

### 사례연구#3 통계기반 데이터 분석

A2팀(임성현, 이순규, 한호종)

B2팀(문현진, 남원식, 오준서)

C2팀(천성한, 황윤수, 문태웅, 유준)

팀별로 다음 과제를 수행하고 결과와 R code를 포함한 사례연구보고서를 제출하시오.

1. 다음 사항을 적용하여 다중회귀분석을 실시하시오.

R의 내장 데이터셋인 state data sets 내 stat.x77 데이터셋 사용

(1) state 데이터셋을 load하고, state.x77 dataset을 데이터프레임으로 변환하고, Life Exp 변수를 Life.Exp로 HS Grad변수를 HS.Grad로 변경하시오.

(2) Life Expectancy 변수를 종속변수로 설정하고 나머지 변수를 독립변수로 설정하여 회귀분석을 실시하시오. 실시 후 결과에 대해 해석하시오.

(3) (2)번 회귀모형에서 Income, Illiteracy, Area 변수를 제외하고 회귀분석을 실시하고 결과에 대해 해석하시오.

(4) Life Expectancy 변수를 종속변수로 설정하고 HS.Grad와 Murder 변수를 예측변수(predictor variable)로 설정하여 회귀분석을 실시하시오.

(5) 전 인구의 55%가 고졸이고 살인비율이 10만명당 8명일 때 Life Expectancy 결과값을 예측하시오.

(6) (4)번에서 처럼 2개의 독립변수, 1개의 종속변수의 데이터와 fit된 회귀평면(fitted regression plane)을 3D 그래프로 시각화하시오.

2. 과거 10년간 일별 KOSPI 지수(종가기준) 데이터를 기준으로 시계열분석을 실시하시오.

데이터: <http://data.krx.co.kr/contents/MDC/MDI/mdiLoader/index.cmd?menuId=MDC0201010101#>

- 1) 추세선 확인
- 2) 4가지 시계열 자료의 변동요인을 분해
- 3) 시각화
- 4) 결과 해석

제출기한: 2021년 11월 12일(금) 09:30

제출처: 카페 내 과제제출 게시판

제출물: 사례연구보고서, R코드

제출양식: 자유(워드, 한글 등)

발표는 없음