

SEM 2

Simon Roth

21.4.2017

1 Datensatz einladen

a. german: Staatsbürgerschaft

- 1 Ja, ausschließlich
- 2 Ja, neben 2. Staatsbürgerschaft
- 3 Nein

b. eastwest: Erhebungsgebiet

- 1 West
- 2 Ost

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
 - 6
 - 7 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

Nationalbewusstsein:

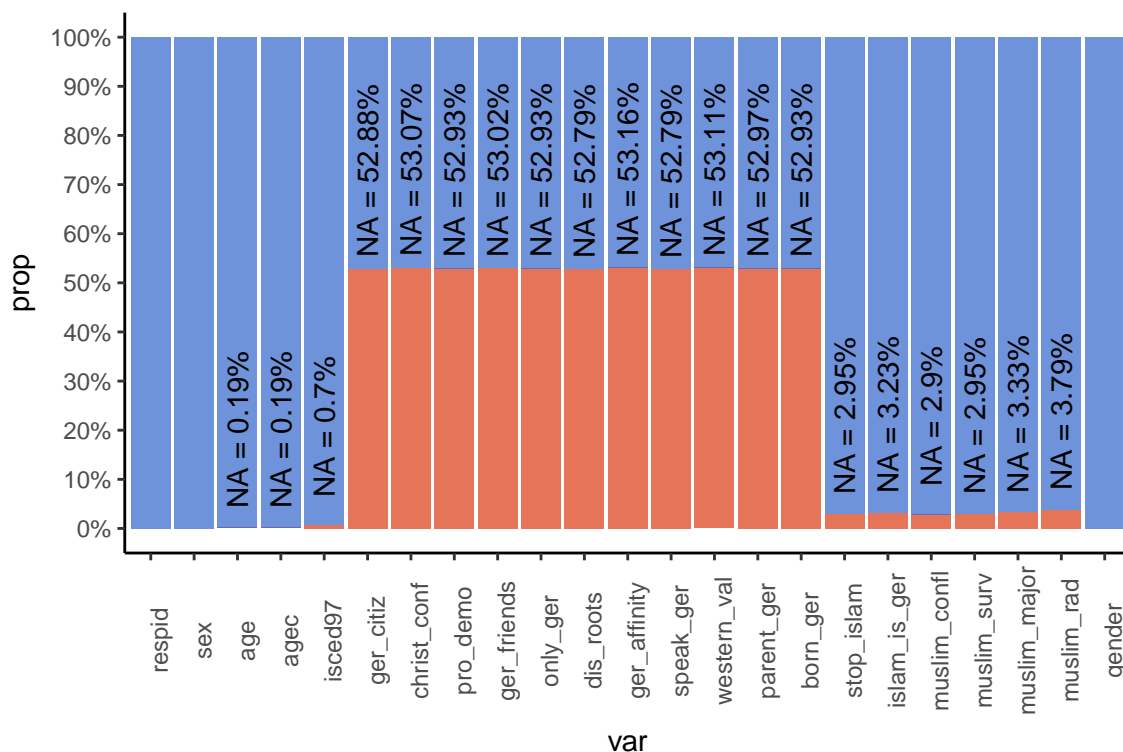
- mn11: DEUTSCH SEIN: DEUTSCHE STAATSBUERGERSCH.
- mn12: DEUTSCH SEIN: CHRISTL.RELIGION ZUGEHUER.
- mn13: DEUTSCH SEIN: BEKENNTNIS ZUR DEMOKRATIE
- mn14: DEUTSCH SEIN: VIELE DEUTSCHE BEKANNT
- mn15: DEUTSCH SEIN: ALTE STAATSANGEH.AUFGEBEN
- mn16: DEUTSCH SEIN: VERBUNDENHEIT ZU DEUTSCHL.
- mn17: DEUTSCH 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

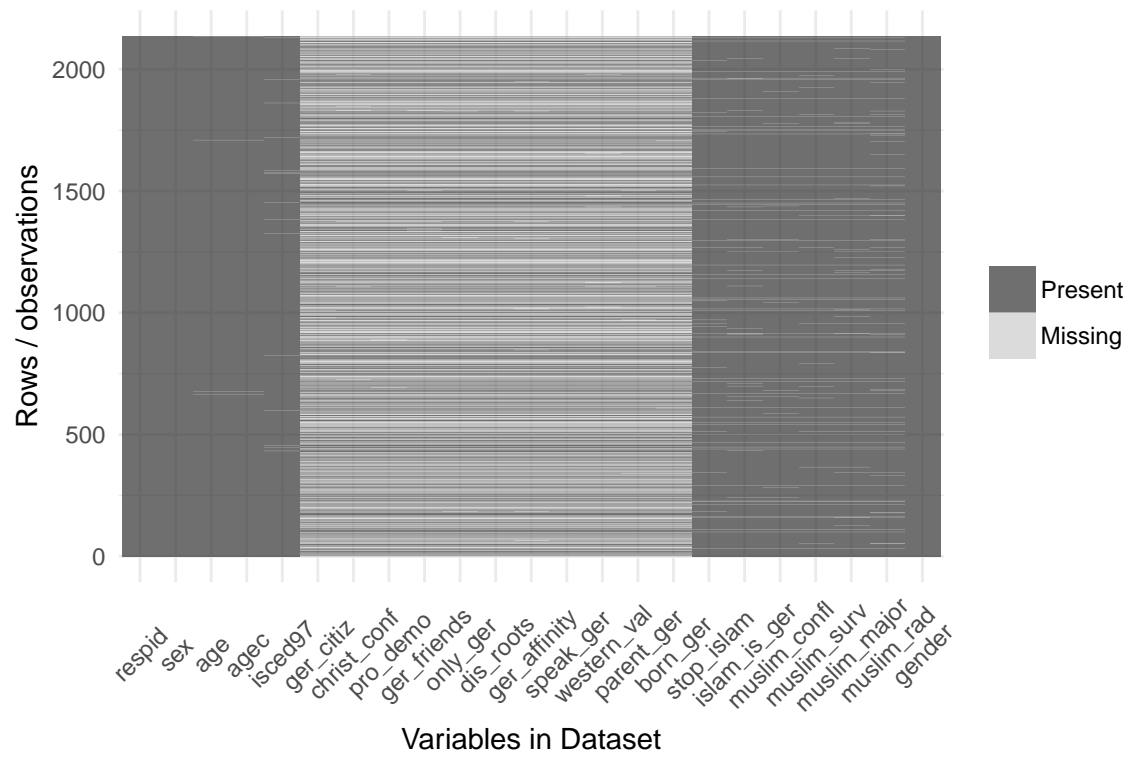
2 Daten partitionieren

- 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
- isced97: 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
- Allgemeiner Bildungsabschluss?

2.1 Missing Pattern

NZ (Fragebogen- Split): Befragter hat aufgrund eines methodischen oder inhaltlichen Splits eine Frage oder Item- batterie nicht gestellt bekommen.





2.2 Load VIM package for Visualization and Imputation of Missing Values

	age	agec	iscd97	ger_citiz	christ_conf	pro_demo	ger_friends	only_ger	dis_roots	ger_affinity	speak_ger	western_val	parent_ger	born_ger	stop_islam	islam_is_ger	muslim_confl	muslim_surv	muslim_major	muslim_rad
age	1	1	0	0	0	0	0	0	0	0	0	0	0	0	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
agec		1	0	0	0	0	0	0	0	0	0	0	0	0	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
iscd97			1	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	0.05	0.05	0.05	0.05	0.05	0.04
ger_citiz				1	0.99	1	0.99	1	1	0.99	1	0.99	0.99	1	0.01	0	0	0.01	0.01	0.02
christ_conf					1	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.01	0.01	0	0.01	0.01	0.01
pro_demo						1	0.99	0.99	1	0.99	1	0.99	0.99	0.99	0.01	0	0	0.01	0.01	0.02
ger_friends							1	0.99	1	0.99	1	0.99	0.99	0.99	0.01	0.01	0	0.01	0.01	0.01
only_ger								1	1	0.99	1	0.99	0.99	0.99	0.01	0	0	0.01	0.01	0.02
dis_roots									1	0.99	1	0.99	1	1	0.01	0	0	0.02	0.01	0.02
ger_affinity										1	0.99	0.99	0.99	0.99	0.02	0.01	0.01	0.02	0.01	0.02
speak_ger											1	0.99	1	1	0.01	0	0	0.02	0.01	0.02
western_val												1	0.99	0.99	0.01	0.01	0	0.01	0.01	0.01
parent_ger													1	1	0.01	0	0	0.01	0.01	0.02
born_ger														1	0.01	0	0	0.01	0.01	0.02
stop_islam															1	0.78	0.81	0.85	0.8	0.76
islam_is_ger																1	0.84	0.78	0.75	0.73
muslim_confl																	1	0.81	0.76	0.75
muslim_surv																		1	0.82	0.79
muslim_major																			1	0.74
muslim_rad																				1

Table 1: Pearson's correlation matrix of missing values

2.3 Little Test

A Test of Missing Completely at Random for Multivariate Data with Missing Values Roderick J. A. Little (1988)

Nullhypothesis: missing values pattern are missing completely at random

```
# install.packages("BaylorEdPsych")
# install.packages("mvmle")
```

```
library(BaylorEdPsych)
little_islam <- LittleMCAR(islam)
little_islam$amount.missing
little_islam$p.value

little_nation <- LittleMCAR(nation)
little_nation$amount.missing
little_nation$p.value
```

2.4 2.2 T-Tests

```
### FEHLT Variable
nation <- dat[,6:16]
islam <- dat[,17:22]

# Islamophobie
m <- c()
for(i in 1:nrow(nation)){
  m[i] <- ifelse(any(is.na(nation[i,]) > 0), 1, 0)
}
dat$FEHLT_nation <- m

# Nationalbewusstsein
m <- c()
for(i in 1:nrow(islam)){
  m[i] <- ifelse(any(is.na(islam[i,]) > 0), 1, 0)
}
dat$FEHLT_islam <- m

names(dat)
t_dat <- dat %>%
  select(respid, gender, age, isced97, FEHLT_nation, FEHLT_islam)

head(t_dat)

### Bildung
library(tidyr)

gg_bildung <- t_dat %>%
  select(isced97, FEHLT_nation, FEHLT_islam) %>%
  gather(key, value, -isced97) %>%
```

```

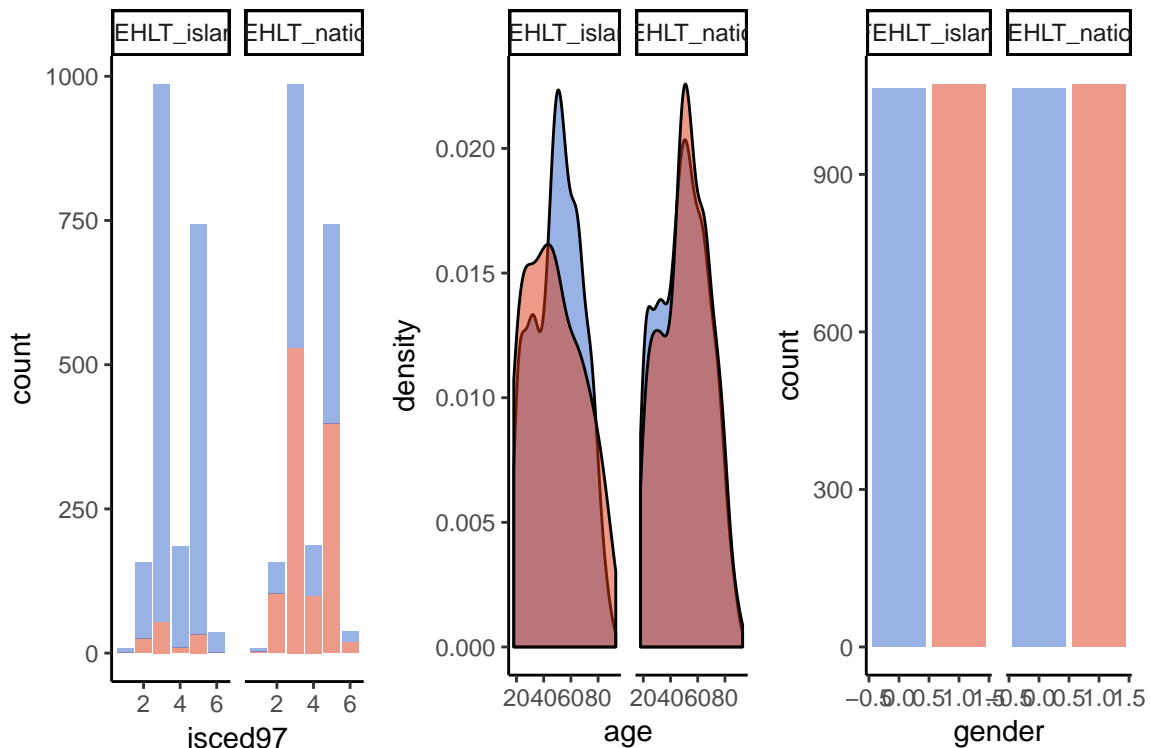
ggplot(aes(isced97, fill = factor(value))) +
  geom_bar(alpha = .5) +
  scale_fill_gdocs(name = "Missing Value", labels = c("No", "Yes")) +
  theme_classic() +
  theme(legend.position = "none") +
  facet_wrap(~ key)

gg_age <- t_dat %>%
  select(age, FEHLT_nation, FEHLT_islam) %>%
  gather(key, value, -age) %>%
  ggplot(aes(age, fill = factor(value))) +
  geom_density(alpha = .5) +
  scale_fill_gdocs(name = "Missing Value", labels = c("No", "Yes")) +
  theme_classic() +
  theme(legend.position = "none") +
  facet_wrap(~ key)

gg_gender <- t_dat %>%
  select(gender, FEHLT_nation, FEHLT_islam) %>%
  gather(key, value, -gender) %>%
  ggplot(aes(gender, fill = factor(gender))) +
  geom_bar(alpha = .5) +
  scale_fill_gdocs(name = "Missing Value", labels = c("No", "Yes")) +
  theme_classic() +
  theme(legend.position = "none") +
  facet_wrap(~ key)

library(gridExtra)
grid.arrange(gg_bildung, gg_age, gg_gender, ncol = 3)

```



```

### isced
# fit_isced_nation <-
summary(aov(FEHLT_nation ~ isced97, data=dat))
# fit_isced_islam <-
summary(aov(FEHLT_islam ~ isced97, data=dat))

### age
# fit_age_nation <-
print(summary(aov(FEHLT_nation ~ age, data=dat)))
# fit_age_islam <-
print(summary(aov(FEHLT_islam ~ age, data=dat)))

### gender
# fit_gender_nation <-
print(t.test(FEHLT_nation ~ gender, data = dat))
# fit_gender_islam <-
print(t.test(FEHLT_islam ~ gender, data = dat))

```

2.5 2.3 Logistic Regression

Table 2: Logistic Regression for Missing Values

	<i>Dependent variable:</i>	
	FEHLT_nation	FEHLT_islam
	(1)	(2)
gender	1.007 (1.092)	1.066 (1.206)
age	1.005 (1.003)	0.994 (1.005)
isc97	0.949 (1.042)	0.727 (1.096)
Constant	1.131 (1.239)	0.256 (1.543)
Observations	2,116	2,116
Log Likelihood	-1,455.950	-464.681
Akaike Inf. Crit.	2,919.900	937.362
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01	