

데이터과학입문

Spring 2024

교과목 개요



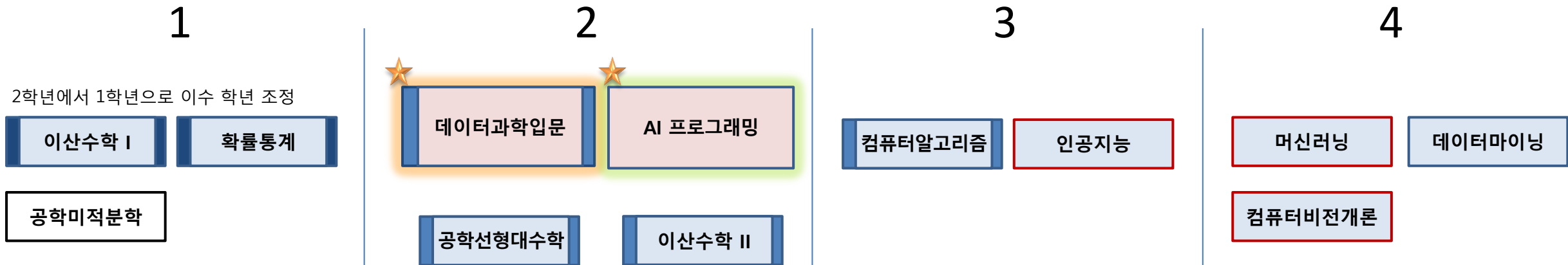
부산대학교 정보·의생명공학대학
정보컴퓨터공학부



교과목 개요

❖ CB35533: 데이터 과학 입문 (Introduction to Data Science)

- 2020학년 교육과정부터 추가된 신규 전공 필수/3학점 교과목



- 인공지능/데이터과학 분야의 기초 교육 강화
- “컴퓨터및프로그래밍입문”과 “확률통계” 교과목 이수 필요
- 2020학번부터 전공 필수
 - 2020 학번 이전 학생은 전공 선택으로 인정
- 이론교과목, but 과제/실습을 통한 학습

주교재

- ❖ Computational and Inferential Thinking: The Foundations of Data Science
 - 2nd Edition by [Ani Adhikari](#), [John DeNero](#), [David Wagner](#).
 - originally developed for the UC Berkeley course.
 - <https://inferentialthinking.com/chapters/intro.html>



참고도서

❖ 여러 교재를 복합적으로 참고하여 강의 내용을 준비

- 아래 교재들이 중심 교재 (한글번역판도 존재)
- [부산대학교 도서관 전자책](#) – O'Reilly for Higher Education

1. [Doing Data Science](#) – R. Schutt & C. O'Neil, O'Reilly, ISBN 9781449358655

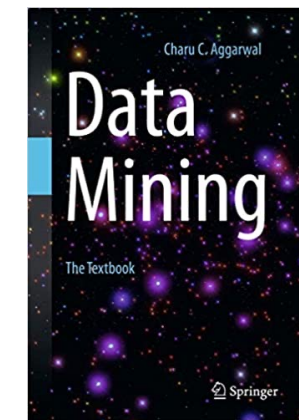
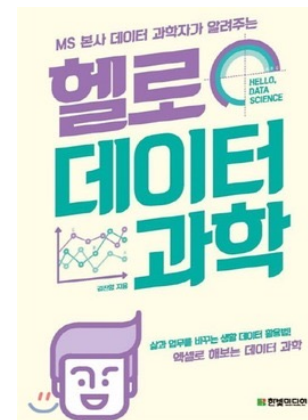
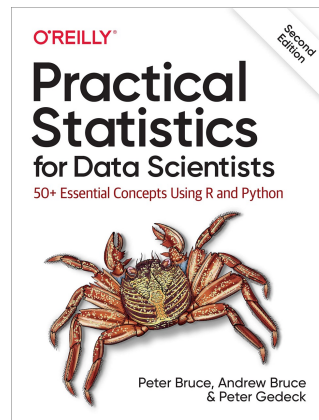
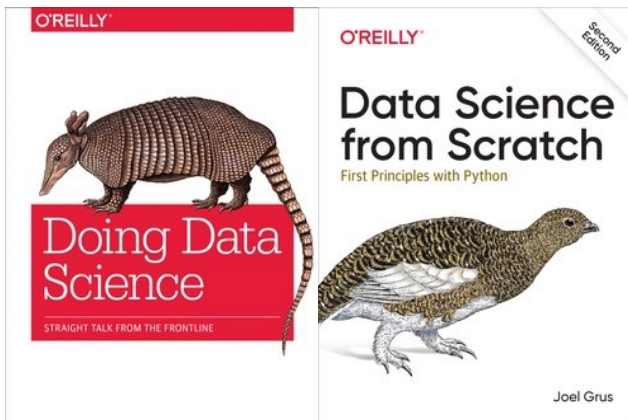
- 데이터 과학 입문, 한빛미디어

2. [Data Science from Scratch](#) 2nd Edition, J. Grus, O'Reilly, ISBN 9781492041139

- 밑바닥부터 시작하는 데이터과학 2판, 인사이트

3. [Practical Statistics for Data Scientists](#), 2nd Edition, P. Bruce, A. Bruce, P. Gedeck, O'Reilly, ISBN 9781492072942

- 데이터 과학을 위한 통계 2판, 한빛미디어



O'Reilly for Higher Education via PNU library website

The screenshot shows the PNU library website (lib.pusan.ac.kr) with the search results for 'doing data science'. The page is titled '전자책' (Electronic Book). The search bar shows the query 'doing data science' and the results are displayed as a list of books. A large blue arrow points to the search results, with the text '전자책 검색' (Electronic Book Search) overlaid. The search results show the book 'doing data science' by John D. Cook, published by O'Reilly. The results also show the number of search results (3 items) and the number of items in the list (1~20).

부산대학교 도서관
PUSAN NATIONAL UNIVERSITY LIBRARY

doing data science

교외접속 ON | MY ?

통합검색 소장자료 학술논문 전자저널 학술DB 전자책

자료검색 컬렉션 연구·학습지원 자료이용안내 시설이용안내 도서관소개

통합검색 소장자료 학술논문 전자저널 학술DB 전자책 온라인강좌

전자책

서명 doing data science

가 나 다 라 마 바 사 아 자 차 카 타 파 하
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

전자책 DB(제공사) 전자책

Yes24/교보문고/북큐브

검색결과 3건 | 목록 1~20

강의 사이트 및 도구

❖ [PLATO](#)

- 부산대학교 스마트 교육 플랫폼
- 강의 자료, 강의 동영상, 강의 공지, Q&A (쪽지), 출석, 퀴즈 등



❖ Google Colab, Docs, Sheet

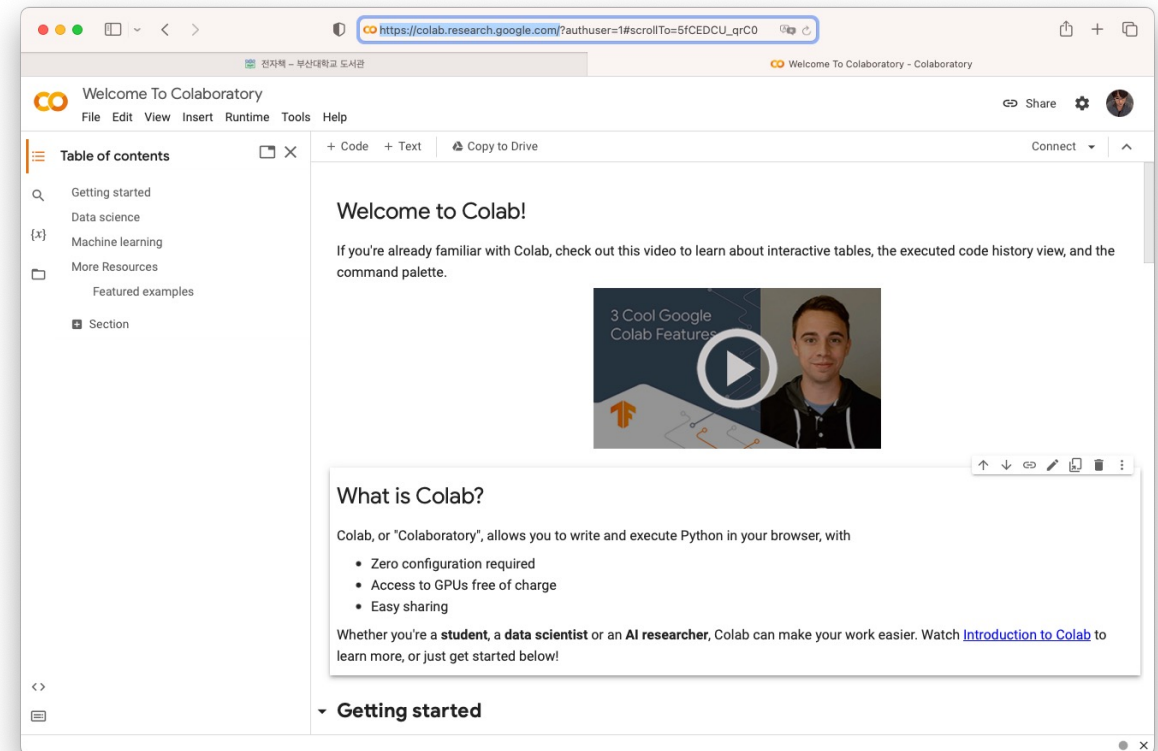
- Google Cloud 기반 도구를 활용한 과제 수행 및 평가
- [Colab](#) / Google Sheet & Doc
- 슬라이드 후반 및 PLATO 를 통해 공지

❖ @pusan.ac.kr 계정 사용



❖ Python

❖ Markdown



Plato 화면 예시 1

The screenshot shows the Plato LMS interface for a course titled "데이터과학입문 (CB3500533-060)". The browser address bar shows the URL "plato.pusan.ac.kr/course/view.php?id=118872". The page header includes the Pusan National University logo and navigation links for "교과과정", "자율강좌", "지정강좌", and "교수·학습". The course title is prominently displayed in the center. Below the title, there are icons for "공유" (Share) and "편집" (Edit). The left sidebar contains a "강의실 홈" (Classroom Home) section with links to "강의정보" (Class Information), "성적/출석관리" (Grade/Attendance Management), and "기타 관리" (Other Management). The main content area is divided into two sections: "과목공지" (Course Notice) and "강의 개요" (Course Overview). The "과목공지" section lists several notices with their dates. The "강의 개요" section features three circular icons representing "과목공지" (Course Notice), "질의응답" (Q&A), and "자료실" (Resource Room). A blue arrow points from the text "공지, Q&A, 기타자료" (Notice, Q&A, Other Materials) to the "과목공지" icon. At the bottom, there is a section for "주차 별 학습 활동" (Learning Activity by Week) with a dropdown menu set to "전체" (All) and a link for "1주차 [3월02일 - 3월08일]" (Week 1 [March 02 - March 08]).

강좌: 데이터과학입문

plato.pusan.ac.kr/course/view.php?id=118872

부산대학교 | 교수학습지원센터 | 학생지원시스템 | 업무포털 | 도서관

최종 접속시간 : 2023-03-01 22:11:05 / 최종 접속IP : 110.13.149.211 | 이명호 | 로그아웃

PLATO 부산대학교 스마트 교육플랫폼

교과과정 자율강좌 지정강좌 교수·학습

데이터과학입문 (CB3500533-060)

이명호 | 쪽지 | 조교

공유 편집

강의실 홈

강의정보

- 교수계획표 (국문)
- 교수계획표 (영문)
- 참여자목록

성적/출석관리

- 학습이력현황
- 동영상이수현황
- 스마트출석부
- 화상강의목록
- 성적부

이메일 보내기

기타 관리

학생화면으로 전환

학습활동

과목공지

- [중요] 중간고사 기말고사 성적 등록 확인 2022/06/20 21:18
- [홍보] 데이터 청년 캠퍼스 2022/06/12 19:45
- [중요] Notice for Final Exam 2022/06/08 21:36
- 기말고사 일정 공지 2022/06/02 11:57
- 강의 오늘 저녁 업로드 2022/05/10 14:37

강의 개요

과목공지 질의응답 자료실

주차 별 학습 활동

전체

1주차 [3월02일 - 3월08일]

공지, Q&A, 기타자료

Plato 화면 예시 2

The screenshot shows the Plato LMS interface for a course titled "강좌: 데이터과학입문" (Course: Introduction to Data Science) on the URL "plato.pusan.ac.kr/course/view.php?id=118872".

Left Sidebar:

- 강의실 홈** (Classroom Home)
- 강의정보** (Course Information)
 - 교수계획표 (국문)
 - 교수계획표 (영문)
 - 참여자목록
- 성적/출석관리** (Grade/Attendance Management)
 - 학습이력현황
 - 동영상이수현황
 - 스마트출석부
 - 화상강의목록
 - 성적부
 - 평가비율관리
- 수강생 알림** (Student Notification)
 - 쪽지 보내기
 - 이메일 보내기
- 기타 관리** (Other Management)
 - 학생화면으로 전환
- 학습활동** (Learning Activity)
 - 과제 (Assignment) - Highlighted with a blue arrow and the label "과제"
 - 더보기 (More)

Main Content Area:

- 데이터과학의 절차** (2022-03-09 00:00:00 ~ 2022-03-19 23:59:59, 34:33)
- 데이터의 분류** (PDF)
- 데이터의 분류** (2022-03-09 00:00:00 ~ 2022-03-19 23:59:59, 01:17:30)
- 3주차 [3월16일 - 3월22일]**
 - 탐색적 데이터 분석 01 (PDF)
 - 탐색적 데이터 분석 01 (2022-03-16 00:00:00 ~ 2022-03-26 23:59:59, 35:21)
 - 탐색적 데이터 분석 02 (PDF)
 - 탐색적 데이터 분석 02 (2022-03-16 00:00:00 ~ 2022-03-26 23:59:59, 01:03:06)
- 4주차 [3월23일 - 3월29일]**
 - 탐색적 데이터 분석 03 (PDF)
 - 탐색적 데이터 분석 03 (2022-03-23 00:00:00 ~ 2022-04-02 23:59:59, 57:01)
 - 나이브 베이즈 01 (PDF)
 - 나이브 베이즈 01 (2022-03-23 00:00:00 ~ 2022-04-02 23:59:59, 43:48)
 - HW01 EDA** (Assignment)
 - Check Google Classroom for the details.
- 5주차 [3월30일 - 4월05일]**
 - 나이브 베이즈 02 (PDF)

Bottom Right Corner:

- 공유 (Share)
- 편집 (Edit)
- ↑ TOP

분반과 주요일정

❖ 4개 분반 개설

- 059, 060, 061: 이명호
- 062: 전상률

❖ 059~061 분반 동일 내용 교육을 지향

- 교육 내용의 일관성 및 교육 품질 제고 기대
- 강의 자료/내용과 평가 도구 공유
- 강의 일정의 동기화

❖ 주요일정

- ~ 3월 8일(금) : 수강 정정 기간
- 공휴일
 - 5월 6일(월): 대체공휴일(어린이날)
 - 5월 15일(수): 부처님오신날
 - 6월 6일(목): 현충일
- 보강은 결강 주 금요일 (공지사항 확인 필수)
- 중간고사 기간 (4월 22일 ~ 28일) 중 중간고사
 - 26일(금) 예정
- 기말 고사 기간 (6월 17일 ~ 6월 21일) 중 기말고사
 - 21일(금) 예정

교수목표 및 강의개요

❖ 교수 목표

- 데이터 과학의 기초와 그 응용에 대한 이해
- 데이터 기반 의사 결정을 위한 문제구성, 데이터수집/처리/분석 방법 학습

❖ 강의 개요

- 인터랙티브하고 반복적인 데이터과학의 절차
- 데이터 과학을 위한 컴퓨팅 도구들 (google sheet, python)
- 데이터 기반 의사결정을 위한 통계 및 알고리즘
- 데이터, 분석모델 및 결과 해석
- 효과적인 데이터 시각화

❖ 강의 내용

- 데이터 과학 절차
- 데이터 타입 분류
- 탐색적 데이터 분석
- 통계적 분석 및 추론
- 머신러닝 기법:
필터링, 클러스터링, 분류, 예측
- 데이터 랭글링
- [기타주제 e.g., 수요반영]

1. Data Science
2. Causality and Experiments
3. Programming in Python
4. Data Types
5. Sequences
6. Tables
7. Visualization
8. Functions and Tables
9. Randomness
10. Sampling and Empirical Distributions
11. Testing Hypotheses
12. Comparing Two Samples
13. Estimation
14. Why the Mean Matters
15. Prediction
16. Inference for Regression
17. Classification
18. Updating Predictions

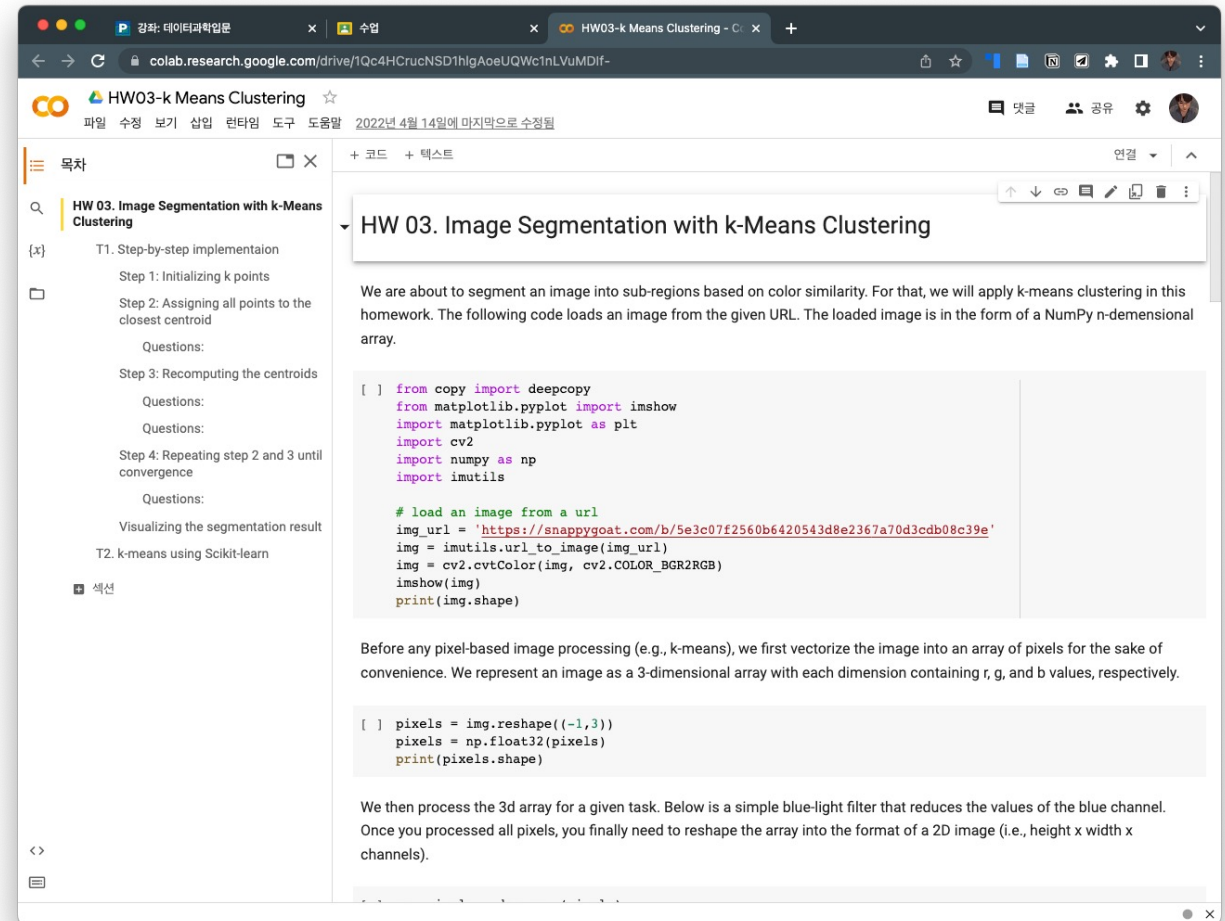
평가

❖ 평가 (100점 만점)

- 출석 및 참여 : 10점
- quiz/lab/homework : 30점
- 중간고사 30점, 기말고사 30점
- 평가 단위 결합 여부는 미정; 원칙은 분반 별 평가

❖ HOW TO GET “F”

- 출석일 1/3이상 결석
- 중간불참
- 기말불참
- 다른 사람 과제 그대로 카피



강사 및 조교 연락처

❖ 강사: 이명호 (Myungho Lee)

- myungho@pusan.ac.kr
- Office hours (Flexible):
 - 13:00 – 15:00, 월수
 - Please make an appointment first
- 자연대연구실험동 410호

❖ 조교

- 자연대연구실험동 418호 확장현실연구실
- 059: 김태연 bigguy31@pusan.ac.kr
- 060: 김동현 emboob@pusan.ac.kr
- 061: 김민성 kim.min.seong@pusan.ac.kr

❖ 이메일 제목

- [IDS_분반_이름] 용건
- e.g., [IDS_059_유재석] 코로나 자가격리
- pusan.ac.kr 이메일 사용

❖ 수업 관련 질문은 되도록 플라토 게시판 활용

❖ 플라토 쪽지 연락 지양