

Python을 이용한 데이터 분석

박중양

경상대학교 자연과학대학 정보통계학과
joongyang.park@gmail.com

강의 시간

- 9월 28일(화) 10:00 12:00
- 9월 28일(화) 14:00 18:00
- 9월 29일(수) 10:00 12:00
- 9월 29일(수) 14:00 18:00

수강생 준비 사항

- 인터넷에 연결된 컴퓨터
- Chrome web browser
- Google 계정
- Google Driver 사용법 숙지

목차(Table of Contents) I

- ❶ 목차
- ❷ 0. 강의와 수강 준비
- ❸ 1. Python 설치와 이용법
- ❹ 2. 정리된 데이터와 데이터 파일
- ❺ 3. Jupyter Notebook과 Markdown
 - 3.1 Jupyter Notebook
 - 3.2 Markdown
- ❻ 4. Python programming 기초
 - 4.1 객체 이름, 할당문
 - 4.2 기본 자료형
 - 4.3 표현식
 - 4.4 자료구조
- ❼ 5. 통계 분석, 기계학습, 빅 데이터 분석, 인공지능을 위한 파이썬 패키지
- ❽ 6. CSV 파일을 데이터프레임(dataframe) 객체로
- ❾ 7. 데이터에 관한 정보 탐색
 - 7.1 전반적인 정보 살펴보기

목차(Table of Contents) II

- 7.2 열 이름 조회와 변경
- 7.3 행과 열의 수 조회
- 7.4 결측치 탐색
- 7.5 자료의 일부분 추출
- 7.6 열의 추가/삭제
- 7.8 척도의 지정

⑩ 8. 데이터프레임 객체를 파일로 저장하기

⑪ 9. 기본적인 통계분석

⑫ 10. 그래프 작성하기

- 10.1 분포를 표현하는 그래프 작성
- 10.2 관계를 알아보기 위한 그래프 작성
- 10.3 집단별 분석

⑬ 11. 회귀분석

강의 자료 내려받기

- Web browser에서 **Python을 이용한 데이터 분석:**
<https://github.com/joongyang/Python-for-Data-Analysis>에 접속한다.
- **Code** 버튼을 눌러 나타나는 메뉴에서 **Download ZIP**을 눌러서 강의 자료 압축 파일을 내려받는다.
- 내려받은 압축 파일의 압축을 푼다.
- Google 사이트에 로그인 한다.
- Google Drive에 연결한다.
- Google Drive에서 오른쪽 마우스 버튼을 눌러서 나타나는 메뉴에서 **새폴더**선택하여 강의 자료를 저장할 폴더를 만든다.
- 새로 만든 폴더를 더블 클릭한다.
- 새로 만든 폴더에서 오른쪽 마우스 버튼을 눌러 나타나는 메뉴에서 **파일 업로드**를 선택한 다음 압축을 푼 강의 자료 파일(확장자가 ipynb인 파일들)업로드 한다.

강의 자료로 배포된 주피터 노트북 실행하기

- 강의 자료가 저장된 Google Drive의 폴더에 접속한다.
- Google Drive에서 오른쪽 마우스 버튼을 누르면 나타나는 메뉴에서 **더보기**와 **연결할 앱 더보기**를 순차적으로 선택한다.
- **Colaboratory**를 검색한 다음 설치한다. (이 작업은 한 번만 하면 된다.)
- 기존 주피터 노트북 사용할 때: 주피터 노트북을 선택한 다음 오른쪽 마우스 버튼을 눌러 나타나는 메뉴에서 **연결 앱**과 **Google Colaboratory**를 순차적으로 선택한다.
- 새 주피터 노트북을 생성할 때: Google Drive의 빈 곳에서 오른쪽 마우스 버튼을 눌러서 나타나는 메뉴에서 **더 보기**와 **Google Colaboratory**를 순차적으로 선택한다.

내려받은 주피터 노트북에 모든 강의 내용이 담겨있다.

