Einflussfaktoren auf die Veröffentlichung von Open Data

Wolfgang Peter

2019-09-22

```
require(stpvers)
```

```
## Loading required package: stpvers
```

```
# HTML_Start("Test")
 #Projekt("html")


 Head("Überschrift 1", style=1)
```

```
## [1] "# Überschrift 1"
```

```
 Head("Überschrift 2", style=2)
```

```
## [1] "## Überschrift 2"
```

```
 Head("Überschrift 3", style=3)
```

```
## [1] "### Überschrift 3"
```

```
 Head("Überschrift 4", style=4)
```

```
## [1] "#### Überschrift 4"
```

```
 Text("Die von mir verwendeten statistischen Methoden basieren vor
      allem auf den Empfehlungen von (Bortz 2006) sowie (Sachs 2006).")
```

```
## Die von mir verwendeten statistischen Methoden basieren vor
##       allem auf den Empfehlungen von (Bortz 2006) sowie (Sachs 2006).
```

```
 Anmerkung( "Beispiel für eine Anmerkung")
```

```
## <p style="color: #0000FF"><b>Anmerkung:</b> <br>
##            Beispiel für eine Anmerkung </p><br>
```

```
  res<-  Tabelle(hkarz,  tzell, lai, gruppe, APA=TRUE)
  Output(res, add_row =c("<b>Erste Zeile</b>" = 1, "Dritte Zeile" = 3) )
```

Tabelle 1: Charakteristik

Item	n	m
Erste Zeile		
tzell (mean)	45	67.31 (6.41)
lai (mean)	45	0.42 (0.50)
Dritte Zeile		
gruppe	45	
krank		53% (24)
gesund		47% (21)

```
Tabelle2(~tzell+gruppe , hkarz)
```

Tabelle 2: Charakteristik

Item	Wert
tzell	67.31 (SD 6.41, range 48.50 to 78.50)
gruppe (krank/gesund)	24/21

```
Tabelle2(tzell~gruppe , hkarz)
```

Tabelle 3: Charakteristik

Item	gruppe	Wert
tzell	krank	63.85 (SD 5.61, range 48.50 to 73.50)
tzell	gesund	71.27 (SD 4.84, range 61.10 to 78.50)

```
#  Tabelle2(~. , hkarz)
```

```
Tabelle2(hkarz, gruppe , tzell, lai, APA=TRUE)
```

Tabelle 4: Charakteristik

Item	n	m
gruppe	45	
krank		53% (24)
gesund		47% (21)
tzell (mean)	45	67.31 (6.41)
lai (mean)	45	0.42 (0.50)

```
APA2(npk.aov <- aov(yield ~ block + N+P+K, npk))
```

Tabelle 5: ANOVA

Quelle	meansq	Df	statistic	eta.sq	eta.sq.part	p
block	68.66	5	4.29	0.39	0.59	.013
N	189.28	1	11.82	0.22	0.44	.004
P	8.40	1	0.52	0.01	0.03	.480
K	95.20	1	5.95	0.11	0.28	.028
Residuals	16.01	15		0.27		NA

```
APA_Xtabs(~lai+gruppe,hkarz)
```

```
fit<- lm(tzell~gruppe, hkarz)
APA_Table(fit)
```

```
#HTML_End()
```

Tabelle 6:

	gruppe	
lai	krank	gesund
0	47% (21)	11% (5)
1	7% (3)	36% (16)

Tabelle 7:

Quelle	b	SE
(Intercept)	63.9***	1.07
gruppegesund	7.41***	1.57
r.squared	0.34	
adj.r.squared	0.33	
AIC	281.18	
BIC	286.60	
RMSE	5.15	
Obs	45	

```
APA_Validation(fit)
```

ende Schleife Test 2 F-Statistic 3 Deviance Residuals 4 R-Squared 5 Heteroskedasticity (Breusch-Pagan) 6 Autocorrelation (Durbin-Watson) 7 Shapiro-Wilk normality test 8 AIC 9 BIC 10 Var: Residual 1 Obs. statistic 2 F(1, 43)=22.19, p<.001 3 1192.5 4 R2=.34, adj.R2=.33 5 BP(1)=0.41, p=.522 6 DW=0.45, p<.001 7 W=0.96, p=.168 8 281.2 9 286.6 10 27.73 1 45

Tabelle 8: Testing Regression Models

Test	statistic
F-Statistic	F(1, 43)=22.19, p<.001
Deviance Residuals	1192.5
R-Squared	R²=.34, adj.R²=.33
Heteroskedasticity (Breusch-Pagan)	BP(1)=0.41, p=.522
Autocorrelation (Durbin-Watson)	DW=0.45, p<.001
Shapiro-Wilk normality test	W=0.96, p=.168
AIC	281.2
BIC	286.6
Var: Residual	27.73
Obs.	45

```
pc <- psych::principal(Harman74.cor$cov,4,rotate="varimax")
mr <- psych::fa(Harman74.cor$cov,4,rotate="varimax")  #minres factor analysis
pa <- psych::fa(Harman74.cor$cov,4,rotate="varimax",fm="pa")  # principal axis factor analysis



round(psych::factor.congruence(list(pc,mr,pa)),2) %>%
  Output("The coefficient of factor congruence between two sets of factor loadings")
```

```
##
##
## Table:
##
## Quelle     RC1    RC3    RC2    RC4    MR1    MR3    MR2    MR4    PA1    PA3    PA2    PA4
## -------  -----  -----  -----  -----  -----  -----  -----  -----  -----  -----  -----  -----
## RC1       1.00   0.53   0.43   0.46   1.00   0.61   0.46   0.54   1.00   0.61   0.46   0.54
## RC3       0.53   1.00   0.43   0.47   0.54   0.99   0.44   0.54   0.54   0.99   0.44   0.54
## RC2       0.43   0.43   1.00   0.47   0.44   0.50   1.00   0.55   0.44   0.50   1.00   0.55
## RC4       0.46   0.47   0.47   1.00   0.47   0.53   0.49   0.99   0.47   0.53   0.49   0.99
## MR1       1.00   0.54   0.44   0.47   1.00   0.61   0.46   0.55   1.00   0.61   0.46   0.55
## MR3       0.61   0.99   0.50   0.53   0.61   1.00   0.50   0.61   0.61   1.00   0.50   0.61
## MR2       0.46   0.44   1.00   0.49   0.46   0.50   1.00   0.57   0.46   0.50   1.00   0.57
## MR4       0.54   0.54   0.55   0.99   0.55   0.61   0.57   1.00   0.55   0.61   0.57   1.00
## PA1       1.00   0.54   0.44   0.47   1.00   0.61   0.46   0.55   1.00   0.61   0.46   0.55
## PA3       0.61   0.99   0.50   0.53   0.61   1.00   0.50   0.61   0.61   1.00   0.50   0.61
## PA2       0.46   0.44   1.00   0.49   0.46   0.50   1.00   0.57   0.46   0.50   1.00   0.57
## PA4       0.54   0.54   0.55   0.99   0.55   0.61   0.57   1.00   0.55   0.61   0.57   1.00
##
##
```

```
APA2(pc)
```

Tabelle 9: Principal Components Analysis

Item	Nr	PC1	PC2	PC3	PC4	h2
SentenceCompletion	7	0.85				0.77
WordMeaning	9	0.84				0.78
PargraphComprehension	6	0.81				0.73
GeneralInformation	5	0.79				0.70
WordClassification	8	0.64	0.31			0.57
ProblemReasoning	22	0.42	0.41			0.45
VisualPerception	1		0.71			0.60
PaperFormBoard	3		0.66			0.47
Flags	4		0.62			0.45
Cubes	2		0.59			0.37
SeriesCompletion	23	0.42	0.52			0.55
Deduction	20	0.43	0.43			0.47
Addition	10			0.83		0.76
CountingDots	12			0.80		0.67
Code	11			0.63	0.37	0.57
StraightCurvedCapitals	13		0.41	0.62		0.59
ArithmeticProblems	24	0.40		0.55		0.55
NumericalPuzzles	21		0.42	0.50		0.49
WordRecognition	14				0.68	0.52
ObjectNumber	17				0.68	0.54
NumberRecognition	15				0.67	0.48
FigureRecognition	16		0.46		0.58	0.55
NumberFigure	18		0.32	0.40	0.50	0.51
FigureWord	19				0.42	0.30

Fit based upon off diagonal values =0.97

Tabelle 10: Erklaerte Gesamtvarianz (Eigenwerte)

Quelle	PC1	PC2	PC3	PC4
SS loadings	4.16	3.31	3.22	2.74
Proportion Var	0.17	0.14	0.13	0.11
Cumulative Var	0.17	0.31	0.45	0.56
Proportion Explained	0.31	0.25	0.24	0.20
Cumulative Proportion	0.31	0.56	0.80	1.00

Tabelle 11: Test of the hypothesis that 4 components are sufficient.

Measures	Statistic
Mean item complexity	1.7
RMSR	0.06

```
APA2(mr)
```

Tabelle 12: Factor Analysis using method = minres

Item	Nr	PC1	PC2	PC3	PC4	h2
SentenceCompletion	7	0.81				0.73
WordMeaning	9	0.81				0.74
PargraphComprehension	6	0.76				0.68
GeneralInformation	5	0.73				0.64
WordClassification	8	0.57	0.34			0.51
VisualPerception	1		0.68			0.55
PaperFormBoard	3		0.55			0.34
Flags	4		0.53			0.35
SeriesCompletion	23	0.37	0.52			0.51
Cubes	2		0.45			0.23
Deduction	20	0.38	0.42			0.42
ProblemReasoning	22	0.37	0.41			0.40
Addition	10			0.82		0.74
CountingDots	12			0.71		0.55
StraightCurvedCapitals	13		0.42	0.54		0.51
Code	11			0.54	0.37	0.47
ArithmeticProblems	24	0.36		0.49		0.49
NumericalPuzzles	21		0.40	0.43		0.42
ObjectNumber	17				0.58	0.41
WordRecognition	14				0.56	0.36
NumberRecognition	15				0.52	0.31
FigureRecognition	16		0.42		0.52	0.45
NumberFigure	18		0.31	0.34	0.45	0.41
FigureWord	19				0.35	0.23

Test of the hypothesis that4factors aresufficient.The degrees of freedom for the null model are 276 and the objective function was 11.44The degrees of freedom for the model are186 and the objective function was 1.72 The df corrected root mean square of the residuals is 0.05 Fit based upon off diagonal values =0.98 Measures of factor score adequacy

Tabelle 13: Erklaerte Gesamtvarianz (Eigenwerte)

Quelle	PC1	PC2	PC3	PC4
SS loadings	3.64	2.93	2.67	2.23
Proportion Var	0.15	0.12	0.11	0.09
Cumulative Var	0.15	0.27	0.38	0.48
Proportion Explained	0.32	0.26	0.23	0.19
Cumulative Proportion	0.32	0.57	0.81	1.00

PC1 PC2 PC3 PC4 Correlation of scores with factors 0.93 0.87 0.91 0.82 Multiple R square of scores with factors 0.87 0.76 0.83 0.68 Minimum correlation of possible factor scores 0.74 0.52 0.65 0.36

Tabelle 14: Test of the hypothesis that 4 factors are sufficient.

Measures	Statistic
Mean item complexity	1.9
RMSR	0.04

```
APA2(pa)
```

Tabelle 15: Factor Analysis using method = pa

Item	Nr	PC1	PC2	PC3	PC4	h2
SentenceCompletion	7	0.81				0.73
WordMeaning	9	0.81				0.74
PargraphComprehension	6	0.76				0.68
GeneralInformation	5	0.73				0.64
WordClassification	8	0.57	0.34			0.51
VisualPerception	1		0.68			0.55
PaperFormBoard	3		0.55			0.34
Flags	4		0.53			0.35
SeriesCompletion	23	0.37	0.52			0.51
Cubes	2		0.45			0.23
Deduction	20	0.38	0.42			0.42
ProblemReasoning	22	0.37	0.41			0.40
Addition	10			0.82		0.74
CountingDots	12			0.71		0.55
StraightCurvedCapitals	13		0.42	0.54		0.51
Code	11			0.54	0.37	0.47
ArithmeticProblems	24	0.36		0.49		0.49
NumericalPuzzles	21		0.40	0.43		0.42
ObjectNumber	17				0.58	0.41
WordRecognition	14				0.56	0.36
NumberRecognition	15				0.52	0.31
FigureRecognition	16		0.42		0.52	0.45
NumberFigure	18		0.31	0.34	0.45	0.41
FigureWord	19				0.35	0.23

Test of the hypothesis that4factors aresufficient.The degrees of freedom for the null model are 276 and the objective function was 11.44The degrees of freedom for the model are186 and the objective function was 1.72

Tabelle 16: Erklaerte Gesamtvarianz (Eigenwerte)

Quelle	PC1	PC2	PC3	PC4
SS loadings	3.64	2.93	2.67	2.23
Proportion Var	0.15	0.12	0.11	0.09
Cumulative Var	0.15	0.27	0.38	0.48
Proportion Explained	0.32	0.26	0.23	0.19
Cumulative Proportion	0.32	0.57	0.81	1.00

The df corrected root mean square of the residuals is 0.05 Fit based upon off diagonal values =0.98 Measures of factor score adequacy
PC1 PC2 PC3 PC4 Correlation of scores with factors 0.93 0.87 0.91 0.82 Multiple R square of scores with factors 0.87 0.76 0.82 0.68 Minimum correlation of possible factor scores 0.74 0.52 0.65 0.36

Tabelle 17: Test of the hypothesis that 4 factors are sufficient.

Measures	Statistic
Mean item complexity	1.9
RMSR	0.04

```
df <- data.frame(
  term = c("A", "B", "C", "D"),
  n = c(23, 14, 56, 2),
  m = c("4.7 (2.4)", "4.1 (2.3)", "8.9 (3.6)", NA)
)

df2 <-
  stp25stat::prepare_output(df,
                            caption = "Überschrift",
                            note = "Anmerkung",
                            N = 256)

res1 <-  Tabelle2(hkarz,
                  gruppe, tzell, lai, APA = TRUE, output=FALSE)
res2 <-  Tabelle2(hkarz,
                  tzell, lai, by =  ~ gruppe, APA = TRUE, output=FALSE)
res3 <- APA_Table(lm( tzell ~ lai, hkarz),
                  lm( tzell ~ lai +gruppe, hkarz),
                  names = c("M 1", "M 2"),
                  output=FALSE)

which_output()
```

[1] “markdown”

```
HTML_(
knitr::kable(df, caption ="Überschrift kable"))
```

```
Output(df2)
```

```
Output(res1)
```

Tabelle 18: Überschrift kable

term	n	m
A	23	4.7 (2.4)
B	14	4.1 (2.3)
C	56	8.9 (3.6)
D	2	NA

Tabelle 19: Überschrift

Quelle	n	m
A	23	4.7 (2.4)
B	14	4.1 (2.3)
C	56	8.9 (3.6)
D	2	

Tabelle 20: Charakteristik

Item	n	m
gruppe	45	
krank		53% (24)
gesund		47% (21)
tzell (mean)	45	67.31 (6.41)
lai (mean)	45	0.42 (0.50)

```
Output(res2)
```

Tabelle 21: gruppe Charakteristik

	krank		gesund	
Item	n	m	n	m
tzell (mean)	24	63.85 (5.61)	21	71.27 (4.84)
lai (mean)	24	0.12 (0.34)	21	0.76 (0.44)

```
Output(res3)
```

```
End()
```

Reset Kontraste

Tabelle 22:

	M 1		M 2	
Quelle	b	SE	b	SE
(Intercept)	64.2***	1.03	63.3***	1.06
lai	7.47***	1.59	4.53*	1.98
gruppegesund			4.53*	1.96
r.squared	0.34		0.41	
adj.r.squared	0.32		0.39	
AIC	281.26		277.89	
BIC	286.68		285.11	
RMSE	5.15		4.85	
Obs	45		45	