

# □ Using Python for micro-data construction and preparation for empirical analysis: Python을 활용하여 실증 분석을 위한 미시 데이터 제작 및 준비하기

## 1. 과정개요

| 담당강사      | 성명                                                                                                                                                                                                                                                                                                                                                                                                                                       | E-mail             |
|-----------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------|
|           | 양현주                                                                                                                                                                                                                                                                                                                                                                                                                                      | hyang@sogang.ac.kr |
| 강의시간      |                                                                                                                                                                                                                                                                                                                                                                                                                                          |                    |
| 강의방식      | 온라인 녹화 강의 4시간 및 실시간 2시간 강의 진행                                                                                                                                                                                                                                                                                                                                                                                                            |                    |
| 강의목표 및 개요 | <p>Python 초보자로서 빅데이터 분석 등 실증 분석을 위해 미시 데이터 전처리 작업이 할 때 많이 사용하게 되는 아래 작업에 대한 설명 및 실습 (pandas library 사용)</p> <ol style="list-style-type: none"> <li>1) 많은 수의 csv, 엑셀 데이터 및 엑셀 tab을 하나의 파일로 취합하는 방법 실습</li> <li>2) 압축된 csv 파일을 압축을 풀지 않은 상태에서 불러오기 및 저장하기</li> <li>3) 데이터 병합 (merge) 작업</li> <li>4) 전체 데이터에서 필요한 변수 및 자료만 추출</li> <li>5) 데이터를 STATA/SAS/SPSS/CSV/EXCEL 등 데이터 파일 형태로 import / export 작업</li> <li>6) API 데이터 끌어오기</li> </ol> |                    |
| 참가 대상     | 대학원생, 연구자<br>(실증 분석을 위해 빅데이터 등 사용 시 효율적인 미시 데이터 전처리를 필요로 하는 연구자)                                                                                                                                                                                                                                                                                                                                                                         |                    |
| 선수 과목     | 선수과목은 없으나 김수현 박사의 "Python을 활용한 데이터 분석" 방법론 강의 선수강 또는 병행 수강 추천함.                                                                                                                                                                                                                                                                                                                                                                          |                    |
| 수업운영방식    | 강의와 실습 병행, 충분한 질문, 답변 기회 제공, 강의 완료 후 실습 과제 제출                                                                                                                                                                                                                                                                                                                                                                                            |                    |
| 교재 및 참고문헌 | 교재 없음.<br>Lecture slides와 python code 제공                                                                                                                                                                                                                                                                                                                                                                                                 |                    |
| 실습자료      | 실습용 자료 제공 (추후 공지)                                                                                                                                                                                                                                                                                                                                                                                                                        |                    |
| 사용프로그램    | Anaconda Individual Edition (open-source software로 누구나 설치 가능)<br><a href="https://www.anaconda.com/products/individual#Downloads">https://www.anaconda.com/products/individual#Downloads</a>                                                                                                                                                                                                                                             |                    |
| 담당교수 정보   | <p><b>[학력]</b><br/> 2016 Brown University 경제학 박사<br/> 2010 London School of Economics 경제학 석사<br/> 2008 인천대학교 국제통상학 학사</p> <p><b>[경력]</b><br/> 2020-현재 서강대학교 경제학부 조교수<br/> 2017-2020 인천대학교 동북아국제통상학부 조교수<br/> 2016-2017 KDI 재정복지부 부연구위원<br/> 2013 World Bank Short-term consultant<br/> 2009 International Crops Research Institute for Semi-Arid Tropics (ICRISAT) research scholar</p> <p><b>[연구]</b></p>                               |                    |

|  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  | <p>Crowding out effects of an old-age pension program on intergenerational transfers: Evidence from South Korea (with Kanghyock Koh), <b>Economic Development and Cultural Change</b> (accepted)</p> <p>Learning to score economic development from satellite imagery, (with Sungwon Han, Donghyun Ahn, Sungwon Park, Jeasurk Yang, Susang Lee, Jihee Kim, Sangyoon Park and Meeyoung Cha), <b>2020, Proceedings of the Association for Computing Machinery (ACM) SIGKDD 2020 Conference on Knowledge Discovery and Data Mining (KDD)</b></p> <p>Is there regional inequality in medical accessibility of the severely injured? Application of driving time data in South Korea (with Elliot Kang, Min Gyung Kim, and Kanghyock Koh), <b>2020, Journal of Market Economy</b>, 49(1)</p> <p>Family clans and public goods: Evidence from the New Village Beautification Project in South Korea, <b>2019, Journal of Development Economics</b>, 136</p> <p>The impact of customer loyalty on restaurant sanitation grades and the importance of narrative information: The case of New York Restaurant Sanitation Grading System (with Min Gyung Kim and Anna Mattila), <b>2018, Cornell Hospitality Quarterly</b>, 59(3)</p> <p><b>[수상실적]</b><br/>         2019 한국경제학회 한국경제학술상 거시부문 수상<br/>         2019 인천대학교 학술연구상<br/>         2016 한국학 부문 Best Paper Award, University of Hawaii at Manoa, School of Pacific and Asian Studies Graduate Student Conference</p> |
|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

## 2. 강의계획

| 일 시 | 강의 및 실습 내용                                                                                                                                                                                                                                                                                                                                       |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|     | <p><b>Part 1: Python 사용을 위한 소프트웨어와 라이브러리 설치</b></p> <ul style="list-style-type: none"> <li>- Anaconda와 Jupyter Notebook 설치</li> <li>- Python 라이브러리 설치 및 업데이트 방법 설명</li> <li>- Python environment 설명</li> <li>- Jupyter Notebook 사용 방법 설명</li> </ul>                                                                                              |
|     | <p><b>Part 2: 데이터 전처리 작업</b></p> <ul style="list-style-type: none"> <li>- CSV 및 Excel 단일 파일 불러오기</li> <li>- 동일 폴더의 여러 CSV 및 Excel 파일을 한꺼번에 불러오기</li> <li>- 압축된 CSV, STATA/SAS/SPSS/CSV/EXCEL 파일 불러오기 및 저장</li> <li>- 데이터 병합 (merge) 작업</li> <li>- 전체 데이터에서 필요한 변수 및 자료만 추출</li> <li>- Regular expressions 소개</li> <li>- API data 끌어오기</li> </ul> |