Participant Data Analysis

Boleslav Khodakov

22 June, 2023

Contents

1 Introduction

This script produces a quick analysis of Participants. We calculate the gender distribution, mean age, Standard Deviation of the age, age range, and how many people recommend using their data. These statistics are about all participants. So, 72 participants for half 1, 82 participants for half 2.

We use the following packages:

```
library(haven)     # Use read_sav function
library(knitr) # Get markdown file
library(tinytex) # Use TeX environment
library(rticles) # Use CTeX documents template
```

2 Round 1

First, we load the two Prolific data files (for half 1 and half 2). We combine them into one file.

```
# load the data files from Prolific
participants_first_half <- read.csv("prolific_export_final_eval_german_first_half_v1_2023_06_22_anonym.
participants_second_half <- read.csv("prolific_export_final_eval_german_second_half_v1_2023_06_22_anony

# combine the data files into one
participants_data <- rbind(participants_first_half, participants_second_half)
```

2.1 Gender distribution

We calculate the amount/percentage of participants who identify as one of the available options: Man (including Trans Male/Trans Man), Woman (including Trans Female/Trans Woman), Non-binary. **Please note: Due to how cat works, the printed results are in the codechunk itself. Look at the auto-generated comments beneath each cat command for the printed statistics.**

```
# Count the amount of participants who identify as ...
count_male <- length(which(participants_data$Gender ==
                             "Man (including Trans Male/Trans Man)"))
count_female <- length(which(participants_data$Gender
                             == "Woman (including Trans Female/Trans Woman)"))
count_non_binary <- length(which(participants_data$Gender
                                 == "Non-binary (would like to give more detail)"))


cat("Amount of participants who identified as male:", count_male,
    "Expressed as percent: ", (count_male/nrow(participants_data)*100),"%\n")
## Amount of participants who identified as male: 82 Expressed as percent:  53.24675 %
cat("Amount of participants who identified as female:", count_female,
    "Expressed as percent: ", (count_female/nrow(participants_data)*100),"%\n")
## Amount of participants who identified as female: 69 Expressed as percent:  44.80519 %
cat("Amount of participants who identified as non-binary:", count_non_binary,
    "Expressed as percent: ", (count_non_binary/nrow(participants_data)*100),"%\n")
## Amount of participants who identified as non-binary: 3 Expressed as percent:  1.948052 %
```

2.2 Age

We calculate the age range, mean age, Standard Deviation of age. **Please note: Due to how cat works, the printed results are in the codechunk itself. Look at the auto-generated comments beneath each cat command for the printed statistics.**

```
mean_age <- mean(participants_data$Age)
SD_age <- sd(participants_data$Age)
min_age <- min(participants_data$Age)
max_age <- max(participants_data$Age)

cat("Age range of participants:", min_age,
    "-", max_age,"\n")
## Age range of participants: 19 - 69
cat("Mean age (rounded): ", round(mean_age),"\n")
## Mean age (rounded):  31
cat("Standard Deviation of age (rounded):", round(SD_age),"\n")
## Standard Deviation of age (rounded): 10
```

2.3 Recommended Data usage

We calculate the amount of people who recommended using their data **Please note: Due to how cat works, the printed results are in the codechunk itself. Look at the auto-generated comments beneath each cat command for the printed statistics.**

```r
# read .sav files for both halves and calculate how many people recommended using their data
data01 <- read_sav("Final_ASA_German_Summative_First_Half_v1_2023_06_22_anonym.sav")
data02 <- read_sav("Final_ASA_German_Summative_Second_Half_v1_2023_06_22_anonym.sav")
recommended1 <- data01[data01$Use_Data == '2',]
recommended2 <- data02[data02$Use_Data == '2',]

cat("Participants who recommended using their data:", nrow(recommended1) + nrow(recommended2), sep=" ")
## Participants who recommended using their data: 141
```