

# SEM 2

*Simon Roth*

*02.05.2017*

## 0.1 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 ZUGEHOER.
- 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

## 1 Daten partitionieren

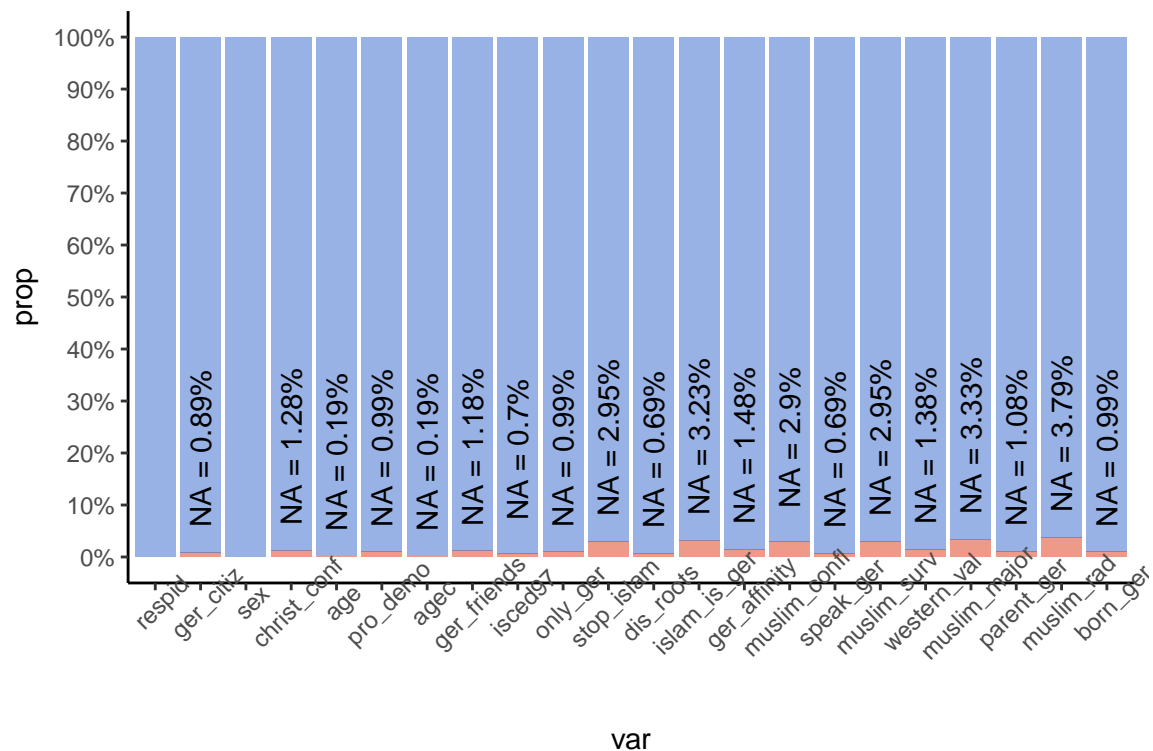
- sex: GESCHLECHT (Int.: Geschlecht der befragten Person ohne Befragen eintragen!)
  - 1 Männlich
  - 2 Weiblich
- age: ALTER: metrisch
- agec: ALTER: KATEGORISIERT 6
  - 18 - 29 Jahre
  - 30 - 44 Jahre
  - 45 - 59 Jahre
  - 60 - 74 Jahre
  - 75 - 89 Jahre
  - Über 89 Jahre

- 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?

## 1.1 Missing Pattern

Wie deutlich an den unten stehenden Schaubildern abzulesen ist, hat die Itembatterie Nationalbewusstsein deutlich mehr missing values. Dies ist auf einen Fragebogensplit (NZ) zurückzuführen, wobei einer befragten Person aufgrund eines methodischen oder inhaltlichen Splits eine Frage oder Itembatterie nicht gestellt worden ist. An dieser Stelle wäre zu diskutieren ob Personen welche nicht befragt worden sind ausgeschlossen oder gesondert betrachtet werden sollten. Klar ist, imputiert werden könne diese Fälle nicht, doch könnte es allein aufgrund des Splits zu Verzerrungen des Samples kommen. Der durchschnittliche Anteil an Item-Nonresponse beträgt zwischen 2 und 4 %. Der Nutzen oder “Schaden” den eine Imputation der fehlenden Werte nach sich ziehen könnte, bleibt wegen der geringen Ausfälle vermutlich gering.

fehlt prop var ord mis 1 FALSE 100.00 respid 1 0.00 2 FALSE 100.00 sex 2 0.00 3 FALSE 99.81 age 3 0.19 4 FALSE 99.81 agec 4 0.19 5 FALSE 99.30 isced97 5 0.70 6 FALSE 97.05 stop\_islam 6 2.95 7 FALSE 96.77 islam\_is\_ger 7 3.23 8 FALSE 97.10 muslim\_confl 8 2.90 9 FALSE 97.05 muslim\_surv 9 2.95 10 FALSE 96.67 muslim\_major 10 3.33 11 FALSE 96.21 muslim\_rad 11 3.79 12 FALSE 99.11 ger\_citiz 1 0.89 13 FALSE 98.72 christ\_conf 2 1.28 14 FALSE 99.01 pro\_demo 3 0.99 15 FALSE 98.82 ger\_friends 4 1.18 16 FALSE 99.01 only\_ger 5 0.99 17 FALSE 99.31 dis\_roots 6 0.69 18 FALSE 98.52 ger\_affinity 7 1.48 19 FALSE 99.31 speak\_ger 8 0.69 20 FALSE 98.62 western\_val 9 1.38 21 FALSE 98.92 parent\_ger 10 1.08 22 FALSE 99.01 born\_ger 11 0.99



## 1.2 1.1 Variablen FEHLT\_\_

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

# Islamophobie
m <- c()
mn <- c()
for(i in 1:nrow(nation)){
  if(dat$sp_nation[i] == 0){
    mn[i] <- sum(is.na(nation[i,]))
    m[i] <- ifelse(any(is.na(nation[i,]) > 0), 1, 0)
  } else {
    mn[i] <- NA
    m[i] <- NA
  }
}

as.data.frame(table(mn))
```

```
mn Freq 1 0 978 2 1 22 3 2 6 4 3 1 5 4 1 6 11 7
```

```
dat$FEHLT_nation <- m

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

as.data.frame(table(mn))
```

```
mn Freq 1 0 2009 2 1 57 3 2 11 4 3 4 5 4 3 6 6 51
```

```
dat$FEHLT_islam <- m
```

Die Variablen `Fehlt_nation/_islam` sind beides Indikatoren für missing values der ganzen Item-Batterie. Doch können wir die Informationen einfach so aggregieren? Personen die nur 1 mal oder auch 6-11 mal nicht geantwortet haben werden in einen Topf gesteckt, wobei die Motive sehr unterschiedlich sein könnten. Z.B. bei Fragen zum Nationalbewusstsein haben mehr Personen nur eine Frage nicht beantwortet, gegenüber denen die die Antwort komplett verweigert haben.

## 1.3 2.1 Little Test

*A Test of Missing Completely at Random for Multivariate Data with Missing Values* Roderick J. A. Little (1988). Eine Weiterentwicklung findet sich bei Jamshidian and Jalal (2010), die die MCAR Annahme basierend auf dem Vergleich der Kovarianzen zwischen Gruppen. Die Gruppen bilden sich aus gleichen Missing Data Patterns.

**Nullhypothesis: missing values pattern are missing completely at random**

```
library(MissMech)

p1 <- TestMCARNormality(islam)
```

Warning: 51 Cases with all variables missing have been removed

from the data.

```
p1$pnormality
```

```
[1] 0.01141182
```

```
p2 <- TestMCARNormality(nation)
```

Warning: 1127 Cases with all variables missing have been removed

from the data.

Warning: More than one missing data pattern should be present.

```
## Error in TestMCARNormality(nation):
```

Der Test auf auf MCAR ergibt eine signifikante Abweichung ( $p = 0.01$ ) von einer multivariat und normalverteilten Streuung der Missing Values für die Itembatterie Islamophobie. Damit muss die Annahme von MCAR zurück gewiesen werden. An dieser Stelle kann leider kein weiterer MCAR Test für die Itembatterie Nationalbewusstsein durchgeführt werden, da neben dem Item-Non-Response aufgrund des Splits keine weiteren Missing Data Pattern in den Daten zu finden sind.

## 1.4 2.3 Logistic Regression

Die logistische Regression weist auf keinen signifikanten Effekt von Bildung, Alter und Geschlecht hin. Zwar weist die Variable ISCED, einen visuell wahrnehmbaren Unterschied bei der Beantwortung der Islamophobie Item-Batterie auf, dennoch besitzt sie kaum Vorhersagekraft (hoher SE). Ein Vergleich der AIC für Modell 1 und 2 bestätigt diese Vermutung und verringert sich bei Hinzunahme von Bildung nur marginal. Neben schwachen Tendenzen bei der Bildung gibt es keine weiteren Assoziationen zwischen den hier diskutierten sozioökonomischen Variablen und Missings in der Item-Batterie für Islamophobie. Im Gegensatz dazu gibt es keine relevanten Effekte bei der Item-batterie Nationalbewusstsein.

Table 1: Logistic Regression for Missing Values

|                | FEHLT_islam      |                  | FEHLT_nation     |
|----------------|------------------|------------------|------------------|
|                | Model 1          | Model 2          | Model 3          |
| gender         | 1.165<br>(1.202) | 1.066<br>(1.206) | 0.878<br>(1.404) |
| age            | 0.993<br>(1.005) | 0.994<br>(1.005) | 1.007<br>(1.009) |
| isc97          |                  | 0.727<br>(1.096) | 0.773<br>(1.177) |
| Constant       | 0.082<br>(1.336) | 0.256<br>(1.543) | 0.072<br>(2.241) |
| N              | 2131             | 2116             | 1003             |
| Log Likelihood | -477.271         | -464.681         | -156.800         |
| AIC            | 960.541          | 937.362          | 321.599          |

\*\*\* $p < .01$ ; \*\* $p < .05$ ; \* $p < .1$