

Introduction to Data Science

Lecture 1.1 Introduction

Instructor: Sangryul Jeon

School of Computer Science and Engineering

srjeonn@pusan.ac.kr



About me



Sangryul Jeon

Education and Professional experience

Pusan National University (2023.09~Present)

Assistant professor, School of Computer Science and Engineering

University of Michigan, Ann arbor (2023.03~2023.08)

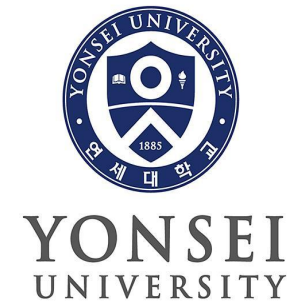
Postdoctoral scholar, Electrical Engineering and Computer Sciences Department

University of California, Berkeley (2022.03~2023.02)

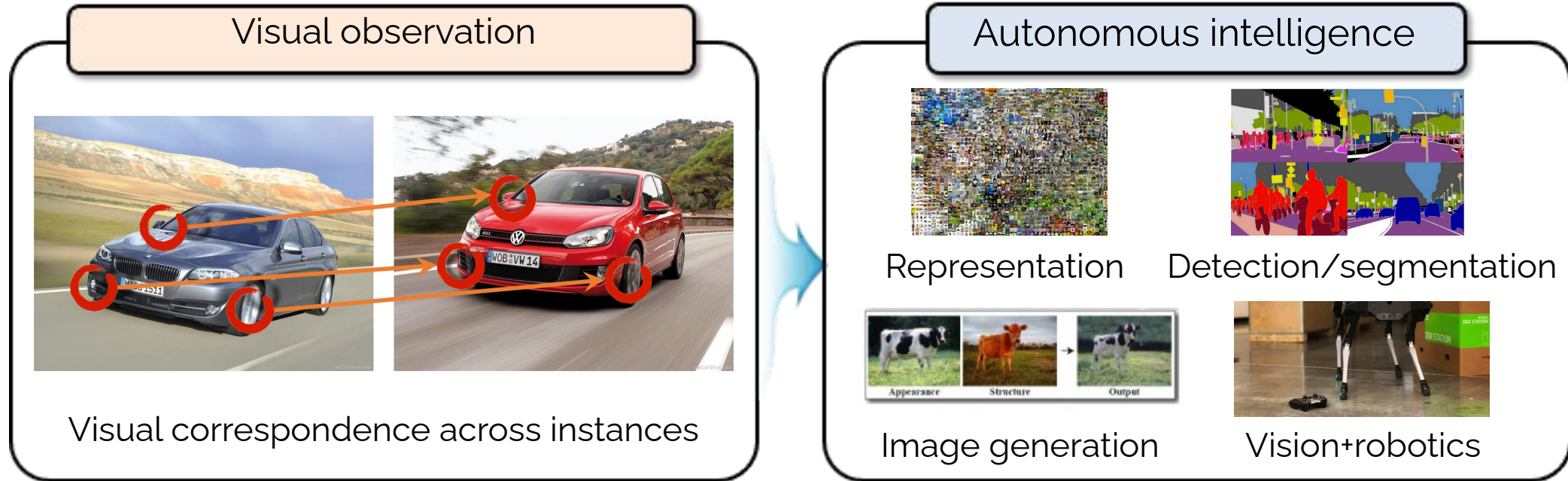
Postdoctoral scholar, Electrical Engineering and Computer Sciences Department

Adobe Research (2021)

Research Intern, Creative Intelligence Lab



About me

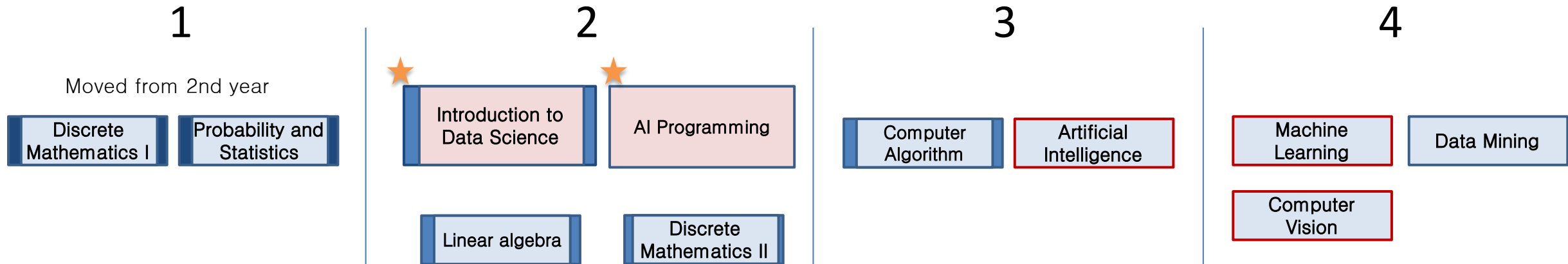


1. **Jeon, S.**, et al., "Pyramidal Semantic Correspondence Networks", **TPAMI' 2022 - IF 24.314, JCR ranking 0.7%**
2. **Jeon, S.**, et al., "Guided Semantic Flow", **ECCV' 2020**
3. **Jeon, S.**, et al., "PARN: Pyramidal Affine Regression Network for Semantic Correspondence", **ECCV' 2018**
4. **Jeon, S.**, et al., "Mining Better Samples for Contrastive Learning of Temporal Correspondence", **CVPR' 2021**
5. **Jeon, S.**, et al., "Joint Learning of Semantic Alignment and Object Landmark Detection", **ICCV' 2019**

Course Introduction

❖ CB35533: Introduction to Data Science

- Mandatory Major class with 3 credits, newly added from 2020



- Required courses: “Introduction to computer and programming” and “Probability and Statistics” courses
- This course will be acknowledged as Elective course for those whose student number starts from 2019
- Theory-based course, but we will practice it through the implementations during the class and HW

Text book

❖ Couple of text books are referred to prepare slides

- PNU library eBook – [O'Reilly for Higher Education](#)

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

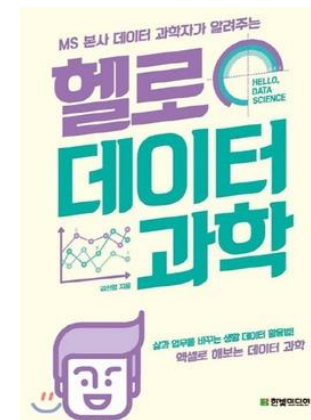
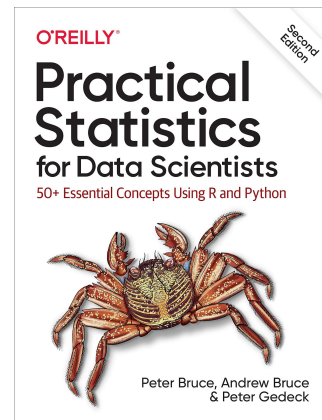
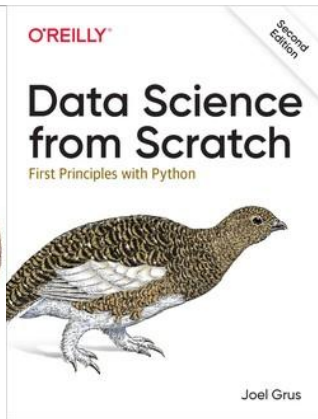
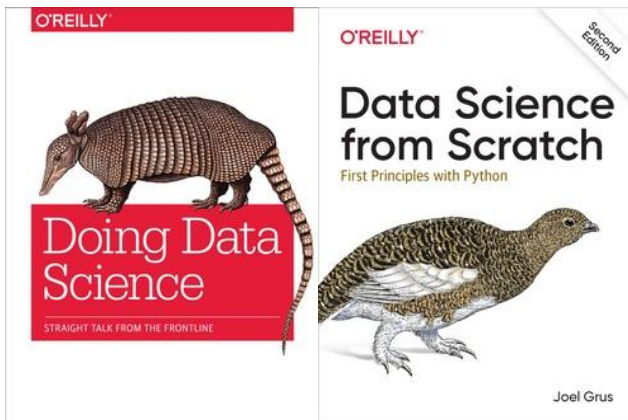
- 데이터 과학 입문, 한빛미디어

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

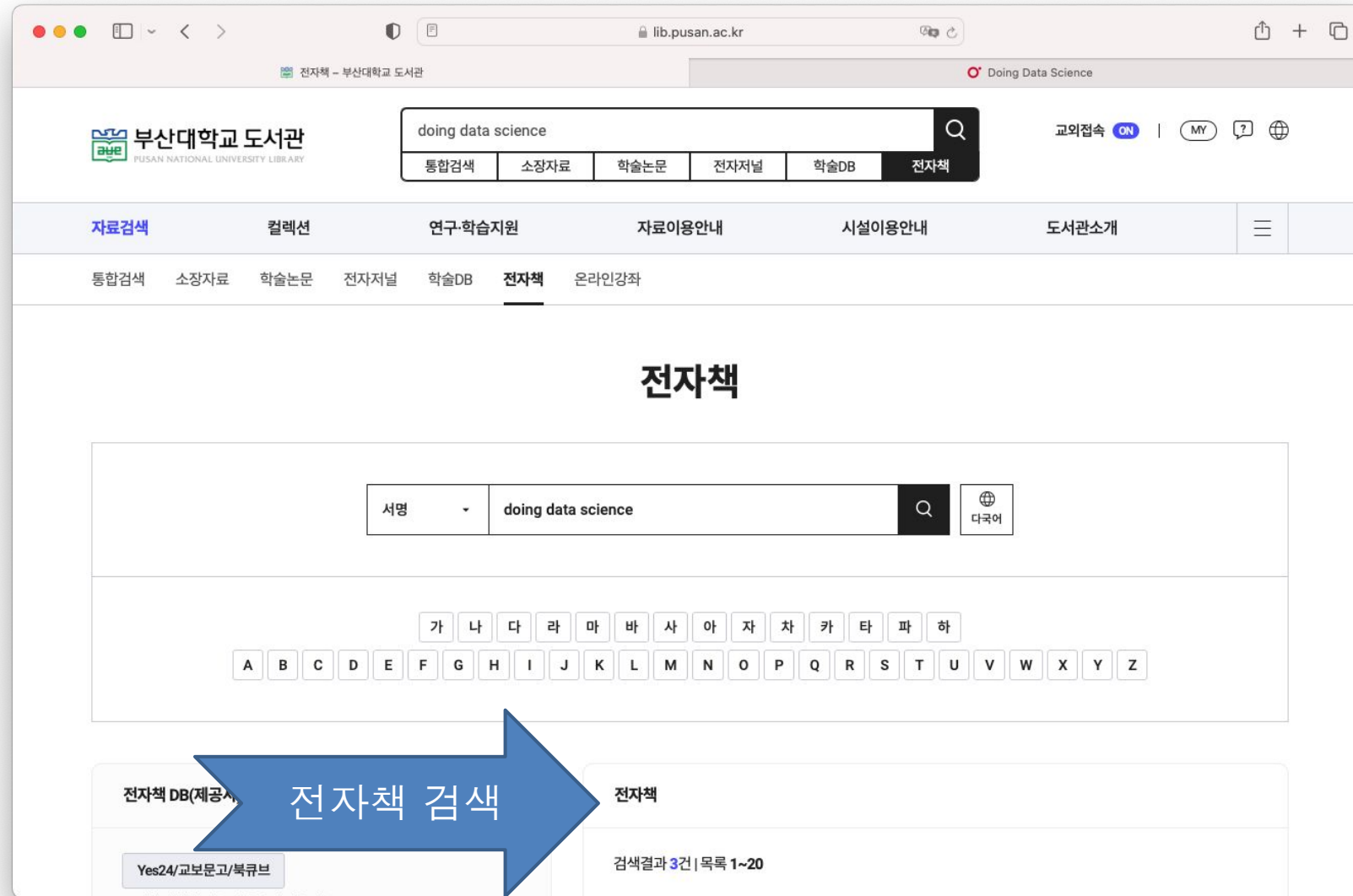
- 밑바닥부터 시작하는 데이터과학 2판, 인사이트

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



Our tools

❖ PLATO

- Attendances, Notices, Slide materials, HW, Lecture videos for make-up classes, etc



❖ Google Colab, Docs, Sheet

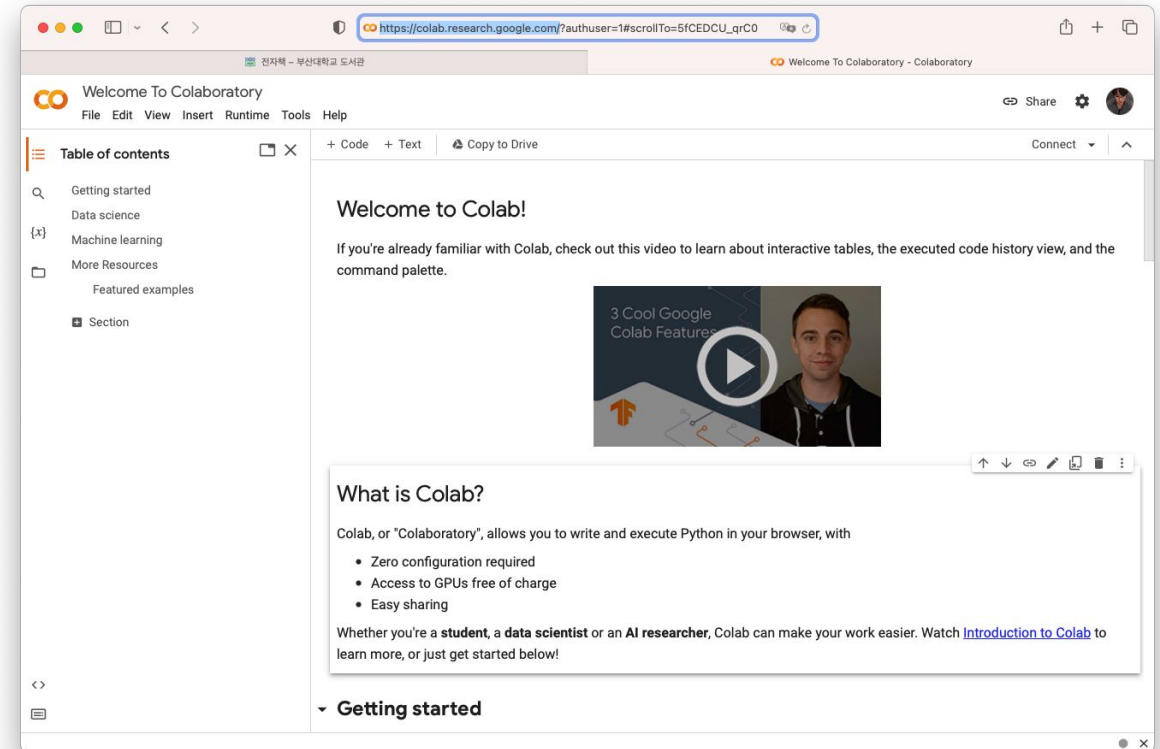
- Based on Google Cloud
- We will utilize these for conducting HWs and their evaluations
- [Colab](#) / Google Sheet & Doc

❖ Use @pusan.ac.kr account for colab



❖ Python

❖ Markdown



Example of Plato website

The screenshot shows a web browser window displaying the Plato website. The 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 main content area is titled '데이터과학입문 (CB3500533-060)' and features a sidebar with a '강의실 홈' menu. The sidebar menu includes '강의정보' (Course Information), '성적/출석관리' (Grade/Attendance Management), and '기타 관리' (Other Management). The '강의정보' section lists '교수계획표 (국문)', '교수계획표 (영문)', and '참여자목록'. The '성적/출석관리' section lists '학습이력현황', '동영상이수현황', '스마트출석부', '화상강의목록', and '성적부'. The '기타 관리' section includes '이메일 보내기' and '학생화면으로 전환'. The main content area displays '과목공지' (Course Notice) with a list of notices and dates, and '강의 개요' (Course Overview) with three icons: '과목공지' (Course Notice), '질의응답' (Q&A), and '자료실' (Resource Room). A blue arrow points from the text 'Notice Q&A, Materials' to the '과목공지' icon.

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]

Notice Q&A, Materials

Example of Plato website

The screenshot displays the Plato website interface for a course titled "데이터과학입문" (Introduction to Data Science). The browser address bar shows the URL "plato.pusan.ac.kr/course/view.php?id=118872".

Left Sidebar:

- 강의실 홈**
- 강의정보**
 - 교수계획표 (국문)
 - 교수계획표 (영문)
 - 참여자목록
- 성적/출석관리**
 - 학습이력현황
 - 동영상이수현황
 - 스마트출석부
 - 화상강의목록
 - 성적부
 - 평가비율관리
- 수강생 알림**
 - 쪽지 보내기
 - 이메일 보내기
- 기타 관리**
 - 학생화면으로 전환
- 학습활동**
- 과제** (Add button: 추가)
- 더보기** (Dropdown arrow)

Main Content Area:

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

Bottom Right:

- 공유 (Share icon)
- 편집 (Edit icon)
- ↑ TOP (Back to top icon)

A large blue arrow labeled "HW" points to the "HW01 EDA" assignment in the 4th week.

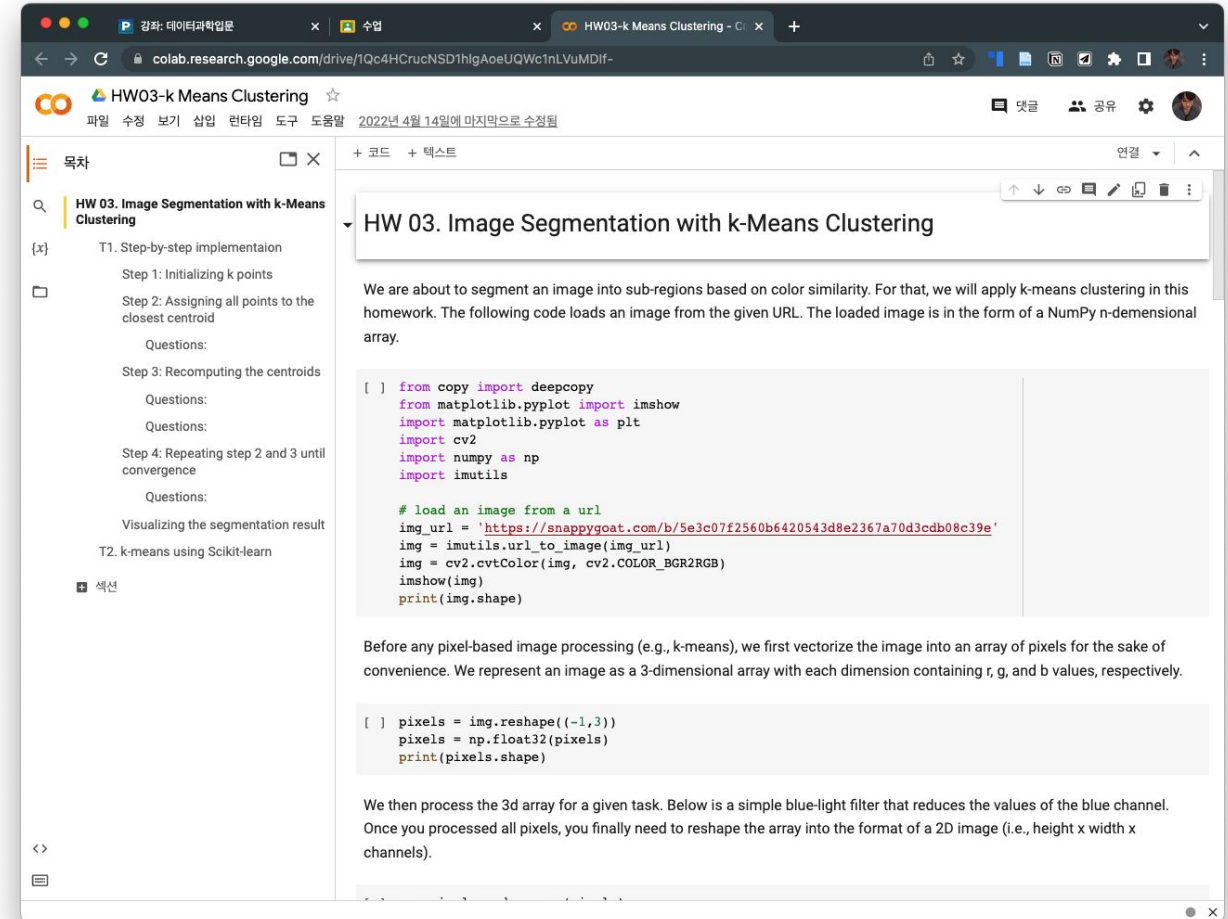
Grading policy

❖ Criteria

- Attendance 10%
- Quiz/lab/homework 30% - Half credit will be given for late submissions (until the release of solution)
- Midterm exam 30%, Final exam 30%

❖ HOW TO GET “F”

- Being absent over 1/3 of the total classes
- Being absent at the midterm exam
- Being absent at the final exam
- Copying HW from other students



Grading policy

❖ Ratio

Standard of PNU

- A+~A0 less than 50%

Internally just for CSE department

- A+ less than 15%
- A+~A0 less than 40%
- A+~B0 less than 70%
- Below C+ over 30%

Q&A

Prof. Jeon, Sangryul

Computer Vision Lab.

Pusan National University, Korea

Tel: +82-51-510-2423

Web: <http://sr-jeon.github.io/>

E-mail: srjeonn@pusan.ac.kr