

# Fallstudie zur Vorlesung Multivariate Methoden

## Herleitung von Profilen deutscher Kreise und Städte

Prof. Dr. Michael Bucker, Niels Schlüsener

13.10.2022

### Kontext

In der rasanten Dynamik des 21. Jahrhunderts sind Städte und Gemeinden konfrontiert mit immer komplexeren Herausforderungen. Eine effiziente und nachhaltige Stadtentwicklung verlangt heute mehr denn je ein fundiertes Verständnis für die vielen Variablen, die unsere Lebensräume prägen. In dieser Hinsicht bietet die Clusteranalyse ein geeignetes Werkzeug, um strukturelle Muster und Trends aufzudecken. Die vorliegende Ausschreibung widmet sich einer eingehenden Untersuchung aller Kreise und Städte in Deutschland, basierend auf Datenpunkten, die verschiedene Aspekte des städtischen und ländlichen Lebens umfassen.

Im Zuge dieser Untersuchung hat der in Deutschland zuständige Verband für städtische Entwicklung Ihrem Team den Auftrag erteilt, einen aktuellen Überblick über die spezifischen Gegebenheiten und Herausforderungen der Kreise und Städte in Deutschland zu erstellen. Dieser Überblick soll als Grundlage für zukünftige Strategien zur Verbesserung der Lebensqualität und der nachhaltigen Entwicklung dienen. Hierfür steht Ihnen ein umfangreicher Datensatz zur Verfügung, der Daten aus allen Kreisen und Städten in Deutschland umfasst.

Mithilfe dieser Daten und den im Rahmen der Studie vorgestellten Analysemethoden sind Sie nun beauftragt, Profile der einzelnen Kreise und Städte zu erstellen. Dabei sollen die Kreise und Städte hinsichtlich ihrer unterschiedlichen Merkmale geclustert werden. Dies wird ein tiefergehendes Verständnis der einzigartigen Gegebenheiten und Herausforderungen in jedem Bezirk ermöglichen und so die Grundlage für evidenzbasierte Strategien zur Verbesserung der Lebensqualität in den Regionen schaffen.

### Daten

In ILIAS finden Sie die Datei `regional_daten_de.RData`. Diese enthält Daten von 537 deutschen Kreisen und kreisfreie Städten. Zu den betrachteten Dimensionen zählen Ackerfläche, Beschäftigte (insbesondere Niederlassungen und verschiedene Betriebsgrößen), die Verteilung der Bodenfläche, die Wahlbeteiligung, Tourismusdaten und die Siedlungsfläche.

## Aufgabenstellung

Ziel dieser Fallstudie ist, zunächst aus den 69 Attributen eine geringere Anzahl Dimensionen mit Hilfe von Methoden zur Dimensionsreduktion abzuleiten, die anschließend zur Bestimmung der Profile mit Hilfe einer Clusteranalyse genutzt werden können. Außerdem können vorab Dimensionen ausgeschlossen werden wenn sie anhand von inhaltlicher Kriterien nicht weiter zur Analyse genutzt werden sollen.

Die konkreten Analyseschritte, welche Sie zur Lösung der Aufgabenstellung in **R** lösen sollen, lauten wie folgt:

1. Verschaffen Sie sich einen Überblick und ein Verständnis der vorliegenden Daten durch deskriptive Analysen und grafische Darstellungen.
2. Reduzieren Sie die Anzahl der Dimensionen mit Methoden der Dimensionsreduktion. Sie sollten dabei nicht mehr als zehn Dimensionen extrahieren. Eine Interpretation der Dimensionen ist nicht zwingend notwendig.
3. Nutzen Sie Ihre neuen Dimensionen als Basis, um mit Hilfe einer Clusteranalyse Profile der Gemeinden herzuleiten. Sie sollten nicht mehr als zehn Profile herausarbeiten. Kennzeichnen Sie ihre endgültige Lösung.
4. Interpretieren und beschreiben Sie die gebildeten Cluster. Entwickeln Sie außerdem Empfehlungen auf Basis der Profile, wie die Gemeinden innerhalb der Profile ihr Abfallmanagement verbessern könnten.

## Format

Ihre Abgabe besteht aus drei Teilen:

### R Code

Geben Sie den von Ihnen erstellten R Code zur Datenanalyse, Dimensionsreduktion und Clustererstellung ab. Der Code sollte nur die relevanten Teile enthalten, welche die Ergebnisse in Ihrer Präsentation erzeugen. Stellen Sie sicher, dass der Code sauber, nachvollziehbar, gut dokumentiert und vor allem auch lauffähig ist. Die aus dem Code resultierenden Ergebnisse sollten denen in Ihrem Bericht entsprechen!

### Management-Präsentation

Erstellen Sie zusätzlich zum \*.R Skript eine Management-Präsentation, in der Ihre Ergebnisse dokumentiert werden. Diese Präsentation sollte maximal 8 Folien enthalten (+ je eine pro ermitteltes Profil) und Ihre Vorgehensweise und die Ergebnisse Ihrer Analysen zusammenfassen. Folgende Elemente sollte Ihre Präsentation abdecken:

- **Titelseite:** Name und Beschreibung der Präsentation, Namen der Autorinnen und Autoren. (1 Folie)
- **Executive Summary:** Zusammenfassung der Ergebnisse. (1 Folie)
- **Kurzbeschreibung der Methodik:** Erläuterung der durchgeführten Analysen und der verwendete Datenbasis inklusiver kritischer Würdigung ihres Vorgehens. (3-4 Folien)
- **Vorstellung der Profile:** Nennung der durch die Analysen hergeleiteten Profile mit einem treffenden Namen und einer Kurzbeschreibung je Profil. (1-2 Folien)
- **Detaillierte Profilierung der Cluster:** Detaillierte Analyse der unterschiedlichen Charakteristika der Profile mit Hilfe von Grafiken oder Tabellen zur Erläuterung der Profilnamen, Beschreibungen und Empfehlungen. (je Profil eine Folie)

## Abgabe

Abgabe des Codes in Form eines \*.R Skripts und der Präsentation als \*.pdf Datei über ILIAS bis zum **19.01.2023**.

## **Bewerkungskriterien**

Bewertet werden alle Teile der Abgabe nach folgenden Kriterien:

- Korrektheit der fachlichen Lösung und Umfang der durchgeführten Analysen
- Adäquater Einsatz und kritische Reflexion von Methoden und Werkzeugen aus den Vorlesungen
- Übersichtlichkeit und Nachvollziehbarkeit des eingereichten Codes
- Strukturierung, Gestaltung und fachliche Korrektheit der Präsentation
- Einordnung und Reflexion der Analyseergebnisse im wirtschaftlichen Kontext

## Merkmale des Datensatzes

ID	Spaltenbezeichnung	Erläuterung
1	ID	Eindeutige Identifikationsnummer der Gemeinde.
2	Region	Region in der die Gemeinde liegt.
3	Provinz	Provinz in der die Gemeinde liegt.
4	Gemeinde	Name der Gemeinde.
5	Flaeche	Fläche der Gemeinde in km <sup>2</sup> .
6	Bevoelkerung	Population der Gemeinde.
7	Bevoelkerungsdichte	Population der Gemeinde pro km <sup>2</sup> .
8	Strassen	Gesamtlänge aller Straßen der Gemeinde in km.
9	Inselgemeinde	Indikator ob Gemeinde auf einer Insel ist (1 ja, 0 nein).
10	Kuestengemeinde	Indikator ob Gemeinde an der Küste liegt (1 ja, 0 nein).
11	Urbanisierungsgrad	Grad der Gemeindeurbanisierung (1 niedrig, 3 hoch).
12	Geologischer_Indikator	Geologische Lage der Gemeinde (1 Süden, 2 Mitte, 3 Norden).
13	Abfaelle_gesamt	Gesamtmenge der Gemeindeabfälle in Kilotonnen (kt).
14	Abfaelle_sortiert	Sortierte Gemeindeabfälle in Kilotonnen (kt).
15	Abfaelle_unsortiert	Unsortierte Gemeindeabfälle in Kilotonnen (kt).
16	Sortierungsgrad	Anteil der sortierten Gemeindeabfälle an den Gesamten.
17	Sort_Bio	Anteil der Bio-Abfällen unter den Sortierten.
18	Sort_Papier	Anteil der Papier-Abfällen unter den Sortierten.
19	Sort_Glas	Anteil der Glas-Abfällen unter den Sortierten.
20	Sort_Holz	Anteil der Holz-Abfällen unter den Sortierten.
21	Sort_Metall	Anteil der Metall-Abfällen unter den Sortierten.
22	Sort_Plastik	Anteil der Plastik-Abfällen unter den Sortierten.
23	Sort_Elektrik	Anteil der Elektrik-Abfällen unter den Sortierten.
24	Sort_Textil	Anteil der Textil-Abfällen unter den Sortierten.
25	Sort_Rest	Anteil der Rest-Abfällen unter den Sortierten.
26	Verwendung_Energie	Anteil des Abfalls, der Energiegewinnung verwendet wird.
27	Verwendung_Deponie	Anteil des Abfalls, der auf Deponien eingelagert wird.
28	Verwendung_Recycling	Anteil des Abfalls, der recyclet wird.
29	Verwendung_Unbekannt	Anteil des Abfalls, dessen Weiterverwendung unbekannt ist.
30	Steuern_gewerblich	Pro Kopf Steuereinnahmen aus gewerblichen Quellen.
31	Steuern_privat	Pro Kopf Steuereinnahmen aus privaten Quellen.
32	Kosten_Basis	Pro Kopf Basiskosten der Abfallwirtschaften.
33	Kosten_Sortierung	Pro Kopf Sortierungskosten der Abfallwirtschaften.
34	Kosten_sonstiges	Pro Kopf sonstige Kosten der Abfallwirtschaften.
35	Gebuehrenregelung	Abfallgebührenregelung der Gemeinde.
36	Region_PAYT	Indikator ob Region PAYT (Pay as you throw) anbietet.