



Statistik II: Klassifikation und Segmentierung

Dr. Andreas Vlašić

Medien Institut

☎ (0621) 52 67 44

💻 vlasic@medien-institut.de



Gliederung

- ▶ **1. Faktorenanalyse**
- 2. Clusteranalyse
- 3. Key Facts

Ziel der Faktorenanalyse

- Ziel der Faktorenanalyse ist es, eine Vielzahl von Variablen durch möglichst wenige Faktoren abzubilden
- Faktorenanalyse basiert auf den Interkorrelationen der Variablen
 - Zusammenfassen von Variablen, die *untereinander stark* und gleichzeitig *mit anderen Variablen schwach* korrelieren
 - ⇒ Extraktion *rechnerisch* und *inhaltlich* möglichst unabhängiger Faktoren
- Grundsätzlich zu klären:
 - Anzahl der Faktoren?
 - Inhaltliche Interpretation der Faktoren?

Ein Beispiel

- In einer Befragung wurden vier Statements erhoben
 - S1: Ich will in Prüfungen immer möglichst gut abschneiden.
 - S2: Ich möchte später einmal eine Arbeit haben, in der ich meine Fähigkeiten verwirklichen kann
 - S3: Ich liebe Country-Music
 - S4: Ich gehe gerne abends im Sonnenschein am Strand der Parkinsel spazieren.

- Die Zustimmung zu diesen Statements ergibt folgende Korrelationen:

	v_pruef	v_job	v_country	v_lupark
v_pruef	1,00			
v_job	0,80	1,00		
v_country	0,10	0,15	1,00	
v_lupark	- 0,05	0,05	0,70	1,00

F1: „Leistungsmotiv, Karrierewunsch...“

F2: „Romantische Ader, Sentimentalität...“

Grundzüge der Faktorenanalyse

- Zweck der Faktorenanalyse:
 - Reduktion von Daten
 - Überprüfung der Dimensionalität komplexer Merkmale
 - Hypothesengenerierung

- Grundlegendes Vorgehen der Faktorenanalyse
 1. Konstruktion einer synthetischen Variable (Faktor), die mit allen anderen Variablen so hoch wie möglich korreliert
 - ⇒ Erklärung eines Teils der Interkorrelationen
 2. Konstruktion einer weiteren synthetischen Variable, um die verbleibenden Interkorrelationen zu erklären
 - ⇒ Erklärung eines weiteren Teils der Interkorrelationen
 3. ...

Vorgehen der Faktorenanalyse

1. Erstellung der Korrelationsmatrizen

- Überprüfung der Variablen auf Eignung
 - Kaiser-Meyer-Olkin-Maß (KMO)
 - Anti-Image-Korrelationsmatrix (MSA)
 - Bartlett-Test auf Sphärizität

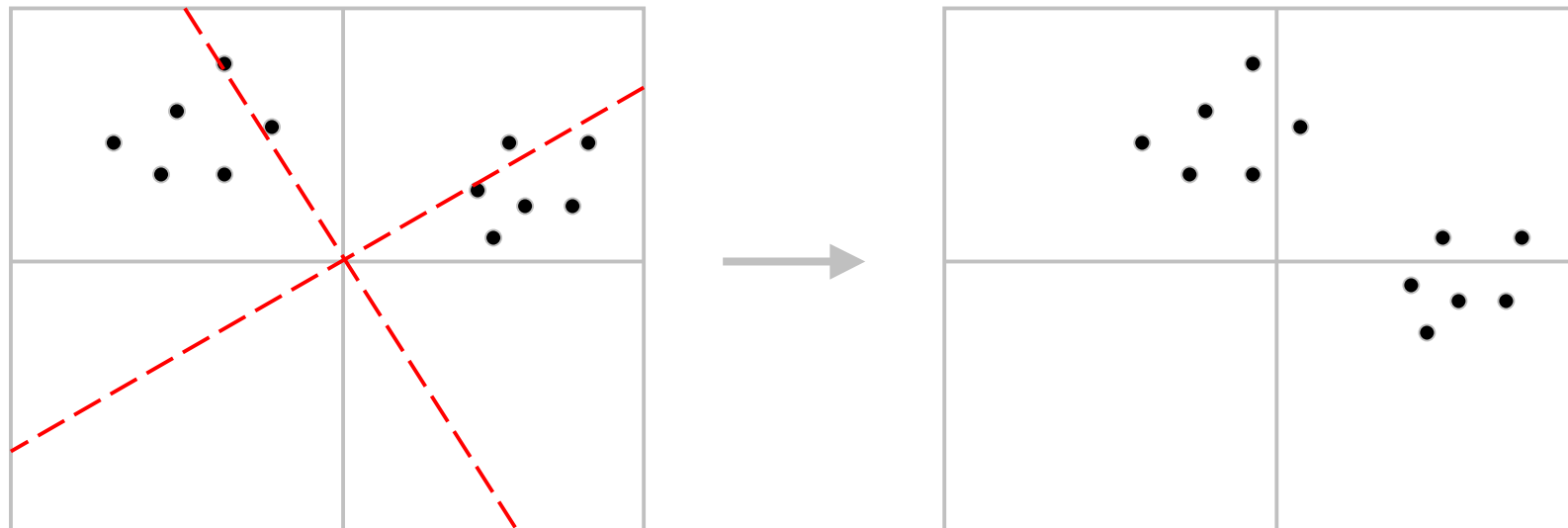
2. Extraktion der Faktoren

- Verschiedene Methoden, am gängigsten: *Hauptkomponentenanalyse*
- *Kommunalität*: Betrag der erklärten Streuung *einer Variable* durch *alle Faktoren*
- *Eigenwert* eines Faktors: Betrag der erklärten Streuung *aller Variablen* durch *einen Faktor* (⇒ aufgeklärte Varianz)
- Bestimmung der Anzahl von Faktoren: Theoretisch oder rechnerisch (Kaiser-Kriterium)
- *Komponentenmatrix*: Faktorladungen

Vorgehen der Faktorenanalyse /2

3. Rotation

- Rotation der Faktorladungsmatrix bzw. der Achsen des Koordinatensystems erleichtert Interpretation
- *Varimax*-Rotation: Orthogonale (rechtwinklige) Rotation

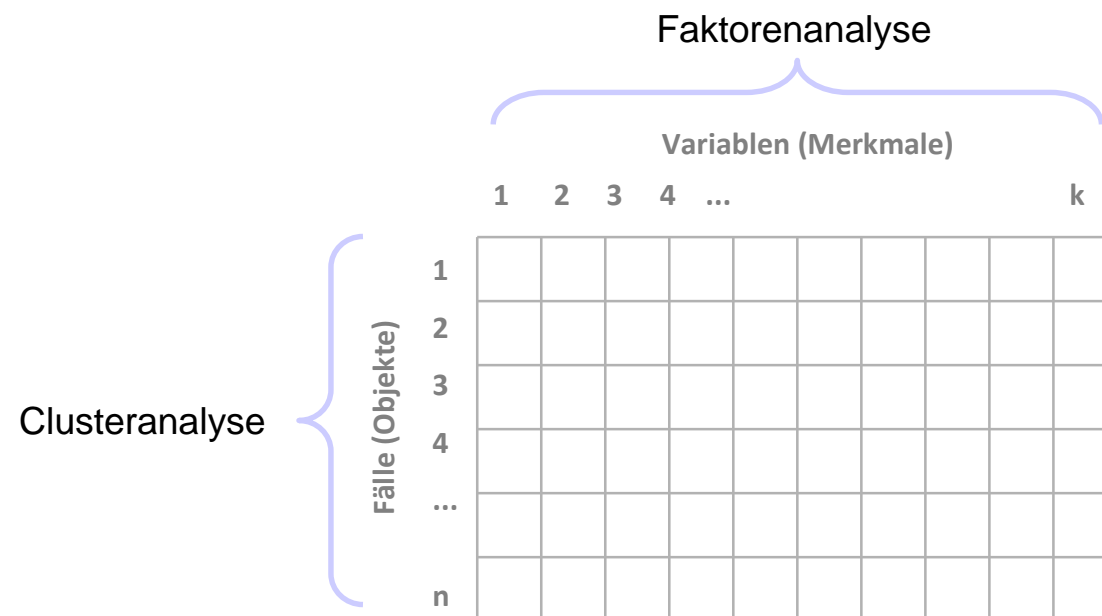


Gliederung

1. Faktorenanalyse
- ▶ 2. **Clusteranalyse**
3. Key Facts

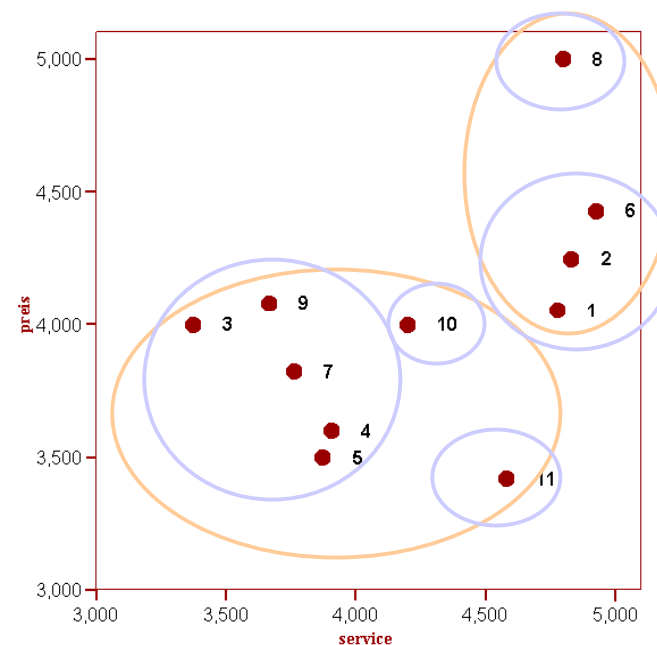
Ziel der Clusteranalyse

- Ziel der Clusteranalyse ist es, eine Vielzahl von Fällen zu Gruppen zusammen zu fassen, die nach innen homogen und nach außen heterogen sind



Ein Beispiel

- In einer Befragung wurden für die 11 Automarken die Wahrnehmung verschiedener Leistungsaspekte erhoben, u.a. das Preis-Leistungs-Verhältnis und die Servicequalität
- Die Ergebnisse der Marken werden auf einem Streudiagramm abgebildet: Welche sinnvollen Gruppen lassen sich bilden?
- Fragen:
 - Vorgehen/Algorithmus?
 - Überschneidungsfreiheit?
 - Maß für Ähnlichkeit?
 - Anzahl der Gruppen/Cluster?



Gliederung

1. Faktorenanalyse
2. Clusteranalyse
- ▶ 3. **Key Facts**

Key Facts

- Was ist das Ziel der Faktorenanalyse?
- Welche drei Schritte umfasst das Verfahren?
- Was ist der Eigenwert der Faktoren, was die Kommunalität?
- Welche Voraussetzungen hat die Durchführung einer Faktorenanalyse?
- Was ist bei der Auswertung von Faktorenanalysen zu beachten?
- Was ist das Ziel der Clusteranalyse?
- Welche Schritte umfasst die hierarchisch-agglomerative Clusteranalyse?
- Wie wird die Ähnlichkeit von Objekten/Cluster bemessen?

Übungsaufgaben

- **Faktorenanalyse:**
 - Auf welche Dimensionen lassen sich die dem Unternehmen X zugeschriebenen Eigenschaften reduzieren?
 - Auf welche Dimensionen lassen sich die dem Unternehmen X von den Befragten zugeschriebenen Aufgaben reduzieren?

- **Clusteranalyse:**
 - Zu wie vielen Clustern können die Automarken sinnvoll zusammengefasst werden?
 - Sollte evtl. eine Marke aus der Analyse ausgenommen werden?