Project Data Proof

Emily Han and Junyi Hui

2025-02-12

```r
suppressMessages({
  library(dplyr)
  library(tidyr)
  library(reshape2)
  library(rstatix)
  library(ggplot2)
  library(ggcorrplot)
  library(lessR)
})
```

Load the data for your replication project into R. Produce the following:

1. A histogram of the dependent variable
2. A correlation matrix for the DV and IVs that the original authors included in the model you are replicating
3. A visual or tabular depiction of the missingness in the data from part (2); see p. 251-255 of the text.

Data load

```r
s2 <- read.csv("data/Study2.csv")
```

Data prep (Extracted from the published code)

```r
# Recoding the variables

s2$home= factor(s2$home, levels = c(1, 2,3,4), labels = c("home-answered","away-notanswered","home-retu
s2$consent= factor(s2$consent, levels = c(1, 2), labels = c("yes","no"))
s2$etnic_ascr= factor(s2$etnic_ascr, levels = c(1, 2), labels = c("slovak","roma"))
s2$etnicita= factor(s2$etnicita, levels = c(1, 2,3), labels = c("slovak","roma","other"))
s2$age <- as.numeric(s2$age)
s2 <- lessR::recode(age, new_vars="roky", old=9:72, new=1996:1933, data = s2)
s2$years <- 2019 - s2$roky
s2$gender <- factor(s2$gender, levels = c(1, 2), labels = c("Male","Female"))

# Filter the observations

s2_consent <- dplyr::filter(s2, consent=="yes")
s2_consent <- dplyr::filter(s2_consent, etnic_ascr != "NA")
s2_consent <- dplyr::filter(s2_consent, eurofondy_iv<5 & skolka_iv<5 & skolka_agree<5 & skolka_vote<5 &
```

```
# Add a Unique Identifier
s2_consent <- mutate(s2_consent, ID = row_number())

# Reshape the Data (Long Format)
s2_mains <- melt(s2_consent,
        # ID variables - all the variables to keep but not split apart on
    id.vars=c("etnic_ascr", "ID","eurofondy_iv","skolka_iv" ),
        # The source columns
    measure.vars=c("control", "suma", "praca","potreba","pila"),
        # Name of the destination column that will identify the original
        # column that the measurement came from
    variable.name="condition",
    value.name="measurement"
)

# Remove a Specific Condition
s2_mains <- filter(s2_mains, condition!="pila")

# Compute Summary Statistics
s2_dvs <- s2_mains %>%
  group_by(etnic_ascr,condition)%>%
  dplyr::select(-ID:-skolka_iv) %>%
  get_summary_stats(type = c("mean_sd"))%>%
  dplyr::select(-variable)%>%
  dplyr::mutate(condition = dplyr::recode(condition,
    "control" = "Control",
    "suma" = "Equality",
    "praca" = "Reciprocity",
    "potreba" = "Need"))%>%
  dplyr::mutate(sd = round(sd,2))
```

Note: s2_main is the long pivot data that can be used for running the model

```
# Recoding the condition names

 s2_mains <- s2_mains %>%
  mutate(condition = dplyr::recode(condition,
    "control" = "Control",
    "suma" = "Equality",
    "praca" = "Reciprocity",
    "potreba" = "Need"))
```

1. A histogram of the dependent variable

```
ggplot(s2_dvs, aes(x = condition, y = mean, fill = etnic_ascr)) +
  geom_bar(stat = "identity", position = "dodge", color = "black") +
  scale_fill_manual(values = c("lightblue", "lightcoral")) +
  labs(title = "Support for Housing Project by Condition and Ethnicity",
       x = "Condition",
       y = "Mean Score of Support for Housing Project",
```
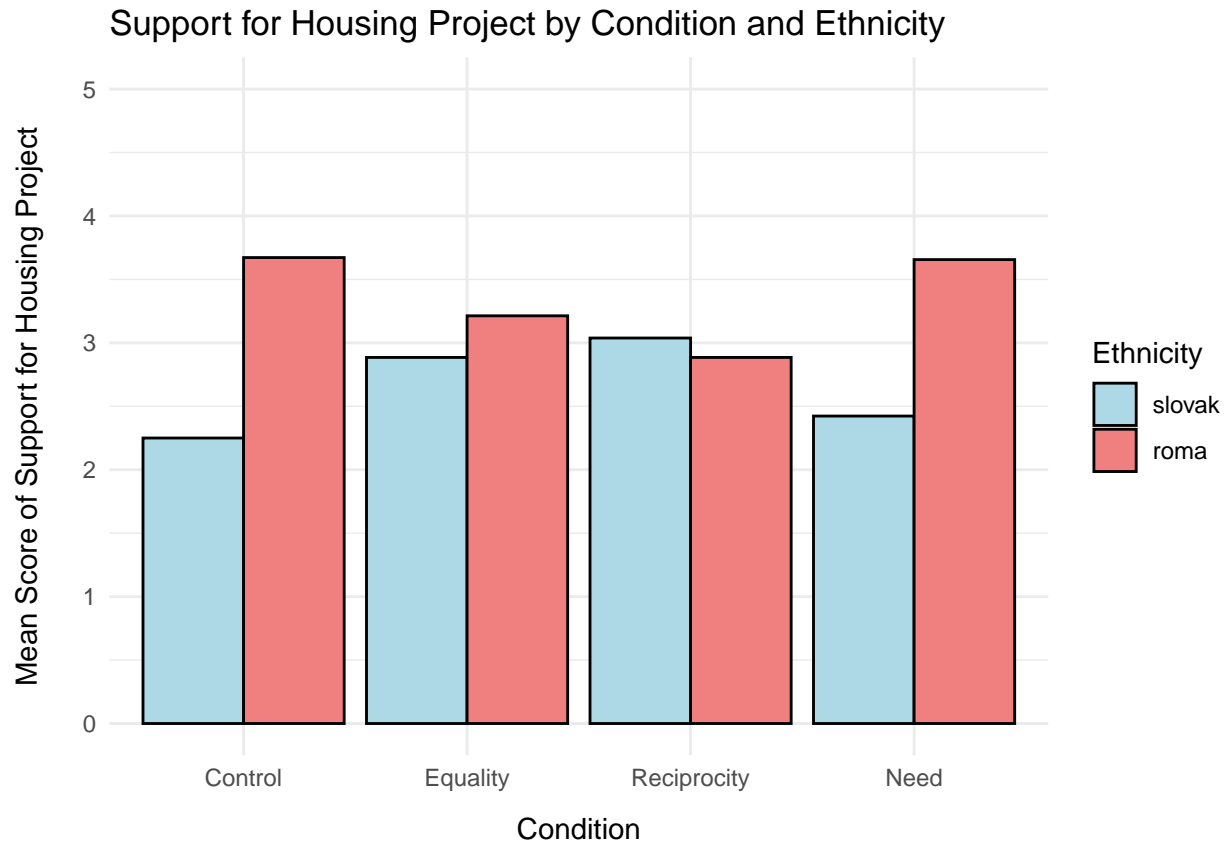
```
        fill = "Ethnicity") +
  ylim(0,5) +
  theme_minimal() +
  theme(
    axis.title.x = element_text(margin = margin(t = 10)), # Adjust X-axis title spacing
    axis.title.y = element_text(margin = margin(r = 15))  # Adjust Y-axis title spacing
  )
```

Grouped Bar Plot for the DVs

Support for Housing Project by Condition and Ethnicity

Histogram for the DVs

```
ggplot(subset(s2_dvs, etnic_ascr == "slovak"), aes(x = condition, y = mean)) +
  geom_bar(stat = "identity", fill = "lightblue", color = 'black', width = 0.6) +
  labs(title = "Support for Housing Project across Condition for Slovak",
       x = "Condition",
       y = "Mean Score of Support for Housing Project") +
  ylim(0, 5) +
  theme_minimal() +
  theme(
    axis.title.x = element_text(margin = margin(t = 10)), # Adjust X-axis title spacing
```
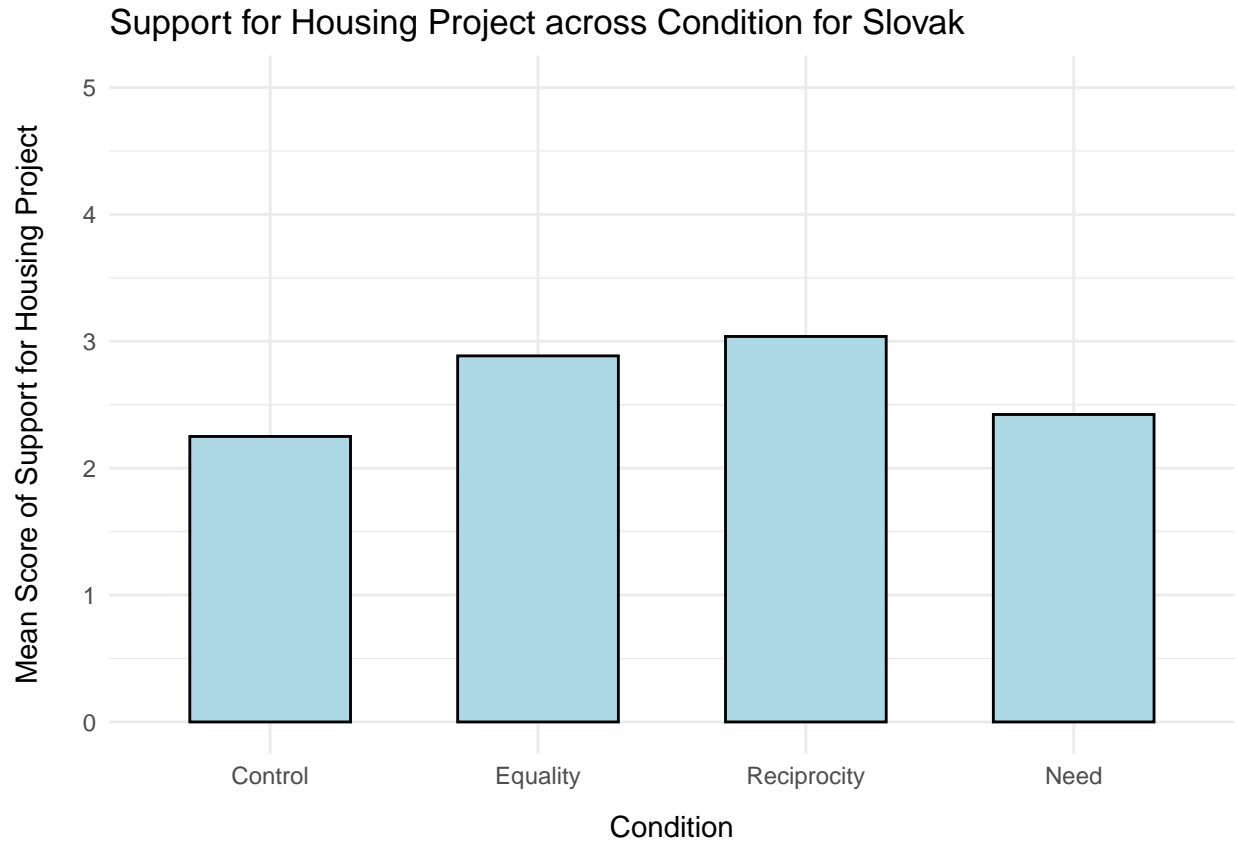
```
    axis.title.y = element_text(margin = margin(r = 15))  # Adjust Y-axis title spacing
  )
```

Plot for Slovak

Support for Housing Project across Condition for Slovak


```
ggplot(subset(s2_dvs, etnic_ascr == "roma"), aes(x = condition, y = mean)) +
  geom_bar(stat = "identity", fill = "lightcoral", color = 'black', width = 0.6) +
  labs(title = "Support for Housing Project across Condition for Roma",
       x = "Condition",
       y = "Mean Score of Support for Housing Project") +
  ylim(0, 5) +
  theme_minimal() +
  theme(
    axis.title.x = element_text(margin = margin(t = 10)), # Adjust X-axis title spacing
    axis.title.y = element_text(margin = margin(r = 15))  # Adjust Y-axis title spacing
  )
```

Plot for Roma

Support for Housing Project across Condition for Roma
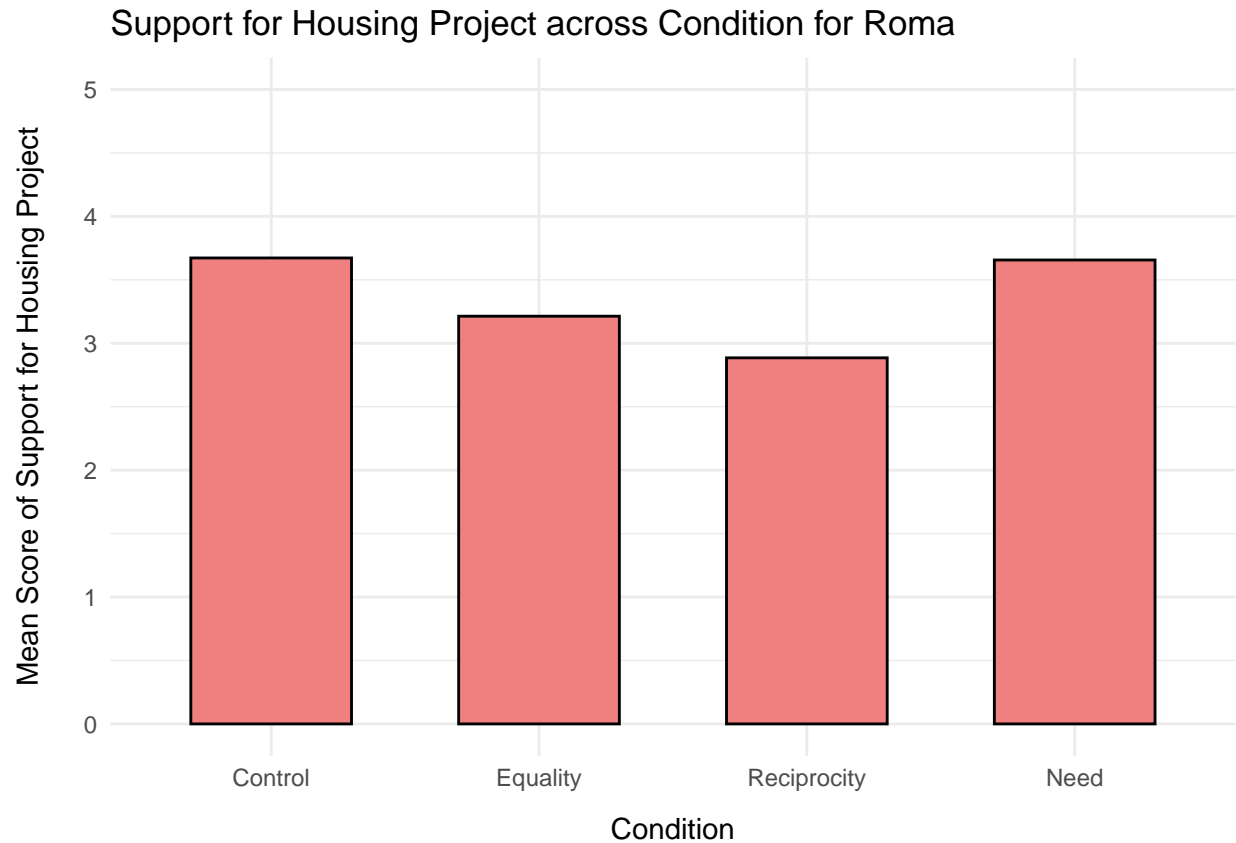
2. Correlation matrix for the DV and IVs

```r
s2_wide <- s2_mains %>%
  pivot_wider(names_from = condition, values_from = measurement)

s2_slovak <- s2_mains %>%
  filter(etnic_ascr == "slovak") %>%
  pivot_wider(names_from = condition, values_from = measurement)

s2_roma <- s2_mains %>%
  filter(etnic_ascr == "roma") %>%
  pivot_wider(names_from = condition, values_from = measurement)
```

```r
# Correlation Matrix Across Different Condition for both Solvak and Roma
cor_matrix <- cor(s2_wide %>% select(Control, Equality, Reciprocity, Need), use = "pairwise.complete.obs

# Correlation Matrix Across Different Condition for Solvak
cor_matrix_slovak <- cor(s2_slovak %>% select(Control, Equality, Reciprocity, Need), use = "pairwise.con

# Correlation Matrix Across Different Condition for Roma
cor_matrix_roma <- cor(s2_roma %>% select(Control, Equality, Reciprocity, Need), use = "pairwise.complet
```
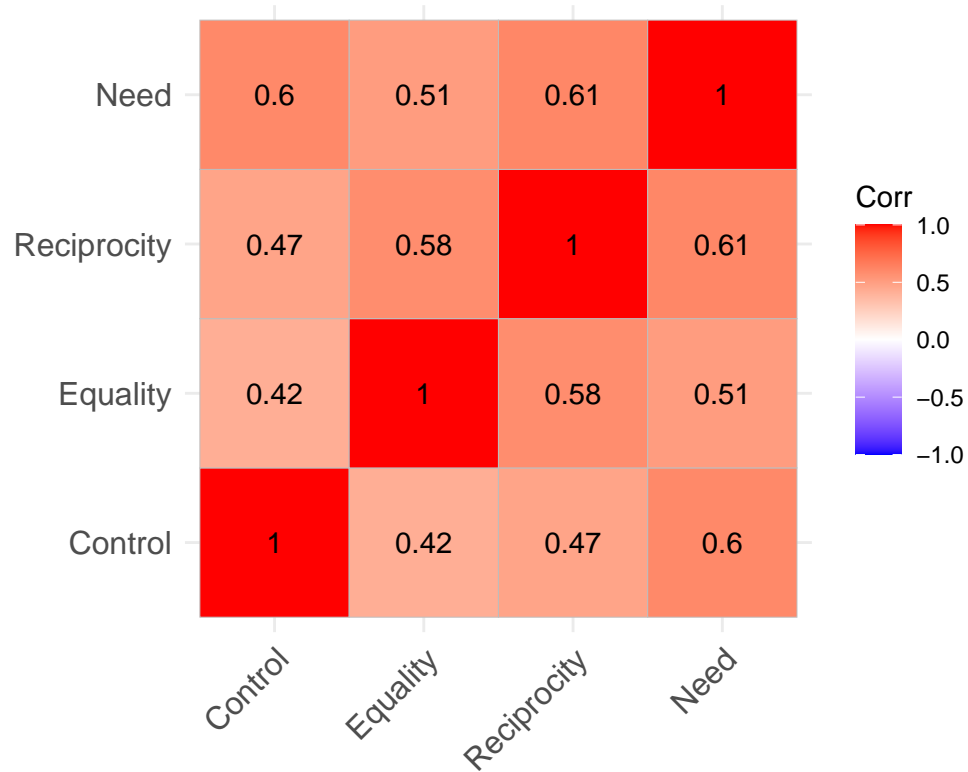
```
ggcorrplot(cor_matrix, lab = TRUE) +
  ggtitle("Correlation Matrix - All Participants") +
  theme(plot.title = element_text(size = 14, margin = margin(b = 10)))
```

Correlation Matrix – All Participants
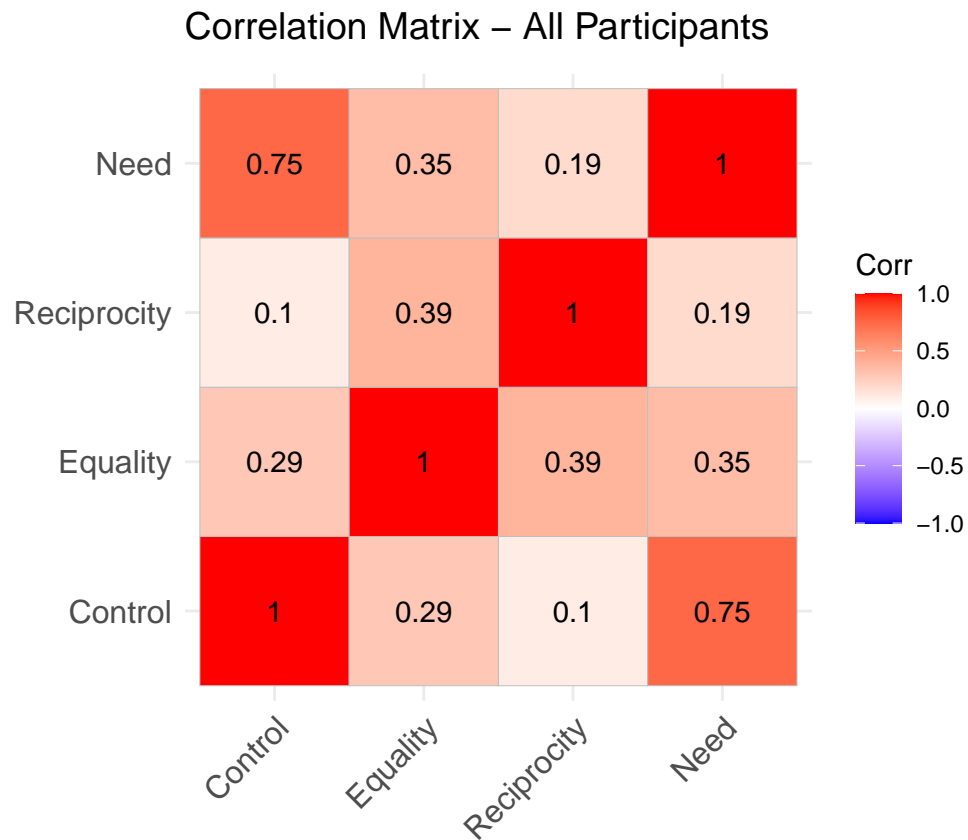

```
ggcorrplot(cor_matrix_slovak, lab = TRUE) +
  ggtitle("Correlation Matrix - Slovak") +
  theme(plot.title = element_text(size = 14, margin = margin(b = 10)))
```

Correlation Matrix – Slovak


```
ggcorrplot(cor_matrix_roma, lab = TRUE) +
  ggtitle("Correlation Matrix - Roma") +
  theme(plot.title = element_text(size = 14, margin = margin(b = 10)))
```

Correlation Matrix – Roma