EVS Data Analysis (ZA7500)

Kangrui Liu

2025-03-23

Introduction

This report explores the 2017 European Values Study (EVS), focusing on two variables measuring attitudes toward gender roles (v72) and immigration (v80). The analysis includes descriptive statistics, visualization of how these attitudes change with age, and regression modeling incorporating age, age squared, sex, and education.

Data Preparation

```
evs <- read_dta("data/ZA7500_v5-0-0.dta")

evs_clean <- evs %>%
  dplyr::select(v72, v80, age = v242, education = v243_r, sex = v225, country = c_abrv) %>%
  mutate(
    age_sq = age^2
  ) %>%
  filter(v72 >=0, v80>=0, age >=0, education>=0, sex >=0, country>=0)

# save(evs_clean, file = "data/evs_clean.RData")
```

Descriptive Statistics

Continuous Variables

```
evs_clean %>%
  dplyr::select(age, v72, v80) %>%
  summary()
```

```
##       age             v72             v80
##  Min.   : 0.00   Min.   :1.000   Min.   :1.000
##  1st Qu.:17.00   1st Qu.:2.000   1st Qu.:1.000
##  Median :19.00   Median :3.000   Median :2.000
##  Mean   :20.53   Mean   :2.719   Mean   :2.322
##  3rd Qu.:23.00   3rd Qu.:3.000   3rd Qu.:3.000
##  Max.   :70.00   Max.   :4.000   Max.   :5.000
```

Categorical Variables

```r
evs_clean %>%
  count(sex) %>%
  kable(caption = "Frequency Table for Sex")
```

Table 1: Frequency Table for Sex

sex	n
1	24044
2	30369

```r
evs_clean %>%
  count(education) %>%
  kable(caption = "Frequency Table for Education")
```

Table 2: Frequency Table for Education

| education	n
1	10650
2	24797
3	18890
66	76

Visualizations

v72 by Age

```r
ggplot(evs_clean, aes(x = cut_width(age, width=10), y = v72)) +
  geom_boxplot(fill = "lightblue") +
  labs(title = "v72 by Age Group", x = "Age Group", y = "v72")
```

v72 by Age Group
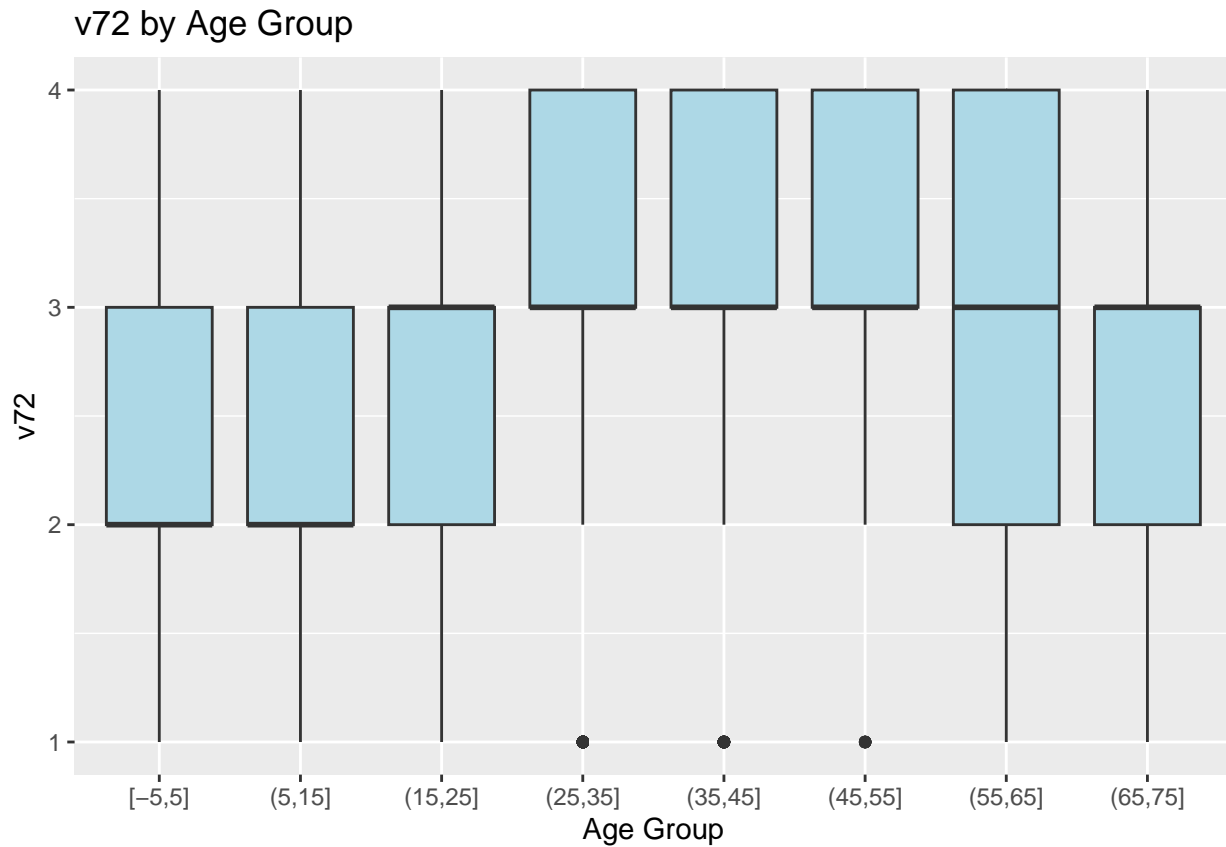
Interpretation: The boxplot illustrates how attitudes towards gender roles (v72) vary across different age groups. Middle-aged groups (approximately 25–65 years old) tend to hold slightly higher median scores, suggesting relatively distinct attitudes toward gender roles compared to the younger and older groups.

v80 by Age

```
ggplot(evs_clean, aes(x = cut_width(age, width=10), y = v80)) +
  geom_boxplot(fill = "lightblue") +
  labs(title = "v80 by Age Group", x = "Age Group", y = "v80")
```

v80 by Age Group
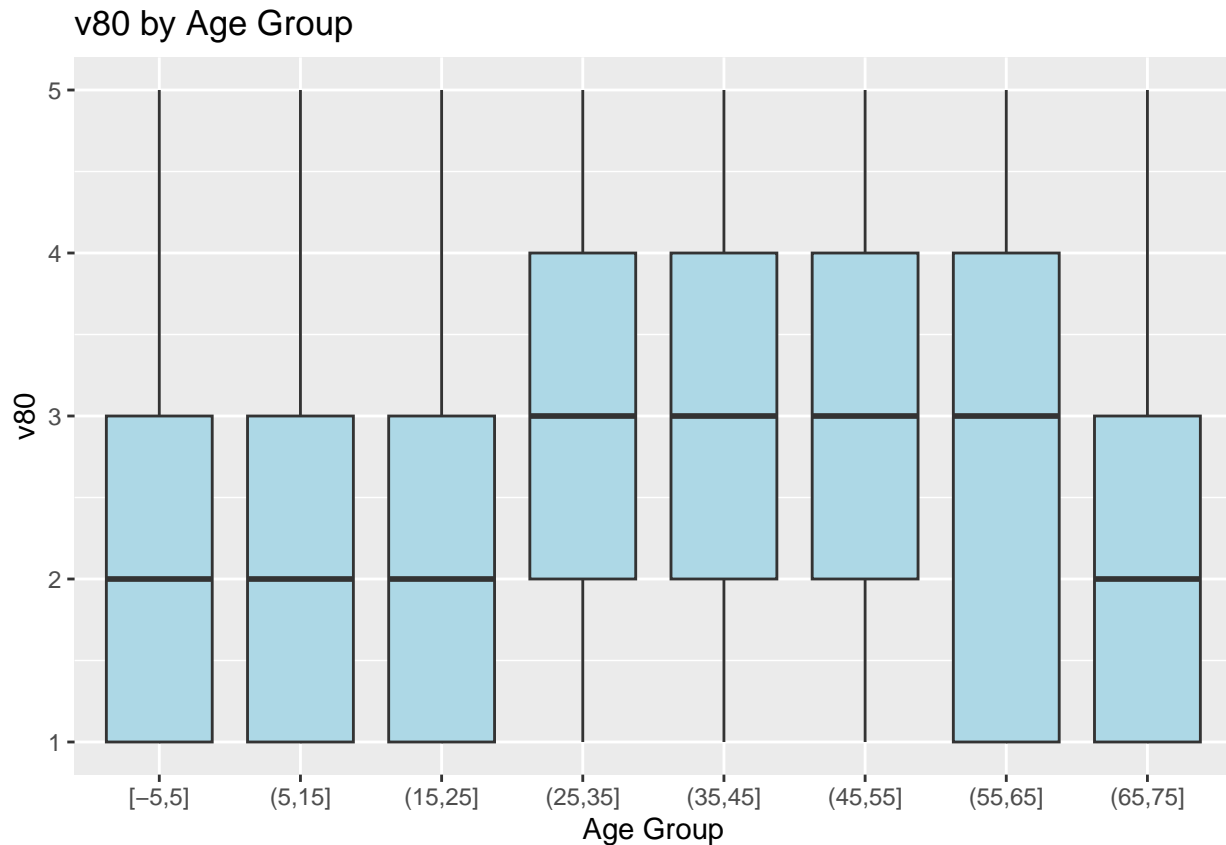
Interpretation: This plot displays how attitudes towards immigration (v80) differ across age groups. Similar to v72, middle-aged individuals appear to have higher median values, indicating that generational factors might influence attitudes towards immigration policies or perceptions.

Regression Analysis

We perform two OLS regression models using attitudes towards gender roles (v72) and immigration (v80) as dependent variables, incorporating age, age squared, sex, and education as predictors.

```
mod1 <- lm(v72 ~ age + age_sq + sex + education, data = evs_clean)
mod2 <- lm(v80 ~ age + age_sq + sex + education, data = evs_clean)

screenreg(list("v72" = mod1, "v80" = mod2))
```

```
##
## =====================================
##                 v72           v80
## -------------------------------------
## (Intercept)     1.68 ***      1.34 ***
##                (0.03)        (0.04)
## age             0.06 ***      0.06 ***
##                (0.00)        (0.00)
## age_sq         -0.00 ***     -0.00 ***
##                (0.00)        (0.00)
```

4

```
## sex               0.07 ***      -0.02
##                   (0.01)        (0.01)
## education         0.01 ***       0.01 ***
##                   (0.00)        (0.00)
## -------------------------------------
## R^2               0.04           0.03
## Adj. R^2          0.04           0.03
## Num. obs.     54413          54413
## =====================================
## *** p < 0.001; ** p < 0.01; * p < 0.05
```

Interpretation: The regression tables show that these predictors are statistically significant to two dependent variables while the age squared is not that significant to v80.

Summary

This report presented a clear overview of attitudes towards gender roles and immigration from the EVS data. Descriptive analyses highlighted demographic characteristics, visualizations demonstrated how these attitudes vary by age, and regression analyses quantified the influence of key variables such as age, sex, and education on these attitudes.