, 데이터를 수집하기 위한 항목(항목은 주제 및 가설과 관련이 있도록 정해야 한다. 상관분석의 경우 수치로 되어야만 나타낼 수 있으며, 차트의 경우에도 수치화 되어 있는 경우가 자료를 나타내기 쉽기 때문에 항목은 될 수 있으면 수치화 시킬 수 있는 자료가 나오도록 정한다.)을 정해보고, 데이터를 직접 수집해서 테이블 형태로 변환하고 자료 중 수치화 시킬 수

있는 것들을 수치화시키는 등 데이터를 준비하도록 하였다.

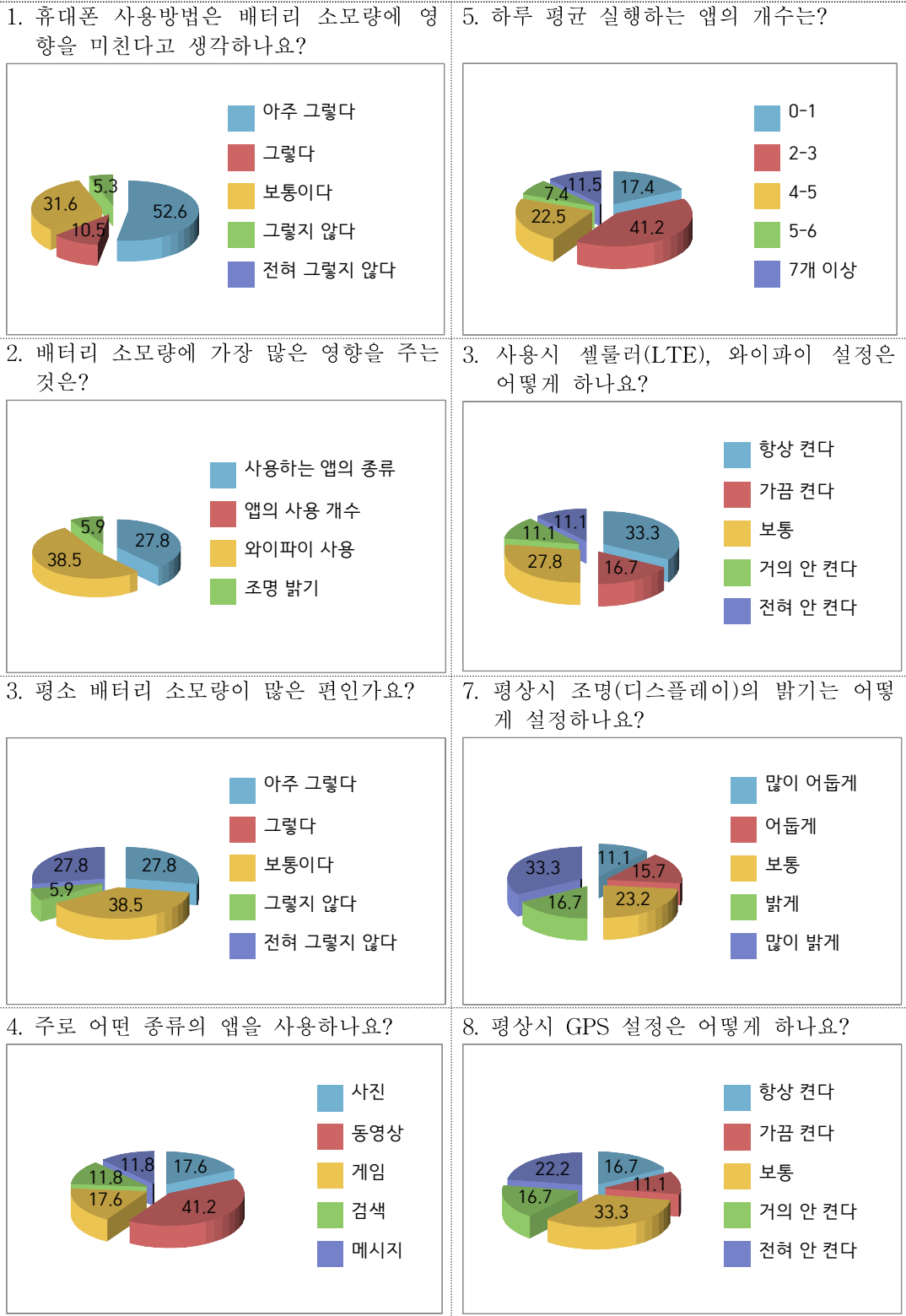
수집 방법

- 대상자: 나와 교육기부 315 강의실 4, 5, 6학년 학생들
- 어떤 방법: 구글 설문지 작성, 직접 실험하기

설문지 내용(구글 설문지 사용)

<휴대폰 배터리의 진실 문항>

1. 휴대폰 사용방법은 배터리 소모량에 영향을 미친다고 생각하나요?
① 매우 그렇다 ② 그렇다 ③ 보통이다 ④ 그렇지 않다 ⑤ 전혀 그렇지 않다
2. 배터리 소모량에 가장 많은 영향을 미치는 것은 무엇이라고 생각하십니까?
① 사용하는 앱의 종류 ② 사용하는 앱의 개수 ③ ④ 와이파이 사용
⑤ 조명(디스플레이)의 밝기
3. 평소 배터리 소모량이 많은 편인가요? (배터리가 금방 나가나요?)
① 매우 그렇다 ② 그렇다 ③ 보통이다 ④ 그렇지 않다 ⑤ 전혀 그렇지 않다
4. 주로 어떤 종류의 앱을 사용하십니까?
① 사진(갤러리) ② 동영상(유튜브) ③ 게임 ④ 검색도구(네이버 등)
⑤ 메세지(카톡)
5. 평상시 하루 평균 사용하는 앱 실행개수는 어느 정도입니까?
① 0~1개 ② 2~3개 ③ 4~5개 ④ 5~6개 ⑤ 7개 이상
6. 사용시 와이파이 설정은 어떻게 했나요?
① 항상 켜다 ② 가끔 켜다 ③ 보통(켄 때도 있고 아닐 때도) ④ 거의 안 켜다
⑤ 전혀 안 켜다.
7. 평상시 조명(디스플레이)의 밝기는 어떻게 설정하십니까?
① 많이 어둡게 ② 어둡게 ③ 보통 ④ 밝게 ⑤ 많이 밝게
8. 평상시 GPS 설정은 어떻게 하십니까?
① 항상 켜다 ② 가끔 켜다 ③ 보통(켄 때도 있고 아닐 때도) ④ 거의 안 켜다
⑤ 전혀 안 켜다.



[그림 IV-3] 데이터 수집하기(이○○)

(3) 현상 이해하기: 탐색적 데이터 분석

이 단계에서는 탐색적 데이터 분석을 해 보는 단계로, 원본 데이터 살펴보고 시각화하기, 피벗 테이블을 작성하고 관계를 중심으로 개별 속성 분석하기, 속성간의 관계 분석하기 등을 수행하도록 하였다.

1. 원본데이터 살펴보고 시각화하기 (조건부 서식, 차트)

번호	1. 사용법은 영향 미칠까?	2. 가장 많은 영향은 무엇?	3. 소모량 많다?	4. 사용앱 종류	5. 사용앱 수	6. 와이파이	7. 조명 설정	8. GPS
1번	5	사용하는 앱의 종류	1	사진	1	2	3	1
2번	5	사용하는 앱의 종류	1	게임	7	2	1	2
3번	2	조명(디스플레이)의 밝기	1	사진	1	1	2	2
5번	3	앱의 사용 개수	3	게임	3	5	3	3
6번	5	앱의 사용 개수	3	동영상	4	3	5	1
7번	5	앱의 사용 개수	5	동영상	6	3	5	2
8번	3	사용하는 앱의 종류	5	동영상	3	3	3	3
9번	5	앱의 사용 개수	5	동영상	5	4	5	3
10번	3	조명(디스플레이)의 밝기	1	메세지	3	5	4	3
11번	5	조명(디스플레이)의 밝기	5	동영상	6	5	5	5
12번	3	와이파이 사용	3	검색도구	3	3	3	2
13번	5	사용하는 앱의 종류	3	동영상	5	4	4	4
14번	5	사용하는 앱의 종류	1	검색도구	1	1	1	1
15번	5	사용하는 앱의 종류	3	사진	3	5	3	1
16번	2	사용하는 앱의 종류	3	사진	3	4	5	3
17번	3	조명(디스플레이)의 밝기	2	메세지	3	3	3	3
18번	5	앱의 사용 개수	3	동영상	5	5	4	5
19번	4	앱의 사용 개수	5	게임	3	5	5	5

- 거의 대부분 학생들이 휴대폰 사용법이 배터리 소모량에 영향을 준다고 답하였다.
- 가장 많은 영향을 주는 것은 '사용하는 앱의 종류'가 가장 많았고, 그 뒤로 '앱의 사용 개수', 그리고 '조명'이 그 뒤를 이었다.
- 반 이상의 학생이 소모량이 중간보다 많다고 했다.
- '사용 앱 종류'는 동영상이 제일 많았다.
- '사용 앱 개수'는 대부분 학생들이 3개 이상을 쓰고 있었다.
- 와이파이는 대부분 학생들이 사용하고 있었다.
- 조명은 대부분 학생들이 중간 이상의 밝기를 사용하고 있다.
- GPS는 비슷하긴 하지만 끄는 학생보다 켜는 학생이 많았다.

[그림 IV-4] 현상 이해하기: 탐색적 데이터 분석(이○○)

(4) 현상 일반화하기

이 단계에서는 분석한 자료를 바탕으로 내가 세운 가설과 비교하여 결론을 내리고 통계적 추론을 해 보도록 하였다.

결론: 휴대폰 사용방법에 따라 휴대폰 배터리 소모량은 달라진다. 배터리 소모량에 가장 많은 영향을 주는 것은 사용하는 앱의 종류이다. 동영상을 썼을 때 배터리 소모량이 제일 많았다. 그 외 조명, GPS 등도 영향을 준다.

통계적 추론: 휴대폰 배터리 소모량을 줄이려면 동영상을 덜 봐야 한다. 휴대폰 배터리 소모량을 줄이려면 실행하는 앱의 수는 적은 것이 좋다. 또 GPS는 꺼두고 조명은 보통 정도로 해야 한다.

[그림 IV-5] 현상 일반화하기(이○○)

(5) 현상 예측하기

이 단계에서는 결론을 실험 연구를 통해 확인하고 인과 관계의 통계적 오류를 방지하기 위해 확인 실험 등을 통해 상관관계를 확인해 보며 앞으로의 예측을 해보도록 하였다.

확인 실험: 결론에서 용량이 큰 앱을 사용하는 경우 배터리가 빨리 나간다는 내용이 있어, 이것을 실험을 통해 다시 내 눈으로 확인하고 싶어졌다. 그래서 직접 실험을 해 보았다.

앱 \ 시간	사용 전	5분 후	앱 \ 시간	사용전	10분	13분
유튜브	33%	32%	유튜브	33%	31%	30%
갤러리	76%	75%	갤러리	69%	67%	66%
게임	76%	74%	게임	75%	72%	71%
네이버	40%	39%	네이버	44%	41%	40%
메세지	76%	76%	메세지	66%	61%	61%
그냥 켜놓기	76%	76%	그냥 켜놓기	71%	70%	70%

사용 핸드폰= 스마트폰

와이파이사용=네이버, 용량=유튜브, 사용자 수=게임 갤러리, 메시지, 그냥 켜놓기=기타
5분은 차이가 별로 없어 10분, 13분으로 했다. 대부분 게임, 네이버, 유튜브에서 많이 줄었다. 이처럼 화면이 계속 움직이는 앱이 대부분 소모량이 컸다.

[그림 IV-6] 현상 예측하기(이○○)

(6) 데이터 스토리텔링

최종 발표(데이터 스토리텔링)는 각자 제작한 최종 프로젝트 보고서를 앞에 나와서 학생들과 학부모님들 앞에서 발표하는 시간으로 운영하였다.



[그림 IV-7] 데이터 스토리텔링 장면

4) 양적 연구 결과 분석

(1) 컴퓨팅 사고력 변화

① 정규성 검정

두 집단이 정규성을 확보하고 있는지 확인하기 위하여 두 집단의 컴퓨팅 사고력 사전 검사 결과로 각각 정규성 검정을 실시하였다. 사전 검사 데이터를 바탕으로 후속 통계의 비모수/모수 통계 여부를 결정하기 위해 정규성 검정 방법으로 Shapiro-Wilks 검정을 실시하였고 <표 IV-5>에 그 결과를 제시하였다.

<표 IV-5> 컴퓨팅 사고력 검사 정규성 검정

구분	기술통계(N=40)				t	p
	M	SD	Max	Min		
개인 데이터반	12.05	3.734	18	11	.924	.119
공공 데이터반	10.72	5.156	19	2	.956	.410

*p<.05

두 집단에 대해 각각 Shapiro-Wilks 정규성 검정을 실시한 결과, 개인 데이터반은 유의도가 .119로 유의수준 .05보다 높게 나타나 정규분포임이 확인되었다. 공공 데이터반의 경우에도 유의도가 .410으로 유의수준 .05보다 높게 나타나 정규분포임이 확인되었다.

② 사전 검사의 집단 간 비교

본 연구의 교육 프로그램이 학습자의 컴퓨팅 사고력에 미친 영향을 살펴보기 위해 컴퓨팅 사고력 검사에 의해 측정된 총점인 컴퓨팅 사고력 지수를 이용하여 비교하였다. 두 집단 모두 정규성이 검증되었으므로 <표 IV-6>과 같이 모수통계인 독립표본 t 검정을 실시하였다.

<표 IV-6> 컴퓨팅 사고력 사전 검사 결과 비교(독립표본 t 검정)

대상	N	M	SD	t	p
개인 데이터반	20	12.05	3.734	1.139	.262
공공 데이터반	20	10.40	5.295		

*p<.05

<표 IV-6>의 독립표본 t 검정의 결과를 살펴보면, 사전 검사의 개인 데이터반 평균은 12.05, 표준편차는 3.734이고 공공 데이터반의 평균은 10.40, 표준편차는 5.295이다. 독립표본 t 검정 결과의 t 통계값은 1.139이고 유의확률은 .262로 유의수준 .05에서 두 집단은 통계적으로 유의미한 차이가 없고 동질집단임이 확인되었다.

③ 사후 검사의 집단 간 비교

본 연구의 교육 프로그램을 활용하여 학습을 진행한 후 학습자의 컴퓨팅 사고력에 미친 영향을 살펴보기 위해 사후 검사를 실시하고 <표 IV-7>과 같이 모수통계인 독립표본 t 검정을 실시하였다.

<표 IV-7> 컴퓨팅 사고력 사후 검사 결과 비교(독립표본 t 검정)

대상	N	M	SD	t	p
개인 데이터반	20	13.15	2.906	1.410	.167
공공 데이터반	20	11.35	4.912		

*p<.05

<표 IV-7>의 독립표본 t 검정의 결과를 살펴보면, 사후 검사의 개인 데이터반 평균은 13.15, 표준편차는 2.906이고 공공 데이터반의 평균은 11.35, 표준편차는 4.912이다. 또한, t 통계값은 1.410이고 유의확률은 .167로 유의수준 .05에서 <표 IV-6>의 사전 검사보다 두 집단 모두 컴퓨팅 사고력이 향상되었지만 특정 집단이 통계적으로 유의미하게 더 향상되지는 않았다.

④ 사전·사후 검사 결과 집단 내 비교

개인 데이터반과 공공 데이터반의 사전·사후 검사 결과 컴퓨팅 사고력의 변화를 알아보기 위하여 <표 IV-8>과 같이 모수통계인 대응표본 t 검정을 실시하였다.

<표 IV-8> 사전·사후 검사 결과 집단 내 비교(대응표본 t 검정)

대상	N	사전 검사		사후 검사		t	p
		M	SD	M	SD		
개인 데이터반	20	12.05	3.734	13.15	2.906	-2.685	.015*
공공 데이터반	20	10.40	5.295	11.35	4.912	-2.263	.036*

*p<.05

<표 IV-8>의 대응표본 t 검정의 결과를 살펴보면, 개인 데이터반의 t 통계값은 -2.685이고 유의확률은 .015로 유의수준 .05에서 컴퓨팅 사고력이 유의미하게 향상된 것으로 나타났다. 또한, 공공 데이터반의 t 통계값은 -2.263이고 유의확률은 .036으로 유의수준 .05에서 개인 데이터반과 마찬가지로 컴퓨팅 사고력이 유의미하게 향상된 것으로 나타났다.

(2) 창의성 변화

① 정규성 검정

개인 데이터반과 공공 데이터반의 창의성 사전 검사 결과를 이용하여 각 집단이 정규성을 확보하고 있는지 확인하기 위하여 정규성 검정을 실시하였다.

사전 검사 결과를 바탕으로 후속 통계의 비모수/모수 통계 여부를 선정하기 위해 정규성 검정 방법으로 Shapiro-Wilks 검정을 실시하였고 <표 IV-9>, <표 IV-10>에 그 결과를 제시하였다.

<표 IV-9> 개인 데이터반 창의성 검사 정규성 검정

하위 요소	기술통계(N=20)				t	p
	M	SD	Max	Min		
유창성	128.95	21.813	149	81	.826	.002**
독창성	123.90	23.554	150	85	.876	.015*
제목의 추상성	80.55	40.654	150	0	.962	.578
정교성	102.00	33.688	150	57	.873	.013*
성급한 종결에 대한 저항	79.80	26.078	125	40	.951	.384
창의성 평균	102.68	20.227	139.6	65	.973	.819
창의성 지수	104.75	21.586	146.6	66	.965	.658

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

개인 데이터반의 창의성 사전 검사에 대해 Shapiro-Wilks 정규성 검정을 실시한 결과, ‘유창성’과 ‘독창성’, ‘정교성’ 영역의 유의도가 각각 .002, .015, .013으로 정규성이 만족되지 않는 것으로 나타났다. 나머지 영역은 유의도가 유의수준 .05보다 높게 나타나 정규분포임이 확인되었다.

<표 IV-10> 공공 데이터반 창의성 검사 정규성 검정

하위 요소	기술통계(N=20)				t	p
	M	SD	Max	Min		
유창성	111.80	24.673	149	68	.939	.234
독창성	117.60	19.701	150	75	.974	.841
제목의 추상성	97.00	28.233	150	40	.973	.808
정교성	104.35	31.823	152	57	.887	.024*
성급한 종결에 대한 저항	74.20	40.107	142	0	.961	.574
창의성 평균	99.88	18.089	130.80	67.60	.967	.690
창의성 지수	102.31	19.682	137.80	67.60	.969	.744

*p<.05

공공 데이터반의 창의성 사전 검사에 대해 Shapiro-Wilks 정규성 검정을 실시한 결과, ‘정교성’ 영역의 유의도가 .024로 정규성이 만족되지 않는 것으로 나타났다. 나머지 영역은 유의도가 유의수준 .05보다 높게 나타나 정규분포임이 확인되었다. 개인 데이터반과 공공 데이터반의 정규성 확보 여부에 따라, 이후의 통계는 모수적 통계방법과 비모수적 통계방법을 사용하였다.

② 사전 검사의 집단 간 비교

본 연구를 통해 개발된 교육 프로그램이 학습자의 창의성에 미친 영향을 살펴보기 위해 개인 데이터반과 공공 데이터반의 창의성 사전 검사 결과를 비교하였다. 먼저, 두 집단의 동질집단 여부를 알아보기 위하여 사전 검사 결과를 비교한 후, 정규성이 확보된 영역은 독립표본 t 검정을 실시하여 <표 IV-11>에 제시하였고, 정규성이 확보되지 못한 ‘유창성’과 ‘독창성’, ‘정교성’은 Mann-Whitney U검정을 실시하여 <표 IV-12>에 제시하였다.

<표 IV-11> 창의성 사전 검사 결과(독립표본 t 검정)

하위요소	대상	N	M	SD	t	p
제목의 추상성	개인 데이터반	20	80.55	40.654	-1.486	.145
	공공 데이터반	20	97.00	28.233		
성급한 종결에 대한 저항	개인 데이터반	20	79.80	26.078	.523	.604
	공공 데이터반	20	74.20	40.107		
창의성 평균	개인 데이터반	20	102.68	20.227	.516	.609
	공공 데이터반	20	99.55	18.089		
창의성 지수	개인 데이터반	20	104.75	21.586	.374	.711
	공공 데이터반	20	102.31	19.682		

*p<.05

<표 IV-12> 창의성 사전 검사 결과(Mann-Whitney U 검정)

하위요소	대상	N	M	SD	평균순위	Mann-Whitney의 U	p
유창성	개인 데이터반	20	128.95	21.813	24.25	125.000	.043*
	공공 데이터반	20	111.80	24.673	16.75		
독창성	개인 데이터반	20	123.90	23.554	22.45	161.000	.301
	공공 데이터반	20	117.60	19.701	18.55		
정교성	개인 데이터반	20	102.00	33.688	19.83	186.500	.718
	공공 데이터반	20	104.35	31.823	21.18		

*p<.05

개인 데이터반과 공공 데이터반의 창의성 사전 검사 결과를 비교한 결과 ‘유창성’을 제외하고 모든 하위 요소가 동질집단으로 확인되었다.

③ 사후 검사 결과 집단 간 비교

본 연구의 교육 프로그램을 활용하여 학습을 진행한 후 학습자의 창의성에 미친 영향을 살펴보기 위해 사후 검사를 실시였다. 사후 검사를 비교하기 위해서 ‘유창성’을 제외하고 동질집단이 확인된 모든 하위 요소에 대하여 모수통계인 독립표본 t 검정을 이용하여 비교한 후 그 결과를 <표 IV-13>에 제시하였다.

<표 IV-13> 창의성 사후 검사 결과(독립표본 t 검정)

하위요소	대상	N	M	SD	t	p
독창성	개인 데이터반	20	130.35	21.069	1.579	.123
	공공 데이터반	20	119.80	21.179		
제목의 추상성	개인 데이터반	20	72.75	30.480	-1.661	.105
	공공 데이터반	20	91.95	41.771		
정교성	개인 데이터반	20	100.45	25.599	-.210	.835
	공공 데이터반	20	102.20	27.101		
성급한 종결에 대한 저항	개인 데이터반	20	92.25	32.314	.259	.797
	공공 데이터반	20	89.35	38.276		
창의성 평균	개인 데이터반	20	107.50	18.317	.525	.602
	공공 데이터반	20	104.34	19.700		
창의성 지수	개인 데이터반	20	111.85	20.306	.368	.715
	공공 데이터반	20	109.39	21.976		

*p<.05

<표 IV-13>의 독립표본 t 검정의 결과를 살펴보면, ‘독창성’의 t 통계값은 1.579이고 유의확률은 .123, ‘제목의 추상성’의 t 통계값은 -1.661이고 유의확률은 .105, ‘정교성’의 t 통계값은 -.210이고 유의확률은 .835, ‘성급한 종결에 대한

저항'의 t 통계값은 .259이고 유의확률은 .797, '창의성 평균'의 t 통계값은 .525이고 유의확률은 .602, '창의성 지수'의 t 통계값은 .368이고 유의확률은 .715로 두 집단은 유의수준 .05에서 특정 집단의 창의성 하위요소가 통계적으로 유의미하게 더 향상되지는 않았다.

④ 사전·사후 검사 결과 집단 내 비교

• 개인 데이터반

개인 데이터반과 공공 데이터반에서 실시한 사전·사후 검사 결과 창의성의 변화를 알아보았다. 개인 데이터반의 경우 정규성을 확보하지 못한 '유창성', '독창성', '정교성'을 제외한 나머지 하위 요소에 대하여 <표 IV-14>와 같이 모수통계인 대응표본 t 검정을 실시하였다. 정규성을 확보하지 못한 '유창성', '독창성', '정교성'에 대해서는 비모수 통계로 Wilcoxon 부호 순위 검정을 실시한 후 그 결과를 <표 IV-15>에 제시하였다.

<표 IV-14> 개인 데이터반 사전·사후 검사 결과 집단 내 비교(대응표본 t 검정)

하위요소	대상	N	사전 검사		사후 검사		t	p
			M	SD	M	SD		
제목의 추상성	개인 데이터반	20	80.55	40.654	72.75	30.480	1.182	.252
성급한 종결에 대한 저항	개인 데이터반	20	79.80	26.078	92.25	32.314	-2.701	.014*
창의성 평균	개인 데이터반	20	102.68	20.227	107.50	18.317	-2.712	.014*
창의성 지수	개인 데이터반	20	104.75	21.586	111.85	20.306	-3.699	.002**

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

<표 IV-15> 개인 데이터반 사전·사후 검사 결과 집단 내 비교(Wilcoxon 부호 순위 검정)

하위요소	대상	N	사전 검사		사후 검사		z	p
			M	SD	M	SD		
유창성	개인 데이터반	20	128.95	21.813	133.95	18.391	-1.774	.076
독창성	개인 데이터반	20	123.90	23.554	130.35	21.069	-1.570	.116
정교성	개인 데이터반	20	102.00	33.688	100.45	25.599	-.240	.810

*p<.05

<표 IV-14>과 <표 IV-15>의 대응표본 t 검정, Wilcoxon 부호 순위 검정의 결과를 살펴보면, ‘성급한 종결에 대한 저항’의 t 통계값은 -2.701이고 유의확률은 .014, ‘창의성 평균’의 t 통계값은 -2.712이고 유의확률은 .014로 나타나 각각 유의수준 .05에서 사전 검사 점수와 비교했을 때 사후 검사 점수가 유의미하게 상승한 것으로 나타나 창의성의 하위요소가 향상된 것을 알 수 있다. 또한, ‘창의성 지수’의 t 통계값은 -3.699이고 유의확률은 .002로 나타나, 유의수준 .01에서 사전 검사 점수와 비교했을 때 사후 검사 점수가 유의미하게 상승하여 창의성의 하위요소가 향상되었음을 알 수 있다.

- 공공 데이터반

또한, 공공 데이터반의 경우 정규성을 확보하지 못한 ‘정교성’을 제외한 나머지 하위 요소에 대하여 <표 IV-16>과 같이 모수통계인 대응표본 t 검정을 실시하였다. 정규성을 확보하지 못한 ‘정교성’에 대해서는 비모수 통계로 Wilcoxon 부호 순위 검정을 실시한 후 그 결과를 <표 IV-17>에 제시하였다.

<표 IV-16> 공공 데이터반 사전·사후 검사 결과 집단 내 비교(대응표본 t 검정)

하위요소	대상	N	사전 검사		사후 검사		t	p
			M	SD	M	SD		
유창성	공공 데이터반	20	111.80	24.673	118.40	22.684	-1.497	.151
독창성	공공 데이터반	20	117.60	19.701	119.80	21.179	-.526	.605
제목의 추상성	공공 데이터반	20	97.00	28.233	91.95	41.771	1.003	.329
성급한 종결에 대한 저항	공공 데이터반	20	74.20	40.107	89.35	38.276	-2.636	.016*
창의성 평균	공공 데이터반	20	99.55	18.089	89.35	38.276	1.509	.148
창의성 지수	공공 데이터반	20	102.31	19.682	109.39	21.986	-3.772	.001**

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

<표 IV-17> 공공 데이터반 사전·사후 검사 결과 집단 내 비교(Wilcoxon 부호 순위 검정)

하위요소	대상	N	사전 검사		사후 검사		z	p
			M	SD	M	SD		
정교성	공공 데이터반	20	104.35	31.823	102.20	27.101	-.610	.542

*p<.05

<표 IV-16>과 <표 IV-17>의 대응표본 t 검정, Wilcoxon 부호 순위 검정의 결과를 살펴보면, ‘성급한 종결에 대한 저항’의 t 통계값은 -2.636이고 유의확률은 .016, ‘창의성 지수’의 t 통계값은 -3.772이고 유의확률은 .001로 나타나, 각각 유의수준 .05와 .01에서 사전 검사 점수와 비교했을 때 사후 검사 점수가 유의미하게 상승한 것으로 나타나 창의성의 하위요소가 향상된 것을 알 수 있다.

(3) 컴퓨팅 사고력과 창의성 검사 결과의 분석

컴퓨팅 사고력 사전 검사 결과 개인 데이터반과 공공 데이터반은 정규분포를 갖춘 표본임이 검증되었으며, 컴퓨팅 사고력 사전 검사에 대한 집단 간 독립표본 t 검정 결과 두 집단은 동질집단인 것으로 나타났다.

창의성 사전 검사 결과 개인 데이터반과 공공 데이터반은 ‘유창성’, ‘독창성’,

‘정교성’을 제외한 창의성의 하위요소에서 정규분포를 갖춘 표본임이 검증되었으며, 창의성 사전 검사에 대한 집단 간 독립표본 t 검정, Mann-Whitney U검정에서 ‘유창성’을 제외하고 모든 하위 요소가 동질집단임이 확인되었다.

본 연구에서 개발한 데이터 과학 교육 프로그램을 적용한 후 실시한 컴퓨팅 사고력 사후 검사의 집단 간 독립표본 t 검정과 집단 내 대응표본 t 검정의 결과, 두 집단 간에 컴퓨팅 사고력의 향상도는 유의미한 차이가 없이 두 집단 모두 사전 검사보다 컴퓨팅 사고력이 향상된 것으로 나타났다.

또한, 창의성 사후 검사의 집단 간 독립표본 t 검정, Mann-Whitney U검정과 집단 내 대응표본 t 검정, Wilcoxon 부호 순위 검정의 결과, 두 집단 간에 창의성의 향상도는 유의미한 차이가 없이 두 집단 모두 사전 검사보다 창의성이 향상된 것으로 나타났다.

이로써 데이터 과학 교육에서 개인 데이터와 공공 데이터는 모두 데이터 문제 해결에 활용할 수 있으며, 본 연구를 통해 개발된 개인 데이터와 공공 데이터를 활용한 2가지의 데이터 과학 교육 프로그램은 모두 학습자의 컴퓨팅 사고력과 창의성을 향상시킬 수 있음을 입증하였다. 또한, 본 연구를 통해 개발된 엑셀을 활용한 데이터 과학 교육 프로그램은 학습자의 컴퓨팅 사고력과 창의성을 향상시킬 수 있음을 입증하였다.

이는 두 집단에서 각각 사용한 개인 데이터와 공공 데이터가 모두 데이터 과학 교육의 문제 해결에 적합하였기 때문인 것으로 생각된다. 또한, 학생들이 스스로 흥미와 관심을 갖고 있는 주제를 선정하고 구글 설문지나 통계청 누리집 등을 활용하여 개인 데이터 및 공공 데이터를 수집하여 분석하고 일반화하여 통계적 추론을 해보는 학습활동을 꾸준히 진행했다. 이처럼 스스로 자료를 수집, 조직, 표현하고 해석하는 직접적인 활동 중심 데이터 문제 해결 학습 과정에서 창의성의 하위 요소들이 향상된 것으로 분석되며, 선행 연구에서 엑셀을 활용하였을 때 올바른 판단력과 분석력이 길러지고(김은정, 2004) 문제해결력이 향상되며(김윤석, 2005) 귀납적인 관찰을 통해 일반적인 규칙을 발견하려는 성향을 보인다(이광상, 2005)는 연구 결과와도 일치한다.

5) 질적 연구 결과 분석

본 연구의 자료 수집 및 분석은 질적 사례 연구에서 대표적으로 사용되는 방법인 내용분석을 주로 활용하였는데, 내용분석에는 총 40명의 학습자들이 6일간 총 42차시의 집합 강의 및 실습을 통해 데이터 문제 해결 단계에 따라 직접 작성한 40편의 데이터 과학 교육 보고서 자료를 활용하였다.

(1) 주제 선정 이유(동기) 및 내용

대부분의 학생들이 평소에 내 자신이나 주변에서 궁금했던 점을 개인 데이터와 공공 데이터를 활용하여 해결해 보기 위해 다양한 주제를 선정했다. 예를 들면, ‘TV를 많이 보면 성적이 떨어질까?’, ‘나는 왜 학원에서 푼 문제보다 집에서 푼 문제를 더 많이 틀릴까?’, ‘사용방법에 따라 휴대폰 배터리 소모량이 다를까?’, ‘선생님을 좋아하면 성적이 오를까?’, ‘선생님이 무서우면 숙제를 잘 해올까?’, ‘신발 가격은 달리기 실력과 관계가 있을까?’, ‘왜 공부 할 때는 집중이 안 되다가 독서를 할 때는 집중이 되는 것인가?’, ‘내 키는 왜 안 클까?’처럼 교육 프로그램 적용 초기에 비해 최종 보고서에서는 정말 초등학생이 아니면 생각해낼 수 없는 창의적이고 기발한 주제들이 많이 나와서 창의성 향상에 따른 긍정적 효과라고 생각되었다.

(2) 단계별 데이터 과학 분석

문제점이나 궁금한 점을 해결하여 얻고 싶은 점과 문제점이나 궁금한 점에 대해 학습자들이 생각한 이유에 대해 학습자들 대부분이 나름대로 논리적으로 가설 설정을 하였고, 직접 작성한 구글 설문지 등을 사용하여 다양한 대상을 상대로 자료를 수집하였다.

또한, 엑셀의 특징을 잘 활용하여 조건부 서식, 차트 등을 원본 데이터의 특징에 맞게 시각화 자료로 활용하여 한 눈에 이해하기 쉽게 제시하였고 피벗 테이블을 활용하여 개별 속성을 잘 분석하였다.

3. 속성간의 관계 분석하기(상관도분석)

	이번 달 내 키는 얼마나 컸나요?	충분한 수면	물은 하루에 1.5L 이상	인스턴트 음식을 하루에 1번 이상	하루 30분 이상 스트레칭	하루 30분 이상 운동
이번 달 내 키는 얼마나 컸나요?						
충분한 수면을 취했나요?	0					
물은 하루에 1.5L 이상 마셨나요?	-0.5	0.866025404				
인스턴트 음식을 하루에 1번 이상 먹었나요?	-0.944911183	-0.327326835	0.188982237			
하루 30분 이상 스트레칭을 했나요?	-0.5	-0.866025404	-0.5	0.755928946		
하루 30분 이상 꾸준히 했나요?	0	1	0.866025404	-0.327326835	-0.866025404	1

위의 표를 보면 키와 관련해서 수면, 물, 인스턴트 음식, 스트레칭, 운동, 이런 것들이 그렇게 크게 관련은 없다고 나왔으나 인스턴트 음식의 경우 음의 관계가 나와서 많이 먹을수록 키에는 좋지 않다는 것을 알 수 있다. **운동과는 그다지 상관성이 없는 것으로 나왔다.** 물과 수면은 상관관계가 높게 나온 것으로 보아 물을 하루에 1.5L를 먹었을 때 충분한 수면도 취한 편이었고 스트레칭을 한 날은 운동을 좀 게을리 한 것으로 보인다.

◆ 4단계- 현상 일반화하기, 통계적 추론 ◆

1. 결론: 키가 안 크는 이유 중 인스턴트 식품이 제일 많은 영향을 주었다. **또한 운동이나 스트레칭은 키 성장에 크게 관련이 없었다.**
2. 통계적 추론: 인스턴트 식품을 계속 먹으면 키가 클 확률도 그만큼 작아질 것이다. 키가 커지려면 인스턴트 식품 섭취를 줄여야 한다.

[그림 IV-8] 속성 간의 관계 분석(상관관계 분석) 후 현상 일반화하기(김○○)

그러나 속성 간의 관계 분석하기에서는 [그림 IV-8]처럼 상관관계 분석이 제대로 이루어지지 않은 학습자들이 대부분이었다. 이는 초등학생들의 특성상 설문 대상이 가족·학교 친구 등에 국한되어 수집한 설문 표본 수가 작고 현상 일반화와 현상 예측을 하기에는 데이터 수집 기간이 짧아서 양질의 데이터를 수집하지 못한 데서 기인한다고 할 수 있다. 또한, 평소 상관관계보다는 인과 관계 위주로 문제 해결을 했던 경험 때문에 상관관계를 원인과 결과의 필연적인 관계인 인과 관계로 생각하려는 경향이 있기 때문인 것으로 생각된다.

이로 인해, [그림 IV-8]의 <속성 간의 관계 분석하기>에서 ‘운동과는 그다지 상관이 없는 것으로 나왔다.’ 또는, <현상 일반화하기>에서 ‘운동이나 스트레칭은 키 성장에 크게 관련이 없었다.’처럼 검증되지 않은 결론을 내려버리는 경우가 있었다. 다만, 본 연구는 데이터 과학 교육 프로그램 개발 원리에 따라 초등학생을 대상으로 한 교육 프로그램이기 때문에 맞고 틀리고 보다는 스스로 데이터를 수집하고 문제를 해결해 나가는 과정에 더 큰 의미를 부여하였다.

2. 애플벤터를 활용한 초등학생 대상 데이터 과학 교육 프로그램

1) 요구 분석

제주대학교에서 실시하는 ‘창의컴퓨터교실’에 지원한 지원자 표본 26명의 학생을 대상으로 아래와 같은 내용에 대하여 요구 분석을 실시하였다.

- 소프트웨어교육 참여 경험
- 교육용 프로그래밍 언어, 언플러그드 활동, 피지컬컴퓨팅에 대한 관심도
- 데이터 과학 교육을 통한 능력 향상
- 데이터 과학 교육 활용 데이터
- 데이터 과학 교육 학습 방법

요구 분석 결과를 종합하여 초등학생들이 가장 선호한 팀별 프로젝트 학습 중심으로 애플벤터를 활용한 데이터 과학 교육 프로그램을 설계 및 개발하여 적용한 후 그 결과를 분석하였다.

2) 설계 및 개발

(1) 연구 가설

- ① 애플벤터를 활용한 데이터 과학 교육 프로그램은 학습자의 컴퓨팅 사고력을 향상시킬 수 있다.
- ② 애플벤터를 활용한 데이터 과학 교육 프로그램은 학습자의 창의성을 향상시킬 수 있다.

(2) 연구 대상

본 연구에서 적용한 프로그램의 교육적 효과를 살펴보기 위해 <표 IV-5>와 같이 제주대학교에서 실시한 교육기부 프로그램의 지원자 표집에 의한 지원자 표본 26명의 학생을 대상으로 오리엔테이션, 사전·사후 컴퓨팅 사고력 검사, 창의성 검사를 포함하여 총 6일 동안 42차시 수업으로 진행되었다.

강사 1인은 전체 학습을 진행하였고 보조 강사 2인은 학습자 출석 관리 및 여러 가지 문제해결 활동에 보조 강사로서 도움을 주었다. 모든 학생들은 ‘데스크 탑 PC’와 ‘MIT App Inventor 2 Beta 버전’, ‘스마트폰(안드로이드)’을 사용하였다. 연구 대상에 대한 구체적인 내용은 <표 IV-18>에 제시하였다.

<표 IV-18> 연구 대상

구분	학생수(명)		
	남	여	소계
4학년	5	5	10
5학년	7	3	10
6학년	2	4	6
소계	14	12	26

(3) 연구 설계

본 연구에서 개발한 교육 프로그램은 집중이수제 형식으로 6일간 총 42차시의 집합 강의 및 실습으로 이루어졌다. 첫 날에는 사전 검사 및 오리엔테이션을 실시하고, 마지막 날에는 사후 검사 및 최종 프로젝트 결과물을 학생들과 학부모를 대상으로 발표하는 시간으로 이루어졌다. 교육 프로그램은 매일 오전 9시부터 오후 15시까지 40분 단위 수업 7차시와 휴식 시간, 점심 시간으로 운영되었고, 학습의 내면화를 위하여 매일 배운 내용을 가정에서 복습 및 연습을 할 수 있도록 과제를 제시하였다.

<표 IV-19> 프로그램 일정

일차	1	2~5	6
일정	사전 검사, 오리엔테이션, 앱인벤터 기초	데이터 과학 교육	사후 검사 데이터 스토리텔링

(4) 교육 내용의 선정

요구 분석 결과 및 데이터 과학 교육 도구의 특징에 따라 <표 IV-20>과 같이 앱인벤터를 활용한 교육 내용을 선정하였다.

<표 IV-20> 교육 프로그램 주제(앱인벤터)

차시	학습 주제
1~7	데이터 과학의 의미를 이해하고 이를 위한 앱인벤터의 기본 기능 익히기
8~14	자동으로 데이터를 수집하여 수학여행 계획을 세우기
15~21	진동을 감지하여 스스로 데이터를 수집하는 앱을 만들고 지진에 대처하는 생활을 실천하기
22~28	스스로 데이터를 수집하여 습관이 건강에 미치는 영향을 알아보고 바람직한 생활습관에 대해 생각해 보기
29~35	교육기부 참가 학생들을 대상으로 평소에 궁금했던 점을 생각하여 주제를 정하고 스스로 데이터를 수집하여 이를 해결하기
36~42	최종보고서에 대한 피드백을 주고받은 후 수정, 보충을 통해 최종발표(데이터 스토리텔링)하기

(5) 검사 도구

① 양적 연구

<표 IV-21> 실험 설계

대상	사전 검사	처치	사후 검사
앱인벤터반	O ₁	X ₁	O ₂

X₁ : 앱인벤터를 활용한 데이터 과학 교육 프로그램

O₁ : 사전검사(컴퓨팅 사고력, 창의성 검사)

⇒ 정규성 검정

O₁, O₂ : 사전·사후 검사(컴퓨팅 사고력, 창의성 검사)

⇒ 정규성 검정 결과에 따른 대응 표본 t 검정 또는 Wilcoxon 부호 순위 검정

컴퓨팅 사고력과 창의성 측정은 사전-사후검사 통제집단 설계를 사용하였다. 컴퓨팅 사고력 검사 도구는 계산적 인지력 검사 A, B형(김병수, 2014)을 활용하였고, 창의성 검사 도구는 TTCT 검사지 도형 A, B형을 사용하였다.

② 질적 연구

총 26명의 학습자들이 6일간 총 42차시의 집합 강의 및 실습을 통해 작성한 최종 데이터 과학 교육 보고서 각 1편씩을 대상으로 내용분석을 하였다.

3) 적용

앞서 제시한 6단계의 데이터 과학 교육 문제 해결 단계에 따른 실제 학습 과정은 다음과 같이 진행되었다. 5일차 학습 주제인 ‘생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 보기’에 따라, ‘석○○’ 학생이 작성한 ‘주제와 주제 선정 이유(동기) 및 내용’은 [그림 IV-9]와 같다.

주제: 초등학생들의 하루 수면시간은 몇 시간일까?

주제 선정 이유 및 내용: 요즘 초등학생들이 잠을 많이 자지 못해 학교에서 하루 종일 졸고 있는 모습을 보았다. 그래서 나는 학생들이 평소에 얼마나 잠을 자는지와 자기 전에 무엇을 하는지, 그리고 수면시간이 건강에 영향을 미친다는 것을 학생들이 얼마나 알고 있는지 궁금하여 이 주제로 선정하였다.

[그림 IV-9] 주제 선정 이유(동기) 및 내용(석○○)

(1) 데이터 문제 정의하기(문제 탐색하기)

이 단계에서는 문제(문제점이나 궁금한 점이 무엇인가?), 목표(문제점이나 궁금한 점을 해결하여 얻고 싶은 점은 무엇인가?), 가설(문제나 궁금한 점에 대해 내가 생각한 이유는 무엇인가?)을 [그림 IV-10]과 같이 정해 보도록 하였다.

문제: 초등학생들이 수면시간이 부족해서 하루 종일 졸고 있다.

목표: 초등학생들의 수면 시간이 건강에 어떤 영향을 끼치는지 알아보자.

가설: 대부분의 초등학생들은 10시간 이상 자지 못하고 수면 시간이 건강에 어떤 영향을 끼치는지 자세히 모를 것이다.

[그림 IV-10] 데이터 문제 정의하기(석○○)

(2) 데이터 수집하기

이 단계에서는 필요한 데이터를 스스로 만들어보는 단계로 먼저, ‘체크리스트 등을 사용하여 내 자신의 누적 자료로 수집할 것인가?, 구글 설문지 등을 사용하여 내 주변에서 자료를 수집할 것인가?’등의 데이터 수집 방법을 정해 보도록 하였다. 다음으로는, 데이터를 수집하기 위한 항목을 정해보고, 데이터를 직접 수집해서 수치화 시킬 수 있는 것들을 수치화시키는 등 데이터를 준비하도록 하였다.

수집 방법

- 대상자: 초등학생
- 어떤 방법: 설문조사

설문지 내용

-평일에 몇 시쯤 잠을 자나요?
(단답형)

-평일에 몇 시쯤 일어나나요?
(단답형)

-주말에 몇 시쯤 잠을 자나요?
(단답형)

-주말에 몇 시쯤 일어나나요?
(단답형)

-잠을 자기 직전에 무엇을 하나요?
>독서 >휴대폰
>컴퓨터나 TV보기 >학교/학원 숙제

-잠을 자는 시간이 건강에 영향을 끼칠 것이라고 생각하나요?
>영향을 끼칠 것이다. >영향을 끼치지 않을 것이다.
>잘 모르겠다.

-잠을 자는 시간이 건강에 미치는 영향을 자세히 알고 있나요?
>자세히 알고 있다. >아예 모른다.

[그림 IV-11] 데이터 수집하기(석○○)

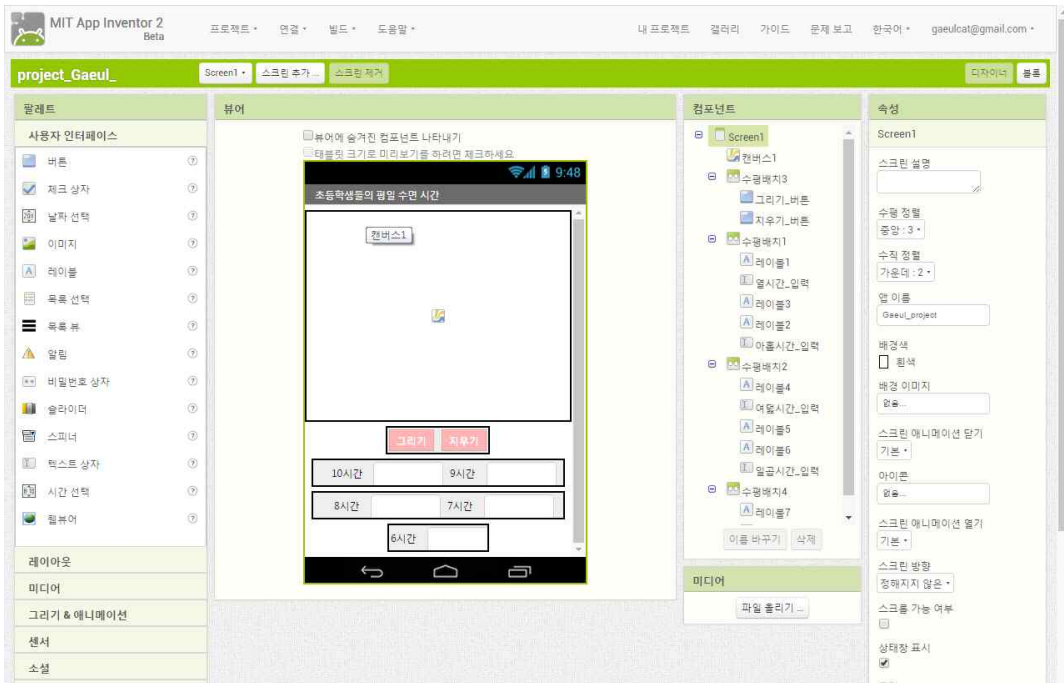
(3) 현상 이해하기: 탐색적 데이터 분석

이 단계에서는 탐색적 데이터 분석을 해 보는 단계로, 원본 데이터 살펴보고 시각화하기, 개별 속성 분석하기, 속성간의 관계 분석하기 등을 수행하도록 하였다.

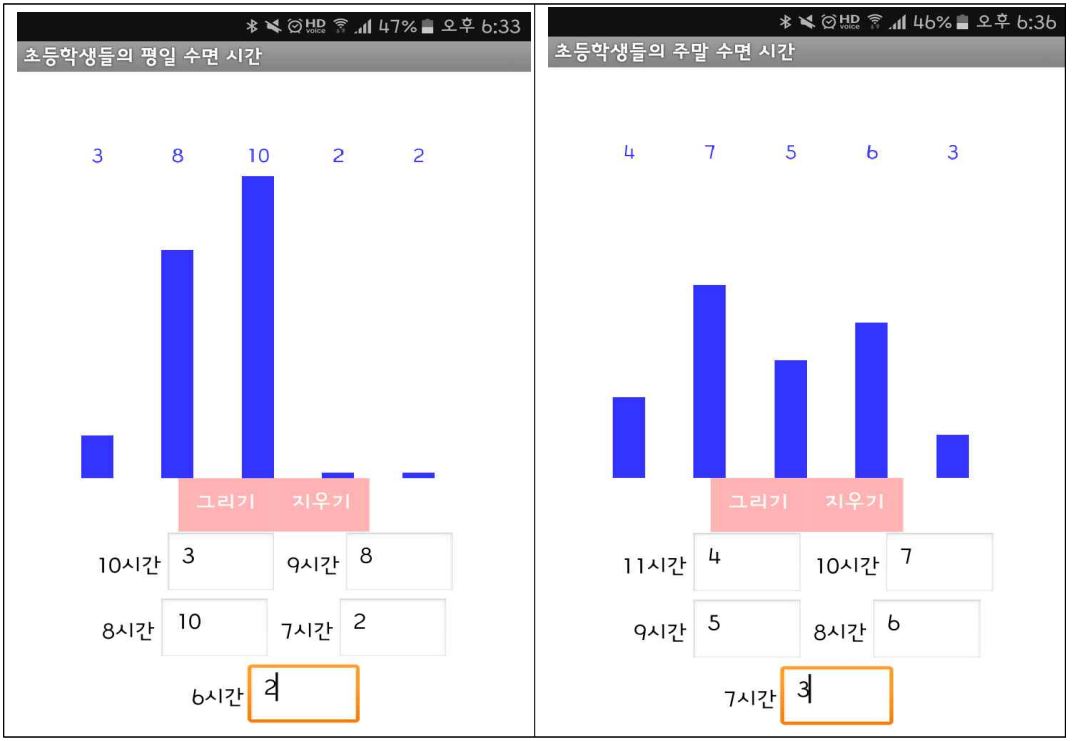
1. 원본데이터 살펴보고 시각화하기 & 개별 속성 분석하기

3. 평일에 못시름 잠을 자나요?	4. 평일에 못시름 일어나나요?	5. 주말에 못시름 잠을 자나요?	6. 주말에 못시름 일어나나요?
12		8	12
11	7:30		11
10		7	10
10		7	10
10		7	10
10		7	11
10		7	11
13		7	12
10		7	11
10	7:30		11
10		7	12
11		9	11
12		8	11
11		7	12
13		9	13
11		9	11
11		7	11
11		9	11
11		7	11
9		7	9
12		7	12

색이 진할수록 늦게 일어나고 늦게 잔다는 것을 나타내고 있다.
(13은 새벽1시를 나타낸다)



원래 초등학생이 건강하기 위해서는 10시간 이상을 자야 하지만 초등학생들의 평일 수면 시간을 나타낸 그래프를 보면 8시간만 자는 학생들이 가장 많고, 10시간 이상 자는 학생은 3명밖에 없다는 것을 알 수 있다. 반대로 주말 수면시간은 대부분의 학생들이 10시간 이상이다.

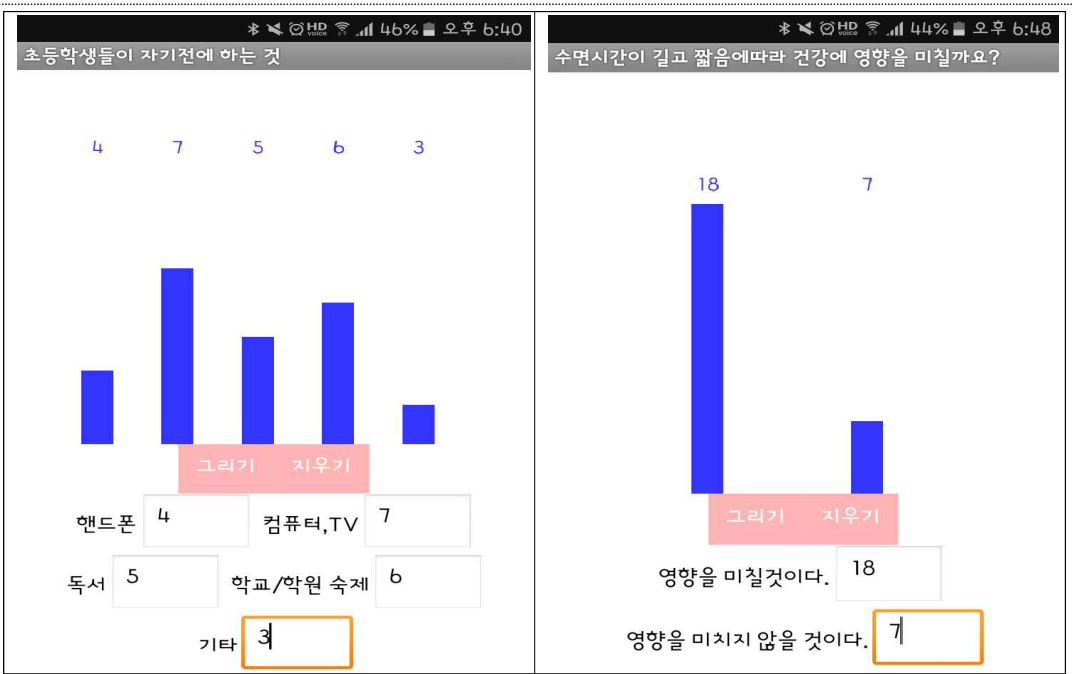


```

언제 그리기_버튼 > 클릭
실행 지정하기 캔버스1 > .보이기 > 값 > 참
지정하기 global list > 값 > 리스트 만들기
출시간_입력 > 텍스트 >
아홉시간_입력 > 텍스트 >
여덟시간_입력 > 텍스트 >
일곱시간_입력 > 텍스트 >
여섯시간_입력 > 텍스트 >
지정하기 캔버스1 > .포인트_작성 > 값 >
지정하기 캔버스1 > .펜_두께 > 값 > 20
호출 드로잉 >
x1 > 50
y1 > 리스트에서 항목 선택하기 리스트 > 가져오기 global list > 위치 > 1
호출 드로잉 >
x1 > 150
y1 > 리스트에서 항목 선택하기 리스트 > 가져오기 global list > 위치 > 2
호출 드로잉 >
x1 > 200
y1 > 리스트에서 항목 선택하기 리스트 > 가져오기 global list > 위치 > 3
호출 드로잉 >
x1 > 250
y1 > 리스트에서 항목 선택하기 리스트 > 가져오기 global list > 위치 > 4
호출 드로잉 >
x1 > 300
y1 > 리스트에서 항목 선택하기 리스트 > 가져오기 global list > 위치 > 5

언제 지우기_버튼 > .클릭
실행 지정하기 캔버스1 > .보이기 > 값 > 거짓
지정하기 열시간_입력 > 텍스트 > 값 > ""
지정하기 아홉시간_입력 > 텍스트 > 값 > ""
지정하기 여덟시간_입력 > 텍스트 > 값 > ""
지정하기 일곱시간_입력 > 텍스트 > 값 > ""
지정하기 여섯시간_입력 > 텍스트 > 값 > ""
언제 Screen1 > .초기화
실행 호출 캔버스1 > .지우기
  
```

이 두 그래프를 보고 많은 학생들이 평일에는 수면시간이 적지만 주말에는 수면시간이 많다는 것을 알 수 있다.



수면시간이 길고 짧음에 따라 건강에 영향을 끼칠 거라고 생각하는지 물어본 결과 영향을 미칠 것이라고 생각하는 학생들은 반을 넘었다. 대부분 학생들은 자기 전 컴퓨터나 TV를 보고 학교/학원 숙제를 한다는 것을 이 그래프를 보고 알 수 있다.



수면시간이 건강에 영향을 미친다고 생각하는 학생은 18명이었지만 자세히 알고 있다고 응답한 사람은 15명이다. 또 아예 모른다고 답한 학생은 10명이었다.

[그림 IV-12] 현상 이해하기: 탐색적 데이터 분석(석○○)

(4) 현상 일반화하기

이 단계에서는 분석한 자료를 바탕으로 내가 세운 가설과 비교하여 결론을 내리고 통계적 추론을 해 보도록 하였다.

결론: 내 생각처럼 설문결과 많은 학생들이 10시간이상 잠을 자지 못하고, 수면시간이 건강에 미치는 영향을 모르는 학생들이 있었다.

통계적 추론: 앞으로 많은 학생들이 수면시간이 건강에 미치는 영향을 모른다면 점점 더 학생들의 수면시간이 줄어들어 건강이 좋지 않을 것이므로 이와 관련된 교육을 실시하거나 많은 학생들에게 이를 알리는 등의 노력이 필요할 것이다.

[그림 IV-13] 현상 일반화하기(석○○)

(5) 현상 예측하기

이 단계에서는 결론을 실험 연구를 통해 확인하고 인과 관계의 통계적 오류를 방지하기 위해 확인 실험 등을 통해 상관관계를 확인해 보며 앞으로의 예측을 해보도록 하였다.

(6) 데이터 스토리텔링

최종 발표(데이터 스토리텔링)는 각자 제작한 최종 프로젝트 보고서를 앞에서는 학생들과 학부모님들 앞에서 발표하는 시간으로 운영하였다.



[그림 IV-14] 데이터 스토리텔링 장면

4) 양적 연구 결과 분석

(1) 컴퓨팅 사고력 변화

① 정규성 검정

실험 집단이 정규성을 확보하고 있는지 확인하기 위하여 정규성 검정을 실시하였다. 사전 검사 데이터를 바탕으로 후속 통계의 비모수/모수 통계 여부를 결정하기 위해 정규성 검정 방법으로 Shapiro-Wilks 검정을 실시하였고 <표 IV-22>에 그 결과를 제시하였다.

<표 IV-22> 컴퓨팅 사고력 검사 정규성 검정

구분	기술통계(N=26)				t	p
	M	SD	Max	Min		
앱인벤터반	12.23	3.972	17	3	.924	.056

*p<.05

앱인벤터반에 대해 Shapiro-Wilks 정규성 검정을 실시한 결과, 유의도가 .056으로 유의수준 .05보다 높게 나타나 정규분포임이 확인되었다.

② 사전·사후 검사 결과 집단 내 비교

앱인벤터반에서 실시한 사전·사후 검사 결과 컴퓨팅 사고력의 변화를 알아보기 위하여 <표 IV-23>과 같이 모수통계인 대응표본 t 검정을 실시하였다.

<표 IV-23> 사전·사후 검사 결과 집단 내 비교(대응표본 t 검정)

대상	N	사전 검사		사후 검사		t	p
		M	SD	M	SD		
앱인벤터반	26	12.23	3.972	13.5	3.455	-2.826	.009**

**p<.01

<표 IV-23>의 대응표본 t 검정의 결과를 살펴보면, 앱인벤터반의 t 통계값

은 -2.826이고 유의확률은 .009로 유의수준 .01에서 컴퓨팅 사고력이 유의미하게 향상된 것으로 나타났다..

(2) 창의성 변화

① 정규성 검정

앱인벤터반의 창의성 사전 검사 결과를 이용하여 정규성 검정을 실시하였다. 사전 검사 결과를 바탕으로 후속 통계의 비모수/모수 통계 여부를 선정하기 위해 정규성 검정 방법으로 Shapiro-Wilks 검정을 실시하였고 <표 IV-24>에 그 결과를 제시하였다.

<표 IV-24> 앱인벤터반 창의성 검사 정규성 검정

하위 요소	기술통계(N=26)				t	p
	M	SD	Max	Min		
유창성	105.84	20.356	145	66	.954	.285
독창성	109.42	20.333	150	68	.959	.367
제목의 추상성	101.34	43.809	158	0	.906	.021*
정교성	107.88	15.079	133	80	.959	.374
성급한 종결에 대한 저항	67.42	33.761	115	0	.917	.038*
창의성 평균	98.38	20.723	139.4	44.8	.972	.681
창의성 지수	102.61	22.074	145.4	44.8	.967	.554

*p<.05

앱인벤터반의 창의성 사전 검사에 대해 Shapiro-Wilks 정규성 검정을 실시한 결과, ‘제목의 추상성’과 ‘성급한 종결에 대한 저항’ 영역의 유의도가 각각 .021, .038로 나타나 정규성이 만족되지 않은 것으로 나타났다. 나머지 영역은 유의도가 유의수준 .05보다 높게 나타나 정규분포임이 확인되었다.

② 사전·사후 검사 결과 집단 내 비교

앱인벤퍼반에서 실시한 사전·사후 검사 결과 창의성의 변화를 알아보았다. 정규성을 확보한 창의성의 하위 요소에 대해서는 <표 IV-25>와 같이 모수통계인 대응표본 t 검정을 실시하였고, 정규성을 확보하지 못한 ‘제목의 추상성’과 ‘성급한 종결에 대한 저항’은 <표 IV-26>과 같이 비모수 통계로 Wilcoxon 부호 순위 검정을 실시하였다.

<표 IV-25> 앱인벤퍼반 사전·사후 검사 결과 집단 내 비교(대응표본 t 검정)

하위요소	N	사전 검사		사후 검사		t	p
		M	SD	M	SD		
유창성	26	105.84	20.356	111.84	20.586	-1.781	.087
독창성	26	109.42	20.333	119.07	19.114	-2.635	.014*
정교성	26	107.88	15.079	128.03	14.572	-8.739	.000**
창의성 평균	26	98.38	20.723	110.57	15.151	-4.590	.000**
창의성 지수	26	102.61	22.074	115.50	15.859	-4.502	.000**

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

<표 IV-26> 앱인벤퍼반 사전·사후 검사 결과 집단 내 비교(Wilcoxon 부호 순위 검정)

하위요소	N	사전 검사		사후 검사		z	p
		M	SD	M	SD		
제목의 추상성	26	101.34	43.809	110.615	30.060	-1.029	.303
성급한 종결에 대한 저항	26	67.42	33.761	81.000	29.640	-2.416	.016*

*p<.05

<표 IV-25>와 <표 IV-26>의 대응표본 t 검정, Wilcoxon 부호 순위 검정의 결과를 살펴보면, ‘독창성’의 t 통계값은 -2.635이고 유의확률은 .014, ‘성급한 종결에 대한 저항’의 t 통계값은 -2.416이고 유의확률은 .016으로 나타나 각각

유의수준 .05에서 사전 검사 점수와 비교했을 때 사후 검사 점수가 유의미하게 상승한 것으로 나타나 창의성의 하위요소가 향상된 것을 알 수 있다.

또한, ‘정교성’의 t 통계값은 -8.739이고 유의확률은 .000, ‘창의성 평균’의 t 통계값은 -4.590이고 유의확률은 .000, ‘창의성 지수’의 t 통계값은 -4.502이고 유의확률은 .000으로 나타나, 데이터 과학 교육 프로그램 투입 결과 유의수준 .01에서 사전 검사 점수와 비교했을 때 사후 검사 점수가 유의미하게 상승하여 창의성의 하위요소가 향상되었음을 알 수 있다.

(3) 컴퓨팅 사고력과 창의성 검사 결과의 분석

컴퓨팅 사고력 사전 검사 결과, 애플벤터반은 정규분포를 갖춘 표본임이 검증되었고 창의성 사전 검사 결과, 애플벤터반은 ‘제목의 추상성’과 ‘성급한 종결에 대한 저항’을 제외하고 정규분포를 갖춘 표본임이 검증되었다.

본 연구에서 개발한 데이터 과학 교육 프로그램을 적용한 후, 사전·사후의 집단 내 대응표본 t 검정의 결과 사전 검사보다 컴퓨팅 사고력이 향상된 것으로 나타났다. 또한, 창의성 사전·사후 검사의 집단 내 대응표본 t 검정, Wilcoxon 부호 순위 검정의 결과 창의성의 유의미한 향상을 보였다.

이로써 본 연구의 애플벤터를 활용한 데이터 과학 교육 프로그램은 학습자의 컴퓨팅 사고력과 창의성을 향상시킬 수 있음을 입증하였다.

이는 데이터 과학 교육에서의 도구 선택이 초등학생 및 초보자를 대상으로 하는 경우 편리한 사용 환경과 사용법이 간단하고 비교적 적은 프로그래밍 경험으로도 사용할 수 있어야 하는데(김진영, 2016), 애플벤터는 프로그래밍 지식이 없는 초보자도 쉽게 드래그 앤 드롭으로 블록을 조합하여 원하는 앱을 제작할 수 있기 때문에(CSTA & ISTE, 2011b) 데이터 과학 교육의 문제 해결 활동에 쉽게 사용법을 익혀 활용할 수 있는 도구로써 적합하였기 때문인 것으로 생각된다.

또한, 학생들 스스로 흥미와 관심을 갖고 있는 주제를 선정하고 구글 설문지나 통계청 누리집 등을 활용하여 개인 데이터 및 공공 데이터를 수집하여 분석하고 일반화하여 통계적 추론을 해보는 학습활동을 꾸준히 진행했다. 이처럼 데이터 문제 해결 단계에 따라 스스로 자료를 수집, 조직, 표현하고 해석하

는 직접적인 활동 중심 문제 해결 학습 과정에서 컴퓨팅 사고력의 구성 요인들과 창의성의 하위 요소들이 향상되어 컴퓨팅 사고력과 창의성 향상에 긍정적인 영향을 미친 것으로 생각된다.

또한, 학생들 스스로 흥미와 관심을 갖고 데이터 문제 해결 활동을 지속적으로 수행할 수 있었던 것은 선행 연구에서 앱인벤터를 활용하여 컴퓨팅 사고력 교육 과정에 기반한 교육프로그램을 개발하여 적용한 결과 학습에 대한 지속적인 흥미와 몰입이 향상된 것으로 나타났다(황성진, 2013)는 연구 결과와도 일치한다.

5) 질적 연구 결과 분석

본 연구의 자료 수집 및 분석은 질적 사례 연구에서 대표적으로 사용되는 방법인 내용분석을 주로 활용하였는데, 내용분석에는 총 26명의 학습자들이 6일간 총 42차시의 집합 강의 및 실습을 통해 데이터 문제 해결 단계에 따라 직접 작성한 26편의 데이터 과학 교육 보고서 자료를 활용하였다.

(1) 주제 선정 이유(동기) 및 내용

대부분의 학생들이 평소에 내 자신이나 주변에서 궁금했던 점을 해결해 보기 위해 다양한 주제를 선정했다. 예를 들면, ‘사람은 어떤 기억을 더 오랫동안 기억할까?’, ‘스마트폰과 게임’, ‘제주도 내진 설계’, ‘초등학생의 수면 시간은 키와 관계가 있을까?’, ‘학생들에게 충치가 생기는 이유는 무엇일까?’, ‘현재 학생들은 욕을 얼마나 사용하고 있을까?’처럼 교육 프로그램 적용 초기에 비해 최종 보고서에서는 초등학생이 아니면 생각해낼 수 없는 다양하고 창의적인 주제들이 많이 나와서 교육 프로그램을 통한 창의성 향상에 따른 긍정적 효과라고 생각되었다.

(2) 단계별 데이터 과학 분석

문제점이나 궁금한 점을 해결하여 얻고 싶은 점과 문제점이나 궁금한 점에 대해 학습자들이 생각한 이유에 대해 학습자들 대부분이 나름대로 논리적으로

가설 설정을 하였고, 직접 작성한 구글 설문지 등을 사용하여 다양한 대상을 상대로 자료를 수집하였다.

또한, 앱인벤터의 특징을 잘 활용하여 원본 데이터를 분석하여 차트 등의 시각화 자료로 활용하여 한 눈에 이해하기 쉽게 제시하였고 개별 속성을 잘 분석하였다.

◆ 3단계- 탐색적 데이터 분석 ◆

1. 원본데이터 살펴보고 시각화하기 (조건부 서식, 차트 그림)

- 요약 -

2. 개별 속성 분석하기 (피벗 테이블, 피벗 차트 넣고 설명)

- 그림 생략 -

- 12시간 잔 학생의 키는 160 이상이었다.
- 하지만 5시간 밖에 안 잔 학생들의 키가 155 이상으로 평균적이었다.
- 그리고 키가 가장 작은 학생도 10시간이나 잤으며, 11시간 잔 학생의 키가 146 밖에 안 되는 경우도 있었다.

◆ 4단계- 현상 일반화하기, 통계적 추론 ◆

1. 결론: 내 가설과는 다르게 위 차트를 통하여 수면시간 만으로는 키 성장과 관련이 그렇게 있지는 않은 것 같다. 수면시간과 평소의 식생활, 운동 등을 비교한다면 키 성장과의 관련이 있을 것이다.

[그림 IV-15] 속성 간의 관계 분석(상관관계 분석) 후 현상 일반화하기(김○○)

그러나 속성 간의 관계 분석하기에서는 [그림 IV-16]처럼 상관관계 분석이 제대로 이루어지지 않은 학습자들이 대부분이었다. 엑셀을 활용한 데이터 과학 교육 프로그램처럼 초등학생들의 특성상 설문 대상이 가족·학교 친구 등에 국한되어 수집한 설문 표본 수가 작고 현상 일반화와 현상 예측을 하기에는 데이터 수집 기간이 짧아서 양질의 데이터를 수집하지 못한 데서 기인한다고 할 수 있다. 또한, 평소 상관관계보다는 인과 관계 위주로 문제 해결을 했던

경험 때문에 상관관계를 원인과 결과의 필연적인 관계인 인과 관계로 생각하려는 경향이 있기 때문인 것으로 생각된다.

하지만, [그림 IV-16]의 학습자처럼 상관관계 분석을 통해 본인의 가설에 문제가 있음을 알고 ‘수면시간’만이 아니라, ‘평소의 식생활’, ‘운동’ 등 다른 요인을 대안으로 제시하는 발전된 학습자도 있어 교육 프로그램의 효과를 체감할 수 있었다.

3. 스크래치를 활용한 대학생 및 예비 코딩 강사 대상 데이터 과학 교육 프로그램

1) 요구 분석

제주대학교 1학년 전공필수 ‘초등컴퓨터교육 I’을 수강하는 19명의 학생과 컴퓨터교육 전공 초등 현직교사 20명을 대상으로 아래와 같은 내용에 대하여 요구 분석을 실시하였다.

- 소프트웨어교육 참여 경험
- 교육용 프로그래밍 언어, 언플러그드 활동, 피지컬컴퓨팅에 대한 관심도
- 데이터 과학 교육을 통한 능력 향상
- 데이터 과학 교육 활용 데이터
- 데이터 과학 교육 학습 방법

요구 분석 결과를 종합하여 대학생들이 가장 선호한 개별 프로젝트 학습 중심으로 스크래치를 활용한 데이터 과학 교육 프로그램을 설계 및 개발하여 적용한 후 그 결과를 분석하였다.

2) 설계 및 개발

- (1) 연구 가설: 스크래치를 활용한 데이터 과학 교육 프로그램은 학습자의 창의성을 향상시킬 수 있다.

(2) 연구 대상

본 연구에서 적용한 프로그램의 교육적 효과를 확인하기 위해 <표 IV-27>과 같이 제주대학교 1학년 전공필수 ‘초등컴퓨터교육 I’을 수강하는 학생 중 19명의 학생과 제주창조경제혁신센터에서 주관하고 제주대학교에서 진행한 ‘청년 코딩 사업’에 지원한 지원자 표본 13명 총 32명을 대상으로 오리엔테이션, 사전·사후 컴퓨팅 사고력 검사·창의성 검사를 포함하여 21차시 수업으로 진행되었다. 모든 학습자들은 ‘데스크 탑 PC’를 각자 사용하였고 ‘Scratch 2.0’을 사용하였다. 연구 대상에 대한 구체적인 내용은 <표 IV-27>에 제시하였다.

<표 IV-27> 연구 대상

구분	학습자수(명)		
	남	여	소계
대학생	3	16	19
예비 코딩 강사	9	4	13
소계	12	20	32

대학생 참여자 중 1명은 창의성 사전 검사는 실시하였으나 사후 검사를 시행하지 못하였으므로 결측값으로 처리한 후 통계를 실시하였다.

(3) 연구 설계

본 연구에서 개발한 교육 프로그램은 제주대학교 1학년 전공필수 ‘초등컴퓨터교육 I’ 45차시 중 21차시와 제주창조경제혁신센터에서 주관하고 제주대학교에서 진행한 ‘청년 코딩 사업’ 강의 72차시 중 21차시 동안 집합 강의 및 실습으로 이루어졌다. 첫 시간에는 사전 검사와 오리엔테이션 중심으로, 마지막 시간에는 사후 검사 및 학습자들이 만든 프로젝트를 최종 발표하는 시간으로 이루어졌고, 각각 주 1회 3시간씩 7주와 주 2회 3시간씩 4주 동안 운영되었다.

<표 IV-28> 프로그램 일정

주차	1~2	3~6	7
대학생 일정	사전 검사, 오리엔테이션, 스크래치 기초	데이터 과학 교육	최종발표 사후 검사
주차	1	2~3	4
예비 코딩 강사 일정	사전 검사, 오리엔테이션, 스크래치 기초	데이터 과학 교육	최종발표 사후 검사

(4) 교육 내용의 선정

요구 분석 결과 및 데이터 과학 교육 도구의 특징에 따라 <표 IV-29>와 같이 교육 내용을 선정하였다.

<표 IV-29> 교육 프로그램 주제(스크래치)

차시	학습 주제
1~3	데이터 과학의 의미를 이해하고 이를 위한 스크래치의 기본 기능 익히기
4~6	반복 구조, 조건 구조, 변수와 리스트, 펜 기능의 이해, 블록 정의하기(재귀)
7~9	생활 속에서 궁금한 내용을 공공 데이터를 수집하여 해결해 보기 1 (막대그래프 구현하기)
10~12	생활 속에서 궁금한 내용을 공공 데이터를 수집하여 해결해 보기2 (꺾은선 그래프 구현하기)
13~15	생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 보기 1 (띠그래프 구현하기)
16~18	생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 보기 2 (원그래프 구현하기)
19~21	데이터 스토리텔링하기

(5) 검사 도구

① 양적 연구

컴퓨팅 사고력 측정 도구인 김병수(2014)의 계산적 인지력 검사 A, B형은

초등학생 수준의 검사이기 때문에 대학생 및 예비 코딩 강사를 대상으로 사용하기에는 무리가 있어 사용하지 않았다.

따라서, 본 연구에서는 창의성만을 대상으로 측정하였다. 창의성 측정은 사전-사후검사 통제집단 설계를 활용하였다. 창의성 검사 도구는 TTCT 검사지 도형 A, B형을 사용하였다.

<표 IV-30> 실험 설계

대상	사전 검사	처치	사후 검사
대학생 예비 코딩 강사	O ₁	X ₁	O ₂

X₁ : 스크래치를 활용한 데이터 과학 교육 프로그램

O₁ : 사전검사(창의성 검사)

⇒ 정규성 검정

O₁, O₂ : 사전·사후 검사(창의성 검사)

⇒ 정규성 검정 결과에 따른 대응 표본 t 검정 또는 Wilcoxon 부호 순위 검정

③ 질적 연구

총 32명의 학습자들이 총 21차시의 집합 강의 및 실습을 통해 작성한 최종 데이터 과학 교육 보고서 각 1편씩을 대상으로 내용분석을 하였다.

3) 적용

앞서 제시한 6단계의 데이터 과학 교육 문제 해결 단계에 따른 실제 학습 과정은 다음과 같이 진행되었다. 16~18차시 학습 주제인 ‘생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 보기(원그래프 구현하기)’에 따라, ‘양○○’ 학생이 작성한 ‘주제와 주제 선정 이유(동기) 및 내용’은 [그림 IV-16]과 같다.

주제: 외모에 만족하는 사람이 더 많이 꾸밀까?

주제와 주제 선정 이유 및 내용: 교육대학을 돌아다니면서 혹은 시청 거리를 지나가다가 주위를 둘러보면 거의 다 나보다 더 꾸미고 다니지, 덜 꾸미는 사람은 눈 씻고 찾아볼 수가 없다. 그리고 친구와 이야기를 나누거나 다양한 경험들을 통해서, 알고 지내는 친구가 외모에 만족하는지 아니면 불만족하는지에 대해 대략적으로 알 수 있는데, 내 경험상 외모에 만족하는 사람들이 자신을 더 많이 꾸미고 신경 쓴다는 느낌이 들었다. 나는 예전부터 확신해 온 내 생각이 맞는지 확인해보고 싶어서 이 주제를 선정하게 되었다.

한편, 슬프게도 나는 나의 외모에 대해 약간 불만족하는 편이다. 그래서인지는 모르겠지만 ‘내가 꾸며 봤자 얼마나 달라지겠어?’ 라는 생각에 치장하는 데에 신경을 많이 쓰지 않는다. 이와 반대로 내 주위의 나보다 옷을 다양하게 입고, 화장을 더 많이 하는 사람들은 자신들의 외모에 만족해서 더 많이 꾸미는 것인지, 아니면 나와 같이 외모에 불만족하지만 그것을 감추기 위해 꾸며서 예쁘게(잘 생기게) 보이려고 하는 것인지 궁금해졌다. 최종적으로는 어쩔 수 없는 대한민국의 외모지상주의 사회에서 자기 만족감을 높이고 자신에게 더 신경 쓸 수 있는 방법을 모색한다.

[그림 IV-16] 주제와 주제 선정 이유(동기) 및 내용(양○○)

(1) 데이터 문제 정의하기(문제 탐색하기)

이 단계에서는 문제(문제점이나 궁금한 점이 무엇인가?), 목표(문제점이나 궁금한 점을 해결하여 얻고 싶은 점은 무엇인가?), 가설(문제나 궁금한 점에 대해 내가 생각한 이유는 무엇인가?)을 [그림 IV-17]과 같이 정해 보도록 하였다.

문제: 외모에 만족하는 사람이 더 많이 꾸미고 다닐까?

목표: 외모에 만족하는 사람이 더 많이 꾸미는 지에 대한 궁금증을 해결하고, 그들이 외모에 만족하는 이유를 분석한다. 이를 통해 나뿐만이 아니라 다른 사람들도 외모지상주의 사회에서 타인에게 상처받지 않으면서 자신의 외모에 만족할 수 있는 방안을 찾고, 더 많이 자신을 꾸미고 가꿔서 자기 홍보의 시대에서 살아남을 수 있다.

가설: 외모에 만족하는 사람이 그렇지 않은 사람보다 더 많이 꾸미고 다닌다.

[그림 IV-17] 데이터 문제 정의하기(양○○)

(2) 데이터 수집하기

이 단계는 필요한 데이터를 스스로 만들어보는 단계로 ‘체크리스트 등을 사용하여 내 자신의 누적 자료로 수집할 것인가?, 구글 설문지 등을 사용하여 내 주변에서 자료를 수집할 것인가?’ 등의 데이터 수집 방법을 정해 보도록 하였다. 다음으로는, 데이터를 수집하기 위한 항목을 정해보고, 데이터를 직접 수집해서

수치화시킬 수 있는 것들을 수치화시키는 등 데이터를 준비하도록 하였다.

수집 방법

- 대상: 교육대학(영어과&음악과) 친구들, 타 대학 친구들, 가족 등 지인 다수
- 방법: 구글 설문지 (<https://goo.gl/forms/fsPDgJIJth1Ub4W12>)

설문지 내용 (구글 설문지 사용)

<외모에 만족하는 사람이 더 많이 꾸밀까?>

-외모에 만족하는 사람이 그렇지 않은 사람보다 더 많이 꾸밀지, 아니면 덜 꾸밀지 알아본다.

1. 본인의 외모에 만족하십니까?

1	2	3	4	
□	□	□	□	매우 만족

2. 본인의 외모에 만족하는 이유가 무엇입니까? (1에서 3, 4번 선택자만/복수응답가능)

- 1 - 거울을 볼 때 내 모습이 만족스러워서
- 2 - 주변에서 외모 칭찬을 많이 들어서
- 3 - 원래 자존감이 높아서 (매사에 긍정적)
- 4 - 외모에 만족을 안 하는 것보단 주어진 외모에 만족하는 것이 낫다고 생각해서
- 5 - 본인의 외모가 보통 이상이라고 생각해서
- 6 - 기타: _____

3. 한 달에 보통 의류를 몇 개 정도 구입하십니까? (신발, 상하의, 악세사리 등을 모두 포함해서)

0개 1~2개 3~4개 5개 이상

4. 보통 화장을 어느 단계까지 하십니까? (여자만 응답)

- 1단계: 스킨 or 로션
- 2단계: 선크림 or 파운데이션(쿠션) or 비비크림
- 3단계: 눈썹그리기 or 아이라인 or 섀도우 or 마스크라
- 4단계: 웨딩 or 하이라이트 or 블러셔 (얼굴에 양감과 음감을 주는 것)

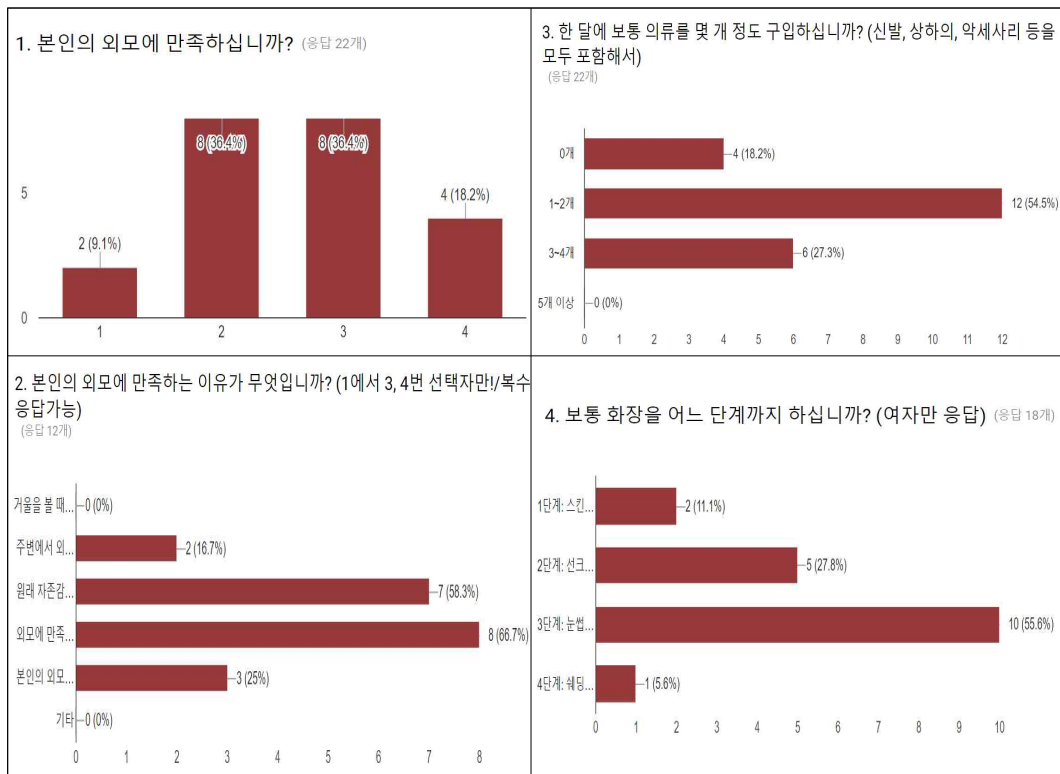
[그림 IV-18] 데이터 수집하기(양○○)

(3) 현상 이해하기: 탐색적 데이터 분석

이 단계에서는 탐색적 데이터 분석을 해 보는 단계로, 원본 데이터 살펴보고 시각화하기, 개별 속성 분석하기, 속성간의 관계 분석하기 등을 수행하도록 하였다.

1. 원본데이터 살펴보고 시각화하기 (스크래치)

<구글 설문지 원본데이터>



문제	응답 수 총 22개										
1번	2	2	4	3	2	3	1	3	1	4	3
2번	X	X	3	3,4	X	4	X	2,4,5	X	2,3,5	3,4
3번	0개	1~2개	1~2개	1~2개	1~2개	3~4개	0개	1~2개	1~2개	3~4개	3~4개
4번	2단계	3단계	3단계	2단계	2단계	3단계	X	3단계	1단계	4단계	3단계
1번	2	3	2	2	4	3	3	2	4	3	2
2번	X	4	X	X	3,4,5	3	3	X	4	4	X
3번	1~2개	1~2개	0개	1~2개	3~4개	1~2개	3~4개	1~2개	1~2개	0개	3~4개
4번	X	3단계	1단계	X	3단계	2단계	3단계	3단계	X	2단계	3단계

<스크래치로 시각화>

161129 외모에 만족하는 이유? (1에서 3,4번 선택자/복수응답)
by antelopes9 (공유되지 않음)

데이터

1	0
2	2
3	7
4	8
5	3
6	0

길이: 6

외모에 만족하는 이유의 비율이 이렇게 나오는군요!!

<외모에 만족하는 이유의 비율>

161129 외모에 만족하는 사람이 꾸미는 비율은? (옷, 화장)
by antelopes9 (공유되지 않음)

외모에 만족하는 사람들은 이만큼 꾸미는군요. (옷 or 화장)

<외모에 만족하는 사람이 꾸미는 비율(옷)>

161129 외모에 만족하는 사람이 꾸미는 비율은? (옷, 화장)
by antelopes9 (공유됨)

외모에 만족하는 사람들은 이만큼 꾸미는군요. (옷 or 화장)

<외모에 만족하는 사람이 꾸미는 비율(화장)>

161129 외모에 불만족하는 사람이 꾸미는 비율은? (옷, 화장)
by antelopes9 (공유되지 않음)

외모에 불만족하는 사람들은 이만큼 꾸미는군요. (옷 or 화장)

<외모에 불만족하는 사람이 꾸미는 비율(옷)>

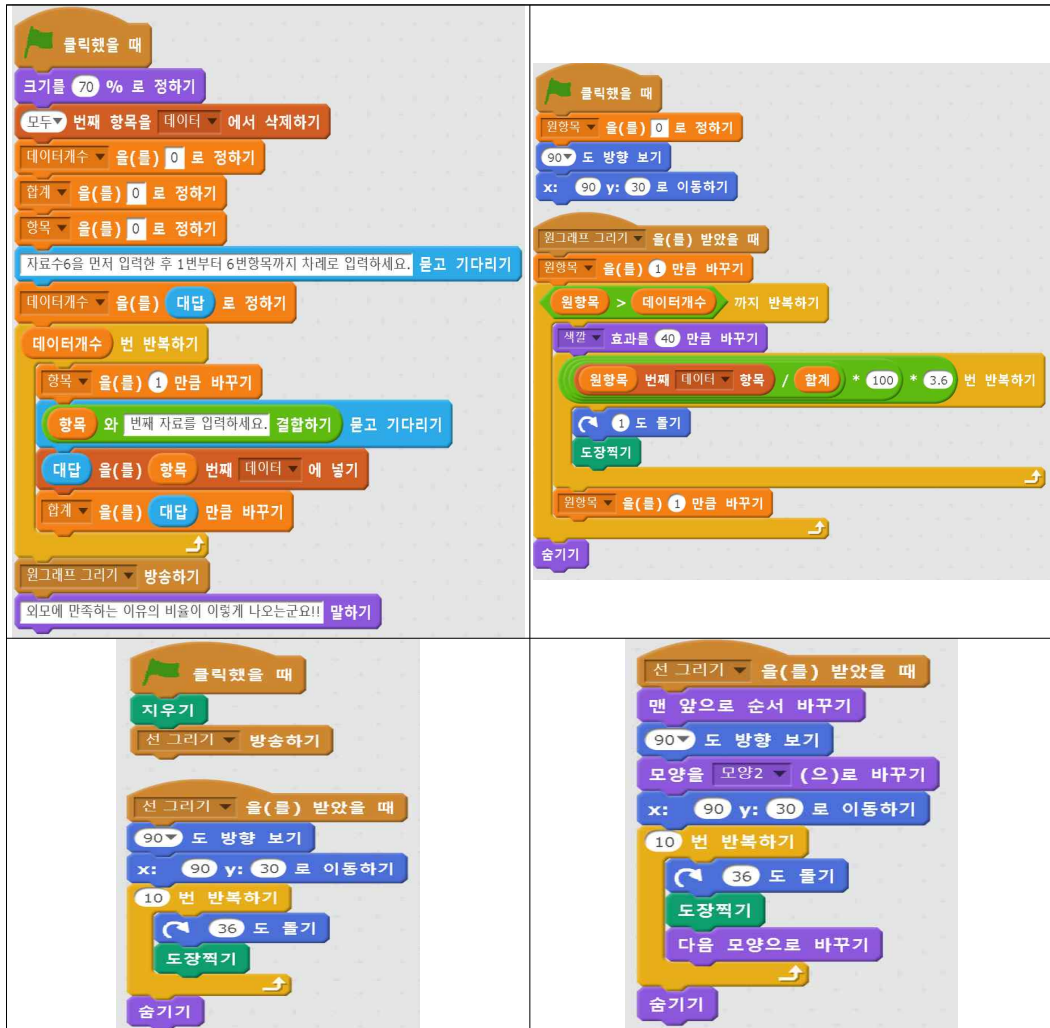
161129 외모에 불만족하는 사람이 꾸미는 비율은? (옷, 화장)
by antelopes9 (공유됨)

외모에 불만족하는 사람들은 이만큼 꾸미는군요. (옷 or 화장)

	더 꾸밈
	덜 꾸밈

<외모에 불만족하는 사람이 꾸미는 비율(화장)>

<주요 스크래치 스크립트- '외모에 만족하는 이유'>



2. 개별 속성 분석하기(스크래치)

'외모에 만족하는 이유'에는 거울을 볼 때 내 모습이 만족스러워서, 주변에서 외모 칭찬을 많이 들어서, 원래 자존감이 높아서(매사에 긍정적), 외모에 만족을 안 하는 것보다는 주어진 외모에 만족하는 것이 낫다고 생각해서, 본인의 외모가 보통 이상이라고 생각해서, 기타의 선택 항목이 있었다. 이것들은 순서대로, 0명, 2명, 7명, 8명, 3명, 0명의 응답수를 보여주었다. 스크래치로 시각화해본 결과 외모에 만족을 안 하는 것보다는 주어진 외모에 만족하는 것이 낫다고 판단해서 외모에 만족하는 경우가 제일 많았고, 이와 근소한 차이를 보인 원래 자존감이 높은 것(매사에 긍정적)이 두 번째로 높은 외모에 만족하는 이유였다. 나머지 선택 항목들은 이와 큰 차이를 보였다.

한편, 본인의 외모에 만족하는 사람은 12명, 불만족하는 사람은 10명이었다. 그리고 유형을 크게 4가지로 분류한다면, 외모에 만족하면서 더 꾸미는 사람, 외모에 만족하면서 덜 꾸미는 사람, 외모에 불만족하면서 더 꾸미는 사람, 외모에 불만족하면서 덜 꾸미는 사람으로 나눌 수 있다. 꾸미는 정도에 대한 분류 기준을 옷으로 잡는다면 이는 각각 5명, 7명, 1명, 9명이다. 꾸미는 정도에 대한 분류 기준을 화장으로 잡는다면 이는 각각 8명, 3명, 3명, 4명이다.

3. 속성간의 관계 분석하기(스크래치)

외모에 만족하는 이유는 외모에 만족을 안 하는 것보다는 주어진 외모에 만족하는 것, 원래 자존감이 높은 것(매사에 긍정적), 본인의 외모가 보통 이상이라고 생각하는 것, 주변에서 외모 칭찬을 많이 듣는 것 순으로 많은 응답 수를 보였다. 아무래도 현실과 타협하는 것이 제일 쉬운 방법이라서 이런 결과가 나온 것 같다. 주어진 외모에 만족하는 것과 자존감이 높은 것(매사에 긍정적)도 상관관계가 있다고 볼 수 있는데, 주어진 외모에 만족하는 것은 다소 긍정적인 성격이라는 것과 의미가 상통하기 때문이다. 따라서 비슷한 응답 수를 보인 것으로 추측된다. 또한 본인의 외모가 보통 이상이라고 생각하는 것과 주변에서 외모 칭찬을 많이 듣는 것도 비슷한 응답 수를 보였는데, 만약 본인의 외모가 보통 이상인 것이 사실이라면 주변에서 외모 칭찬을 많이 들을 수 있기 때문이라고 생각한다. 실제 설문 결과에서도 주변에서 칭찬을 들어서 외모에 만족한다는 사람 2명은 자신의 외모가 보통 이상이라고 생각하는 것에도 복수 응답하였다.

꾸미는 정도에 대한 분류 기준이 옷일 때 외모에 불만족하면서 덜 꾸미는 사람은 9명이 고, 외모에 불만족하면서 더 꾸미는 사람은 1명이었는데 이 차이가 엄청 많이 나고, 꾸미는 정도에 대한 분류 기준을 화장으로 했을 때 외모에 만족하면서 더 꾸미는 사람이 8명, 외모에 만족하면서 덜 꾸미는 사람이 3명인 것도 차이가 많이 나는데, 이러한 설문 결과들이 나의 가설이 참이라는 것을 지지하는 좋은 근거가 된다.

개인 안에서도 상관관계를 찾을 수가 있었는데, 옷과 화장 모두 더 꾸미거나, 모두 덜 꾸미는 쪽이 13명이고, 더 꾸미는 것 하나, 덜 꾸미는 것 하나 선택한 사람이 5명이기 때문이다.(남자 제외) 그리고 여자 11명 만족, 여자 7명 불만족, 남자 1명 만족, 남자 3명 불만족인 결과가 나왔는데, 남자의 응답수도 별로 없고, 그 차이도 별로 없지만 여자는 보통 자신의 얼굴에 만족하고 남자는 불만족 한다고 추론해볼 수도 있다고 생각한다. 외모에 만족하는 정도에 따라서도 상관관계를 찾아보려 하였으나 응답 수가 적었기 때문에 결론을 내릴 수가 없었다.

[그림 IV-19] 현상 이해하기: 탐색적 데이터 분석(양○○)

(4) 현상 일반화하기

이 단계에서는 분석한 자료를 바탕으로 내가 세운 가설과 비교하여 결론을 내리고 통계적 추론을 해 보도록 하였다.

결론: 나의 가설은 ‘외모에 만족하는 사람이 그렇지 않은 사람보다 더 많이 꾸미고 다닌다.’ 였는데 이번에 조사하고 정리해보니 가설이 맞았다는 것을 알 수 있었다. 그리고 자신의 외모에 만족하려면 있는 자신을 있는 그대로 받아들이려는 넓은 마음과 자기 이해가 필요하다. 이에 못지않게 매사에 긍정적으로 생각하고 자존감을 가지는 것도 자신의 외모에 만족하기 위해 필요하다.

통계적 추론: 위와 같은 결론을 내린 후, 외모에 만족하는 사람이 그렇지 않은 사람보다 더 많이 꾸미고 다닌다는 것을 깨달았기 때문에, 내가 나의 외모에 만족을 하고 행복해지려면 일단 내 얼굴을 있는 그대로 인정해야 한다.

‘피할 수 없다면 즐겨라!’라는 말도 있듯이 내 얼굴을 바꿀 수 없다면 거울을 보면서 예쁘다고 해주면 될 것이다. 그리고 자존감이 낮고 매사에 부정적인 것도 외모 만족도에 영향을 많이 미치기 때문에 이것들도 긍정적인 방향으로 바꾸도록 노력해야겠다. 나와 비슷한 다른 사람들도 마찬가지이다. 이렇게 노력한다면 외모 만족도가 높아질 것이고, 더 꾸미고 다녀서 나중에는 더 높은 외모 만족도를 가지고, 외면부터 내면까지 더 아름다운 사람이 될 수 있을 것이다.

[그림 IV-20] 현상 일반화하기(양○○)

(5) 현상 예측하기

이 단계에서는 결론을 실험 연구를 통해 확인하고 인과 관계의 통계적 오류를 방지하기 위해 확인 실험 등을 통해 상관관계를 확인해 보며 앞으로의 예측을 해보도록 하였다.

(6) 데이터 스토리텔링

최종 발표(데이터 스토리텔링)는 각자 제작한 최종 프로젝트 보고서를 앞에 나와서 발표하는 시간으로 운영하였다.

4) 양적 연구 결과 분석

(1) 창의성 변화

① 정규성 검정

대학생 및 예비 코딩 강사의 창의성 사전 검사 결과를 이용하여 정규성 검정을 실시하였다. 사전 검사 결과를 바탕으로 후속 통계의 비모수/모수 통계 여부를 선정하기 위해 정규성 검정 방법으로 Shapiro-Wilks 검정을 실시하였고 <표 IV-31>에 그 결과를 제시하였다.

<표 IV-31> 대학생 및 예비 코딩 강사 창의성 검사 정규성 검정

하위 요소	기술통계(N=31)				t	p
	M	SD	Max	Min		
유창성	107.90	24.399	150	63	.964	.379
독창성	119.74	20.200	154	79	.963	.360
제목의 추상성	117.48	26.910	160	66	.949	.143
정교성	118.35	18.902	139	68	.880	.002**
성급한 종결에 대한 저항	93.58	21.455	130	40	.933	.055
창의성 평균	111.41	15.378	133.4	82	.943	.101
창의성 지수	119.03	16.214	143.40	86	.949	.148

**p<.01

대학생 및 예비 코딩 강사의 창의성 사전 검사에 대해 Shapiro-Wilks 정규성 검정을 실시한 결과, ‘정교성’ 영역은 유의도가 .002로 나타나 정규성이 만족되지 않은 것으로 나타났다. 나머지 영역은 유의도가 유의수준 .05보다 높게 나타나 정규분포임이 확인되었다.

② 사전·사후 검사 결과 집단 내 비교

대학생 및 예비 코딩 강사를 대상으로 실시한 사전·사후 검사 결과 창의성의 변화를 알아보았다. 정규성을 확보한 하위 요소에 대하여 <표 IV-32>와 같이 모수통계인 대응표본 t 검정을 실시하였고, 정규성을 확보하지 못한 ‘정교성’은 <표 IV-33>과 같이 비모수 통계로 Wilcoxon 부호 순위 검정을 실시하였다.

<표 IV-32> 사전·사후 검사 결과 집단 내 비교(대응표본 t 검정)

하위요소	N	사전 검사		사후 검사		t	p
		M	SD	M	SD		
유창성	31	107.90	24.399	118.70	23.261	-2.673	.012*
독창성	31	119.74	20.200	136.38	18.787	-4.028	.000**
제목의 추상성	31	117.48	26.910	130.87	23.836	-3.540	.001**
성급한 종결에 대한 저항	31	93.58	21.455	97.74	14.921	-1.092	.283
창의성 평균	31	111.41	15.378	123.80	12.806	-6.520	.000**
창의성 지수	31	119.03	16.214	133.34	14.404	-7.010	.000**

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

<표 IV-33> 사전·사후 검사 결과 집단 내 비교(Wilcoxon 부호 순위 검정)

하위요소	N	사전 검사		사후 검사		z	p
		M	SD	M	SD		
정교성	31	118.35	18.902	135.29	13.770	-4.150	.000**

**p<.01

<표 IV-32>와 <표 IV-33>의 대응표본 t 검정, Wilcoxon 부호 순위 검정의 결과를 살펴보면, ‘유창성’의 t 통계값은 -2.673이고 유의확률은 .012로 나타나 유의수준 .05에서 사전 검사 점수와 비교했을 때 사후 검사 점수가 유의미하게 상승한 것으로 나타났다. 또한, ‘독창성’의 t 통계값은 -4.028이고 유의확률은 .000, ‘제목의 추상성’의 t 통계값은 -3.540이고 유의확률은 .001, ‘창의성 평균’의 t 통계값은 -6.520이고 유의확률은 .000, ‘창의성 지수’의 t 통계값은 -7.010이고 유의확률은 .000, ‘정교성’의 t 통계값은 -4.150이고 유의확률은 .000으로 나타나, 유의수준 .01에서 사전 검사 점수와 비교했을 때 사후 검사 점수가 유의미하게 상승한 것으로 나타나 창의성의 하위 요소가 향상된 것을 알 수 있다.

③ 창의성 검사 결과의 분석

창의성 사전 검사 결과, 대학생 및 코딩 강사 대상 스크래치반은 ‘정교성’을 제외하고 정규분포를 갖춘 표본임이 검증되었다.

본 연구에서 개발한 데이터 과학 교육 프로그램을 적용한 후 창의성 사전·사후의 집단 내 대응표본 t 검정, Wilcoxon 부호 순위 검정의 결과 창의성의 유의미한 향상을 보였다. 이로써 본 연구의 스크래치를 활용한 데이터 과학 교육 프로그램은 학습자의 창의성을 향상시킬 수 있음을 입증하였다.

데이터 과학 교육에서의 도구 선택은 초등학생 및 초보자를 대상으로 하는 경우 편리한 사용 환경과 사용법이 간단하고 비교적 적은 프로그래밍 경험으로도 사용할 수 있어야 하는데(김진영, 2016), 스크래치는 프로그래밍을 처음 접하는 학생들을 위한 교육용 프로그래밍 언어로 누구나 쉽게 사용할 수 있고 다양한 유형, 개인적인 관심·경험 등에 대한 프로젝트 제작이 가능하다(Resnick et al, 2009). 따라서, 소프트웨어교육 경험이 전무하거나 적은 초보자들이 대부분이었던 대학생 및 예비 코딩 강사를 대상으로 한 데이터 과학 교육의 문제 해결 활동에 쉽게 사용법을 익혀 활용할 수 있는 도구로써 적합하였기 때문인 것으로 생각된다.

또한, 스크래치를 활용하여 데이터 문제 해결 단계에 따라 스스로 자료를 수집, 조직, 표현하고 해석하는 직접적인 활동 중심 문제 해결 학습 과정을 통해 창의성이 향상된 것은 선행 연구에서 스크래치 프로그래밍 학습 후 창의성에 긍정적인 영향을 미쳤다는 연구(김종진 외, 2010)와 스크래치 프로그래밍 학습이 창의적 문제해결력에 긍정적인 효과를 미쳤다는 연구 결과(류충구, 2012)와도 일치한다.

5) 질적 연구 결과 분석

본 연구의 자료 수집 및 분석은 질적 사례 연구에서 대표적으로 사용되는 방법인 내용분석을 주로 활용하였는데, 내용분석에는 총 32명의 학습자들이 21차시의 집합 강의 및 실습을 통해 데이터 문제 해결 단계에 따라 직접 작성한 32편의 데이터 과학 교육 보고서 자료를 활용하였다.

(1) 주제 선정 이유(동기) 및 내용

대부분의 학습자들이 평소에 내 자신이나 주변에서 궁금했던 점을 해결해 보기 위해 개인 데이터와 공공 데이터를 적절하게 선택하여 다양한 주제를 선정했다. 예를 들면, ‘여자 친구 혹은 남자 친구가 있으면, 학교생활을 하는데 행복한 비율이 높은가?’, ‘기숙사와 자취 중 어느 것이 더 만족도가 높을까?’, ‘야식을 자주 먹으면 그만큼 아침을 안 먹을까?’, ‘술을 마시는 빈도수와 지각의 빈도수의 상관관계’, ‘외모에 만족하는 사람이 더 많이 꾸밀까?’, 글씨가 예쁘면 필기를 많이 할까?’, ‘요즘 내가 하는 고민의 종류는 무엇인가?’, ‘다이어트를 왜 하기 힘든가?’, ‘야식은 끊을 수가 없나?’처럼 교육 프로그램 적용 초기에 비해 최종 보고서에서는 대학생이나 예비 코딩 강사들의 연령대에 한 번씩은 고민해 봤을 만한 다양하고 창의적인 주제들이 많이 나와서 교육 프로그램을 통한 창의성 향상에 따른 긍정적 효과라고 생각되었다.

(2) 단계별 데이터 과학 분석

문제점이나 궁금한 점을 해결하여 얻고 싶은 점과 문제점이나 궁금한 점에 대해 학습자들이 생각한 이유에 대해 학습자들 대부분이 나름대로 논리적으로 가설 설정을 하였고, 직접 작성한 구글 설문지 등을 사용하여 다양한 대상 상대로 자료를 수집하였다.

또한, 스크래치의 특징을 잘 활용하여 원본 데이터를 분석하여 다양한 그래프 등의 시각화 자료로 활용하여 한 눈에 이해하기 쉽게 제시하였고 개별 속성을 잘 분석하였다.

속성 간의 관계 분석하기에서는 앞서 초등학생을 대상으로 한 데이터 과학 교육 프로그램처럼 설문 대상이 가족·학교 친구 등에 국한되어 수집한 설문 표본 수가 작고 현상 일반화와 현상 예측을 하기에는 데이터 수집 기간이 짧아서 양질의 데이터를 수집하지 못하였기 때문에 상관관계 분석이 충분히 이루어지지 않는 못하였지만 상관관계에 있는 다양한 요인들을 나름 잘 찾아내고 제시하였다.

V. 결론 및 제언

본 연구는 초등학생, 대학생, 예비 코딩 강사를 대상으로 컴퓨팅 사고력과 창의성을 향상시키기 위하여 ADDIE 모형과 Dick & Carey 모형의 단계에 따라 데이터 과학 교육 프로그램을 개발·적용하였다.

요구 분석은 문헌 연구 및 선행연구 분석, 학습자에 대한 요구 분석을 실시하였다. 먼저 컴퓨팅 사고력, 창의성, 데이터 과학 교육에 대한 이론적 고찰을 통해 문헌 연구 및 선행연구 분석을 실시하였다. 다음으로 학습자 요구 분석은 Rossett의 요구 분석 모형에 따라 초등학생, 대학생, 교사를 대상으로 설문을 통한 조사연구를 수행하였다. 요구 분석 결과 다음과 같은 요구를 추출 및 정리하였다.

첫째, 초등학생, 대학생, 교사 모두 교육용 프로그래밍 언어에 대한 관심도가 가장 높게 나타나, 데이터 과학 교육 프로그램 교수 전략으로 교육용 프로그래밍 언어를 사용하는 것이 효과적일 것으로 분석되었다. 한편, 대학생이 상대적으로 소프트웨어교육 경험이 적은 것으로 나타나 대학생과 예비 코딩 강사를 대상으로 하는 교육 프로그램의 경우 초보자들이 다루기에 가장 쉬운 교육용 프로그래밍 언어를 활용할 필요가 있는 것으로 분석되었다.

둘째, 초등학생, 대학생, 교사 모두 데이터 과학 교육 프로그램을 통하여 컴퓨팅 사고력과 창의성 향상을 기대하는 것으로 나타났다. 따라서, 컴퓨팅 사고력과 창의성이 본 연구에서 개발된 교육 프로그램의 적용을 통해 향상되었는지에 대한 검증이 필요한 것으로 분석되었다.

셋째, 데이터 과학 교육 프로그램에 활용할 데이터에 대해 초등학생, 대학생, 교사에 따라 개인 데이터와 공공 데이터 사이에 선호도 차이가 뚜렷이 나타나 설계 단계에서 이를 고려한 설계가 필요한 것으로 분석되었다.

따라서, 요구 분석 내용을 바탕으로 데이터 과학 교육 프로그램을 설계 및 개발하였다.

설계 단계에서는 Dick & Carey 모형의 단계에 따라 성취 목표 진술, 교수

전략을 개발하였다.

개발 단계에서는 설계 단계의 내용을 바탕으로 Dick & Carey 모형의 단계에 따라 4가지의 데이터 과학 교육 프로그램에 대한 교수 매체 선정 및 개발을 하였다.

적용 단계에서는 데이터 과학 교육 도구 및 교육 대상에 따라 개발된 4가지의 데이터 과학 교육 프로그램을 총 5회에 걸쳐 초등학생, 대학생, 예비 코딩 강사에게 적용하여 프로그램의 교육적 효과를 검증하였다. 4가지의 교육 프로그램은 ‘엑셀을 활용한 초등학생 대상 데이터 과학 교육 프로그램(개인 데이터)’, ‘엑셀을 활용한 초등학생 대상 데이터 과학 교육 프로그램(공공 데이터)’, ‘앱인벤터를 활용한 초등학생 대상 데이터 과학 교육 프로그램’, ‘스크래치를 활용한 대학생 및 예비 코딩 강사 대상 데이터 과학 교육 프로그램’으로 구성되어 있다.

평가 단계에서는 컴퓨팅 사고력을 측정하기 위한 도구로 계산적 인지력 검사 A, B형(김병수, 2014)을 활용하였고, 창의성 검사 도구는 TTCT 검사지 도형 A, B형을 사용하였다.

평가 단계를 통해 분석된 본 연구의 데이터 과학 교육 프로그램 학습 효과는 다음과 같다.

첫째, 본 연구에서 개발한 ‘엑셀을 활용한 초등학생 대상 데이터 과학 교육 프로그램(개인 데이터)’, ‘엑셀을 활용한 초등학생 대상 데이터 과학 교육 프로그램(공공 데이터)’, ‘앱인벤터를 활용한 초등학생 대상 데이터 과학 교육 프로그램’은 모두 초등학생의 컴퓨팅 사고력과 창의성 향상에 효과적인 것으로 나타났다.

둘째, 본 연구에서 개발한 ‘스크래치를 활용한 대학생 및 예비 코딩 강사 대상 데이터 과학 교육 프로그램’은 대학생과 예비 코딩 강사의 창의성 향상에 효과적인 것으로 나타났다.

본 연구의 시사점은 다음과 같다.

첫째, 데이터 과학 교육 프로그램은 컴퓨팅 사고력과 창의성 향상에 효과적이다. 데이터 문제 정의하기(문제 탐색하기), 데이터 수집하기, 현상 이해하기(탐색적 데이터 분석), 현상 일반화하기, 현상 예측하기, 데이터 스토리텔링 등

직접적인 활동 중심의 데이터 과학 교육 문제 해결 과정을 통해 컴퓨팅 사고력과 창의성이 향상되었다.

둘째, 개인 데이터와 공공 데이터를 모두 데이터 과학 교육에 활용할 수 있음을 입증하였다.

셋째, 대표적인 통계용 프로그램인 ‘엑셀’ 뿐만 아니라 ‘엡인벤티’, ‘스크래치’와 같은 교육용 프로그래밍 언어를 데이터 과학 교육 프로그램의 도구로써 활용할 수 있음을 제시하였다.

넷째, 데이터 과학 교육 프로그램은 초등학생, 대학생, 예비 코딩 강사를 대상으로 그 효과가 입증되어, 앞으로 교육현장에서 다양한 연령층을 대상으로 활용할 수 있게 되었다.

본 연구의 결과 다음과 같이 제언하고자 한다.

첫째, 본 연구의 대학생 및 예비 코딩 강사 대상 데이터 과학 교육 프로그램은 창의성에 대한 검증만 이루어지고 컴퓨팅 사고력 향상에 대한 검증은 이루어지지 않았다. 컴퓨팅 사고력에 대해 많은 연구가 이루어지고 있으나 아직까지 컴퓨팅 사고력 향상 검증을 위한 타당도와 신뢰도가 갖춰진 검사 도구가 없기 때문이기도 하지만 본 연구 전반에 걸쳐 활용한 계산적 인지력 검사 A, B형(김병수, 2014)은 초등학생을 대상으로 하는 검사 도구이기 때문에 대학생 및 예비 코딩 강사를 대상으로 적용하기에는 무리가 있어 창의성에 대한 검증만 이루어졌다. 앞으로 다양한 연령층을 대상으로 컴퓨팅 사고력을 측정할 수 있는 검사 도구를 개발하고 적용해 보는 후속 연구가 진행되어야 할 것이다.

둘째, 본 연구의 5회에 걸친 데이터 과학 교육 프로그램 중 일부 교육 프로그램의 각 실험집단은 상관연구에 필요한 30명 이상의 참여자를 확보하지 못하여 실험결과의 상관관계를 분석할 수 없었으며 일반화하는 데에도 한계가 있다. 추후의 연구에서는 다수의 참여자를 대상으로 연구 결과에 대한 각 요인들 간의 상관관계를 분석할 필요가 있다고 생각된다.

인류가 처음으로 맞이하게 될 제2차 정보혁명이라고도 일컬어지는 제4차 산업혁명 시대는 인간이 천 년 동안 쉬지 않고 학습해야만 습득할 수 있는 바둑의 기보를 단숨에 학습해버리는 알파고, 인공지능 법률가 로스(ROSS), 알고리즘에 의해 8분 만에 작곡하는 인공지능 등 거의 모든 분야에서 인공지능이 인

간의 영역을 빠르게 대체해 나갈 것으로 여겨지고 있어 교육에 엄청난 변화를 요구하고 있다.

하지만, 아직 교육 현장에서는 이러한 엄청난 변화의 물결에 대응할 만큼 소프트웨어교육을 통한 미래 인재 양성 교육에 대한 인식과 관심, 준비가 충분하지 못한 것이 현실이다. 본 논문은 이러한 미래 인재 양성 교육이라는 소프트웨어교육의 목적 달성에 기여하고자 한, 컴퓨팅 사고력과 창의성 향상을 위한 데이터 과학 교육 프로그램에 대한 실험 연구로 4차 산업혁명을 효과적으로 주도할 학생들을 길러내는 데에 조금이라도 도움이 됐으면 하는 바람이다.

참 고 문 헌

- 강문봉 (2009). 컴퓨터를 활용한 수학적 사고력 신장 방안. *과학교육논총*, 22(1), 37-54.
- 강이철 (2011). *교육방법 및 교육공학의 입문*. 서울: 양서원.
- 강현영 (2012). 통계적 소양의 교육적 의미 고찰. *한국수학사학회지*, 25(4), 121-137.
- 구덕훈 (2007). *초등학교 5학년의 컴퓨터 음악 프로그램을 활용한 창작지도 방안 연구*. 석사 학위논문, 대구교육대학교 교육대학원.
- 권대용 (2011). *Algorithmic brick based tangible robot and hybrid programming environments for enhancing computational thinking*. 박사 학위논문, 고려대학교 대학원.
- 길순희 (2011). *컴퓨터를 활용한 학원 음악교육의 활성화방안*. 석사 학위논문, 제주대학교 교육대학원.
- 김동일 (2016). 스마트 제조를 위한 데이터사이언스-기법, 응용 및 이슈. *ie 매거진*, 23(3), 20-30.
- 김병수 (2014). *계산적 사고력 신장을 위한 PPS기반 프로그래밍 교육 프로그램*. 박사 학위논문, 제주대학교 대학원.
- 김병수, 김종훈 (2013). CPS 모형 기반 스크래치 프로그래밍 학습이 언어 창의성에 미치는 영향. *컴퓨터교육학회논문지*, 16(6), 11-19.
- 김병호 (2013). 안드로이드 앱 인벤터를 활용한 컴퓨터 프로그래밍 교육. *한국정보통신학회논문지* 17, 467-472.
- 김선태 (2011). *특성화고·마이스터고 수업혁신에 활용할 STEAM 수업 매뉴얼*. 한국 직업능력개발원.
- 김성범, 강성현 (2016). 제4차 산업혁명을 주도하는 데이터 사이언스. *ie 매거진*, 23(3), 9-13.
- 김성훈 (2010). *LOGO프로그래밍 언어가 초등학생의 창의성 발달에 미치는 영향*. 석사 학위논문, 제주대학교 교육대학원.

- 김영정 (2005). 한국의 미래를 위한 기반 교육으로서의 창의성과 비판적 사고 교육. *한국연구재단(NRF) 연구성과물*, 1-21.
- 김영채 (2010a). *TTCT(도형) A형 검사 요강*, 대구: 창의력 한국 FPSP.
- 김영채 (2010b). *TTCT(도형) B형 검사 요강*, 대구: 창의력 한국 FPSP.
- 김용민, 김종훈 (2015). 초등학교 여학생의 컴퓨팅적 사고력 신장을 위한 앱인벤터 활용 S/W 교육 프로그램 개발 및 적용, *정보교육학회논문지*, 19(4). 385-397.
- 김용민, 김종훈 (2017a). 스프레드시트를 활용한 데이터 과학 교육 프로그램이 초등 학생의 컴퓨팅 사고력 향상에 미치는 효과, *정보교육학회논문지*, 21(2). 219-230.
- 김용민, 김종훈 (2017b). 스크래치를 활용한 데이터 시각화 활동이 예비 코딩 강사의 창의성 향상에 미치는 효과, *정보교육학회논문지*, 21(3). 309-320.
- 김용민, 김태훈, 김종훈 (2015). 초등학교 여학생의 창의성 신장을 위한 로봇 활용 프로그래밍 교육 프로그램 개발 및 적용, *정보교육학회논문지*, 19(1). 31-44.
- 김유리, 문용은 (2015). 소프트웨어(SW)교육을 위한 공공 데이터 활용 방안에 관한 연구. *인터넷전자상거래연구*, 15(6), 245-261.
- 김윤석 (2005). *엑셀을 활용한 '확률과 통계'의 효율적인 교수·학습 과정안 연구*. 석사 학위논문, 아주대학교 교육대학원.
- 김은경, 이경화, 유경훈 (2010). 대학생들의 창의적 능력과 창의적 성향에 관한 인식 조사. *창의력교육연구*, 10(2), 5-26.
- 김은정 (2004). *Excel과 SPSS를 이용한 통계지도자료개발*. 석사 학위논문, 아주대학교 대학원.
- 김종진 (2011). *EPL을 이용한 창의성 증진 교육 프로그램 개발 및 적용에 관한 연구 : 로고와 스크래치를 중심으로*. 박사 학위논문, 홍익대학교 대학원.
- 김종진, 현동립, 김승완, 김종훈, 원유현 (2010). 교육용 프로그래밍 언어인 로고와 스크래치 교재 개발 및 비교 실험. *한국콘텐츠학회*, 10(7), 459-469.
- 김중혜 (2009). *정보과학적 사고 기반의 문제 해결 능력 향상을 위한 중등 교육 프로그램*. 박사 학위논문, 고려대학교 대학원.
- 김진영 (2016). *헬로 데이터 과학*. 서울: 한빛미디어.
- 김태훈 (2015). *컴퓨팅 사고력 신장을 위한 프로그래밍 중심 STEAM 교육 프로그램*. 박사학위 논문, 제주대학교 대학원.

- 김형철 (2011). *컴퓨터과학 교육용 계산 원리 학습도구의 기능요소 고찰*. 석사학위 논문, 제주대학교 대학원.
- 류미영, 한선관 (2015). 초등 SW교육을 위한 CT교육 프로그램 개발. *한국정보교육학회지*, 19(1), 11-20.
- 류충구 (2012). *Scratch가 초등 영재들의 창의적 문제해결력에 미치는 효과*. 석사학위논문, 경인교육대학교 교육대학원.
- 박경재 (2010). *EPL과 로봇 프로그램 교육의 창의성 신장 효과 분석*. 석사학위논문, 경인교육대학교 교육대학원.
- 박용철, 이수정 (2011). 스크래치 프로그래밍 교육이 초등학생의 자기 주도적 학습 능력에 미치는 효과. *정보교육학회논문지*, 15(1), 93-100.
- 박정호, 김철 (2011). 초등학교에서 로봇활용 미술수업이 창의성 신장에 미치는 효과. *정보교육학회논문지*, 15(2), 277-285.
- 박찬정, 강재희, 김명진, 유예림, 김효상, 고정욱 (2012). 앱인벤터를 이용한 실시간/비실시간 질의-응답 앱 설계 및 개발. *한국컴퓨터교육학회 학술발표대회논문집*, 16(2).
- 배혜진, 이동환 (2016). 통계적 문제해결 과정 관점에 따른 초등 수학교과서 통계 지도 방식 분석. *한국초등수학교육학회지*, 20(1), 55-69.
- 서영민, 이영준 (2010). 초등정보영재의 창의성 신장을 위한 교과 통합 로봇 프로그래밍수업 모형. *컴퓨터교육학회논문지*, 13(1), 19-26.
- 설문규, 손창익 (2013). 워터폴 모델을 적용한 앱 인벤터 프로그래밍 교재개발 연구. *정보교육학회논문지*, 17, 409-419.
- 성은지 (2014). *App Inventor를 활용한 초등학교 프로그래밍 교육 프로그램에 관한 연구*. 석사학위논문, 서울교육대학교 교육전문대학원.
- 손은경 (1995). *취학전 아동의 연령, 과제제시방법 및 과제유형에 따른 인과성 발달 연구*. 박사학위논문, 동아대학교 대학원.
- 송정미 (2011). *스크래치를 활용한 정보과학적 사고 기반 퍼즐 교육*. 석사학위논문, 고려대학교 교육대학원.
- 송정범, 조성환, 이태욱 (2008). 스크래치 프로그래밍 학습이 학습자의 동기와 문제해결력에 미치는 영향. *정보교육학회논문지*, 12(3), 323-332.

- 안경미, 손원성, 최윤철 (2011). 스크래치 프로그래밍 교육이 초등학생의 학습 몰입과 프로그래밍 능력에 미치는 효과. *정보교육학회논문지*, 15(1), 1-10.
- 안기완 (2003). *엑셀을 활용한 중학교 수학통계교육에 관한 연구*. 석사 학위논문, 인제대학교 교육대학원.
- 양혜진 (2003). *엑셀을 활용한 변수 개념 지도에 관한 연구*. 석사 학위논문, 한국교원대학교 대학원.
- 유중현, 김종혜 (2010). 문제 해결과정에서의 정보과학적 사고 능력에 대한 개념적 고찰. *정보창의교육*, 2(2), 15-24.
- 이강범 (2008). *'비판적 사고'에 대한 고찰과 그 활용: Delphi report를 중심으로*. 석사 학위논문, 연세대학교 교육대학원.
- 이광상 (2005). *엑셀을 통한 일차함수의 과정-대상관점 형성*. 박사 학위논문, 한국교원대학교 대학원.
- 이명호 (2016). 데이터 사이언스 교과과정에 대한 연구. *한국비블리아학회지*, 27(1), 263-290.
- 이민희 (2009). *두리틀을 이용한 프로그래밍 수업이 창의성, 문제해결력, 프로그래밍 흥미도 향상에 미치는 영향*. 석사 학위논문, 전주교육대학교 교육대학원.
- 이영준, 백성혜, 신재홍, 유현창, 정인기, 안상진, 최정원, 정성균 (2014). 초중등 단계 Computational Thinking 도입을 위한 기초 연구. 한국과학창의재단.
- 이은경 (2009). *Computational Thinking 능력 향상을 위한 로봇 프로그래밍 교수 학습 모형*. 박사 학위논문, 한국교원대학교 대학원.
- 이은경 (2013). 계산적 사고 향상을 위한 창의적 스크래치 프로그래밍 학습. *컴퓨터교육학회논문지*, 16(1), 1-9.
- 이은희, 김원경 (2015). 국내외 통계교육 연구동향 비교 분석. *수학교육*, 54(3), 241-259.
- 이점순 (2008). *LOGO프로그래밍 언어가 초등학생의 창의성 발달에 미치는 영향*. 석사 학위논문, 전주교육대학교 교육대학원.
- 이재무 (2014). Rossett 모형을 적용한 적응형 이러닝 시스템을 위한 요구 분석. *한국콘텐츠학회논문지*, 14(6), 529-538.
- 이재윤 (2015). 데이터 사이언스와 데이터 리터러시. *한국정보관리학회 학술대회 논문*

- 문집, 11-15.
- 이태욱 (2006). *마이크로 로봇 교육을 통한 초등학교 창의성 개발에 대한 연구*. 석사 학위논문, 제주교육대학교 교육대학원.
- 임화경 (2013). 초등학생을 대상으로 앱 인벤터를 활용한 안드로이드용 앱 제작교육. *멀티미디어학회논문지*, 18(7), 1495-1507.
- 전경원 (2000). *유아 종합 창의성 검사*. 서울: 학지사
- 전성해, 최준혁 (2015). 특허와 데이터사이언스, 그 연결은?. *한국지능시스템학회 학술발표 논문집*, 25(1), 3-4.
- 정재삼 (1998). 교육 프로그램 개발 모형의 분석. *교과교육학연구* 2(2), 80-97.
- 조성환, 송정범, 김성식, 이경화 (2008). CPS에 기반한 스크래치 EPL이 문제해결력과 프로그래밍 태도에 미치는 효과. *정보교육학회논문지*, 12(1), 77-88.
- 조준필 (2012). *기술교육에서 중학생의 논리적 사고력 함양을 위한 스크래치 학습 프로그램 개발*. 석사 학위논문, 한국교원대학교 대학원.
- 최윤식 (2012). *미래학자의 통찰법*. 서울: 김영사.
- 최재황 (2016). 정보리터러시 교육의 교수설계 모델 분석. *사회과학연구*, 27(4), 163-180.
- 최훈 (2010). 김영정 교수의 비판적 사고론. *논리연구*, 13(2), 1-26.
- 한선관, 김수환, 서정보 (2010). 스크래치 프로그래밍을 활용한 게임중독 치료 프로그램의 개발. *정보교육학회논문지*, 14(1), 61-68.
- 함성진, 김순화, 박세영, 송기상 (2014). 융합적사고력 신장을 위한 초등학생용 CT 기반 융합인재교육(CT-STEAM) 프로그램 개발. *컴퓨터교육학회논문지*, 17(6), 81-91.
- 함성진, 양창모 (2011). 스크래치를 이용한 초등학교 컴퓨터 교육과정 설계. *정보교육학회논문지*, 15(3), 413-423.
- 허경철, 김홍원, 임선하, 김명숙, 양미경 (1991). *사고력 신장을 위한 프로그램 개발 연구(V)*. 서울: 한국교육개발원.
- 황성진, 최정원, 이영준 (2013). 앱인벤터를 이용한 초등정보영재 앱 제작 교육프로그램 개발. *한국컴퓨터정보학회학술대회논문집*, 21(2).
- Bar, D., Harrison, J., & Conery, L. (2011) Computational thinking: A Digital

- Age Skill for Everyone. *Learning & Learning with Technology*, 38(6), 20-23.
- Barnes, D. J. (2002). Teaching introductory Java through LEGO MINDSTORMS models. *ACM SIGCSE Bulletin, Proceedings of the 33rd SIGCSE technical symposium on Computer science education*, 34(1), 147-151.
- Booth, C. (2011). *Reflective Teaching, Effective Learning: Instructional Literacy for Library Educators*. Chicago: American Library Association.
- Cleveland, W. S. (2001). Data Science: An Action Plan for Expanding the Technical Areas of the Field of Statistics. *International Statistical Review*, 69(1).
- Computer Science Teachers Association(CSTA), & International Society for Technology in Education(ISTE). (2011a). *Computational Thinking in K - 12 Education teacher resource second edition*. Retrieved from http://www.iste.org/docs/ct-documents/ct-teacher-resources_2ed-pdf.pdf?sfvrsn=2.
- Computer Science Teachers Association(CSTA), & International Society for Technology in Education(ISTE). (2011b). *Computational Thinking Teacher Resources*. Retrieved from <http://csta.acm.org/Curriculum/sub/CompThinking.html>.
- Conway, D. (2010). *The Data Science Venn Diagram*. Retrieved from <http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram>.
- Denning, P. J. (2010). Ubiquity symposium 'What is computation?': Opening statement. *Ubiquity*, 1.
- Feldhusen, J. F. (1983). The Purdue Creative Thinking Program. In I. S. Sato (Ed.), *Creativity research and educational planing*. Louisiana: National State Leadership Training Institute for the Gifted and Talented.
- Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report Pre-K-12*. American Statistical Association. Retrieved from http://www.amstat.org/education/gaise/GAISEPreK12_Intro.pdf.
- Gold, J. B. (1981). *Developing the creative problem solving skills of*

- intermediate age educable mentally retarded student*. Doctoral dissertation, Fordham University.
- Grover, S., & Pea, R. (2013). Using a discourse-intensive pedagogy and android's app inventor for introducing computational concepts to middle school students. *Proceeding of the 44th ACM technical symposium on Computer science education*, 723-728.
- Guilford, J. P. (1956). Structure of intellect, *Psychological Bulletin*, 53, 267-293.
- Guilford, J. P. (1967), *The nature of human intelligence*. New York : McGraw-Hill.
- Guo, P. (2010). *What is Computer Science? Efficiently Implementing Automated Abstractions*. Retrieved from <http://www.pgbovine.net/what-is-computer-science.htm>.
- Hsu, Y. C., & Ching, Y. H. (2013). Mobile app design for teaching and learning: Educators' experiences in an online graduate course. *The International Review of Research in Open and Distance Learning*, 14(4), 117-139.
- Jagadish, H. V. (2015). Big Data and Science: Myths and Reality. *Big Data Research*, 2(2), 49-52.
- Kim, Y. M., & Kim, J. H. (2016). Application of a Software Education Program Developed to Improve Computational Thinking in Elementary Schoolgirls, *The Indian Journal of Science and Technology*, 9(44), 1-9.
- Kurdek, L. A. (1977). Structural components and intellectual correlates of cognitive perspective taking in first-through fourth-grade children. *Child Development*, 48, 1503-1511.
- Morelli, R. (2010). *Can Android App Inventor Bring Computational Thinking to K-12?*. Computer Science Department, Trinity College Hartford, CT, USA.
- Morelli, R., de Lanerolle, T., Lake, P., Limardo, N., Tamotsu, E., & Uche, C. (2011). Can android app inventor bring computational thinking to k-12. In Proc. *42nd ACM technical symposium on Computer science education (SIGCSE'11)*.

- National Research Council. (2010). *Report of a Workshop on the Scope and Nature of Computational Thinking*. National Academies Press.
- National Research Council. (2011). *Report of a Workshop on the Pedagogical Aspects of Computational Thinking*. National Academies Press.
- Piotr, B. (2001). *Spreadsheet use in elementary statistics course*. Doctoral Dissertation, Teachers College, Columbia University.
- Resnick, M., Maloney, J., Monroy-Hernandez, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., & Silverman, B. (2009). Scratch: Programming for All 'Digital fluency' should mean designing, creating, and remixing, not just browsing, chatting, and interacting. *Communications of the ACM*, 52(11), 60-67.
- Romeike, R. (2007). Applying creativity in CS high school education: criteria, teaching example and evaluation. *Proceedings of the Seventh Baltic Sea Conference on Computing Education Research*, 8, 87-96. Australian Computer Society, Inc.
- Rossett, A. (1987). *Training needs assessment*. Englewood Cliffs, NJ: Educational Technology Publications.
- Runco, M. A. (2001). Introduction to the special issue: Commemorating Guilford's 1950 presidential address. *Creativity Research Journal*, 13(3-4), 245. doi: 10.1207/S15326934CRJ1334_01.
- Selby, C. C. (2012). Promoting computational thinking with programming. *Proceedings of the 7th Workshop in Primary and Secondary Computing Education*, 74-77.
- Shi, Y., P. S. Yu., Y. Zhu., & Y. Tian. (2014). Explore New Field of Data Science Under Big Data Era: Preface for ICDS 2014. *Procedia Computer Science*, 30, 1-3.
- Smith, P. K., & Trope, Y. (2006). You focus on the forest when you're in charge of the trees: power priming and abstract information processing. *Journal of personality and social psychology*, 90(4), 578.

- Stanton, J. M. (2013). *A Introduction to Data Science*. Syracuse University.
- Torrance, E. P. (1974). *Torrance tests of creativity thinking: Norms, technical manual*. Princeton, New Jersey: Personnel Press/Ginn.
- Torrance, E. P. (1982). Hemisphericity and thinking creative functioning, *Journal of Research and Development in Education*, 15, 29-37.
- Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review*, 110(3), 403-421.
- Walden, J., Doyle, M., Garns, R., & Hart, Z. (2013). An informatics perspective on computational thinking. *Proceedings of the 18th ACM conference on Innovation and technology in computer science education*, 4-9.
- Werner, L., Denner, J., Campe, S., & Kawamoto, D. C. (2012). The Fairy Performance Assessment: Measuring computational thinking in middle school. *Proceedings of the 43rd ACM technical symposium on Computer Science Education*, 215-220.
- Wing, J. M. (2006). Computational Thinking. *Communication of the ACM*, 49(3), 33-35.
- Wing, J. M. (2008). Computational Thinking and Thinking and Thinking about Computing. *Philosophical Transactions of the Royal Society*, 366, 3717-3725.

<ABSTRACT>

Data Science Education Program to Improve Computational Thinking and Creativity

Yongmin Kim

Major of Computer Education, Faculty of Science Education

Graduate School, Jeju National University

Supervised by professor Jonghoon Kim

The purpose of this study is to develop data science education programs with an aim of enhancing computational thinking and creativity and to evaluate the overall program.

Since the “AlphaGo Shock” marked the beginning of the fourth wave of industrial revolution, it has fundamentally altered the way people live and influenced them in every corner of their lives. Amid the big data deluge with unmeasurable and explosive growth of digital data, the transformation is unlike anything humankind had experienced during the first three industrial revolutions.

Against this backdrop, it becomes indispensable to enhance computational thinking and creativity of students with courseware(education software), which focuses on data collection, analysis and efficient problem solving to find answers for different questions. With data science education and “computers,” this will enable talents to cultivate their creativity, allowing decision making based on collected, analyzed and processed data and critical

evaluation required to survive in the Fourth Industrial Revolution. Recognizing their necessity, this present study focuses on the development of data science education programs to advance computational thinking and creativity of students.

The tools used in this study to develop the education programs include the ADDIE model and the Dick & Carey model.

In the initial stage of this study, the needs of students were defined and analyzed with elementary school students, university students and teachers based on the needs analysis model of Rossette and the outcome of analyzing advanced research and reviewing the literature on computational thinking, creativity and data science education.

Then, achievement goals and instructional strategies were established in the design phase in accordance with the Dick & Carey model.

Based on the outcome of the former stage, instructional media were selected and developed for four data science education programs in the development stage, following each phase of the Dick & Carey model. The programs were “Data Science Education Program with Excel for Elementary Students(personal data),” “Data Science Education Program with Excel for Elementary Students(open data),” “Data Science Education Program with App Inventor for Elementary Students,” and “Data Science Education Program with Scratch for University Students and Future Coding Instructors.”

Subsequently, the programs developed according to the defined data science education tools and targets were delivered to elementary students, university students and future coding instructors five times to validate the effects of the educational programs.

Lastly, in the evaluation stage, computational thinking and creativity tests were conducted to measure the skills.

The assessment results revealed that “Data Science Education Program

with Excel for Elementary Students(personal data),” “Data Science Education Program with Excel for Elementary Students(open data),” and “Data Science Education Program with App Inventor for Elementary Students” were adequate tools to improve computational thinking and creativity of elementary students. “Data Science Education Program with Scratch for University Students and Future Coding Instructors,” another program developed by this study, also contributed to enhancing the skills of the university students and the instructors.

The implications of this study are summarized as follows:

First, the data science education programs were effective in boosting computational thinking and creativity. Engaging in problem solving of the data science education programs centered on direct activities such as definition of data sets(problem-searching), collection of data, understanding (Exploratory Data Analysis), generalization of and prediction for phenomena and “data storytelling” enriched computational thinking and creativity.

In addition, it was confirmed that both personal and open data can be utilized for data science education.

Furthermore, this study demonstrated the possible practice and use of educational programming language such as “Scratch” or “App Inventor” as tools for data science education programs along with Excel, the most well-known statistical program.

Lastly, the effectiveness of the data science education programs for elementary students was proved for university students and future coding instructors; the means of the education can be introduced to any students regardless of their age.

A future study can further advance the data science education programs that were originally developed in this present study and enhance validity and confidence of verification tools to validate computational thinking. This will lead to systematic analysis of the relationships among the elements


and contributors to research findings on data science education programs to improve computational thinking and creativity.

Keywords: Computational thinking, Creativity, Data science education, ADDIE model, Needs analysis

부 록

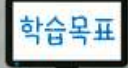
<부록 1> 엑셀을 활용한 데이터 과학 교육 프로그램(개인 데이터)	125
<부록 2> 엑셀을 활용한 데이터 과학 교육 프로그램(공공 데이터)	183
<부록 3> 애플리케이션을 활용한 데이터 과학 교육 프로그램	223
<부록 4> 스크래치를 활용한 데이터 과학 교육 프로그램	258

<부록 1> 엑셀을 활용한 데이터 과학 교육 프로그램(개인 데이터)

 Excel
 로 배우는

데
이
터
과
학

수업일 :
이름 :
1



학습목표




데이터 과학의 의미를 이해하고 이를 위한 엑셀의 기본 기능을 익힐 수 있다.


|
생각해 보기
|

마크 드랭솔트는 미국 워싱턴대학 치대 교수이다. 그는 취미로 칠인 3종 경기에 출전하곤 하는 활동적 인물이었지만 최근 몇 가지 건강문제를 겪었다. 마크가 선택한 해결책은 의사와의 상담에 더해 자신의 건강 상태에 대한 데이터를 직접 수집하여 분석하는 것이다. 그는 자신의 건강 상태를 개선하겠다고 마음먹고, 매일 몸무게와 체지방을 측정함과 동시에 한 번에 한 가지씩 식생활을 개선해 나갔다. 그 결과 그는 20대의 몸무게를 다시 회복하는 데 성공했다. 2008년 그는 경미한 심장 발작을 일으켰다. 그는 심장 발작이 올 때마다 그 직전에 어떤 일이 있었는지 발작의 강도는 어떠했는지 기록하기 시작했다. 이를 바탕으로 격렬한 운동이나 카페인 섭취가 이런 증상의 발생과 관계가 깊다는 것을 알게 되었고 이를 주치의와 공유하여 큰 수술 대신 간단한 절제 시술로 증상을 개선시킬 수 있었다(김진영, 2016).

☑ 데이터 과학이란?
데이터 과학의 의미 알기

1. 데이터 과학의 의미



2. 데이터 과학의 예







3. 데이터 과학을 시작하기 위해 필요한 것







4. 데이터 과학의 단계

가. 데이터 문제 정의하기







나. 데이터 수집하기










수업일 :

이름 :

다. 현상 이해하기: 탐색적 데이터 분석

 _____
 _____
 _____

라. 현상 일반화하기: 통계적 추론

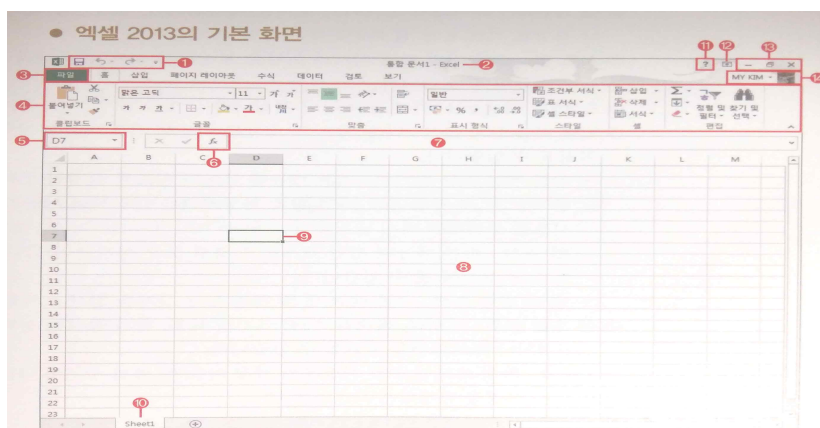
 _____
 _____
 _____



데이터 과학과 엑셀: 엑셀로 데이터 과학 맛보기-기본 기능 익히기

1. 엑셀 실행하기


시작-모든 프로그램-Microsoft Office-Microsoft Office 2013 클릭

2. 엑셀의 화면 구성



 Excel	
수업일 :	이름 :
	1

- ① **빠른 실행 도구 모음:** 사용자가 자주 사용하는 도구를 등록하여 클릭으로 간단히 실행함.
- ② **제목 표시줄:** 현재 작업 중인 엑셀 문서의 이름이 표시됨.
- ③ **파일 탭:** 저장, 열기 등 문서 관리 및 인쇄, 계정 등을 위한 공간. 클릭하면 전체 창으로 열림.
- ④ **리본 메뉴:** 작업에 필요한 명령 도구를 기능별로 묶음.
홈/삽입/페이지 레이아웃/수식/데이터/검토/보기 7가지 탭
- ⑤ **이름 상자:** 워크시트에서 현재 셀 포인터가 위치한 곳의 셀 주소 또는 지정된 이름 표시
- ⑥ **함수 삽입:** 클릭하면 함수 마법사 대화 상자가 나타나 사용할 함수를 선택할 수 있음.
- ⑦ **수식 입력줄:** 현재 셀에 입력된 내용이 표시되며, 데이터 또는 수식을 직접 입력할 수 있음.
- ⑧ **워크시트:** 데이터의 입력과 계산, 편집 등의 모든 문서 작업이 이루어지는 공간. 워크시트는 행과 열이 만나는 '셀(cell)'로 구성되어 있으며 화면의 크기를 확대 또는 축소할 수 있음.
- ⑨ **셀 포인터:** 데이터가 입력되는 위치. 원하는 위치의 셀을 클릭하거나 화살표 키를 눌러 셀 포인터를 이동할 수 있음.
- ⑩ **시트탭:** 문서 내의 모든 시트의 이름 표시. 시트 탭의 색과 시트 이름, 위치 등은 편집 가능함.
- ⑪ **도움말:** 엑셀 2013의 도움말 창이 열림.
- ⑫ **리본 메뉴 표시 옵션:** 리본 메뉴 자동 숨기기, 탭 표시, 탭 및 명령 표시의 메뉴를 이용해 리본 메뉴를 보이거나 숨기도록 설정할 수 있음.
- ⑬ **창 조절 단추:** 최소화, 최대화, 이전 크기로 복원, 닫기 중에서 선택해 창의 크기를 조절하거나 프로그램을 종료할 수 있음.
- ⑭ **사용자:** 현재 로그인 중인 사용자의 계정 이름 및 계정 사진이 나옴. 클릭해서 설정을 변경할 수 있음.

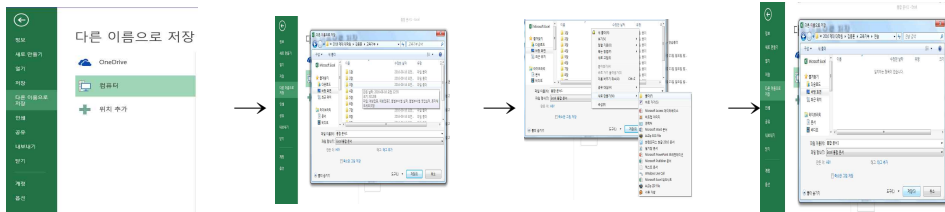


3. 문서 작성하고 저장하기

가. 문서 작성: 셀을 클릭해 셀 포인터가 위치하면 입력(또는 셀을 두 번 더블클릭)하고 'Enter' 키를 누르거나 다른 셀에 클릭

나. 저장하기

빠른 실행 도구 모음에서 '저장'버튼을 클릭하거나 '파일'탭의 '저장'을 클릭한 후



저장버튼 클릭-다른 이름으로 저장

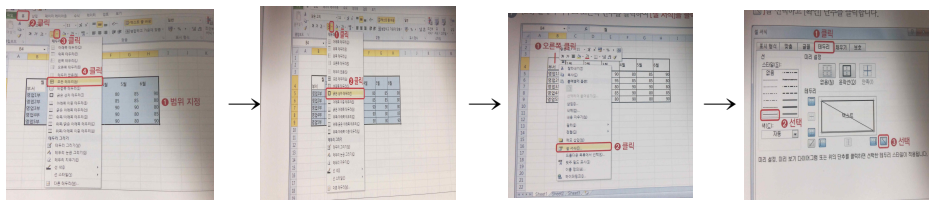
왼쪽 항목에 저장할 드라이브 선택-새 폴더 클릭

새 폴더가 생기면 폴더 이름을 입력하고 '여기' 클릭

파일 이름을 입력하고 '저장' 클릭-문서가 저장되어 제목 표시줄에 파일 이름이 나타남

4. 셀 서식 지정하기

가. 테두리 서식 지정하기



데이터 입력-홈-글꼴-테두리-모든 테두리 클릭

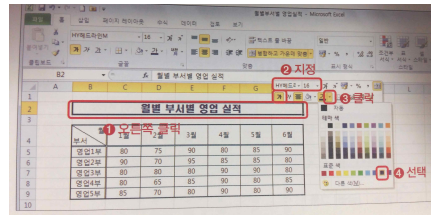
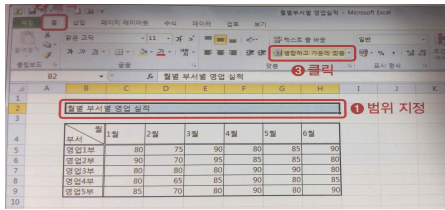
홈-글꼴-테두리-굵은 상자 테두리 클릭

해당 셀에서 오른쪽 버튼-셀 서식 클릭

테두리-선 스타일은 실선, 테두리에서 대각선 선택-확인



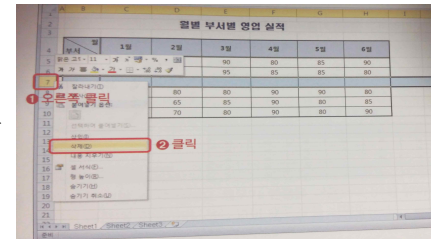
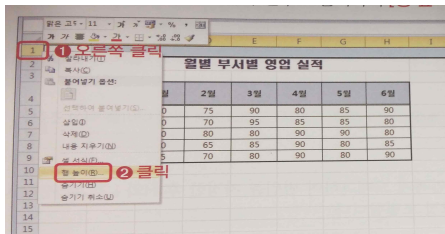
나. 글꼴과 맞춤 서식 지정하기



해당 영역 지정-홈 탭의 맞춤 그룹의 병합하고 가운데 맞춤 클릭

해당 영역 설정하고 오른쪽 버튼-글꼴

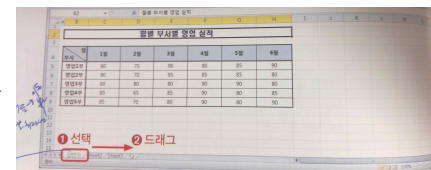
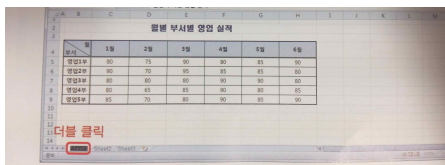
다. 행 높이와 열 너비 조절하고 행 삽입하기



관련 행 선택-오른쪽 버튼- 높이 클릭-숫자 기록-확인(열 너비도 같은 방법)

관련 행 선택-오른쪽 버튼-삽입(또는 삭제)

라. 워크시트 관리하기



시트 이름 바꾸기: 아래 sheet1 더블클릭-글자 입력하고 - Enter

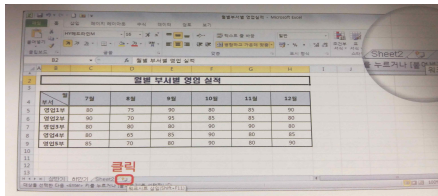
시트 복사: 시트 1을 누른 상태로 뒤로 드래그



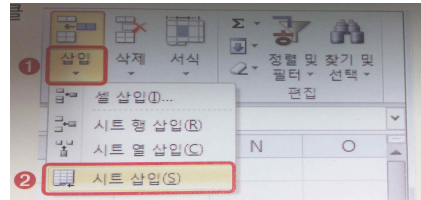
수업일 :

이름 :

1



시트 삽입 1: 워크시트 삽입 단추 클릭



시트삽입 2: 홈-셀 그룹의 삽입-시트삽입

마. 표 서식 / 조건부 서식 적용하기

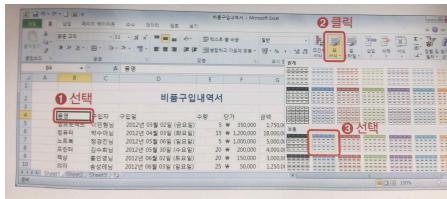
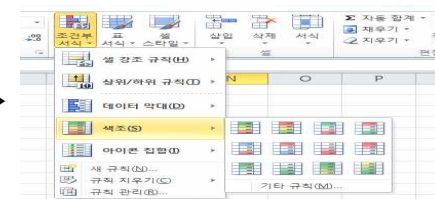
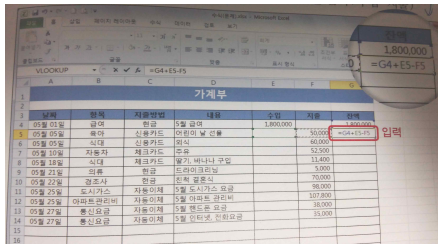


표 서식: 해당 영역 선택-홈-표 서식

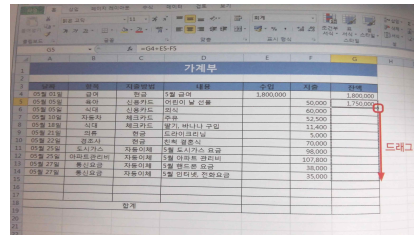


조건부 서식: 홈-조건부 서식-색조

5. 수식 입력하기



해당 셀에 수식 적기
(더하기+, 빼기-, 곱하기*, 나누기/)



드래그 하면 서식이 복사됨



합계 들어갈 셀 선택

영역 드래그-수식-자동합계

6. 차트 만들기

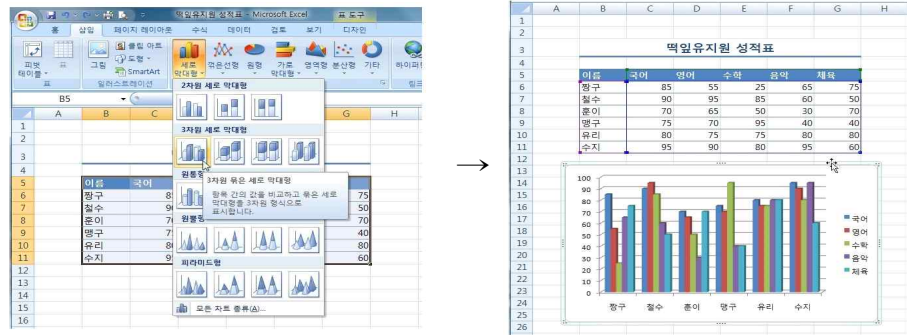
가. 차트의 구성 요소

- ① 차트 영역: 차트의 제목과 배경, 계열 등 모든 요소가 포함된 차트 전체의 영역
- ② 차트 제목: 차트의 제목을 글상자로 표시함.
- ③ 그림 영역: 차트의 계열이 표시되는 영역
- ④ 데이터 계열: 데이터의 계열 값들이 막대나 선, 도형 등으로 표시됨.
- ⑤ 데이터 레이블: 데이터의 계열 값을 표시함.
- ⑥ 눈금선: 그림 영역 안에서 계열이나 데이터 값을 표시하는 단위를 수직선 또는 수평선으로 구분해 표시함.



- ⑦ 축: 차트의 수평 방향(X축)과 수직 방향(Y축)에 각각의 항목이 나타내는 내용 또는 수치 표시
- ⑧ 축 제목: X축 또는 Y축에서 나타내는 내용을 제목으로 표시함.
- ⑨ 범례: 데이터의 계열이나 항목을 구분할 수 있도록 해당 색이나 무늬 등을 표시한 상자

가. 차트 삽입



영역 설정-삽입-차트 그룹 3차형 막은 세로 막대형 그래프

나. 차트 레이아웃과 서식 지정하기



차트 도구-디자인-차트레이아웃 차트 도구-레이아웃-레이블(제목 입력)





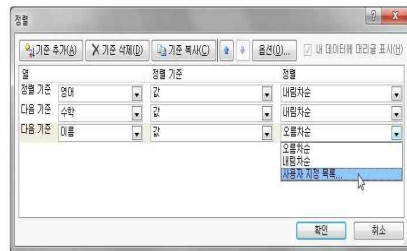
차트의 세로축 변경 (차트 오른 버튼-축 서식에서 주 단위 숫자 변경) 차트 도구-레이아웃-레이블-데이터 레이블-표시 (전체 숫자 표시됨, 일부분만 할 경우는 해당 그래프만 Ctrl을 사용하여 선택)

7. 데이터베이스 관리하기

가. 데이터 정렬 이용하여 오름차순, 내림차순, 사용자 지정 목록 순 정렬하기

떡잎유치원 성적표							
이름	국어	영어	수학	음악	체육		
장구	85	55	25	65	75		
철수	90	95	85	60	50		
훈이	70	65	50	30	70		
맹구	75	70	95	40	40		
유리	80	75	75	80	80		
수지	95	90	80	95	60		

이름	국어	영어	수학	음악	체육
장구	85	55	25	65	75
철수	90	95	85	60	50
훈이	70	65	50	30	70
맹구	75	70	95	40	40
유리	80	75	75	80	80
수지	95	90	80	95	60

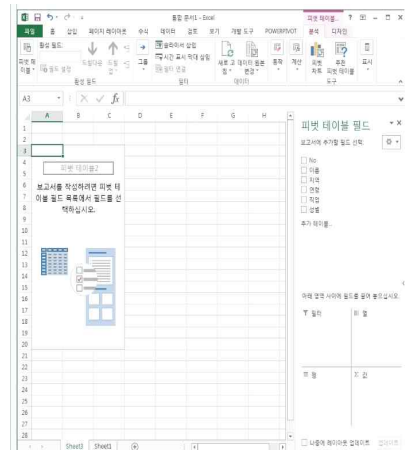
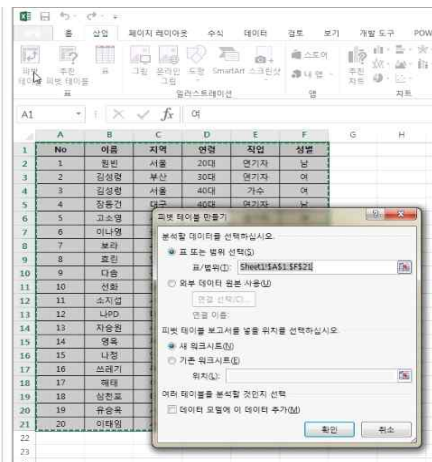


오름차순(1,2,3,4,5...)/내림차순(...5,4,3,2,1) 영역 지정-정렬-사용자 지정 목록



나. 피벗 테이블 이용하기

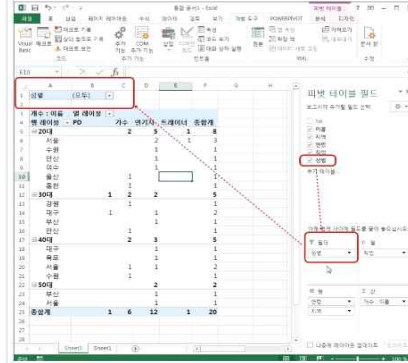
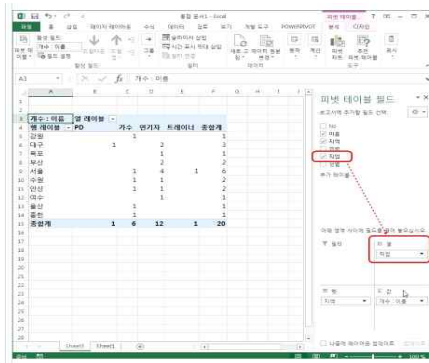
	A	B	C	D	E	F
1	No	이름	지역	연령	직업	성별
2	1	원빈	서울	20대	연기자	남
3	2	김성령	부산	30대	연기자	여
4	3	김성령	서울	40대	가수	여
5	4	장동건	대구	40대	연기자	남
6	5	고소영	서울	50대	연기자	여
7	6	이나영	울산	20대	가수	여
8	7	보라	수원	40대	가수	여
9	8	효린	안산	30대	가수	여
10	9	다솜	강원	30대	가수	여
11	10	선화	충천	20대	가수	여
12	11	소지섭	서울	40대	연기자	남
13	12	나PD	대구	30대	PD	남
14	13	자승원	수원	20대	연기자	남
15	14	영욱	부산	50대	연기자	여
16	15	나정	안산	20대	연기자	여
17	16	쓰레기	목포	40대	연기자	남
18	17	해태	여수	20대	연기자	남
19	18	삼천포	대구	30대	연기자	남
20	19	유승욱	서울	20대	트레이너	여
21	20	이태임	서울	20대	연기자	여



삽입-왼쪽에 있는 피벗 테이블 아이콘 클릭

피벗 테이블 필드가 보임





필요한 항목을 가져옴.
(합계: 이름, 행: 지역, 열: 직업)

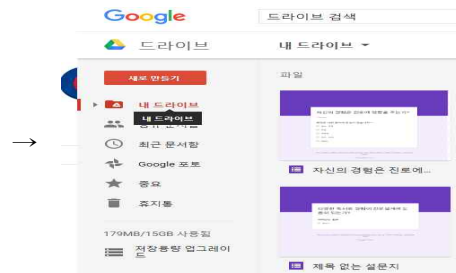
행에 연령 추가
남녀 성비로 나눠서 보고 싶으면
필터에 성별 선택한 후 원하는 성을
선택함.

☑ 데이터 과학과 설문: 구글을 이용하여 설문지 만들기

1. 네이버에 '구글'을 치거나 주소창에 <https://www.google.co.kr/> 를 친다.
2. 아이디와 비번을 넣어 구글 계정을 만든다. 아이디는 기존에 사용하는 아이디를 넣는다.

(내 아이디: _____, 내 비번: _____)

3. 로그인한다.



4. 드라이브를 찾아 클릭한다.

5. 왼쪽 '새로 만들기' 를 클릭한다.

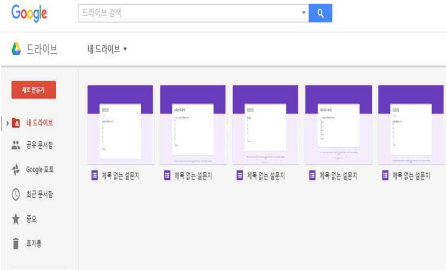


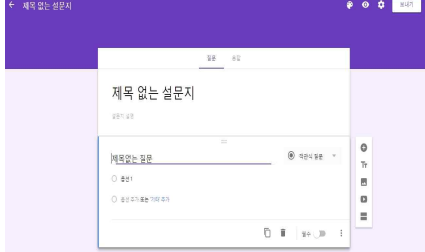
Excel 로 배우는 데 이 터 과 학

수업일 :

이름 :

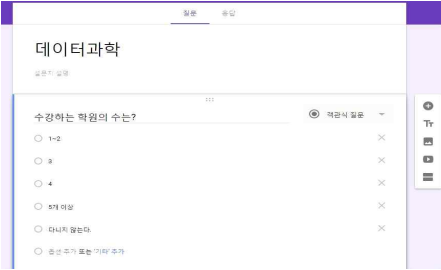
1






6. 구글 설문지를 클릭한다.

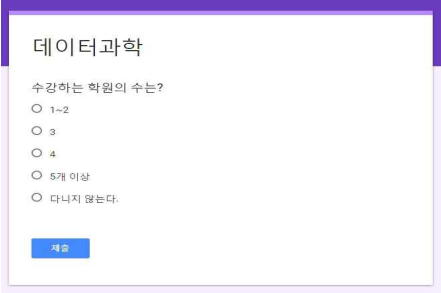
7. 설문지 제목을 입력한다.






8. 설문지 응답 내용을 입력한다.


9. 설문지를 보낸다.(링크시킨다.)





10. 응답자는 주소를 클릭하면 질문에 응답할 수 있다.

11. 응답자가 전부 응답하면 설문 결과를 볼 수 있다.





로 배우는



수업일 :

이름 :

1



★ 과제는 다음과 같이 올립니다.

네이버-창의컴퓨터 교실 가입-왼쪽 탭 10기-데이터 과학 ‘개별’-글쓰기-제목에 8월 16일 과제 (~초 000), 아이디는 실명(원래 이름)-내용에는 구글 주소를 붙임

★ 과제 1

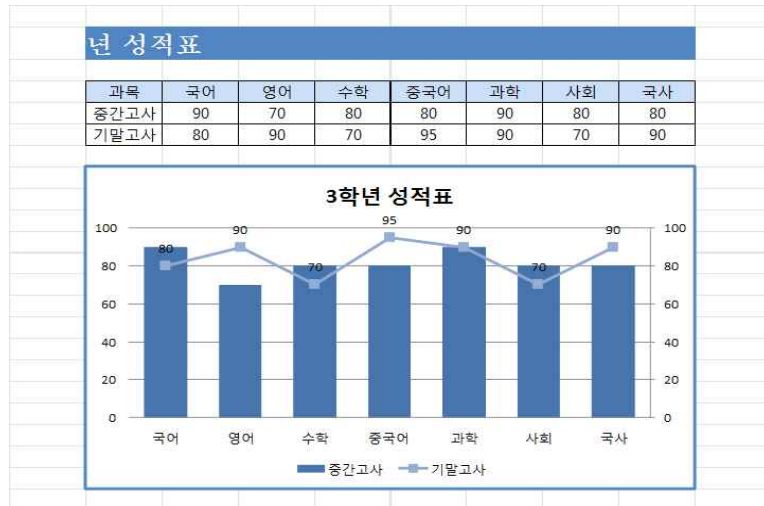
데이터 과학에 대한 내용을 주제로 정하여, 구글 설문지로 객관식 한 문항 (오지선다)을 만들고 완성한 설문지의 주소를 카페에 올리시오.

예시) 1. 데이터 과학의 의미에 대해 어느 정도 알고 있습니까?

- ① 매우 잘 알고 있다. ② 잘 알고 있다. ③ 보통이다. ④ 잘 모른다.
- ⑤ 전혀 모른다.

★ 과제 2

엑셀 프로그램을 사용하여 아래와 같은 표와 차트를 완성하고 카페에 올리시오.



학습목표

스스로 데이터를 수집하여 나에게 안전사고가 자주 일어나는 이유에 알아보고, 안전한 생활을 위해 노력해야 할 점을 실천할 수 있다.

생각해 보기

안전초등학교 4학년 나사랑 학생은 부모님께서 어렵게 얻은 4대 독자이다. 친구들과도 잘 어울리고 공부도 잘하는 나사랑 학생은 학교에서 인기도 많고 부모님과 선생님의 사랑을 듬뿍 받고 있다. 그런데 얼마 전 복도에서 달리다가 앞에서 오던 6학년 선배와 부딪혀 넘어지는 바람에 오른쪽 손목을 다치게 되었다. 병원에서는 인대가 늘어났다고 보호 차원에서 손목 교정용 깁스를 해주었다. 일주일 간 오른쪽 손목을 쓰지 못하게 되는 일은 여간 불편한 일이 아니었고 나사랑 학생은 이제와 자신의 행동을 후회하게 되었다. 생각해 보니 학교에서 지내는 시간 동안 큰 사고만 없었을 뿐이지 한 달에 두 세 번씩 작은 사고들은 많았었다. 나사랑 학생은 왜 이러한 사고들이 자꾸 생겼는지에 대해 깊은 생각에 잠기게 되었다.

☑ 데이터 문제 정의하기: 학교에서 나에게 자주 ()가 일어난다.

1. 문제 정의하기

- 목표: 나에게 ()가 자주 일어나는 ()를 알아보고, () 생활을 위해 노력해야 할 점을 실천할 수 있다.



- 범위: ()에서 일어나는 안전사고
- 성공 및 실패의 척도: 안전사고 횟수를 한 달에 1회 이하로 줄인다.
- 문제 해결의 제약 조건(시간과 비용): 한 달, 비용은 무료.

2. 데이터 정의하기

- 데이터의 제한: 나의 학교 생활 및 학교에서 발생하는 안전사고로 제한한다.
- 데이터를 어떻게 수집할 것인가: () (워드 또는 한글 문서 사용)
- 데이터 처리 및 분석을 위한 최적의 방법과 도구: () 프로그램을 활용한 분석 및 시각화
- 최종 결과물은 어떤 형태로 누구에게 전달되어야 하는가: 결과보고서, ()

3. 연구 설문 및 가설 정의하기

- 연구 설문(무엇을 알아내고 싶은가?): 학교에서 안전사고가 자주 일어나는 (), () 방법
- 가설: 안전사고가 자주 일어나는 이유는 ()



☑ 데이터 수집하기: 필요한 데이터 스스로 만들기

1. 데이터 수집의 기본

- 관찰결과를 개별 항목으로 하는 테이블로 만들기
- 각 열은 개별 항목이 갖는 속성, 각 행은 수집단위가 되는 개별 항목

○ 관련된 항목 생각해 보기 ○

 _____

 _____

 _____

2. 수집방법 결정하기(수동수집-만들어 가는 데이터)

○ 항목별 체크리스트 예시 ○

나의 안전생활 체크리스트						
일자 (일주일)	2016년 월 주	사고 횟수	회			
항목	내용	평 가				
		5 (매우 그렇다)	4 (자주 그렇다)	3 (보통 이다)	2 (거의 그렇지 않다.)	1 (전혀 그렇지 않다.)
행동	1. 물건을 함부로 던진다. 2. 잘 뛰어다닌다. 3. 계단이나 난간에 잘 걸터앉는다.					
마음가짐	1. 안전수칙을 지키는 것은 귀찮다.					
안전수칙 인식	1. 안전수칙을 잘 모른다.					



3. 데이터 품질 점검하기

- 완전성(문제 해결에 필요한 내용이 들어 있느냐?)
- 정확성(문제가 되는 상황이 들어 있느냐?)
- 일관성(데이터 항목들이 연관이 있고 일치하느냐?)

4. 데이터 준비하기

- 테이블 형태로 변환하기

안전사고, 왜 나에게는?							
내용	사고횟수	물건 던지기	뛰어다님	걸터앉음	지키는 것 귀찮음	안전수칙 잘 모름	
기간							
1주	1	3	4	1	3	2	
2주	2	3	4	2	3	2	
3주	3	4	5	1	4	2	
4주	4	5	5	2	5	2	

()을 열에 배치한다.

5. 데이터 추가, 선택, 집계하기

- 데이터 추가 및 선택은 필요시에 함.
- 집계는 피벗테이블을 사용하는 것이 좋으나 탐색적 데이터 과정에서 사 용함.



☑ 현상 이해하기: 탐색적 데이터 분석

1. 원본 데이터 잘 살펴보고 시각화 시키기

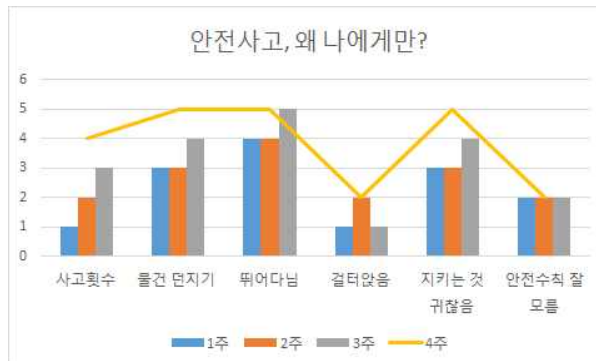
(데이터를 전체적으로 살펴보고 시각화를 통해 나타나는 패턴을 한눈에 살펴봄)

- 원본 데이터 잘 살펴보기(엑셀 홈-조건부 서식 사용)

안전사고, 왜 나에게만?						
기간 \ 내용	사고횟수	물건 던지기	뛰어다님	걸터앉음	지키는 것 귀찮음	안전수칙 잘 모름
1주	1	3	4	1	3	2
2주	2	3	4	2	3	2
3주	3	4	5	1	4	2
4주	4	5	5	2	5	2

○ 알 수 있는 점 ○

- 시각화 시키기(패턴 살피기)



○ 알 수 있는 점 ○

 _____
 _____
 _____



○ 알 수 있는 점 ○

 _____
 _____
 _____



2. 개별 속성 분석하기

- 피벗 테이블 만들기

행 레이블	합계 : 사고횟수
4	3
5	7
총합계	10

행 레이블	합계 : 사고횟수
3	3
4	3
5	4
총합계	10

합계 : 사고횟수	열 레이블			
행 레이블	3	4	5	총합계
4	3			3
3	3			3
5		3	4	7
4		3		3
5			4	4
총합계	3	3	4	10

○ 알 수 있는 점 ○

 _____
 _____
 _____



3. 속성간의 관계 분석하기(의미 있는 상관관계 찾기)

	사고횟수	물건던지기	뛰어다님	걸터앉음	지키는 것 귀찮음	안전수칙잘 모름
사고횟수	1					
물건던지기	1	1				
뛰어다님	0.866025	0.8660254	1			
걸터앉음	0	0	-0.5	1		
지키는 것 귀찮음	1	1	0.866025	0	1	
안전수칙잘 모름	0.234523	0.234523	0.234523	0.151426	0.234523	1

○ 알 수 있는 점 ○

 _____
 _____
 _____



현상 일반화하기: 통계적 추론

1. 내가 세운 가설을 대상으로 비교 및 결론 내리기

○ 결론 ○

 _____

 _____

 _____

2. 앞으로의 예측(노력해야 할 점 포함)

○ 예측 ○

 _____

 _____

 _____

결과 소통: 발표하고 공유하기

- 서로 의견 나누기



☑ 데이터 과학 보고서: 오늘 실습한 내용을 바탕으로 보고서 써보기

1. 주제 선택 및 탐구 동기

2. 탐구과정

가. 문제 정의

- 문제:

- 목표:

- 가설:

나. 데이터 수집

- 방법:

- 항목(내용):



다. 탐색적 데이터 분석 방법(실제 완성 때는 차트 등도 삽입해야 함. 이번은 연습)

- 원본 데이터 살펴보고 시각화:
- 개별 속성 분석:
- 속성간의 관계:

라. 통계적 추론(결론 및 예측, 짧게 간추려서)

- 결론:
- 예측:

3. 느낀 점(소감)





Excel

로 배우는



수업일 :

이름 :

2



★ 과제는 다음과 같이 올립니다.

네이버-창의컴퓨터 교실 가입-왼쪽 탭 10기-데이터 과학 '개별' -글쓰기

★ 과제 1

개인 데이터를 이용하여 본인이 해결하고 싶은 주제 1가지를 정한 후, 그와 관련된 구글 설문지를 만드시오. 설문지의 문항은 3문항~5문항 정도로 객관식 문항으로 만드시오. 만들고 난 뒤에 완성한 설문지의 주소를 카페에 올리시오.

예시) 주제가 나의 식습관과 관련된 것이라면

1. 나는 매일 패스트푸드 음식을 먹는다.
 - ① 매우 잘 알고 있다. ② 잘 알고 있다. ③ 보통이다.
 - ④ 잘 모른다. ⑤ 전혀 모른다.

★ 과제 2

한글 프로그램을 이용하여 다음의 내용이 들어가도록 한글 문서를 작성하여 카페에 올리시오.(주제는 과제 1번과 동일함. 목요일 오전 수업에 사용할 내용이므로 오래 생각해서 잘 정리해 올 것)

◇ 들어갈 내용: 제목 (번호 없이 처음 시작 가운데에)

1. 주제 선택 및 탐구 동기
2. 탐구 과정 중 가. 문제 정의 나. 데이터 수집

◇ 양식은 자유임. (여백, 글자체, 그림 삽입, 표 등은 자유. 내용만 잘 들어가면 됨)





스스로 데이터를 수집하여 학원은 학업성적 향상에 도움이 되는지 알아보고, 바람직한 학습 방법에 대해 생각해 볼 수 있다.

생각해 보기

으뜸초등학교 6학년 학생 부진아는 요즘 고민이 많다. 5학년이 될 때까지 본인이 원하지 않으면 학원에 가지 않아도 된다던 부모님 덕에 방과 후에는 취미활동도 하고 친구들과 어울리며 자유로이 시간을 보냈는데 6학년이 올라오면서 부모님 생각이 바뀐 것이다. 1년 후면 중학교에 진학해야 하니 지금부터라도 학원을 다니며 꾸준히 실력을 쌓아야 한다는 부모님 말에 진아는 동의할 수 없다. 진아네 반에서 1, 2등을 다투는 친구는 나일등과 최최고인데 나일등은 밤 11시까지 4개의 학원을 다니는 반면, 최최고는 1개의 학원도 다니지 않는다. 주변 친구들도 살펴보니 공부를 잘하는 친구 중에는 학원을 많이 다니는 친구도 있고, 그렇지 않은 친구들도 있었다. 진아는 문득 학원과 학업성적이 어떤 관계가 있는지, 학원은 학업성적에 정말 도움이 되는지 궁금해졌다.

☑ 데이터 문제 정의하기

1. 문제 정의하기

- 목표: ()
- 범위: ()
- 문제 해결의 제약 조건(시간과 비용): 즉시 또는 전날, 비용은 무료.



2. 데이터 정의하기

- 데이터의 제한: 학원은 교육과정 내에서 ()평가와 직접적 관련이 있는 것들로 제한한다(국수사과영).
- 데이터를 어떻게 수집할 것인가? () 설문
- 데이터 처리 및 분석을 위한 최적의 방법과 도구: 엑셀 프로그램을 활용한 분석 및 시각화
- 최종 결과물은 어떤 형태로 누구에게 전달되어야 하는가? 결과보고서, 보호자와 학생

3. 연구 설문 및 가설 정의하기

- 연구 설문(무엇을 알아내고 싶은가?): ()
- 가설: ()

데이터 수집하기: 필요한 데이터 스스로 만들기

1. 데이터 수집의 기본

- 관찰결과를 개별 항목으로 하는 테이블로 만들기
- 각 열은 개별 항목이 갖는 속성, 각 행은 수집단위가 되는 개별 항목

○ 관련된 항목 생각해 보기 ○

 _____

 _____

 _____





로 배우는



수업일 :

이름 :

3

2. 수집방법 결정하기(수동수집-만들어가는 데이터)

○ 설문지 내용 예시 ○

1. 본인은 몇 학년인가요?

- ① 4학년 ② 5학년 ③ 6학년

2. 수강하는 학원의 수는?

- ① 다니지 않는다 ② 1개 ③ 2개 ④ 3개 ⑤ 4개 이상

3 하루 평균 학원에서 보내는 시간은?

- ① 없다. ② 1시간 ③ 2시간 ④ 3시간 ⑤ 4시간 이상

4. 학업 성취도 평가 평균 점수는? (올해 1학기 기준)

- ① 80점미만 ② 80~85점 미만 ③ 85~90점 미만 ④ 90~95점미만
- ⑤ 95점 이상

5. 학원을 다니기 전후 성적변화는?

- ① 떨어짐 ② 변화 없음 ③ +1~+4점 ④ +5~+9점 ⑤ +10점 이상

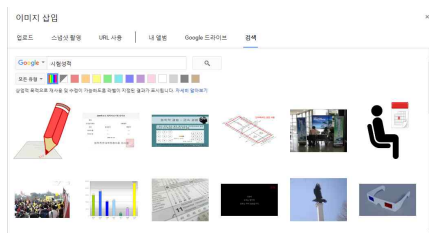
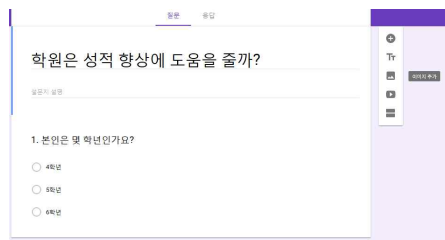
6. 학원을 다니면 성적에 도움이 된다고 생각하나요?

- ① 도움이 된다. ② 도움이 되지 않는다.



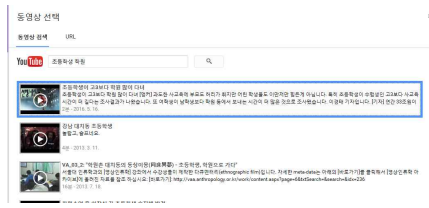
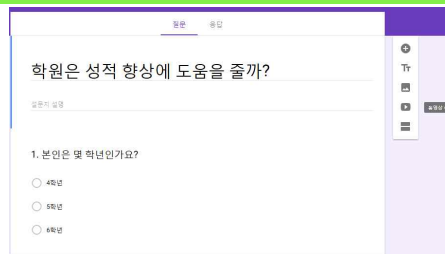
▣ 구글을 사용한 설문

- 이미지 및 동영상 추가, 전체 배경 바꾸기



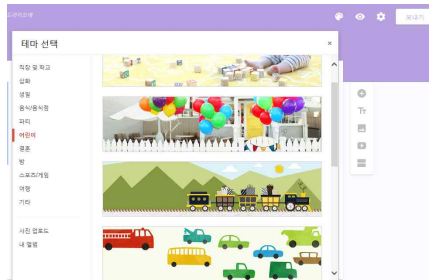
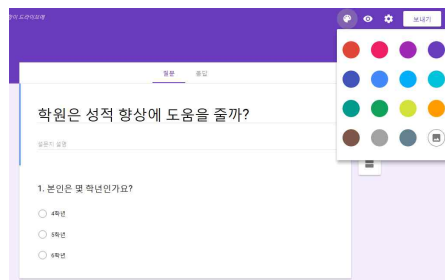
구글 설문 작성 중 오른쪽 메뉴-이미지 추가

검색 등을 하여 관련 이미지 찾기-선택



구글 설문 작성 중 오른쪽 메뉴-동영상 추가

동영상 검색 후-선택

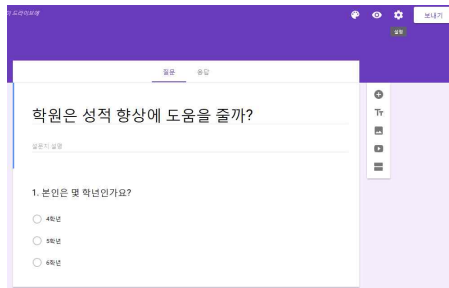


색상 팔레트-색깔 또는 이미지 가져오기

이미지 가져오기-선택 (전체배경 바뀜)



- 설정 바꾸기(1인당 1번만 설문에 응시할 수 있게 하여 오류 방지, 응답자가 결과 확인 가능)



위 탭에서 설정 클릭 응답횟수 1회 제한, 요약 차트와 텍스트 응답 확인에 체크

- 엑셀 자료 가져오기(응답한 결과 보기 및 결과를 엑셀 파일에 가져오기)



설문지 응답 클릭-결과 확인 오른쪽 우측 스프레드시트에서 응답보기 클릭



성적다운

연도	월	일	요일	점수	과목	합계
2016.1	1	1	월	100	영어	100
2016.1	1	2	화	100	영어	100
2016.1	1	3	수	100	영어	100
2016.1	1	4	목	100	영어	100
2016.1	1	5	금	100	영어	100
2016.1	1	6	토	100	영어	100
2016.1	1	7	일	100	영어	100
2016.1	1	8	월	100	영어	100
2016.1	1	9	화	100	영어	100
2016.1	1	10	수	100	영어	100
2016.1	1	11	목	100	영어	100
2016.1	1	12	금	100	영어	100
2016.1	1	13	토	100	영어	100
2016.1	1	14	일	100	영어	100
2016.1	1	15	월	100	영어	100
2016.1	1	16	화	100	영어	100
2016.1	1	17	수	100	영어	100
2016.1	1	18	목	100	영어	100
2016.1	1	19	금	100	영어	100
2016.1	1	20	토	100	영어	100
2016.1	1	21	일	100	영어	100

구글 설문지 엑셀파일로 열림 파일-다른 이름으로 다운로드-마이크로소프트 엑셀 선택

설문지 결과

응답자	성별	연령	응답률	응답 시간
응답자1	남	20대	100%	1분
응답자2	여	30대	100%	1분
응답자3	남	40대	100%	1분
응답자4	여	50대	100%	1분
응답자5	남	60대	100%	1분
응답자6	여	70대	100%	1분
응답자7	남	80대	100%	1분
응답자8	여	90대	100%	1분
응답자9	남	10대	100%	1분
응답자10	여	11대	100%	1분
응답자11	남	12대	100%	1분
응답자12	여	13대	100%	1분
응답자13	남	14대	100%	1분
응답자14	여	15대	100%	1분
응답자15	남	16대	100%	1분
응답자16	여	17대	100%	1분
응답자17	남	18대	100%	1분
응답자18	여	19대	100%	1분
응답자19	남	20대	100%	1분
응답자20	여	21대	100%	1분

다른 이름으로 저장하기-폴더 선택 엑셀 2013에서 파일 열기-필요 없는 자료 삭제, 항목 알맞게 바꾸기

3. 데이터 품질 점검하기

- 완전성(문제 해결에 필요한 내용이 들어 있는냐?)
- 정확성(문제가 되는 상황이 들어 있는냐?)
- 일관성(데이터 항목들이 연관이 있고 일치하느냐?)



4. 데이터 준비하기

- 테이블 형태로 변환하기
- 열에 넣을 항목: ()
- 행에 넣을 항목: ()

★ 스스로 해 보기 1★ 설문 내용을 테이블 형식으로 만들어 보세요.

1. 본인도 몇 학원인가?	2. 수강하는 학원의 수는?	3. 하루 평균 학원에서 보내는 시간은?	4. 학업 성취도 평가 평균 점수는? (올해 1학기)	5. 학원을 다니기 전후 성적 변화는?	6. 학원을 다니면 성적이 도움이 될까요?
6학년	0개	0시간	90-95점 미만	변화없음	그렇지 않다
5학년	0개	0시간	95점 이상	변화없음	그렇지 않다
6학년	1개	3시간	85점 이상	1-4점	그렇다
4학년	2개	3시간	90-95점 미만	5-9점	그렇지 않다
6학년	2개	2시간	85점 이상	5-9점	그렇다
4학년	1개	2시간	95점 이상	변화없음	그렇지 않다
5학년	0개	0시간	85점 이상	변화없음	그렇지 않다
6학년	1개	2시간	95점 이상	변화없음	그렇지 않다
6학년	0개	0시간	85점 이상	변화없음	그렇지 않다
6학년	4개 이상	3시간	95점 이상	1-4점	그렇다
5학년	3개	2시간	85점 이상	1-4점	그렇다
4학년	4개 이상	4시간 이상	85-90점 미만	1-4점	그렇다
6학년	1개	2시간	85점 이상	변화없음 1-4점 5-9점	그렇다, 그렇지 않다
6학년	2개	2시간	80-85점 미만	말어짐	그렇지 않다
6학년	3개	3시간	90-95점 미만	5-9점	그렇다

학년	학원 수	학원에서 보내는	평가 점수	성적변화	학원은 도움이 되
6학년	0	0	92	0	그렇지 않다
5학년	0	0	95	0	그렇지 않다
6학년	1	3	95	2	그렇다
4학년	2	3	92	7	그렇지 않다
6학년	2	2	95	7	그렇다
4학년	1	2	95	0	그렇지 않다
5학년	0	0	95	0	그렇지 않다
6학년	1	2	95	0	그렇지 않다
6학년	0	0	95	0	그렇지 않다
6학년	4	3	95	2	그렇다
5학년	3	2	95	2	그렇다
4학년	4	4	87	2	그렇다
6학년	1	2	95	0	그렇다, 그렇지 않다
6학년	2	2	82	-2	그렇지 않다
6학년	3	3	92	7	그렇다



5. 데이터 추가, 선택, 집계하기

- 데이터 추가 및 선택은 필요시에 함.
- 집계는 피벗테이블을 사용하는 것이 좋으나 탐색적 데이터 과정에서 사용함.

☑ 현상 이해하기: 탐색적 데이터 분석

1. 원본 데이터 잘 살펴보고 시각화 시키기

(데이터를 전체적으로 살펴보고 시각화를 통해 나타나는 패턴을 한눈에 살펴봄)

- 원본 데이터 잘 살펴보기(엑셀 홈-조건부 서식 사용)

- ★ 스스로 해 보기 2★ 조건부 서식을 사용하여 원본 데이터를 살펴보세요.

학년	학원 수	학원에서 보내는	평가 점수	성적변화	학원은 도움이 되
6학년	0	0	92	0	그렇지 않다
5학년	0	0	95	0	그렇지 않다
6학년	1	3	95	2	그렇다
4학년	2	3	92	7	그렇지 않다
6학년	2	2	95	7	그렇다
4학년	1	2	95	0	그렇지 않다
5학년	0	0	95	0	그렇지 않다
6학년	1	2	95	0	그렇지 않다
6학년	0	0	95	0	그렇지 않다
6학년	4	3	95	2	그렇다
5학년	3	2	95	2	그렇다
4학년	4	4	87	2	그렇다
6학년	1	2	95	0	그렇다, 그렇지 않다
6학년	2	2	82	-2	그렇지 않다
6학년	3	3	92	7	그렇다





로 배우는



수업일 :

이름 :

3

○ 알 수 있는 점 ○

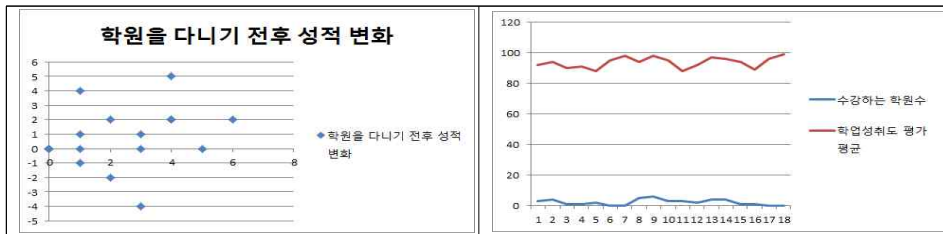
 _____

 _____

 _____

- 시각화 시키기(패턴 살피기)

★ 스스로 해 보기 3 ★ 데이터를 시각화시켜 보세요.



○ 알 수 있는 점(본인의 자료를 보고 작성해 볼 것) ○

 _____

 _____

 _____



2. 개별 속성 분석하기

- 피벗 테이블 만들기

★ 스스로 해 보기 4 ★ 학원수와 평균 점수, 학원에서 보내는 시간과 평균 점수, 학원수와 성적이 오른 정도를 피벗테이블로 나타내 보세요.



○ 알 수 있는 점 ○



3. 속성간의 관계 분석하기(의미 있는 상관관계 찾기)

	수강하는 학원 수	하루 평균 학원에서 보내는 시간	학업성취도 평가 평균	학원을 다니기 전후 성적변화
수강하는 학원 수	1			
하루 평균 학원에서 보내는 시간	0.976506947	1		
학업성취도 평가 평균	0.075887301	0.035163401	1	
학원을 다니기 전후 성적변화	0.25036523	0.182119455	0.387807788	1

○ 알 수 있는 점 ○

 _____

 _____

 _____

현상 일반화하기: 통계적 추론

1. 내가 세운 가설을 대상으로 비교 및 결론 내리기

○ 결론 ○

 _____

 _____

 _____



2. 앞으로의 예측(노력하고 싶은 점, 앞으로 추가해서 분석해 보고 싶은 것 등)

○ 예측 ○



결과 소통: 발표하고 공유하기

- 느낀 점 서로 이야기하며 공유하기



데이터 과학 보고서: 오늘 실습한 내용을 바탕으로 보고서 써보기

1. 주제 선택 및 탐구 동기

2. 탐구과정

가. 문제 정의

- 문제:

- 목표:

- 가설:

나. 데이터 수집

- 방법:

- 항목(내용):

다. 탐색적 데이터 분석 방법(차트 등도 삽입. 표를 한글에 붙이고 싶으면 Ctrl+ C(복사)→엑셀 붙여넣기 메뉴에서 그림으로 저장하기→다시 Ctrl+ C (복사)→한글에서 Ctrl+ V)



- 원본 데이터 살펴보고 시각화:
- 개별 속성 분석:
- 속성간의 관계: (이번은 그림 생략, 내용 정리만 할 것)

라. 통계적 추론(결론 및 예측, 짧게 간추려서)

- 결론:
- 예측:

3. 느낀 점(소감)





로 배우는

데

이

터

과

학

수업일 :

이름 :

3



★ 과제는 다음과 같이 올립니다.

네이버-창의컴퓨터 교실 가입-왼쪽 탭 10기-데이터 과학 '개별' -글쓰기

★ 과제 1

주변 사람들을 대상으로 데이터를 만들어 분석하고 싶은 주제 1가지를 정한 후, 그와 관련된 구글 설문지를 만드시오. 설문지의 문항은 3문항 ~ 5문항 정도로 객관식 문항으로 만드시오. 만들고 난 뒤에 완성한 설문지의 주소를 카페에 올리시오.

★ 과제 2

한글 프로그램을 이용하여 다음의 내용이 들어가도록 한글 문서를 작성하여 카페에 올리시오. (주제는 과제 1번과 동일함. 양식은 카페 참고.)

계획서			
주제			
학교		이름	
<주제 선정 이유(동기) 및 내용>			
<문제해결 방향/계획 기술 - 탐구과정, 자료수집 방법, 분석방법 등을 중심으로>			





로 배우는



수업일 :

이름 :

4



스스로 데이터를 수집하여 독서가 창의력 신장에 도움을 주는지 알아볼 수 있다.

생각해 보기

으뜸초등학교 5학년 학생 독서노가 가장 싫어하는 활동 중 하나는 책읽기다. 어렸을 때부터 엄마는 독서가 좋다며 시간이 날 때마다 책을 읽으라고 하셨지만 독서노는 책 읽기가 너무 귀찮아도망치기 일쑤였다. 그런데 독서노가 독서를 해야만 하는 일이 발생했다. 새 학기가 되어 만난 담임 선생님 독서광이 책 읽기의 중요성을 강조하며 일주일에 한권의 책을 읽는 과제를 낸 것이다. 선생님은 독서가 창의력에 많은 영향을 주신다고 하셨다. 독서노가 주변 친구들도 살펴보니 창의적인 친구들 중에 독서를 잘하는 친구도 있었고 그렇지 않은 친구도 있었다. 독서노는 정말 독서가 이해력, 창의력에 많은 영향을 주는지 알고 싶어졌다.

☐ 데이터 문제 정의하기 ()

1. 문제 정의하기

- 목표: ()
- 범위: 초등학교()학년 학생
- 문제 해결의 제약 조건(시간과 비용): 즉시 또는 전날, 비용은 무료.



2. 데이터 정의하기

- 데이터를 어떻게 수집할 것인가: _____ 으로 _____ 를(을) 대상으로 함.
- 데이터 처리 및 분석을 위한 최적의 방법과 도구: (_____) 프로그램을 활용한 분석 및 시각화
- 최종 결과물은 어떤 형태로 누구에게 전달되어야 하는가?: 결과보고서, 보호자와 학생

3. 연구 설문 및 가설 정의하기

- 연구 설문(무엇을 알아내고 싶은가?): (_____)
- 가설: (_____)

☑ 데이터 수집하기: 필요한 데이터 스스로 만들기

1. 데이터 수집의 기본

- 관찰결과를 개별 항목으로 하는 테이블로 만들기
- 각 열은 개별 항목이 갖는 속성, 각 행은 수집단위가 되는 개별 항목

★ 스스로 해 보기 1★ 관련된 항목 생각하기(어떤 내용이 들어가야 될까요?)

2. 수집방법 결정하기(수동수집- 만들어가는 데이터)

★ 스스로 해 보기 2★ 설문지 내용을 완성해 보세요.





Excel

로 배우는



수업일 :

이름 :

4

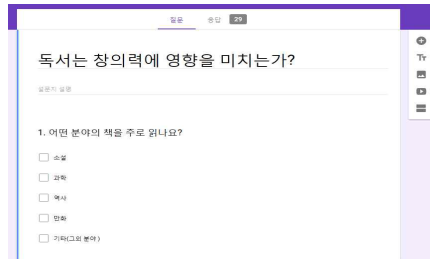
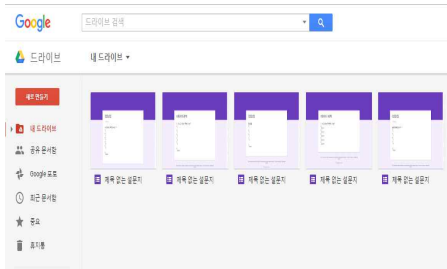
○ 설문지 내용 예시 ○

1. 본인은 몇 학년인가요? ① 4학년 ② 5학년 ③ 6학년
2. 일주일 평균 몇 권 정도의 책을 읽나요?
① 0~1권 ② 2권 ③ 3권 ④ 4권 ⑤ 5권 이상
3. 주로 어떤 분야의 책을 읽나요?
① 동화(소설) ② 과학 ③ 역사 ④ 만화 ⑤ 기타(그 외 분야)
4. 책을 읽을 때 내용을 생각하며 자세히 읽는 편인가요?
① 매우 그렇다 ② 그렇다 ③ 보통이다 ④ 그렇지 않다 ⑤ 매우 그렇지 않다
5. 문제를 해결할 때 다른 친구들이 생각하지 못한 방법으로 잘 해결하는 편인가요?
① 매우 그렇다 ② 그렇다 ③ 보통이다 ④ 그렇지 않다 ⑤ 매우 그렇지 않다

※ 그 외에 책을 읽는 방법에 따른 이해력을 알고 싶다면 책을 읽는 방법(모르는 낱말이 나올 때 어떻게 하는지, 한권의 책을 몇 번 읽어보는지, 어떤 종류의 책을 읽는지, 읽을 때 내용을 생각하며 읽는지)과 이해력(모르는 내용을 들었을 때 이해가 빨리 되는지)를 물어보는 것도 좋은 방법입니다.

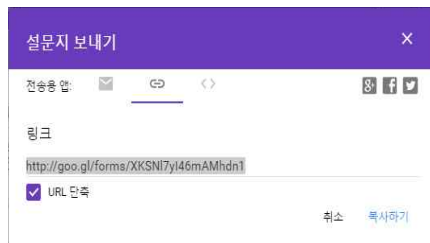
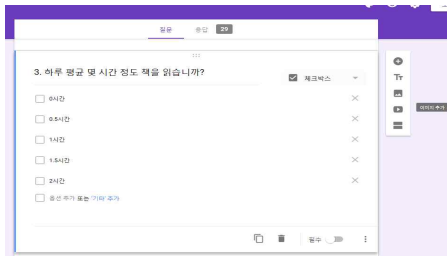


구글을 사용한 설문



구글 설문 제작하기

설문지 제목 입력하기



설문지 응답 입력하기

설문지 보내기

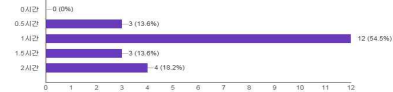
5. 문제를 해결할 때 다른 친구들이 생각하지 못한 방법으로 잘 해결하는 편이가요?

- 매우 그렇다
- 그렇다
- 보통이다
- 그렇지 않다
- 매우 그렇지 않다



→

3. 하루 평균 몇 시간 정도 책을 읽습니까? (응답 22회)



설문 응답하기(학생)

설문 결과 분석하기(교사)

3. 데이터 품질 점검하기

- 완전성(문제 해결에 필요한 내용이 들어 있느냐?)
- 정확성(문제가 되는 상황이 들어 있느냐?)
- 일관성(데이터 항목들이 연관이 있고 일치하느냐?)





로 배우는



수업일 :

이름 :

4

4. 데이터 준비하기

- 테이블 형태로 변환하기

- 열에 넣을 항목: ()

- 행에 넣을 항목: ()

★ 스스로 해 보기 3★ 설문 결과 내용을 테이블 형식으로 만들어 보세요.

독서는 창의력에 도움을 주는가?(예시)					
학년	내용	권 수	분야	읽는 스타일	창의력
4		2	동화	보통이다	그렇다
4		3	과학	그렇다	매우 그렇다
5		2	역사	그렇다	그렇다
5		4	역사	매우 그렇다	그렇다
6		5	기타	매우 그렇다	그렇다
6		1	만화	아니다	매우 아니다
4		1	동화	보통이다	보통이다
5		3	만화	보통이다	아닌다
6		1	기타	그렇다	그렇다
4		1	역사	매우 그렇다	매우 그렇다
5		4	만화	매우 아니다	보통이다
6		3	만화	아니다	아니다

독서는 창의력에 도움을 주는가?(예시)					
학년	내용	권 수	분야	읽는 스타일	창의력
4		2	동화	3	4
4		3	과학	4	5
5		2	역사	4	3
5		4	역사	5	3
6		5	기타	5	3
6		1	만화	2	1
4		1	동화	3	3
5		3	만화	3	2
6		1	기타	4	3
4		1	역사	5	5
5		4	만화	1	3
6		3	만화	2	2



5. 데이터 추가, 선택, 집계하기

- 데이터 추가 및 선택은 필요시에 함.
- 집계는 피벗테이블을 사용하는 것이 좋으나 탐색적 데이터 분석 과정에서 나오므로, 내용 중복으로 인해 탐색적 데이터 분석 과정에서 사용하겠음.

☑ 현상 이해하기: 탐색적 데이터 분석

1. 원본 데이터 잘 살펴보고 시각화 시키기

(데이터를 전체적으로 살펴보고 시각화를 통해 나타나는 패턴을 한눈에 살펴봄)

- 원본 데이터 잘 살펴보기(엑셀 홈-조건부 서식 사용)
- ★ 스스로 해 보기 4★ 조건부 서식을 사용하여 원본 데이터를 살펴보세요.

독서는 창의력에 도움을 주는가?(예시)				
학년 \ 내용	권 수	분야	읽는 스타일	창의력
4	2	동화	3	4
4	3	과학	4	5
5	2	역사	4	4
5	4	역사	5	4
6	5	기타	5	4
6	1	만화	2	1
4	1	동화	3	3
5	3	만화	3	2
6	1	기타	4	4
4	1	역사	5	5
5	4	만화	1	3
6	3	만화	2	2

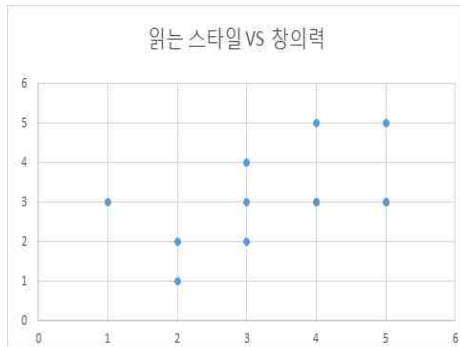
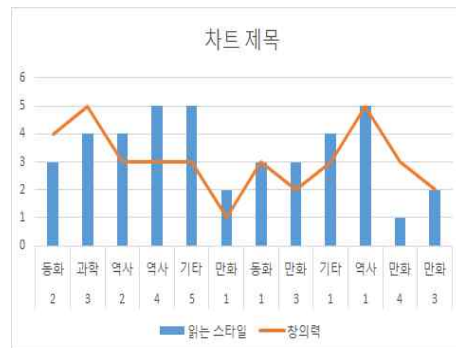
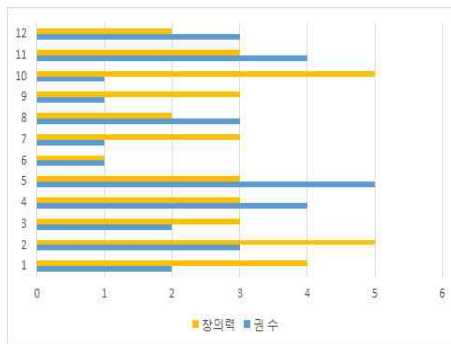


○ 알 수 있는 점 ○

Three horizontal lines for writing, each preceded by a colored pencil icon (blue, red, green).

- 시각화 시키기(패턴 살피기)

★ 스스로 해 보기 3 ★ 데이터를 아래 자료를 참고하여 시각화시켜 보세요.



○ 알 수 있는 점(본인의 시각화 자료를 보고 작성해 볼 것) ○

 _____

 _____

 _____

2. 개별 속성 분석하기

- 피벗 테이블 만들기

★ 스스로 해 보기 4 ★ **알아보고 싶은 항목을 선택하여 피벗테이블로 나타내 보세요.**

개수: 사람수	권수								
창의력	1	2	3	4	5	총합계			
1	1								1
2			2						2
3	2	1		2	1				6
4		1							1
5	1		1						2
총합계	4	2	3	2	1				12

○ 알 수 있는 점(본인의 피벗 테이블 자료를 보고 작성해 볼 것) ○

 _____

 _____

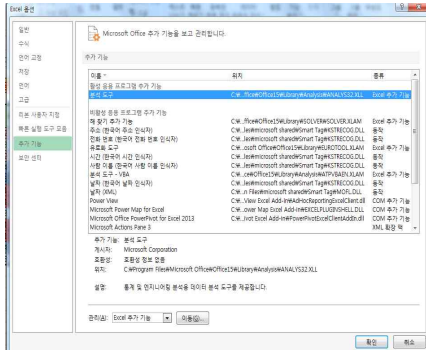
 _____



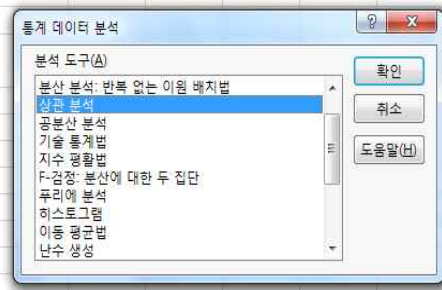
3. 속성간의 관계 분석하기(의미 있는 상관관계 찾기)

□ 분석 방법

1) 파일-옵션-추가기능-분석도구 선택하고 확인



2) 데이터-데이터 분석도구-상관 분석 확인



3) 해당 표 영역 선택- 확인 클릭

	권수	읽는 스타일	창의력
권수	1		
읽는 스타일	0.07525915	1	
창의력	0.079583716	0.722055783	1

○ 알 수 있는 점 ○



현상 일반화하기: 통계적 추론

1. 내가 세운 가설을 대상으로 비교 및 결론 내리기

○ 결론 ○

 _____

 _____

 _____

2. 앞으로의 예측

(따라서 앞으로 ~한다면 ~게 될 것으로 생각되므로 ~기 위한 ~에 대한 노력이 필요하겠다. 이번 조사를 바탕으로 이런 부분을 추가해서 더 알고 싶다 등)

○ 예측 ○

 _____

 _____

 _____



☑ 결과 소통: 발표하고 공유하기

- 느낀 점 서로 이야기하며 공유하기

☑ 데이터 과학 보고서: 오늘 실습한 내용을 바탕으로 보고서 써보기



※ 홈페이지 '데이터 과학 보고서' 양식 참고하여 직접 만든 차트와 표, 그림 등을 붙여서 정리한 후 창의컴퓨터 교실 카페에 제목 쓰고 첨부할 것.

★ 과제 1

주변 사람들을 대상으로 데이터를 만들어 분석하고 싶은 주제 1가지에 따른 설문 문항을 선생님의 조언을 참고하여 구글로 작성합니다. 작성한 주소를 올리고 참여해 달라는 글을 남기세요.

(전부 저녁 8시 전까지 올립니다.)

★ 과제 2

과제 1을 하고 난 뒤, 저녁 8시가 되면 카페에 들어가 친구들의 올린 글을 방문하여 설문을 응시한 후 댓글로 '00초 000 응시함' 이라고 글을 남깁니다. 반 친구 설문을 모두 해 주고 댓글로 전부 달아주어야 합니다. 좀 번거롭겠지만 친구들의 발표를 위해서 협조해 주세요!!

(같은 설문은 한 번만 하고 제출해 주세요. 여러 번 하면 신뢰도가 떨어져요~)





로 배우는



수업일 :

이름 :

5



평소에 내 자신이나 또는 내 주변에서 고쳤으면 하는 점이나 궁금한 점을 생각하여 주제를 정하고 스스로 데이터를 수집하여 이를 해결할 수 있다.

생각해 보기

평소에 내 자신이 고쳤으면 하는 점이나 궁금한 점은 무엇이 있을까? 또는 내 주변에서 고쳤으면 하는 점이나 궁금한 점은 무엇이 있을까? 생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 보자.

주제 정하기

나와 내 주변에서 일어나는 일을 중심으로 주제 선정하기

1. 내 자신이 고쳤으면 하는 점이나 궁금한 점을 생각해 보자.

○ 예시 ○

- 나는 왜 공부하는 시간이 많은데 성적이 안 오를까?
- 나는 왜 많이 먹어도 살이 안 찌까?
- 나는 왜 조금만 먹어도 살이 찌까?
- 나는 왜 배탈이 잘 날까?
- 나는 왜 아침마다 늦잠을 잘까?
- 나는 왜 친구들과 자주 싸울까?



2. 나와 관련된 일 중 고쳤으면 하는 점 또는 궁금한 점이 있으면 생각해 보자.

○ 예시 ○

- 악기 연주는 스트레스 해소에 도움을 줄까?
- 바른 자세로 앉으면 건강(공부)에 도움이 될까?
- 나이가 젊은 선생님이 우리들을 더 잘 이해해 줄까?
- 여자일수록 수다가 많을까?
- 엄한 선생님일수록 숙제를 잘 해 올까?
- 형제가 많으면 성격이 좋을까? (다른 사람들을 더 잘 배려하고)
- 남자들은 여자보다 무서움이 없을까?
- 성적이 좋을수록 스트레스가 적을까?

3. 자신이 탐구해 보고 싶은 주제 및 그 주제를 결정한 이유를 써 보자.

(★ 붙임 보고서에 정리)

☑ 문제 정의: 필요한 데이터 스스로 만들기

1. 정한 주제에 대해 문제, 목표, 가설을 정해 보자.

(★ 붙임 보고서에 정리)

- 문제(문제점이나 궁금한 점이 무엇인가?)
- 목표(문제점이나 궁금한 점을 해결하여 얻고 싶은 점은 무엇인가?)
- 가설(문제나 궁금한 점에 대해 내가 생각한 이유는 무엇인가?)



☑ 데이터 수집: 필요한 데이터 스스로 만들기

1. 데이터 수집 방법을 정해 보자.(★ 붙임 보고서에 정리)

- 체크리스트 등을 사용하여 내 자신의 누적 자료로 수집할 것인가?
- 구글 설문지 등을 사용하여 내 주변에서 자료를 수집할 것인가?

2. 데이터를 수집하기 위한 항목을 정해 보자.

(★ 붙임 보고서에 정리)

- 항목은 5개~8개 사이로 정한다.
- 항목은 주제 및 가설과 관련 있도록 정하는 것이 좋다.
- 항목은 될 수 있으면 수치화 시킬 수 있도록 한다.

(상관분석의 경우 수치로 되어야만 나타낼 수 있으며, 차트의 경우에도 수치화 되어 있는 경우가 자료를 나타내기 쉽다.)

3. 데이터를 직접 수집해 보자.

- 체크리스트 또는 구글 설문지 등을 통하여 자료를 수집한다.

(★ 붙임 보고서에 정리)

4. 데이터 준비하기

- 테이블 형태로 변환하기(★ 붙임 보고서에 정리)
- 열과 행 중 열에 항목의 내용을 적어준다.

- 자료 중 수치화 시킬 수 있는 것들을 수치화 시킨다.

(매우 그렇다 5점, 그렇다 4점 등)



☑ 현상 이해하기: 탐색적 데이터 분석

1. 원본 데이터 잘 살펴보고 시각화 시키기

(★ 붙임 보고서에 정리)

- 원본 데이터 잘 살펴보기(엑셀 홈-조건부 서식 사용)
- 시각화 시키기(패턴 살피기)

2. 개별 속성 분석하기(★ 붙임 보고서에 정리)

- 열, 행, 값에 어떤 항목을 넣으면 좋을지 생각하고 피벗테이블을 작성한다. (가설을 기준으로)
- 피벗테이블을 분석할 때는 관계를 중심으로 하여 살펴본다.

3. 속성간의 관계 분석하기(★ 붙임 보고서에 정리)

- 수치화 된 자료만 선택하여 상관분석으로 나타낸다.

☑ 현상 일반화하기: 통계적 추론

1. 내가 세운 가설을 대상으로 비교 및 결론 내리기

(★ 붙임 보고서에 정리)

2. 앞으로의 예측(★ 붙임 보고서에 정리)

☑ 결과 소통: 발표하고 공유하기

- 이번 발표를 통해 느꼈던 점, 또 더 추가해서 알아보고 싶은 점 등에 대해 서로 이야기를 나눈다.

(★ 붙임 보고서에 정리)



내가 만드는 데이터 과학 보고서

주제				제출일	
학교		번호		이름	

< 주제 선정 이유(동기) 및 내용 >

< 단계별 데이터 과학 분석 >

◆ 1단계-문제 정의 ◆

1. 문제:
2. 목표:
3. 가설:

◆ 2단계-데이터 수집 방법 ◆

1. 수집 방법
 - 대상자:
 - 어떤 방법:
2. 설문지 내용 (적어도 좋고, 구글 설문지 그림을 붙여도 됨)

◆ 3단계-탐색적 데이터 분석 ◆

1. 원본데이터 살펴보고 시각화하기(조건부 서식, 차트 그림 넣고 설명)
2. 개별 속성 분석하기(피벗 테이블, 피벗차트 넣고 설명)
3. 속성간의 관계 분석하기(상관도 그림 넣고 설명, 할 수 있는 사람만 하세요)

◆ 4단계-현상 일반화하기, 통계적 추론 ◆

1. 결론:
2. 통계적 추론:

< 소감 >





★ 과제는 다음과 같이 올립니다.

네이버-창의컴퓨터 교실 가입-왼쪽 탭 10기-데이터 과학
'개별' -글쓰기

★ 과제 1


완성한 보고서를 카페에 올립니다. (양식은 최종보고서 양식을 사용할 것)

★ 과제 2

작성한 보고서의 내용 중 필요한 부분을 간추려 큰 목소리로 또박또박 발표하는 연습합니다.(시간은 5분 이내)



<부록 2> 엑셀을 활용한 데이터 과학 교육 프로그램(공공 데이터)

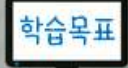
 Excel
 로 배우는

데
이
터
과
학

수업일 :

이름 :

1



학습목표


데이터 과학의 의미를 이해하고 이를 위한 엑셀의 기본 기능을 익힐 수 있다.


| 생각해 보기 |


마크 드랭솔트는 미국 워싱턴대학 치대 교수이다. 그는 취미로 칠인 3중 경기에 출전하곤 하는 활동적 인물이었지만 최근 몇 가지 건강문제를 겪었다. 마크가 선택한 해결책은 의사와의 상담에 더해 자신의 건강 상태에 대한 데이터를 직접 수집하여 분석하는 것이다. 그는 자신의 건강 상태를 개선하겠다고 마음먹고, 매일 몸무게와 체지방을 측정함과 동시에 한 번에 한 가지씩 식생활을 개선해 나갔다. 그 결과 그는 20대의 몸무게를 다시 회복하는 데 성공했다. 2008년 그는 경미한 심장 발작을 일으켰다. 그는 심장 발작이 올 때마다 그 직전에 어떤 일이 있었는지 발작의 강도는 어떠한지 기록하기 시작했다. 이를 바탕으로 격렬한 운동이나 카페인 섭취가 이런 증상의 발생과 관계가 깊다는 것을 알게 되었고 이를 주치의와 공유하여 큰 수술 대신 간단한 절제 시술로 증상을 개선시킬 수 있었다(김진영, 2016).


☑ 데이터 과학이란? 데이터 과학의 의미 알기

1. 데이터 과학의 의미









2. 데이터 과학의 예

 _____

 _____

 _____

3. 데이터 과학을 시작하기 위해 필요한 것


 _____


 _____


 _____

4. 데이터 과학의 단계

가. 데이터 문제 정의하기

 _____

 _____

 _____

나. 데이터 수집하기

 _____

 _____

 _____



다. 현상 이해하기: 탐색적 데이터 분석

 _____

 _____

 _____

라. 현상 일반화하기: 통계적 추론

 _____

 _____

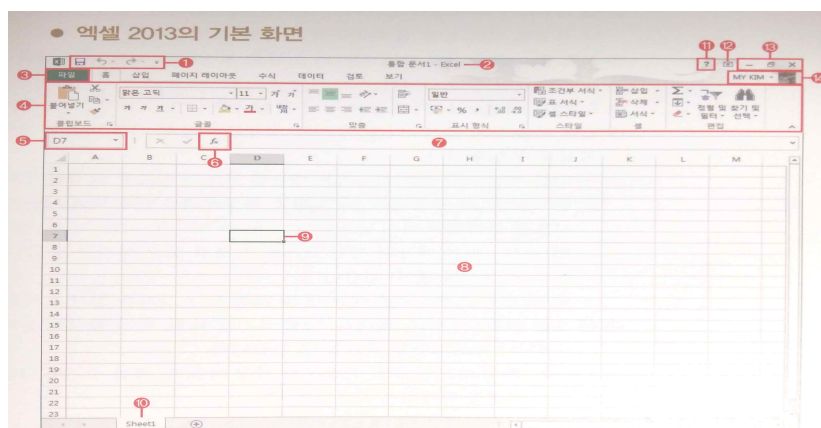
 _____

데이터 과학과 엑셀: 엑셀로 데이터 과학 맛보기-기본 기능 익히기

1. 엑셀 실행하기

시작-모든 프로그램-Microsoft Office-Microsoft Office 2013 클릭


2. 엑셀의 화면 구성



Excel 로 배우는 데이터 이터 과 학

수업일 : 이름 : 1

- ① 빠른 실행 도구 모음: 사용자가 자주 사용하는 도구를 등록하여 클릭으로 간단히 실행함.
- ② 제목 표시줄: 현재 작업 중인 엑셀 문서의 이름이 표시됨.
- ③ 파일 탭: 저장, 열기 등 문서 관리 및 인쇄, 계정 등을 위한 공간. 클릭하면 전체 창으로 열림.
- ④ 리본 메뉴: 작업에 필요한 명령 도구를 기능별로 묶음.
홈/삽입/페이지 레이아웃/수식/데이터/검토/보기 7가지 탭
- ⑤ 이름 상자: 워크시트에서 현재 셀 포인터가 위치한 곳의 셀 주소 또는 지정된 이름 표시
- ⑥ 함수 삽입: 클릭하면 함수 마법사 대화 상자가 나타나 사용할 함수를 선택할 수 있음.
- ⑦ 수식 입력줄: 현재 셀에 입력된 내용이 표시되며, 데이터 또는 수식을 직접 입력할 수 있음.
- ⑧ 워크시트: 데이터의 입력과 계산, 편집 등의 모든 문서 작업이 이루어지는 공간. 워크시트는 행과 열이 만나는 '셀(cell)'로 구성되어 있으며 화면의 크기를 확대 또는 축소할 수 있음.
- ⑨ 셀 포인터: 데이터가 입력되는 위치. 원하는 위치의 셀을 클릭하거나 화살표 키를 눌러 셀 포인터를 이동할 수 있음.
- ⑩ 시트탭: 문서 내의 모든 시트의 이름 표시. 시트 탭의 색과 시트 이름, 위치 등은 편집 가능함.
- ⑪ 도움말: 엑셀 2013의 도움말 창이 열림.
- ⑫ 리본 메뉴 표시 옵션: 리본 메뉴 자동 숨기기, 탭 표시, 탭 및 명령 표시의 메뉴를 이용해 리본 메뉴를 보이거나 숨기도록 설정할 수 있음.
- ⑬ 창 조절 단추: 최소화, 최대화, 이전 크기로 복원, 닫기 중에서 선택해 창의 크기를 조절하거나 프로그램을 종료할 수 있음.
- ⑭ 사용자: 현재 로그인 중인 사용자의 계정 이름 및 계정 사진이 나옴. 클릭해서 설정을 변경할 수 있음.

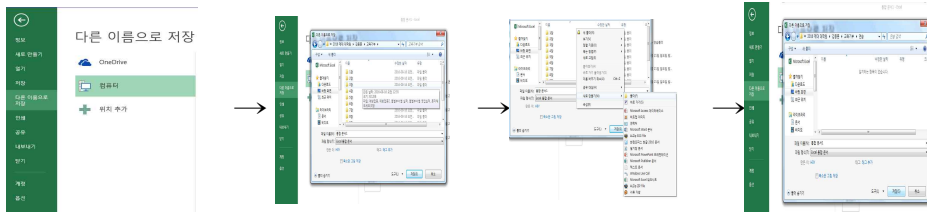


3. 문서 작성하고 저장하기

가. 문서 작성: 셀을 클릭해 셀 포인터가 위치하면 입력(또는 셀을 두 번 더블클릭)하고 'Enter' 키를 누르거나 다른 셀에 클릭

나. 저장하기

빠른 실행 도구 모음에서 '저장'버튼을 클릭하거나 '파일'탭의 '저장'을 클릭한 후



저장버튼
클릭-다른
이름으로 저장

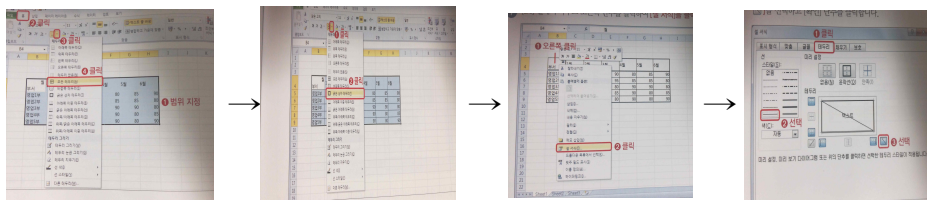
왼쪽 항목에
저장할 드라이브
선택-새 폴더
클릭

새 폴더가 생기면
폴더 이름을
입력하고 '여기'
클릭

파일 이름을
입력하고 '저장'
클릭-문서가
저장되어 제목
표시줄에 파일
이름이 나타남

4. 셀 서식 지정하기

가. 테두리 서식 지정하기



데이터
입력-홈-글꼴-테
두리-모든 테두리
클릭

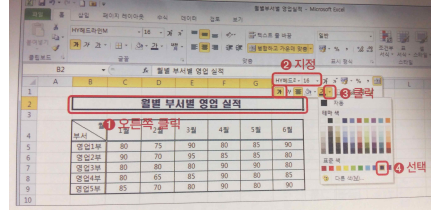
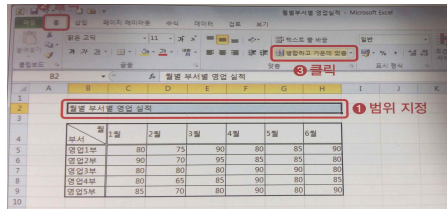
홈-글꼴-테두리-
굵은 상자 테두리
클릭

해당 셀에서
오른쪽 버튼-셀
서식 클릭

테두리-선
스타일은 실선,
테두리에서 대각선
선택-확인



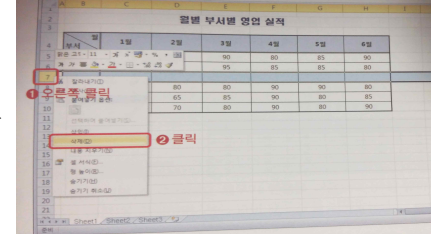
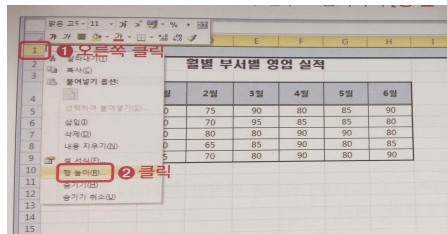
나. 글꼴과 맞춤 서식 지정하기



해당 영역 지정-홈 탭의 맞춤 그룹의 병합하고 가운데 맞춤 클릭

해당 영역 설정하고 오른쪽 버튼-글꼴

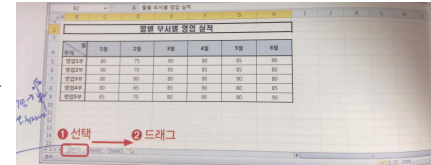
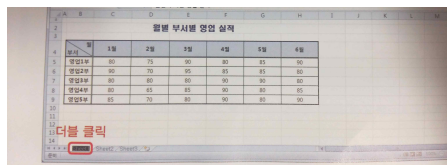
다. 행 높이와 열 너비 조절하고 행 삽입하기



관련 행 선택-오른쪽 버튼-높이 클릭-숫자 기록-확인(열 너비도 같은 방법)

관련 행 선택-오른쪽 버튼-삽입(또는 삭제)

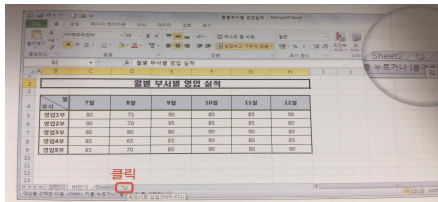
라. 워크시트 관리하기



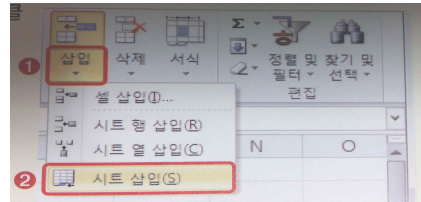
시트 이름 바꾸기: 아래 sheet1 더블클릭-글자 입력하고-Enter

시트 복사: 시트 1을 누른 상태로 뒤로 드래그





시트 삽입 1: 워크시트 삽입 단추 클릭



시트삽입 2: 홈-셀 그룹의 삽입-시트삽입

마. 표 서식 / 조건부 서식 적용하기

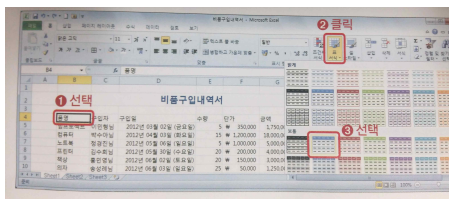
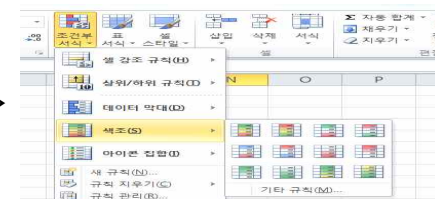
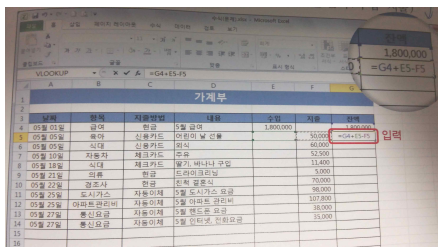


표 서식: 해당 영역 선택-홈-표 서식

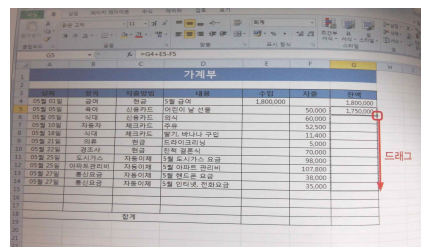


조건부 서식: 홈-조건부 서식-색조

5. 수식 입력하기



해당 셀에 수식 적기
(더하기+, 빼기-, 곱하기*, 나누기/)

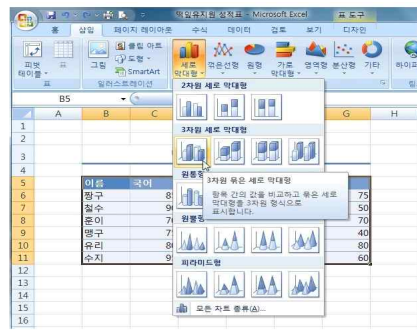


드래그 하면 서식이 복사됨



- ⑦ 축: 차트의 수평 방향(X축)과 수직 방향(Y축)에 각각의 항목이 나타내는 내용 또는 수치 표시
- ⑧ 축 제목: X축 또는 Y축에서 나타내는 내용을 제목으로 표시함.
- ⑨ 범례: 데이터의 계열이나 항목을 구분할 수 있도록 해당 색이나 무늬 등을 표시한 상자

가. 차트 삽입



영역 설정-삽입-차트 그룹

3차형 묶은 세로 막대형 그래프

나. 차트 레이아웃과 서식 지정하기

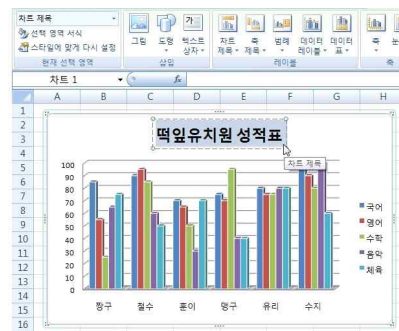


차트 도구-디자인-차트레이아웃

차트 도구-레이아웃-레이블(제목 입력)





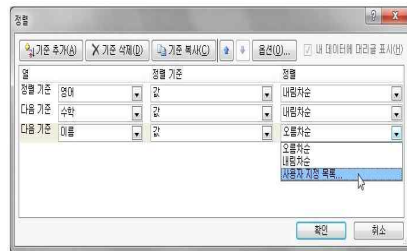
차트의 세로축 변경 (차트 오른 버튼-축 서식에서 주 단위 숫자 변경)
 차트 도구-레이아웃-레이블-데이터 레이블-표시 (전체 숫자 표시됨, 일부분만 할 경우는 해당 그래프만 Ctrl을 사용하여 선택)

7. 데이터베이스 관리하기

가. 데이터 정렬 이용하여 오름차순, 내림차순, 사용자 지정 목록 순 정렬하기

떡잎유치원 성적표							
이름	국어	영어	수학	음악	체육		
장구	85	55	25	65	75		
철수	90	95	85	60	50		
훈이	70	65	50	30	70		
맹구	75	70	95	40	40		
유리	80	75	75	80	80		
수지	95	90	80	95	60		

이름	국어	영어	수학	음악	체육
장구	85	55	25	65	75
철수	90	95	85	60	50
훈이	70	65	50	30	70
맹구	75	70	95	40	40
유리	80	75	75	80	80
수지	95	90	80	95	60

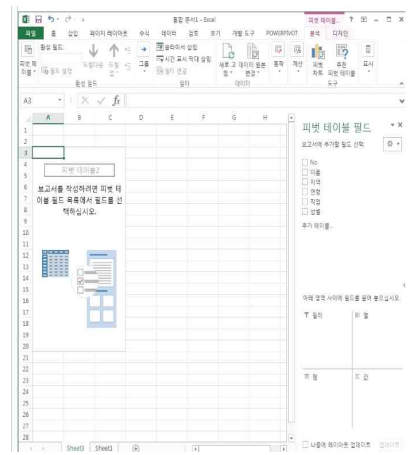
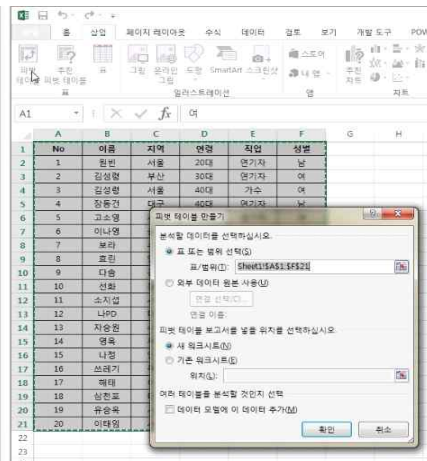


오름차순(1,2,3,4,5...)/내림차순(...5,4,3,2,1)
 영역 지정-정렬-사용자 지정 목록



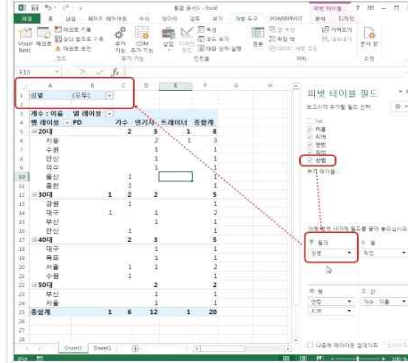
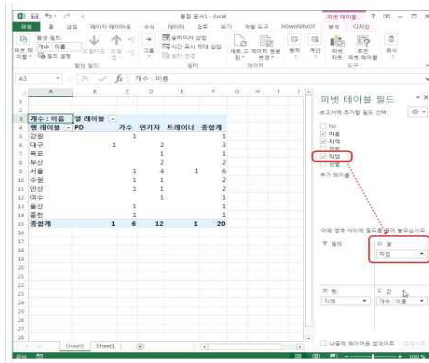
나. 피벗 테이블 이용하기

	A	B	C	D	E	F
1	No	이름	지역	연령	직업	성별
2	1	원빈	서울	20대	연기자	남
3	2	김성령	부산	30대	연기자	여
4	3	김성령	서울	40대	가수	여
5	4	장동건	대구	40대	연기자	남
6	5	고소영	서울	50대	연기자	여
7	6	이나영	울산	20대	가수	여
8	7	보라	수원	40대	가수	여
9	8	효린	안산	30대	가수	여
10	9	다솜	강원	30대	가수	여
11	10	선화	홍천	20대	가수	여
12	11	소지섭	서울	40대	연기자	남
13	12	나PD	대구	30대	PD	남
14	13	자승원	수원	20대	연기자	남
15	14	영욱	부산	50대	연기자	여
16	15	나정	안산	20대	연기자	여
17	16	쓰레기	특포	40대	연기자	남
18	17	해태	여수	20대	연기자	남
19	18	삼천포	대구	30대	연기자	남
20	19	유승옥	서울	20대	트레이너	여
21	20	이태임	서울	20대	연기자	여
22						



삽입-왼쪽에 있는 피벗 테이블 아이콘 클릭 피벗 테이블 필드가 보임



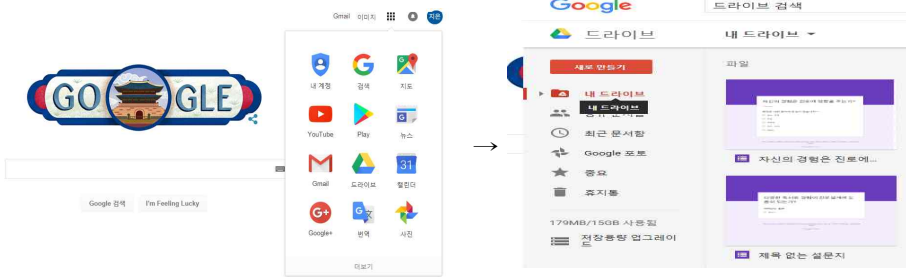


필요한 항목을 가져옴.
(합계: 이름, 행: 지역, 열: 직업)

행에 연령 추가
남녀 성비로 나눠서 보고 싶으면
필터에 성별 선택한 후 원하는 성을
선택함.

☑ 데이터 과학과 설문: 구글을 이용하여 설문지 만들기

- 네이버에 '구글'을 치거나 주소창에 <https://www.google.co.kr/> 를 친다.
- 아이디와 비번을 넣어 구글 계정을 만든다. 아이디는 기존에 사용하는 아이디를 넣는다.
(내 아이디: _____ , 내 비번: _____)
- 로그인한다.



- 드라이브를 찾아 클릭한다.
- 왼쪽 '새로 만들기' 를 클릭한다.



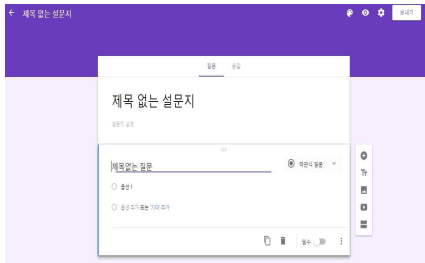
Excel 로 배우는 데 이 터 과 학

수업일 :

이름 :


1






6. 구글 설문지를 클릭한다.


7. 설문지 제목을 입력한다.






8. 설문지 응답 내용을 입력한다.


9. 설문지를 보낸다. (링크시킨다.)





10. 응답자는 주소를 클릭하면 설문에 응답할 수 있다.

11. 응답자가 전부 응답하면 설문 결과를 볼 수 있다.





로 배우는



수업일 :

이름 :

1



★ 과제는 다음과 같이 올립니다.

네이버-창의컴퓨터 교실 가입-왼쪽 탭 10기-데이터 과학 '공공' -글쓰기-제목에 8월 16일 과제 (~초 000), 아이디는 실명(원래 이름)-내용에는 구글 주소를 붙임

★ 과제 1

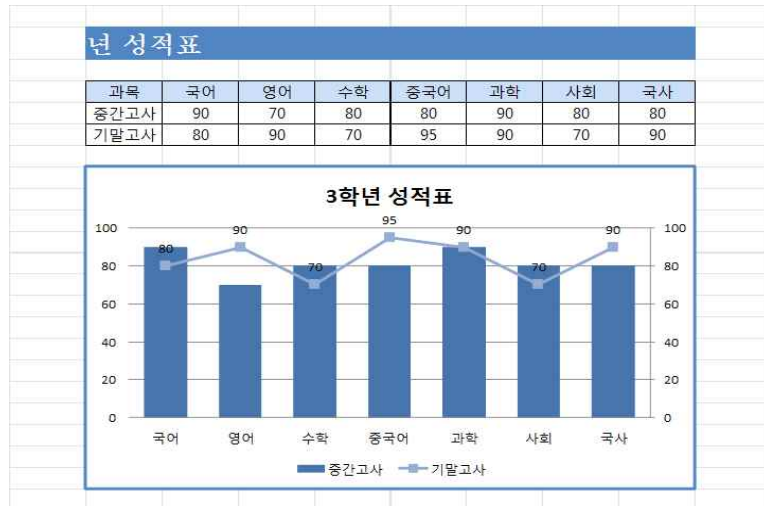
데이터 과학에 대한 내용을 주제로 정하여, 구글 설문지로 객관식 한 문항 (오지선다)을 만들고 완성한 설문지의 주소를 카페에 올리시오.

예시) 1. 데이터 과학의 의미에 대해 어느 정도 알고 있습니까?

- ① 매우 잘 알고 있다. ② 잘 알고 있다. ③ 보통이다. ④ 잘 모른다.
- ⑤ 전혀 모른다.

★ 과제 2

엑셀 프로그램을 사용하여 아래와 같은 표와 차트를 완성하고 카페에 올리시오.



학습목표

데이터를 수집, 가공하여 알뜰한 여행코스를 계획할 수 있다.

생각해 보기

푸른초등학교 6학년 학생은 여름방학을 맞이하여 가족여행을 떠나고 싶다. 토털검색과 블로그 등 인터넷을 이용해 교통, 숙박, 여행지에 대한 정보를 찾아보고 있다. 하지만 경비를 좀 더 절약하여 알뜰한 여행을 떠나고 싶다. 이 많은 정보를 어떻게 정리하여 최적의 여행코스를 계획할 수 있을까?

☑ 데이터 문제 정의하기

여행에 필요한 경비를 절약하는 다양한 방법을 이해하여 합리적으로 소비하는 자세를 가질 수 있는가?

1. 문제 정의하기

- 목표: 여행에 필요한 경비를 절약하는 다양한 방법을 이해하여 합리적으로 소비하는 자세를 가질 수 있다.
- 범위: 초등학교 4-6학년 학생
- 문제 해결의 제약 조건(시간과 비용): 즉시 또는 전날, 비용은 무료.



2. 데이터 정의하기

- 데이터의 제한: 여행은 제주도내, 4인 가족 기준으로 제한한다.
- 데이터를 어떻게 수집할 것인가: 각 체험장소별 홈페이지, 블로그 등 활용 (참가 학생 모두 스마트폰이 있을 때는 SNS를 활용한 방법도 생각해 볼 수 있다.)
- 데이터 처리 및 분석을 위한 최적의 방법과 도구: 엑셀 프로그램을 활용한 분석 및 시각화
- 최종 결과물은 어떤 형태로 누구에게 전달되어야 하는가: 결과보고서, 보호자와 학생

3. 연구 설문 및 가설 정의하기


- 주어진 많은 정보를 어떻게 정리하면 최적의 여행코스를 계획할 수 있을까?


☑ 데이터 수집하기: 필요한 데이터 만들기-자동수집


1. 데이터 수집의 기본

- 관찰결과를 개별 항목으로 하는 테이블로 만들기
- 각 열은 개별 항목이 갖는 속성, 각 행은 수집단위가 되는 개별 항목

○ 관련된 항목 생각해 보기 ○

 _____

 _____

 _____



2. 수집방법 결정하기(자동수집)

step 1) 인터넷에서 정보를 찾아 셀 서식의 맞춤과 글꼴을 이용하여 정리하기
 제주 주요 박물관 홈페이지를 검색하여 정보를 수집한다.

step 2) 응용 및 개별 과제: 통계청에서 건강과 관련된 통계 자료를 찾아 조건부서식을 이용한 문서꾸미기

3. 데이터 품질 점검하기

- 완전성(문제 해결에 필요한 내용이 들어 있느냐?)
- 정확성(문제가 되는 사항이 들어 있느냐?)
- 일관성(데이터 항목들이 연관이 있고 일치하느냐?)

4. 데이터 준비하기

- 테이블 형태로 변환하기

입장료와 할인표를 엑셀의 셀 서식 중 표시형식과 테두리, 무늬를 이용하여 정리한다.

제주 주요 박물관 입장료								
1								
2								
3	박물관 이름	관람 시간	입장료		청소년		휴관 요일	
4			일반		개인	단체		
5		최대 시간	변동	개인	단체	개인	단체	
6	국립제주박물관	09:00~19:00	요일	2,000	1,500	1,000	500	월요일
7	제주민속자연사박물관	09:00~18:00	계절, 휴일	1,000	700	500	300	월요일
8	제주민속촌박물관	09:00~19:00	계절, 휴일	3,000	2,400	1,500	1,200	화요일
9	제주현대미술관	09:00~19:00	요일	1,000	700	500	300	월요일
10	제주교육박물관	09:00~19:00	요일	1,000	700	500	300	월요일
11	항공우주박물관	09:00~19:00	요일	1,000	700	500	300	월요일
12	세계자동차 박물관	09:00~19:00	요일	1,000	700	500	300	월요일



☑ 현상 이해하기: 탐색적 데이터 분석

1. 원본 데이터의 시각화(별 속성값 관찰, 속성간의 관계와 패턴 찾기)

- 표 선택 -> 마우스 오른쪽 클릭 -> 빠른 분석 -> 묶은 세로 막대형 선택

○ 알 수 있는 점 ○









수업일 :

이름 :

☑ 데이터 과학 보고서: 오늘 실습한 내용을 바탕으로 보고서 써보기

1. 주제 선택 및 탐구 동기

2. 탐구과정

가. 문제 정의

- 문제:

- 목표:

- 가설:

나. 데이터 수집

- 방법:

- 항목(내용):



다. 탐색적 데이터 분석 방법(실제 완성 때는 차트 등도 삽입해야 함. 이번은 연습)

- 원본 데이터 살펴보고 시각화:
- 개별 속성 분석:
- 속성간의 관계:

라. 통계적 추론(결론 및 예측, 짧게 간추려서)

- 결론:
- 예측:

3. 느낀 점(소감)





★ 과제는 다음과 같이 올립니다.

네이버-창의컴퓨터 교실 가입-왼쪽 탭 10기-데이터 과학 '공공' -글쓰기

★ 과제

한국고속철도 홈페이지를 방문하여 KTX 열차시간 및 운임표를 보면서 자료를 정리, 종합하여 시간과 비용을 절약할 수 있는 여행계획을 세워보자.





우리나라의 인구 변화 데이터를 보고 대응하는 방안을 마련해 보자

생각해 보기

우리나라의 인구 정책 포스터는 다음과 같이 바뀌어 왔습니다.



1970년대



1980년대



2000년대



최근

왜 이러한 변화가 발생했을까요? 그리고, 어떤 문제점들이 있는지 알아보고, 대응 방안을 마련해 봅시다.

☑ 데이터 문제 정의하기: 인구 변화가 우리 사회에 미치는 영향

1. 문제 정의하기

- 목표: 우리나라의 인구 변화가 우리 사회의 미치는 영향을 알아볼 수 있다.
- 범위: 초등학교 4~6학년 학생



2. 데이터 정의하기

- 데이터를 어떻게 수집할 것인가: 공공 데이터 활용(국가통계포털 활용)
- 데이터 처리 및 분석을 위한 최적의 방법과 도구: 엑셀 프로그램을 활용한 분석 및 시각화
- 최종 결과물은 어떤 형태로 누구에게 전달되어야 하는가: 결과보고서, 보호자와 학생

3. 연구 설문 및 가설 정의하기

- 인구 변화에 따른 문제점을 알아보고, 다양한 해결 방안을 탐색한다.

☑ 데이터 수집하기: 필요한 데이터 수정하기

1. 데이터 수집의 기본

- 인구 변화를 탐색할 수 있는 자료를 구분하기
- 필요 없는 데이터 제거하기

2. 데이터 품질 점검하기

- 완전성, 정확성, 일관성

3. 데이터 준비하기

- 테이블 형태로 변환하기

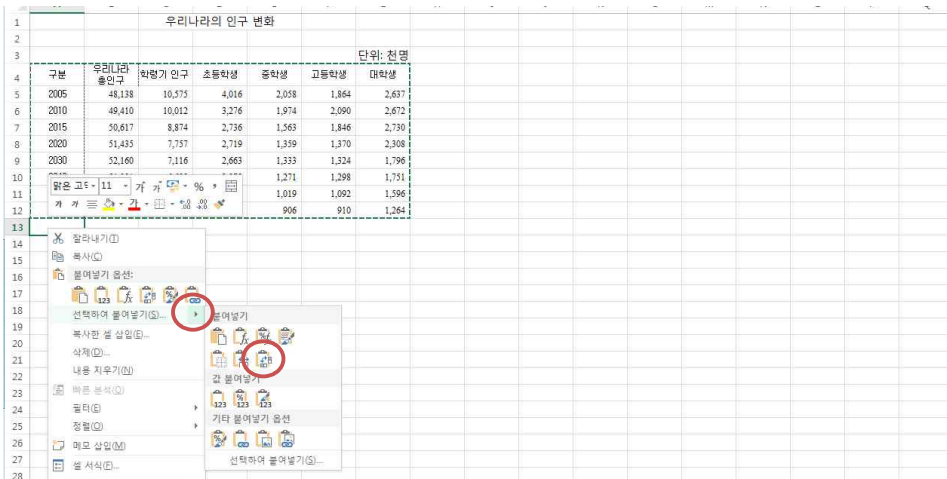
구분	2005	2010	2015	2020	2030	2040	2050	2060
우리나라 총인구	48,138	49,410	50,617	51,435	52,160	51,091	48,121	43,959
학령기 인구	10,575	10,012	8,874	7,757	7,116	6,698	5,618	4,884
초등학생	4,016	3,276	2,736	2,719	2,663	2,378	1,912	1,805
중학생	2,058	1,974	1,563	1,359	1,333	1,271	1,019	906
고등학생	1,864	2,090	1,846	1,370	1,324	1,298	1,092	910
대학생	2,637	2,672	2,730	2,308	1,796	1,751	1,596	1,264



- 1) 제목 글씨체와 글씨 크기 바꾸기
- 2) 표 가로 축과 세로 축 바꾸기
 - 표 전체 영역을 드래그하여 설정하고, 오른쪽 마우스를 클릭하여 복사한다.

	A	B	C	D	E	F	G
1	우리나라의 인구 변화						
2							
3							단위: 천명
4	구분	우리나라 총인구	학령기 인구	초등학생	중학생	고등학생	대학생
5	2005	48,138	10,575	4,016	2,058	1,864	2,637
6	2010	49,410	10,012	3,276	1,974	2,090	2,672
7	2015	50,617	8,874	2,736	1,563	1,846	2,730
8	2020	51,435	7,757	2,719	1,359	1,370	2,308
9	2030	52,160	7,116	2,663	1,333	1,324	1,796
10	2040	51,091	6,698	2,378	1,271	1,298	1,751
11	2050	48,121	5,618	1,912	1,019	1,092	1,596
12	2060	43,959	4,884	1,805	906	910	1,264

- A13 셀에 가서 오른쪽 마우스 클릭 -> 선택하여 붙여넣기



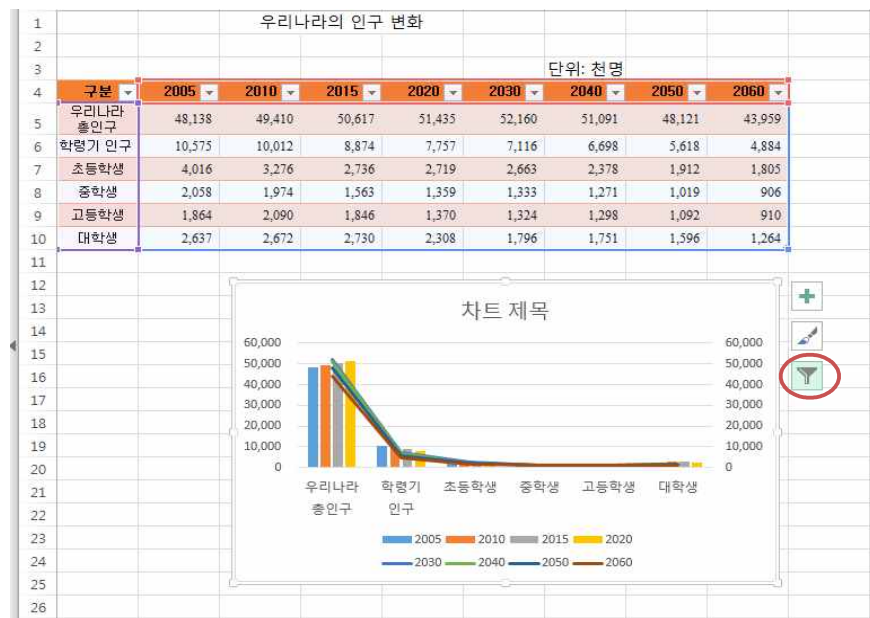
- 처음 표 영역 삭제하기
- 표 영역 안 글자체 및 글자 크기 바꾸기
- 표 영역 오른쪽 마우스 클릭-셀 서식-테두리-테두리 선 넣기
- 홈 메뉴-스타일-표 서식에서 표 서식을 선택



☑ 현상 이해하기: 탐색적 데이터 분석

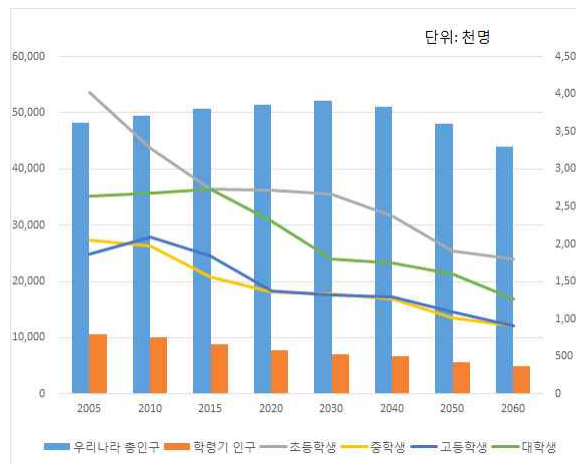
1. 원본 데이터의 시각화(개별 속성값 관찰, 속성간의 관계와 패턴 찾기)

- 표 영역 선택 -> 삽입 메뉴 -> 추천차트 -> 모든 차트 -> 콤보 -> 묶은 세로 막대형, 꺾은선 형, 보조 축 선택
- 만약, 계열이 잘못 선택 되었을 경우에는 차트 필터 선택 -> 데이터 선택 -> 행/열 전환 버튼 클릭



- 차트에서 마우스 오른쪽 클릭-> 차트 종류 변경 -> 초등학생, 중학생 차트 종류를 꺾은선 그래프로 바꾸고 보조 축 체크하기





2. 속성간의 관계 분석하기

- 1) 우리나라 총인구는 어떻게 변화하고 있나요?
- 2) 우리나라 학령기 인구는 어떻게 변화하고 있나요?
- 3) 초등학생, 중학생, 고등학생, 대학생 중 가장 변화가 심하게 나타나는 것은 어느 것인가요? 그리고 그 이유는 무엇이라고 생각하나요?
- 4) 우리나라 인구로 인한 어떤 문제가 발생할지 적어보세요.
- 5) 위 문제의 원인은 무엇이라고 생각하나요?
- 6) 문제의 원인을 알기 위해서는 어떤 자료가 더 필요할까요?

3. 더 나아가기

- 1) 국가별 출산율 변화



국가별	1995	2000	2005	2010	2015	2020
세계	3.04	2.74	2.62	2.56	2.51	2.47
한국	1.70	1.51	1.22	1.23	1.26	1.33
중국	2.00	1.48	1.50	1.53	1.55	1.59
일본	1.48	1.37	1.30	1.34	1.40	1.46
터키	2.87	2.62	2.35	2.18	2.10	2.01
캐나다	1.69	1.56	1.52	1.64	1.61	1.56
미국	2.03	2.00	2.04	2.06	1.89	1.90
덴마크	1.75	1.76	1.76	1.85	1.73	1.76
핀란드	1.82	1.74	1.75	1.84	1.75	1.77
프랑스	1.71	1.76	1.88	1.97	2.00	1.99
독일	1.30	1.35	1.35	1.36	1.39	1.44
이탈리아	1.27	1.22	1.30	1.42	1.43	1.49
노르웨이	1.89	1.86	1.81	1.92	1.80	1.81
폴란드	1.95	1.51	1.26	1.37	1.37	1.33
스페인	1.28	1.19	1.29	1.39	1.32	1.38
영국	1.78	1.74	1.66	1.88	1.92	1.91

- 2) 위의 자료를 다양한 방법으로 시각화 해 봅시다. 그리고, 위에서 알아보았던 인구변화와 관련된 내용을 분석하여 봅시다.
- 3) 위의 변화에 대한 대응방안에 대하여 생각해 보고, 위의 자료들을 종합하여 발표하여 봅시다.





로 배우는



수업일 :

이름 :

3



★ 과제는 다음과 같이 올립니다.

네이버-창의컴퓨터 교실 가입-왼쪽 탭 10기-데이터 과학 '공공' -글쓰기

★ 과제

※ 다음은 다문화 가정 학생 수의 변화를 나타낸 것입니다.

(단위 : 명)

연도	2006년	2010년	2015년
전체 학생	7,776,211	7,236,248	6,097,297
초등학교 다문화 가정 학생 수	7,910	24,701	60,283
중학교 다문화 가정 학생 수	1,139	5,260	13,865
고등학교 다문화 가정 학생 수	340	1,827	8,388

- 표를 선택하고 마우스 오른쪽 클릭 빠른 분석을 통하여 다양한 방법으로 표현해 볼 수 있습니다.



<기본과제>

- 다양한 방법으로 시각화 해봅시다.
(표를 선택하고 마우스 오른쪽 클릭 빠른 분석을 통하여 다양한 방법으로 표현해 볼 수 있습니다.)



<심화과제>

- 시각화한 내용을 통해 알게된 점을 적어봅시다.
- 미래에는 어떤 상황이 발생하고 우리는 어떻게 대비하여야할지 생각해 봅시다.





청소년들의 안전 문제에 대한 데이터를 보고 대안을 마련해 보자

생각해 보기

청소년들 중에 안타깝게 소중한 생명을 잃는 친구들이 있습니다. 그런데 우리나라의 통계 중 매해 1,000명에 가까운 학생들이 ○○ 때문에 소중한 생명을 잃을 뿐만 아니라 최근 몇 년 동안 이 것 때문에 소중한 생명을 잃는 학생들이 가장 많다고 합니다. 이것이 무엇인지 여러분들은 알 수 있겠습니까?

여러분들은 OECD라고 들어보셨나요? OECD는 경제협력개발기구로 세계 여러 나라들 중 협력을 통해 경제와 사회발전을 공동으로 모색해 나가는 단체입니다. 우리나라는 OECD에 1996년에 가입하였으며, 현재 회원국수는 34개국입니다. 대표적인 가입국으로 미국, 일본, 영국, 프랑스, 독일, 이탈리아, 스페인, 네덜란드, 스웨덴, 노르웨이, 덴마크 등이 있습니다. OECD에서는 다양한 통계를 제공하고 있습니다. 그런데 10년째 우리나라는 ○○률이 가장 높은 나라로 나오고 있습니다. 이 ○○은 과연 무엇일까요?

이 ○○을 줄일 수 있는 청소년들의 안전 문제에 대하여 대안을 마련해 봅시다.



☑ 데이터 문제 정의하기

청소년들의 안전을 위한 사망 원인을 분석하고 대안 마련하기

1. 문제 정의하기

- 목표: 우리나라 청소년들의 사망 원인을 알아보고, 그 대안을 마련할 수 있다.
- 범위: 초등학교 4~6학년 학생

2. 데이터 정의하기

- 데이터를 어떻게 수집할 것인가: 공공 데이터 활용(국가통계포털 활용)
- 데이터 처리 및 분석을 위한 최적의 방법과 도구: 엑셀 프로그램을 활용한 분석 및 시각화
- 최종 결과물은 어떤 형태로 누구에게 전달되어야 하는가: 결과보고서, 보호자와 학생

3. 연구 설문 및 가설 정의하기

- 우리나라 청소년들의 사망 원인을 알아보고, 다양한 해결 방안을 탐색한다.

☑ 데이터 수집하기: 필요한 데이터 스스로 만들기

1. 데이터 수집의 기본

- 사망 원인을 탐색할 수 있는 자료를 구분하기
- 필요 없는 데이터 제거하기

2. 데이터 품질 점검하기

- 완전성, 정확성, 일관성

3. 데이터 준비하기

- 원본 파일



(단위 : 명, 인구10만 명당)

9-24세	1위	2위	3위	4위	5위
	사망자(사망률)	사망자(사망률)	사망자(사망률)	사망자(사망률)	사망자(사망률)
2000	운수 사고	고의적 자해(자살)	악성신생물(암)	불의의 익사 및 익수	심장 질환
	1,655 (14.3)	697 (6.0)	523 (4.5)	394 (3.4)	164 (1.4)
2005	고의적 자해(자살)	운수 사고	악성신생물(암)	불의의 익사 및 익수	심장 질환
	888 (7.9)	693 (6.2)	427 (3.8)	164 (1.5)	115 (1.0)
2010	고의적 자해(자살)	운수 사고	악성신생물(암)	심장 질환	불의의 익사 및 익수
	926 (8.8)	634 (6.0)	344 (3.3)	101 (1.0)	93 (0.9)
2014	고의적 자해(자살)	운수 사고	악성신생물(암)	심장 질환	불의의 익사 및 익수
	742 (7.4)	485 (4.9)	288 (2.9)	80 (0.8)	60 (0.6)

- 사망원인별 / 연도별로 변환하기

(단위 : 명, 인구10만 명당)

사망원인	고의적 자해(자살)	운수 사고	악성신생물(암)	심장 질환	불의의 익사 및 익수
2000	6.0	14.3	4.5	1.4	3.4
2005	7.9	6.2	3.8	1.0	1.5
2010	8.8	6.0	3.3	1.0	0.9
2014	7.4	4.9	2.9	0.8	0.6

☑ 현상 이해하기: 탐색적 데이터 분석

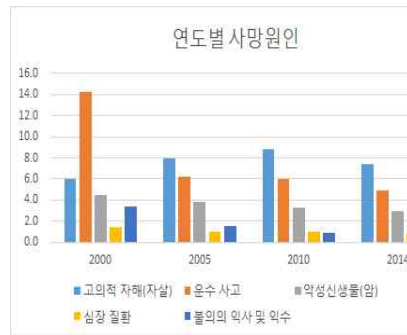
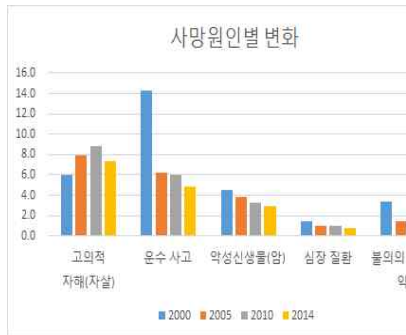
1. 원본 데이터의 시각화 (개별 속성값 관찰, 속성간의 관계와 패턴 찾기)

- 표 선택 -> 마우스 오른쪽 클릭 -> 빠른 분석 -> 묶은 세로 막대형 선택(2가지 종류)



2. 속성간의 관계 분석하기

- 1) ○○은 무엇인가요?
- 2) 다음 항목을 분류하여 봅시다. 사망원인은 크게 질병과 사고로 나눌 수 있습니다. 위의 다섯 개의 항목을 사고와 질병으로 나누어 봅시다. 자살, 운수 사고, 악성신생물(암), 심장 질환, 불의의 익사 및 익수



사고	질병

2. 속성간의 관계 분석하기

- 1) ○○은 무엇인가요?
- 2) 다음 항목을 분류하여 봅시다. 사망원인은 크게 질병과 사고로 나눌 수 있습니다. 위의 다섯 개의 항목을 사고와 질병으로 나누어 봅시다. 자살, 운수 사고, 악성신생물(암), 심장 질환, 불의의 익사 및 익수
- 3) 고의적 자해(자살)은 어떻게 변화하고 있나요?
- 4) 연도별로 가장 많이 줄어든 사망 원인은 무엇인가요?
- 5) 왜 그것이 가장 많이 줄어든 것 같나요?
- 6) 가장 적게 줄어든 사망 원인은 무엇인가요?
- 7) 왜 그것이 가장 많이 줄어든 것 같나요?
- 8) 위 문제의 원인은 무엇이라고 생각하나요?
- 9) 문제를 해결하기 위해서는 어떤 자료가 더 필요할까요?





Excel

로 배우는



수업일 :

이름 :

4

3. 더 나아가기

1) 자살에 대한 충동 여부 및 이유

(단위 : %)

연도	있다	자살하고 싶었던 주된 이유							
		계	경제적 어려움	이성 문제	직장 문제	외로움 고독	가정 불화	성적, 진학문제	기타
2008	8.9	100.0	12.3	4.4	9.4	16.6	10.3	35.1	12.0
2010	8.8	100.0	17.0	3.9	6.8	12.7	11.8	37.8	9.9
2012	11.2	100.0	20.5	2.8	6.7	14.1	13.6	28.0	14.3
2014	7.9	100.0	19.7	6.4	8.5	12.4	10.2	27.3	15.5

2) 위의 자료를 다양한 방법으로 시각화 해 봅시다. 그리고, 위에서 알아 보았던 자살과 관련된 내용을 분석하여 봅시다.

3) 위의 원인을 분석하여 대응방안에 대하여 생각해 보고, 위의 자료들을 종합하여 발표하여 봅시다.





로 배우는



수업일 :

이름 :

4



※ 홈페이지 '데이터 과학 보고서' 양식을 참고하여 직접 만든 차트와 표, 그림 등을 붙여서 정리한 후 창의컴퓨터 교실 카페에 제목 쓰고 첨부할 것.

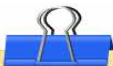
★ 과제

※ 다음은 OECD 국가의 자살률 변화를 나타낸 것입니다.

(단위 : 명, 10만 명당)

연도	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
멕시코	4.6	4.5	4.6	4.5	4.5	4.7	5	4.8	5.2	5
영국	6.6	6.9	6.7	6.7	6.3	6.9	6.8	6.7	6.9	7
덴마크	11.5	12	11.3	11.6	10.3	10.7	10.8	9.8	10.2	11.3
독일	12.6	12	11.4	10.7	10.2	10.3	10.3	10.8	10.8	10.5
호주	10.9	10.6	-	10.5	10.6	10.9	10.7	10.9	10.1	10.1
노르웨이	11.1	11.8	11.6	11.5	10.5	10.6	11.9	11.2	12.1	10.2
미국	11.1	11.3	11.2	11.3	11.7	12	12.2	12.5	12.5	12.5
체코	16.7	15	14.8	13.1	12.7	12.6	13.2	13.5	14.3	14.7
핀란드	20.1	20	18.3	19.6	18.2	19	18.9	17.3	16.4	15.6
일본	23.3	21.9	22.1	21.6	22.1	21.8	22.2	21.2	20.9	19.1
헝가리	27.1	26.5	25.2	23.4	23.3	23.4	23.3	23.4	22.8	22
한국	28.1	29.5	29.9	26.2	28.7	29	33.8	33.5	33.3	29.1

- 표를 선택하고 마우스 오른쪽 클릭 빠른 분석을 통하여 다양한 방법으로 표현해 볼 수 있습니다.



<기본과제>

- 다양한 방법으로 시각화 해 봅시다. (표를 선택하고 마우스 오른쪽 클릭 빠른 분석을 통하여 다양한 방법으로 표현해 볼 수 있습니다.)



<심화과제>

- 시각화한 내용을 통해 알게 된 점을 적어봅시다.
- 미래에는 어떤 상황이 발생하고 우리는 어떻게 대비하여야할지 생각해 봅시다.





미래의 진로에 관한 다양한 주제를 접해 보고, 나의 진로를 탐색해 보자

생각해 보기

여러분들은 미래의 진로에 대하여 생각해 본적이 있습니까? 중, 고등학생들을 대상으로 선호 직업 순위를 조사한 결과는 다음 표와 같습니다.

순위	전체	남자	여자
1	초등학교 교사	공무원	초등학교 교사
2	의사	운동선수	간호사
3	공무원	의사	중고등학교 교사
4	중고등학교 교사	경찰	의사
5	요리사	회사원	요리사
6	간호사	요리사	연예인
7	경찰	초등학교교사	공무원
8	운동선수	중고등학교교사	스튜어디스
9	연예인	직업군인	디자이너
10	회사원	건축가	연기자

여러분들도 위의 자료에 선호하는 직업이 있습니까? 그럼 어떤 기준으로 그것을 선택했습니까? 어떤 기준으로 위의 직업을 선택했는지 한번 알아보시다.

☐ 문제 정의

직업 선택을 위한 요인을 파악하기

1. 문제 정의하기

- 목표: 우리나라 청소년들의 직업 선택 요인을 알아보고, 그에 따라 나의 진로를 선택할 수 있다.
- 범위: 초등학교 4~6학년 학생



2. 데이터 정의하기

- 데이터를 어떻게 수집할 것인가: 공공 데이터 활용(국가통계포털 활용)
- 데이터 처리 및 분석을 위한 최적의 방법과 도구: 엑셀 프로그램을 활용한 분석 및 시각화
- 최종 결과물은 어떤 형태로 누구에게 전달되어야 하는가: 결과보고서, 보호자와 학생

3. 연구 설문 및 가설 정의하기

- 우리나라 청소년들의 직업 선택 원인을 알아보고, 그에 따라 나의 진로를 선택해 봅시다.

☑ 데이터 수집: 필요한 데이터 수정하기

1. 데이터 수집의 기본

- 진로 선택 요인 탐색할 수 있는 자료를 구분하기
- 필요 없는 데이터 제거하기

2. 데이터 품질 점검하기

- 완전성, 정확성, 일관성

3. 데이터 준비하기

- 원본 파일

(단위 : %)

연도	성별	계	명명 예성	안정성	수 입	적 흥 성 미	보람 자아성취	발전성 장래성	기 타
2011	남자	100.0	5.0	21.5	28.9	29.1	6.1	7.4	2.1
2011	여자	100.0	2.6	19.1	25.9	38.8	6.2	6.0	1.3
2013	남자	100.0	4.9	22.2	28.1	31.6	5.4	6.4	1.4
2013	여자	100.0	2.3	20.3	25.9	36.9	7.4	5.7	1.4
2015	남자	100.0	5.0	22.4	29.2	31.0	6.4	5.5	0.7
2015	여자	100.0	3.3	23.2	24.7	35.5	7.1	5.2	1.0



- 남자, 여자 평균값 만들기(표 서식 넣기)

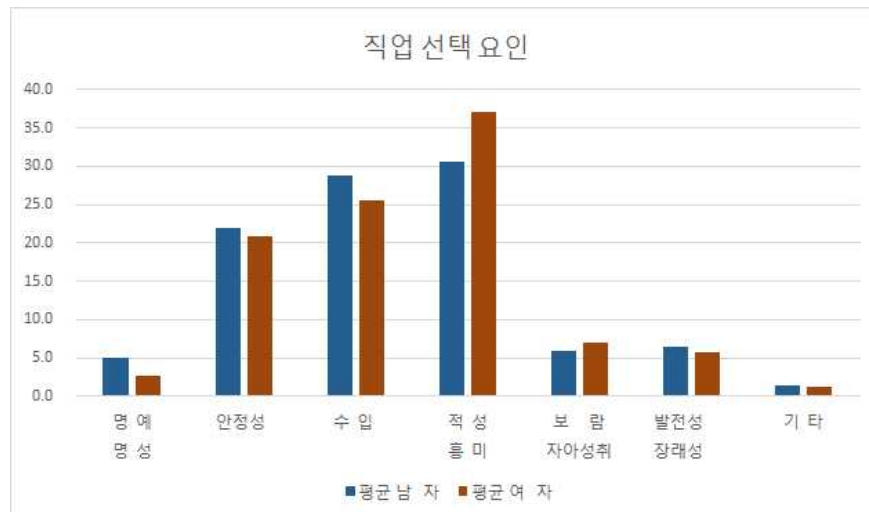
(단위 : %)

연도	성별	계	명 명	예 성	안정성	수 입	적 성 흥 미	보 람 자아성취	발전성 장래성	기 타
2011	남 자	100.0	5.0	21.5	28.9	29.1	6.1	7.4	2.1	
2011	여 자	100.0	2.6	19.1	25.9	38.8	6.2	6.0	1.3	
2013	남 자	100.0	4.9	22.2	28.1	31.6	5.4	6.4	1.4	
2013	여 자	100.0	2.3	20.3	25.9	36.9	7.4	5.7	1.4	
2015	남 자	100.0	5.0	22.4	29.2	31.0	6.4	5.5	0.7	
2015	여 자	100.0	3.3	23.2	24.7	35.5	7.1	5.2	1.0	
평균	남 자	100.0	5.0	22.0	28.7	30.6	6.0	6.4	1.4	
평균	여 자	100.0	2.7	20.9	25.5	37.1	6.9	5.6	1.2	

☑ 현상 이해하기: 탐색적 데이터 분석

1. 원본 데이터의 시각화 (개별 속성값 관찰, 속성간의 관계와 패턴 찾기)

- 표 선택 -> 마우스 오른쪽 클릭 -> 빠른 분석 -> 묶은 세로 막대형 선택



- 범주에서 계 체크 해제

2. 속성간의 관계 분석하기

- 1) 남자들이 직업 선택을 위한 요인을 순서대로 적어봅시다.
- 2) 여자들이 직업 선택을 위한 요인을 순서대로 적어봅시다.
- 3) 선호 직업에 대한 선택 요인을 추측하여 봅시다.(명예, 안정성, 수입, 적성, 보람, 장래성)
 - 중복하여 적을 수 있음

남자 선호 직업	직업 선택 요인 추측	여자 선호 직업	직업 선택 요인 추측
공무원		초등학교 교사	
운동선수		간호사	
의사		중고등학교 교사	
경찰		의사	
회사원		요리사	
요리사		연예인	
초등학교교사		공무원	
중고등학교교사		스튜어디스	
직업군인		디자이너	
건축가		연기자	

- 4) 내가 중요하게 생각하는 선택 요인을 차례대로 적어 봅시다.
- 5) 위의 선택 요인에 따라 선호 직업을 적어 봅시다.(인터넷을 활용하여 봅시다. 커리어넷 등)

3. 더 나아가기

- 1) 선호하는 직장 유형

(단위 : %)

연도	성별	계	국가 기관	공기업 (공사)	대기업	벤처 기업	외국계 기업	전문직 기업	중소 기업	해외 취업	자영업	기타
2011	남자	100.0	28.1	13.2	25.5	3.8	4.0	8.3	3.0	2.1	9.8	2.3
2011	여자	100.0	28.5	13.0	20.2	2.5	6.9	12.1	2.2	3.9	8.1	2.6
2013	남자	100.0	25.9	16.0	25.9	3.8	4.6	7.3	3.6	2.4	7.9	2.7
2013	여자	100.0	31.3	14.8	18.3	1.9	7.4	9.7	2.5	3.6	7.2	3.1
2015	남자	100.0	22.3	18.7	22.4	5.0	5.0	7.2	3.4	3.3	11.2	1.6
2015	여자	100.0	25.1	17.6	17.6	2.4	8.3	11.6	3.0	4.1	8.3	2.1





로 배우는



수업일 :

이름 :

5

- 2) 위의 자료를 다양한 방법으로 시각화 해 봅시다. 그리고, 위에서 알아 보았던 직업 선택 요인과 관련지어 분석하여 봅시다.
- 3) 위의 자료들을 종합하여 발표하여 봅시다.



★ 과제는 다음과 같이 올립니다.

네이버-창의컴퓨터 교실 가입-왼쪽 탭 10기-데이터 과학 '공공' -글쓰기

★ 과제

※ 다음은 청소년의 여가활용에 관한 자료입니다.

주말이나 휴일의 여가활동(복수응답)

(단위 : %)

성별	TV및 DVD 시청	문화 예술 관람	문화 예술 참여	스포츠 관람	스포츠 활동	관광 활동	컴퓨터 게임 인터넷 검색 등	취미·자기 개발 활동	휴식 활동	사회 및 기타 활동	기타
남자	52.7	20.9	4.7	7.7	24.8	3.7	58.7	21.4	35.4	6.8	1.4
여자	66.4	31.5	5.3	3.3	4.5	6.8	38.2	22.8	49.7	11.8	1.4

앞으로 하고 싶은 여가활동(복수응답)

(단위 : %)

성별	TV및 DVD 시청	문화 예술 관람	문화 예술 참여	스포츠 관람	스포츠 활동	관광 활동	컴퓨터 게임 인터넷 검색 등	취미·자기 개발 활동	휴식 활동	사회 및 기타 활동	기타
남자	14.6	35.5	11.5	18.2	38.4	42.2	25.6	34.6	17.0	7.0	0.2
여자	13.5	58.6	15.2	7.5	19.9	57.8	8.2	45.0	19.0	10.0	0.2

- 2개의 자료를 보고 다양한 분석 방법을 활용하여 분석하고 알게 된 점을 표현해 봅시다.





<기본과제>

- 다양한 방법으로 시각화 해봅시다.
(표를 선택하고 마우스 오른쪽 클릭 빠른 분석을 통하여 다양한 방법으로 표현해 볼 수 있습니다.)



<심화과제>

- 시각화한 내용을 통해 알게 된 점을 적어봅시다.
- 여가활용에 대한 현재와 하고 싶은 점의 차이를 알아봅시다.



<부록 3> 앱인벤터를 활용한 데이터 과학 교육 프로그램

MIT App Inventor 로 배우는 데이터 과학 학

수업일 : 이름 : 1

학습목표


데이터 과학의 의미를 이해하고 이를 위한 앱인벤터의 기본 기능을 익힐 수 있다.

☑ 데이터 과학이란? 데이터 과학의 의미 알기

1. 데이터 과학의 의미

2. 데이터 과학의 예

3. 데이터 과학을 시작하기 위해 필요한 것





로 배우는 **데** **이** **터** **과** **학**


수업일 : 이름 : 1

4. 데이터 과학의 단계


가. 데이터 문제 정의하기


 _____


 _____

 _____


나. 데이터 수집하기


 _____


 _____

 _____


다. 현상 이해하기: 탐색적 데이터 분석


 _____


 _____


 _____

라. 현상 일반화하기: 통계적 추론

 _____

 _____

 _____



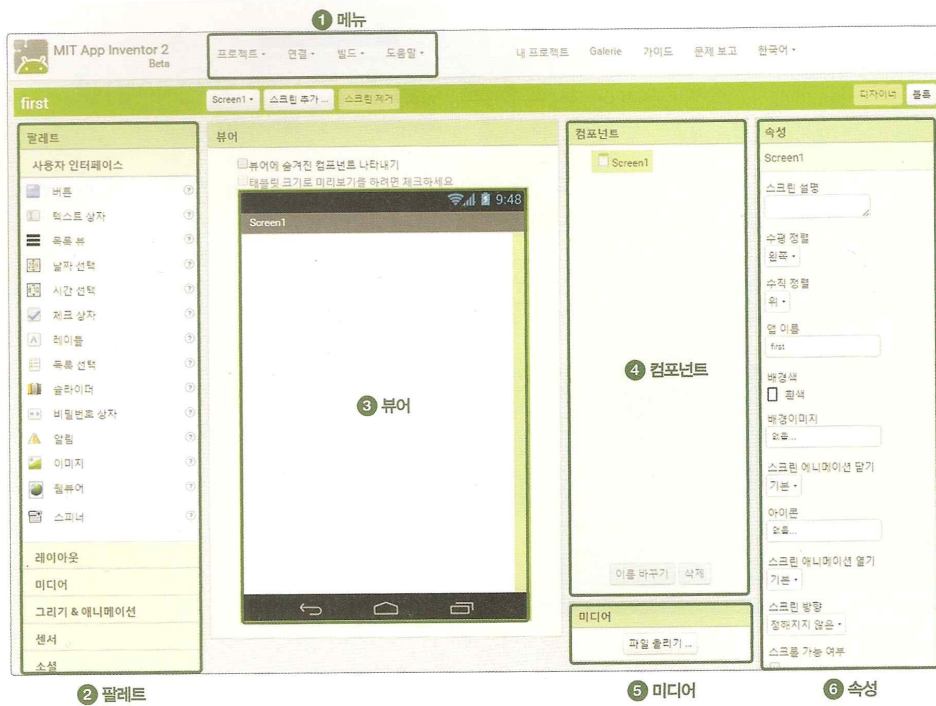
☑ 데이터 과학과 앱인벤터: 앱인벤터로 데이터 과학 맛보기-기본 기능 익히기

<앱인벤터로 그래프 만들기>

1. 앱인벤터(<http://appinventor.mit.edu/>) 사이트 접속하기

2. 앱인벤터 화면 구성

가. 디자이너 화면

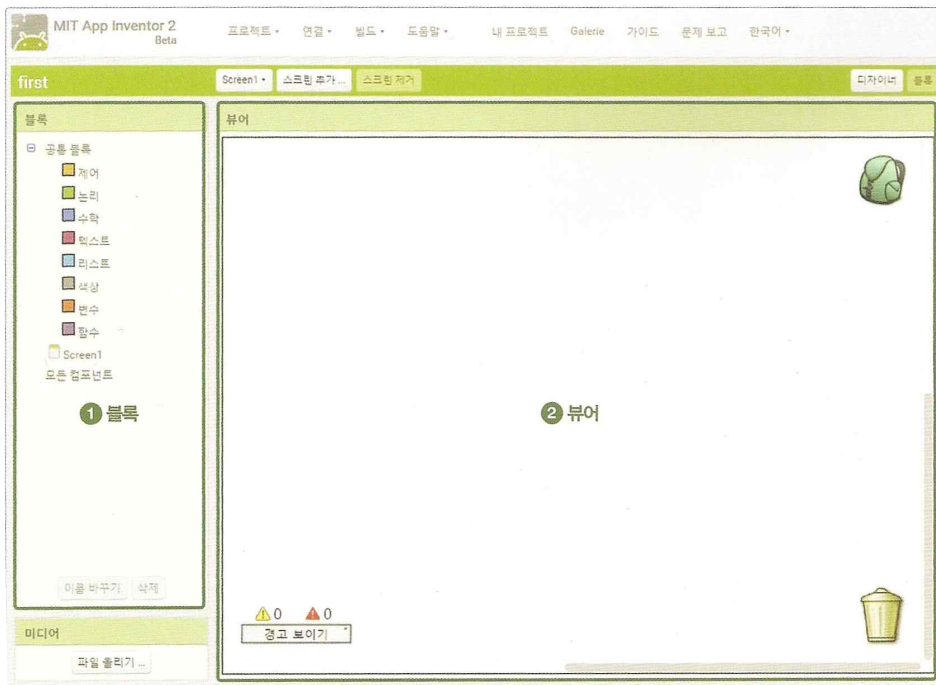





- ① 메뉴: 새 프로젝트 시작하기, 가져오기, 저장하기, 테스트해보기, 앱 만들기 등의 기능을 이용할 수 있음.
- ② 팔레트: 앱에 필요한 요소, 즉 컴포넌트를 모아 놓은 곳.
- ③ 뷰어: 사용자가 만들 앱에 필요한 컴포넌트들을 배치하는 곳
- ④ 컴포넌트: 사용자가 만들 앱에 필요한 컴포넌트 목록을 나열한 곳
- ⑤ 미디어: 사용자가 만들 앱에 필요한 미디어 목록을 나열한 곳
- ⑥ 속성: 컴포넌트에 대한 속성을 설정하는 곳

나) 블록 에디터 화면

- ① 블록: 앱의 동작을 설정하기 위해 사용 가능한 블록들을 모아놓은 곳
- ② 뷰어: 블록을 이용해서 앱의 동작을 설정하는 영역





로 배우는

데

이

터

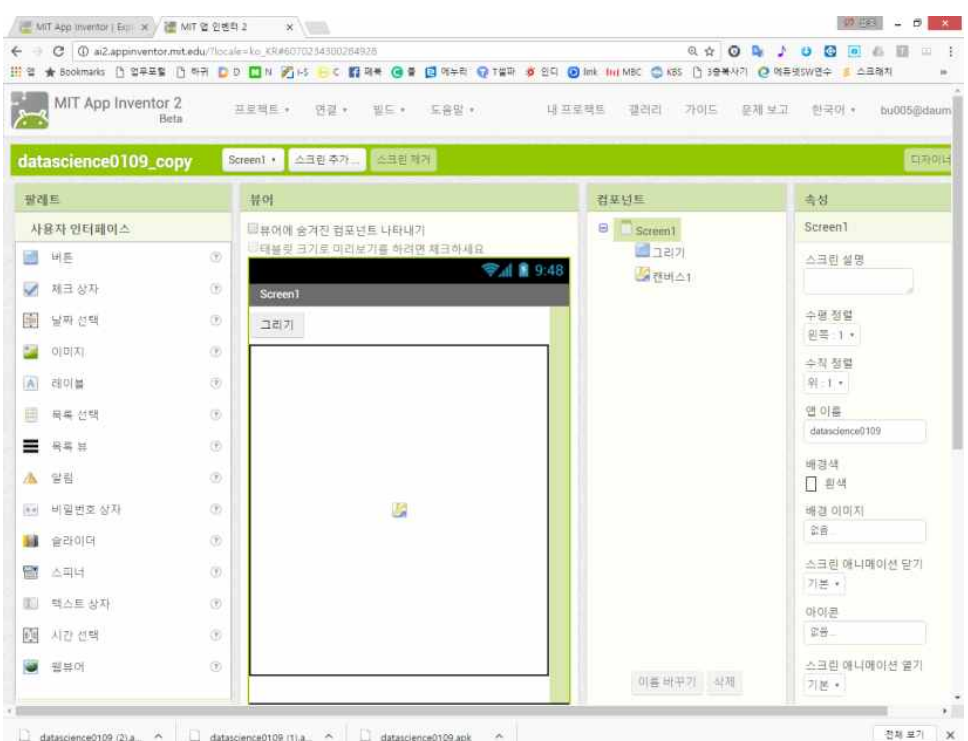
과

학



수업일 :


이름 :

1



The screenshot shows the MIT App Inventor 2 web interface. The browser address bar shows 'ai2.appinventor.mit.edu'. The project name is 'datascience0109_copy'. The interface is divided into several panels: '팔레트' (Palette) on the left with various UI components like buttons, checkboxes, and text boxes; '뷰어' (Viewer) in the center showing a mobile device simulation with a '그리기' (Canvas) component; '컴포넌트' (Components) on the right showing a tree view of the app's structure; and '속성' (Properties) on the far right showing settings for the selected 'Screen1' component, such as '스크린 설명' (Screen description) and '배경색' (Background color).



전역변수 초기화 (list) 값

```

언제 Screen1 초기화
실행
호출 [선택박스1] 지우기
지정하기 global list 값
리스트 만들기 50
70
80
85
95
    
```

```

함수 (x1, y1)
실행
호출 [선택박스1] 선 그리기
x1 가져오기 x1
y1 300 - 가져오기 y1 * 2
호출 [선택박스1] 글자 쓰기
텍스트 가져오기 y1
x 가져오기 x1
y 가져오기 y1
    
```

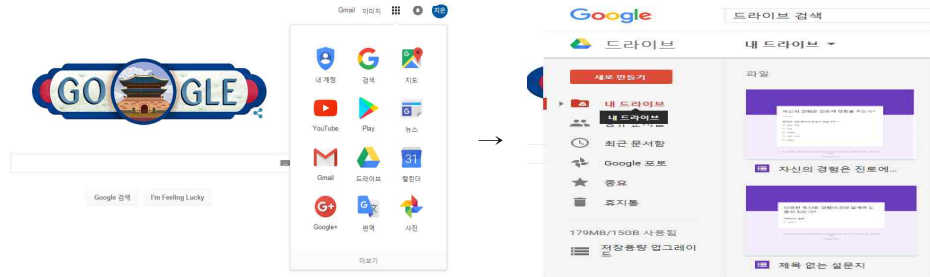
```

언제 [선택박스1] 클릭
실행
지정하기 [선택박스1] [선택박스1] 값
지정하기 [선택박스1] [선택박스1] 값 20
호출 [선택박스1] 지우기
호출 [선택박스1]
x1 50
y1 리스트에서 항목 선택하기 리스트 가져오기 global list 위치 1
호출 [선택박스1]
x1 100
y1 리스트에서 항목 선택하기 리스트 가져오기 global list 위치 2
호출 [선택박스1]
x1 150
y1 리스트에서 항목 선택하기 리스트 가져오기 global list 위치 3
호출 [선택박스1]
x1 200
y1 리스트에서 항목 선택하기 리스트 가져오기 global list 위치 4
호출 [선택박스1]
x1 250
y1 리스트에서 항목 선택하기 리스트 가져오기 global list 위치 5
    
```



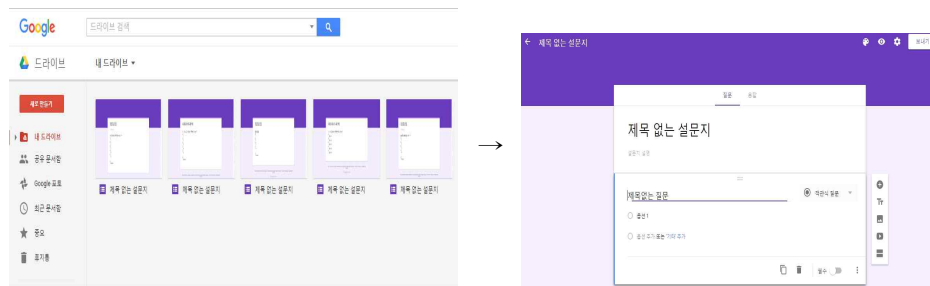
☑ 데이터 과학과 설문: 구글을 이용하여 설문지 만들기

1. 네이버에 '구글'을 치거나 주소창에 <https://www.google.co.kr/> 를 친다.
2. 아이디와 비밀번호를 넣어 구글 계정을 만든다. 아이디는 기존에 사용하는 아이디를 넣는다.
(내 아이디: _____, 내 비밀번호: _____)
3. 로그인한다.



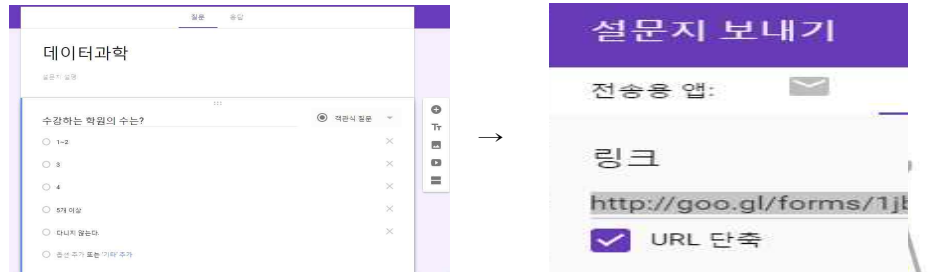
4. 드라이브를 찾아 클릭한다.

5. 왼쪽 '새로 만들기' 를 클릭한다.



6. 구글 설문지를 클릭한다.

7. 설문지 제목을 입력한다.



8. 설문지 응답 내용을 입력한다.

10. 설문지를 보낸다. (링크시킨다.)

데이터과학

수강하는 학원의 수는?

1-2

3

4

5개 이상

다니지 않는다.

저장

응답 10개

요약 개별 보기

수강하는 학원의 수는? (응답

11. 응답자는 주소를 클릭하면 질문에 응답할 수 있다.

12. 응답자가 전부 응답하면 설문 결과를 볼 수 있다.



★ 과제는 다음과 같이 올립니다.

네이버카페/창의컴퓨터 교실 가입/왼쪽 탭 교육기부11기/데이터 과학/“1일차 학습내용 및 과제” 글의 답글로 글쓰기- 제목에(~ 초 000), 아이디는 실명(원래 이름)-내용에는 구글 주소를 붙임

★ 과제 1

데이터 과학에 대한 내용을 주제로 정하여, 구글 설문지로 객관식 한 문항(오지선다)을 만들고 완성한 설문지의 주소를 카페에 올리시오.

예시) 1. 데이터 과학의 의미에 대해 어느 정도 알고 있습니까?

- ① 매우 잘 알고 있다. ② 잘 알고 있다. ③ 보통이다. ④ 잘 모른다.
- ⑤ 전혀 모른다.

★ 과제 2

앱인벤터를 활용하여 시험 결과를 입력하여 그래프를 그리는 앱을 완성하고 카페에 올리시오.





로 배우는

데

이

터

과

학

수업일 :

이름 :

2

학습목표

자동으로 데이터를 수집하여 수학여행 계획을 세울 수 있다.

생각해 보기

제주초등학교는 수학여행으로 1박2일 제주도 일주를 하기로 하였다. 여러 박물관과 아름다운 자연경관 등 제주도의 관광명소가 많고, 맛있는 음식들도 많다.

기본경비와 시간 등을 고려하여 실속 있는 여행을 떠나기 위해 포털검색과 블로그 등 인터넷을 이용해 교통, 숙박, 여행지에 대한 정보를 찾아보고, 이 많은 정보를 어떻게 정리하여 최적의 여행코스를 계획할 수 있을까?

☑ 데이터 문제 정의하기

여행에 필요한 경비를 절약하는 다양한 방법을 이해하여 합리적으로 소비하는 자세를 가질 수 있는가?

1. 문제 정의하기

- 목표: 여행에 필요한 경비를 절약하는 다양한 방법을 이해하여 합리적으로 소비하는 자세를 가질 수 있다.
- 범위: 초등학교 4-6학년 학생
- 문제 해결의 제약 조건(시간과 비용): 즉시 또는 전날, 비용은 무료.

2. 데이터 정의하기

- 데이터의 제한: 여행은 제주도내, 4인 가족 기준으로 제한한다.
- 데이터를 어떻게 수집할 것인가: 각 체험장소별 홈페이지, 블로그 등 활용 (참가 학생 모두 스마트폰이 있을 때는 SNS를 활용한 방법도 생각해 볼 수 있다.)





로 배우는

데

이

터

과

학

수업일 :

이름 :

2

- 데이터 처리 및 분석을 위한 최적의 방법과 도구: 엑셀 프로그램을 활용한 분석 및 시각화
- 최종 결과물은 어떤 형태로 누구에게 전달되어야 하는가: 결과보고서, 보호자와 학생

3. 연구 설문 및 가설 정의하기

- 주어진 많은 정보를 어떻게 정리하면 최적의 여행코스를 계획할 수 있을까?

데이터 수집하기: 필요한 데이터 만들기-자동수집

1. 데이터 수집의 기본

○ 관련된 항목 생각해 보기 ○







2. 수집방법 결정하기(자동수집)

step 1) 인터넷에서 정보를 찾아 셀 서식의 맞춤과 글꼴을 이용하여 정리하기
제주 주요 박물관 홈페이지를 검색하여 정보를 수집한다.

step 2) 응용 및 개별 과제: 통계청에서 건강과 관련된 통계 자료를 찾아 조건부서식을 이용한 문서꾸미기



3. 데이터 품질 점검하기

- 완전성(문제 해결에 필요한 내용이 들어 있느냐?)
- 정확성(문제가 되는 상황이 들어 있느냐?)
- 일관성(데이터 항목들이 연관이 있고 일치하느냐?)

4. 데이터 준비하기

5. 데이터 추가, 선택, 집계하기

- 데이터 추가 및 선택은 필요시에 함.

☑ 현상 이해하기: 탐색적 데이터 분석

1. 원본 데이터 잘 살펴보고 시각화하기

(데이터를 전체적으로 살펴보고 시각화를 통해 나타나는 패턴을 한눈에 살펴봄)

○ 알 수 있는 점 ○

 _____

 _____

 _____





로 배우는

데

이

터

과

학

수업일 :

이름 :

2

2. 개별 속성 분석하기

○ 알 수 있는 점 ○







3. 속성간의 관계 분석하기(의미 있는 상관관계 찾기)

○ 알 수 있는 점 ○







현상 일반화하기: 통계적 추론

1. 내가 세운 가설을 대상으로 비교 및 결론 내리기

○ 결론 ○











로 배우는

데

이

터

과

학

수업일 :

이름 :

2

2. 앞으로의 예측(따라서 앞으로 ~한다면 ~게 될 것으로 생각되므로 ~기 위한 ~에 대한 노력이 필요하겠다. 이번 조사를 바탕으로 이런 부분을 추가해서 더 알아보고 싶다 등)

○ 예측 ○








결과 소통: 발표하고 공유하기

- 서로 의견 나누기

데이터 과학 보고서: 오늘 실습한 내용을 바탕으로 보고서 써보기

1. 주제 선택 및 탐구 동기





로 배우는

데

이

터

과

학

수업일 :

이름 :

2

2. 탐구과정

가. 문제 정의

- 문제:

- 목표:

- 가설:

나. 데이터 수집

- 방법:

- 항목(내용):

다. 탐색적 데이터 분석 방법(실제 완성 때는 차트 등도 삽입해야 함. 이번은 연습)

- 원본 데이터 살펴보고 시각화:


- 개별 속성 분석:

- 속성간의 관계:

라. 통계적 추론(결론 및 예측, 짧게 간추려서)

- 결론:

- 예측:



3. 느낀 점(소감)



★ 과제는 다음과 같이 올립니다.

네이버-창의컴퓨터 교실 가입-왼쪽 탭 11기-데이터 과학 “2
일차 과제” 글의 답글로 글쓰기

★ 과제 1

개별 데이터: 평소에 내 자신이 고쳤으면 하는 점이나 궁금한 점은 무엇이 있을까?

개별 데이터를 이용하여 본인이 해결하고 싶은 주제 1가지를 정하고 그와 관련된 구글 설문지를 만드시오. 설문지의 문항은 3문항 ~ 5문항 정도로 객관식 문항으로 만드시오. 만들고 난 뒤에 완성한 설문지의 주소를 카페에 올리시오.

예시) 주제가 나의 생활습관과 관련된 것이라면

1. 나는 1주일에 늦잠을 몇 번 잘까?

- ① 늦잠을 자지 않는다. ② 1회 ③ 2회 ④ 3회 ⑤ 4회 이상





로 배우는

데

이

터

과

학

수업일 :

이름 :

2

★ 과제 2

한글 프로그램을 이용하여 다음의 양식에 맞게 한글 문서를 작성하여 카페에 올리시오.(주제는 과제 1번과 동일함. 5일차 수업에 사용할 내용이므로 오래 생각해서 잘 정리해 올 것.)

내가 만드는 데이터 과학 보고서			
주제		제출일	
		이름	
< 주제 선정 이유(동기) 및 내용 >			
< 단계별 데이터 과학 분석 >			

◆ 1단계- 문제 정의 ◆

1. 문제:
2. 목표:
3. 가설:

◆ 2단계- 데이터 수집 방법 ◆

1. 수집 방법
 - 가. 대상:
 - 나. 방법:
2. 설문지 내용

◇ 양식 파일은 카페에 있음





로 배우는

데

이

터

과

학

수업일 :

이름 :

3



진동을 감지하여 스스로 데이터를 수집하는 앱을 만들고 지진에 대처하는 생활을 실천할 수 있다.

생각해 보기

제주초등학교에 다니는 ○○이는 최근 우리나라에도 지진이 많이 발생하고 있고 우리가 사는 제주도에도 근처에서도 지진이 발생하고 있다는 뉴스를 보았다. 학교에서 지진발생시 대처훈련을 많이 해오고 있기 때문에 우리가 생활하는 낮 시간에 지진이 발생한다면 훈련받은 대로 행동을 하면 되지만, 우리가 자고 있는 시간에 지진이 발생한다면 끔찍한 결과가 초래될 수도 있다는 생각에 ○○이는 걱정이 많다.

☑ 데이터 문제 정의하기: 우리나라에 자주 ()이 일어난다.

1. 문제 정의하기

- 목표: 나에게 ()가 자주 일어나는 ()를 알아보고, () 생활을 위해 노력해야 할 점을 실천할 수 있다.
- 범위: ()에서 일어나는 지진
성공 및 실패의 척도: 안전사고 횟수를 한 달에 1회 이하로 줄인다.
- 문제 해결의 제약 조건(시간과 비용): 한 달, 비용은 무료.



2. 데이터 정의하기

- 데이터의 제한: ()에서 발생하는 지진으로 제한한다.
진도는() 이상으로 제한한다.
- 데이터를 어떻게 수집할 것인가: () (워드 또는 한글 문서 사용)
- 데이터 처리 및 분석을 위한 최적의 방법과 도구: () 프로그램을 활용한 분석 및 시각화
- 최종 결과물은 어떤 형태로 누구에게 전달되어야 하는가?
: 결과보고서, ()

3. 연구 설문 및 가설 정의하기

- 연구 설문(무엇을 알아내고 싶은가?): 지진이 자주 발생하는 (), 진도 5 이상 지진이 일어난 ()와 횟수
- 가설: ()지역에 지진이 발생할 것이다. 그 이유는 ()

☑ 데이터 수집하기: 필요한 데이터 스스로 만들기

1. 데이터 수집의 기본

- 관찰결과를 개별 항목으로 하는 테이블로 만들기
- 각 열은 개별 항목이 갖는 속성, 각 행은 수집단위가 되는 개별 항목

○ 관련된 항목 생각해 보기 ○

 _____

 _____

 _____





로 배우는

데

이

터

과

학

수업일 :

이름 :

3

2. 수집방법 결정하기(수동수집-만들어가는 데이터)

○ 항목별 체크리스트 예시 ○

지진 발생 목록										
내용	년도									
	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
()										
제주도에 발생한 진도 2~3 지진 횟수										
()										

3. 데이터 품질 점검하기

- 완전성(문제 해결에 필요한 내용이 들어 있느냐?)
- 정확성(문제가 되는 상황이 들어 있느냐?)
- 일관성(데이터 항목들이 연관이 있고 일치하느냐?)

4. 데이터 준비하기

- 테이블 형태로 만들기

5. 데이터 추가, 선택, 집계하기

- 데이터 추가 및 선택은 필요시에 함.

☑ 현상 이해하기: 탐색적 데이터 분석

1. 원본 데이터 잘 살펴보고 시각화 시키기

(데이터를 전체적으로 살펴보고 시각화를 통해 나타나는 패턴을 한눈에 살펴봄)





원본 데이터 잘 살펴보기

안전사고, 왜 나에게만?						
기간	내용	사고횟수	물건 던지기	뛰어다님	걸터앉음	지키는 것 귀찮음
1주		1	3	4	1	3
2주		2	3	4	2	3
3주		3	4	5	1	4
4주		4	5	5	2	5

○ 알 수 있는 점 ○


 _____

 _____


 _____

2. 개별 속성 분석하기

○ 알 수 있는 점 ○

 _____

 _____

 _____





로 배우는

데

이

터

과

학

수업일 :

이름 :

3

3. 속성간의 관계 분석하기(의미 있는 상관관계 찾기)

○ 알 수 있는 점 ○







현상 일반화하기: 통계적 추론

1. 내가 세운 가설을 대상으로 비교 및 결론 내리기

○ 결론 ○







2. 앞으로의 예측(따라서 앞으로 ~한다면 ~게 될 것으로 생각되므로 ~기 위한 ~에 대한 노력이 필요하겠다. 이번 조사를 바탕으로 이런 부분을 추가 해서 더 알아보고 싶다 등)

○ 예측 ○









결과 소통: 발표하고 공유하기

- 서로 의견 나누기

데이터 과학 보고서: 오늘 실습한 내용을 바탕으로 보고서 써보기

네이버카페/창의컴퓨터 교실/교육기부11기/데이터 과학/ “3일차 학습내용” 글의 답글로 글쓰기-제목에(~초 000), 아이디는 실명(원래 이름), 첨부파일에 작성한 보고서를 넣음



로 배우는

데

이

터

과

학

수업일 :

이름 :

3



★ 과제는 다음과 같이 올립니다.

네이버카페/창의컴퓨터교실/교육기부11기/데이터 과학/“3일차 과제” 글의 답글로 글쓰기

★ 과제 1

개별 데이터: 생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 보기
<평소에 나와 관련된 일 중 고쳤으면 하는 점 또는 궁금한 점은 어떤 것이 있을까?>

개별 데이터를 이용하여 본인이 해결하고 싶은 주제 1가지를 정하고 그와 관련된 구글 설문지를 만드시오. 설문지의 문항은 3문항 ~ 5문항 정도로 객관식 문항으로 만드시오. 만들고 난 뒤에 완성한 설문지의 주소를 카페에 올리시오.

예시) 주제가 바른 자세와 학습의 관계에 관련된 것이라면,

1. 바른 자세로 앉으면 건강(공부)에 도움이 된다.

① 전혀 아니다. ② 아니다. ③ 보통이다. ④ 그렇다.

⑤ 매우 그렇다.

★ 과제 2

한글 프로그램을 이용하여 보고서 양식에 맞게 한글 문서를 작성하여 카페에 올리시오.(주제는 과제 1번과 동일함. 5일차 수업에 사용할 내용이므로 오래 생각해서 잘 정리해 올 것.)

◇ 양식 파일은 카페에 있음





로 배우는

데

이

터

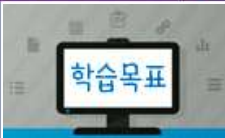
과

학

수업일 :

이름 :

4



스스로 데이터를 수집하여 습관이 건강에 미치는 영향을 알아보고 바람직한 생활 습관에 대해 생각해 볼 수 있다.

생각해 보기

○○이는 요즘 부쩍 살이 찌고 있어서 걱정이다. 처음에 어른들은 건강해 보인다고 좋아하셨지만 점점 더 불어가면서 몸을 보며 요즘은 살 좀 빼야겠다고 이야기 하신다. 초등학생들의 생활 습관과 비만과의 관계를 파악하여 어떤 관계가 있는지 알아보고, 생활 습관을 고쳐나가 보려고 한다.

☑ 데이터 문제 정의하기

청소년들의 안전을 위한 사망 원인을 분석하고 대안 마련하기

1. 문제 정의하기

- 목표: ()
- 범위: ()
- 문제 해결의 제약 조건(시간과 비용): 즉시 또는 전날, 비용은 무료.

2. 데이터 정의하기

- 데이터의 제한:
- 데이터를 어떻게 수집할 것인가? () 설문
- 데이터 처리 및 분석을 위한 최적의 방법과 도구:
()을 활용한 분석 및 시각화
- 최종 결과물은 어떤 형태로 누구에게 전달되어야 하는가?
결과보고서, ()



3. 연구 설문 및 가설 정의하기

- 연구 설문(무엇을 알아내고 싶은가?): ()
- 가설: ()

☑ 데이터 수집하기: 필요한 데이터 스스로 만들기

1. 데이터 수집의 기본

- 관찰결과를 개별 항목으로 하는 테이블로 만들기
- 각 열은 개별 항목이 갖는 속성, 각 행은 수집단위가 되는 개별 항목

○ 관련된 항목 생각해 보기 ○

 _____

 _____

 _____

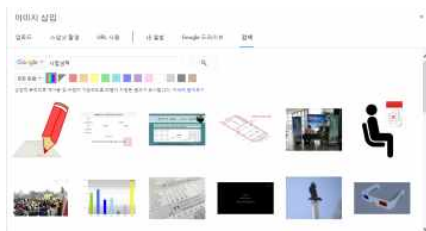
2. 수집방법 결정하기(수동수집- 만들어가는 데이터)

○ 설문지 내용 예시 ○

- 본인은 몇 학년인가요?
 ① 4학년 ② 5학년 ③ 6학년
- 수강하는 학원의 수는?
 ① 다니지 않는다. ② 1개 ③ 2개 ④ 3개 ⑤ 4개 이상
- 하루 평균 학원에서 보내는 시간은?
 ① 없다. ② 1시간 ③ 2시간 ④ 3시간 ⑤ 4시간 이상
- 학업 성취도 평가 평균 점수는? (올해 1학기 기준)
 ① 80점미만 ② 80~85점 미만 ③ 85~90점 미만 ④ 90~95점미만
 ⑤ 95점 이상
- 학원을 다니기 전후 성적변화는?
 ① 떨어짐 ② 변화 없음 ③ +1~+4점 ④ +5~+9점 ⑤ +10점 이상
- 학원을 다니면 성적에 도움이 된다고 생각하나요?
 ① 도움이 된다. ② 도움이 되지 않는다.

▣ 구글을 사용한 설문

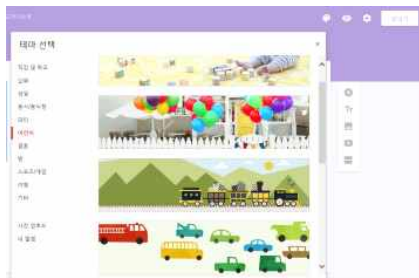
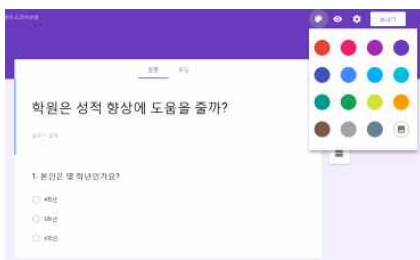
이미지 및 동영상 추가, 전체 배경 바꾸기



구글 설문 작성 중 오른쪽 메뉴 - 이미지 추가 검색 등을 하여 관련 이미지 찾기 - 선택



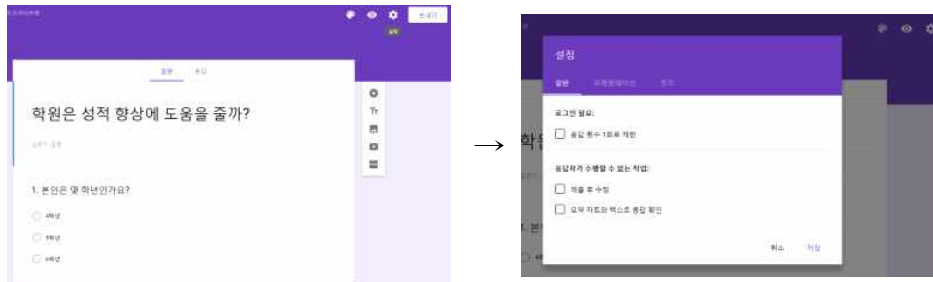
구글 설문 작성 중 오른쪽 메뉴 - 동영상 추가 동영상 검색 후 - 선택



색상 팔레트-색깔 또는 이미지 가져오기 이미지 가져오기-선택 (전체배경 바뀜)



설정 바꾸기(1인당 1번만 설문에 응시할 수 있게 하여 오류 방지, 응답자가 결과 확인 가능)



위 탭에서 설정 클릭 응답횟수 1회 제한, 요약 차트와 텍스트 응답 확인에 체크

응답 결과 가져오기



설문지 응답 클릭-결과 확인





로 배우는

데

이

터

과

학

수업일 :

이름 :

3. 데이터 품질 점검하기

- 완전성(문제 해결에 필요한 내용이 들어 있느냐?)
- 정확성(문제가 되는 상황이 들어 있느냐?)
- 일관성(데이터 항목들이 연관이 있고 일치하느냐?)

4. 데이터 준비하기

- 테이블 형태로 변환하기
- 열에 넣을 항목: ()
- 행에 넣을 항목: ()

★ 스스로 해 보기 1 ★ 설문 내용을 한글 문서를 이용하여 테이블 형식으로 만들어 보세요.

5. 데이터 추가, 선택, 집계하기

- 데이터 추가 및 선택은 필요시에 함.


☑ 현상 이해하기: 탐색적 데이터 분석

1. 원본 데이터 잘 살펴보고 시각화 시키기

(데이터를 전체적으로 살펴보고 시각화를 통해 나타나는 패턴을 한눈에 살펴봄)

- 시각화 시키기(패턴 살피기)

학년	학원 수	학원에서 보내는	평가 점수	성적변화	학원은 도움이 되
6학년	0	0	92	0	그렇지 않다
5학년	0	0	95	0	그렇지 않다
6학년	1	3	95	2	그렇다
4학년	2	3	92	7	그렇지 않다
6학년	2	2	95	7	그렇다
4학년	1	2	95	0	그렇지 않다
5학년	0	0	95	0	그렇지 않다
6학년	1	2	95	0	그렇지 않다
6학년	0	0	95	0	그렇지 않다
6학년	4	3	95	2	그렇다
5학년	3	2	95	2	그렇다
4학년	4	4	87	2	그렇다
6학년	1	2	95	0	그렇다, 그렇지 않다
6학년	2	2	82	-2	그렇지 않다
6학년	3	3	92	7	그렇다



로 배우는

데

이


터


과


학

수업일 : 이름 : 4

○ 알 수 있는 점 ○










★ 스스로 해 보기 2 ★ 데이터를 시각화시켜 보세요.

○ 알 수 있는 점 ○











2. 개별 속성 분석하기

○ 알 수 있는 점 ○









3. 속성간의 관계 분석하기(의미 있는 상관관계 찾기)

○ 알 수 있는 점 ○

 _____

 _____

 _____

☑ 현상 일반화하기: 통계적 추론

1. 내가 세운 가설을 대상으로 비교 및 결론 내리기

○ 결론 ○

 _____

 _____

 _____

2. 앞으로의 예측(따라서 앞으로 ~한다면 ~게 될 것으로 생각되므로 ~기 위한 ~에 대한 노력이 필요하겠다. 이번 조사를 바탕으로 이런 부분을 추가 해서 더 알아보고 싶다 등)

○ 예측 ○

 _____

 _____

 _____





결과 소통: 발표하고 공유하기

네이버 카페/창의컴퓨터 교실/교육기부11기/데이터 과학/ “4일차 학습내용” 글의 답글로 글쓰기-제목에(~초 000), 아이디는 실명(원래 이름), 첨부파일에 작성한 보고서를 넣음



※ 과제제출방법: 네이버카페/창의컴퓨터 교실/교육기부11기/ 데이터 과학/ “4일차 과제” 글의 답글로 글쓰기

★ 과제 1

개별 데이터: 생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 보기 2

<평소에 궁금했던 점 해결해 보기>

교육기부에 참여한 학생들을 대상으로 데이터를 만들어 분석하고 싶은 주제 1가지에 따른 설문 문항을 구글 설문지로 작성합니다. 작성한 주소를 게시판에 올리고 참여해 달라는 글을 남기세요.

(전부 저녁 8시 전까지 올립니다.)

★ 과제 2

과제 1을 하고 난 뒤, 저녁 8시가 되면 카페에 들어가 친구들이 올린 글을 방문하여 설문을 응시한 후 댓글로 ‘00초 000 응시함’ 이라고 글을 남깁니다. 반 친구 모두 설문을 모두 해 주고 댓글로 전부 달아주어야 합니다. 좀 번거롭겠지만 친구들의 발표를 위해서 협조해 주세요!!

(설문은 한번만 하고 제출해 주세요. 여러 번 하면 신뢰도가 떨어져요~)





로 배우는

데

이

터

과

학

수업일 :

이름 :

5



교육기부 참가 학생들을 대상으로 평소에 궁금했던 점을 생각하여 주제를 정하고 스스로 데이터를 수집하여 이를 해결할 수 있다.

생각해 보기

평소에 궁금했던 점은 무엇이 있을까? 또는 내 주변에서 고쳤으면 하는 점이나 궁금한 점은 무엇이 있을까? 생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 보자.

☐ 문제 정의

나와 내 주변에서 일어나는 일을 중심으로 주제 선정하기

1. 내 자신이 고쳤으면 하는 점이나 궁금한 점을 생각해 보자.

○ 예시 ○

- 나는 왜 공부하는 시간이 많은데 성적이 안 오를까?
- 나는 왜 많이 먹어도 살이 안 찌까?
- 나는 왜 조금만 먹어도 살이 찌까?
- 나는 왜 배탈이 잘 날까?
- 나는 왜 아침마다 늦잠을 잘까?
- 나는 왜 친구들과 자주 싸울까?

2. 나와 관련된 일 중 고쳤으면 하는 점 또는 궁금한 점이 있으면 생각해 보자.





로 배우는

데

이

터

과

학

수업일 :

이름 :

5

○ 예시 ○

- 악기 연주는 스트레스 해소에 도움을 줄까?
- 바른 자세로 앉으면 건강(공부)에 도움이 될까?
- 나이가 젊은 선생님이 우리들을 더 잘 이해해 줄까?
- 여자일수록 수다가 많을까?
- 엄한 선생님일수록 숙제를 잘 해 올까?
- 형제가 많으면 성격이 좋을까?(다른 사람들을 더 잘 배려하고)
- 남자들은 여자보다 무서움이 없을까?
- 성적이 좋을수록 스트레스가 적을까?

3. 자신이 탐구해 보고 싶은 주제 및 그 주제를 결정한 이유를 써 보자.

(★ 붙임 보고서에 정리)

데이터 수집: 필요한 데이터 스스로 만들기

(★ 붙임 보고서에 정리)

현상 이해하기: 탐색적 데이터 분석

1. 원본 데이터 잘 살펴보고 시각화 시키기

(★ 붙임 보고서에 정리)



2. 개별 속성 분석하기

(★ 붙임 보고서에 정리)

3. 속성간의 관계 분석하기(의미 있는 상관관계 찾기)

(★ 붙임 보고서에 정리)

현상 일반화하기: 통계적 추론

1. 내가 세운 가설을 대상으로 비교 및 결론 내리기

(★ 붙임 보고서 자료에 정리하기)

2. 앞으로의 예측


(따라서 앞으로 ~한다면 ~게 될 것으로 생각되므로 ~기 위한 ~에 대한 노력이 필요하겠다. 이번 조사를 바탕으로 이런 부분을 추가해서 더 알아보고 싶다 등)

(★ 붙임 보고서에 정리)

결과 소통: 발표하고 공유하기 (느낀 점)

(★ 붙임 보고서에 정리)





로 배우는

데

이

터

과

학


수업일 :

이름 :

5


☑ 데이터 과학 보고서: 오늘 실습한 내용을 바탕으로 보고서 써보기

네이버 카페/창의컴퓨터 교실/교육기부11기/데이터 과학/ **“5일차 학습내용”**
 글의 답글로 글쓰기-제목에(~초 000), 아이디는 실명(원래 이름), 첨부파일에
 작성한 보고서를 넣음




★ **과제 1**
 완성한 보고서를 카페에 올립니다.

★ **과제 2**
 작성한 보고서의 내용 중 필요한 부분을 간추려 큰 목소리로 또박또박 발표
 하는 연습을 합니다. (시간은 5분 이내)



<부록 4> 스크래치를 활용한 데이터 과학 교육 프로그램



로 배우는

데 이 터 과 학

수업일 : 이름 : 1

학습목표


데이터 과학의 의미를 이해하고 이를 위한 스크래치의 기본 기능을 익힐 수 있다.

데이터 과학이란? 데이터 과학의 의미 알기

1. 데이터 과학의 의미

2. 데이터 과학의 예

3. 데이터 과학을 시작하기 위해 필요한 것





로 배우는

데

이

터

과

학

수업일 :

이름 :

1

4. 데이터 과학의 단계

가. 데이터 문제 정의하기







나. 데이터 수집하기







다. 현상 이해하기: 탐색적 데이터 분석







라. 현상 일반화하기: 통계적 추론











로 배우는

데

이

터

과

학

수업일 :

이름 :

1

☑ 데이터 과학과 엑셀: 스크래치로 데이터 과학 맛보기-기본 기능 익히기

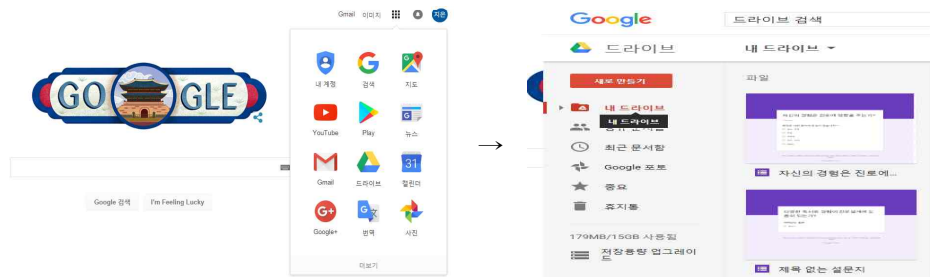
1. 스크래치 웹사이트 회원가입하기
2. 스크래치 화면 살펴보기
3. 스크래치 기본 기능 익히기

☑ 데이터 과학과 설문: 구글을 이용하여 설문지 만들기

1. 네이버에 '구글'을 치거나 주소창에 <https://www.google.co.kr/> 를 친다.
2. 아이디와 비번을 넣어 구글 계정을 만든다. 아이디는 기존에 사용하는 아이디를 넣는다.

(내 아이디: _____ , 내 비번: _____)

3. 로그인한다.



4. 드라이브를 찾아 클릭한다.

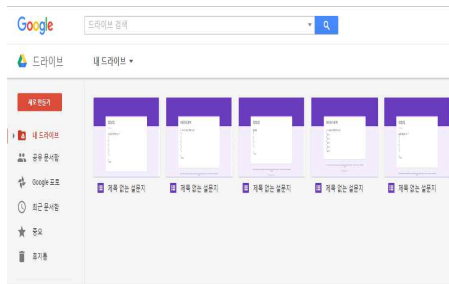
5. 왼쪽 '새로 만들기' 를 클릭한다.



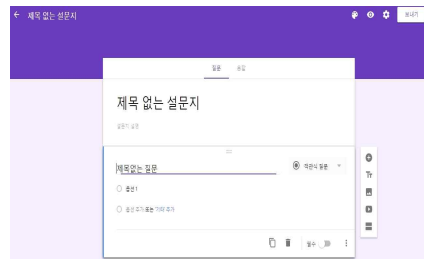
수업일 :

이름 :

1



6. 구글 설문지를 클릭한다.



7. 설문지 제목을 입력한다.



8. 설문지 응답 내용을 입력한다.



9. 설문지를 보낸다. (링크시킨다.)



10. 응답자는 주소를 클릭하면 설문에 응답할 수 있다.



11. 응답자가 전부 응답하면 설문 결과를 볼 수 있다.





로 배우는

데

이

터

과

학

수업일 :

이름 :

1



★ 과제

데이터 과학에 대한 내용을 주제로 정하여, 구글 설문지로 객관식 한 문항 (오지선다)을 만들고 완성한 설문지의 주소를 카페에 올리시오.

예시) 1. 데이터 과학의 의미에 대해 어느 정도 알고 있습니까?

- ① 매우 잘 알고 있다. ② 잘 알고 있다. ③ 보통이다. ④ 잘 모른다.
- ⑤ 전혀 모른다.





로 배우는

데

이

터

과

학

수업일 :

이름 :

2



생활 속에서 궁금한 내용을 공공데이터를 수집하여 해결해 볼 수 있다
 물건 운반 로봇(공 넘기기 로봇)이 생활 속에 필요한 이유를 알고 스크래치를 이용하여 막대 그래프를 그려 볼 수 있다.
 물건 운반 로봇(공 넘기기 로봇) 프로젝트를 순서대로 나타낼 수 있다.

생각해 보기

- ▶ 평소에 궁금했던 점은 무엇이 있을까? 또는 내 주변에서 고쳤으면 하는 점이나 궁금한 점은 무엇이 있을까? 생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 보자.
- ▶ 생활 속에서 물건을 운반하는 로봇 사례를 통해 물건 운반 로봇이 생활 속에 필요한 이유는 무엇일까?

- 아마존 창고를 책임지는 로봇 집꾼 ‘키바’

<https://www.youtube.com/watch?v=gQpMDdJmbNs>(약 2분 41초)





로 배우는

데

이

터

과

학

수업일 :

이름 :

- "택배 왔어요"...택배 로봇 현실 되나

<https://www.youtube.com/watch?v=lch7OFARIlc>(약 37초)



- 대한민국 청소년 로봇스포츠 공 넘기기

<https://www.youtube.com/watch?v=CrR06sbS6X>(약 1분 38초)





로 배우는

데

이

터

과

학

수업일 :

이름 :

☑ 데이터 문제 정의하기

1. 문제 정의하기

- * 문제(문제점이나 궁금한 점이 무엇인가?)
- * 목표(문제점이나 궁금한 점을 해결하여 얻고 싶은 점은 무엇인가?)
- * 가설(문제나 궁금한 점에 대해 내가 생각한 이유는 무엇인가?)

☑ 데이터 수집하기: 필요한 데이터 만들기-자동수집

1. 데이터 수집의 기본

○ 관련된 항목 생각해 보기 ○

 _____

 _____

 _____

2. 수집방법 결정하기(자동수집)

 _____

 _____

 _____





로 배우는

데

이

터

과

학

수업일 :

이름 :

2

3. 데이터 품질 점검하기

- 완전성(문제 해결에 필요한 내용이 들어 있느냐?)
- 정확성(문제가 되는 상황이 들어 있느냐?)
- 일관성(데이터 항목들이 연관이 있고 일치하느냐?)

4. 데이터 준비하기

5. 데이터 추가, 선택, 집계하기

- 데이터 추가 및 선택은 필요시에 함.

☑ 현상 이해하기: 탐색적 데이터 분석

1. 원본 데이터 잘 살펴보고 시각화하기

(데이터를 전체적으로 살펴보고 시각화를 통해 나타나는 패턴을 한눈에 살펴봄)

○ 알 수 있는 점 ○











로 배우는

데

이

터

과

학

수업일 :

이름 :

2

2. 항목 1(빨강)

```

when green flag clicked
  erase
  move to x: -140 y: -135
  hide
  when clicked
    repeat 1 times
      stamp
      change y coordinate by 20
  
```

3. 항목 2(파랑)

```

when green flag clicked
  erase
  move to x: -45 y: -135
  hide
  when clicked
    repeat 2 times
      stamp
      change y coordinate by 20
  
```

4. 항목 3(초록)

```

when green flag clicked
  erase
  move to x: 50 y: -135
  hide
  when clicked
    repeat 3 times
      stamp
      change y coordinate by 20
  
```

5. 항목 4(주황)

```

when green flag clicked
  erase
  move to x: 145 y: -135
  hide
  when clicked
    repeat 4 times
      stamp
      change y coordinate by 20
  broadcast message
  
```





로 배우는

데

이

터

과

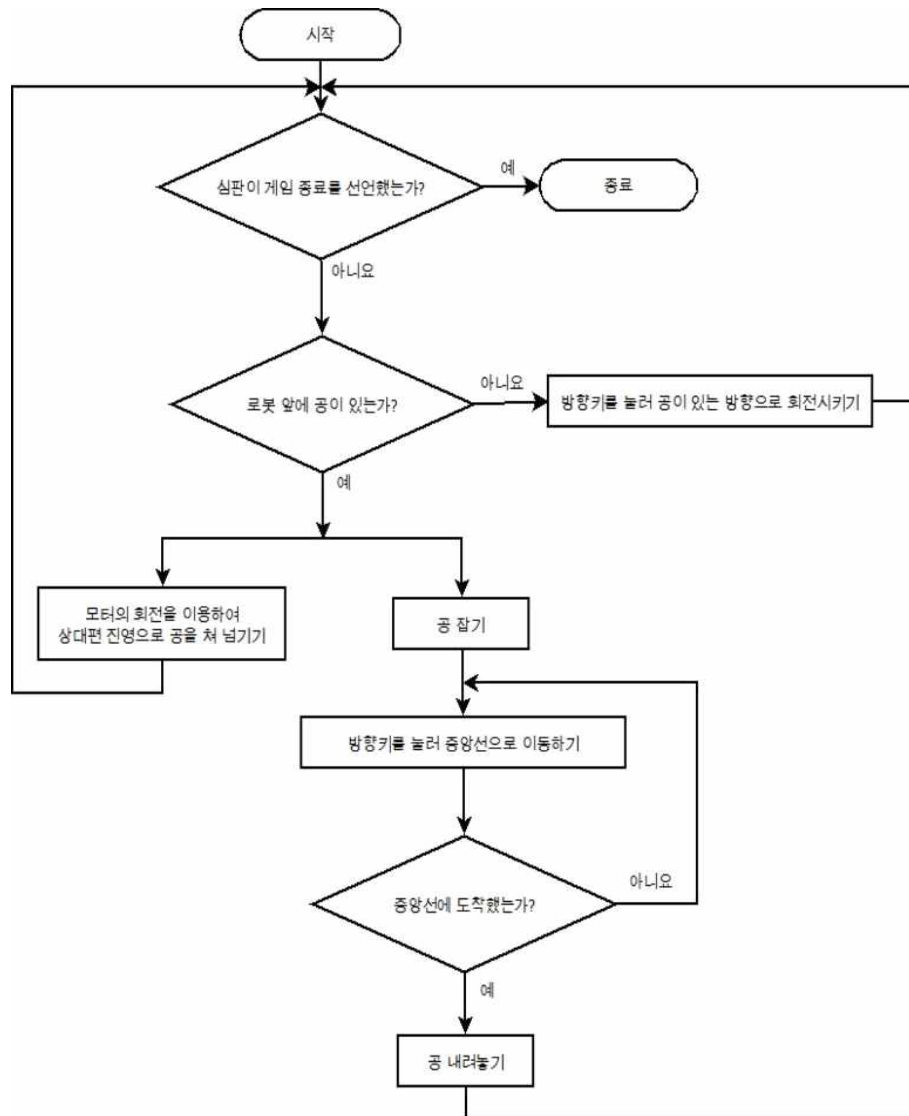
학

수업일 :

이름 :

2

<물건 운반 로봇(공 넘기기 로봇) 프로젝트를 순서도로 나타내기-센서 활용>





로 배우는



수업일 :

이름 :

2

2. 개별 속성 분석하기

○ 알 수 있는 점 ○







3. 속성간의 관계 분석하기(의미 있는 상관관계 찾기)

○ 알 수 있는 점 ○







현상 일반화하기

통계적 추론

1. 내가 세운 가설을 대상으로 비교 및 결론 내리기

○ 결론 ○











로 배우는

데

이

터

과

학

수업일 :

이름 :

2

2. 앞으로의 예측(따라서 앞으로 ~한다면 ~게 될 것으로 생각되므로 ~기 위한 ~에 대한 노력이 필요하겠다. 이번 조사를 바탕으로 이런 부분을 추가해서 더 알아보고 싶다 등)

○ 예측 ○







결과 소통: 발표하고 공유하기

- 서로 의견 나누기

데이터 과학 보고서: 오늘 실습한 내용을 바탕으로 보고서 써보기

1. 주제 선택 및 탐구 동기





로 배우는



수업일 :

이름 :

2

2

2. 탐구과정

가. 문제 정의

- 문제:

- 목표:

- 가설:

나. 데이터 수집

- 방법:

- 항목(내용):

다. 탐색적 데이터 분석 방법(실제 완성 때는 차트 등도 삽입해야 함. 이번은 연습)

- 원본 데이터 살펴보고 시각화:

- 개별 속성 분석:

- 속성간의 관계:

라. 통계적 추론(결론 및 예측, 짧게 간추려서)

- 결론:

- 예측:





로 배우는

데

이

터

과

학

수업일 :

이름 :

2

3. 느낀 점(소감)



★ 과제 1

공공 데이터를 이용하여 본인이 해결하고 싶은 주제 1가지를 정하고 데이터를 수집한 후, 스크래치를 이용하여 탐색적 데이터 분석을 한 결과를 붙임 양식에 기록하여 답글로 올리기

★ 과제 2

<선택 심화> 물건 운반 로봇(공 넘기기 로봇) 프로젝트를 스크래치로 구현해보기





로 배우는

데

이

터

과

학

수업일 :

이름 :

3



생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 볼 수 있다.
 로봇청소기가 생활 속에 필요한 이유를 알고 스크래치를 이용하여 꺾은선그래프를 그려 볼 수 있다.
 로봇청소기 프로젝트를 순서도로 나타낼 수 있다.

생각해 보기

- ▶ 평소에 궁금했던 점은 무엇이 있을까? 또는 내 주변에서 고쳤으면 하는 점이나 궁금한 점은 무엇이 있을까? 생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 보자.
- ▶ 추석을 맞이해서 우리 집에 친척들이 오기로 하였다. 그런데, 엄마는 음식 준비, 아빠는 차례상 준비로 바쁘셔서 나에게 청소를 하라고 하셨다. 하지만 나는 사촌들과 재미있게 놀려면 학원 숙제를 미리 해 놓아야 한다. 만약, 로봇청소기가 있다면 시간을 뺏기지 않고 숙제를 할 수 있을 텐데…….

☑ 데이터 문제 정의하기

1. 문제 정의하기

* 문제(문제점이나 궁금한 점이 무엇인가?)

평소에 내 자신이 고쳤으면 하는 점이나 궁금한 점은 무엇이 있을까?

또는 내 주변에서 고쳤으면 하는 점이나 궁금한 점은 무엇이 있을까?





로 배우는

데

이

터

과

학

수업일 :

이름 :

3

○ 예시 ○

- 나는 왜 공부하는 시간이 많은데 성적이 안 오를까?
- 나는 왜 많이 먹어도 살이 안 찢까?
- 나는 왜 조금만 먹어도 살이 찢까? - 나는 왜 배탈이 잘 날까?
- 나는 왜 아침마다 늦잠을 잘까? - 나는 왜 친구들과 자주 싸울까?
- * 목표(문제점이나 궁금한 점을 해결하여 얻고 싶은 점은 무엇인가?)
- * 가설(문제나 궁금한 점에 대해 내가 생각한 이유는 무엇인가?)

☑ 데이터 수집하기

○ 관련된 항목 생각해 보기 ○

 _____

 _____

 _____

2. 수집방법 결정하기(자동수집)

 _____

 _____

 _____





로 배우는

데

이

터

과

학

수업일 :

이름 :

3

3. 데이터 품질 점검하기

- 완전성(문제 해결에 필요한 내용이 들어 있느냐?)
- 정확성(문제가 되는 상황이 들어 있느냐?)
- 일관성(데이터 항목들이 연관이 있고 일치하느냐?)

4. 데이터 준비하기

5. 데이터 추가, 선택, 집계하기

- 데이터 추가 및 선택은 필요시에 함.

☑ 현상 이해하기: 탐색적 데이터 분석

1. 원본 데이터 잘 살펴보고 시각화 시키기

(데이터를 전체적으로 살펴보고 시각화를 통해 나타나는 패턴을 한눈에 살펴봄)

○ 알 수 있는 점 ○











로 배우는

데

이

터

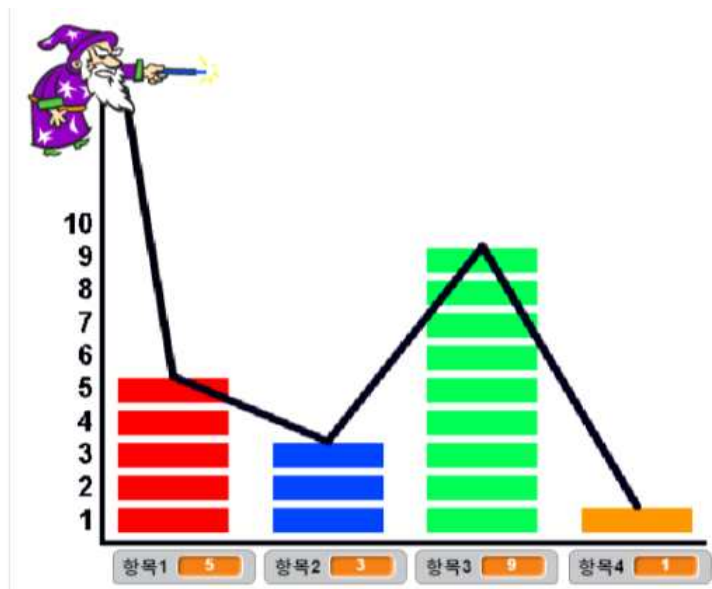
과

학

수업일 :

이름 :

3



1. 마법사

```

클릭했을 때
x: -170 y: 130 로 이동하기
항목1의 응답수는? 묻고 기다리기
항목1 을(은) 대답 로 정하기
항목1그리기 방송하기
항목2의 응답수는? 묻고 기다리기
항목2 을(은) 대답 로 정하기
항목2그리기 방송하기
항목3의 응답수는? 묻고 기다리기
항목3 을(은) 대답 로 정하기
항목3그리기 방송하기
항목4의 응답수는? 묻고 기다리기
항목4 을(은) 대답 로 정하기
항목4그리기 방송하기

읽은선그래프 을(은) 받았을 때
항목1점 위치로 이동하기 x: -170
펜 색깔을 1 부터 100 사이의 난수 (으)로 정하기
펜 굵기를 5 (으)로 정하기
펜 내리기
0.5 초 기다리기
항목2점 위치로 이동하기
0.5 초 기다리기
항목3점 위치로 이동하기
0.5 초 기다리기
항목4점 위치로 이동하기
0.5 초 기다리기
펜 올리기
x: -170 y: 130 로 이동하기
  
```





로 배우는

데

이

터

과

학

수업일 :

이름 :

3

2. 세로축: 명령블록 없이 숫자만

3. 항목1(빨강)

```

when green flag clicked
  clear
  x: -140 y: -135 to position
  hide
  when clicked on item 1
    repeat 1 times
      click
      change y coordinate by 20
  
```

4. 항목 2(파랑)

```

when green flag clicked
  clear
  x: -45 y: -135 to position
  hide
  when clicked on item 2
    repeat 2 times
      click
      change y coordinate by 20
  
```

5. 항목 3(초록)

```

when green flag clicked
  clear
  x: 50 y: -135 to position
  hide
  when clicked on item 3
    repeat 3 times
      click
      change y coordinate by 20
  
```

6. 항목 4

```

when green flag clicked
  clear
  x: 145 y: -135 to position
  hide
  when clicked on item 4
    repeat 4 times
      click
      change y coordinate by 20
  when green flag clicked
    show
  
```





로 배우는

데

이

터

과

학

수업일 :

이름 :

3

2. 개별 속성 분석하기

○ 알 수 있는 점 ○







3. 속성간의 관계 분석하기(의미 있는 상관관계 찾기)

○ 알 수 있는 점 ○







☑ 현상 일반화하기: 통계적 추론

1. 내가 세운 가설을 대상으로 비교 및 결론 내리기

○ 결론 ○











로 배우는

데

이

터

과

학

수업일 :

이름 :

3

2. 앞으로의 예측(따라서 앞으로 ~한다면 ~게 될 것으로 생각되므로 ~기 위한 ~에 대한 노력이 필요하겠다. 이번 조사를 바탕으로 이런 부분을 추가 해서 더 알아보고 싶다 등)

○ 예측 ○







결과 소통: 발표하고 공유하기

-서로 의견 나누기

데이터 과학 보고서: 오늘 실습한 내용을 바탕으로 보고서 써보기



★ 과제 1

개별 데이터를 이용하여 본인이 해결하고 싶은 주제 1가지를 정한 후, 그와 관련된 구글 설문지를 만들고, 데이터를 수집한 후, 스크래치를 이용하여 탐색적 데이터 분석을 한 결과를 붙임 양식에 기록하여 답글로 올리기

<참고> 구글 설문지 만들 때 오른 쪽 위에 색상팔레트/테마를 선택하면 다양한 배경을 추가할 수 있습니다.

★ 과제 2

<선택 심화> 청소기 로봇 프로젝트를 스크래치로 구현해보기





로 배우는

데

이

터

과

학

수업일 :

이름 :

4

학습목표

생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 볼 수 있다.
제주 정낭 프로젝트를 순서대로 나타낼 수 있다.

생각해 보기

- ▶ 평소에 궁금했던 점은 무엇이 있을까? 또는 내 주변에서 고쳐 으면 하는 점이나 궁금한 점은 무엇이 있을까? 생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 보자.
- ▶ 걸쳐져 있는 정낭 수에 따른 의미에 대해 이야기해 보기
3개: 좀 멀리 갔어요.
2개: 저녁 때쯤 돌아와요.
1개: 금방 돌아와요.
0개: 사람 있어요.
- ▶ 제주특별자치도청 홈페이지/분야별 정보 바로가기/문화·역사
<http://www.jeju.go.kr/culture/folklore/samda/stone/townDol.htm>





로 배우는

데

이

터

과

학

수업일 :

이름 :

▶ Korean National Parks-Jeongnang

<https://www.youtube.com/watch?v=OqKJOM2vo6o>



☑ 데이터 문제 정의하기

1. 문제 정의하기

- 목표: ()

○ 예시 ○

- 악기 연주는 스트레스 해소에 도움을 줄까?
- 바른 자세로 앉으면 건강(공부)에 도움이 될까?
- 나이가 젊은 선생님이 우리들을 더 잘 이해해 줄까?
- 여자일수록 수다가 많을까?
- 엄한 선생님일수록 숙제를 잘 해 올까?
- 형제가 많으면 성격이 좋을까? (다른 사람들을 더 잘 배려하고)
- 남자들은 여자보다 무서움이 없을까?
- 성적이 좋을수록 스트레스가 적을까?





로 배우는

데

이

터

과

학

수업일 :

이름 :

4

- 범위: ()
- 문제 해결의 제약 조건:

2. 데이터 정의하기

- 데이터의 제한:
- 데이터를 어떻게 수집할 것인가: () 설문
- 데이터 처리 및 분석을 위한 최적의 방법과 도구: ()을 활용한 분석 및 시각화
- 최종 결과물은 어떤 형태로 누구에게 전달되어야 하는가
: 결과보고서, ()

3. 연구 설문 및 가설 정의하기

- 연구 설문(무엇을 알아내고 싶은가?): ()
- 가설: ()

데이터 수집하기: 필요한 데이터 스스로 만들기

1. 데이터 수집의 기본

- 관찰결과를 개별 항목으로 하는 테이블로 만들기
- 각 열은 개별 항목이 갖는 속성, 각 행은 수집단위가 되는 개별 항목

○ 관련된 항목 생각해 보기 ○











로 배우는

데

이

터

과

학

수업일 :

이름 :

4

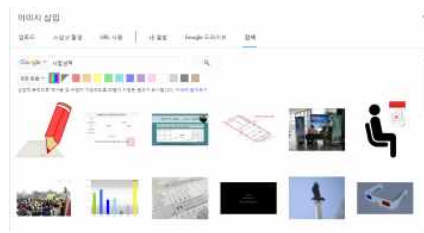
2. 수집방법 결정하기(수동수집-만들어가는 데이터)

▣ 구글을 사용한 설문

이미지 및 동영상 추가, 전체 배경 바꾸기



구글 설문 작성 중 오른쪽 메뉴 -
이미지 추가



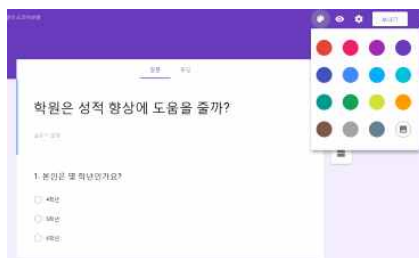
검색 등을 하여 관련 이미지 찾기-
선택



구글 설문 작성 중 오른쪽 메뉴 -
동영상 추가



동영상 검색 후 - 선택



색상 팔레트-색깔 또는 이미지
가져오기



이미지 가져오기-선택 (전체배경
바뀜)





로 배우는

데

이

터

과

학

수업일 :

이름 :

4

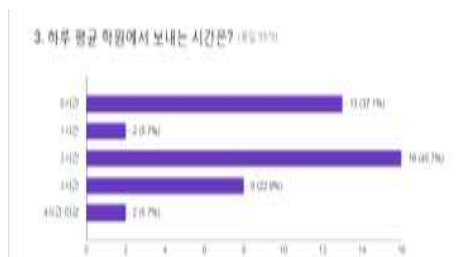
설정 바꾸기(1인당 1번만 설문에 응시할 수 있게 하여 오류 방지, 응답자가 결과 확인 가능)



위 탭에서 설정 클릭

응답횟수 1회 제한, 요약 차트와 텍스트 응답 확인에 체크

응답 결과 가져오기



설문지 응답 클릭-결과 확인





로 배우는



수업일 :

이름 :

3. 데이터 품질 점검하기

- 완전성(문제 해결에 필요한 내용이 들어 있느냐?)
- 정확성(문제가 되는 상황이 들어 있느냐?)
- 일관성(데이터 항목들이 연관이 있고 일치하느냐?)

4. 데이터 준비하기

- 테이블 형태로 변환하기

열에 넣을 항목: ()

행에 넣을 항목: ()

★ 스스로 해 보기 1 ★ 설문 내용을 한글 문서를 이용하여 테이블 형식으로 만들어 보세요.

5. 데이터 추가, 선택, 집계하기

- 데이터 추가 및 선택은 필요시에 함.

현상 이해하기: 탐색적 데이터 분석

1. 원본 데이터 잘 살펴보고 시각화 시키기

(데이터를 전체적으로 살펴보고 시각화를 통해 나타나는 패턴을 한눈에 살펴봄)

- 시각화 시키기(패턴 살펴보기)





로 배우는

데

이

터

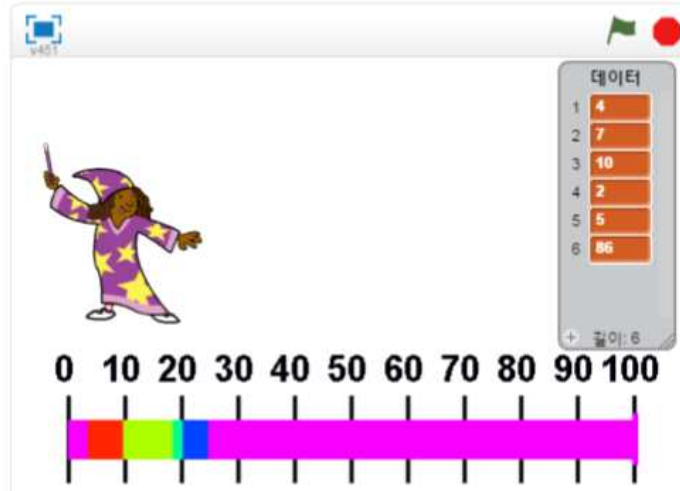
과

학

수업일 :

이름 :

4



1. 마법사

2. 스프라이트1(파랑)

```

클릭했을 때
크기를 70% 로 정하기
트레이드백 항목을 데이터 에서 삭제하기
데이터 개수 음(음) 0 로 정하기
합계 음(음) 0 로 정하기
항목 음(음) 0 로 정하기
몇개의 자료가 남아있나요? 묻고 기다리기
데이터 개수 음(음) 대답 로 정하기
데이터 개수 번 반복하기
항목 음(음) 1 만큼 바꾸기
항목 와 번째 자료를 입력하셔 결합하기 묻고 기다리기
대답 음(음) 항목 번째 데이터 에 넣기
합계 음(음) 대답 만큼 바꾸기
피그렛프 그리기 발송하기
  
```

```

클릭했을 때
피 항목 음(음) 0 로 정하기
x: <200> y: <90> 로 이동하기
피그렛프 그리기 음(음) 발악을 때
피 항목 음(음) 1 만큼 바꾸기
피 항목 > 데이터 개수 까지 반복하기
색깔 요소를 40 만큼 바꾸기
피 항목 번째 데이터 항목 / 합계 * 100 * 4 번 반복하기
1 만큼 움직이기
도움말하기
피 항목 음(음) 1 번 바꾸기
  
```





3. 스프라이트2(검정)

```

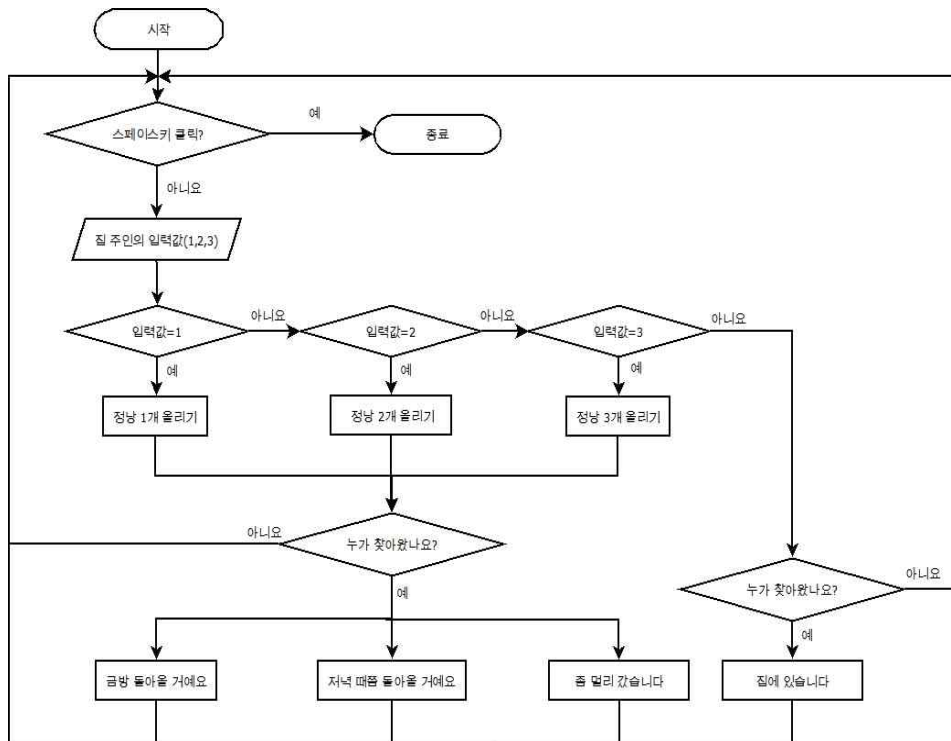
when green flag clicked
  clear
  turn right 90 degrees
  when green flag clicked
  x: -200 y: -90 move to
  repeat 11 times
    rotate
    move 40 pixels
  hide
  
```

4. 스프라이트3

```

when green flag clicked
  shape to oval
  x: -200 y: -40 move to
  repeat 11 times
    rotate
    move to next shape
    move 40 pixels
  hide
  
```

<제주 정낭 프로젝트를 순서도로 나타내기-센서 활용>





로 배우는

데

이

터

과

학

수업일 :

이름 :

4

○ 알 수 있는 점 ○







★ 스스로 해 보기 2 ★ 데이터를 시각화시켜 보세요.

○ 알 수 있는 점 ○







2. 개별 속성 분석하기

○ 알 수 있는 점 ○











로 배우는

데

이

터

과

학

수업일 :

이름 :

4

3. 속성간의 관계 분석하기(의미 있는 상관관계 찾기)

○ 알 수 있는 점 ○







현상 일반화하기: 통계적 추론

1. 내가 세운 가설을 대상으로 비교 및 결론 내리기

○ 결론 ○







2. 앞으로의 예측(따라서 앞으로 ~한다면 ~게 될 것으로 생각되므로 ~기 위한 ~에 대한 노력이 필요하겠다. 이번 조사를 바탕으로 이런 부분을 추가해서 더 알아보고 싶다 등)





로 배우는

데

이

터

과

학

수업일 :

이름 :

4

○ 예측 ○



결과 소통: 발표하고 공유하기

- 서로 의견 나누기

데이터 과학 보고서: 오늘 실습한 내용을 바탕으로 보고서 써보기



★ 과제 1

개별 데이터를 이용하여 본인이 해결하고 싶은 주제를 정한 후, 붙임 데이터 과학 보고서를 완성하여 답글로 올리기

★ 과제 2

<선택 심화> 제주 정낭 프로젝트를 스크래치로 구현해보기





로 배우는

데

이

터

과

학

수업일 :

이름 :

5



생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 볼 수 있다.
 무인자동차(미로 탈출 로봇)의 원리가 생활 속에 필요한 이유를 알고 스크래치를 이용하여 원그래프를 그려 볼 수 있다.
 미로 탈출 로봇을 순서도로 나타낼 수 있다.

생각해 보기

- ▶ 평소에 궁금했던 점은 무엇이 있을까? 또는 내 주변에서 고쳤으면 하는 점이나 궁금한 점은 무엇이 있을까? 생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 보자.
- ▶ 자동차가 자동으로 거리를 인식해서 운전하면 어떤 점이 편리할까?
- ▶ 자율주행 자동차 영상 보여주기

<https://www.youtube.com/watch?v=V1IP1fYMsNo&feature=youtu.be>(약 1분 23초)





로 배우는

데

이

터

과

학

수업일 :

이름 :

https://www.youtube.com/watch?v=-j5BIUL_sPc(약 3분 32초)



- ▶ 자율주행 기술이 가져올 사회적 변화에 대해 이야기해 보기
- ▶ 자율주행 기술에 필요한 센서의 역할에 대해 이야기해 보기

☑ 문제 정의

나와 내 주변에서 일어나는 일을 중심으로 주제 선정하기

1. 내 자신이 고쳤으면 하는 점이나 궁금한 점을 생각해 보자.

○ 예시 ○

- 나는 왜 공부하는 시간이 많은데 성적이 안 오를까?
- 나는 왜 많이 먹어도 살이 안 찢까?
- 나는 왜 조금만 먹어도 살이 찢까?
- 나는 왜 배탈이 잘 날까?
- 나는 왜 아침마다 늦잠을 잘까?
- 나는 왜 친구들과 자주 싸울까?





로 배우는

데

이

터

과

학

수업일 :

이름 :

5

2. 나와 관련된 일 중 고쳤으면 하는 점 또는 궁금한 점이 있으면 생각해 보자.

○ 예시 ○

- 악기 연주는 스트레스 해소에 도움을 줄까?
- 바른 자세로 앉으면 건강(공부)에 도움이 될까?
- 나이가 젊은 선생님이 우리들을 더 잘 이해해 줄까?
- 여자일수록 수다가 많을까?
- 엄한 선생님일수록 숙제를 잘 해 올까?
- 형제가 많으면 성격이 좋을까?(다른 사람들을 더 잘 배려하고)
- 남자들은 여자보다 무서움이 없을까?
- 성적이 좋을수록 스트레스가 적을까?

3. 자신이 탐구해 보고 싶은 주제 및 그 주제를 결정한 이유를 써 보자.

(★ 붙임 보고서에 정리)

데이터 수집: 필요한 데이터 스스로 만들기

(★ 붙임 보고서에 정리)

현상 이해하기: 탐색적 데이터 분석

1. 원본 데이터 잘 살펴보고 시각화 시키기

(★ 붙임 보고서에 정리)





로 배우는

데

이

터

과

학

수업일 :

이름 :

5



1. 마법사

2. 스프라이트 1

```

클릭했을 때
크기를 60%로 정하기
지우기
드롭다운 메뉴 항목을 응답결과에서 삭제하기
응답수행계 음(증) 0로 정하기
보기개수 음(증) 0로 정하기
순서 음(증) 1로 정하기
문장의 보기는 몇 개인가요? 묻고 기다리기
보기개수 음(증) 대답로 정하기
보기개수 번 반복하기
순서와 번째 응답수를 입력하세요. 경합하기 묻고 기다리기
대답 음(증) 순서 번째 응답결과에 넣기
순서 음(증) 1만큼 바꾸기
응답수행계 음(증) 대답만큼 바꾸기
원 그래프 그리기 방송하기
  
```

```

원 그래프 그리기 방송할 때
x: 0 y: 0로 이동하기
60도 발췌 보기
순서 음(증) 1로 정하기
보기개수 번 반복하기
색상 효과를 10만큼 바꾸기
순서 번째 응답결과 항목 / 응답수행계 * 100 * 3.14번 반복하기
도정하기
1도 옮기기
순서 음(증) 1만큼 바꾸기
  
```





로 배우는

데

이

터

과

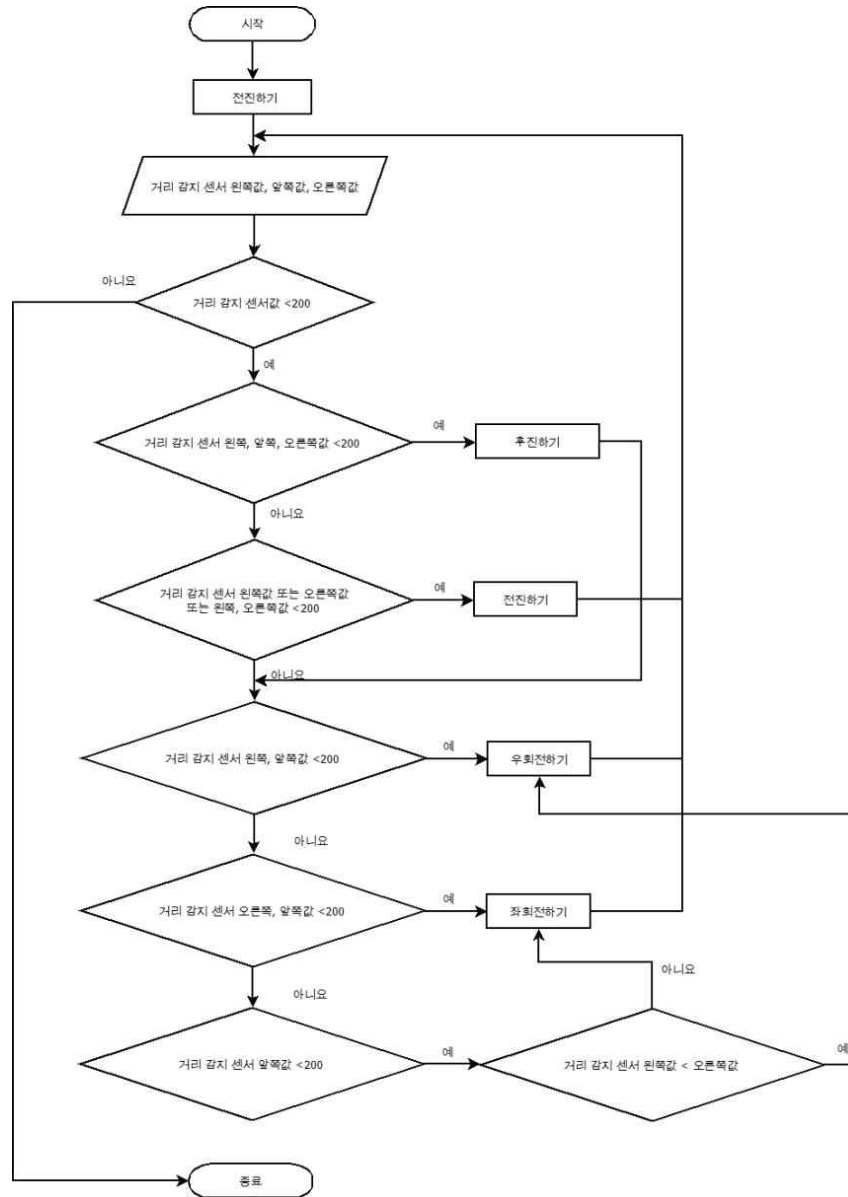
학

수업일 :

이름 :

5

<미로 탈출 로봇 프로젝트를 순서도로 나타내기-센서 활용>





로 배우는



수업일 :	이름 :	5
-------	------	---

2. 개별 속성 분석하기

(★ 붙임 보고서에 정리)

3. 속성간의 관계 분석하기(의미 있는 상관관계 찾기)

(★ 붙임 보고서에 정리)

현상 일반화하기: 통계적 추론

1. 내가 세운 가설을 대상으로 비교 및 결론 내리기

(★ 붙임 보고서에 정리)

2. 앞으로의 예측

(따라서 앞으로 ~한다면 ~게 될 것으로 생각되므로 ~기 위한 ~에 대한 노력이 필요하겠다. 이번 조사를 바탕으로 이런 부분을 추가해서 더 알아보고 싶다 등)

(★ 붙임 보고서에 정리)

결과 소통: 발표하고 공유하기(느낀 점)

(★ 붙임 보고서에 정리)





로 배우는

데

이

터

과

학

수업일 :

이름 :

5

☑ 데이터 과학 보고서: 오늘 실습한 내용을 바탕으로 보고서 써보기



★ 과제 1

개별 데이터를 이용하여 본인이 해결하고 싶은 주제를 정한 후, 붙임 데이터 과학 보고서를 완성하여 답글로 올리기

★ 과제 2

<선택 심화> 미로 탈출 로봇 프로젝트를 스크래치로 구현해보기





로 배우는

데

이

터

과

학

수업일 :

이름 :

6



생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 볼 수 있다.
 무인 경비 로봇(강아지 로봇)의 원리가 생활 속에 필요한 이유를 알고 스크래치를 이용하여 원그래프를 그려 볼 수 있다.
 무인 경비 로봇(강아지 로봇)을 순서도로 나타낼 수 있다.

생각해 보기

- ▶ 평소에 궁금했던 점은 무엇이 있을까? 또는 내 주변에서 고쳤으면 하는 점이나 궁금한 점은 무엇이 있을까? 생활 속에서 주제를 찾아 스스로 데이터를 수집하여 해결해 보자.
- ▶ 국산 무인정찰/전투 견마로봇 '견마'[Korea Unmanned Combat Vehicle]

<https://www.youtube.com/watch?v=FWBnCFOWGtE>(약 1분 53초)





로 배우는

데

이

터

과

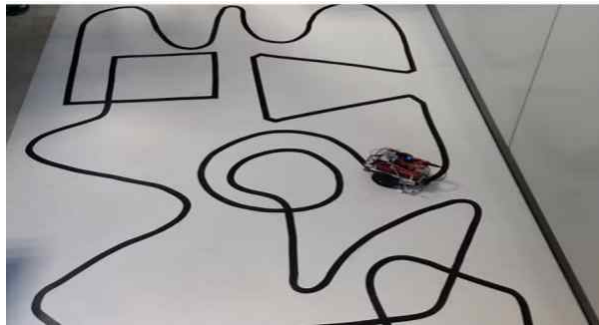
학

수업일 :

이름 :

6

- 라인트레이서 <https://www.youtube.com/watch?v=qVsINWssRU4>
(약 2분 33초)



문제 정의

1. 문제 정의하기

* 문제(문제점이나 궁금한 점이 무엇인가?)

평소에 내 자신이 고쳤으면 하는 점이나 궁금한 점은 무엇이 있을까?

또는 내 주변에서 고쳤으면 하는 점이나 궁금한 점은 무엇이 있을까?











로 배우는

데

이

터

과

학

수업일 :

이름 :

6

☑ 데이터 수집

1. 데이터 수집의 기본

○ 관련된 항목 생각해 보기 ○







2. 수집방법 결정하기(자동수집)







3. 데이터 품질 점검하기

- 완전성(문제 해결에 필요한 내용이 들어 있느냐?)
- 정확성(문제가 되는 상황이 들어 있느냐?)
- 일관성(데이터 항목들이 연관이 있고 일치하느냐?)

4. 데이터 준비하기

5. 데이터 추가, 선택, 집계하기

- 데이터 추가 및 선택은 필요시에 함.





로 배우는

데

이

터

과

학

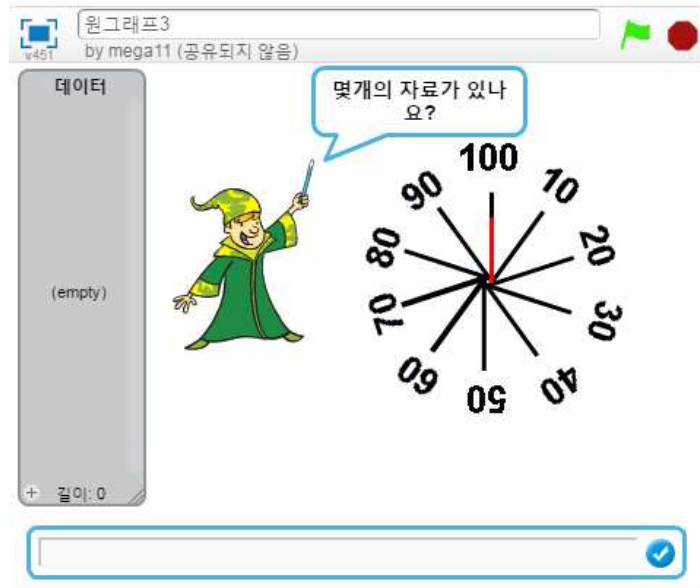
수업일 :

이름 :

☑ 현상 이해하기: 탐색적 데이터 분석

1. 원본 데이터 잘 살펴보고 시각화 시키기

(★ 붙임 보고서에 정리)



(데이터를 전체적으로 살펴보고 시각화를 통해 나타나는 패턴을 한눈에 살펴봄)

○ 알 수 있는 점 ○

Three horizontal lines for writing, each preceded by a colored pencil icon (blue, red, green).





로 배우는

데

이

터

과

학

수업일 :

이름 :

6

1. 마법사

```

클릭했을 때
크기를 70% 로 정하기
오른쪽 번째 항목은 데이터 에서 삭제하기
데이터 개수 음(증) 0 로 정하기
합계 음(증) 0 로 정하기
항목 음(증) 0 로 정하기
몇개의 자료가 남아있나요? 묻고 기다리기
데이터 개수 음(증) 대답 로 정하기
데이터 개수 번 반복하기
항목 음(증) 1 만큼 바꾸기
항목 와 번째 자료를 일치하리 결합하기 묻고 기다리기
대답 음(증) 항목 번째 데이터 에 넣기
합계 음(증) 대답 만큼 바꾸기
원그래프 그리기 ► 발송하기
  
```

2. 스프라이트 1

```

클릭했을 때
원 항목 음(증) 0 로 정하기
90도 방향 보기
x: 60 y: 60 로 이동하기

원그래프 그리기 ► 음(증) 받았을 때
원 항목 음(증) 1 만큼 바꾸기
원 항목 > 데이터 개수 까지 반복하기
색깔 ► 요소를 40 만큼 바꾸기
원 항목 번째 데이터 ► 항목 / 합계 * 100 * 0.5 번 반복하기
3도 돌기
도장찍기
원 항목 음(증) 1 만큼 바꾸기
  
```

3. 스프라이트 2

```

클릭했을 때
지우기
선 그리기 ► 발송하기

선 그리기 ► 음(증) 받았을 때
90도 방향 보기
x: 90 y: 30 로 이동하기
10 번 반복하기
36도 돌기
도장찍기
숨기기
  
```

4. 스프라이트 3

```

선 그리기 ► 음(증) 받았을 때
90도 방향 보기
모양을 모양2 (으)로 바꾸기
x: 90 y: 30 로 이동하기
10 번 반복하기
36도 돌기
도장찍기
다음 모양으로 바꾸기
숨기기
  
```





로 배우는

데

이

터

과

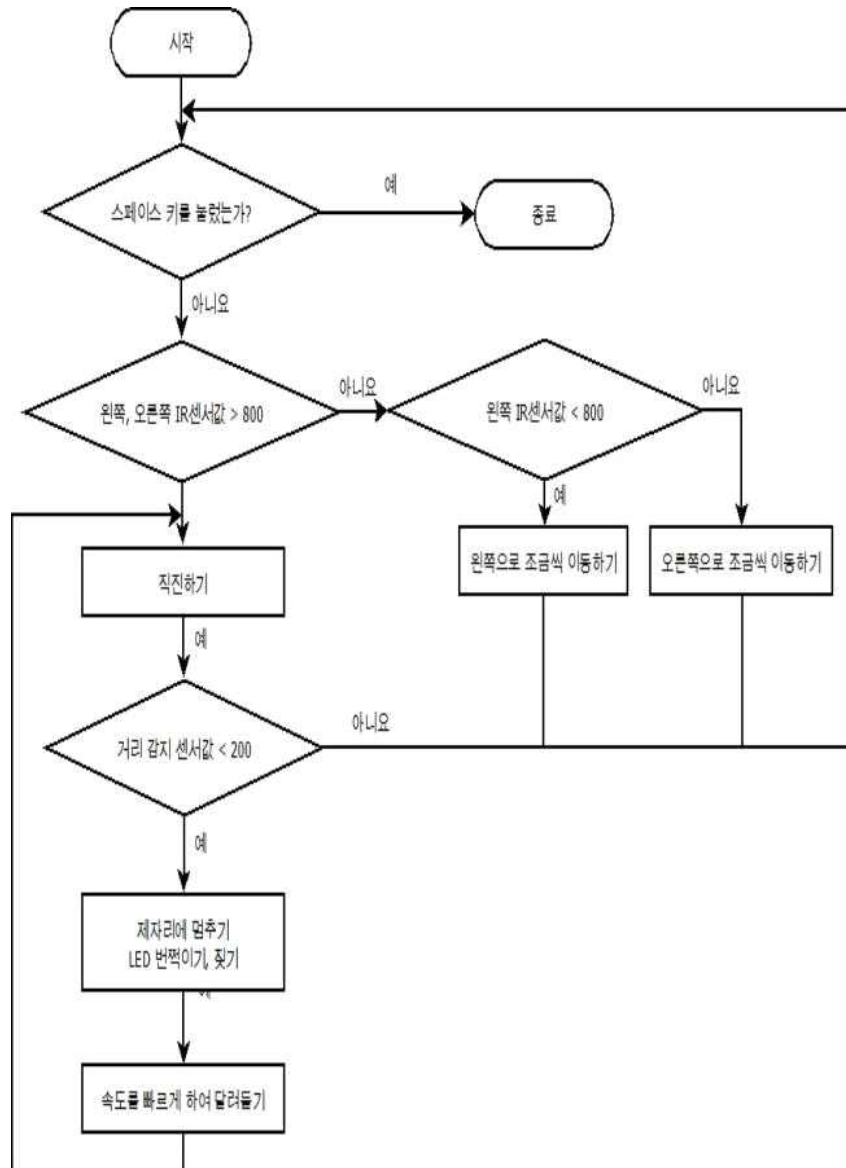
학

수업일 :

이름 :

5

<무인 경비 로봇 프로젝트를 순서도로 나타내기-센서 활용>





로 배우는

데

이

터

과

학

수업일 :

이름 :

6

2. 개별 속성 분석하기

○ 알 수 있는 점 ○







3. 속성간의 관계 분석하기(의미 있는 상관관계 찾기)

○ 알 수 있는 점 ○







☑ 현상 일반화하기: 통계적 추론

1. 내가 세운 가설을 대상으로 비교 및 결론 내리기

○ 결론 ○











로 배우는

데

이

터

과

학

수업일 :

이름 :

5

2. 앞으로의 예측

(따라서 앞으로 ~한다면 ~게 될 것으로 생각되므로 ~기 위한 ~에 대한 노력이 필요하겠다. 이번 조사를 바탕으로 이런 부분을 추가해서 더 알아보고 싶다 등)

○ 결론 ○







☑ 결과 소통: 발표하고 공유하기(느낀 점)

- 서로 의견 나누기

☑ 데이터 과학 보고서: 오늘 실습한 내용을 바탕으로 보고서 써보기



★ 과제 1

개별 데이터를 이용하여 본인이 해결하고 싶은 주제 1가지를 정한 후, 그와 관련된 구글 설문지를 만들고, 데이터를 수집한 후, 스크래치를 이용하여 탐색적 데이터 분석을 한 결과를 최종보고서 양식에 기록하여 답글로 올리기

★ 과제 2

<선택 심화> 무인 경비로봇 프로젝트를 스크래치로 구현해보기

