

KDI 국제정책대학원

2019년 1분기 연구방법론세미나 강의계획서

□ 사회정책 연구자를 위한 빅데이터 분석

1. 과정개요

담당강사	성명	E-mail
	강정한	jhk55@yonsei.ac.kr
강의시간 / 강의실	2019년 2월 11일(월) ~ 2월 14일(목) (19:00~22:00) / KDI 대학원 3층 Computer Lab	
강의목표 및 개요	본 강좌는 생산자로서가 아니라 소비자, 혹은 협업자로서 빅데이터 분석을 이해하는데 초점을 둔다. 따라서 직접적 코딩 실습보다는, 데이터 과학의 언어를 이해하고 전문적 데이터 과학자와 소통능력을 키우는데 목적이 있다. 행렬, 벡터, 네트워크 분석, 토픽 모델링, Word2Vec, 온라인 실험 등이 사회정책적 문제와 결합해 구현되는 사례들을 소개한다.	
참가 대상	대학원생 및 일반 연구자	
선수 과목	기초 통계학 (t-test, one-way ANOVA 정도)	
수업운영방식	사례 중심 강의	
교재 및 참고자료	1. Salganik, Matthew (2018). Bit by Bit. Princeton University Press. 2. Angrist, J. D., & Pischke, J. S. (2008). Mostly harmless econometrics: An empiricist's companion. Princeton university press. 3. 스티븐 래빗. 괴짜경제학. 웅진지식하우스. (강의용 슬라이드 제공하므로 구입이 의무는 아님)	
실습자료	필요에 따라 수업시간에 배포	
사용프로그램	R, Python, STATA 등 (강사의 시연 목적이므로 수강자가 설치할 의무는 없음)	
기타 사항	평소 분석 소프트웨어를 사용할 때 구글검색을 자주 활용하기를 추천합니다.	
담당교수 정보	학력 2006. 6 미국 시카고대학 대학원 사회학 박사 1998. 2 서울대학교 대학원 사회학 석사 1996. 2 서울대학교 수학과 학사 경력 2008. 3 - 현재 연세대학교 사회학과 조교수, 부교수, 교수 2017, 1학기 KAIST 전산학부 방문교수 2014 미국 코넬대학교 사회학과 방문교수 논문 및 저서 송준모·강정한. 2018. "메갈리아의 두 딸들 : 익명성 수준에 따른 온라인 커뮤니티의	

	정체성 분화.” 『한국사회학』52(4): 161-206. Kang, Jeong-han and Da Young Chung. 2017. “Homophily in an Anonymous Online Community: Sociodemographic vs. Personality Traits.” <i>Cyberpsychology, Behavior, and Social Networking</i> 20(6): 376-381. Kang, Jeong-han, Jibum Kim, and Min-Ah Lee. 2016. “Marital Status and Mortality: Does Family Structure in Childhood Matter?” <i>Social Science and Medicine</i> 159: 152-160. Kang, Jeong-han, Seung-Yeal Ha, Kyungkeun Kang, and Eunhee Jeong. 2014. “How Do Cultural Classes Emerge from Assimilation and Distinction?: An Extension of the Cucker-Smale Flocking Model.” <i>Journal of Mathematical Sociology</i> 38(1): 47-71. 강정한·조연정. 2013. “온라인 연결망 불평등에 대한 학제적 이론화 모색(Toward an Interdisciplinary Perspective of Online Network Research).” 『한국사회학』47(4): 107-137. Bothner, Matthew S., Jeong-han Kang, and Toby E. Stuart. 2007. “Competitive Crowding and Risk Taking in a Tournament: Evidence from NASCAR Racing.” <i>Administrative Science Quarterly</i> 52 (2): 208-247. Kang, Jeong-han. 2007. “The Usefulness and Uselessness of the Decomposition of Tobit Coefficients.” <i>Sociological Methods and Research</i> 35 (4): 572-582. 조연정·강정한. 2015.『카카오톡은 어떻게 공동체가 되었는가』다산 출판사.
--	---

2. 강의계획

일 시	강의 및 실습 내용
2월 11일 (월)	데이터 사이언스 개괄 - Language: Python, R - Useful Sites: Stack Overflow, 생활코딩 - Self-help
2월 12일 (화)	빅데이터 분석과 기계학습 - Nowcasting, Precision & Recall - Term Frequency - Topic Modeling - Word2Vec
2월 13일 (수)	(온라인) 실험의 이해 - Lab experiment - Online experiment - Amazon Mechanical Turk (MTurk), Survey Money
2월 14일 (목)	정책평가와 데이터 사이언스 - Field experiment, A/B test - Counter-factual analysis with non-experimental data - Performative feedback loops

□ QGIS를 활용한 GIS데이터 분석

1. 과정개요

담당강사	성명	E-mail
	양현주	hyang263@gmail.com
강의시간 / 강의실	2019년 2월 18일(월) ~ 2월 19일(화) (19:00~22:00) / KDI 대학원 3층 Computer Lab	
강의목표 및 개요	1. GIS data의 기본적인 구조와 사용 방법 이해 2. 사회과학 분야에서 GIS data 이용 분석 사례 소개 3. QGIS software를 이용하여 GIS data 정보 탐색 4. QGIS로 다양한 지도 작성 5. QGIS로 위성데이터, 행정빅데이터 자료 등을 연구 분석에 활용	
참가 대상	학부, 대학원생, 연구자 (GIS data를 활용하여 지도를 그리거나 통계 분석에 사용하려는 사람)	
선수 과목	없음	
수업운영방식	강의와 실습 병행. 충분한 질문, 답변 시간 제공	
교재 및 참고자료	강의 slides	
실습자료	실습용 자료 제공 (추후 공지)	
사용프로그램	QGIS	
기타 사항	강의실 컴퓨터 혹은 개인노트북 (QGIS는 open-source software로 누구나 설치 가능)	
담당교수 정보	학력 2016 Brown University 경제학 박사 2010 London School of Economics 경제학 석사 2008 인천대학교 국제통상학 학사 경력 2017-현재 인천대학교 동북아통상학부 조교수 2016-2017 KDI 재정복지부 부연구위원 2013 World Bank Short-term consultant 2009 International Crops Research Institute for Semi-Arid Tropics (ICRISAT) research scholar 연구 Family clans and public goods: Evidence from the New Village Beautification Project in South Korea, 2019, Journal of Development Economics, 136 The impact of customer loyalty on restaurant sanitation grades and the importance of narrative information: The case of New York Restaurant Sanitation Grading System, 2018 (with Min Gyung Kim and Anna Mattila), Cornell Hospitality Quarterly, 59(3) Crowding out effects of an old-age pension program on intergenerational transfers: Evidence from South Korea, with Kanghyock Koh (revision requested at Economic	

	Development and Cultural Change) Did high speed trains promote local economic activities?: Evidence from the introduction of Korea Train Express, with Kanghyock Koh (under review) Is there regional inequality in medical accessibility? Application of driving time data to measure trauma center accessibility in South Korea, with Minkoo Kang, Min Gyung Kim, and Kanghyock Koh (under review) Effects of visual cues and social density on beverage consumption: A Field experiment in a bar, with Min Gyung Kim and Anna Mattila (under review)
--	---

2. 강의계획

일 시	강의 및 실습 내용
2월 18일 (월)	QGIS로 GIS data 익숙해지기 - 사회과학에서 GIS data 이용 분석 사례 소개 - GIS 데이터의 종류 및 구조 설명 - QGIS로 GIS 자료 불러오기 및 자료 정보 탐색 - QGIS를 활용한 논문 및 발표용 지도 작성
2월 19일 (화)	QGIS를 이용하여 위성데이터, 행정빅데이터 자료 등을 연구 분석에 활용 - 건물 및 행정구역의 면적 계산 - 두 지점 간의 직선 거리 측정 - 위/경도 계산 - 야간 조도(night time light intensity) 위성 데이터를 활용한 지역 경제활동 변수 만들기

□ Panel Data Analysis

1. 과정개요

담당강사	성명	E-mail
	김근태	sochyunsik@khu.ac.kr
강의시간 / 강의실	2019년 2월 20일(수) ~ 2월 28일(목) (19:00~22:00)(2월 22일(금) ~ 2월 24일(일) 휴강)/ KDI 대학원 3층 Computer Lab	
강의목표 및 개요	패널 데이터를 이해하고 고정효과(fixed-effect) 및 확률효과(random effect) 개념의 이해 및 활용에 대해 논의하며 도구변수를 활용한 패널모형 추정법을 다룬다. 종속변수가 연속변수가 아닌 경우 패널 분석기법을 이해한다. 민인식·최필선(2012) 교재에서 다루는 모형들을 이해하고 Stata에서 실습하여 다양한 패널 모형을 습득한다.	
참가 대상	학부, 대학원생 및 일반 연구자 (사회과학에서 사용하는 통계적 연구방법에 관심 있는 사람들은 누구나)	
선수 과목	기초 통계학, 최소제곱법(least squares estimation), 로짓(logit) 및 프로빗(probit) 모형	
수업운영방식	강의와 실습을 병행합니다.(강의 2/3, 실습 1/3)	
교재 및 참고문헌	강사의 PPT 노트가 제공됩니다. 기본적으로 다음의 책에 기반하여 논의하기 때문에 이 책을 읽고 오면 많은 도움이 됩니다; 민인식, 최필선(2012), 패널데이터분석: STATA Panel Data Analysis. (쥬지필미디어).	
실습자료	강사가 제공하는 실습용 자료 (추후 공지)	
사용프로그램	STATA	
기타 사항	강의실에 컴퓨터와 STATA 프로그램이 준비되어 있으니, 개인적으로 지참하지 않으셔도 됩니다.	
담당교수 정보	학력 2011 University of Wisconsin-Madison 사회학 박사 *Dissertation title: "Dynamic Development of Cognitive Skills and Non-Cognitive Traits in Childhood." 2009 University of Wisconsin-Madison 통계학 석사 2007 University of Wisconsin-Madison 사회학 석사 2003 서울대학교 사회학 석사 2001 서울대학교 사회학 학사 경력 2018. 03. - 현재 경희대학교 정경대학 사회학과 부교수 2013. 03. - 2018. 02. 경희대학교 정경대학 사회학과 조교수 2012. 02. - 2013. 02. 부연구위원, 저출산고령사회연구실, 한국보건사회연구원 2011. 08. - 2012. 01. 박사후과정 Wisconsin Scholars Longitudinal Study, Wisconsin Center for Education Research, University of Wisconsin-Madison 저서 *Kim, Hyun Sik. 2018. Fertility differentials between Korean and international marriage couples in	

	South Korea. Asian Population Studies, 14(1), 43-60. http://dx.doi.org/10.1080/17441730.2018.1426188 *Kim, Hyun Sik. 2016. Sexual debut and mental health among South Korean adolescents. Journal of Sex Research, 53(3), 313-320. DOI: http://dx.doi.org/10.1080/00224499.2015.1055855 *Kim, Hyun Sik. 2015. Active life expectancy of elderly Koreans, 1994-2011. Asian Population Studies, 11(3), 312-327. DOI: http://dx.doi.org/10.1080/17441730.2015.1096096 *Kim, Hyun Sik. 2015. Women's wages and fertility hazards in South Korea. Asian Women, 31(2), 1-27. DOI: http://dx.doi.org/10.14431/aw.2015.06.31.2.1 *Kim, Hyun Sik. 2015. Effects of sexual intercourse on suicidal behaviors among adolescents in South Korea. Journal of Korean Academy of Nursing, 45(2), 183-191. DOI: http://dx.doi.org/10.4040/jkan.2015.45.2.183 *Kim, Hyun Sik. 2015. Foregone opportunities: Unveiling teacher expectancy effects in kindergarten using counterfactual predictions. Social Psychology of Education, 18, 273-296. DOI: http://dx.doi.org/10.1007/s11218-014-9284-4 *Kim, Hyun Sik. 2014. Female labour force participation and fertility in South Korea. Asian Population Studies, 10(3), 252-273. DOI: http://dx.doi.org/10.1080/17441730.2014.943941
--	---

2. 강의계획

일 시	강의 및 실습 내용
2월 20일 (수)	인과적 추론 -역인과관계, 혼동변수, 매개변수, 조절변수, 상충변수
2월 21일 (목)	종단 자료 구축 기초분석 회귀 분석 기초(최소제곱법)
2월 25일 (월)	패널GLS Between Effects 모형 고정효과모형
2월 26일 (화)	1차차분모형 확률효과모형
2월 27일 (수)	가설검정 기본가정위반: AR(1)오차항과 설명변수 내생성
2월 28일 (목)	동적패널모형 이항 종속변수 모형