Participant Data Analysis

Boleslav Khodakov

13 June, 2023

Contents

1 Introduction

This script produces a quick analysis of Participants for all translation rounds. We calculate the gender distribution, mean age, Standard Deviation of the age, age range.

We use the following packages:

```
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_german_first_half_anonym.csv")
participants_second_half <- read.csv("prolific_export_german_second_half_anonym.csv")

# 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: 30 Expressed as percent:  50 %
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: 27 Expressed as percent:  45 %
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:  5 %
```

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 - 73
cat("Mean age (rounded): ", round(mean_age),"\n")
## Mean age (rounded):  35
cat("Standard Deviation of age (rounded):", round(SD_age),"\n")
## Standard Deviation of age (rounded): 13
```

3 Round 2

First, we load the Prolific data file

```
# load the data files from Prolific
participants_data <- read.csv("prolific_export_german_anonym.csv")
```

3.1 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.**

```r
# 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: 15 Expressed as percent:  50 %
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: 13 Expressed as percent:  43.33333 %
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: 2 Expressed as percent:  6.666667 %
```

3.2 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.**

```r

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: 22 - 70
cat("Mean age (rounded):", round(mean_age),"\n")
## Mean age (rounded): 35
cat("Standard Deviation of age (rounded):", round(SD_age),"\n")
## Standard Deviation of age (rounded): 12
```

4 Round 3

First, we load the Prolific data file

```r
# load the data files from Prolific
participants_data <- read.csv("prolific_export_german_round3_anonym.csv")
```

4.1 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.**

```r
# 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: 15 Expressed as percent:  50 %
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: 15 Expressed as percent:  50 %
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: 0 Expressed as percent:  0 %
```

4.2 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.**

```r
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: 21 - 46
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): 6
```