

## 7 aus 49 – Ergebnis einer Clusteranalyse niedersächsischer Regionen

Ein wesentliches Element von regionalstatistischen Fragestellungen ist der Vergleich von Gebietseinheiten. Die Interpretation der Ergebnisse ist zum Teil schwierig: Ist die Einordnung im zeitlichen Vergleich auch meist simpel, ist die Frage des regionalen Vergleichsmaßstabes keineswegs trivial. Sollten sich zwei benachbarte Landkreise überhaupt miteinander vergleichen und aneinander messen oder sind sie strukturell so verschieden, dass – auch weiter – entfernte Kreise vielleicht einen viel besseren Vergleichspartner darstellen? Dieser Frage wurde mit Hilfe einer Clusteranalyse anhand von Strukturdaten für Landkreise, kreisfreie Städte sowie ausgewählte Großstädte und ihr Umland nachgegangen.

### Zur Methode der Clusteranalyse

Ziel einer Clusteranalyse ist es, systematische Ähnlichkeiten zwischen Objekten zu identifizieren und diese aufgrund ihrer Ähnlichkeiten zu Gruppen zusammenzufassen. Dabei sollten die Objekte innerhalb einer Gruppe möglichst homogen in Bezug auf die betrachteten Merkmale sein. Zwischen den Gruppen sollten die Unterschiede dagegen möglichst groß sein.

Nachfolgend wird eine Clusteranalyse für die niedersächsischen Kreise und kreisfreien Städte vorgestellt, die mit Hilfe von 10 allgemeinen Strukturindikatoren der Regionen, wie z.B. der Arbeitslosenquote, durchgeführt wurde (siehe Tabelle). Um Verzerrungen in den Ergebnissen durch zufällige Schwankungen zu vermeiden, wurden mehrjährige Durchschnitte der Daten gebildet. Vorab wurde auch geprüft, ob bestimmte Variablen miteinander stark korrelieren, d.h. ob ein linearer Zusammenhang besteht. Dies war für die Arbeitslosenquote und die Quote der Mindestsicherungsempfänger der Fall. Hier liegt der Korrelationskoeffizient  $r$  nach Bravais-Pearson bei 0,9. Da für die Arbeitslosenquote aktuellere Werte vorlagen, wurde die Mindestsicherungsempfängerquote aus der Analyse ausgeschlossen. Die grundlegende Vorgehensweise der Clusteranalyse soll im Folgenden kurz erläutert werden.

Die Clusteranalyse lässt sich in drei Ablaufschritte gliedern. Im ersten Schritt wurden die Distanzen bzw. Abstände zwischen den Kreisen hinsichtlich aller betrachteten Merkmale bestimmt. Für jedes mögliche Paar von Kreisen wurden dafür die Differenzen der einzelnen Merkmale berechnet und dann quadriert. Das gebildete Distanzmaß ergibt sich aus der Summe der quadrierten

### Clustermittelwerte der einzelnen Indikatoren

	Struktur-schwacher Raum	Groß-städti-sches Umland	Ost-friesland	Westen	Prosperie-render ländlicher Raum	Groß-städte	Wolfs-burg	Nieder-sachsen
Anzahl der Clustermitglieder	14	8	3	4	12	7	1	49
<b>Indikatoren<sup>1)</sup></b>								
SVB am Arbeitsort je 1 000 Einwohner	268	209	212	296	235	464	755	291
Arbeitslosenanteil an der Bevölkerung im erwerbsfähigen Alter in Prozent	7,3	5,7	6,8	4,0	4,7	7,4	5,5	5,9
Lebendgeborene je 1 000 Einwohner	7,2	7,4	8,1	9,4	8,1	8,6	8,0	8,1
Wanderungssaldo je 1 000 Einwohner	-3,0	-1,7	0,6	1,6	1,0	2,7	4,2	0,3
Kinder-Seniorenverhältnis (Bevölkerung im Alter von unter 15 Jahren je 100 Senioren über 75 Jahren)	132,4	161,0	189,4	242,9	197,0	142,3	126,8	165,8
Kinderbetreuungsquote 0- bis unter 6-Jähriger	49,8	52,6	43,1	45,3	49,1	51,8	57,0	49,5
Creditreform-Schuldnerquote in Prozent	12,1	10,5	11,7	8,9	9,5	11,3	8,9	10,6
Schulabbrecherquote in Prozent	7,6	5,8	9,1	7,0	6,3	8,8	5,8	7,0
Sachinvestitionen der Gemeinden je Einw. in €	117	140	185	257	175	153	362	160
Verfügbares Einkommen pro Kopf in €	17 476	16 602	15 541	16 608	18 428	16 947	18 482	17 814

1) Für alle Indikatoren wurden die Durchschnittswerte der Jahre 2007 bis 2009 verwendet. Ausgenommen sind die Sachinvestitionen und das Verfügbare Einkommen mit Durchschnitten für 2006 bis 2008.





Werteabstände und wird als *quadrierte euklidische Distanz* bezeichnet. Da den einbezogenen Merkmalen unterschiedliche Messniveaus zu Grunde liegen, z.B. Quoten und Absolutwerte in Euro, müssen die Werte vor der Berechnung des Distanzmaßes standardisiert, d.h. auf ein Niveau gebracht werden, um eine Vergleichbarkeit herzustellen. Es wurde dafür eine *z-Transformation* der Ausgangswerte durchgeführt, so dass die standardisierten Werte den Mittelwert 0 und die Standardabweichung 1 aufweisen. Im zweiten Schritt der Analyse wurden die Kreise auf Basis der berechneten Distanzen zu Gruppen zusammengefasst. Dazu wurde das *Ward-Verfahren* angewandt, ein hierarchisches Verfahren mit agglomerativem Algorithmus. Bei der Agglomeration werden solche Kreise zu Gruppen zusammengefügt, die die Varianz innerhalb der Gruppe, d.h. die Abweichungen vom Durchschnitt, möglichst wenig erhöhen. Die Agglomeration wird ausgehend von der feinsten Gliederung, bei der alle Kreise eine eigene Gruppe darstellen, so oft wiederholt, bis alle Kreise in einer Gruppe vereinigt sind. Dabei nimmt die Homogenität innerhalb der Gruppen mit steigendem Agglomerationsgrad ab. Im letzten Schritt muss nun die optimale Clusterzahl bestimmt werden. Dazu wurden verschiedene statistische Kriterien zur Beurteilung der Clustergüte herangezogen, wie z.B. das Ellbow-Kriterium, die Stopping-Rule von Calinski/Harabasz, der Test von Mojena und die Pseudo-t<sup>2</sup>-Statistik von Duda/Hart. Auf die einzelnen Kriterien soll an dieser Stelle nicht weiter eingegangen werden. Es ist jedoch zu beachten, dass bei der Entscheidung für eine Clusterlösung stets ein Konflikt zwischen der Forderung nach einer möglichst hohen Homogenität und einer handhabbaren, d.h. einer nicht zu großen Clusterzahl besteht, so dass es nicht immer eine „beste“ Lösung gibt.

Bei der Analyse der Basisindikatoren wiesen die Gütekriterien gute Werte bei einer Clusterzahl von 3 auf. Vergleichsweise gute Werte lagen auch bei 7 Clustern vor. Bei 3 Clustern waren die Kreise innerhalb einer Gruppe noch relativ heterogen, und die Zusammenfassung ist nach inhaltlichen Gesichtspunkten betrachtet noch vergleichsweise grob. Daher wurde als Lösung eine Aufteilung in 7 Cluster bevorzugt.

### Ergebnisse der Strukturindikatoren-Clusteranalyse

Ergebnis dieser mit zehn Strukturindikatoren durchgeführten Clusteranalyse sind die bereits erwähnten sieben Cluster, d.h. sieben mehr oder weniger große Gruppen mit einander ähnlichen Mitgliedern (siehe Karte).

- Strukturschwacher Raum (14 Mitglieder)
- Großstädtisches Umland (8 Mitglieder)
- Ostfriesland (3 Mitglieder)
- Westen (4 Mitglieder)
- Prosperierender ländlicher Raum (12 Mitglieder)

- Großstädte (7 Mitglieder)
- Wolfsburg / monostrukturelle Großstadt (1 Mitglied)

Diese sollen im Folgenden anhand der Clustermittelwerte für die einzelnen Indikatoren (siehe Tabelle) näher beschrieben werden. Bessere Werte als im Landesdurchschnitt weisen für die einzelnen Indikatoren vor allem der Westen, der Prosperierende ländliche Raum sowie Wolfsburg auf. Das Großstädtische Umland und der Strukturschwache Raum weichen im Gegensatz dazu eher zum Negativen vom Landesdurchschnitt ab.

**Strukturschwacher Raum:** Dieses Cluster weist die größte Mitgliederzahl auf und enthält als einziges sowohl städtische Gebiete als auch Landkreise. Zu Wilhelmshaven, Delmenhorst und Salzgitter gesellen sich die Heidekreise Soltau-Fallingb., Celle und Uelzen, die Landkreise Lüchow-Dannenberg und Helmstedt an der östlichen Landesgrenze, sowie die Kreise im Harz und im Weserbergland. Betrachtet man die eingeflossenen Strukturindikatoren, zeigt sich, dass neben der Kinderbetreuungsquote einzig das verfügbare Einkommen pro Kopf leicht über dem Landesdurchschnitt liegt – und dieses sind Indikatoren, die gewöhnlich erst zeitverzögert auf Veränderungen reagieren. In allen anderen Bereichen, d.h. vor allem Arbeitsmarkt, Demographie und Bildung, sind die Werte – zum Teil sogar deutlich – unterdurchschnittlich. Das ist umso problematischer, als Arbeitsplätze, Wanderungen und generatives Verhalten zusammenwirken und Positiv- oder Negativspiralen begründen können.

**Großstädtisches Umland:** Dieses Cluster hat acht Mitglieder und ist räumlich dreigeteilt: Im Norden gehören Friesland, Wesermarsch und Cuxhaven, im Süden das Göttinger Umland und schließlich weite Teile des Umlands der Städte Hannover, Hildesheim und Braunschweig dazu. Diese Gebiete sind zum einen durch geringe Zahlen von Arbeitslosen sowie niedrige Schuldnerquoten gekennzeichnet. Der leicht unterdurchschnittliche Arbeitsplatzbesatz ist typisch, weil hier zum Teil enge Verflechtungen mit der jeweiligen Großstadt bestehen und zur Arbeit in die Stadt gependelt wird. Die Kinderbetreuungsquoten sind über-, die Schulabbrecherquoten unterdurchschnittlich. Es handelt sich um „Speckgürtellandkreise“, die vergleichsweise gut dastehen. Das unterdurchschnittliche verfügbare Einkommen pro Kopf passt auf den ersten Blick nicht recht dazu. Es wird maßgeblich durch die drei Kreise im Norden reduziert. Eingetrübt wird dieses positive Bild durch die demographischen Strukturdaten: Unterdurchschnittliche Geburtenzahlen und ein davon beeinflusstes unterdurchschnittliches Kinder-Senioren-Verhältnis gehen mit Wanderungsverlusten einher. Hier zeigt sich auch der in den letzten Jahren zu beobachtende Trend der Rückkehr in die Städte.

**Ostfriesland:** Die drei Mitglieder dieses Clusters sind mit Ausnahme der Stadt Emden, die dem Großstädte-Cluster

zugeordnet wurde, genau die Landkreise, die Ostfriesland im historisch-politischen Sinne bilden. Das Cluster weist einen vergleichsweise geringen Arbeitsplatzbesatz und leicht überdurchschnittliche Anteile von Arbeitslosen und überschuldeten Personen auf. Das verfügbare Pro-Kopf-Einkommen ist mit Abstand das niedrigste aller gebildeten Cluster und liegt sehr deutlich unter dem Landesdurchschnitt. Deutlich unterdurchschnittlich sind auch die Kinderbetreuungsquoten 0- bis unter 6-Jähriger. Unter dem Aspekt frühkindlicher Sozialisation und Förderung ist dies negativ zu sehen, zumal die Schulabbrecherquote klar höher liegt als im Landesdurchschnitt und die größte aller Cluster darstellt. In den Bereichen demographische Entwicklung sowie den Sachinvestitionen der Gemeinden weisen diese Gebiete jedoch bessere Werte als im Landesdurchschnitt auf.

**Westen:** Es handelt sich bei den Mitgliedern mit den Landkreisen Vechta, Cloppenburg, Emsland und Grafschaft Bentheim um eng kooperierende Gebiete und eine Zone intensivster agrarindustrieller Produktion und Weiterverarbeitung derer Erzeugnisse. Der Boom dieser Gewerbe tritt in Form eines überdurchschnittlichen Arbeitsplatzbesatzes, vergleichsweise niedrigen bzw. sehr niedrigen Arbeitslosenanteilen und Schuldnerquoten zu Tage. Die Gewerbesteuererinnahmen werden von den Gemeinden offensichtlich vor Ort investiert. So gehören die Sachinvestitionen der Gemeinden je Einwohner zu den höchsten aller Cluster und überragen den Landesdurchschnitt. Demographisch steht dieses Cluster mit hohen Geburtenzahlen, Wanderungsgewinnen und einem günstigen Kinder-Senioren-Verhältnis so gut da wie kein anderes. Nach unten weichen allein das verfügbare Einkommen pro Kopf und die Kinderbetreuungsquoten vom Durchschnitt ab. Ersteres ist angesichts der Stärke des Produzierenden Gewerbes mit vielfach einfachen Tätigkeiten allerdings nicht überraschend.

**Prosperierender ländlicher Raum:** Dieses aus zwölf Landkreisen bestehende Cluster ist strukturell dem Cluster Westen sehr ähnlich. Es handelt sich insbesondere um den sich in den letzten Jahren positiv entwickelnden Raum, der das Land von den Landkreisen Osnabrück und Ammerland bis vor die Tore Hamburgs durchzieht. Hinzu kommt mit dem Landkreis Gifhorn das „Speckpolster“ der Städte Wolfsburg und Braunschweig. Strukturell sind Arbeitsplatzbesatz und Kinderbetreuungsquote leicht unterdurchschnittlich. Das verfügbare Einkommen ist hingegen stark über dem Landesdurchschnitt und bildet hier mit dem Cluster Wolfsburg die Spitze. Das ist durchaus typisch. Im Gegensatz zum Westen ist dieses Cluster nämlich keineswegs zentrenfern sondern enthält eine Reihe von Kreisen, die unmittelbar an Großstädte angrenzen und zumindest in Teilen auch schlicht das (sehr) wohlhabende Umland eines Zentrums darstellen.

**Großstädte:** Mitglieder dieses Clusters sind demographisch gut aufgestellte Großstädte, die sich als Zentren von Wirtschaft, Ausbildung und Forschung durch ihren Arbeitsplatzbesatz positiv abheben. Von einer traditionell größeren Dichte an Kinderbetreuungseinrichtungen einmal abgesehen, sind alle anderen Parameter eher leicht bis deutlich negativ vom Landesdurchschnitt abweichend. Neben einem vergleichsweise großen Angebot sozialer, kultureller und sonstiger Einrichtungen sind eine soziale Dichotomie sowie eine Ballung sozialer Problemlagen (wie sie hier in Arbeitslosenanteil, Schulabbrecher- und Schuldenquote zum Ausdruck kommt) typisch.

**Wolfsburg / monostrukturelle Großstadt:** Wolfsburg nimmt eine Sonderstellung ein. Würde man nur sechs Cluster bilden, fände sich die VW-Stadt mit den Großstädten in einer Gruppe wieder. Damit würde man aber eine eigenständige Struktur überdecken. Problematisch ist, dass der Stadt so in Niedersachsen ein geeigneter Vergleichsmaßstab fehlt.<sup>1)</sup> Das Arbeitsplatzangebot, das verfügbare Einkommen und die Sachinvestitionen suchen vom Niveau her ihresgleichen, sind aber überaus stark durch die dominierende Automobilindustrie beeinflusst und durch diese Monostruktur – an deren Veränderung gearbeitet wird – letztlich auch rapide wandelbar. Auch Werte aus den Bereichen Demographie, Kinderbetreuung und Überschuldung privater Haushalte stechen positiv hervor. Das Kinder-Senioren-Verhältnis ist allerdings das schlechteste aller Cluster.

## Ausblick

Die Ergebnisse der Clusteranalyse sind einerseits wenig spektakulär, weil verschiedene Gebiete bei diversen regionalstatistischen Fragestellungen immer wieder ähnlich „abschneiden“ und eine Unterscheidung in Stadt und Land ohnehin eine gängige Abgrenzungslinie darstellt. Andererseits sind die Ergebnisse aber doch neu und überraschend, wenn es um konkrete Clusterzugehörigkeiten und Gebietsabgrenzungen geht: Es ergeben sich nämlich räumlich zusammenhängende Gebiete der Cluster und kein „Flickenteppich“. Die 7 aus Strukturindikatoren gebildeten Cluster sind als „ein erster Aufschlag“ zu verstehen. Sie geben zweifellos Hinweise zu der Frage, welche Gebiete sich grundsätzlich miteinander vergleichen sollten. Hier wird mit einem gewissen Abstand jedoch zu überprüfen sein, ob die Ergebnisse konstant sind. Es bietet sich auch an, zusätzlich eine Clusteranalyse anhand von Entwicklungsindikatoren vorzunehmen und den Kreis der Strukturindikatoren zu erweitern.

1) Hier ist bundesweit nach geeigneten „Partnern“ zu suchen. Möglicherweise könnten Leverkusen und Mannheim/Ludwigshafen geeignet sein.