# SEM 1

 $Simon\ Roth$ 

21.4.2017

## 1 Datensatz einladen

### 1.1 Itembatterien

### Islamophobie:

- mm01: ISLAMAUSUEBUNG IN BRD BESCHRAENKEN
  - -10 Befragter gehört einer islamischen Religionsgemeinschaft an (Code 1 in rd03)
  - -9
  - 1 Stimme überhaupt nicht zu
  - -2
  - 3
  - -4
  - 5
  - 67 Stimme voll und ganz zu
- mm02: ISLAM PASST IN DIE DEUTSCHE GESELLSCHAFT
- mm03: ANWESENHEIT VON MUSLIMEN BRINGT KONFLIKT
- mm04: STAAT SOLLTE ISLAM. GRUPPEN BEOBACHTEN
- mm05: MUSLIMISCHER BUERGERMEISTER IN ORDNUNG
- mm06: UNTER MUSLIMEN SIND VIELE REL. FANATIKER

#### Deutschein:

• mn11: DEUTSCH SEIN: DEUTSCHE STAATSBUERGERSCH.

```
• mn12: DEUTSCH SEIN: CHRISTL.RELIGION ZUGEHOER.
```

- mn13: DEUTSCH SEIN: BEKENNTNIS ZUR DEMOKRATIE
- mn14: DEUTSCH SEIN: VIELE DEUTSCHE BEKANNTE
- mn15: DEUTSCH SEIN: ALTE STAATSANGEH.AUFGEBEN
- mn16: DEUTSCH SEIN: VERBUNDENHEIT ZU DEUTSCHL.
- mn17: DEUTSCH SEIN: ALTE GEBRAEUCHE ABLEGEN
- mn18: DEUTSCH SEIN: GUT DEUTSCH SPRECHEN
- mn19: DEUTSCH SEIN: WESTLICHE WERTE TEILEN
- mn20: DEUTSCH SEIN: MIND. 1 ELTERNTEIL DEUTSCH
- mn21: DEUTSCH SEIN: IN DEUTSCHLAND GEBOREN

# 2 Daten partitionieren

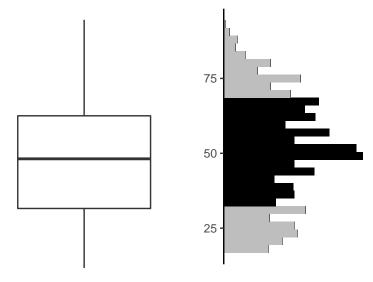
- sex: GESCHLECHT (Int.: Geschlecht der befragten Person ohne Befragen eintragen!)
  - 1 Männlich
  - 2 Weiblich
- age: ALTER: metrisch
- agec: ALTER: KATEGORISIERT 6
  - 18 29 Jahre
  - 30 44 Jahre
  - -45 59 Jahre
  - -60 74 Jahre
  - 75 89 Jahre
  - Über 89 Jahre
- isced 97: BEFR.: ISCED 1997 - 6 STUFEN: International Standard Classification of Education (ISCED) 1997, 6 Stufen
  - 1. Level Primary education or first stage of basic education
  - 2. Level Lower secondary or second stage of basic education
  - 3. Level (Upper) secondary education
  - 4. Level Post-secondary non-tertiary education
  - 5. Level First stage of tertiary education
  - 6. Level Second stage of tertiary education

```
### subset data
all_2 <- all_1 %>%
  select(respid, sex,
         age, agec, isced97,
         num_range('mn', 11:21),
         num_range('mm0', 1:6))
islam <- c('stop_islam', 'islam_is_ger', 'muslim_confl', 'muslim_surv', 'muslim_major', 'muslim_rad')</pre>
deutsch <- c('ger_citiz', 'christ_conf', 'pro_demo', 'ger_friends', 'only_ger', 'dis_roots', 'ger_affini</pre>
colnames(all_2) <- c(c('respid', 'sex', 'age', 'agec', 'isced97'), deutsch, islam)</pre>
### clean from -11 and other NA
all_3 <- all_2
all_3 \leftarrow lapply(all_2, FUN = function(x){ifelse(x > 0, x, NA)}) %>%
  as.data.frame()
### recode variables
all_3 <- all_3 %>%
  mutate(gender = ifelse(sex %in% 2, 1, 0)) # F = 1, M = 0)
```

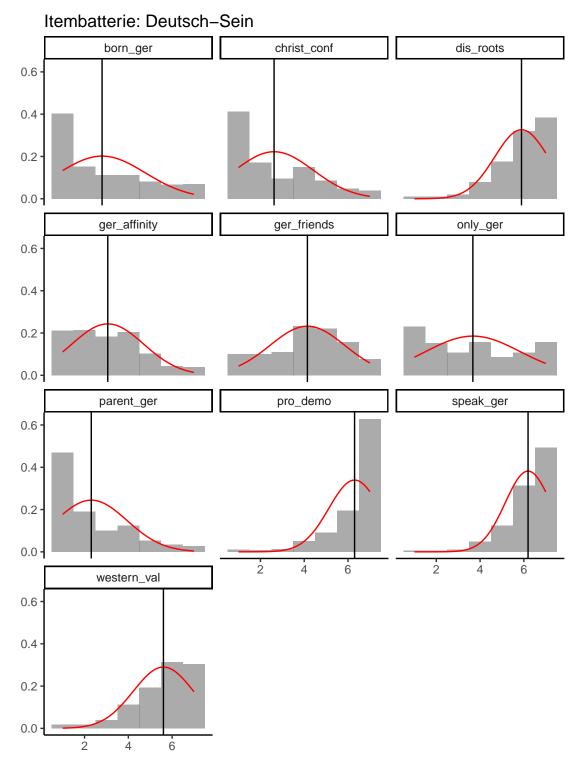
# 3 Visualisierungen

Der Boxplot zeigt die quantile der **Altersverteilung**. Der Median liegt bei 50 Lebensjahren, wobei 50% der Befragten zwischen 30 und 70 Jahre alt sind. Die Verteilung des Alters nähert sich einer gaussischen Normalverteilung an.

```
library(ggplot2)
library(ggthemes)
### Boxplot
gg_1 <- all_3 %>%
  ggplot(aes(x = factor(""), y = age)) +
  geom_boxplot() +
  theme_map()
mea <- mean(all_3$age, na.rm = T)</pre>
s \leftarrow sd(all_3 age, na.rm = T)
mask <- ifelse(all_3$age > (mea + s) |
                 all_3$age < (mea - s), T, F)
gg_2 <- cbind(all_3, mask) %>%
  ggplot(aes(age)) +
  geom_histogram(aes(fill = mask)) +
  scale_y_discrete() +
  scale_fill_manual(values = c("black", "grey")) +
  coord_flip() +
  theme_classic() +
  theme(legend.position = "none") +
  labs(x = "", y = "")
library(gridExtra)
grid.arrange(gg_1, gg_2, ncol = 2)
```



```
library(tidyr)
### help function
norm_dens <- function(dat) {</pre>
    mi <- min(dat$value)</pre>
    ma <- max(dat$value)</pre>
    grid \leftarrow seq(mi, ma, by = .1)
    return(data.frame(
    predicted = grid,
    density = dnorm(grid, mean(dat$value), sd(dat$value))
  }
### Islamophobie
#names(all_3)
gg_data <- all_3[, c(1,7:16)] %>%
  gather(key, value, -respid) %>% # , -trt
  na.omit()
# melt(
# messy,
# variable.name = "key",
# value.names = "value",
# id.vars = c("id", "trt")
normaldens <- gg_data %>%
  group_by(key) %>%
  do(norm_dens(.))
m <- gg_data \%>%
  group_by(key) %>%
  summarise(m1 = mean(value))
gg_h1 <- ggplot(data = gg_data, aes(x = value)) +
    facet_wrap(~ key, ncol = 3) +
    geom_histogram(aes(y = ..density..), binwidth = 1, alpha = .5) +
    geom_line(data = normaldens,
              aes(x = predicted, y = density),
              colour = "red") +
    geom_vline(data = m, aes(xintercept = m$m1)) +
    labs(x = "", y = "", title = "Itembatterie: Deutsch-Sein") + theme_classic()
gg_h1
```

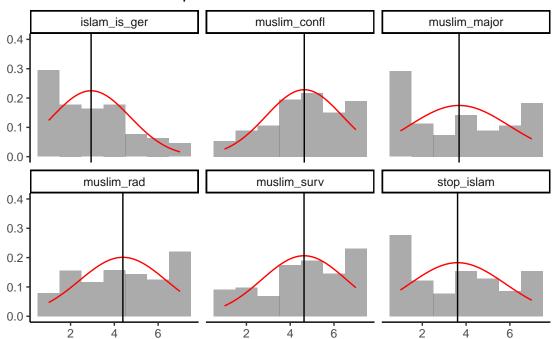


Wie wichtig sind die folgenden Dinge, damit Sie sagen würden, dass eine Person mit ausländischen Wurzeln ein wirklicher Deutscher/eine wirkliche Deutsche wird? Religiöse sowie ethnische Voraussetzungen scheinen für die Mehrheit der Befragten irrelevant. Entscheidend sind hingegen kulturelle Faktoren, wie Demokratieverständnis und soziale Assimilation. Nur 2 von 10 Variablen sind annäherend normalverteilt und weisen ansonst eine Tendenz zu den Skalenrändern auf.

#### Itembatterie: Deutsch-Sein christ\_conf pro\_demo ger\_friends ger\_affinity speak\_ger western\_val only\_ger dis\_roots parent\_ger born\_ger 400 christ\_conf 300 -Corr: Corr: Corr: Corr: Corr: Corr: Corr: Corr: Corr: 200 -0.548 -0.01410.319 0.261 0.08 0.391 0.0272 0.216 0.422 100 -7.2 pro\_demo 6.8 Corr: Corr: Corr: Corr: Corr: Corr: Corr: Corr: 6.4 0.145 0.0755 0.329 -0.00540.25 0.329 -0.0929-0.08036.0 -6.0 -5.5 -5.0 -4.5 -4.0 -3.5 ger\_friends Corr: Corr: Corr: Corr: Corr: Corr: Corr: 0.083 0.261 0.167 0.164 0.249 0.236 0.204 only\_ger Corr: Corr: Corr: Corr: Corr: Corr: 0.223 0.454 0.12 0.245 0.27 0.28 3 -6.5 Corr: Corr: Corr: Corr: Corr: 6.0 0.0925 0.391 0.394 0.0112 0.0515 5.5 Corr: Corr: Corr: Corr: 0.125 0.23 0.419 0.362 7.00 6.75 -Corr: Corr: Corr: 6.50 6.25 0.427 -0.0268 0.0149 6.00 7.0 6.5 Corr: Corr: 6.0 0.119 0.0994 Corr: 3 -0.646

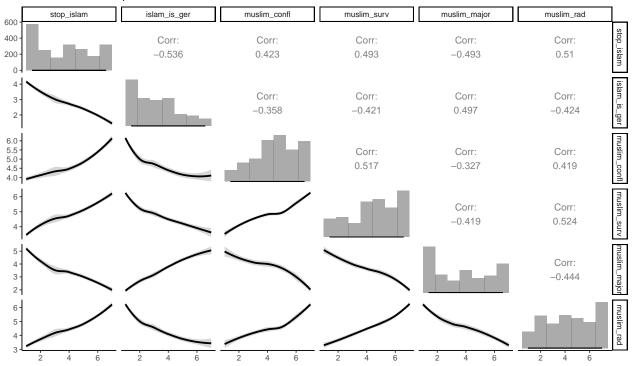
```
### Deutsch-Sein
gg_data <- all_3[, c(1,17:22)] %>%
  gather(key, value, -respid) %>%# , -trt
  na.omit()
normaldens <- gg_data \%>\%
  group_by(key) %>%
  do(norm_dens(.))
m <- gg_data %>%
  group_by(key) %>%
  summarise(m1 = mean(value))
gg_h2 <- ggplot(data = gg_data, aes(x = value)) +
    facet_wrap(~ key, ncol = 3) +
    geom_histogram(aes(y = ..density..), binwidth = 1, alpha = .5) +
    geom_line(data = normaldens,
              aes(x = predicted, y = density),
              colour = "red") +
    geom_vline(data = m, aes(xintercept = m$m1)) +
    ylim(0, .4) + labs(x = "", y = "", title = "Itembatterie: Islamophobie") +
    theme classic()
gg_h2
```

# Itembatterie: Islamophobie



Auch bei den Fragen zur Akzeptanz von Muslimen und des Islams zeigt sich die kulturelle Komponente. Während Muslime von einer Mehrheit der Deutschen als nicht besonders gefährlich oder ungefährlich eingestuft werden, wird die Deutsche Identität nicht mit dem Islam in Verbindung gebracht wird. Eine vermeintliche Islamisierung wird von den meisten Befragten nicht beobachtet. Im Gegenzug zu den Verteilungen zum Deutsch-sein zeigt sich beim Thema Islam eine stärkere Polarisierung (heavy-tailed).

### Itembatterie: Islamophobie



```
library(xtable)
print(xtable(mcor_table, caption="Pearson's correlation matrix"),
    rotate.colnames = T, size = "small", comment = F, missing = "?")
```

	age	isced97	ger_citiz	christ_conf	pro_demo	ger_friends	only_ger	dis_roots	ger_affinity	speak_ger	western_val	parent_ger	born_ger	${ m stop\_islam}$	islam_is_ger	muslim_confl	muslim_surv	muslim_major	muslim_rad
age	1.00	0.00	0.11	0.17	0.12	0.13	0.21	0.14	0.12	0.09	0.18	0.14	0.13	0.16	-0.18	0.11	0.26	-0.16	0.30
isced 97		1.00	-0.09	-0.12	0.14	0.03	-0.14	0.03	-0.18	-0.03	-0.04	-0.17	-0.17	-0.20	0.17	-0.10	-0.15	0.21	-0.22
$\operatorname{ger\_citiz}$			1.00	0.19	0.29	0.10	0.36	0.27	0.20	0.22	0.30	0.12	0.18	0.13	-0.15	0.18	0.18	-0.11	0.17
$christ\_conf$				1.00	-0.01	0.32	0.26	0.08	0.39	0.03	0.22	0.55	0.42	0.38	-0.27	0.24	0.32	-0.32	0.33
$pro\_demo$					1.00	0.15	0.08	0.33	-0.01	0.25	0.33	-0.09	-0.08	-0.05	0.06	0.06	0.06	0.06	-0.01
$ger\_friends$						1.00	0.08	0.26	0.17	0.16	0.25	0.24	0.20	0.09	-0.00	0.14	0.16	-0.07	0.10
$only\_ger$							1.00	0.22	0.45	0.12	0.25	0.27	0.28	0.30	-0.27	0.21	0.26	-0.30	0.31
$\operatorname{dis}_{\operatorname{roots}}$								1.00	0.09	0.39	0.39	0.01	0.05	0.05	0.01	0.17	0.15	-0.04	0.08
$ger\_affinity$									1.00	0.12	0.23	0.42	0.36	0.43	-0.32	0.26	0.31	-0.38	0.36
$speak\_ger$										1.00	0.43	-0.03	0.01	0.08	-0.10	0.17	0.16	-0.05	0.16
$western\_val$											1.00	0.12	0.10	0.20	-0.18	0.20	0.26	-0.17	0.23
$parent\_ger$												1.00	0.65	0.33	-0.23	0.23	0.28	-0.28	0.30
$born\_ger$													1.00	0.30	-0.22	0.19	0.27	-0.27	0.28
$stop\_islam$														1.00	-0.54	0.42	0.49	-0.49	0.51
$islam\_is\_ger$															1.00	-0.36	-0.42	0.50	-0.42
$muslim\_confl$																1.00	0.52	-0.33	0.42
$muslim\_surv$																	1.00	-0.42	0.52
$muslim\_major$																		1.00	-0.44
muslim_rad																			1.00

Table 1: Pearson's correlation matrix

Wie die Korrelationstabelle deutlich offen legt gibt es einen linearen Zusammenhang zwischen Alter und der islamophobischen Eeinstellungen. Während die Korrelation für die Itembatterie: Islamophobie sinnvoll interpretiert werden kann, liegen bei den *Deutsch-Sein* Variablen oft eine non-lineare Assoziation vor, wodurch Pearson-R seine Aussagekraft verliert.