Week 2 Report

Diana Hilleshein

1/24/2022

General information on the data

Data was downloaded from: https://www.europeansocialsurvey.org/download.html?file=ESS9e03_1&y=2018.
Codebook fro the data can be found by the following link: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/viewer.html?pdfurl=https%3A%2F%2Fwww.europeansocialsurvey.org%2Fdocs%2Fround9%2Fsurvey%2FESS9_appendix_a7_e03_1.pdf&clen=2111156.

The European Social Survey (ESS) is an academically-driven multi-country survey, which has been administered in over 35 countries to date. Its three aims are, firstly – to monitor and interpret changing public attitudes and values within Europe and to investigate how they interact with Europe's changing institutions, secondly – to advance and consolidate improved methods of cross-national survey measurement in Europe and beyond, and thirdly – to develop a series of European social indicators, including attitudinal indicators. For the course we will use round 9 ESS.

Variables of the datases are separated for the following groups:
Country
Weights
Media and social trust
Politics
Subjective well-being, social exclusion, religion, national and ethnic identity
Timing of life
Gender, Year of birth and Household grid
Socio-demographics
Justice and Fairness
Human values
Administrative variables
Sample Design variables
User defined variables

"'{r include=FALSE} setwd("C:/R scripts/DWA2022")

load/attach package

library(haven)

import data

data <- read_sav("C:/R scripts/DWA2022/ESS9e03_1.sav")

#View(data)

save(data,file="data.csv")

```
## Part 1. Explore the European Social Survey (ESS) data.
I am interested in the following self-made groups of variables:
*Involovemt in politics*
1) *polintr* -How interested in politics (scale from 1 to 4, 1= Very interested, 1 = Not at all, 4 = Ver
2) *psppipla* -Political system allows people to have influence on politics (scale from 1 to 5, 1 = Not
*Attitude to immigrants*
3) *imwbcnt* -Immigrants make country worse or better place to live (scale from 0 to 10, 0 = Worse place
4) *imbgeco* -Immigration bad or good for country's economy  (scale from 0 to 10, 0 = Bad for the econor
5) *imueclt* -Country's cultural life undermined or enriched by immigrants (scale from 0 to 10, 0 = Cult
6) *imsmetn* -Allow many/few immigrants of same race/ethnic group as majority (scale from 1 to 4, 1 = Al
7) *imdfetn* -Allow many/few immigrants of different race/ethnic group from majority (scale from 1 to 4
8) *impcntr* -Allow many/few immigrants from poorer countries outside Europe (scale from 1 to 4, 1 = All
*General information on a participants*
9) *stflife* -How satisfied with life as a whole (from 0 to 10, 0 = extremely dissatisfied, 10 = extreme
10) *gndr*-Gender (1 = Male, 2 = Female)
11) *agea* -Age of respondent, calculated (integer)

Comments on the variables:
- there is just one continuous variable that is agea, while gender is binary. The rest of the variables

I  will display descriptive statistics by variable groups:
*General information on a participants*

'''{r}
#load necessary libraries
library(ggplot2)
library(GGally)
library(haven)
library(tidyverse)
library(corrplot)
library(tidyverse)
library(dplyr)

#make one dataframe that contains all the cat variables. Dont do the same for age and gender since there
data <- as.data.frame(data)
interested_var_cat <- select(data, "stflife", "psppipla", "polintr", "imwbcnt", "imbgeco", "imueclt", "i
data$gndr = as.factor(data$gndr)
interested_var <- select(data, "stflife", "psppipla", "polintr", "imwbcnt", "imbgeco", "imueclt", "imsme
# "gndr"
# agea

gather(interested_var_cat) %>% ggplot(aes(value)) + geom_bar(col ="dark red", fill = "dark red", alpha =
```

Bar plot indicates that: - there are people who thinks that political system doesnt allow people to have influence on politics or allows on issuficient level (*psppipla*). While participants are rather interested in politics that not interested (*polintr*); - in majority people are satisfied with life as whole (*stflife*); - participants tend to be neurtal about their opinion regarding immigrants influence on a country in general as well and in economic and cultural sence (*imwbcnt*, *imbgeco*, *imueclt*). Hovewer there are sligtly more poeple who thiok of positive influence of immigrants. - participants tend to be more accepteble towards immigrand of similar ethnicity (*imsmetn* vs *imdfetn*), and tend to accept the least immigrants from poorer countries outside Europe (*impcntr*).

The level of acceptence can be pictured as following: same athnicity > different ethnicity > from poorer countries out of Europe.

Next, disctiptive ststistics for age and gender. As we can see there are slightly more women who participates in the survey.
Average age of participants is 51 years. People from 15 to 90 years old took part in the survey. 50% of all ages are from approximetely 37 to 67 years.

```
#rstudio was complaining that gndr was a list, so I unlisted it
data$gndr <- unlist(data$gndr)
#summary fo rgndr
summary(data$gndr)
#some vizualization to make the eyes happy
col <- c(rep("dark blue", 1), rep("dark red", 1))
ggplot(data = data, aes(data$gndr)) + geom_histogram(col = col, fill = col, alpha = 0.7, bins = "free",
#descriptives for agea
summary(data$agea)
#vizualization to make the eyes happy again
ggplot(data = data, aes(agea)) + geom_histogram(breaks = seq(15,90,1), col = "dark orange", fill = "dar
col2 = rep("dark orange", alpha = 0.7)

ggplot(data = data, aes(agea, y = " ")) + geom_boxplot(fill = "dark orange", alpha = 0.7)
```

##Part 2. Copy a small text file data (week2data) from Moodle

```
#load the data for part 2 ex
library(readxl)
part2_data <- read_excel("C:/R scripts/DWA2022/part_two_data.xlsx")

View(part2_data) #good!
save(part2_data,file="part2_data.csv")
```

Task 1: Add variable labels: id = "ID number", agea = "Age of respondent", q1 = "How often you read newspaper?", q2 ="Government should reduce differences in income levels?"

```
#install.packages("expss")
library(expss)
part2_data = apply_labels(part2_data, id = "ID number",
                          agea = "Age of respondent",
                          q1 = "How often you read newspaper?",
                          q2 = "Government should reduce differences in income levels?")
View(part2_data) #we can see labelles in the dataset
get_label(part2_data$id) #also worked
```

Task 2: Add value labels to variables q1 and q2
Q1: 1 = "Never", 2 = "Only occasionally", 3 = "A few times a week", 4 = "Most days", 5 = "Every day, 8 ="Don't know", 9 = "No answer"
Q2: 1 = "Agree strongly", 2 = "Agree", 3 = "Neither agree nor disagree", 4 = "Disagree", 5 = "Disagree strongly", -1 = "missing"

```
part2_data$q1 <- ordered(part2_data$q1,
                         levels = c(1, 2, 3, 4, 5, 8, 9),
                         labels = c("Never", "Only occasionally", "A few times a week", "Most days", "E
```

```
get_labels(part2_data$q1) #worked!
View(part2_data) #I can see the labels fro q1 in the table

part2_data$q2 <- ordered(part2_data$q2,
                         levels = c(1, 2, 3, 4, 5, -1),
                         labels = c("Agree strongly", "Agree", "Neither agree nor disagree", "Disagree"

View(part2_data) # can see the lables fro q1 and q2 in the table
```