# ANOVA Lecture Material

Prof. Dr. Stephan Huber[1,2]

[1] Fresenius University of Applied Science
[2] Charlotte Fresenius University

Abstract

In this paper, I illustrate the process of making nice tables and graphics that are related to ANOVA and had been shown in the lecture. The paper adheres to the APA style, implementing the R template provided by the 'papaja' package (Aust & Barth, 2023).

*Keywords:* papaja, descriptive statistics

**Contents**



---

All files related to that paper are hostes on github. see: https://github.com/hubchev/ewa.

Correspondence concerning this article should be addressed to Prof. Dr. Stephan Huber, Im Mediapark 4e. E-mail: stephan.huber@hs-fresenius.de

## 1 Summary

I create many of the tables and figures of lecture in this report. In particular, I show the full dataset in Table 1. Table 2 contains summary statistics for all variables and Table 3 for all values of the combinations of variables of `modus` and `kognition`. Table 4 shows the ANOVA results. Table 5 also shows ANOVA results but with more interactions.

Figure 1 shows boxplots for all combinations of variables of `modus` and `kognition`. Figure 3 shows an interaction plot of `dauer` and `modus`. Figure 2 shows an interaction plot of `dauer` and `kognition`. Figure 4 shows boxplots for all combinations of variables of `modus`, `kognition`, and interviewer.

## 2 Data Preperation

```
if (!require(pacman)) install.packages("pacman")
pacman::p_load(tidyverse, janitor, psych, car, knitr, papaja, kableExtra, stargazer)
rm(list = ls())

ModKogDat <- read.csv("../data/ModKogDat.csv", header=TRUE, sep=",")
# Read in data
df <- ModKogDat |>
  mutate(
    modus = as.factor(modus),
    kognition = as.factor(kognition)
    ) |>
  group_by(modus, kognition) |>
  mutate(
    id_num = cur_group_id(),
    m_str = substr(modus, 1, 2),
    k_str = substr(kognition, 1, 2),
    id = paste(m_str, k_str, sep = "_")
  ) |>
  select(-m_str, -k_str) |>
  tibble() |>
  ungroup()
```

## 3 Inspect Data

*Figure 1*. Boxplots of all combinations of `modus` and `kognition`

Table 1
*Full Dataset*

| dauer | modus | kognition | id_num | id |
|---|---|---|---|---|
| 8 | a1 telefon | b1 altersadaequat | 1 | a1_b1 |
| 16 | a1 telefon | b1 altersadaequat | 1 | a1_b1 |
| 12 | a1 telefon | b1 altersadaequat | 1 | a1_b1 |
| 7 | a1 telefon | b1 altersadaequat | 1 | a1_b1 |
| 17 | a1 telefon | b1 altersadaequat | 1 | a1_b1 |
| 20 | a1 telefon | b2 lkb | 2 | a1_b2 |
| 26 | a1 telefon | b2 lkb | 2 | a1_b2 |
| 20 | a1 telefon | b2 lkb | 2 | a1_b2 |
| 14 | a1 telefon | b2 lkb | 2 | a1_b2 |
| 20 | a1 telefon | b2 lkb | 2 | a1_b2 |
| 10 | a1 telefon | b3 begdem | 3 | a1_b3 |
| 7 | a1 telefon | b3 begdem | 3 | a1_b3 |
| 11 | a1 telefon | b3 begdem | 3 | a1_b3 |
| 9 | a1 telefon | b3 begdem | 3 | a1_b3 |
| 13 | a1 telefon | b3 begdem | 3 | a1_b3 |
| 15 | a2 besuch | b1 altersadaequat | 4 | a2_b1 |
| 25 | a2 besuch | b1 altersadaequat | 4 | a2_b1 |
| 22 | a2 besuch | b1 altersadaequat | 4 | a2_b1 |
| 19 | a2 besuch | b1 altersadaequat | 4 | a2_b1 |
| 29 | a2 besuch | b1 altersadaequat | 4 | a2_b1 |
| 32 | a2 besuch | b2 lkb | 5 | a2_b2 |
| 27 | a2 besuch | b2 lkb | 5 | a2_b2 |
| 26 | a2 besuch | b2 lkb | 5 | a2_b2 |
| 20 | a2 besuch | b2 lkb | 5 | a2_b2 |
| 25 | a2 besuch | b2 lkb | 5 | a2_b2 |
| 30 | a2 besuch | b3 begdem | 6 | a2_b3 |
| 21 | a2 besuch | b3 begdem | 6 | a2_b3 |
| 33 | a2 besuch | b3 begdem | 6 | a2_b3 |
| 39 | a2 besuch | b3 begdem | 6 | a2_b3 |
| 27 | a2 besuch | b3 begdem | 6 | a2_b3 |

Table 2
*Summary Statistics*

| Variables | n | mean | sd | median | min | max | se |
|---|---|---|---|---|---|---|---|
| dauer | 30.00 | 20.00 | 8.44 | 20.00 | 7.00 | 39.00 | 1.54 |
| modus* | 30.00 | 1.50 | 0.51 | 1.50 | 1.00 | 2.00 | 0.09 |
| kognition* | 30.00 | 2.00 | 0.83 | 2.00 | 1.00 | 3.00 | 0.15 |
| id_num | 30.00 | 3.50 | 1.74 | 3.50 | 1.00 | 6.00 | 0.32 |
| id* | 30.00 | 3.50 | 1.74 | 3.50 | 1.00 | 6.00 | 0.32 |

*Note.* This table contains all variables.

Table 3
*Summary Statistics for the Variable 'dauer'*

| id | count | mean | sd | COV (sd/mean) | min | q25 | median | q75 | max |
|---|---|---|---|---|---|---|---|---|---|
| a1_b1 | 5 | 12.00 | 4.53 | 0.38 | 7 | 8.00 | 12 | 16.00 | 17 |
| a1_b2 | 5 | 20.00 | 4.24 | 0.21 | 14 | 20.00 | 20 | 20.00 | 26 |
| a1_b3 | 5 | 10.00 | 2.24 | 0.22 | 7 | 9.00 | 10 | 11.00 | 13 |
| a2_b1 | 5 | 22.00 | 5.39 | 0.24 | 15 | 19.00 | 22 | 25.00 | 29 |
| a2_b2 | 5 | 26.00 | 4.30 | 0.17 | 20 | 25.00 | 26 | 27.00 | 32 |
| a2_b3 | 5 | 30.00 | 6.71 | 0.22 | 21 | 27.00 | 30 | 33.00 | 39 |

*Note.* This table contains summary statistics for each combination of 'modus' and 'kognition'

Table 4
*A beautiful ANOVA table.*

| Effect | $\hat{\eta}^2_G$ | 90% CI | $F$ | $df$ | $df_{\text{res}}$ | $p$ |
|---|---|---|---|---|---|---|
| Modus | .665 | [.462, .778] | 47.65 | 1 | 24 | $< .001$ |
| Kognition | .249 | [.014, .450] | 3.97 | 2 | 24 | .032 |
| Modus × Kognition | .323 | [.062, .517] | 5.74 | 2 | 24 | .009 |

*Note.* Bli bla blubb.

## 4 Interaction Plots

*Figure 2*. Interaction Plot: `dauer` and `modus`

*Figure 3*. Interaction Plot: `dauer` and `kognition`

## 5 Contrast Matrix

```r
contrasts(df$kognition) <- cbind(c(2, -1, -1), c(0, 1,-1))
```

## 6 Data `ModKogDat3F.csv`

```r
df3 <- read.csv("../data/ModKogDat3F.csv", header=TRUE, sep=",") |>
  mutate(
    modus = as.factor(modus),
    kognition = as.factor(kognition),
    interviewer = as.factor(interviewer)
  ) |>
    group_by(modus, kognition, interviewer) |>
  mutate(
    id_num = cur_group_id(),
    m_str = substr(modus, 1, 2),
    k_str = substr(kognition, 1, 2),
    i_str = substr(interviewer, 1, 2),
    id = paste(m_str, k_str, i_str, sep = "_")
  ) |>
  select(-m_str, -k_str, -i_str) |>
  tibble()
```

Table 5
*A beautiful ANOVA table.*

| Effect | $\hat{\eta}^2_G$ | 90% CI | $F$ | $df$ | $df_{\text{res}}$ | $p$ |
|---|---|---|---|---|---|---|
| Modus | .665 | [.462, .778] | 47.65 | 1 | 24 | $< .001$ |
| Kognition | .249 | [.014, .450] | 3.97 | 2 | 24 | .032 |
| Kognition × altersadäquat vs beeinträchtigt | .199 | [.018, .419] | 5.96 | 1 | 24 | .022 |
| Kognition × LKB vs beginnende Demenz | .076 | [.000, .282] | 1.99 | 1 | 24 | .172 |
| Modus × Kognition | .323 | [.062, .517] | 5.74 | 2 | 24 | .009 |
| Modus × Kognition × altersadäquat vs beeinträchtigt | .199 | [.018, .419] | 5.96 | 1 | 24 | .022 |
| Modus × Kognition × LKB vs beginnende Demenz | .187 | [.013, .407] | 5.51 | 1 | 24 | .027 |

*Note.* Bli bla blubb.

*Figure 4*. Boxplots of all combinations of `modus`, `kognition`, and `interviewer`

## 7 Excercises

1. In Table 7 is an error. What is wrong here? Please correct.
2. Table 6 is too long. Please split it up to two tables by interviewer.
3. Tables that relate to `ModKogDat3F.csv` data are not yet mentioned in the summary. Please add them, because, according to APA rules, each Figure and Table, respectively, must be mentioned in the text.

Table 6
*Full Dataset: 'ModKogDat3F.csv'*

| dauer | modus | kognition | interviewer | id_num | id |
|---|---|---|---|---|---|
| 8 | telefon | altersadaequat | ehrenamt | 7 | te_al_eh |
| 16 | telefon | altersadaequat | ehrenamt | 7 | te_al_eh |
| 12 | telefon | altersadaequat | ehrenamt | 7 | te_al_eh |
| 7 | telefon | altersadaequat | ehrenamt | 7 | te_al_eh |
| 17 | telefon | altersadaequat | ehrenamt | 7 | te_al_eh |
| 20 | telefon | lkb | ehrenamt | 11 | te_lk_eh |
| 26 | telefon | lkb | ehrenamt | 11 | te_lk_eh |
| 20 | telefon | lkb | ehrenamt | 11 | te_lk_eh |
| 14 | telefon | lkb | ehrenamt | 11 | te_lk_eh |
| 20 | telefon | lkb | ehrenamt | 11 | te_lk_eh |
| 10 | telefon | begdem | ehrenamt | 9 | te_be_eh |
| 7 | telefon | begdem | ehrenamt | 9 | te_be_eh |
| 11 | telefon | begdem | ehrenamt | 9 | te_be_eh |
| 9 | telefon | begdem | ehrenamt | 9 | te_be_eh |
| 13 | telefon | begdem | ehrenamt | 9 | te_be_eh |
| 15 | hausbesuch | altersadaequat | ehrenamt | 1 | ha_al_eh |
| 25 | hausbesuch | altersadaequat | ehrenamt | 1 | ha_al_eh |
| 22 | hausbesuch | altersadaequat | ehrenamt | 1 | ha_al_eh |
| 19 | hausbesuch | altersadaequat | ehrenamt | 1 | ha_al_eh |
| 29 | hausbesuch | altersadaequat | ehrenamt | 1 | ha_al_eh |
| 32 | hausbesuch | lkb | ehrenamt | 5 | ha_lk_eh |
| 27 | hausbesuch | lkb | ehrenamt | 5 | ha_lk_eh |
| 26 | hausbesuch | lkb | ehrenamt | 5 | ha_lk_eh |
| 20 | hausbesuch | lkb | ehrenamt | 5 | ha_lk_eh |
| 25 | hausbesuch | lkb | ehrenamt | 5 | ha_lk_eh |
| 30 | hausbesuch | begdem | ehrenamt | 3 | ha_be_eh |
| 21 | hausbesuch | begdem | ehrenamt | 3 | ha_be_eh |
| 33 | hausbesuch | begdem | ehrenamt | 3 | ha_be_eh |
| 39 | hausbesuch | begdem | ehrenamt | 3 | ha_be_eh |
| 27 | hausbesuch | begdem | ehrenamt | 3 | ha_be_eh |
| 8 | telefon | altersadaequat | profi | 8 | te_al_pr |
| 16 | telefon | altersadaequat | profi | 8 | te_al_pr |
| 12 | telefon | altersadaequat | profi | 8 | te_al_pr |
| 7 | telefon | altersadaequat | profi | 8 | te_al_pr |
| 17 | telefon | altersadaequat | profi | 8 | te_al_pr |
| 20 | telefon | lkb | profi | 12 | te_lk_pr |
| 26 | telefon | lkb | profi | 12 | te_lk_pr |
| 20 | telefon | lkb | profi | 12 | te_lk_pr |
| 14 | telefon | lkb | profi | 12 | te_lk_pr |
| 20 | telefon | lkb | profi | 12 | te_lk_pr |
| 10 | telefon | begdem | profi | 10 | te_be_pr |
| 7 | telefon | begdem | profi | 10 | te_be_pr |
| 11 | telefon | begdem | profi | 10 | te_be_pr |
| 9 | telefon | begdem | profi | 10 | te_be_pr |
| 13 | telefon | begdem | profi | 10 | te_be_pr |
| 15 | hausbesuch | altersadaequat | profi | 2 | ha_al_pr |
| 25 | hausbesuch | altersadaequat | profi | 2 | ha_al_pr |
| 22 | hausbesuch | altersadaequat | profi | 2 | ha_al_pr |

Table 7
*Summary Statistics: 'ModKogDat3F.csv'*

| Variables | n | mean | sd | median | min | max | se |
|---|---|---|---|---|---|---|---|
| dauer | 30.00 | 20.00 | 8.44 | 20.00 | 7.00 | 39.00 | 1.54 |
| modus* | 30.00 | 1.50 | 0.51 | 1.50 | 1.00 | 2.00 | 0.09 |
| kognition* | 30.00 | 2.00 | 0.83 | 2.00 | 1.00 | 3.00 | 0.15 |
| id_num | 30.00 | 3.50 | 1.74 | 3.50 | 1.00 | 6.00 | 0.32 |
| id* | 30.00 | 3.50 | 1.74 | 3.50 | 1.00 | 6.00 | 0.32 |

*Note.* This table contains all variables.

Table 8
*Summary Statistics for the Variable 'dauer': 'ModKogDat3F.csv'*

| id | count | mean | sd | COV (sd/mean) | min | q25 | median | q75 | max |
|---|---|---|---|---|---|---|---|---|---|
| ha_al_eh | 5 | 22.00 | 5.39 | 0.24 | 15 | 19.00 | 22 | 25.00 | 29 |
| ha_al_pr | 5 | 22.00 | 5.39 | 0.24 | 15 | 19.00 | 22 | 25.00 | 29 |
| ha_be_eh | 5 | 30.00 | 6.71 | 0.22 | 21 | 27.00 | 30 | 33.00 | 39 |
| ha_be_pr | 5 | 30.00 | 6.71 | 0.22 | 21 | 27.00 | 30 | 33.00 | 39 |
| ha_lk_eh | 5 | 26.00 | 4.30 | 0.17 | 20 | 25.00 | 26 | 27.00 | 32 |
| ha_lk_pr | 5 | 26.00 | 4.30 | 0.17 | 20 | 25.00 | 26 | 27.00 | 32 |
| te_al_eh | 5 | 12.00 | 4.53 | 0.38 | 7 | 8.00 | 12 | 16.00 | 17 |
| te_al_pr | 5 | 12.00 | 4.53 | 0.38 | 7 | 8.00 | 12 | 16.00 | 17 |
| te_be_eh | 5 | 10.00 | 2.24 | 0.22 | 7 | 9.00 | 10 | 11.00 | 13 |
| te_be_pr | 5 | 10.00 | 2.24 | 0.22 | 7 | 9.00 | 10 | 11.00 | 13 |
| te_lk_eh | 5 | 20.00 | 4.24 | 0.21 | 14 | 20.00 | 20 | 20.00 | 26 |
| te_lk_pr | 5 | 20.00 | 4.24 | 0.21 | 14 | 20.00 | 20 | 20.00 | 26 |

*Note.* This table contains summary statistics for each combination of 'modus', 'kognition', and 'interviewer'

Table 9
*A beautiful ANOVA table with 'interviewer'.*

| Effect | $\hat{\eta}^2_G$ | 90% CI | $F$ | $df$ | $df_{\text{res}}$ | $p$ |
|---|---|---|---|---|---|---|
| Modus | .665 | [.533, .751] | 95.29 | 1 | 48 | $< .001$ |
| Kognition | .249 | [.077, .398] | 7.94 | 2 | 48 | .001 |
| Interviewer | .000 | [.000, .000] | 0.00 | 1 | 48 | $> .999$ |
| Modus × Kognition | .323 | [.140, .469] | 11.47 | 2 | 48 | $< .001$ |
| Modus × Interviewer | .000 | [.000, .000] | 0.00 | 1 | 48 | $> .999$ |
| Kognition × Interviewer | .000 | [.000, .000] | 0.00 | 2 | 48 | $> .999$ |
| Modus × Kognition × Interviewer | .000 | [.000, .000] | 0.00 | 2 | 48 | $> .999$ |

*Note.* Bli bla blubb.

Table 10
*Summary Statistics: 'ModKogDat3F.csv'*

| Variables | n | mean | sd | median | min | max | se |
|---|---|---|---|---|---|---|---|
| dauer | 60.00 | 20.00 | 8.36 | 20.00 | 7.00 | 39.00 | 1.08 |
| modus* | 60.00 | 1.50 | 0.50 | 1.50 | 1.00 | 2.00 | 0.07 |
| kognition* | 60.00 | 2.00 | 0.82 | 2.00 | 1.00 | 3.00 | 0.11 |
| interviewer* | 60.00 | 1.50 | 0.50 | 1.50 | 1.00 | 2.00 | 0.07 |
| id_num | 60.00 | 6.50 | 3.48 | 6.50 | 1.00 | 12.00 | 0.45 |
| id* | 60.00 | 6.50 | 3.48 | 6.50 | 1.00 | 12.00 | 0.45 |

*Note.* This table contains all variables.

## 8 Solutions

1. The wrong dataframe was used here. The correct data is `df3`. Table 10 is the correct one.

```
tabsumstat3 <- df3 |>
  psych::describe()   |>
  as_tibble(rownames="Variables")  |>
  select(-skew, -kurtosis, -range, -vars, -trimmed, -mad)

apa_table(
  tabsumstat3
  , caption = "Summary Statistics: `ModKogDat3F.csv`"
  , note = "This table contains all variables."
  , escape = TRUE
)
```

2. The splitted tables are shown in Tables 11 and 12 and here is the corresponding code:

```
df3p <- df3 |>
  filter(interviewer == "profi")
df3e <- df3 |>
  filter(interviewer == "ehrenamt")
```

```
apa_table(df3p,  caption = "Interviews by Professionals")
```

```
apa_table(df3e,  caption = "Interviews by Volunteers (Ehrenamt)")
```

3. The unmentioned tables are Tables 6, 7, 8, 9, and Figure 4.

Table 11
*Interviews by Professionals*

| dauer | modus | kognition | interviewer | id_num | id |
|---|---|---|---|---|---|
| 8 | telefon | altersadaequat | profi | 8 | te_al_pr |
| 16 | telefon | altersadaequat | profi | 8 | te_al_pr |
| 12 | telefon | altersadaequat | profi | 8 | te_al_pr |
| 7 | telefon | altersadaequat | profi | 8 | te_al_pr |
| 17 | telefon | altersadaequat | profi | 8 | te_al_pr |
| 20 | telefon | lkb | profi | 12 | te_lk_pr |
| 26 | telefon | lkb | profi | 12 | te_lk_pr |
| 20 | telefon | lkb | profi | 12 | te_lk_pr |
| 14 | telefon | lkb | profi | 12 | te_lk_pr |
| 20 | telefon | lkb | profi | 12 | te_lk_pr |
| 10 | telefon | begdem | profi | 10 | te_be_pr |
| 7 | telefon | begdem | profi | 10 | te_be_pr |
| 11 | telefon | begdem | profi | 10 | te_be_pr |
| 9 | telefon | begdem | profi | 10 | te_be_pr |
| 13 | telefon | begdem | profi | 10 | te_be_pr |
| 15 | hausbesuch | altersadaequat | profi | 2 | ha_al_pr |
| 25 | hausbesuch | altersadaequat | profi | 2 | ha_al_pr |
| 22 | hausbesuch | altersadaequat | profi | 2 | ha_al_pr |
| 19 | hausbesuch | altersadaequat | profi | 2 | ha_al_pr |
| 29 | hausbesuch | altersadaequat | profi | 2 | ha_al_pr |
| 32 | hausbesuch | lkb | profi | 6 | ha_lk_pr |
| 27 | hausbesuch | lkb | profi | 6 | ha_lk_pr |
| 26 | hausbesuch | lkb | profi | 6 | ha_lk_pr |
| 20 | hausbesuch | lkb | profi | 6 | ha_lk_pr |
| 25 | hausbesuch | lkb | profi | 6 | ha_lk_pr |
| 30 | hausbesuch | begdem | profi | 4 | ha_be_pr |
| 21 | hausbesuch | begdem | profi | 4 | ha_be_pr |
| 33 | hausbesuch | begdem | profi | 4 | ha_be_pr |
| 39 | hausbesuch | begdem | profi | 4 | ha_be_pr |
| 27 | hausbesuch | begdem | profi | 4 | ha_be_pr |

Table 12
*Interviews by Volunteers (Ehrenamt)*

| dauer | modus | kognition | interviewer | id_num | id |
|---|---|---|---|---|---|
| 8 | telefon | altersadaequat | ehrenamt | 7 | te_al_eh |
| 16 | telefon | altersadaequat | ehrenamt | 7 | te_al_eh |
| 12 | telefon | altersadaequat | ehrenamt | 7 | te_al_eh |
| 7 | telefon | altersadaequat | ehrenamt | 7 | te_al_eh |
| 17 | telefon | altersadaequat | ehrenamt | 7 | te_al_eh |
| 20 | telefon | lkb | ehrenamt | 11 | te_lk_eh |
| 26 | telefon | lkb | ehrenamt | 11 | te_lk_eh |
| 20 | telefon | lkb | ehrenamt | 11 | te_lk_eh |
| 14 | telefon | lkb | ehrenamt | 11 | te_lk_eh |
| 20 | telefon | lkb | ehrenamt | 11 | te_lk_eh |
| 10 | telefon | begdem | ehrenamt | 9 | te_be_eh |
| 7 | telefon | begdem | ehrenamt | 9 | te_be_eh |
| 11 | telefon | begdem | ehrenamt | 9 | te_be_eh |
| 9 | telefon | begdem | ehrenamt | 9 | te_be_eh |
| 13 | telefon | begdem | ehrenamt | 9 | te_be_eh |
| 15 | hausbesuch | altersadaequat | ehrenamt | 1 | ha_al_eh |
| 25 | hausbesuch | altersadaequat | ehrenamt | 1 | ha_al_eh |
| 22 | hausbesuch | altersadaequat | ehrenamt | 1 | ha_al_eh |
| 19 | hausbesuch | altersadaequat | ehrenamt | 1 | ha_al_eh |
| 29 | hausbesuch | altersadaequat | ehrenamt | 1 | ha_al_eh |
| 32 | hausbesuch | lkb | ehrenamt | 5 | ha_lk_eh |
| 27 | hausbesuch | lkb | ehrenamt | 5 | ha_lk_eh |
| 26 | hausbesuch | lkb | ehrenamt | 5 | ha_lk_eh |
| 20 | hausbesuch | lkb | ehrenamt | 5 | ha_lk_eh |
| 25 | hausbesuch | lkb | ehrenamt | 5 | ha_lk_eh |
| 30 | hausbesuch | begdem | ehrenamt | 3 | ha_be_eh |
| 21 | hausbesuch | begdem | ehrenamt | 3 | ha_be_eh |
| 33 | hausbesuch | begdem | ehrenamt | 3 | ha_be_eh |
| 39 | hausbesuch | begdem | ehrenamt | 3 | ha_be_eh |
| 27 | hausbesuch | begdem | ehrenamt | 3 | ha_be_eh |

## References

Aust, F., & Barth, M. (2023). *papaja: Prepare reproducible APA journal articles with R markdown*. Retrieved from https://github.com/crsh/papaja