

교육 과정 소개서.

논문/보고서 작성을 위한 통계분석 11기 CAMP



코스요약

코스명	논문/보고서 작성을 위한 통계분석 CAMP
기간	2020.07. 04 ~ 2020. 08. 29 (2020. 08. 15 휴강)
일정	매주 토요일 10:00 ~ 13:00 (주 1회, 총 24시간)
장소	패스트캠퍼스 강남강의장
준비물	개인 노트북
담당자	02-568-9886 / help-ds@fastcampus.co.kr
수강료	1,100,000
상세페이지 url	fastcampus.co.kr/data_camp_anova/

코스목표

스스로 통계분석을 진행하고, 논문 및 보고서에 활용할 수 있으며, 논문에 나오는 통계분석 결과를 해석할 수 있습니다. 또한, 현업에 필요한 통계개념을 확실하게 익히고, 이를 업무에 활용할 수 있습니다.

코스정보

논문/보고서를 작성할 때 통계분석이 걱정되신다면, 필요한 내용만 핵심적으로 짚어주는 8주 커리큘럼을 통해 분석 결과 도출, 해석 그리고 R 활용법까지 익힐 수 있습니다.



코스특징

지금 바로 써먹을 수 있는 실용적인 통계

확률과 통계는 대학에서도 필수교양으로 수강할만큼 많은 사람들이 찾고 중요하게 생각하는 부분입니다. 하지만, 중요도에 비해 그 개념을 제대로 이해하고 활용하는 사람은 드물죠. 배우긴 배웠지만 가물가물한 통계, 이제는 언제라도 써먹을 수 있게 실습 중심으로 핵심만 짚어 드립니다.

통계 분석 실습, 결과 해석, 진짜 작성해보는 논문/보고서

가설을 세워 R로 통계분석을 하고 결과를 해석하는 방법을 배우고, 분석한 통계 내용을 바탕으로 함께 논문을 작성해봅니다. 논문 작성을 위한 템플릿 제공과 구조적 글쓰기에 대해 알려드리는 것은 기본! 주의점 등에 대해서도 꼼꼼히 알려 드려서, 실수하지 않도록 기초를 잡아 드립니다.

통계분석을 넘어 머신러닝, 회귀분석까지

수강생 분들이 조금 더 심도있는 데이터 분석을 할 수 있도록 논문, 연구 및 실무에서 활용되는 통계분석을 넘어, 머신러닝과 회귀분석까지 폭넓게 다룹니다. 앞으로 통계분석의 활용법은 물론, 수업시간에 배운 내용들을 기반으로 하여 데이터 분석 분야에서 새로운 기회를 잡으실 수 있도록 방향성까지 잡아 드립니다.

매주 제공되는 퀴즈와 꼼꼼한 피드백

아무리 강의를 듣고 혼자 책을 들여다봐도 그 개념이 머릿속에서 정리되지 않을까봐 걱정되신다면, 매주 제공되는 퀴즈와 강사님의 꼼꼼한 피드백으로 수업 내용을 100% 본인의 것으로 만들어 보세요.

수강생-강사간의 네트워킹 지원

8주간의 과정을 함께한 수강생들과 강사님, 조교님과의 네트워킹을 적극 지원합니다. 다양한 분야에서 통계분석이라는 같은 주제로 모인 분들과의 적극적인 네트워킹으로 한 단계 성장할 수 있는 기회입니다.



커리큘럼

	<Step 1. 통계분석의 준비단계, 데이터 접근과 이해>
1회차	데이터의 통계적인 접근법 소개 [이론] - 기존과 현재의 통계학의 차이 - 데이터 분석시 고려해야 할 4가지 : adjust/facet/outlier/missing - 기술통계와 추론통계의 차이 [실습] - R 및 R studio 설치 - 원하는 데이터 접근 및 추출 - R에서의 기본적인 문법 소개 - data outline 파악 - Rmarkdown
2회차	통계분석을 위한 데이터 이해 및 전처리 [이론] - 표와 시각화를 통한 데이터 탐색 : 탐색적 자료분석(EDA) - 문자형, 숫자형 데이터에 따른 분석 및 접근법의 차이 - 데이터에서 인사이트 발견하기 : 쪼개보기의 중요성 [실습] - Tidyverse package를 활용한 데이터 전처리 - 분석을 위한 데이터 정리 - 인사이트 도출을 위한 데이터 시각화
3회차	분석 전 고려해야 할 사항들 및 통계분포 / 확률의 소개 [이론] - 비교대상을 동질하게 만드는 것의 중요성 (Propensity score matching:PSM) : 두개의 집단을 비교할 때 고려해야 할 사항은? - 이상치/결측값 확인 및 처리 방법 - 필수적인 확률 및 확률분포 : 정규분포, 포아송분포, 이항분포 : 특정 교통사고가 어떤 패턴을 가질까? [실습] - 반복문/조건문 활용하는 방법 - R에서의 빈도분석 - 데이터 merge하기 - Propensity score matching:PSM 실습



커리큘럼

	〈Step 2. 통계분석의 해석 및 적용〉
4회차	집단 간의 차이 비교, 연관성 파악을 위한 통계분석 [이론] - 통계분석의 기초, 가설검정/추정치/신뢰구간 - 상관관계분석(cor/partial cor) : 광고를 많이하면 매출이 커지는가? - 2개 집단의 평균비교 : 라이벌! A 제품 or B 제품 누가 더 잘할까? - 3개 이상의 평균비교와 유의사항 : A,B제품만 라이벌이냐! C반도 같이 보자. [실습] - 상관관계분석 실습 및 해석 - t.test, anova / adjusted t.test 실습 및 해석 - 변수명변경/for문으로 index변경하여 값할당하기 - rbind 또는 matrix를 통해 값 채워넣기
5회차	현업에서 활용가능한 통계방법론 [이론] - 문자형 데이터에서의 통계분석(Chisq/Fisher' Exact test) - 경향분석 (Trend test) : 저체중/정상/비만에 따라 과연 트렌드가 존재할까? - 데이터가 적거나 흔히 아는 분포를 따르지 않을 때의 통계분석(Non Parametric test) : 데이터가 너무 적거나, 우리가 아는 분포가 아닌데. - 어떤 분석이 가능할까? - 시계열 자료/공간분석/Machine Learning 소개 [실습] - 대량의 값 한번에 처리하는 방법 - 중복 값 처리 후 병합하기 - 사용자 정의 함수 사용
6회차	예측/추론을 위한 통계분석의 기초, 회귀분석 [이론] - 단순/다중 회귀분석 - 변수 선택방법 : 이 많은 변수 중, 분석을 위해서 어떤 변수를 사용해야 할까? - 상대적 중요성(표준화 회귀계수) : 변수 중에 어떤 변수가 더 중요하려나? - 로지스틱 회귀분석 (logistic regression)/ 포아송 회귀분석(poisson regression) : 사망/생존의 예측을 하고싶다! 그리고 /각 도로별 사고 건수도 예측하고 싶다. [실습] - 각 도로별 사고 건수 예측, 사망/생존의 예측 - 공부시간과 성적을 통한 회귀분석 실습 - 예제를 통한 변수 선택 실습



커리큘럼

7회차 • <Step 3. 기계학습 적용 및 함께 논문 작성하기> 기계 학습의 기초 및 활용

[이론]

- 기계학습의 기초 : 기계학습이 대체 무엇일까?
- Supervised learning / Tree, random forest : 어떤 변수가 제일 중요한지 알수 있을까?
- Unsupervised learning / Clustering : 비슷한 사람들끼리 묶고 싶은데 어떻게 할까?
- Evaluation / ROC curve : 어떤 모델이 제일 좋은지 어떻게 평가를 할까?
- parameter tuning, Deep learning 기초

[실습]

- 의료데이터를 활용한 사망/생존예측
- 암 발생 미발생 예측

8회차 • 함께 해보는 논문 작성

[이론]

- 논문 작성 시 큰 흐름 이해하기
- 예시 데이터를 활용하여 논문 작성 A to Z까지 함께 해보기
- 논문 작성 시 흔히하는 Statistical, Reporting 오류 점검하기

[실습]

- 논문 작성을 위한 구조적 글쓰기
- Data loading 및 기본 전처리, Table1 구성 이해하기
- method & result 작성하는 방법 익히기



강사소개



유재용 강사님

[약력]

- 서울시립대 통계학 석사
- 성균관대학교 삼성융합의과학원 디지털헬스 박사과정

- 삼성서울병원 임상의학연구원 (현)
- 서울대병원 응급의료연구실 통계연구원(전)
- 서울국립병원 통계연구원(전)

[교육이력]

- 기상청/충남대/뉴스젤리/분당서울대병원
내과/서울대 건축학과/ 아산병원 등

“실제 분석에 들어가면 마주하게 될 실무적인 고민들과 연결해서 조금 더 입체적으로 통계를 이해하실 수 있도록 가르쳐 드리려고 합니다. 연구를 하시는 분이나 회사에서 업무를 진행하시는 분들 모두에게 유용한 내용을 다루는 수업이 될 수 있도록 저의 실무 노하우를 아낌없이 전달드리겠습니다.”



수강환경

패스트캠퍼스 강남강의장



❖ 강의에 따라 강의장이 변경될 수 있습니다.