

한국복지패널 데이터 분석: 실습



실제 공공 데이터를 이용한 분석:

한국복지패널 데이터

- 한국보건사회연구원에서 가구의 경제활동을 연구해 정책 지원에 반영할 목적으로 발간하는 조사 자료
- 전국에서 7000여 가구를 선정해 2006년부터 매년 추적 조사한 자료로, 경제활동, 생활실태, 복지욕구 등 천여 개 변수로 구성되어 있음
- 다양한 분야의 연구자와 정책 전문가들이 복지패널데이터를 활용하여 논문과 연구보고서를 발표하고 있음
- 한국복지패널데이터를 분석하면 대한민국 사람들이 어떻게 살아가고 있는지 알 수 있음
- <https://www.koweps.re.kr:442/data/book/list.do>






The screenshot shows the KOWEPS website interface. At the top, there's a navigation bar with links like HOME, CONTACT US, and ENGLISH. Below that, a main banner area features three large boxes: '2019년 제14차 한국복지패널 조사안내' (2019 14th Korea Welfare Panel Survey Guide), '제12회 한국복지패널 학술대회 논문공모' (12th Korea Welfare Panel Academic Conference Paper Submission), and '보고서 및 데이터 업로드 안내' (Report and Data Upload Guide). Below the banner, there's a '한국복지패널 공지사항' (KOWEPS Notice) section with a list of recent announcements. To the right is a 'LOG-IN' section. At the bottom, there are several categorized sections: '조사설계' (Survey Design), '데이터다운' (Data Download - circled in red), '설문지다운' (Questionnaire Download), and '연구자료' (Research Materials). The '데이터다운' section has a red circle around it, indicating the area of interest for data analysis.

<https://www.koweps.re.kr:442/main.do>

Data



(2019년 14차 한국복지패널조사) 데이터(beta1)_spss

 14차 데이터파일설명(beta1)	2020-03-30 오전 1...	텍스트 문서	1KB
 koweps_h14_2019_beta1.sav	2020-04-06 오후 8...	SAV 파일	8,710KB
 koweps_hpwc14_2019_beta1.sav	2020-04-06 오후 8...	SAV 파일	15,742KB
 koweps_p14_2019_beta1.sav	2020-04-06 오후 8...	SAV 파일	3,849KB
 koweps_wc14_2019_beta1.sav	2020-04-06 오후 8...	SAV 파일	490KB

<데이터 파일 설명 - Beta1 Version>

*****< 데이터 명 >*****

1. 가구용데이터(SAS, SPSS, STATA):koweps_h14_2019_beta1
2. 가구원 용데이터(SAS, SPSS, STATA):koweps_p14_2019_beta1
3. 복지인식설문용데이터(SAS, SPSS, STATA):koweps_wc14_2019_beta1
4. 가구용, 가구원 용, 복지인식설문용 머지데이터(SAS, SPSS, STATA):koweps_hpwc14_2019_beta1

*****< Beta1 Version >*****

※ 2020년 3월 31일자로 수정한 버전


설문지 & Codingbook





14차 웨이브





- (유형1)14차년도 조사표_가구용.pdf
- (유형2)14차년도 조사표_신규가구용.pdf
- (유형3)14차년도 조사표_가구원용.pdf
- (유형4)14차년도 조사표_신규가구원용.pdf
- (유형5)14차년도 조사표_부가조사용(복지인식).pdf

 (2019년 14차 한국복지패널조사) 조사설계서(beta1)

 (2019년 14차 한국복지패널조사) 조사설계서-가구용(beta1)

 (2019년 14차 한국복지패널조사) 조사설계서-가구원용(beta1)

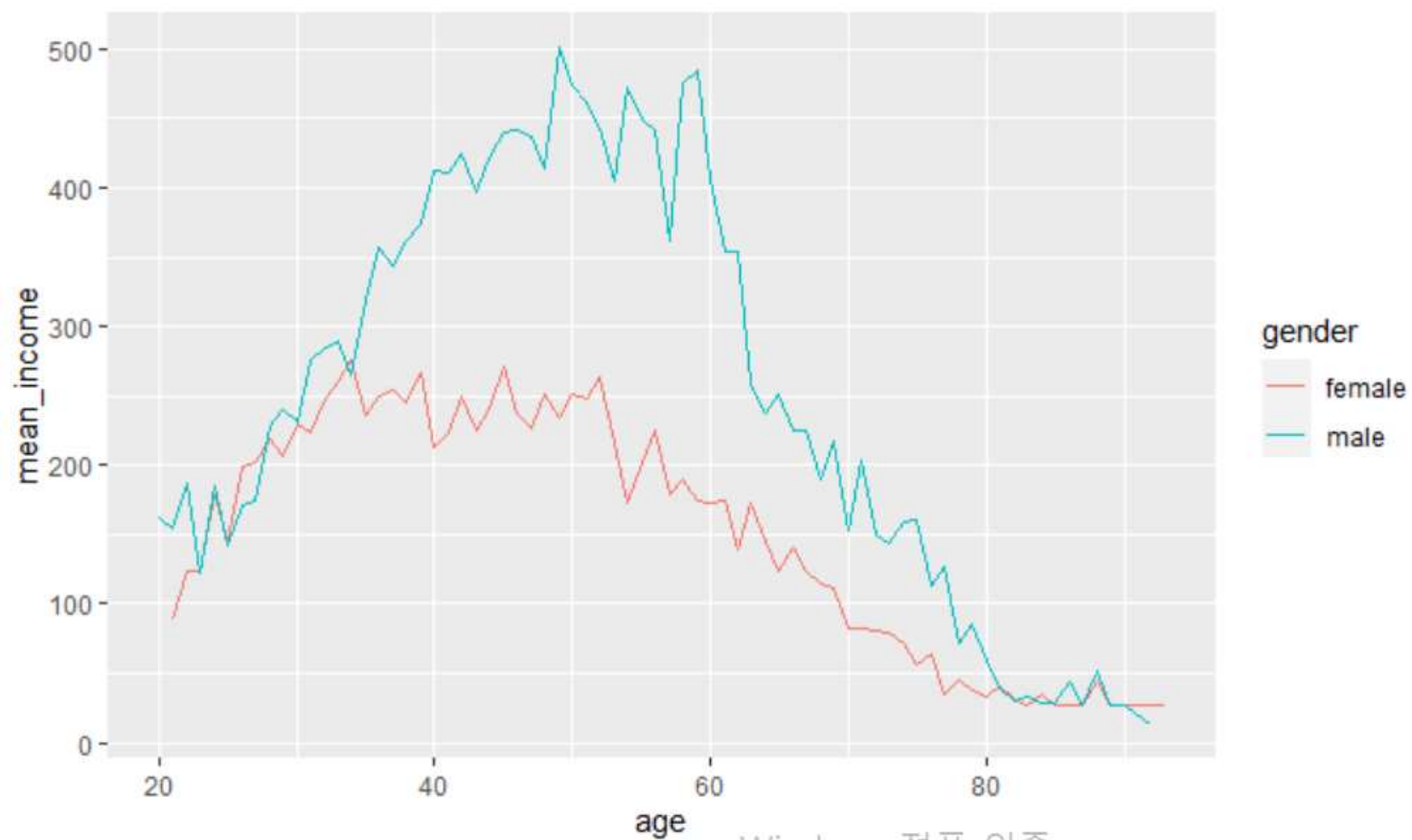
 (2019년 14차 한국복지패널조사) 조사설계서-복지인식용(beta1)

 14차 머지데이터_변수명

HW#III-1



1) 나이, 성별에 따른 월급 평균의 선그래프

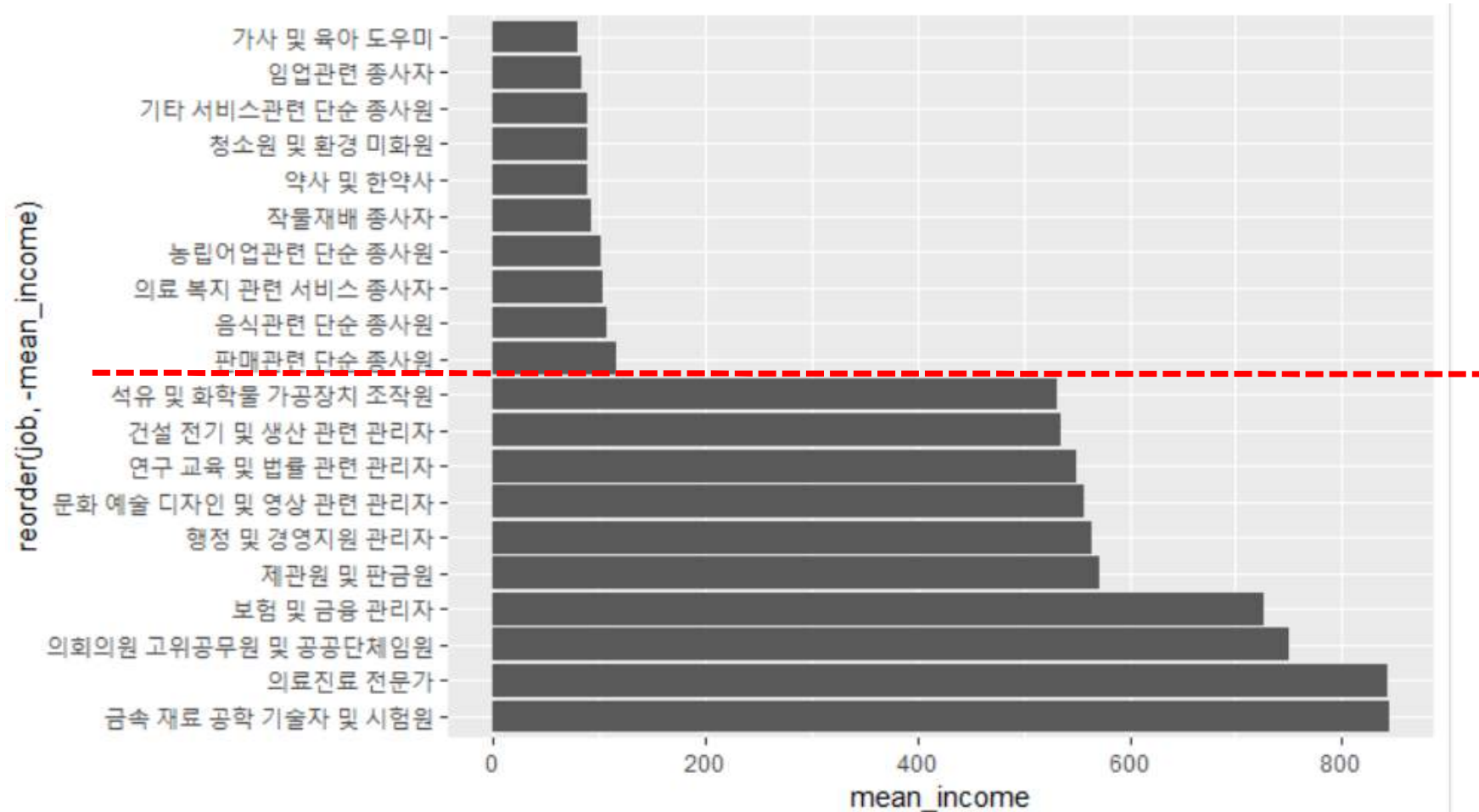


HW#III-2



2) 직업별로 월급 평균을 알아내고, 평균 월급 상위 10개 직업 및 하위 10개 직업을 막대 그래프로 나타내시오.

▪ (참고) 2015



HW#III Tips

1. 변수 검토 및 전처리

⇒ welfare 데이터 마련

3. Job code 검토

=> list_job 데이터 마련

4. Job 변수를 welfare에 결합(left_join)

5. 직업별 월급 평균표 생성

6. 월급 내림차순 정렬

7. 상위 10개 추출

8. 하위 10개 추출

9. 상위 10개+하위10개 bind_rows

10. 그래프 생성

```
1  install.packages("foreign")
2  install.packages("readxl")
3  library(foreign)
4  library(dplyr)
5  library(ggplot2)
6  library(readxl)
7
8  raw_welfare=read.spss(file="koweps_hpwc14_2019_beta1.sav",
9                        to.data.frame=T)
10
11
12  welfare=raw_welfare
13  head(welfare)
14  welfare=rename(welfare,
15                gender=h14_g3,          #gender
16                birth=h14_g4,          #birth year
17                marriage=h14_g10,      #marrital status
18                religion=h14_g11,      #religion
19                income=p1402_8aq1,     #일한달의 평균 월급
20                code_job=h14_eco9,     #job code
21                code_region=h14_reg7) #region code (7)
22
23  welfare=select(welfare, gender, birth, marriage, religion, income, code_job, code_region)

# 직업별 임금 분석
list_job=read_excel("Job_Code_2020.xlsx", col_names=T, sheet=4)
head(list_job)
str(list_job)
list_job$code_job=as.numeric(list_job$job_code)

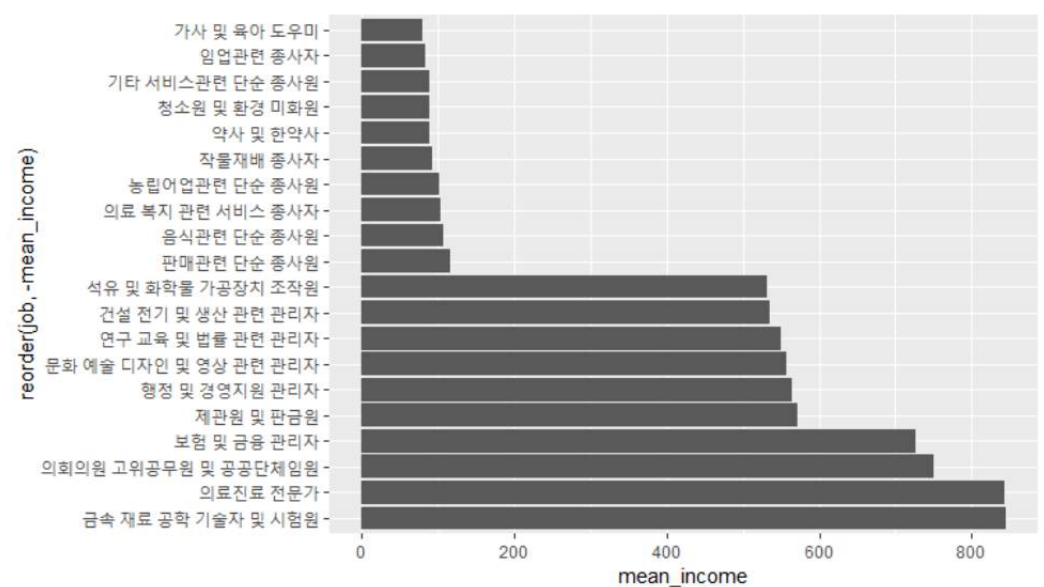
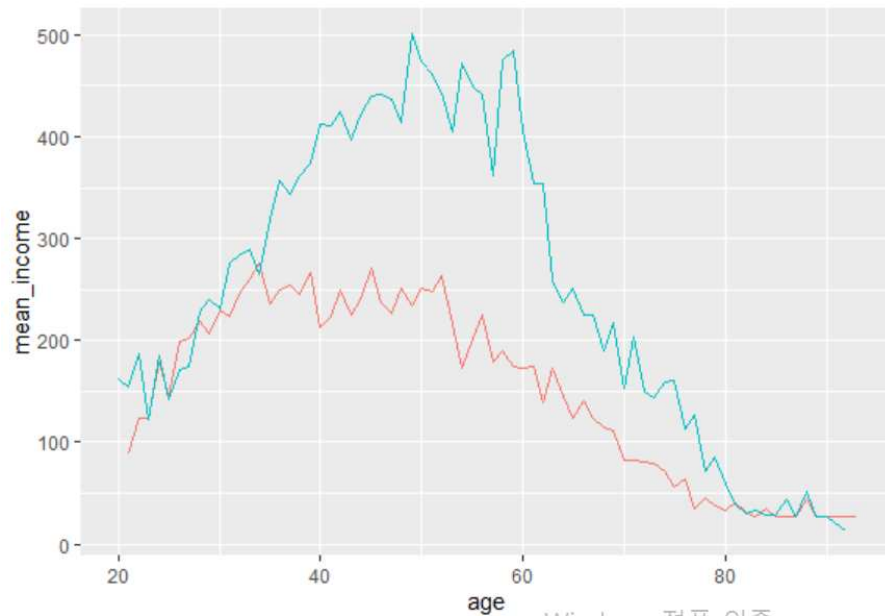
#join job name by job code_job
welfare=left_join(welfare, list_job, id="job_code")

# join결과 확인
welfare %>% filter(!is.na(job_code) & !is.na(job_name)) %>%
  select(job_code, job_name) %>%
  head(10)
```

HW#III



- 1) 나이, 성별에 따른 월급 평균의 선그래프
- 2) 직업별로 월급 평균을 알아내고, 평균 월급 상위 10개 직업 및 하위 10개 직업을 막대 그래프로 나타내시오.



DUE: 5월4일 12:00까지 (e-class 과제함에 제출)