

Lösung zu Kapitel 11: Beispiel 2

Die Daten beziehen sich auf 48 Bewerbungen um eine Position in einem Unternehmen. Diese Bewerbungen wurden anhand von 15 Variablen bewertet (Datenfile: [bewerbung.csv](#)).

- Form des Bewerbungsschreibens
- Erscheinung
- Akademische Fähigkeiten
- Sympathie
- Selbstvertrauen
- Klarheit
- Ehrlichkeit
- Geschäftstüchtigkeit
- Erfahrung
- Schwung
- Ambition
- Auffassungsgabe
- Potenzial
- Eifer
- Eignung

(Je höher der Wert, desto stärker ist die Eigenschaft ausgeprägt.)

- Gibt es einige zugrunde liegende Komponenten, die mit den Haupteigenschaften der Bewerber korrespondieren?
- Falls ja, können die Bewerber leichter verglichen werden. Basierend auf den Ergebnissen der Hauptkomponentenanalyse: Welche Kandidaten würde man auswählen, wenn die zu besetzende Position im Verkauf, im Marketing oder aber in der Abteilungsleitung angesiedelt ist?

R

```
> bewerbung <- read.csv2("bewerbung.csv", header = TRUE)
> itemnam <- c("id", "Form des Bewerbungsschreibens", "Erscheinung",
+             "Akademische Fähigkeiten", "Sympathie", "Selbstvertrauen",
+             "Klarheit", "Ehrlichkeit", "Geschäftstüchtigkeit", "Erfahrung",
+             "Schwung", "Ambition", "Auffassungsgabe", "Potenzial", "Eifer",
+             "Eignung")
> colnames(bewerbung) <- itemnam
> dat.b <- bewerbung[, -1]
```

Ob die vorhandenen Daten für eine Hauptkomponentenanalyse geeignet sind, wurde vorweg mittels der Kaiser-Meyer-Olkin-Statistik überprüft. Diese berechnet das Maß der Korrelation, welches in den Daten steckt - unter Berücksichtigung des Ausmaßes an partieller Korrelation.

R

```
> library("rela")
> test.b <- paf(as.matrix(dat.b))
> cat("KMO Statistik:", test.b$KMO, " Bartlett-Statistik:", test.b$Bartlett,
+     "\n")
```

KMO Statistik: 0.78378 Bartlett-Statistik: 645.32

R

```
> library("psych")
> bart.b <- cortest.bartlett(cor(dat.b), n = nrow(dat.b))
> unlist(bart.b)
```

```
      chisq    p.value      df
6.4532e+02 1.2563e-78 1.0500e+02
```

Um die Vielzahl von Merkmalen auf einige wenige Faktoren zu reduzieren, wird eine Hauptkomponentenanalyse mit anschließender Varimax-Rotation durchgeführt, wobei die Faktoren auf Basis ihrer Korrelation erstellt werden. Werden jene Komponenten verwendet, deren Eigenwert mindestens 1 erreicht, können 81 % der Gesamtvarianz erklärt werden.

R

```
> pca.b <- principal(dat.b, 4, rotate = "none")
> pca.b$criteria <- NULL
> pca.b
```

Principal Components Analysis

Call: principal(r = dat.b, nfactors = 4, rotate = "none")

Standardized loadings based upon correlation matrix

	PC1	PC2	PC3	PC4	h2	u2
Form des Bewerbungsschreibens	0.45	0.61	0.38	-0.10	0.73	0.27
Erscheinung	0.58	-0.05	-0.03	0.29	0.43	0.57
Akademische Fähigkeiten	0.11	0.34	-0.52	0.70	0.88	0.12
Sympathie	0.62	-0.19	0.56	0.38	0.87	0.13
Selbstvertrauen	0.80	-0.36	-0.29	-0.19	0.88	0.12
Klarheit	0.86	-0.19	-0.18	-0.08	0.82	0.18
Ehrlichkeit	0.43	-0.58	0.34	0.46	0.85	0.15
Geschäftstüchtigkeit	0.89	-0.04	-0.22	-0.22	0.89	0.11
Erfahrung	0.37	0.79	0.10	0.07	0.78	0.22
Schwung	0.86	0.07	-0.10	-0.17	0.79	0.21
Ambition	0.87	-0.10	-0.25	-0.22	0.88	0.12
Auffassungsgabe	0.91	-0.03	-0.14	0.08	0.85	0.15
Potenzial	0.91	0.03	-0.09	0.21	0.89	0.11
Eifer	0.71	-0.12	0.56	-0.22	0.89	0.11
Eignung	0.65	0.60	0.11	-0.02	0.79	0.21

	PC1	PC2	PC3	PC4
SS loadings	7.51	2.06	1.46	1.20
Proportion Var	0.50	0.14	0.10	0.08
Cumulative Var	0.50	0.64	0.74	0.81

R

```
> pca.b.rotate <- principal(dat.b, 4, scores = TRUE)
> pca.b.rotate$criteria <- NULL
> print(pca.b.rotate, cut = 0.5, sort = TRUE, digits = 2)
```

Principal Components Analysis

Call: `principal(r = dat.b, nfactors = 4, scores = TRUE)`

Standardized loadings based upon correlation matrix

	item	RC1	RC2	RC3	RC4	h2	u2
Ambition	11	0.92				0.88	0.12
Selbstvertrauen	5	0.92				0.88	0.12
Geschäftstüchtigkeit	8	0.91				0.89	0.11
Klarheit	6	0.86				0.82	0.18
Auffassungsgabe	12	0.81				0.85	0.15
Schwung	10	0.80				0.79	0.21
Potenzial	13	0.75				0.89	0.11
Erscheinung	2					0.43	0.57
Erfahrung	9		0.85			0.78	0.22
Form des Bewerbungsschreibens	1		0.83			0.73	0.27
Eignung	15		0.80			0.79	0.21
Sympathie	4			0.87		0.87	0.13
Ehrlichkeit	7			0.86		0.85	0.15
Eifer	14			0.53	-0.52	0.89	0.11
Akademische Fähigkeiten	3				0.93	0.88	0.12

	RC1	RC2	RC3	RC4
SS loadings	5.77	2.73	2.39	1.35
Proportion Var	0.38	0.18	0.16	0.09
Cumulative Var	0.38	0.57	0.73	0.81

Um die Hauptkomponenten inhaltlich sinnvoll interpretieren zu können, lässt man sich zudem einen Screeplot ► Abbildung 1 als grafische Unterstützung erstellen. In diesem Fall stimmt die Anzahl der Hauptkomponenten, deren Eigenwert größer als 1 ist, mit der Anzahl der Hauptkomponenten, welche im Scree-Plot links von der Knickstelle liegen, überein.

R

```
> VSS.scree(dat.b)
```

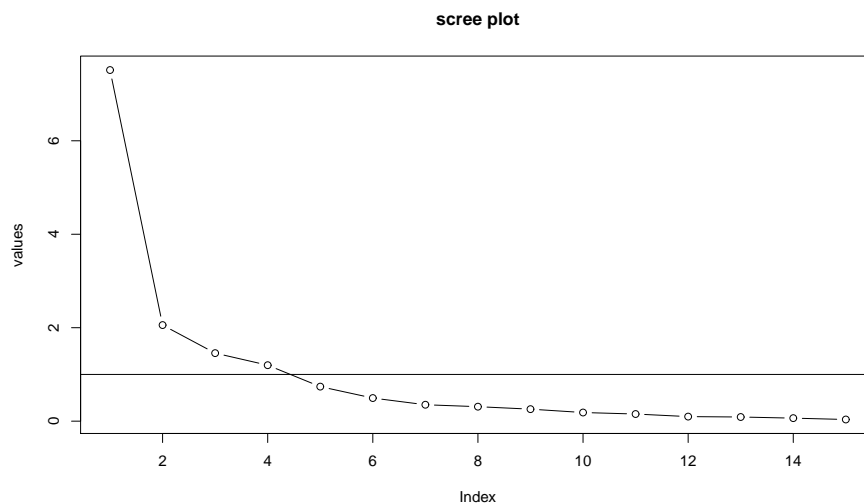
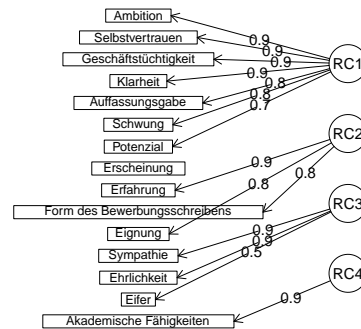


Abbildung 1: Screeplot

R

```
> fa.diagram(pca.b.rotate, cut = 0.5, cex = 0.8, rsize = 0.5, main = "")
```



Nachdem die neu gewonnenen Variablen benannt wurden

- FAKTOR 1 ...Softskills
- FAKTOR 2 ...Fachliche Kompetenz
- FAKTOR 3 ...Soziale Kompetenz
- FAKTOR 4 ...Ausbildung

wird mit diesen weitergerechnet.

Im Folgenden soll nun mit Hilfe der Ergebnisse eine Auswahl der Kandidaten für eine zu besetzende Position

1. im Verkauf,
2. im Marketing oder
3. in der Abteilungsleitung

erfolgen.

Eine Möglichkeit, diese Auswahl zu treffen, besteht darin, die Eignung der Bewerber für die aus-geschriebenen Stellen anhand der wichtigsten Faktoren zu bestimmen. Dafür werden die jeweilige Positionen auf den beiden Achsen eines Streudiagrammes aufgetragen. Je weiter rechts oben die Werte der BewerberInnen liegen, desto besser eignen sich diese für die jeweilige Stelle.

Welche Faktoren für die jeweilige Position als Grundlage zur Entscheidungsfindung herangezogen werden, ist eine subjektive Entscheidung. Unter den Annahmen, dass

1. für die Position im Verkauf die beiden Faktoren **Softskills** und **Soziale Kompetenz**,
2. für die Position im Marketing die beiden Faktoren **Fachliche Kompetenz** und **Ausbildung** und
3. für die Position in der Abteilungsleitung die beiden Faktoren **Fachliche Kompetenz** und **Soziale Kompetenz**

am wichtigsten sind, können die BewerberInnen nun leichter miteinander verglichen werden. Wenn beispielsweise jemand im Verkauf eingestellt werden soll, werden die Faktoren 1 und 3 in einem Streudiagramm einander gegenüber gestellt. Somit gelangt man zu den folgenden Entscheidungen:

Folgenden BewerberInnen sind für die jeweiligen Stellen am geeignetsten:

```

> pca.b.scores <- data.frame(pca.b.rotate$scores)
> colnames(pca.b.scores) <- c("Softskills", "Fachliche_Kompetenz",
+   "Soziale_Kompetenz", "Ausbildung")
> attach(pca.b.scores)
> plot(Softskills, Soziale_Kompetenz)
> abline(v = 0, h = 0, col = "blue", lty = "dotted")
> text(Softskills, Soziale_Kompetenz, labels = bewerbung$id, pos = 2)

```

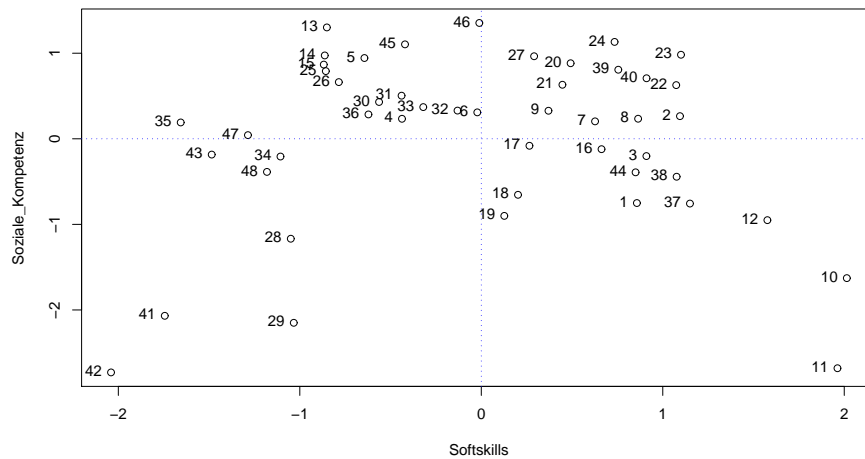


Abbildung 2: Empfehlung für die Position im Verkauf

1. Verkauf: 23 und 24
2. Marketing: 41 und 42
3. Abteilungsleitung: 39 und 40

R

```
> plot(Fachliche_Kompetenz, Ausbildung)
> abline(v = 0, h = 0, col = "blue", lty = "dotted")
> text(Fachliche_Kompetenz, Ausbildung, labels = bewerbung$id,
+      pos = 2)
```

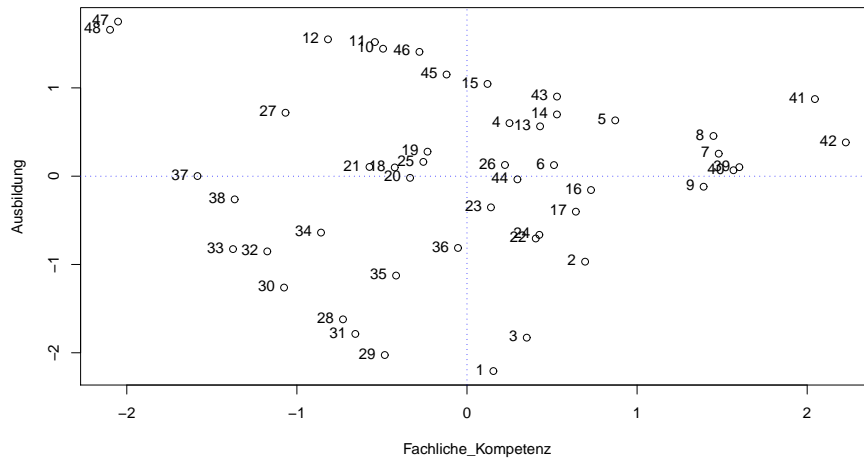


Abbildung 3: Empfehlung für die Position im Marketing

R

```
> plot(Fachliche_Kompetenz, Soziale_Kompetenz)
> abline(v = 0, h = 0, col = "blue", lty = "dotted")
> text(Fachliche_Kompetenz, Soziale_Kompetenz, labels = bewerbung$id,
+      pos = 2)
> detach(pca.b.scores)
```

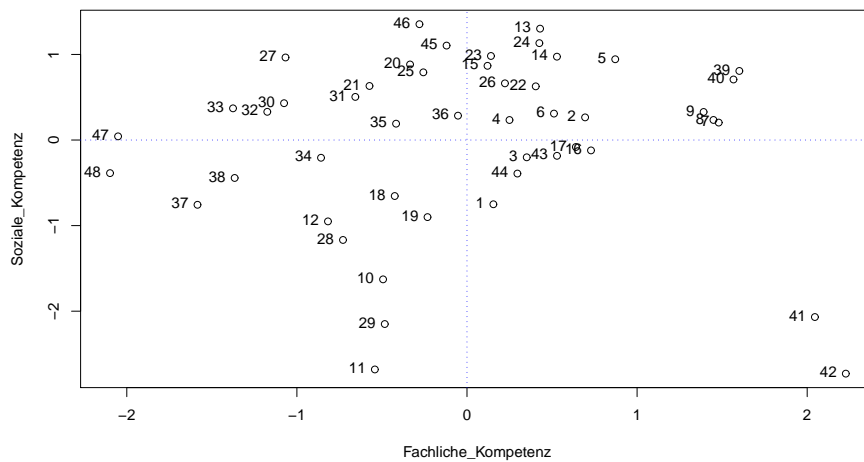


Abbildung 4: Empfehlung für die Position in der Abteilungsleitung