

R Auswertung

Boris T. Podzeit, Yasir Khan

11. Juni 2017

R Auswertung

Boris T. Podzeit, Yasir Khan

Inhaltsverzeichnis

Forschungsdesign und Methode	2
Hypothesen	2
Variablen	3
Befragung	3
Fragebogen	3
Kodierung	3
Empirie	3
Der Datensatz	3
Übersicht NA	4
Demographische Daten	5
Die abhängige Variable	7
Hypothese 1	11
Hypothese 2	11
Hypothese 3	11
Hypothese 4	11
Hypothese 5	11
Hypothese 6	11
Anhang	11

Abstract

Abstract hierher ...

Forschungsdesign und Methode

Hypothesen

Nr.	Hypothese	Dimension
1	LeserInnen von Gratiszeitungen sind AusländerInnen gegenüber eher negativ eingestellt.	
2	Extrovertierte Menschen haben eine positivere Wahrnehmung von Migranten als introvertierte Menschen.	
3	Je mehr Menschen im persönlichen Umfeld (Freunde, Familie) ausländerfeindlich sind, umso negativer ist die eigene Haltung gegenüber AusländerInnen.	
4	Je höher die Zufriedenheit der Studenten mit der Diversität der Studierenden auf der Wirtschafts-universität, umso positiver die eigene Haltung gegenüber AusländerInnen.	
5	Je mehr Kontakt zu ausländischen Mitbürgern besteht, desto besser sind die Einstellungen AusländerInnen gegenüber.	
6	Männer sind fremdenfeindlicher als Frauen.	

Tabelle 1: Übersicht der Hypothesen

Variablen

Allen Hypothesen liegt die gleiche abhängige Variable (AV) zugrunde, diese lautet: "Ausmaß der Fremdenfeindlichkeit". Es wird der Einfluss verschiedener Faktoren auf die Stärke der Fremdenfeindlichkeit untersucht. Aus diesem Grund wird dieselbe AV in der Untersuchung aller Hypothesen zur Verwendung kommen.

Die unabhängigen Variablen (UV) sind für jede Hypothese unterschiedlich gewählt. Im demographischen Teil wurde auf Fragen zu persönlichen Merkmalen bis auf das Geschlecht verzichtet um den Fragebogen möglichst anonym zu halten.

Befragung

Die Befragung wurde schriftlich und im Zeitraum vom 10.5. bis 22.5. mittels Fragebogen bei insgesamt 100 Studierenden der Wirtschaftsuniversität Wien durchgeführt. Die Fragebögen wurden am Campus in den Gebäuden D2 und TC bei den für die Studenten vorgesehenen Lernplätzen und -räumen verteilt. Die Bereitschaft den Fragebogen auszufüllen war hoch, besonders als das Thema "Migration" erwähnt wurde. Die Befragten waren ausschließlich StudentInnen der WU, die Grundgesamtheit besteht daher aus allen Studierenden der WU.

Es wurde ein einmaliger Pretest mit 3 Personen durchgeführt um Feedback zur Qualität des Fragebogens zu erhalten. Der Test hat zu Anpassungen und Präzisierungen bei der Fragenformulierung geführt. Das grundlegende Design wurde positiv aufgenommen. Alle Testpersonen sind in keinem persönlichen Naheverhältnis gestanden.

Fragebogen

In ausgedruckter Form besteht der Fragebogen aus 4 A4-Seiten und umfasst 7 Fragen die größtenteils mit Hilfe von Item-Batterien in Form von Likert-Skalen erhoben wurden. Die der Likert-Skala zugrunde liegenden Intervallskala waren sechsstufig. Die Antwortmöglichkeiten waren die beiden Extrempole "Trifft sehr zu" (1) und "Trifft gar nicht zu" (6) mit dazwischenliegenden Werte die mit Zahlen (2,3,4,5) ohne Beschriftung angegeben worden sind um einerseits den Fragebogen optisch nicht zu überladen und andererseits subjektiv wahrgenommene Unterschiede bei Abstufungen in Textform zu vermeiden.

Kodierung

Beschreibung der Kodierung

Empirie

In diesem Bereich wird der Datensatz aus der Befragung mit verschiedenen statistischen Verfahren ausgewertet. Allen statistischen Untersuchungen wird ein Signifikanzniveau von 5% ($\alpha=0,05$) zugrunde gelegt.

Der Datensatz

```
# Einlesen der Daten
fragebogen <- read.csv("./fragebogen.csv")
```

Ausgabe der ersten 5 Zeilen:

Nr.	wu	publikum	spass	wohl	reden	party	kurier	presse	krone	oesterreich	standard
1	3		5	2	2	3	3	4	3	4	4
2	3		4	2	3	1	3	1	4	3	4
3	2		5	1	5	1	1	1	2	3	2
4	3		6	2	2	4	4	4	4	4	3
5	1		3	2	2	2	4	3	3	2	4

heute	wiener	kleine	orf	zeit	bild	andere	diskut	scherzen	erfahrung	demo	gewalt
4	4	4	1	4	4	4	3	2		4	3
3	4	4	2	4	4	4	3	4		2	5
4	4	3	1	4	4	4	5	5		5	5
4	4	4	3	3	4	1	2	4		6	6
4	4	3	2	4	4	4	4	1		5	3

krimi	fpoe	sicherheit	feindlich	kreis	arbeiten	skype	treffen	engag	lv	heimat	rechte
5	2	3		2	6	1	4	2	6	1	6
6	5	3		4	4	2	6	4	6	3	5
6	5	5		6	6	2	6	2	6	2	6
6	5	6		5	3	3	6	5	6	4	6
3	1	2		3	2	1	6	5	6	2	6

kultur	verlassen	partner	knapp	pflegen	politik	verdienen	egal	sex	var45
6	6	6	6	6	6	6	6	2	NA
5	6	6	2	4	5	3	6	2	NA
6	6	6	3	4	6	4	6	2	NA
6	6	6	6	6	6	6	4	2	NA
6	6	6	3	4	5	5	6	2	NA

Tabelle 2: Ein Auszug der Daten

Übersicht NA

```
na_count <- sapply(fragebogen[-1,], function(y) sum(is.na(y)))
kable(na_count, format="latex", digits=2, longtable=TRUE)
```

Nr.	0
wu	0
publikum	0
spass	0
wohl	0
reden	0
party	0

kurier	0
presse	1
krone	0
oesterreich	0
standard	0
heute	0
wiener	1
kleine	1
orf	1
zeit	1
bild	1
andere	18
diskut	0
scherzen	1
erfahrung	0
demo	0
gewalt	2
krimi	1
fpoe	1
sicherheit	2
feindlich	2
kreis	2
arbeiten	2
skype	2
treffen	2
engag	2
lv	2
heimat	2
rechte	2
kultur	2
verlassen	2
partner	0
knapp	0
pflegen	0
politik	0
verdienen	0
egal	0
sex	0
var45	87

Tabelle 3: Übersicht der NA im Datensatz

```
# Summe aller NA im Datensatz
sum(is.na(fragebogen))
```

```
## [1] 141
```

Demographische Daten

```
# Einlesen der Spalte "Sex"
sex_count <- table(fragebogen$sex)
```

```
# Zuweisen von lesbaren Labels zu den Werten der Spalte
rownames(sex_count) <- c("männlich", "weiblich")
# Balkendiagramm erzeugen
barplot(sex_count, main="Geschlecht", ylim=c(0,80), xlab="Frage: 'Ihr Geschlecht?'", ylab="Anzahl", col=
box(which="figure", lty="solid", col="black")
```

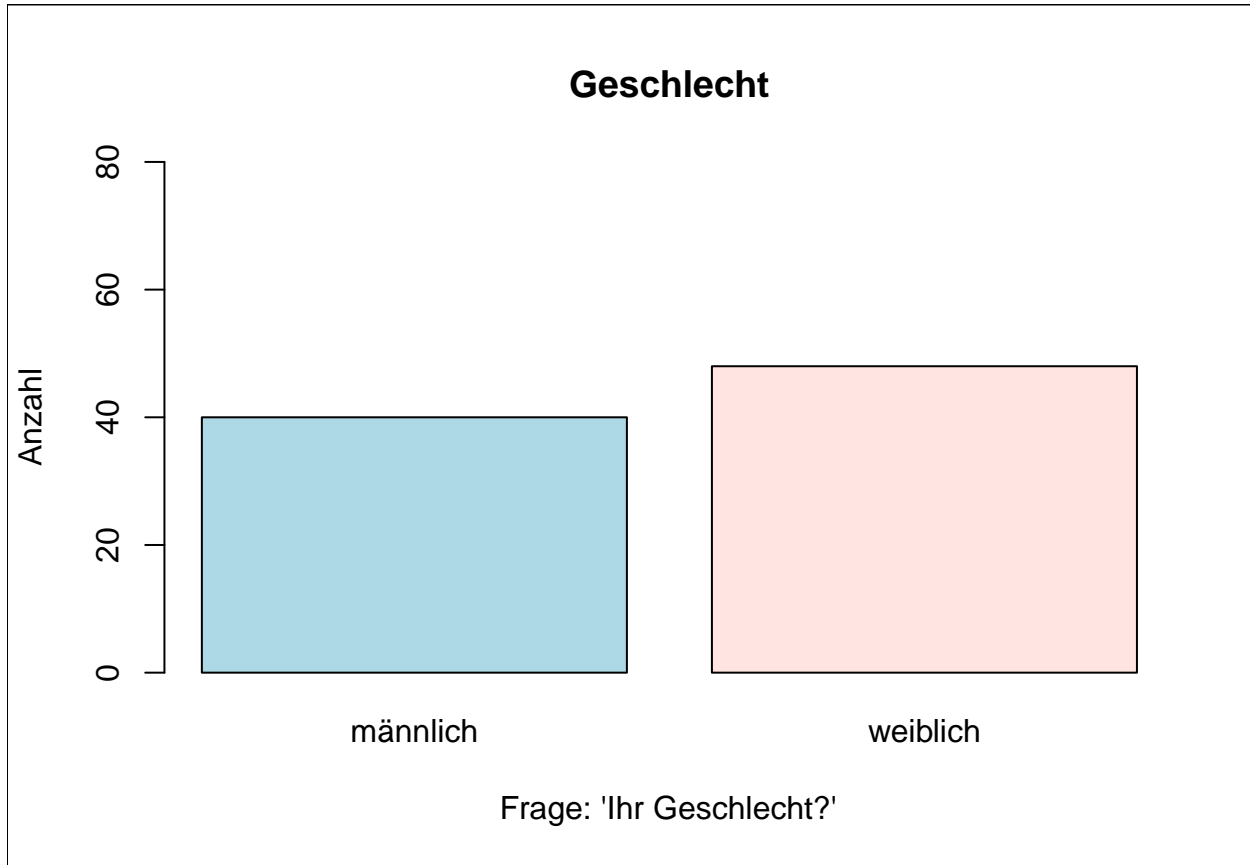


Abbildung 1: Frage nach Geschlecht

In Abbildung 2.1 ist die Verteilung der Befragten nach ihrem Geschlecht dargestellt. Dabei ist zu erkennen, dass sich unter den Befragten Personen 57 männliche Studierende und 39 weibliche Studierende befinden. Dies deutet auf ein Ungleichgewicht hin. Ob die Stichprobe dennoch repräsentativ ist, kann festgestellt werden, wenn wir die Verteilung dieser Stichprobe mit der Verteilung der Grundgesamtheit vergleichen. Die Statistik Austria gibt an, dass im Studienjahr 2013/14 insgesamt 21.157 Personen an der WU studiert haben. 11.137 davon sind männliche und 10.020 sind weibliche Studierende. Wenn wir davon ausgehen, dass sich dieses Verhältnis im Studienjahr 2014/15 nicht verändert hat, ergibt dies eine Verteilung von etwa 52% männlichen und 48% weiblichen Studierenden. Mit einem Chi-Quadrat Test können wir feststellen, ob sich unsere Stichprobe von der erwarteten Verteilung der Grundgesamtheit unterscheidet. Die beobachteten und erwarteten Werte der Verteilung der Geschlechter sehen wir in Tabelle 4.1. Für den Chi-Quadrat Test werden folgende Hypothesen aufgestellt:

H_0 :	Die Verteilung der Geschlechter unterscheidet sich <i>nicht</i> von der erwarteten Verteilung
H_A :	Die Verteilung der Geschlechter unterscheidet sich von der erwarteten Verteilung

Chi-Quadrat Test hierher

```
chisq.test(na.omit(fragebogen$sex))
```

```
## Warning in chisq.test(na.omit(fragebogen$sex)): Chi-squared approximation
## may be incorrect

##
## Chi-squared test for given probabilities
##
## data:  na.omit(fragebogen$sex)
## X-squared = 14.118, df = 87, p-value = 1
```

Die abhängige Variable

Allen Hypothesen liegt die gleiche abhängige Variable (AV) zugrunde und lautet: "Ausmaß der Fremdenfeindlichkeit". Zur Messung der AV wurde den Studierenden eine Likertskala mit 10 Items vorgelegt. Die Antwortmöglichkeiten waren auf einer 6-stufigen Intervallskala mit den Ausprägungen "Stimme sehr zu" bis "Stimme gar nicht zu" vorgegeben. Die Frage lautete "Wie ist deine Meinung zur Migration?" (siehe im Anhang Frage 6 im Fragebogen).

```
library("psych")
av <- na.omit(fragebogen[,35:44])
library("REdaS")
```

```
## Loading required package: grid
```

```
bart_spher(av)
```

```
## Bartlett's Test of Sphericity
##
## Call: bart_spher(x = av)
##
##      X2 = 445.214
##      df = 45
## p-value < 2.22e-16
```

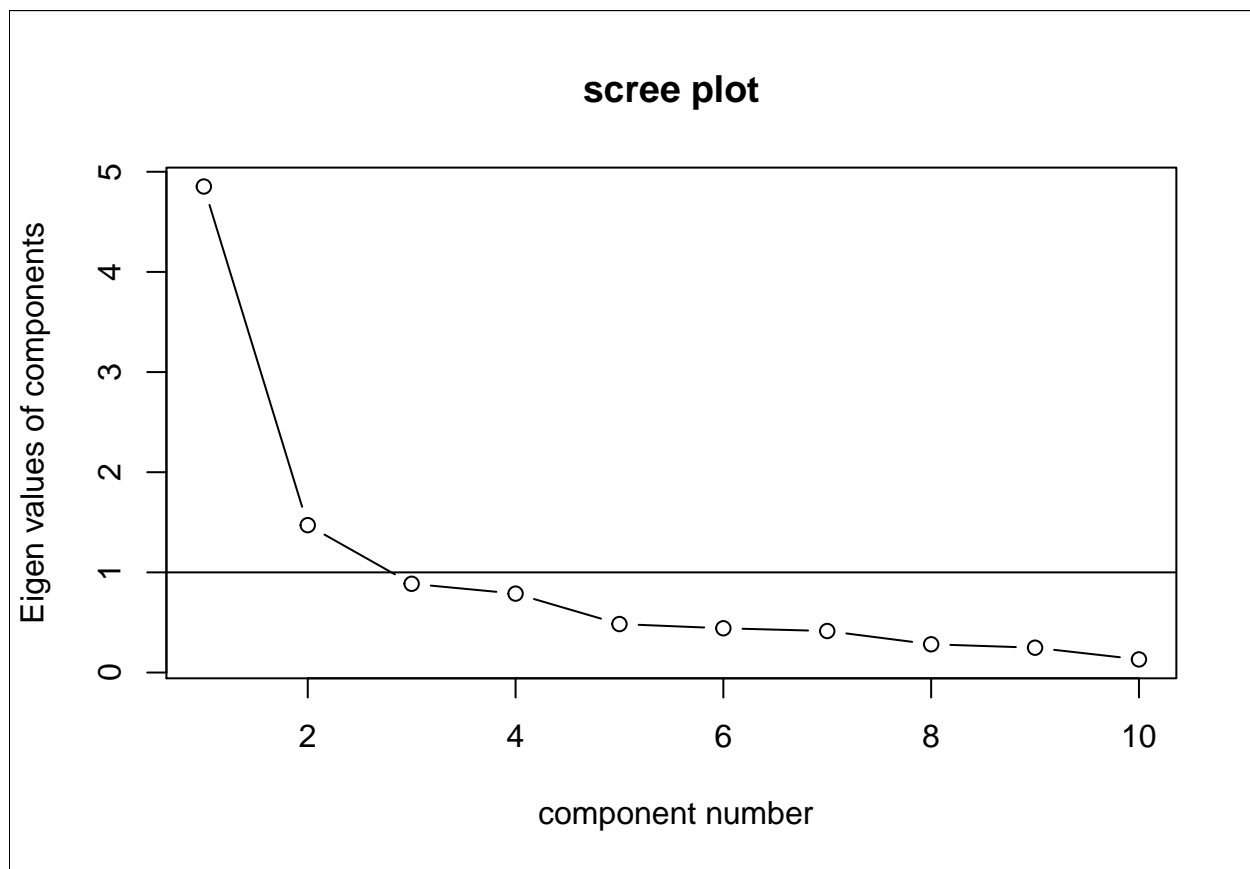
```
kmosmd <- KMOS(av)
print(kmosmd, stats="KMO")
```

```
##
## Kaiser-Meyer-Olkin Statistic
## Call: KMOS(x = av)
##
## KMO-Criterion: 0.7849755
```

```
print(kmosmd, stats="MSA", sort=TRUE, digits=3, show=1:5)
```

```
##
## Kaiser-Meyer-Olkin Statistics
##
## Call: KMOS(x = av)
##
## Measures of Sampling Adequacy (MSA):
## verlassen  pflegen    kultur    egal    partner
##      0.697      0.728      0.733      0.753      0.767
```

```
VSS.scree(av)
box(which="figure", lty="solid", col="black")
```



```
pca.av <- principal(av, 5, rotate="none")
pca.av
```

```
## Principal Components Analysis
## Call: principal(r = av, nfactors = 5, rotate = "none")
## Standardized loadings (pattern matrix) based upon correlation matrix
##      PC1  PC2  PC3  PC4  PC5  h2  u2 com
## heimat  0.80  0.18 -0.29 -0.14  0.20  0.81  0.190  1.6
## rechte  0.65  0.37 -0.01  0.43 -0.37  0.88  0.120  3.1
## kultur  0.80  0.23 -0.24 -0.19 -0.27  0.86  0.145  1.7
## verlassen 0.67  0.54 -0.27 -0.20  0.21  0.89  0.106  2.7
## partner  0.57  0.07  0.60 -0.48  0.03  0.92  0.079  2.9
## knapp   0.67 -0.40  0.02  0.35  0.32  0.84  0.164  2.8
## pflegen  0.71 -0.51  0.13 -0.18 -0.21  0.86  0.145  2.3
## politik  0.78 -0.38  0.07  0.13 -0.09  0.79  0.210  1.6
## verdienen 0.74 -0.39 -0.22  0.01  0.09  0.76  0.241  1.8
## egal     0.54  0.47  0.49  0.33  0.15  0.88  0.118  3.8
##
##      PC1  PC2  PC3  PC4  PC5
## SS loadings      4.85  1.47  0.89  0.79  0.48
## Proportion Var    0.49  0.15  0.09  0.08  0.05
## Cumulative Var    0.49  0.63  0.72  0.80  0.85
## Proportion Explained 0.57  0.17  0.10  0.09  0.06
## Cumulative Proportion 0.57  0.75  0.85  0.94  1.00
##
```



```

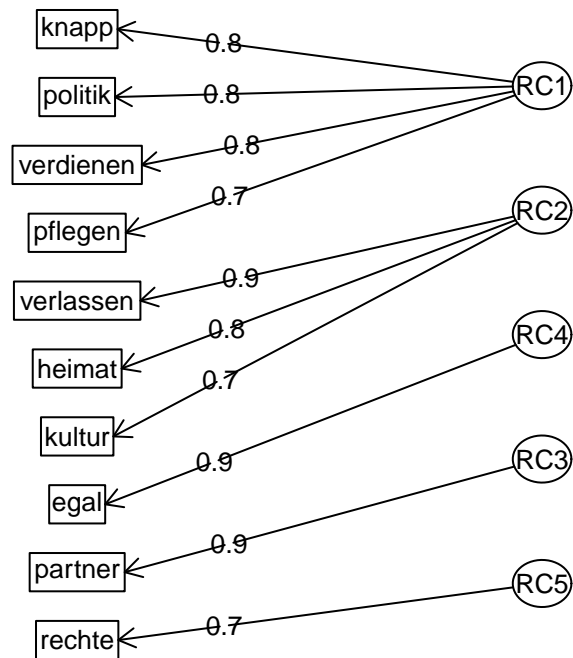
## Mean item complexity = 2.4
## Test of the hypothesis that 5 components are sufficient.
##
## The root mean square of the residuals (RMSR) is 0.06
## with the empirical chi square 23.44 with prob < 0.00028
##
## Fit based upon off diagonal values = 0.98

pca.av_r <- principal(av,5)
print(pca.av_r, cut=0.5, sort=TRUE, digits=2)

## Principal Components Analysis
## Call: principal(r = av, nfactors = 5)
## Standardized loadings (pattern matrix) based upon correlation matrix
##      item RC1  RC2  RC4  RC3  RC5  h2   u2 com
## knapp      6 0.84                0.84 0.164 1.3
## politik     8 0.79                0.79 0.210 1.5
## verdienen   9 0.79                0.76 0.241 1.4
## pflegen     7 0.74                0.86 0.145 2.1
## verlassen   4      0.90                0.89 0.106 1.2
## heimat      1      0.78                0.81 0.190 1.7
## kultur      3      0.66                0.52 0.86 0.145 2.6
## egal       10           0.86                0.88 0.118 1.4
## partner     5           0.89                0.92 0.079 1.4
## rechte      2           0.73 0.88 0.120 2.3
##
##      RC1  RC2  RC4  RC3  RC5
## SS loadings      2.84 2.19 1.21 1.21 1.03
## Proportion Var      0.28 0.22 0.12 0.12 0.10
## Cumulative Var      0.28 0.50 0.62 0.75 0.85
## Proportion Explained 0.33 0.26 0.14 0.14 0.12
## Cumulative Proportion 0.33 0.59 0.74 0.88 1.00
##
## Mean item complexity = 1.7
## Test of the hypothesis that 5 components are sufficient.
##
## The root mean square of the residuals (RMSR) is 0.06
## with the empirical chi square 23.44 with prob < 0.00028
##
## Fit based upon off diagonal values = 0.98

fa.diagram(pca.av_r, cut=0.5, cex=0.8, rsize=0.5, main="")
box(which="figure", lty="solid", col="black")

```



```
pca.av2 <- principal(av,2)
print(pca.av2, cut=0.5, sort=TRUE, digits=2)
```

```
## Principal Components Analysis
## Call: principal(r = av, nfactors = 2)
## Standardized loadings (pattern matrix) based upon correlation matrix
##      item  RC1  RC2   h2   u2 com
## pflegen    7 0.86    0.76 0.24 1.0
## politik    8 0.83    0.76 0.24 1.2
## verdienen  9 0.80    0.70 0.30 1.2
## knapp     6 0.76    0.61 0.39 1.1
## verlassen  4    0.85 0.74 0.26 1.0
## rechte     2    0.72 0.56 0.44 1.2
## kultur     3    0.72 0.69 0.31 1.6
## egal     10    0.71 0.51 0.49 1.0
## heimat     1    0.68 0.66 0.34 1.7
## partner    5    0.33 0.67 1.9
##
##              RC1  RC2
## SS loadings    3.20 3.12
## Proportion Var  0.32 0.31
## Cumulative Var  0.32 0.63
## Proportion Explained 0.51 0.49
## Cumulative Proportion 0.51 1.00
##
## Mean item complexity = 1.3
```

```
## Test of the hypothesis that 2 components are sufficient.  
##  
## The root mean square of the residuals (RMSR) is  0.08  
## with the empirical chi square 55.82 with prob < 6e-04  
##  
## Fit based upon off diagonal values = 0.96
```

Hypothese 1

Hypothese 2

Hypothese 3

Hypothese 4

Hypothese 5

Hypothese 6

Anhang

Was so alles in den Anhang kommt