

‘데이터 분석 역량 키우기’ 수업 제안서

2021 상반기 프로그램

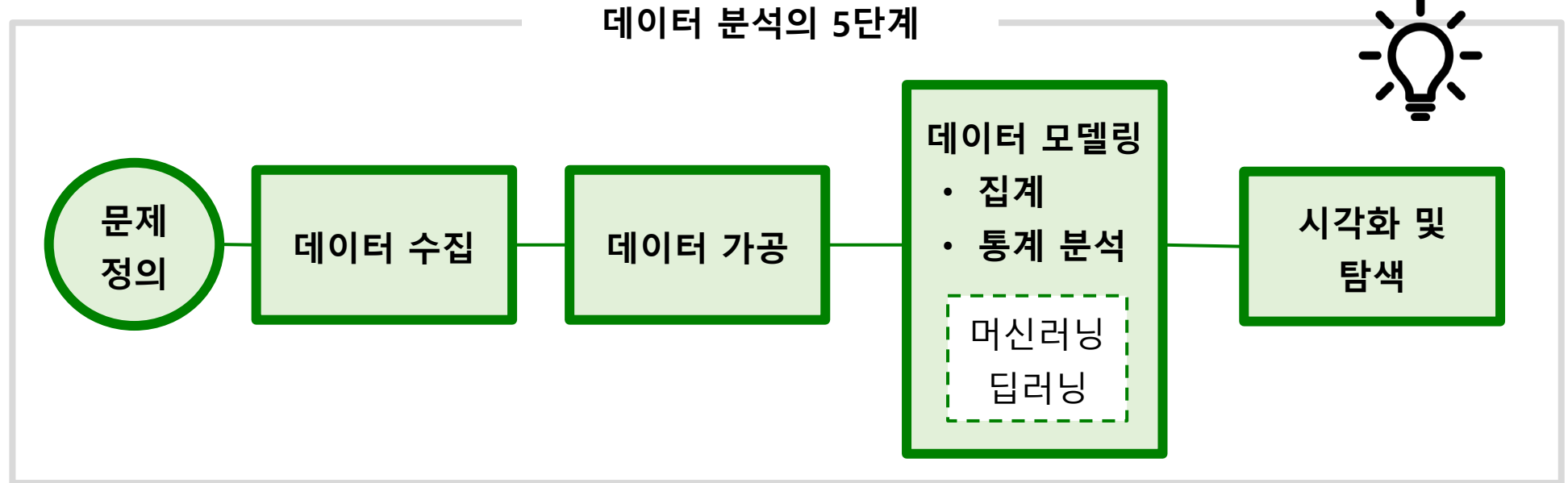


요약

- **데이터 분석**은 데이터 수집, 가공 및 데이터 모델링 등의 과정을 통해 **데이터로부터 통찰력을 얻는 과정**입니다.
- 데이터 분석 과정을 통해 **다양한 전공을 고민하는 학생들의 진로 선택에 도움**을 주고자 합니다.
- 수업은 크게 **데이터 분석 기초 및 응용** 그리고 **머신러닝** 과정으로 구성됩니다.
- 각 학교의 일정과 상황에 맞게 수업을 조합하여 선택할 수 있습니다.
- 기초 수업은 학생들이 데이터를 수집하는 방법부터 시작하여 가공, 집계를 거쳐 그래프를 직접 작성할 수 있게 지도합니다.
- 수업은 **온라인으로 진행**되며 학생 교육 관리 시스템인 프리스왓을 통해 강의, 출석 체크, 채팅 및 퀴즈를 수행합니다.
- 세이지는 **고등학생에게 최적화 된 맞춤형 데이터 분석 교육**을 제공합니다.

데이터 분석이란

데이터 분석은 데이터 수집, 가공 및 데이터 모델링 등의 과정을 통해 데이터로부터 통찰력을 얻는 과정입니다.



- 기존에는 주로 통계 패키지 등을 활용하여 데이터를 분석
- 최근에는 파이썬 언어 및 컴퓨팅 파워의 발전으로 머신러닝 등을 누구나 분석이 가능해 짐

파이썬은 타 언어대비 배우기가 비교적 쉽기 때문에 학생들도 데이터 분석에 접근이 용이함

수업 대상

데이터 분석 수업은 다양한 전공을 고민하는 학생들의 진로 선택에 도움을 주고자 합니다.

관련 학과 예시

다양한 진로에
관심있는
학생들이
들으면
좋습니다.



경영학과

- 데이터 기반 전략 및 마케팅

통계학과

- 통계 분석

사회학과

- 사회 통계
- 사회조사방법

생물학과

- 생물정보학
- 유전자 분석

컴퓨터과학과

- 인공지능
- 프로그래밍

산업공학과

- 품질 공학
- 빅데이터

수학

- 선형대수
- AI 모형 개발

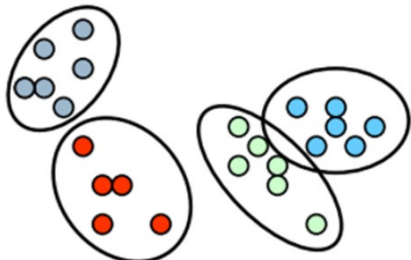

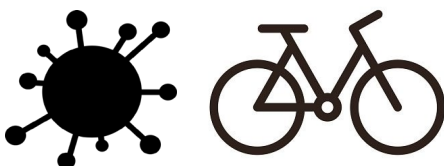
의학

- 의료 빅데이터
- 질병 통계 분석

이과 지망 학생 뿐 아니라 문과를 지망하는 학생들에게도 유용한 수업임

수업 소개

수업은 크게 데이터 분석 기초 및 응용 그리고 머신러닝 과정으로 구성됩니다.

수업 수준				학습 내용
심화	머신러닝	<ul style="list-style-type: none"> • 데이터에서 규칙 추출 • 새로운 데이터를 예측 		<ul style="list-style-type: none"> • 파이썬 및 수학 • 머신러닝 맛보기
응용	데이터 분석 응용 (따릉이)	<ul style="list-style-type: none"> • 데이터에서 규칙 추출 • 새로운 데이터를 예측 		<ul style="list-style-type: none"> • 빅데이터 맛보기 • 다양한 형태의 데이터
기초	데이터 분석 기초 (코로나)	<ul style="list-style-type: none"> • 데이터 분석 		<ul style="list-style-type: none"> • 데이터 수집, 가공, 집계, 시각화 • 코로나, 따릉이 데이터 활용

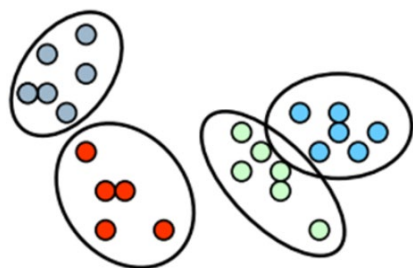
기초부터 심화까지 학생들 수준에 맞는 데이터 분석 수업을 선택할 수 있음

수업 구성

각 학교의 일정과 상황에 맞게 수업을 조합하여 선택할 수 있습니다.

수업 구성

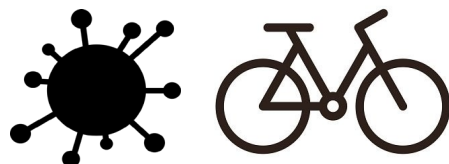
머신러닝
(4회)



데이터 분석
응용 (2회)



데이터 분석
기초 (4회)



수업 안 (예시)

- 데이터 분석 기초 (총 4회)
- 데이터 분석 기초 1 +
데이터 분석 기초 2 (총 8회)
- 데이터 분석 기초 + 응용 +
머신러닝 (총 12회)

동아리 (예시)

- 한 학기 동안 데이터 분석
기초 + 응용 + 머신러닝 총
12회 수업
- 학생들의 조별 과제를 추가
실습으로 지원

문과 지망 학생들은 사회 현상을 분석하는 데이터 분석
기초 수업을, 이과 지망 학생들은 데이터 분석 외에도 벡터
개념을 다루는 머신러닝 수업을 추천함

수업 내용 : 데이터 분석 기초 (1) - 코로나

기초 수업은 학생들이 데이터를 수집하는 방법부터 시작하여 가공, 집계를 거쳐 그래프를 직접 작성할 수 있게 지도합니다.

데이터 수집



어떤
데이터를
어디에서
구하지?

- 코로나 관련 데이터를
국가통계포털 및 뉴욕 타임즈
깃허브에서 수집

데이터 가공 및 집계



그런데
데이터가
조금 이상한
것 같아

- 분석하기 편리하도록
데이터를 가공 및 집계
- 파이썬의 판다스
라이브러리를 활용

시각화



그래프로 보니
복잡한 데이터도
한 눈에 이해할
수가 있네!

- 분석 목적에 맞는 다양한
그래프 유형 학습 및 작성
- 막대, 꺾은선, 원, 상자 그래프
등

수업 내용 : 데이터 분석 기초 (1) – 데이터 수집

실생활과 관련이 있으면서 학생들이 구하기 쉬운 공공 데이터를 사용합니다.

수집 데이터

코로나 발생
현황



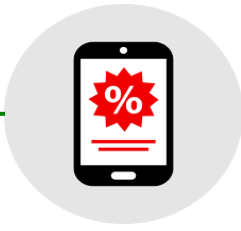
- 코로나 신규
확진자 추이
- 한국 / 미국



코로나가
사회에
미치는 영향



- 코로나 전후
감기 발생 추이



- 코로나 전후
온라인 매출
채널 / 상품별



수집 장소 (공공 데이터 활용)



- 국가통계포털
- 코로나19 데이터

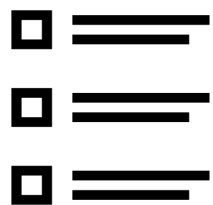
- 건강보험 심평원
- 감기 발생 데이터

- 국가통계포털
- 온라인 매출
데이터

수업 내용 : 데이터 분석 기초 (1) – 데이터 가공 및 집계

데이터 분석에 특화된 파이썬의 판다스 라이브러리를 활용하여 데이터를 가공 및 집계합니다.

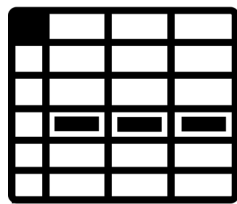
데이터 가공



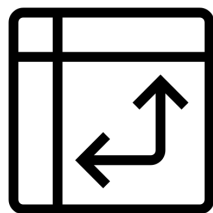
• 인덱스 세팅



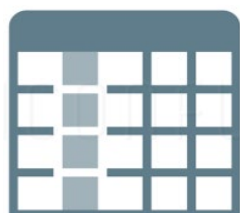
• 자료형 변경



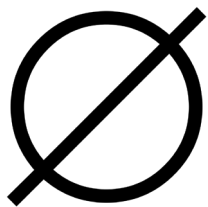
• 행 데이터 추출



• 행과 열 바꾸기

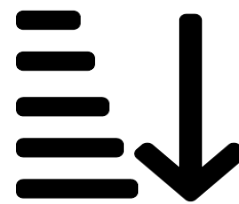


• 열 데이터 추출



• 누락값 처리

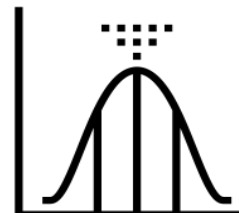
데이터 집계 (및 통계량 산출)



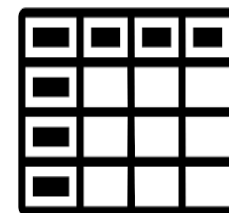
• 정렬



• 그룹 집계



• 통계량 산출



• 피벗

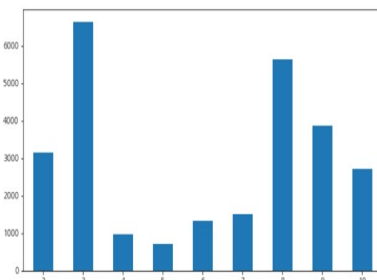
수업 내용 : 데이터 분석 기초 (1) - 시각화

그래프를 작성하며 다양한 시각으로 데이터를 해석하는 연습을 합니다.

코로나 발생 현황



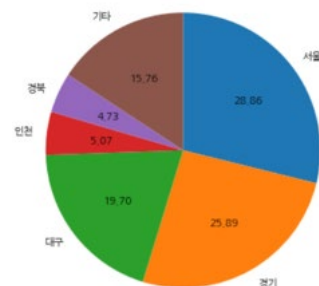
코로나가
많이 발생한
월은?



• 3월



코로나가
많이 발생한
지역은?

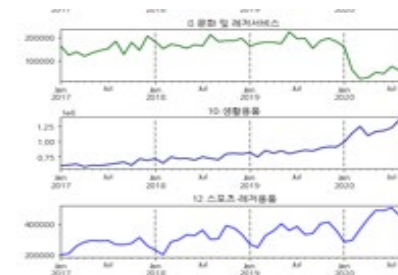


• 서울 및 경기

코로나가 사회에 미치는 영향



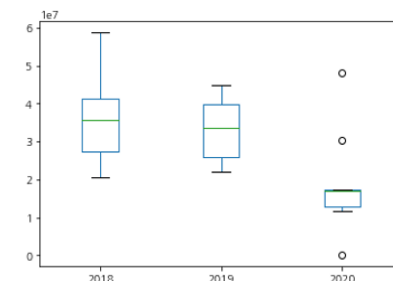
코로나 이후
매출이 줄은
산업은?



• 문화 서비스



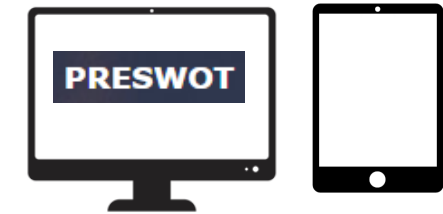
예년 대비
감기가 적게
발생했습니까?



• 그렇습니다

수업 방식 - 강의

수업은 온라인으로 진행되며 학생 교육 관리 시스템인 프리스왓을 통해 강의, 출석 체크, 채팅 및 퀴즈를 수행합니다.



- 가정에서 컴퓨터 및 태블릿 등으로 접속

온라인으로 강의 진행

채팅을 통한 실시간 수업 참여

The central image shows a screenshot of the PRESWOT system interface. It features a list of topics on the left, a grid of line charts in the center, and a chat window on the right. The chat window shows a list of participants and their messages, such as '김 음식', '노 e쿠폰서비스', '이 자동차...?', '이 생활용품', '오 음식', and '이 컴퓨터 및 주변기기'. There are buttons for '입력' (Input), '강습자님과 1:1 통화 요청' (Request 1:1 call with instructor), '질문 인식 테스트' (Question recognition test), and '세션자료' (Session materials).

출석 인증 진행 중
남은 시간 1:33

출석 인증번호 숫자 4자리를 입력해주세요.

1234

확인

- 출석 체크

문항 - 단답

영S의 노래로 2020년 빌보드 HOT 100 처음으로 1위한 곡은?

Dynamite

다음

이전

- 퀴즈 풀이

실시간 채팅 및 퀴즈를 통해 학생들의 수업 참여도를 높여 온라인 수업의 한계를 보완함

수업 방식 - 퀴즈

퀴즈 결과에 따라 다음 수업 보완 및 학생 개별 피드백이 이루어집니다.

퀴즈 문제 예시

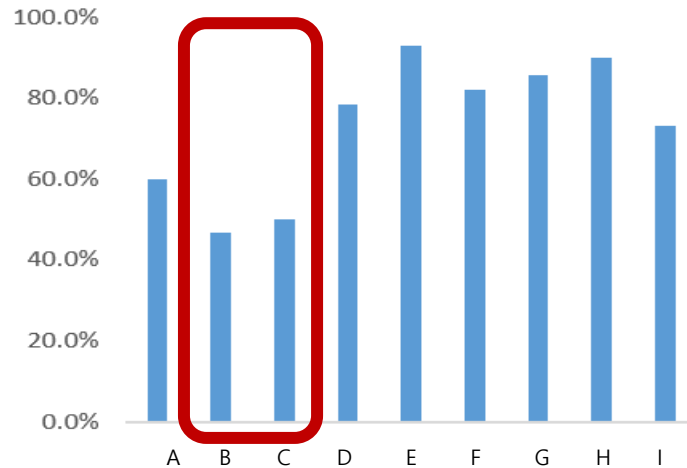


데이터프레임의
행과 열을 바꾸는
명령어를 고르시오.

행 단위 데이터를
추출하는 명령어를
고르시오

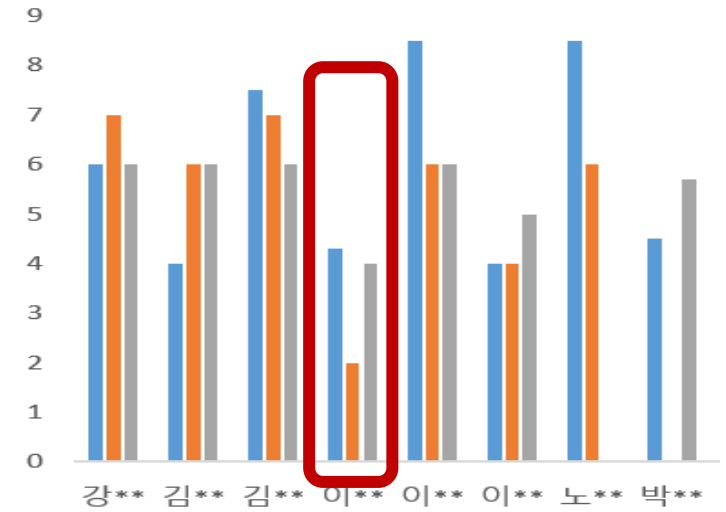
- 수업 1회 당 약 8 ~ 10 문항
정도의 퀴즈 풀이

퀴즈별 정답률



- 정답률이 낮은 문제는 다음
시간에 추가로 설명

학생별 퀴즈 점수



- 수업 이해도가 낮은 학생들에
대한 개별 피드백

수업 방식 - 실습

실습은 구글의 파이썬 실습 환경인 코랩Colaboratory에서 수행됩니다.

코랩이란

- Colaboratory 또는 Colab 코랩은 구글에서 제공하는 클라우드 기반의 파이썬 실습 환경



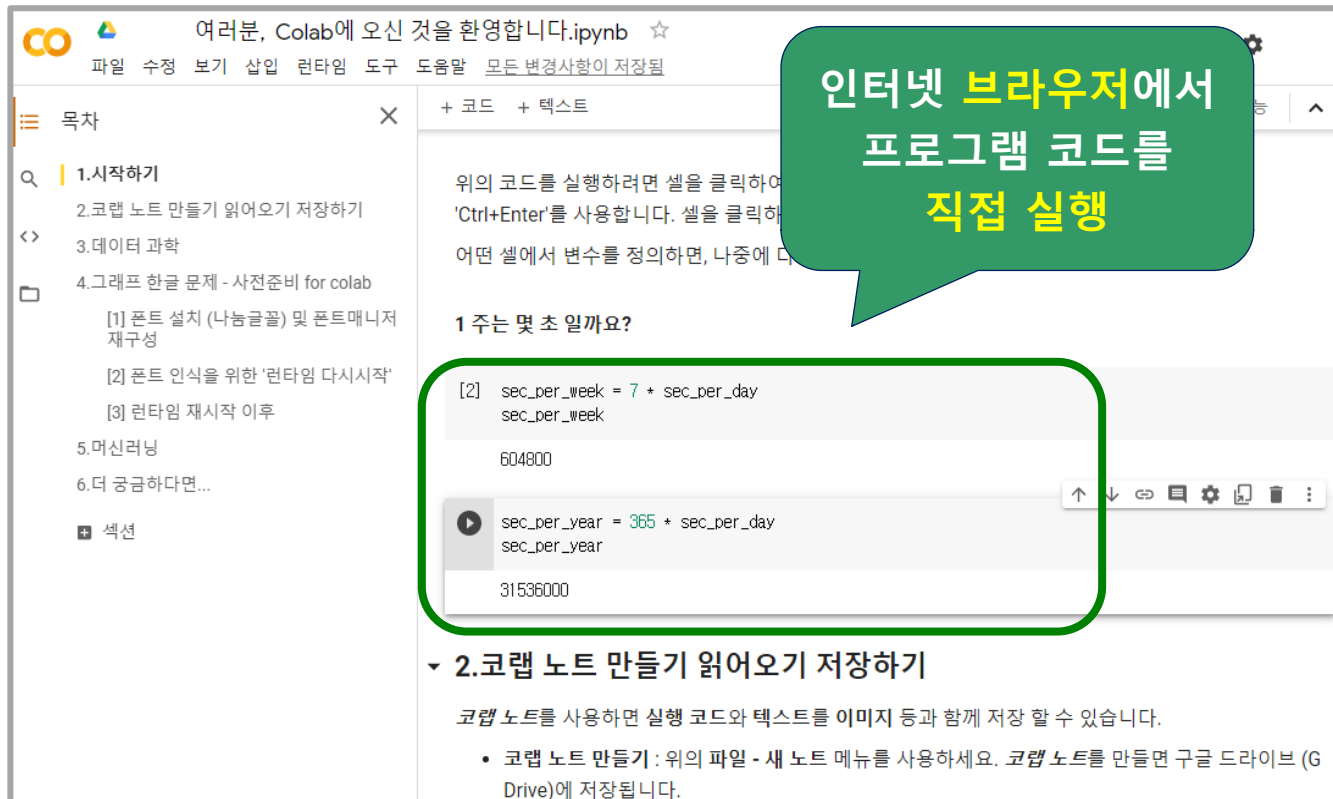
- 크롬 브라우저에서 직접 코드 실행



- 컴퓨터, 태블릿 및 휴대폰 등으로 접속



- 별도 프로그램 설치 필요 없음 (서버에서 실행)



<http://colab.research.google.com>

수업 방식 - 질의 응답

강사 외 전문 조교가 배정되어 학생들의 문의에 실시간으로 응답합니다.

채팅을 통한 질의 응답

수업 중 채팅을 통한 문의



아 날짜 정하는것부터 못따라갔어요ㅠㅠ

노 저도요..


김 저도요..

김 날짜 정하는 게 어디있는지 잘 모르겠어요


김정은 화면 부분 다시 실습으로 할게요.

- 학생들의 피드백을 수업에 실시간으로 반영


개별 카톡을 통한 질의 응답



sorting 할 때
ascending 대신
descending 써도
되나요?



파라미터는
ascending=False
로 사용합니다.



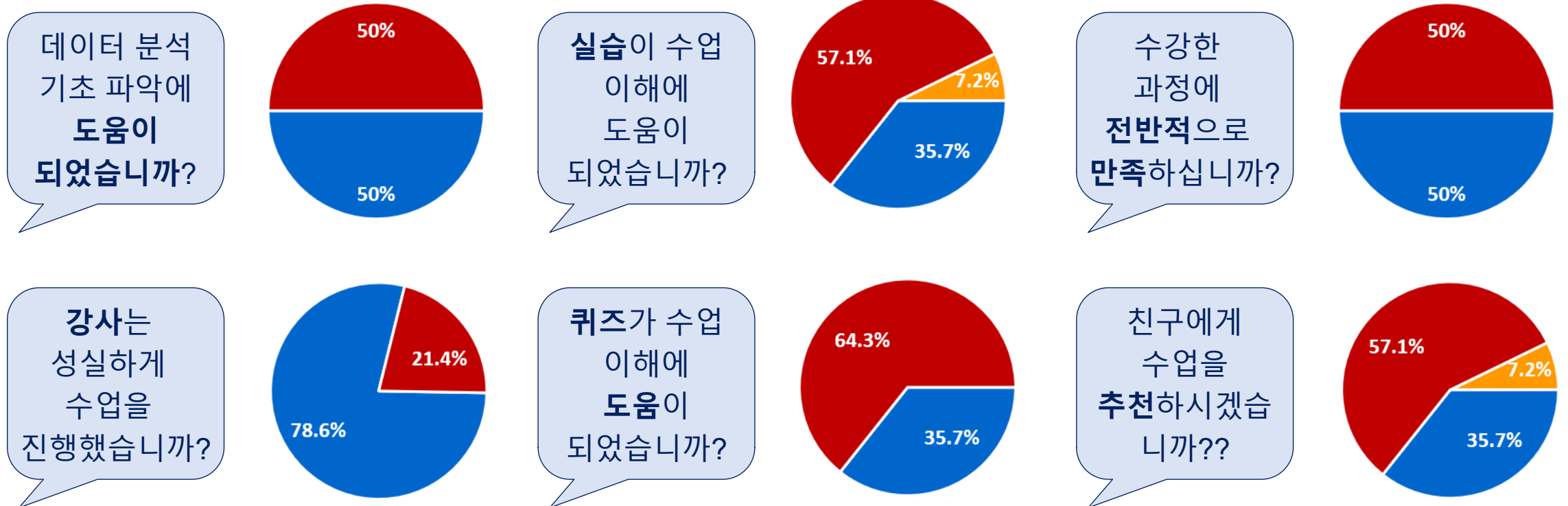
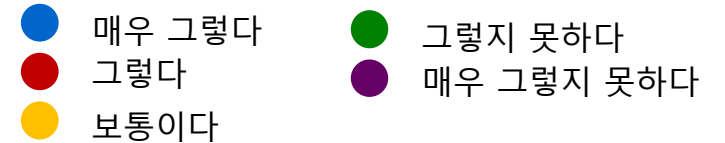
- 학생들의 개별 질문에 상세한 답변

전문 조교가 학생들의 질의 응답에 실시간으로 대응하여 온라인 수업의 한계를 보완

학생 설문 결과

학생들은 수강한 데이터 분석 수업에 전반적으로 만족하는 것으로 조사되었습니다.

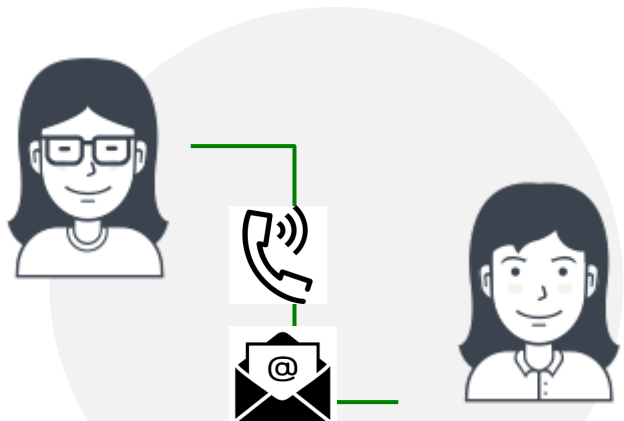
타 고교 설문 조사 현황



출처: A 고교 수강 학생 만족도 설문 조사, 2020.11 월 시행

수업 진행 절차

학생 및 학교 상황에 맞는 수업을 진행할 수 있도록 세이지와 함께 협의합니다.



자세한 내용은 홈페이지를 방문해 주세요.
궁금하신 사항은 편하게 문의 주세요.

- <http://www.sagelab.co.kr>
- 010-4677-8279
- sage@sagelab.co.kr

수업 과정 설계 (예시)



- 문과 혹은
이과 지망
학생



- 학생 수준에
맞는 수업
선정



- 중간고사 등
학교별 상황에
맞는 일정

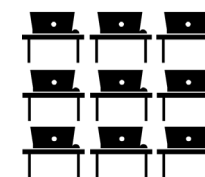
학생 선발 및 환경 조사 (예시)



- 사전 설명회
개최



- 희망 학생 지원
- 학생 선발



- 컴퓨터 환경
조사
- 구글 메일 개설

수업 사후 관리

학생 및 학교에 필요한 다양한 피드백을 수업 후에도 협의하여 제공합니다.



선생님

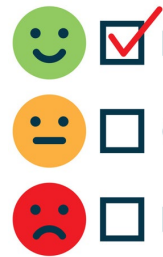
우리
학생들에게
너무 어려웠던
건 아닐까?



학생

수업을 듣기는
했는데 진로에
어떻게
관련되는 걸까?

만족도 조사



☒ Excellent

☐ Good

☐ Poor

- 수업 만족도 조사
- 수업 내용, 난이도, 강사 등 평가

수업 결과 보고서



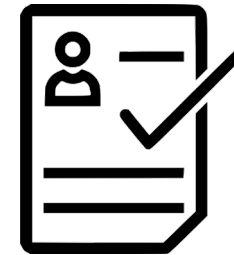
- 출석 및 퀴즈 결과 등 포함한 결과 보고서

학생 피드백



- 수업 내용 리뷰
- 향후 학습 방향 추천

생활 기록부 기술 사항 (협의)



- 학습 내용 리뷰 및 진로 설계

세이지 소개

세이지는 고등학생을 위한 맞춤형 데이터 분석 교육을 제공합니다.

“ 세이지(世利智)는 학생들이
데이터를 직접 다뤄보고
이해하며 이를 진로 설계에
고려할 수 있도록 교육
컨텐츠와 전략을 제공합니다. ”

세이지 소개 영상

<https://youtu.be/6zTlwgkekjk>

세이지 홈페이지

<https://www.sagelab.co.kr>



- 어려울 수도 있는 데이터 분석이란 주제를 학생들의 눈높이에 맞게 제공



- 연세대학교에서 검증된 컴퓨터 교육 관리 시스템을 활용한 효과적 학습 관리



- 프로그래밍 뿐 아니라 데이터 분석 관련 전문 인력 확보

세이지 팀 소개

세이지는 소명의식과 역량을 보유한 인력으로 구성되어 있습니다.

김정은

대표

연세대학교 디지털애널리틱스 융합과정 강사
한국 IBM, 삼성 SDS
KAIST 테크노 MBA
연세대학교 컴퓨터과학

박민재

마케팅 본부장

한국 IBM
PWC 컨설팅
KAIST 경영공학 석사
연세대학교 컴퓨터과학

강형주

컨텐츠 본부장

삼성전자
연세대학교 컴퓨터과학

장윤경

전략 본부장

정보통신산업진흥원 AI 글로벌 창업진흥 프로그램 선정
한국 IBM
액센츄어 코리아
오라클 코리아
연세대학교 컴퓨터과학

박봉주

인력개발 본부장

연세대학교 디지털애널리틱스 융합과정 강사
한국어 교원
삼성 SDS
외국인을 위한 한국어 교원과정 (서울대)
연세대학교 컴퓨터과학

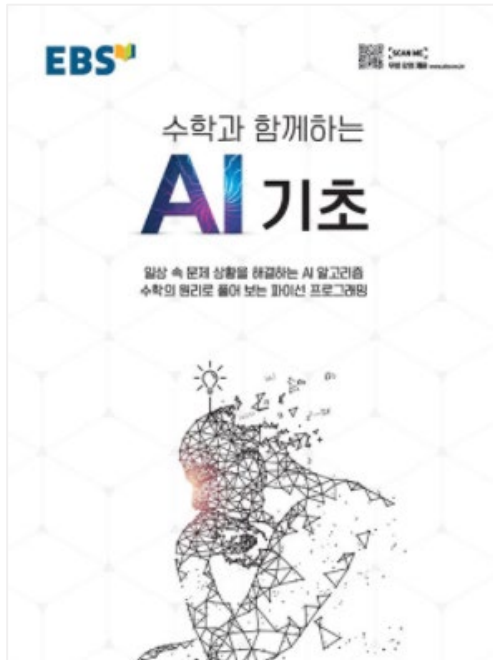
조현진

사업개발 본부장

모니터-딜로이트 컨설팅
한국 IBM
University of Michigan, MBA
연세대학교 컴퓨터과학

별첨. EBS 교재 소개

데이터 분석에 관심있는 학생들에게 EBS에서 출간된 '수학과 함께 하는 AI 기초' 책을 추천합니다.

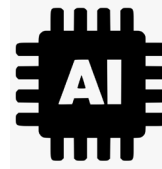


수학과 함께하는

AI 기초

일상에서 접하는 다양한 문제를 해결하는 데 인공지능 기술을 적용하기 위한 프로그래밍 방법 및 수학의 연결고리를 학습할 수 있도록 구성

주요 내용



인공지능 소개 (인공지능의 정의, 활용, 수학과 관련성 등)



데이터의 표현과 가공 (생활, 이미지, 소리 데이터 처리)



다양한 머신러닝 소개 (군집화, 회귀, 로지스틱 회귀 등)