Klaus Pehl

Profilanalyse und Typisierung am Beispiel Weiterbildungsstatistik

Deutsches Institut für Erwachsenenbildung
Oktober 2005

Online im Internet:
URL: http://www.die-bonn.de/esprid/dokumente/doc-2005/pehl05_06.pdf
Dokument aus dem Internetservice texte.online des Deutschen Instituts für Erwachsenenbildung
http://www.die-bonn.de/publikationen/online-texte/index.asp

Abstract

Klaus Pehl (2005): Profilanalyse und Typisierung am Beispiel Weiterbildungsstatistik

Wichtige Eigenschaften von Weiterbildungseinrichtungen lassen sich nicht in einfachen Kennzahlen fassen. Vielmehr sind *Verteilungen* eines Merkmals über seine Ausprägungen (Kategorien) angemessene Beschreibungen. Beispiele dafür sind die Zusammensetzung der Gesamtfinanzierung aus Finanzierungsquellen wie auch die Verteilung des Zeitvolumens von Veranstaltungen auf Programmbereiche. Solche Verteilungen werden hier *Profile* genannt. Das zentrale Anliegen (Problemstellung) ist die statistische Analyse der Profile einer Gesamtheit von Weiterbildungseinrichtungen (Beispiel: Daten von 980 Volkshochschulen in 2003).
Zunächst wird eine Dokumentationsform für Profile vereinbart. Es wird ein Verfahren vorgeschlagen, wie die Ähnlichkeit eines Profils mit einem Vergleichsprofil geprüft und beurteilt wird. Darauf aufbauend wird ein „globaler" Indikator für die Ähnlichkeit entwickelt, der im ersten Schritt zwischen den Kategorien nicht unterscheidet. Wegen des sachlich begründeten Bedarfs, bei der Analyse die Kategorien einzeln zu berücksichtigen, wird der globale Indikator in kategorienspezifische Komponenten zerlegt. Die Ausprägung der Komponenten ermöglicht eine Klassenbildung für alle Profile nach ähnlichen Profiltypen. Diese Typisierung liefert das methodische Rüstzeug, um Untersuchungen des Zusammenhangs zwischen Profilen und anderen Merkmalen von Weiterbildungseinrichtungen anzustellen, lässt sich für die Differenzierung nach Gebieten (Bundesgebiet, Länder, Regionen) nutzen und macht Profile für eine Panelanalyse zugänglich. Der methodische Vorschlag wird an Beispielen aus der Statistik der Volkshochschulen erläutert.

Autor

Klaus Pehl ist Leiter des Programms „Strukturwandel der Weiterbildung" und der Abteilung „Informationszentrum Weiterbildung" am DIE.

Profilanalyse und Typisierung am Beispiel Weiterbildungsstatistik

Klaus Pehl

Inhalt

1 Zusammenfassung 1
2 Einleitung 2
3 Problemstellung 2
4 Profile als Untersuchungsgegenstand – Das Modell 5
5 Prüfung der Ähnlichkeit zwischen einem Profil und einem Vergleichsmodell 6
5.1 Vergleichsprofil als Referenz 7
5.2 Prüfung der Anpassung 7
5.3 Beispiele für die Anwendung auf Finanzierungsprofile und Programmprofile 8
6 Kennzeichnung der Ähnlichkeit zwischen Profilen und einem Vergleichsprofi 10
6.1 Der Kontingenzkoeffizient als Indikator 10
6.2 Kontingenzkomponenten als kategorienspezifische Indikatoren 12
6.3 Typisierung von Profilen mit Hilfe von Kontingenzkomponenten 14
7 Weiterführende Analyseansätze 21
7.1 Zusammenhang zwischen Profiltyp und anderen Merkmalen 21
7.2 Gebietsspezifische Analysen 23
7.3 Zeitreihenanalyse 24
8 Schlussbemerkung 29
9 Tabellen- und Abbildungsverzeichnis 29
10 Literatur 30

1 Zusammenfassung

Wichtige Eigenschaften von Weiterbildungseinrichtungen lassen sich nicht in einfachen Kennzahlen fassen. Vielmehr sind *Verteilungen* eines Merkmals über seine Ausprägungen (Kategorien) angemessene Beschreibungen. Beispiele dafür sind die Zusammensetzung der Gesamtfinanzierung aus Finanzierungsquellen wie auch die Verteilung des Zeitvolumens von Veranstaltungen auf Programmbereiche. Solche Verteilungen werden hier *Profile* genannt.[1] Das zentrale Anliegen (Problemstellung) ist die statistische Analyse der Profile einer Gesamtheit von Weiterbildungseinrichtungen (Beispiel: Daten von 980 Volkshochschulen in 2003).

Zunächst wird eine Dokumentationsform für Profile vereinbart. Es wird ein Verfahren vorgeschlagen, wie die Ähnlichkeit eines Profils mit einem Vergleichsprofil geprüft und beurteilt wird. Darauf aufbauend wird ein „globaler" Indikator für die Ähnlichkeit entwickelt, der im ersten Schritt zwischen den Kategorien nicht unterscheidet. Wegen des sachlich begründeten Bedarfs, bei der Analyse die Kategorien einzeln zu berücksichtigen, wird der globale Indikator in kategorienspezifische Komponenten zerlegt. Die Ausprägung der Komponenten ermöglicht eine Klassenbildung für alle Profile nach ähnlichen Profiltypen. Diese Typisierung liefert das methodische Rüstzeug, um Untersuchungen des Zusammenhangs zwischen Profilen und anderen Merkmalen von Weiterbildungseinrichtungen anzustellen, lässt sich für die Differenzierung nach Gebieten (Bundesgebiet, Länder, Regionen) nutzen und macht Profile für eine Panelanalyse zugänglich. Der methodische Vorschlag wird an Beispielen aus der Statistik der Volkshochschulen erläutert.

[1] Der Begriff „Profile" macht auf der Ebene der Sachanwendung in der Weiterbildung die Art der Information verständlicher als der Begriff „(Häufigkeits-)Verteilung eines statistischen Merkmals" aus der statistischen Methodenlehre.

2 Einleitung

Der wesentliche Akzent dieses Beitrags liegt auf einem methodischen Vorschlag, wie zentrale Fragestellungen von Weiterbildungseinrichtungen und ihren Verbänden einer statistischen Analyse zugänglich gemacht werden können.[2] Dazu gehört sowohl die Zusammensetzung der Finanzierungsquellen als auch das Profil der Veranstaltungen für eine Einrichtung oder für eine Gruppe von Einrichtungen zu beschreiben, Vergleiche anstellen zu können und hierbei auch die zeitliche Entwicklung einzubeziehen. Die unmittelbaren Nutzer der Anwendung der vorgeschlagenen Methoden können also die Akteure in der Weiterbildungspraxis in Einrichtungen und Verbänden sein, die sich bei der zukünftigen Gestaltung ihrer Organisation in der Argumentation auch auf empirische Befunde aus den zurückliegenden Jahren stützen wollen. Dabei ist zu erwarten, dass sie sich fachkundiger Hilfe versichern. Fachkundig heißt genauer: in der Anwendung quantitativer statistischer Methoden ausreichend bewanderte Personen. Selbst wenn diese Kompetenz bei Personen mit Hochschulabschluss z. B. in Erziehungs- oder Sozialwissenschaftlern gut ausgeprägt sein sollte, kann man nicht damit rechnen, dass sie in jedem Fall ausreichende Einblicke in die spezifischen Sachzusammenhänge der Weiterbildung und ihrer statistischen Probleme haben. Insofern wendet sich dieser Beitrag auch an wissenschaftlich Arbeitende mit statistischen Grundkenntnissen, die die Weiterbildungseinrichtungen und ihre Verbände bei statistischen Analysen unterstützen, und fördert den sachgerechten Dialog zwischen ihnen und den unmittelbaren Nutzern in den Einrichtungen und Verbänden. Dabei werden nicht zu viele Kenntnisse vorausgesetzt. Das kommt sowohl den Methodikern zu Gute, die sich nicht zu den Statistikspezialisten zählen, wie auch den Weiterbildungspraktikern, die einen Einblick in die Methoden wünschen.

Im folgenden Abschnitt 3 „Problemstellung“ wird anhand von Beispielmaterial die Problemstellung formuliert. Im Abschnitt 4 „Profile als Untersuchungsgegenstand – Das Modell“, Seite 5, wird vorbereitend der statistisch-methodische Formalismus zur Modellierung von Profilen in kompakter Form vorgestellt. Diese Vorbereitung wird durch eine Einstimmung in den „klassischen“ Ansatzes des Verteilungsvergleichs über die Analyse von Mehrfeldertafeln mit Beispielen aus der Weiterbildungspraxis in Abschnitt 5 „Prüfung der Ähnlichkeit zwischen einem Profil und einem Vergleichsmodell“, Seite 6ff., ergänzt. Der Abschnitt 6 „Kennzeichnung der Ähnlichkeit zwischen Profilen und einem Vergleichsprofil“, Seite 10ff., ist insoweit zentral, als er den Vorschlag für eine Methode zur Typisierung in verschiedenen Varianten auf der Basis von Kontingenzkomponenten umfasst. Im abschließende Abschnitt 7, Seite 21ff., wird versucht, den möglichen Nutzen der vorgestellten Methode für zentrale Fragestellungen zur Weiterbildung aus institutioneller Sicht zu umreißen.

Beispiele zur Sicherung der Nachvollziehbarkeit sind der vom Deutschen Institut für Erwachsenenbildung wissenschaftlich betreuten Volkshochschul-Statistik entnommen und durch kursive Schrift in kleinerem Schriftgrad kenntlich gemacht.

3 Problemstellung

Von Analysen jährlicher statistischer Daten aus Weiterbildungseinrichtungen wird gemeinhin erwartet, dass sie zu Aussagen über komplexe Strukturen wie die Zusammensetzung der Gesamtfinanzierung aus verschiedenen Finanzierungsquellen („Finanzierungsprofil“) oder der Verteilung der Unterrichtsstunden von Veranstaltungen auf Programmbereiche („Programmprofil“) führen. Der Begriff des Profils wird hier verwendet, um auszudrücken, dass es auf die Größenverhältnisse zwischen den Kategorien (Finanzierungsquellen, Programmbereiche) ankommen soll und nicht auf die absolute Werte der Merkmale (in EUR bzw. Unterrichtsstunden).

Im Fall Volkshochschul-Statistik liegen jährlich für ca. 1.000 Einrichtungen entsprechende Informationen vor. Neben Jahresstatistiken als Analysen im Querschnitt sind besonders Analysen der zeitlichen Entwicklung gefragt. Sie können sich zu Nutze machen, dass über einen großen Zeitraum *alle* Volkshochschulen in die Statistik einbezogen sind. Insofern sind alle

[2] Die Methode geht auf eine Idee zurück, die bereits 1987 (Pehl 1987) exploriert wurde.

Möglichkeiten einer Panel-Analyse auf der Basis vieler individueller Entwicklungen („Institutionenpanel“) gegeben.

Das passende mathematisch-statistische Modell für Profile der genannten Art bilden *Häufigkeitsverteilungen*. „Profilanalysen“ in diesem Sinn hätten sich also im Querschnitt wie in der Zeitreihe nicht mit jeweils einer (Kenn-)Zahl, sondern mit komplexeren Gebilden zu befassen. Dafür wäre ein Konzept zur Reduzierung der Komplexität willkommen. Das nahe liegende Konzept, einfache Parameter der Häufigkeitsverteilungen wie Lage- oder Streuungsmaße zu verwenden, scheitert daran, dass Daten wie in den genannten Beispielen „Finanzprofil“ und „Programmprofil“ lediglich kategorial vorliegen. Gesucht ist also ein Verfahren, das aus Häufigkeitsverteilungen einfach strukturierte Informationen herauszieht, um sie in geeignete Klassen einzuteilen. Dies erlaubt dann Analysen

- von vielen Häufigkeitsverteilungen aller Untersuchungseinheiten im Querschnitt,
- von einzelnen Untersuchungseinheiten im Längsschnitt,
- aller Häufigkeitsverteilungen aller Untersuchungseinheiten in einem Zeitraum (Panel),
- des Zusammenhangs zwischen Häufigkeitsverteilungen und anderen Merkmalen.

Beispiel 1 Finanzierungsprofile

Eine zentrale Fragestellung für Weiterbildungseinrichtungen ist, für ihr Finanzierungsvolumen in einem Jahr die wichtigsten Finanzierungsquellen zu unterscheiden, um ihre Anteile am Gesamtvolumen in der Entwicklung verfolgen oder Vergleiche mit anderen Einrichtungen anstellen zu können (s. Tabelle 1 und Abbildung 1).

Tabelle 1 Finanzierung zweier Volkshochschulen 2003 nach Finanzierungsquellen

Finanzierungsquelle	A Teilnahmegebühren	B Zuschüsse der Kommune	C Zuschüsse des Landes	D Drittmittel	Summe
VHS 1: Betrag in 1.000 EUR	718	494	425	423	2.060
VHS 1: Anteil	34,9%	20,5%	24,0%	20,6%	100%
VHS 2: Betrag in 1.000 EUR	6.051	9.166	2.865	3.389	21.471
VHS 2: Anteil	28,2%	42,7%	13,3%	15,8%	100%

Solche Vergleiche müssen sich nicht auf dieselbe Aggregationsstufe beziehen, d. h. solche Vergleiche sind nicht auf direkte Vergleiche zwischen zwei Weiterbildungseinrichtungen beschränkt. Sie können sich auch auf die Profile von die Weiterbildungseinrichtung umfassende Gebiete (definierte Gruppe von Einrichtungen, Einrichtungen Region, Land, Bundesrepublik Deutschland) beziehen. Zur Veranschaulichung sind in Abbildung 1, Seite 4 die Finanzierungsprofile von zwei Einrichtungen und dem Finanzierungsprofil des zugehörigen Landes gegenübergestellt.

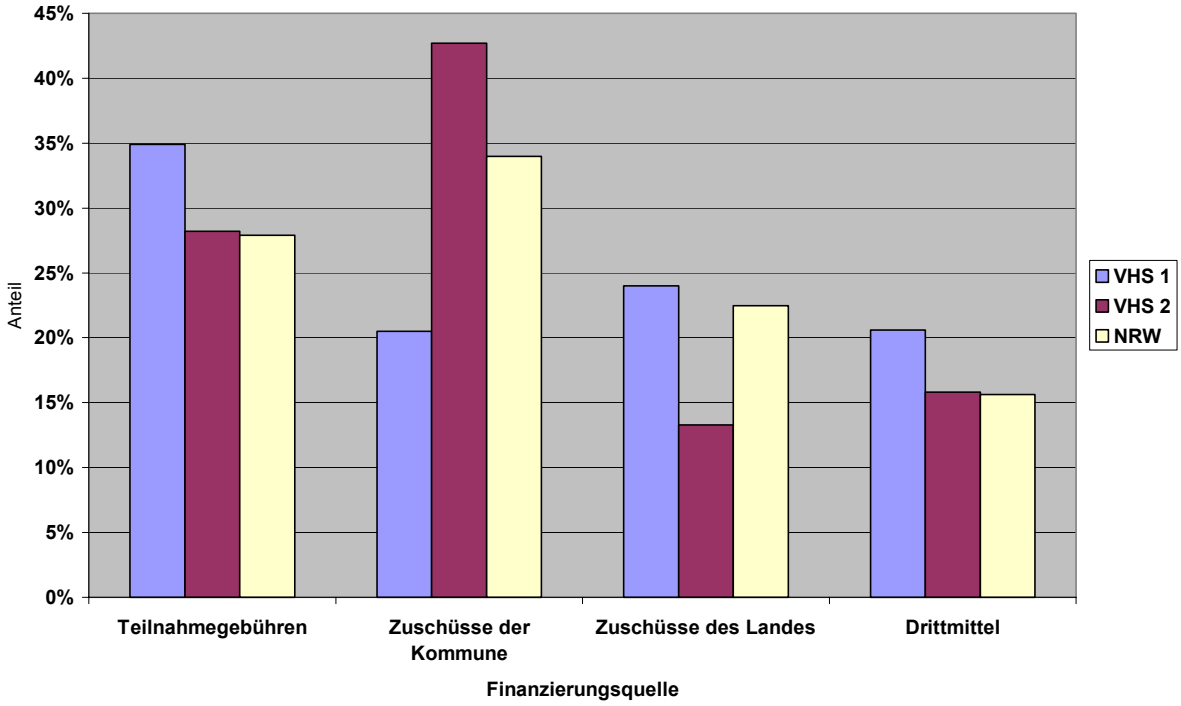


Abbildung 1 Finanzierungsprofile

Fragen sind z. B.

- *Was charakterisiert das Finanzierungsprofil der VHS 1 bzw. VHS 2?*
- *Gibt es unter den Volkshochschulen 2003 zu dem Finanzierungsprofil der VHS 1 bzw. VHS 2 ähnliche Finanzierungsprofile?*
- *Wie haben sich die Finanzierungsprofile der beiden Volkshochschulen im Zeitverlauf entwickelt?*
- *Welche Typen lassen sich für Finanzierungsprofile unterscheiden?*
- *Welche Zusammenhänge zwischen Typ des Finanzierungsprofils und anderen Merkmalen lassen sich feststellen?*

Beispiel 2 Programmprofile

Ebenso zentral wie Analysen von Finanzierungsprofilen sind die Fragen nach den Anteilen von Programmbereichen an dem Gesamtzeitvolumen der durchgeführten Veranstaltungen einer Weiterbildungseinrichtung (s. Tabelle 2, unten). Auch hier werden Analysen sich auf Vergleiche mit weiteren Einrichtungen oder Gebieten sowie auf zeitliche Entwicklungen beziehen.

Tabelle 2 Verteilung der Unterrichtsstunden für zwei Volkshochschulen 2003 auf Programmbereiche

Programmbereiche	A Politik – Gesellschaft – Umwelt	B Kultur – Gestalten	C Gesundheit	D Sprachen	E Arbeit – Beruf	F Grundbildung – Schulabschlüsse	Summe
VHS 1: Anzahl der Unterrichtsstunden	813	1.167	1.205	12.096	5.557	8.334	29.172
VHS 1: Anteil	2,8%	4,0%	4,1%	41,5%	19,0%	28,6%	100%
VHS 2: Anzahl der Unterrichtsstunden	5.155	22.797	13.860	124.481	32.651	11.287	210.231
VHS 2: Anteil	2,5%	10,8%	6,6%	59,2%	15,5%	5,4%	100%

Analog zu Finanzierungsprofilen sind auch im zweiten Beispiel in der Abbildung 2, Seite 5 nicht nur die Programmprofile zweier Einrichtungen gegenübergestellt, sondern auch zum Vergleich das Profil des entsprechenden Landes.

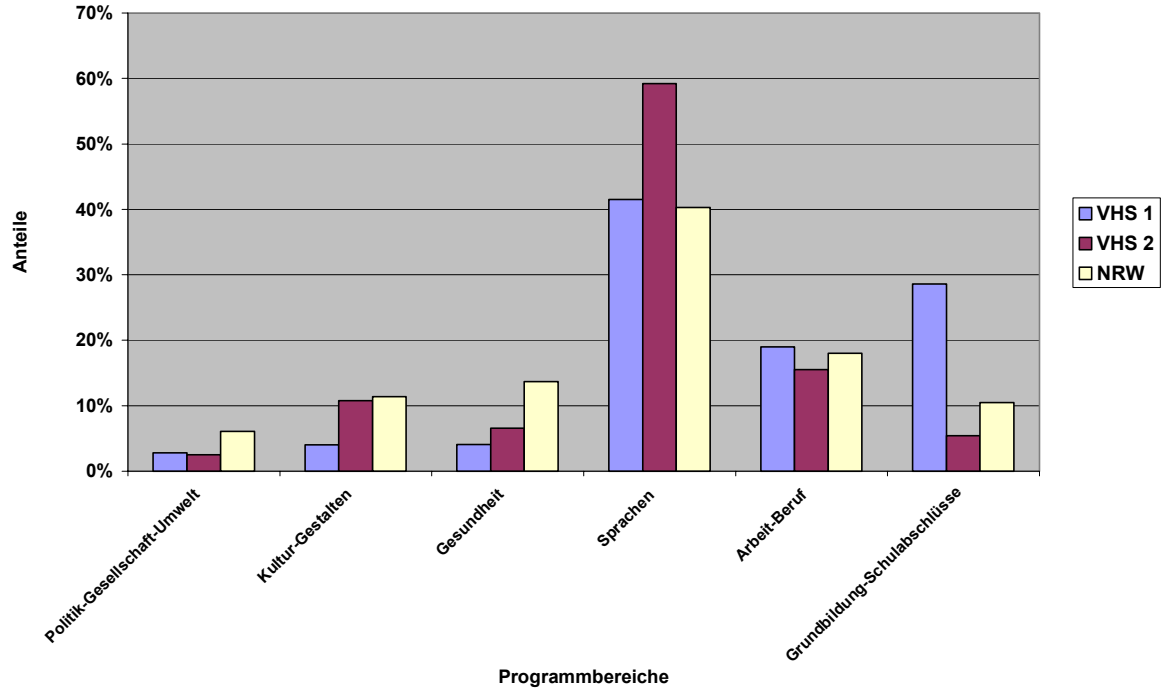


Abbildung 2 Programmprofile

Fragen sind z. B.

- *Was charakterisiert das Programmprofil der VHS 1 bzw. VHS 2?*
- *Gibt es unter den Volkshochschulen 2003 zu dem Programmprofil der VHS 1 bzw. VHS 2 ähnliche Finanzierungsprofile?*
- *Wie haben sich die Programmprofile der beiden Volkshochschulen im Zeitverlauf entwickelt?*
- *Welche Typen lassen sich für Programmprofile unterscheiden?*
- *Welche Zusammenhänge zwischen Typ des Programmprofils und anderen Merkmalen lassen sich feststellen? Gibt es einen Zusammenhang zwischen Programmprofil und Finanzierungsprofil?*

4 Profile als Untersuchungsgegenstand – Das Modell

In diesem Abschnitt wird das übliche mathematisch-statistische Modell für Profile, d. h. Häufigkeitsverteilungen, in Symbolschreibweise in kompakter Form benannt. Das Modell ist die Grundlage, um in den folgenden Abschnitten die vorgeschlagene Analysemethode für Profile zu verstehen und praktisch durchführen zu können, und erübrigt ein Nachschlagen in statistischer Grundlagenliteratur. Dieser Abschnitt kann von den statistisch-methodisch Kundigen überschlagen werden.

Eine Häufigkeitsverteilung P eines kategorialen Merkmals mit k Ausprägungen wird als „Profil“ bezeichnet. Es ist vollständig beschrieben durch die absoluten oder relativen Häufigkeiten der Ausprägungen.

$$\mathrm{P} = \left\{ \frac{n_1}{n}, \ldots, \frac{n_k}{n} \right\}$$

Dabei bezeichnen die n_i die absoluten Häufigkeiten, k (≥ 2) die Zahl der Ausprägungen und N die Summe aller absoluten Häufigkeiten $n = \sum_{i=1}^{k} n_i$. Die Verhältnisse $\frac{n_1}{n}$ sind die relativen

Häufigkeiten (Anteile; als Verhältniszahlen zwischen 0 und 1 oder als Prozentzahlen zwischen 0 und 100), für die gilt

$$\sum_{i=1}^{k} \frac{n_i}{n} = 1$$

Üblich ist eine Darstellung eines Profils mit z. B. 6 Kategorien (s. Beispiel 2, Seite 4) in Tabellenform.

Tabelle 3 Tabellarische Darstellung einer Häufigkeitsverteilung/eines Profils

i	Absolute Häufigkeiten	Relative Häufigkeiten als Verhältniszahlen zwischen 0 und 1	Relative Häufigkeiten als Prozentzahlen
1	n_1	n_1/n	$n_1/n \times 100\%$
2	n_2	n_2/n	$n_2/n \times 100\%$
3	n_3	n_3/n	$n_3/n \times 100\%$
4	n_4	n_4/n	$n_4/n \times 100\%$
5	n_5	n_5/n	$n_5/n \times 100\%$
6	n_6	n_6/n	$n_6/n \times 100\%$
Summe	N	1	100%

In dem Anwendungsfeld Weiterbildungsstatistik, das hier an dem speziellen Beispiel der Volkshochschul-Statistik dargestellt wird, ist ein Profil *eine* (komplexe) Eigenschaft *einer* Untersuchungseinheit (= Volkshochschule) in *einem* Berichtsjahr. Da Analysen weitergehend auf eine Gesamtheit von vielen Einrichtungen und ihre Profile sowie deren zeitliche Entwicklung zielen, benötigt man Schreibweisen, um eine Einrichtung und einen Zeitpunkt zu identifizieren. Es werden deshalb im Folgenden auch Bezeichnungen gebraucht wie

$$\mathrm{P}_{17;2001} = \left\{ \frac{n_{1;17;2001}}{n_{17;2001}}, \ldots, \frac{n_{k;17;2001}}{n_{17;2001}} \right\}$$ für das Profil der Volkshochschule mit der Schlüsselziffer 17 im Berichtsjahr 2001 oder allgemein für die Untersuchungseinheit j im Berichtsjahr y

$$\mathrm{P}_{j;y} = \left\{ \frac{n_{1;j;y}}{n_{j;y}}, \ldots, \frac{n_{k;j;y}}{n_{j;y}} \right\}$$

Aufbauend auf diesem „Rüstzeug“ widmet sich der nächste Abschnitt Vergleichsverfahren für Profile.

5 Prüfung der Ähnlichkeit zwischen einem Profil und einem Vergleichsmodell

Ausgangspunkt für die Entwicklung eines Konzepts für Profilanalysen und Typisierung von Profilen ist die Überlegung, dass es für ein Berichtsjahr möglich sein muss, die Profile einer großen *Gruppe* von Untersuchungseinheiten zu analysieren. Deshalb wird für die Analyse von Profilen *ein* geeignetes Vergleichsprofil herangezogen. Die Grundfrage in dieser Phase ist die Suche nach einem allgemeinen Ähnlichkeitsmaß zwischen einem Profil und dem gewählten Vergleichsprofil. Hier wird ein Verfahren analog dem χ^2-Anpassungstest (Clauß 2002, S. 189f.) vorgeschlagen, ein Verfahren, das die Passung zwischen einem Modell und empirischen Werten prüft. Auch diesen Abschnitt werden statistisch-methodisch Kundige überschlagen oder sich allenfalls anhand der Beispiele in die Weiterbildungspraxis einendenken wollen (Abschnitt 5.3, Seite 8).

5.1 Vergleichsprofil als Referenz

Ein Vergleichsprofil P_0 ist wie das zu vergleichende Profil selbst eine Verteilung. An die Stelle der relativen Häufigkeiten treten (theoretische) Anteile. Sie beschreiben ein modellartiges Gewichtungsschema für die gleichen k Kategorien wie das mit empirischen Werten gewonnene Profil.

$$\mathrm{P}_0 = \{p_{1;0},...,p_{k;0}\} \text{ mit } \sum_{i=1}^{k} p_{i;0} = 1$$

Für das Ergebnis spielt die Auswahl des Vergleichsprofils eine entscheidende Rolle. Ein nahe liegender Ansatz ist, für die Analyse von Profilen einer Gruppe von Untersuchungseinheiten das „aggregierte" Profil heranzuziehen. Das bedeutet, dasjenige Profil zu verwenden, welches sich aus der Summierung aller absoluten Häufigkeiten derselben Kategorie über alle Untersuchungseinheiten der Gruppe ergibt. In Fortsetzung der Beispiele für Volkshochschulen wäre dies bei einer Untersuchung aller Volkshochschulen in der Bundesrepublik Deutschland die Verteilung der Finanzierungssummen nach Finanzierungsquellen bzw. die Verteilung der Unterrichtsstundensummen auf die Programmbereiche. Im Fall einer Untersuchung, die sich auf Nordrhein-Westfalen beschränkt, wären die aus den nordrhein-westfälischen Gesamtsummen gebildeten Vergleichprofile heranzuziehen. Dieses Vorgehen hat den Vorteil, dass die durchschnittliche Streuung zwischen den Weiterbildungseinrichtungen der Gruppe berücksichtigt wird, aber den (bei einer großen Gruppe vernachlässigbaren) formalen Nachteil, dass das Vergleichsmodell nicht unabhängig von den zu analysierenden Profilen ist.

Konzepte zur Etablierung von Vergleichmodellen können auch Zielvorgaben oder Profile aus früheren Bezugsjahren sein. Ein Gleichverteilungsansatz, d. h. Profile mit der Gleichverteilung über die Kategorien zu vergleichen, ist für die praktisch relevanten Fragestellungen (Finanzierungsprofil, Programmprofil) jedoch eher abwegig.

5.2 Prüfung der Anpassung

Das Vergleichsprofil beschreibt die *Erwartung* für empirische Werte im Rahmen der Modellannahmen. Demgegenüber stehen die beobachteten Werte aus der Statistik.

Test der Anpassung eines Profils an ein Vergleichsmodell

Die Testgröße für den χ^2-Anpassungstest ist

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_{i;0})^2}{np_{i;0}} \text{ oder auch } \chi^2 = n\sum_{i=1}^{k} \frac{(\frac{n_i}{n} - p_{i;0})^2}{p_{i;0}}$$

Die erste Form ist gut einzusetzen, wenn absolute Häufigkeiten als empirische Werte vorliegen. Die zweite Form eignet sich für den Fall, dass relative Häufigkeiten und die Gesamtanzahl vorliegen.

Sie ist unter der Annahme der Nullhypothese $H_0 : p_1 = p_{1;0},..., p_k = p_{k;0}$ (Verteilungen sind identisch) näherungsweise χ^2-verteilt mit $k-1$ Freiheitsgraden. Die Näherung gilt als brauchbar, wenn

1. die Summe der absoluten Häufigkeiten $n \geq 40$ ist,
2. alle erwarteten absoluten Häufigkeiten $n \times p_{i;0} \geq 1$ sind und
3. mindestens vier Fünftel (80%) der erwarteten absoluten Häufigkeiten $n \times p_{i;0} \geq 5$ sind.

Die Nullhypothese H_0 wird mit einer Irrtumswahrscheinlichkeit α abgelehnt, wenn der Wert der Testgröße das $1-\alpha$-Quantil der χ^2-Verteilung mit $k-1$ Freiheitsgraden erreicht oder übertrifft. Für die Güte der Anpassung (engl. goodness of fit) wird nach Lienert (in Clauß 2002, S. 191) vorgeschlagen sich an der Irrtumswahrscheinlichkeit zu orientieren, mit der die Nullhypothese H_0 abgelehnt werden könnte.

Tabelle 4 Goodness of fit nach Lienert (1962)

Irrtumswahrscheinlichkeit α	$> 0{,}50$	$0{,}50...0{,}20$	$0{,}19...0{,}05$	$< 0{,}05$
Anpassung	gut	mäßig	schwach	keine

Die kritischen Werte (Quantile der χ^2-Verteilung) zeigt die folgende Tabelle. Dabei werden nur Werte bis 10 Freiheitsgrade tabelliert, was für die relevanten Profile im Anwendungsbeispiel ausreicht.

Tabelle 5 Kritische Werte für den χ^2-Anpassungstest

F.G.[3]	α						
	0,50	**0,30**	**0,10**	**0,05**	**0,025**	**0,01**	**0,001**
1	0,455	1,07	2,71	3,84	5,02	6,64	10,8
2	1,39	2,41	4,61	5,99	7,38	9,21	13,8
3	2,37	3,67	6,25	7,81	9,35	11,3	16,3
4	3,36	4,88	7,78	9,49	11,1	13,3	18,5
5	4,35	6,06	9,24	11,1	12,8	15,1	20,5
6	5,35	7,23	10,6	12,6	14,4	16,8	22,5
7	6,35	8,38	12,0	14,1	16,0	18,5	24,3
8	7,34	9,52	13,4	15,5	17,5	20,1	26,1
9	8,34	10,7	14,7	16,9	19,0	21,7	27,9
10	9,34	11,8	16,0	18,3	20,5	23,2	29,6

Diese Bedingungen sind im Anwendungsbeispiel der Analyse von Finanzierungsprofilen oder Programmprofilen von Volkshochschulen im Vergleich zu dem Bundesprofil in der Regel erfüllbar.

5.3 Beispiele für die Anwendung auf Finanzierungsprofile und Programmprofile

Um das Verfahren zu veranschaulichen wird in diesem Unterabschnitt das Berechnungsverfahren für Finanzierungsprofile und Veranstaltungsprofile skizziert.

[3] Freiheitsgrade

Beispiel 3 Ähnlichkeiten zwischen Finanzierungsprofilen und Vergleichprofil

Das Beispiel setzt das Beispiel 1, Seite 3 fort.

Tabelle 6 Finanzierung zweier Volkshochschulen 2003 nach Finanzierungsquellen („Finanzierungsprofil“)

Finanzierungsquelle	A Teilnahmegebühren	B Zuschüsse der Kommune	C Zuschüsse des Landes	D Drittmittel	Summe
VHS 1: Betrag in 1.000 EUR	718	494	425	423	2.060
VHS 1: Anteil	34,9%	20,5%	24,0%	20,6%	100%
VHS 2: Betrag in 1.000 EUR	6.051	9.166	2.865	3.389	21.471
VHS 2: Anteil	28,2%	42,7%	13,3%	15,8%	100%
NRW: Betrag in 1.000 EUR	60.360	73.485	48.596	33.810	216.251
NRW: Anteil	27,9%	34,0%	22,5%	15,6%	100%

Die Testgröße für den χ^2 -Anpassungstest hat bei der VHS 1 den Wert

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_{i;0})^2}{np_{i;0}} = \frac{(718-2060\times 0{,}279)^2}{2060\times 0{,}279} + \frac{(494-2060\times 0{,}340)^2}{2060\times 0{,}340} + \frac{(425-2060\times 0{,}225)^2}{2060\times 0{,}225} + \frac{(423-2060\times 0{,}156)^2}{2060\times 0{,}156}$$

$$\chi^2 = \frac{(718-575)^2}{575} + \frac{(494-700)^2}{700} + \frac{(425-463)^2}{463} + \frac{(423-322)^2}{322}$$

$$\chi^2 = \frac{143^2}{575} + \frac{(-206)^2}{700} + \frac{(-38)^2}{463} + \frac{101^2}{322}$$

$$\chi^2 \approx 36+61+3+32 = 131 \geq 16{,}3 \quad (\chi^2_{0{,}999;3})$$

Das bedeutet: Das Finanzierungsprofil weicht signifikant vom nordrhein-westfälischen Vergleichsprofil ab.

Für dieselbe Testgröße errechnet man mit Hilfe der Formel, die lediglich die Summe aller absoluten Häufigkeiten und die Anteile verwendet:

$$\chi^2 = n\sum_{i=1}^{k} \frac{(\frac{n_i}{n} - p_{i;0})^2}{p_{i;0}} = 2.060\times\left[\frac{(0{,}349-0{,}279)^2}{0{,}279} + \frac{(0{,}205-0{,}340)^2}{0{,}340} + \frac{(0{,}240-0{,}225)^2}{0{,}225} + \frac{(0{,}206-0{,}156)^2}{0{,}156}\right]$$

$$\chi^2 = 2.060\times\left[\frac{0{,}070^2}{0{,}279} + \frac{(-0{,}135)^2}{0{,}340} + \frac{0{,}015^2}{0{,}225} + \frac{0{,}050^2}{0{,}156}\right]$$

$$\chi^2 \approx 2.060\times[0{,}017+0{,}053+0{,}001+0{,}016] \approx 2.060\times 0{,}088 \approx 181 \geq 16{,}3$$

Die Abweichung vom Ergebnis nach dem ersten Verfahren ergibt sich aus Rechenungenauigkeiten (Runden bei Zwischenergebnissen), für die dieses Verfahren wesentlich empfindlicher ist als das erste.

Auch bei der VHS 2 ergibt sich mit einem Wert von 1.276 eine hochsignifikante Abweichung vom nordrhein-westfälischen Vergleichsmodell.

Beispiel 4 Ähnlichkeiten zwischen Programmprofilen und Vergleichprofil

Das Beispiel setzt das Beispiel 2, Seite 4 fort.

Tabelle 7 Verteilung der Unterrichtsstunden für Nordrhein-Westfalen 2003 auf Programmbereiche ("Programmprofile")

Programmbereiche	A Politik-Gesellschaft-Umwelt	B Kultur-Gestalten	C Gesundheit	D Sprachen	E Arbeit-Beruf	F Grundbildung-Schulabschlüsse	Summe
VHS 1: Anzahl der Unterrichtsstunden	813	1.167	1.205	12.096	5.557	8.334	29.172
VHS 1: Anteil	2,8%	4,0%	4,1%	41,5%	19,0%	28,6%	100%
VHS 2: Anzahl der Unterrichtsstunden	5.155	22.797	13.860	124.481	32.651	11.287	210.231
VHS 2: Anteil	2,5%	10,8%	6,6%	59,2%	15,5%	5,4%	100%
NRW: Anzahl der Unterrichtsstunden	169.643	316.133	379.518	1.119.165	501.029	291.562	2.777.050
NRW: Anteil	6,11%	11,38%	13,67%	40,30%	18,04%	10,50%	100%

Die Testgröße für den χ^2-Anpassungstest hat bei der VHS 1 ungefähr den Wert 13.000, bei der VHS 2 den Wert ungefähr 37.000. Hätte man hier ähnlich wie bei dem Finanzierungsprofil in 1.000 Unterrichtsstunden gerechnet, hätten sich Werte für die Testgröße zu 13,0 bzw. 36,8 ergeben, beides Werte die 11,1 (0,95-Quantil der χ^2-Verteilung mit 5 Freiheitsgraden) übersteigen.

Der χ^2-Anpassungstest ist offensichtlich nicht robust gegen die Verwendung verschiedener Skalen des Zählmerkmals. Außerdem berücksichtigt das Verfahren weder, in welche Richtung beobachtete von den erwarteten Werten abweichen, noch für welche Kategorien große oder kleine Abweichungen festzustellen sind. Es handelt sich deshalb um ein Verfahren, welches, ohne die Kategorien im Einzelnen zu berücksichtigen, nur einen globalen Vergleich für die Prüfung berücksichtigt.

Die Vorgehensweise in Anlehnung an den χ^2-Anpassungstest ist für die Konstruktion eines Analyse- und Typisierungsverfahrens lediglich der erste Schritt. In einem weiteren Schritt wird es nun zu einem globalen Indikator für den Vergleich von Profilen mit einem Vergleichsprofil ausgebaut.

6 Kennzeichnung der Ähnlichkeit zwischen Profilen und einem Vergleichsprofi

6.1 Der Kontingenzkoeffizient als Indikator

Der „Abwesenheit“ einer plausiblen Anpassung zwischen empirischen Werten und einem Modell im Sinne des χ^2-Anpassungstest entspricht das Konzept der „Unabhängigkeit“ im Sinne des χ^2-Unabhängigkeitstest in Kontingenztafeln. Auch dort werden Abweichungen zwischen beobachteten und aufgrund des Unabhängigkeitsmodells erwarteten Werten quadriert und mit den erwarteten Werten gewichtet zu einer χ^2-Testgröße addiert.

Als Maß für die Stärke des Zusammenhangs dient der Pearsonsche Kontingenzkoeffizient C (z. B. Bortz 1989, S. 286f.). Dieses Konzept wird als Indikator für die Ähnlichkeit zwischen Profilen und einem Vergleichsprofil genutzt.

$$C = \sqrt{\frac{\chi^2}{n+\chi^2}}$$

Es gilt: $0 \le C < 1$.

Dieses Maß weist bei Identität zwischen einem Profil und einem Vergleichprofil den Wert Null auf. Je stärker der Unterschied, desto größer ist der Wert des Kontingenzkoeffizienten. Allerdings wird der Wert 1 nicht angenommen, denn der maximal mögliche Wert hängt von der Zahl k der Kategorien ab.

$$C_{\max} = \sqrt{\frac{k-1}{k}}$$

Es bietet sich an, den Wert des Kontingenzkoeffizienten ins Verhältnis zu dem maximal möglichen Wert zu setzten.

$$C_{korr} = \frac{\sqrt{\frac{\chi^2}{n+\chi^2}}}{\sqrt{\frac{k-1}{k}}} = \sqrt{\frac{k}{k-1} \times \frac{\chi^2}{n+\chi^2}}$$

Für den korrigierten Kontingenzkoeffizienten gilt: $0 \le C \le 1$.

Mit Hilfe des Kontingenzkoeffizienten lässt sich beurteilen, wie stark die Abweichung eines Profils von einem Vergleichsprofil ist. Entsprechende Überlegungen liefern ein weiteres, aber trennschärferes Konzept für die Güte der Anpassung (goodness of fit).

Im Anwendungsbeispiel Volkshochschul-Statistik ist es damit beispielsweise möglich, alle Volkshochschulen nach dem Ausmaß der Ähnlichkeit mit dem Vergleichsprofil zu ordnen. Dieses Ziel war mit der Testgröße des χ^2-Anpassungstest (vgl. Abschnitt 5.2, Seite 7) nicht erreichbar. Allerdings beschreibt der Indikator weiterhin das Ausmaß einer *globalen* Ähnlichkeit ohne Hinweis darauf, auf welche Kategorien Unähnlichkeiten in welche Richtung zurückzuführen sind. Deshalb ist das Konzept für ein Analyse- und Typisierungsverfahrens von Profilen noch in einem weiteren Schritt auszubauen.

Für die Anwendungsbeispiele des vorangehenden Abschnitts werden die Ergebnisse als Werte für den korrigierten Kontingenzkoeffizienten präsentiert.

Tabelle 8 Beispielergebnisse

Vergleich	Kategorien	n	χ^2	C	C_{korr}
Finanzierungsprofil VHS 1	4	2.060 (in 1.000 €)	131	0,24	0,57
Finanzierungsprofil VHS 2	4	21.471 (in 1.000 €)	1.276	0,24	0,56
Programmprofil VHS 1	6	29.172	13.000	0,56	0,82
Programmprofil VHS 2	6	210.231	37.000	0,68	0,68

Damit wird das Problem deutlich: Obwohl offensichtlich die Finanzierungsprofile der beiden Volkshochschulen deutlich voneinander abweichen, haben sie etwa den gleichen Indikatorwert von über 0,5. Sowohl die Tatsache, dass die Programmprofile stärker von dem nordrhein-westfälischen Vergleichprofil abweichen als im Fall der Finanzierungsprofile, als auch der Fakt, dass die VHS 1 vergleichsweise stärkere Unterschiede im Programmprofil aufweist als die VHS 2, ist weder an den Ausgangstabellen noch an den Abbildungen auf den ersten Blick zu erkennen. Hier lohnt sich also offensichtlich ein Verfahren, was es ermöglicht aus den komplexen Strukturen in einfacher Weise Schlüsse zu ziehen. Trotzdem greifen die globalen Verfahren, die die Abweichungen nicht nach Kategorien verorten können, zu kurz und es bedarf daher einer Erweiterung des Konzepts.

6.2 Kontingenzkomponenten als kategorienspezifische Indikatoren

In dem χ^2-Ansatz des vorhergehenden Abschnitts ist erkennbar, dass der globale Wert einer Überprüfung der Anpassung und damit auch der Wert des korrigierten Kontingenzkoeffizienten sich summativ aus Komponenten zusammensetzt. Die Erweiterung des Konzepts besteht darin, dass der Übergang von einer χ^2-Summe zu einem Kontingenzkoeffizienten für jeden Summanden einzeln vollzogen wird. Die Ergebnisse werden Kontingenzkomponenten genannt, weil sie in einem ähnlichen Verhältnis zu dem Kontingenzkoeffizient stehen wie die Komponenten eines Vektors im n-dimensionalen Raum. [4]

Kontingenzkomponenten $\{c_1,...,c_k\}$ werden definiert durch $|c_i| = \sqrt{\frac{\frac{(n_i - np_{i;0})^2}{np_{i;0}}}{n + \frac{(n_i - np_{i;0})^2}{np_{i;0}}}}$ für $i = 1,...,k$.

Die Kontingenzkomponenten c_1 werden mit dem Vorzeichen versehen, welches sich aus der Differenz $n_i - np_{i;0}$ ergibt. Insofern beschreibt jede Komponente für eine Kategorie in welche Richtung und in welcher Stärke sich beobachtete von erwarteten Werten unterscheiden.

Sie haben die Eigenschaften:

1. Summeneigenschaft $C = \sqrt{c_1^{\,2} + ... + c_k^{\,2}}$

2. Alle Kontingenzkomponenten liegen zwischen -1 und 1.

3. Wenn alle Kontingenzkomponenten den Wert Null haben ist auch der Kontingenzkoeffizient gleich 0 und das Profil stimmt mit dem Vergleichsprofil überein.

Da in praktischen Anwendungen die Frage nach der maximalen Länge von Kontingenzkomponenten keine besondere Rolle spielt, wird hier darauf verzichtet, das Konzept der Korrektur aus dem vorhergehenden Abschnitt auf die Komponenten zu übertragen. Klar ist jedenfalls, dass die Länge einer Komponente (ihr Betrag) durch den Betrag des Kontingenzkoeffizienten C begrenzt bleibt. Hat eine Komponente dieselbe Länge müssen alle anderen Komponenten den Wert Null haben.

Mit den Kontingenzkomponenten steht jetzt ein Mittel zu Verfügung, die Art der Abweichung eines Profils von einem Vergleichprofil nach Kategorien differenziert zu identifizieren und dies für die Typisierung von Profilen zu nutzen.

Für die Profilvergleiche in den Anwendungsbeispielen werden die Kontingenzkomponenten als Abbildungen präsentiert.

[4] Man entsinne sich an Beispiele aus der Physik: Ein Geschwindigkeitsvektor im 3-dimensionalen Raum lässt sich eindeutig in eine waagrechte Komponente, eine senkrechte Komponente und eine Komponente „in die Tiefe“ zerlegen.

Beispiel 5 Kontingenzkomponenten bei Finanzierungsprofilen
Das Beispiel setzt das Beispiel 3, Seite 9 fort.

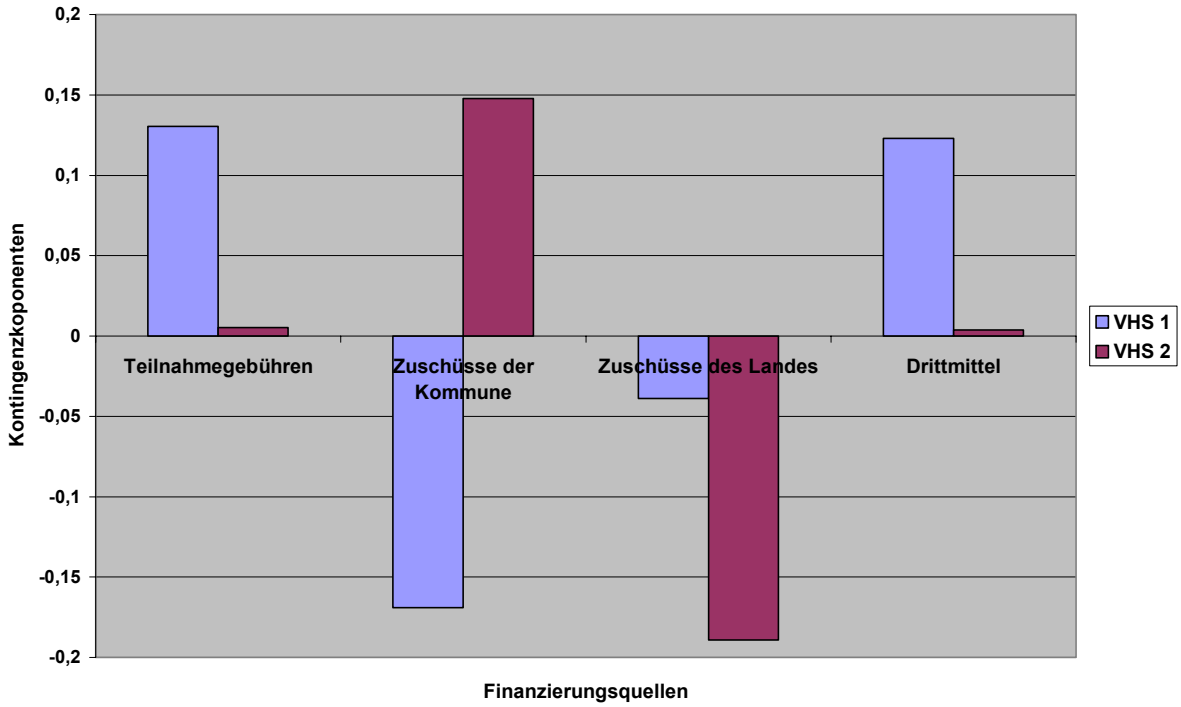


Abbildung 3 Vergleich von Finanzierungsprofilen: Kontingenzkomponenten

Wesentlich deutlicher als in der Abbildung 1, Seite 4 werden die Auffälligkeiten mit Hilfe der Kontingenzkomponenten identifizierbar.

- *Zunächst bleiben im Fall der Finanzierungsprofile für beide Volkshochschulen die Komponenten in einer Zone zwischen -0,2 und +0,2.*
- *Die VHS 1 kompensiert die unterdurchschnittliche Rolle der kommunalen Zuschüsse durch Teilnahmegebühren einerseits und Drittmittel andererseits.*
- *Bei der VHS 2 spielen die Zuschüsse vom Land eine untergeordnete Rolle.[5] Die besondere Auffälligkeit besteht in der überdurchschnittlich starken Rolle der kommunalen Zuschüsse. Die Rolle von Teilnahmegebühren und Drittmittel ist im Vergleich zum Landesprofil nicht auffällig.*

[5] Möglicherweise ist dies so, weil das Unterrichtsstundenvolumen weit über das förderbare Grundangebot nach WbG NW (Weiterbildungsgesetz Nordrhein-Westfalen) hinausgeht.

Beispiel 6 Kontingenzkomponenten bei Veranstaltungsprofilen
Das Beispiel setzt das Beispiel 4, Seite 9fort.

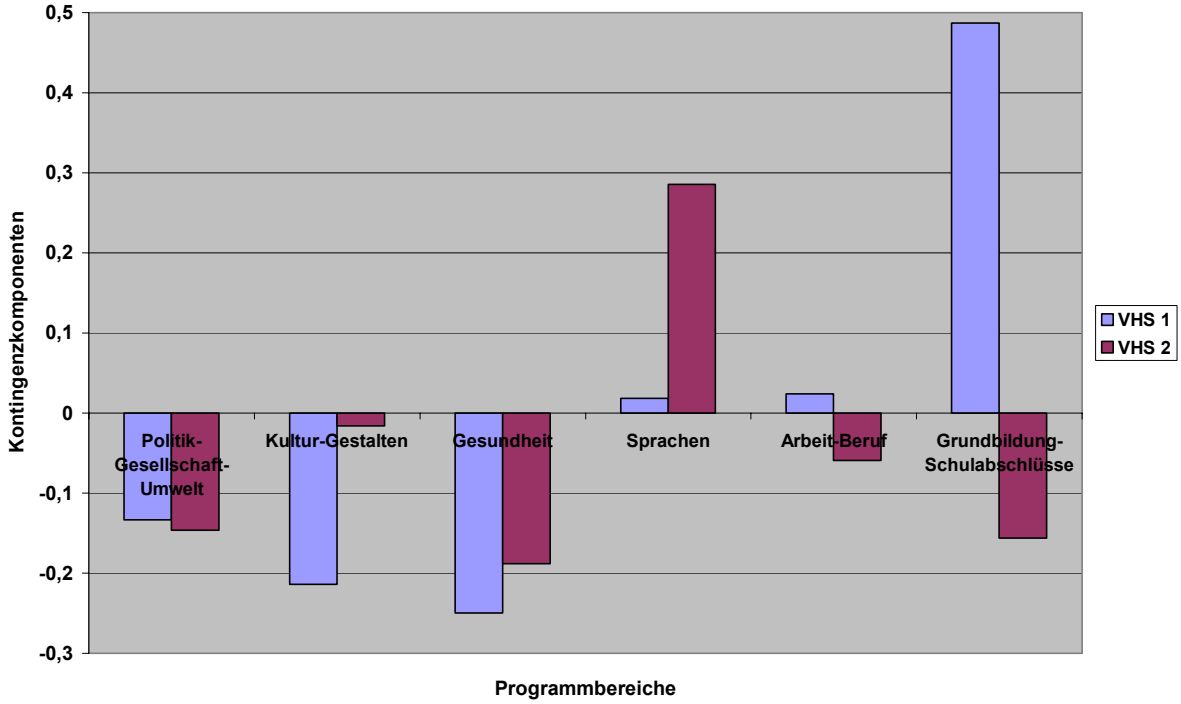


Abbildung 4 Vergleich von Programmprofilen: Kontingenzkomponenten

Mit Hilfe der Kontingenzkomponenten werden die Charakteristika der Programmprofile deutlich:

- *Die stärkste Auffälligkeit bei der VHS 1 besteht in dem Unterrichtsstundenvolumen im Programmbereich Grundbildung-Schulabschlüsse, und zwar im Sinn eines überdurchschnittlichen Ausbaus. Die beiden anderen „qualifizierenden" Programmbereiche Sprachen und Arbeit-Beruf weisen keine spezifische Abweichung vom dem Vergleichsprofil auf. In den allgemeinbildenden Programmbereichen ist das Unterrichtsstundenvolumen jeweils unterdurchschnittlich, und zwar von Politik-Gesellschaft-Umwelt über Kultur-Gestalten zu Gesundheit in zunehmendem Maße.*
- *Bei der VHS 2 ist allein Sprachen überdurchschnittlich stark ausgebaut. Wenig auffallend sind die Programmbereiche Kultur-Gestalten sowie Arbeit-Beruf, während die Programmbereiche Politik-Gesellschaft-Umwelt, Gesundheit und Grundbildung-Schulabschlüsse unterdurchschnittlich ausgebaut sind.*

Es sei an dieser Stelle betont, dass die Auffälligkeiten in Relation zum Vergleichmodell keineswegs der Rangordnung der Anteile (vgl. Abbildung 2, Seite 5) entsprechen. Die Kontingenzkomponenten berücksichtigen die Größenverhältnisse des Vergleichprofils und betonen relative Auffälligkeiten.

In einem abschließenden Abschnitt wird das Konzept um einen Vorschlag erweitert, wie die Kontingenzkomponenten praktisch für eine Typisierung sowie für Zeitreihenanalysen eingesetzt werden können.

6.3 Typisierung von Profilen mit Hilfe von Kontingenzkomponenten

Die Kontingenzkomponenten bilden die vielschichtige Struktur eines Profils gemessen an einem Vergleichsprofil für eine einfache Typisierung noch zu komplex ab. Gesucht ist ein Verfahren, aus den Kontingenzkomponenten die wesentlichen Informationen herauszuziehen, also eine Informationsreduktion zugunsten der Einfachheit der Beschreibung zu betreiben.

Bei dem Vergleich von Finanzprofilen und Programmprofilen fiel auf, dass im einen Fall alle Komponenten beider Untersuchungseinheiten nur in einem Intervall [-0,2; +0,2] variierten,

während im anderen Fall beide Untersuchungseinheiten Komponenten aufwiesen, die aus diesen Intervall herausragten (vgl. Abbildung 3, Seite 13, und Abbildung 4, Seite 14).

Dieser Gedanke wird in diesem Abschnitt aufgegriffen und systematisiert. Es wird ein um Null symmetrisches („Toleranz“-)Intervall [-x;+x] mit der Regel angegeben, dass Kontingenzkomponenten nur beachtet werden sollen, wenn ihre Werte außerhalb des Intervalls liegen.

Beispiel 7 Auf dem Weg zum Profiltyp für Finanzierungsprofile
Das Beispiel setzt das Beispiel 5, Seite 13 fort.

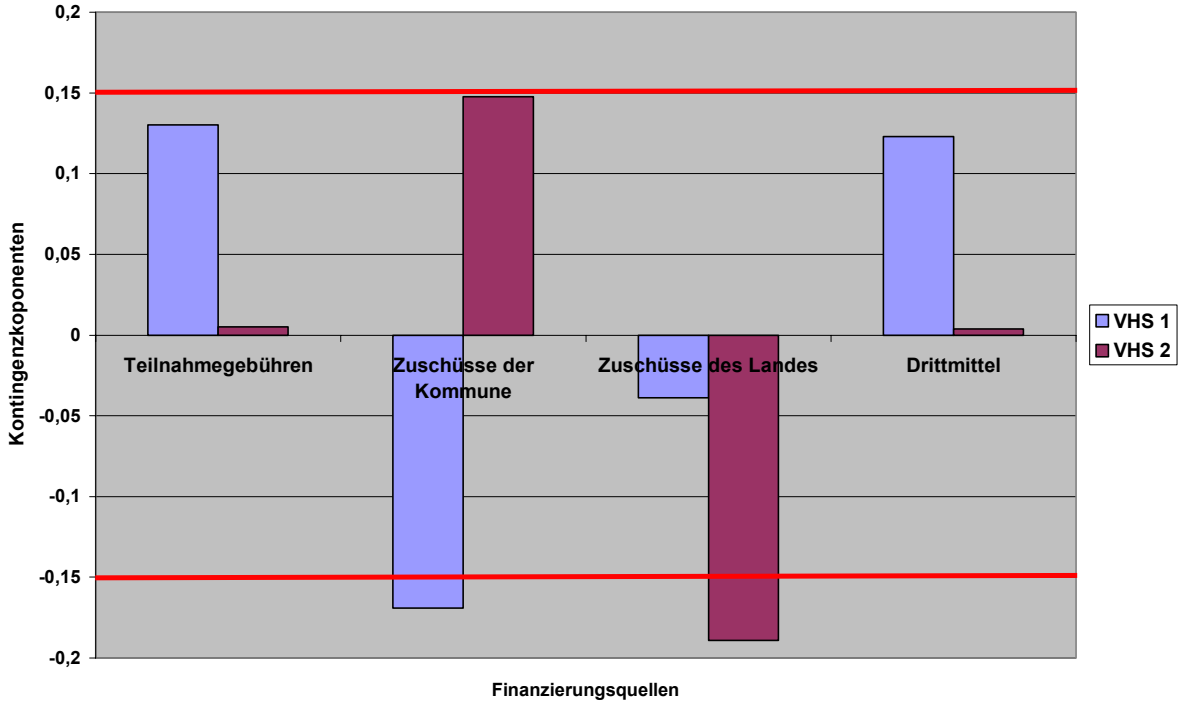


Abbildung 5 Vergleich von Finanzierungsprofilen: Auffällige Kontingenzkomponenten

Bei einem Toleranzintervall von [-0,15;+0,15] ist bei der VHS 1 die geringe Rolle der kommunalen Zuschüsse auffällig, bei der VHS 2 die geringe Rolle der Landeszuschüsse. Dort bleibt die starke Rolle der kommunalen Zuschüsse knapp unter der Auffälligkeit relativ zu den gewählten Toleranzgrenzen.

Beispiel 8 Auf dem Weg zum Profiltyp für Veranstaltungsprofile
Das Beispiel setzt das Beispiel 6, Seite 14 fort.

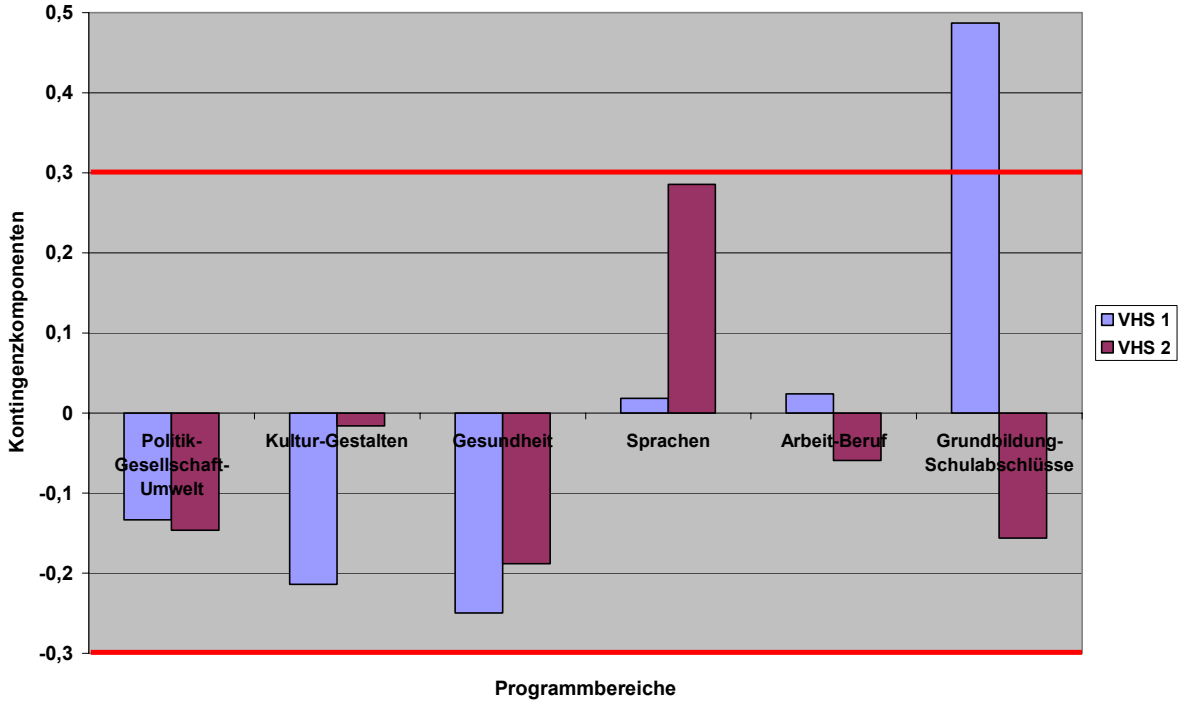


Abbildung 6 Vergleich von Programmprofilen: Auffällige Kontingenzkomponenten

Selbst bei einen „anspruchsvollen" Toleranzintervall von [-0,3;+0,3] ist bei der VHS 1 die starke Rolle des Programmbereichs Grundbildung-Schulabschlüsse auffällig, während bei der VHS 2 der Programmbereich Sprachen gerade noch unterhalb der Schwelle bleibt.

Für die Bestimmung eines geeigneten Toleranzintervalls liefern die zu Grunde gelegten Verfahren keine theoretische Begründung. Es ist jedoch offensichtlich, welche Folgen es hätte, wenn das Intervall zu eng, d. h. die Toleranz zu knapp gewählt würde: Fast alle Komponenten – ob positiv oder negativ, ob betragsmäßig klein oder groß – würden als auffällig betrachtet werden müssen. Zur Beschreibung des entsprechenden Typs müssten alle Kategorien herangezogen werden. Mit dem Vergleichsprofil hinreichend ähnliche Profile würden ausgeschlossen.

Umgekehrt hätte auch ein zu großes Toleranzintervall unerwünschte Folgen. Zwar würden nur wenige Komponenten aus dem Intervall „herausragen". Die meisten Abweichungen würden toleriert. Die Beschreibung des entsprechenden Typs könnte sich auf wenige Kategorien beschränken. Die Kehrseite wäre allerdings, dass zu viele Untersuchungseinheiten als dem Vergleichsprofil zu ähnlich eingestuft würden. Sie wären sozusagen „typenlos".

Die Informationstechnik macht es möglich, das Problem unaufwendig durch Variation der Toleranzbreite zu optimieren. Dabei wird als Ziel verfolgt, dass nur maximal 10% der Untersuchungseinheiten ohne erkennbaren Typ eingestuft werden.

Die Analyse der Finanzierungsprofile und der Programmprofile aller Volkshochschulen 2003 (= Untersuchungseinheiten) wurde mit verschiedenen Toleranzintervallen durchgespielt.

Tabelle 9 Optimierung des Toleranzintervalls für Profilanalysen

Toleranzintervall	Finanzierungsprofil Anteil der VHS ohne erkennbaren Typ	Programmprofil Anteil der VHS ohne erkennbaren Typ
[0,0]	0,0%	0,0%
[-0,1; +0,1]	0,2%	0,4%
[-0,2; +0,2]	2,4%	4,6%
[-0,25;+0,25]	6,0%	**10,5%**
[-0,29;+0,29]	**10,5%**	**20,5%**
[-0,3; +0,3]	12,0%	30,7%
[-0,4; +0,4]	33,7%	50,1%
[-0,5; +0,5]	69,6%	65,0%
[-0,6; +0,6]	87,1%	77,4%
[-0,7; +0,7]	94,6%	87,3%
[-0,8; +0,8]	97,7%	93,5%
[-0,9; +0,9]	99,8%	97,7%
[-1,0;+1,0]	100,0%	99,7%[6]

Die graphische Veranschaulichung der Abhängigkeiten des Anteils nicht erkennbarer Typen in der Profilanalyse von der Toleranzbreite (s. Abbildung 7, unten) kann ähnlich wie die Risikofunktion eines Test interpretiert werden. Je größer die Toleranzbreite gewählt wird, desto höher wird das Risiko, dass der Typ einer Untersuchungseinheit nicht erkannt wird.

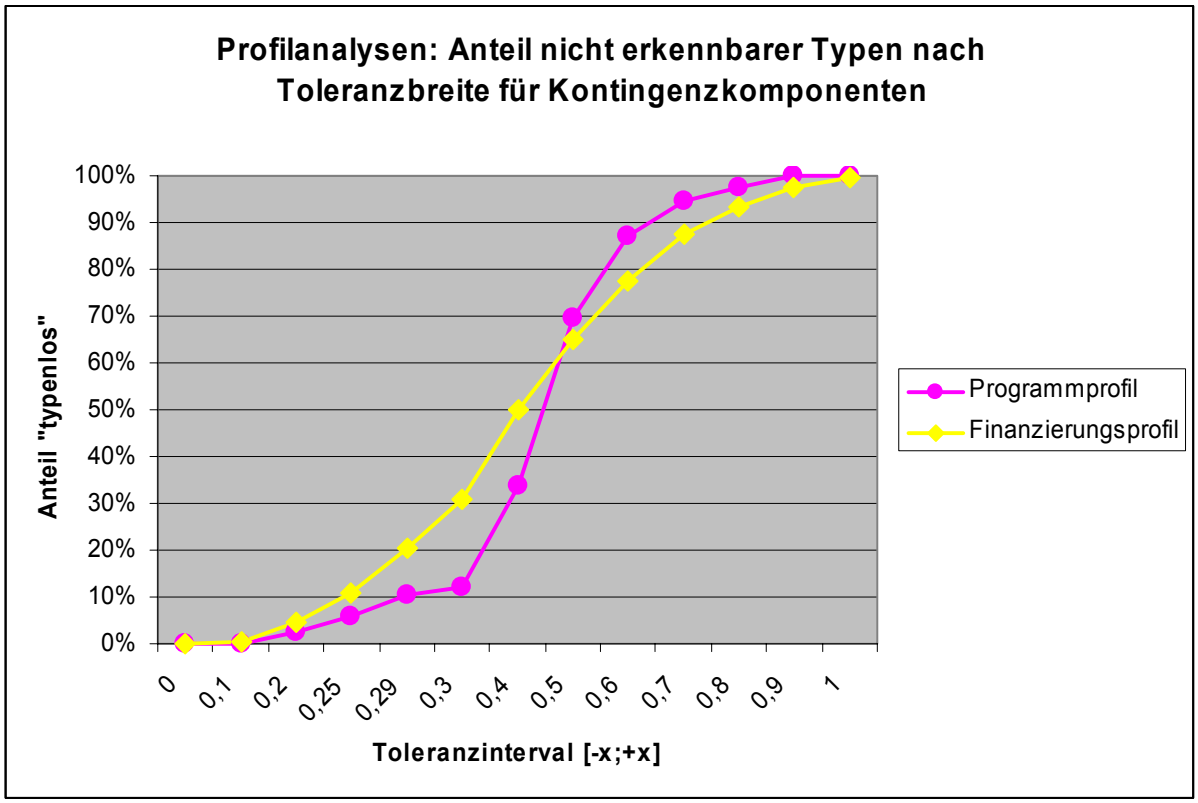


Abbildung 7 Abhängigkeit des Anteil nicht erkennbarer Typen bei Profilanalysen von der Toleranzbreite (980 Volkshochschulen 2003)

Für beide Profilanalysen – Finanzierungsprofil und Programmprofil – ist dasselbe Berichtsjahr (2003) mit denselben Volkshochschulen zu Grunde gelegt. Insofern ist die Risikokurve eine Eigenschaft, die für das analysierte Profil charakteristisch ist. Mit der Zielvorgabe eines Risikos von ca. 10% ist ein optimales Toleranzintervall für Kontingenzkomponenten im Fall

[6] Für 3 Untersuchungseinheiten liegt jeweils eine Kontingenzkomponente aufgrund von Rundungsfehlern in der dritten Dezimale über +1.

der Programmprofilanalyse durch [-0,25;+0,25] gegeben, während im Fall der Finanzierungsprofilanalyse ein breiteres Toleranzintervall [-0,29;+0,29] gewählt werden muss.

Die Kontingenzkomponenten und das Toleranzintervall induzieren für die Profilvergleiche einer Gruppe von Untersuchungseinheiten *Klassen* von Untersuchungseinheiten („Typ“) mit jeweils ähnlichem Muster der Kontingenzkomponenten, die außerhalb des Toleranzintervalls auffallen. Für eine universell taugliche Beschreibung der Klassen wird vorgeschlagen, die Kategorien mit Buchstaben in alphabetischer Reihenfolge zu benennen und für einen Typ diejenigen Kategorien mit einem *großen* Buchstaben zu bezeichnen, deren Kontingenzkomponenten Werte über der oberen Grenze des Toleranzintervalls aufweisen, und entsprechend mit *kleinen* Buchstaben diejenigen Kategorien, deren Kontingenzkomponenten unter der unteren Toleranzgrenze bleiben.

Dabei sind verschieden informative Stufen für die ***Typenbildung*** denkbar:

> ***Variante (1)***: Nennung ausschließlich der Kategorien mit der betragsmäßig größten Kontingenzkomponente, sei sie positiv oder negativ. Die Anzahl verschiedener Typen ist maximal 2 x die Zahl der Kategorien der Profile (s. Tabelle 10, Seite 19, und Tabelle 11, Seite 19), dazu der Typ „typenlos“.
>
> *Beispiele: Finanzierungsprofil – Typ D heißt, dass die Finanzierungsquelle „Drittmittel“ relativ zum Vergleichsprofil am auffälligsten nach oben abweicht. Der Typ b heißt, dass die Finanzierungsquelle „Kommunale Zuschüsse“ am auffälligsten nach unten abweicht.*
>
> ***Variante (2)***: Aufzählung aller auffallenden Kategorien in lexikographischer Reihenfolge
>
> *Beispiele (setzt die Beispiele 7, Seite 15 und 8, Seite 16 fort): Bei einem Toleranzintervall [-0,2; +0,2] – Beim Finanzierungsprofil weisen VHS 1 und VHS 2 keine hinreichend erkennbaren Auffälligkeiten zum Vergleichsprofi auf, beide sind „typenlos“; beim Veranstaltungsprofil ist VHS 1 vom Typ bcF und VHS 2 vom Typ C*
>
> ***Variante (3)***: zunächst alle Kategorien mit auffallend großen, dann alle Kategorien mit auffallend kleinen Kontingenzkomponenten. Diese Variante gibt in Erweiterung der Variante (2) auch die Rangordnung gemessen an der Länge der Kontingenzkomponenten wieder.
>
> *Beispiele (setzt das Beispiel 8, Seite 16 fort): Bei einem Toleranzintervall [-0,2; +0,2] – Beim Veranstaltungsprofil repräsentiert der Typ Fcb der VHS 1 die Kontingenzkomponenten in der Reihenfolge ihrer Auffälligkeit (Länge).*

Für ein Toleranzintervall von [-0,29; +0,29] und die Analyse von Finanzierungsprofilen ist für die Volkshochschulen 2003 unter Berücksichtigung der betragsmäßig größten Kontingenzkomponente die folgende Typenzuordnung feststellbar.

Tabelle 10 Übersicht über festgestellte Typen bei der Finanzierungsprofilanalyse (980 Volkshochschulen 2003) – Variante (1): Nur die größte Kontingenzkomponente ist berücksichtigt

Auffällig viel/wenig Beträge	Finanzierungsquelle	Typ Variante (1)	Anzahl	In %
-	-	„“	*103*	10,5%
viel	Teilnahmegebühren	A	277	28,3%
	Zuschüsse von Kommunen	B	101	10,3%
	Zuschüsse vom Land	C	117	11,9%
	Drittmittel	D	96	9,8%
	zusammen		*591*	*60,3%*
wenig	Teilnahmegebühren	a	5	0,5%
	Zuschüsse von Kommunen	b	29	3,0%
	Zuschüsse vom Land	c	18	1,8%
	Drittmittel	d	234	23,9%
	zusammen		*286*	*29,2%*
insgesamt			*980*	*100%*

103 (= 10,5%) der VHS können in Relation zum Vergleichsprofil nicht als besondere Typen erkannt werden. Wesentlich mehr Volkshochschulen (60,3%) fallen dadurch auf, dass die auffälligste Abweichung in einer überdurchschnittlichen Stärke einer Finanzierungsquelle besteht. Die Schlüsselrolle hierfür haben Teilnahmegebühren als Finanzquelle (277 von 591 VHS). Unter den 23,9% der Volkshochschulen mit unterdurchschnittlicher Rolle einer Finanzierungsquelle bilden die Volkshochschulen mit auffällig niedrigen Drittmittel eine dominierenden Rolle (234 von 286 VHS); die VHS mit vergleichsweise niedrigen Teilnehmergebühren sind sehr selten (5 VHS von 286).

Für ein Toleranzintervall von [-0,25; +0,25] und die Analyse von Programmprofilen ist für die Volkshochschulen 2003 unter Berücksichtigung der betragsmäßig größten Kontingenzkomponente folgende Typenzuordnung feststellbar.

Tabelle 11 Übersicht über festgestellte Typen bei der Programmprofilanalyse (980 Volkshochschulen 2003) – Variante 1: Nur die größte Kontingenzkomponente ist berücksichtigt

Auffällig viel/wenig U.-Stunden	Programmbereich	Typ Variante (1)	Anzahl	In %
-	-	„“	*105*	*10,7%*
viel	Politik-Gesellschaft-Umwelt	A	53	5,4%
	Kultur-Gestalten	B	133	13,6%
	Gesundheit	C	273	27,9%
	Sprachen	D	75	7,7%
	Arbeit-Beruf	E	82	8,4%
	Grundbildung-Schulabschlüsse	F	93	9,5%
	zusammen		*709*	*72,3%*
wenig	Politik-Gesellschaft-Umwelt	a	0	0,0%
	Kultur-Gestalten	b	1	0,1%
	Gesundheit	c	6	0,6%
	Sprachen	d	9	0,9%
	Arbeit-Beruf	e	32	3,3%
	Grundbildung-Schulabschlüsse	f	118	12,0%
	zusammen		*166*	*16,9%*
insgesamt			*980*	*100,0*

Bei der Programmprofilanalyse ist der Anteil der „typenlosen“ Volkshochschulen ähnlich hoch wie bei der Finanzierungsprofilanalyse (10,7%). Dafür hat die Wahl des Toleranzintervalls unter dieser Zielvorgabe gesorgt. Anders als bei dem Finanzierungsprofil sind bei der Programmprofilanalyse Volkshochschulen mit auffälliger unterdurchschnittlicher Ausprägung wesentlich seltener (16,9%). Die Schlüsselrolle hat hier der Programmbereich Grundbildung-Schulabschlüsse (118 von 166 VHS). Bei den fünf Sechstel der Volkshochschulen ist das Auffälligste ein vergleichsweise hoher Ausbau eines Programmbereichs. Hier bilden die Volkshochschulen mit auffällig hohem Ausbau von Gesundheitsbildung die größte Gruppe (273 von 709 VHS).

Im Erweiterung zu der einfachsten Variante (1) bei der Typenbildung ermöglicht insbesondere die Variante (3), die Typen gestuft nach der Rolle der Komponenten aufzubauen: Mittels des Verfahrens (1) werden die Typen als „Grobtypen“ je nach vorherrschender Komponente unterschieden. Das weitere Verfahren wird herangezogen, um „Untertypen“ nach der nächst auffälligen Komponente zu bilden. Das gibt die Möglichkeit, besondere Affinitäten zwischen den Kategorien festzustellen (s. Abbildung 8, unten). Welche Kombinationen von Typkomponenten treten häufiger/weniger häufig auf? Diese Betrachtung schöpft die Information zumindest aus den beiden auffälligsten Kontingenzkomponenten aus, während die Variante (1) nur eine, nämlich die betragsmäßig größte berücksichtigt. Dies hat aber für weiterführende Analysen, insbesondere Langzeitanalysen, Vorteile (s. Abschnitt 7, Seite 21).

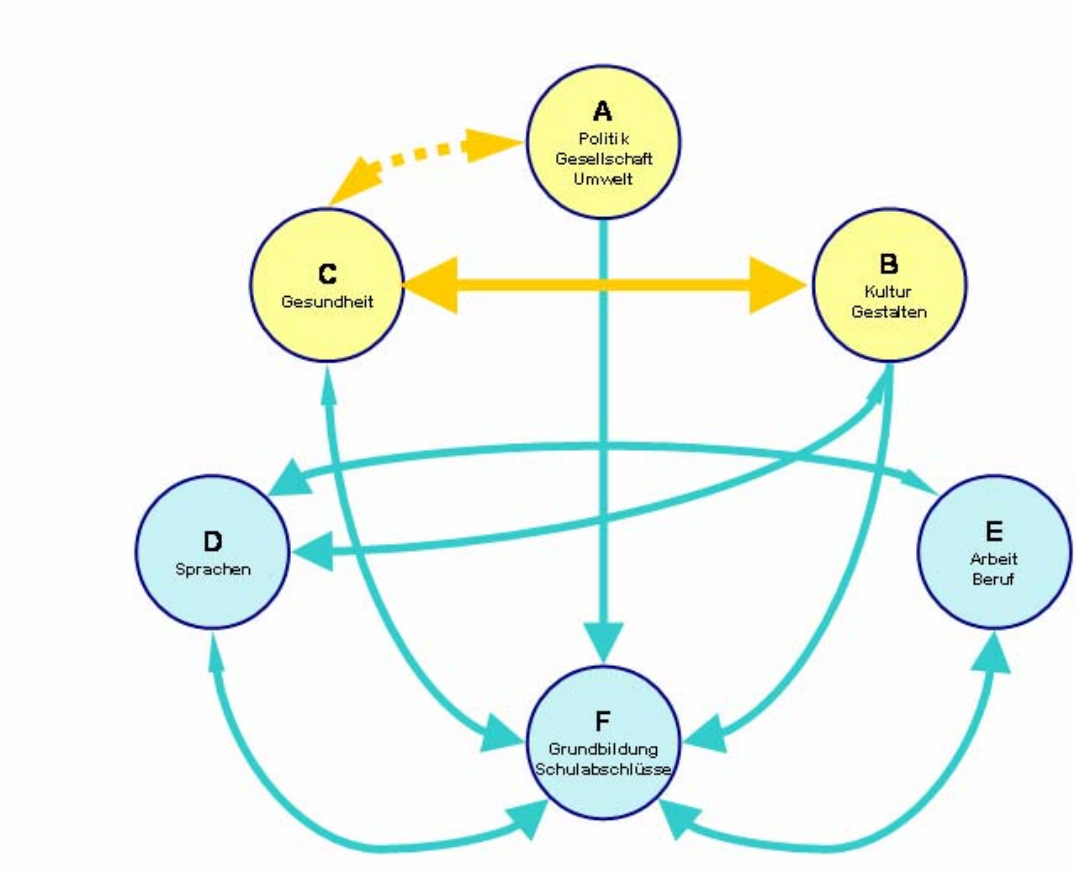


Abbildung 8 Affinitäten zwischen Programmbereichen Volkshochschulen 2003 auf der Basis der beiden auffälligsten Kontingenzkomponenten

Die Abbildung veranschaulicht die häufiger auftretenden Kombinationen der beiden auffälligsten Kontingenzkomponenten. Deutlich wird der Zusammenhang der drei Programmbereiche A, B und C, in denen persönliche Entfaltung und gesellschaftliche Teilhabe im Vordergrund stehen, weniger Qualifizierung wie in den Programmbereichen D, E und F. Sie treten häufig zusammen als auffällig starke Komponenten auf („goldene“ Pfeile). Weitere feststellbare Zusammenhänge sind von der Art, dass ein auffällig starker Programmbereich einen auffällig schwachen Programmbereich bedingt (blaue Pfeile). Diese Beziehungen sind nicht symmetrisch. Die Stärke des Zusammenhangs ist in Abbildung 8 durch

die unterschiedliche Stärke der Pfeilspitzen dargestellt. So ist die hemmende Wirkung eines stark ausgebauten Programmbereichs Arbeit-Beruf auf den Sprachenbereich stärker als umgekehrt. Die Sonderstellung des Programmbereichs Grundbildung-Schulabschlüsse ist daran zu erkennen, dass eine Stärke aller anderen Programmbereiche mit seiner Rücknahme einhergeht.

7 Weiterführende Analyseansätze

7.1 Zusammenhang zwischen Profiltyp und anderen Merkmalen

Die Transformation eines komplexen Konstrukts wie dem Profil als Häufigkeitsverteilung in ein einfaches nominal skaliertes Merkmal, dem „Typ“ nach Variante (1) mit den möglichen Ausprägungen A, B, C, …, a, b, c, … je nach Zahl der Klassen der Häufigkeitsverteilung, macht es möglich, Korrelationen mit anderen Merkmalen einer Weiterbildungseinrichtung zu untersuchen.

Hierfür als Anregung zwei Beispiele. Das erste Beispiel geht der Frage nach, ob die 2003 für VHS festgestellten Typen von Programmprofilen mit der Einwohnerzahl der (politischen) Versorgungsgebiete korrelieren. Das heißt, das Ergebnis der Typenbildung (vgl. Tabelle 11, Seite 19) ist nach Klassen von Einwohnerzahlen (Stichtag 31.12.2003) aufzuschlüsseln. Als Klassierung für die Zahl der Einwohner wird die im Statistischen Jahrbuch der Gemeinden verwendete Einteilung übernommen. Tabelle 12 (s. unten), illustriert das Ergebnis. Wegen der geringen Besetzungszahlen in einigen Zellen müssen auf der einen Seite „seltene“ Typen auf der anderen Seite die Einwohnerklassen über 100.000 Einwohner zusammengefasst werden. Die Zahlen belegen den Zusammenhang. Auffällige Kombinationen sind die besondere Betonung des Programmbereichs Sprachen – vermutlich wegen des starken Anteils der Deutschkurse für Menschen mit Migrationshintergrund in Ballungsgebieten – und des Programmbereichs Grundbildung-Schulabschlüsse. Währenddessen besitzen VHS in kleinen Gemeinden unter 20.000 Einwohner besonders große Profilkomponenten in den Programmbereichen Kultur-Gestalten und Gesundheit und haben den Programmbereich Grundbildung-Schulabschlüsse eher zurückgenommen. Untersuchenswert ist auch die Tatsache, dass solche VHS unter den Typenlosen, d. h. ohne auffällige Abweichung zum Bundesprofil, besonders selten sind.

Tabelle 12 Zusammenhang zwischen Programmprofiltypen und Gemeindegrößenklasse[7]

Auffällig viel/wenig U.-Stunden	Programmbereich	Typ (1)	Einwohnerzahl im Gebiet (in 1000)				Summe
			Über 100	50 bis unter 100	20 bis unter 50	Unter 20	
-		„“	17	37	47	*4*	105
viel	Politik-Gesellschaft-Umwelt	A	2	18	22	11	53
	Kultur-Gestalten	B	1	11	38	*83*	133
	Gesundheit	C	5	32	113	*123*	273
	Sprachen	D	*17*	22	34	2	75
	Arbeit-Beruf	E	2	30	37	13	82
	Grundbildung-Schulabschlüsse	F	*17*	33	39	*4*	93
wenig	Politik-Gesellschaft-Umwelt, Kultur-Gestalten, Gesundheit oder Sprachen	a, b, c oder d	2	1	9	4	16
	Arbeit-Beruf	e	4	4	21	3	32
	Grundbildung-Schulabschlüsse	f	7	20	74	17	118
Summe			74	208	434	264	980

[7] Besonders auffällige Felder sind hervorgehoben.

Das zweite Beispiel nutzt die vorgeschlagene Methode noch weitgehender aus. Die Frage ist hier: Hat das Finanzierungsprofil der Volkshochschulen in der Bundesrepublik Deutschland 2003 etwas mit dem Veranstaltungsprofil zu tun und wenn ja wie? Die Entscheidungsgrundlage ist eine zweidimensionale Zusammenführung der Tabelle 10, Seite 19 und Tabelle 11, Seite 19. Bei der Präsentation der Tabelle 13 (s. unten) sind zu gering besetzte Zellen durch Zusammenlegung von Typen berücksichtigt. Sie belegt einen Zusammenhang zwischen Programmprofiltyp und Finanzierungstyp. Auffällige Kombinationen sind

- eine besondere Betonung des Programmbereichs Gesundheit (Typ Programmbereich „C“) in Verbindung mit einem relativ hohen Anteile der Finanzierung aus Teilnehmergebühren (Typ Finanzierung „A“);

- ein vergleichsweise starker Anteil des Programmbereichs Arbeit-Beruf (Typ Programmbereich „D“) und ein relativ hoher Anteil der Finanzierung aus Drittmitteln (Typ Finanzierung „D“);

- dual dazu ein vergleichsweise geringer Anteil der Finanzierung aus Drittmitteln (Typ Finanzierung „d“) bei den VHS mit auffällig hohem Anteil im Programmbereich Arbeit-Beruf (Typ Programmbereich „D“).

Tabelle 13 Zusammenhang zwischen Veranstaltungsprofil und Finanzierungsprofil 2003[8]

Auffällig viel/ wenig U.-Stunden/ Finanzierung	Programmbereich	Typ (1)	Finanzierungsprofiltypen									Summe
			-	viel				wenig				
				Tn.-Gebühren	Kommune	Land	Drittmittel	Tn.-Gebühren	Kommune	Land	Drittmittel	
			„“	A	B	C	D	a	b	c	d	
-		„“	17	10	14	21	11	1	2	2	27	105
viel	Politik-Gesellschaft-Umwelt	A	8	13	2	11	3	-	2	-	14	53
	Kultur-Gestalten	B	6	53	18	3	5	1	3	1	43	133
	Gesundheit	C	8	136	22	18	6	-	10	8	65	273
	Sprachen	D	9	8	8	16	5	-	3	2	24	75
	Arbeit-Beruf	E	12	9	7	8	39	1	2	2	2	82
	Grundbildung-Schulabschlüsse	F	18	4	16	23	15	1	3	-	13	93
wenig	Politik-Gesellschaft-Umwelt, Kultur-Gestalten, Gesundheit oder Sprachen	a, b, c oder d	2	3	1	1	2	1	1	1	4	16
	Arbeit-Beruf	E	3	9	3	4	1	-	1	-	11	32
	Grundbildung-Schulabschlüsse	F	20	32	10	12	9	-	2	2	31	118
Summe			103	277	101	117	961	5	29	18	234	980

Diese Art der Analysen bezieht sich global auf eine definierte Gruppe von Weiterbildungseinrichtungen. Andere Analysen kommen lokal einzelnen Weiterbildungseinrichtungen zugute. Dazu gehören Vergleiche von Befunden der Einrichtung mit Befunden, die alle Einrichtungen in einem Gebiet – die Einrichtung selbst einschließend – betreffen. Dies war schon durch Berücksichtigung des Profils aller Einrichtungen in einem Land bei den Beispielen des Ab-

[8] Besonders auffällige Felder sind hervorgehoben.

schnitts 3, Seite 2, angedeutet. Der Aspekt der zeitlichen Entwicklung lässt sich zwar prinzipiell in alle Fragestellungen integrieren, wegen mehreren möglichen Varianten ist solchen Fragen jedoch ein weiterer eigener Unterabschnitt gewidmet.

7.2 Gebietsspezifische Analysen

Organisierte Weiterbildung in Einrichtungen vollzieht sich praktisch in Regionen. Zwar unterliegen verschiedene Regionen eines Landes dem gleichen Weiterbildungsgesetz.[9] doch in den neueren Fassungen oder in den vorbereitenden Diskussionen zu Neufassungen wird in allen die dezentrale Verantwortung in den Regionen gestärkt. Insofern ist es ein nahe liegendes Anliegen der Weiterbildungspraxis, Profilvergleiche zwischen der eigenen Einrichtung, ihrer Region, ihrem Land und dem Bundesprofil anzustellen, nicht etwa, um sich den Profilen übergeordneter Gebiete anzugleichen, sondern um sich argumentativ mit den Unterschieden auseinanderzusetzen.

An Beispielmaterial aus der Volkshochschul-Statistik 2003 soll der Ansatz demonstriert werden. Die Argumente im BMBF-Programm „Lernende Regionen – Förderung von Netzwerken“,[10] den Begriff der Region nicht scharf zu definieren oder abgrenzbar zu operationalisieren, sind vor dem Hintergrund seiner Ziele verständlich. Eine statistische Datenanalyse muss jedoch auf der Eigenschaft eines Regionensystems ansetzen, um eine „Zerlegung“ des Landes bzw. der Bundesrepublik zu bilden, d. h. es muss die beiden Forderungen erfüllen:

- die Regionen müssen das Land abdecken;
- zwei Regionen dürfen sich nicht wechselseitig überlappen.

Mit den Raumordnungsregionen, die Siedlungsstruktur sowie Pendlerströme und weitgehend politische Gebietsgrenzen kleinräumig (Kreise) und großräumig (Länder) berücksichtigen, existiert ein mit allen Ländern abgestimmtes System von 97 Regionen (vgl. Pehl 2001). Besonders im Fall der Volkshochschulen als kommunale Weiterbildungszentren, kann dieses System nutzbringend angewendet werden, vor allem, weil über einen Service des Bundesamts für Bauwesen und Raumordnung (BBR 2003) viele Merkmale und Indikatoren für diese Regionen verfügbar sind und eine klein- wie großräumige Typisierung vorliegen.

Mit dieser Hilfe könnte der These nachgegangen werden, dass die Prägung von Profilen durch ein gemeinsames Weiterbildungsgesetz in einem Land weitaus schwächer ist als der siedlungsstrukturelle Typ der Region oder der Kommune einer Weiterbildungseinrichtung. Dies auszuführen, kann allerdings nicht das Ziel dieses Beitrags sein. Hier sollen lediglich beispielhaft die Veranstaltungsprofile von acht kommunalen Einrichtungen des gleichen Gemeindetyps, Kreis- und Regionstyps (Stadtkreise vom Typ Kernstädte in verstädterten Regionen) in acht verschiedenen Ländern zur Anregung dienen (s. Tabelle 14, Seite 24).

[9] Derzeit liegen in allen Ländern außer Berlin, Hamburg und Schleswig-Holstein jeweils spezifische Weiterbildungsgesetze vor.
[10] Vgl. auch die Internetseite des Projektträgers http://www.lernende-regionen.info/dlr/index.php.

Tabelle 14 Gebietsspezifische Typenbildung nach der Variante (2)[11]

Volkshochschule	Lübeck	Oldenburg	Münster	Kassel	Mainz	Augsburg	Rostock	Erfurt
Typ (BRD)	De	Ec	A	Ae	""	Be	F	Dc
Typ (Land)	Dec	E	Af	Ae	Dc	""	F	c
Typ (Region)	D	""	A	eDF	""	""	""	""
Region	**Schleswig-Holstein-Ost**	**Oldenburg**	**Münster**	**Hessen-Nord**	**Rhein-hessen-Nahe**	**Augsburg**	**Mittleres Mecklenburg/ Rostock**	**Mittelthüringen**
Typ (BRD)	E	Ed	""	Ed	""	Cf	F	""
Typ (Land)	""	Ed	""	Ed	""	Cf	F	""
Land	**Schleswig-Holstein**	**Niedersachsen**	**Nordrhein-Westfalen**	**Hessen**	**Rheinland-Pfalz**	**Bayern**	**Mecklenburg-Vorpommern**	**Thüringen**
Typ (BRD)	""	E	""	""	""	C	F	""

Aus der lokalen Sicht zeigt die Tabelle 14, dass der Typ bis auf zwei Ausnahmen (Münster und Kassel) mit dem regionalen Profil besser übereinstimmt, als mit dem des zugehörigen Landes oder im Vergleich zu dem Bundesprofil. Deutlichstes Beispiel ist Erfurt. Im Fall von Münster wird mit der Auffälligkeit des Programmbereichs Politik-Gesellschaft-Umwelt eine lokale Spezifität deutlich, denn das Regional-, das Landes- und das Bundesprofil scheinen gut überein zu stimmen. Im Fall von Kassel machen sich eine lokale Besonderheit und eine (andere) regionale Auffälligkeit bemerkbar. Im Vergleich zum Land wie zum Bundesprofil ist lokal der besonders starke Programmbereich Politik-Gesellschaft-Umwelt maßgebend. Im Vergleich zur Region der schwächer ausgeprägte Programmbereich Arbeit-Beruf. Gerade dieser Programmbereich ist in der Region im Vergleich zum Landes- und Bundesprofil bemerkenswert stark.

Die vergleichsweise enge Bindung zu den Regionen weniger zum Land zeigt sich auch darin, dass die Regionen im Vergleich zum Landes- oder Bundesprofil nahezu die gleichen Unterschiede aufweisen.

7.3 Zeitreihenanalyse

Eine Weiterbildungseinrichtung ermittelt für ein Berichtsjahr bezüglich des Merkmals Unterrichtsstunden nach Programmbereichen („Veranstaltungsprofil“) bezogen auf ein vorliegendes Vergleichsprofil (z. B. Aggregation für eigenes Land) und bei Anwendung eines Toleranzintervalls für Kontingenzkomponenten (z. B. Schwelle 0,25 für Betrag von Kontingenzkomponenten) ihren Profiltyp (Variante (1) bei Berücksichtigung der betragsmäßig größten auffälligen Kontingenzkomponente). Für eine Prognose oder die weitere Steuerung ist die Frage bedeutsam: Belegen die Daten früherer Berichtsjahre, dass sich der Profiltyp verändert hat (wenn ja, wie) oder eher dass das Veranstaltungsprofil unverändert blieb.

Um solchen *lokalen* Fragestellungen nachzugehen, ist es erforderlich, den Profiltyp der Einrichtung für mehrere Berichtsjahre festzustellen und die Ergebnisse als Zeitreihe zu betrachten. Hierfür gibt es keineswegs einen eindeutigen Weg. Mindestens vier Varianten sind hierbei zu durchdenken. Sie unterscheiden sich bei unveränderten anderen Parametern wie Vergleichsgebiet und Toleranzbreite in der Wahl des Bezugsjahrs des Vergleichsprofils:

[11] Auffallende Komponenten in absteigender Ordnung.

(a) Als Vergleichsprofil wird *für jedes Berichtsjahr d*as Profil des Gebiets *in demselben Berichtsjahr* gewählt.

(b) Als Vergleichsprofil wird *für alle Berichtsjahre* das Profil des Gebiets *im ersten Berichtsjahr* des betrachteten Zeitraums gewählt.

(c) Als Vergleichsprofil wird *für alle Berichtsjahre* das Profil des Gebiets im letzten Berichtsjahr des betrachteten Zeitraums gewählt.

(d) Als Vergleichsprofil wird *für alle Berichtsjahre* das Profil des Gebiets in einem geeigneten Berichtsjahr innerhalb des betrachteten Zeitraums gewählt.

Jede Variante macht einen Sinn, aber in der Interpretation ist dem unterschiedlichen Design Rechnung zu tragen. Die Varianten (b) bis (d) sind am ehesten in der Lage, Veränderungen zu Tage treten zu lassen, denn die Profile im Zeitverlauf werden mit jeweils einem konstanten Profil verglichen. Dabei wird im Fall (b) sozusagen „ex post" der Weg des Profils in der Entwicklung hin zur aktuellen Situation analysiert, während der Fall (c) eine „a priori" Betrachtung ist bezüglich der Frage, wie sich das Profil der Einrichtung seit dem Beginn des Untersuchungszeitraums weiterentwickelt hat. Die Variante (d) bezieht sich auf ein geeignet gewähltes Basisjahr innerhalb des Zeitraums. Beispielsweise könnte für einen Zeitraum, der die Integration der VHS in den neuen Ländern in die VHS-Statistik 1991 umfasst, das Jahr 1992 ausgewählt werden, da hier die Erfassungsquote wie schon in den alten Ländern nahezu 100% betrug. In der Variante (a) wird eine (im Fall von umfassenden Gebieten globale) Veränderung des Vergleichsprofils mit eingerechnet und nur die relative (im Fall von umfassenden Gebieten lokale) Änderung berücksichtigt. Sollten da Typveränderungen auftreten, sind sie daher besonders einrichtungsspezifisch zu interpretieren.

Für eine VHS werden die Ergebnisse der Typbildung für das Veranstaltungsprofil im Vergleich zum Landesprofil im Zeitraum 1987 bis 2003 bei gleicher Toleranzschwelle von 0,25 in den ersten drei Varianten präsentiert.

Beispiel 9 Typisierung des Veranstaltungsprofils einer Einrichtung in der Zeitreihe nach Variante (3) einer Volkshochschulen in der Zeitreihe zwischen 1987 und 2003
Die Tabelle 15, Seite 26, ist gut geeignet, um in einem Langzeitvergleich schnell Phasen in der Entwicklung zu erkennen. Bis auf die Jahre 1992 bis 1994 ist das Profil der Einrichtung im Vergleich zum „Bundesprofil" einrichtungsspezifisch. Vor diesem Zeitraum liegt der Hauptunterschied in der Stärke des Sprachenbereichs („D"). Danach wird der Programmbereich Grundbildung-Schulabschlüsse ausgebaut, nur unterbrochen von einem vergleichsweise starken Ausbau im Programmbereich Arbeit-Beruf zwischen 1997 und 1999. Meist geht mit dem Ausbau eines Bereichs die relative Rücknahme eines anderen Bereichs einher. In dieser Einrichtung ist es ab 1995 kontinuierlich der Bereich Gesundheit.

Gerade diese Eigenschaft tritt in der Variante (b) „ex ante" im Vergleich zum Bundesprofil 1987 so gut wie nicht, in der Variante (c) „ex post" im Vergleich zum Bundesprofil 2003 noch stärker zutage. Hier macht sich bemerkbar, dass Gesundheitsbildung sich im Bundesprofil ab Mitte der 1990er Jahre auffällig stark entwickelt hat (vgl. auch Tabelle 16, Seite 27). Offensichtlich macht der Vergleich des lokalen Profils mit dem jeweils aktuellen Bundesprofil nach der Variante (a) unabhängiger von eigenen Entwicklungen dieses Vergleichsprofils. Das ist die Stärke dieser Variante.

Tabelle 15 Veranstaltungsprofile einer Einrichtung nach Variante (1) wischen 1987 und 2003

VHS XYZ	Variante (a) Basis aktuelles Jahr	Variante (b) "ex ante" Basis 1987	Variante (c) "ex post" Basis 2003
1987	Dc	Dc	cD
1988	D	D	Dc
1989	De	De	eD
1990	De	De	Dec
1991	D	D	c
1992		D	c
1993		D	c
1994		D	c
1995	Fc		Fc
1996	Fc		c
1997	E	Ef	E
1998	Ec	Ef	Ec
1999	E	E	Ec
2000	c	E	Ec
2001	Fc	Fb	Fc
2002	Fc	b	Fc
2003	Fc	Fb	Fc

An dieser Stelle gibt es einen Beleg für den Ertrag der Methode im Dialog mit Weiterbildungseinrichtungen, der gleichzeitig eine Validierungsperspektive bietet. Im Gespräch mit der stellvertretenden Leiterin der betreffenden Einrichtung gelang es in kürzester Zeit (8 Minuten) den „Befund“ zu der Profilentwicklung zu formulieren ohne eine einzige Zahl zu verwenden. Sie hat ihn als in allen Punkten zutreffend bezeichnet, so dass sofort eine gemeinsame (empirische) Basis erzielt war, um über die Art der Faktoren zu diskutieren, die die Entwicklung bestimmt haben könnten. Solche Methoden müssten der Beratung in Organisationsentwicklungsprozessen hochwillkommen sein.

Eine weitere Variante der Zeitreihenbildung, die allerdings besonders niedrige Toleranzschwellen erfordert, wird am Beispiel der Entwicklung des Bundesprofils demonstriert. Es handelt sich um einen Vergleich des aktuellen Profils eines Gebiets zum *eigenen* Profil in einem Basisjahr („ex ante“ oder „ex post“). Diese Variante lässt sich auch für einzelne Einrichtungen anwenden.

Tabelle 16 Die Entwicklung des „Bundesveranstaltungsprofils“ im Vergleich zum Basisjahr 1987

Jahr	Typ Toleranzschwelle 0,05	$Kontig_{korr}$	$Koeff_A$	$Koeff_B$	$Koeff_C$	$Koeff_D$	$Koeff_E$	$Koeff_F$
1987		0,000	0,000	0,000	0,000	0,000	0,000	0,000
1988		0,042	-0,016	-0,016	0,012	0,028	-0,014	-0,009
1989	**D**	0,113	-0,033	-0,043	0,011	0,082	-0,036	-0,022
1990	**Db**	0,163	-0,035	-0,071	0,011	0,119	-0,049	-0,032
1991	**Db**	0,140	-0,033	-0,079	0,009	0,092	-0,022	-0,007
1992	**Db**	0,143	-0,025	-0,081	0,025	0,086	-0,004	-0,043
1993	**Dbf**	0,175	-0,034	-0,087	0,044	0,106	-0,026	-0,054
1994	**CDbf**	0,185	-0,033	-0,088	0,071	0,099	-0,025	-0,070
1995	**CDbf**	0,209	-0,045	-0,106	0,085	0,089	0,018	-0,090
1996	**CDbf**	0,244	-0,043	-0,115	0,089	0,099	0,039	-0,128
1997	**CDbf**	0,262	-0,043	-0,129	0,087	0,122	0,030	-0,134
1998	**CDEabf**	0,280	-0,064	-0,145	0,083	0,124	0,081	-0,122
1999	**CDEabf**	0,293	-0,070	-0,164	0,075	0,119	0,088	-0,122
2000	**CDEabf**	0,315	-0,071	-0,181	0,081	0,113	0,110	-0,122
2001	**CDEabf**	0,314	-0,063	-0,187	0,086	0,120	0,096	-0,120
2002	**CDEabf**	0,324	-0,074	-0,186	0,093	0,138	0,076	-0,125
2003	**CDabf**	0,320	-0,081	-0,180	0,114	0,147	0,034	-0,115

Für das Bundesprofil in Tabelle 16 ist bereits am Veranstaltungsprofiltyp (Variante (2) zu erkennen, wann Entwicklungen einsetzen und in den meisten Fällen bis heute anhalten. Ab 1989 der verstärkte Ausbau des Sprachenbereich („D“), ab 1994 der Ausbau des Programmbereichs Gesundheit („C“), ab 1998 bis 2002 der Ausbau des Programmbereichs Arbeit-Beruf. Damit gehen auch im Vergleich zum Profil 1987 Rücknahmen einher: seit 1990 beim Programmbereich Kultur-Gestalten, ab 1993 bei Schulabschlüssen, ab 1998 auch im Programmbereich Politik-Gesellschaft-Umwelt.

In der Tabelle sind Spalten für den Kontingenzkoeffizienten als Maß für die Ähnlichkeit mit dem Vergleichsprofil und seine programmbereichsspezifischen Komponenten ergänzt um zu zeigen, worin der Informationsmehrwert im Vergleich zu der Typenbildung liegt. Die Spalte für den Kontingenzkoeffizienten zeigt, dass die Entfernung vom Profil 1987 kontinuierlich wächst. Aber auch schon der Typ zeigt, dass immer mehr Programmbereiche an dem Kontrast zu 1987 beteiligt sind. An den programmbereichspezifischen Kontingenzkomponenten ist für jeden einzelnen Programmbereich abzulesen, dass sich die begonnenen Entwicklungen mit einer Ausnahme fortsetzen. Im Fall des Programmbereichs Arbeit-Beruf ist an den genauen Zahlen schon 2001 eine Trendwende zu erkennen, die bei der Bewertung des Typs erst 2003 zu identifizieren ist.

Besonders ergiebig könnte bei der Betrachtung einer ganzen Population von Weiterbildungseinrichtungen (z. B. alle Einrichtungen in der Bundesrepublik) sein, auszunutzen, wenn Daten als Panel vorliegen, wie das bei der Volkshochschul-Statistik der Fall ist. „Panel-Mortalität“ spielt hier nur insofern eine Rolle, als dass die wenigen Neugründungen (ohne Vorgängereinrichtungen), Zusammenlegungen oder Trennungen über den betrachteten Zeitraum berücksichtigt werden müssen. Solche Veränderungen sind seit 1987 – ab da liegt eine elektronische Datenbasis für Panelbetrachtungen der Volkshochschul-Statistik vor – selten. Zusammenlegungen wie für die neuen Länder in den 1990er Jahren müssten nicht zu einer Eliminierung aus dem Panel führen, da durch Aggregation über die Vorgängereinrichtungen das Panel vervollständigt werden könnte. Im Rahmen einer Panel-Betrachtung können so spezielle Fragen für Profile und ihre Typisierung Fragen beantworten werden, wie

- Welche Übergänge bei Profiltypen sind wie häufig?
- Häufen sich Wechsel im Profiltyp zu bestimmten Zeitpunkten?
- Welche Profiltypen sind in der Population kurzlebig oder langlebig?

Beispiel 10 Typwechsel beim Veranstaltungsprofil von hessischen Volkshochschulen zwischen 1987 und 2003

Das Basismaterial bilden die Veranstaltungsprofildaten der 33 hessischen Volkshochschulen. In dem Zeitraum zwischen 1987 und 2003 gibt es in Hessen keine institutionellen Veränderungen wie Zusammenlegungen von Einrichtungen oder Hinzutreten von neuen Einrichtungen etwa durch Auflösung einer Einrichtung in mehrere Nachfolgeorganisationen, die die Zusammensetzung des Institutionenpanels beeinträchtigen könnten. Für die Analyse ist die Typenbildung in ihrer einfachsten Variante (1) – nur die betragsmäßig größte Kontingenzkomponente ist berücksichtigt – zugrunde gelegt. Vergleichsprofil ist das Veranstaltungsprofil aller Volkshochschulen in der Bundesrepublik im entsprechenden Berichtsjahr (vgl. Tabelle 16, Seite 27) bei einer Toleranzschwelle von 0,25.

*Bei den 33 Untersuchungseinheiten finden für die 16 Übergänge von einem Berichtsjahr zum nächsten 123 Typwechsel statt. Das sind 23,3% von den 16 * 33 = 528 möglichen Wechsel. Als Indikator bezeichnet dieser Anteil die durchschnittliche Zahl der Wechsel pro Einrichtung pro Jahr. Die Streuung um diesen Wert ist deutlich: In drei Einrichtungen findet kein Typwechsel statt, während in vier Einrichtungen in dem Zeitraum sieben Wechsel gezählt werden (7 / 16 = 43,8%).*

Eine Sonderrolle spielt das „typenlose" Veranstaltungsprofil – keine auffälligen Abweichungen vom Vergleichsprofil. Es könnte sich als zweckmäßig erweisen, Übergänge zu und von diesem Typ bei der Analyse zu vernachlässigen. Allein 31 der 123 Typwechsel wären von dieser Bauart.

Unter den Typen mit auffällig starken Programmbereichen zeigen Einrichtungen, die „Kultur-Gestalten" und „Gesundheit" betonen, eine 90%-ige Beharrungstendenz; auch für „Sprachen" sowie „Arbeit-Beruf" wiederholt sich in 80% der Fälle der Typ. Der Programmbereich Grundbildung-Schulabschlüsse scheint wegen der starken Bindung von Unterrichtsstunden für vergleichsweise wenig Teilnehmende eine Schlüsselrolle einzunehmen. Der Wechsel zu dem Einrichtungstyp, der diesen Programmbereich zurücknimmt, ist ausgehend von nahezu allen Programmbereichen besonders häufig. Umgekehrt ist die Tendenz der Übergänge von diesem Typ zurück zu einem Typ mit Verstärkung der anderen Programmbereiche vergleichsweise häufig.

Zwischen 1989 und 1991 bzw. zwischen 1998 und 2001 scheinen Typwechsel häufiger als in anderen Zeiträumen. Dies müsste mit Einrichtungen in anderen verglichen werden. Für spezifische Schlüsse scheint die Datenbasis im Beispiel nicht ausreichend.

Welche Einrichtungstypen sind bezüglich der Auffälligkeiten von Programmbereichen mit besonders hohem bzw. niedrigem Gewicht lang- oder kurzlebig? Die Programmbereiche Kultur-Gestalten, Sprachen sowie Arbeit-Beruf weisen besonders viele „Ketten" – Typ verändert sich nicht – auf, darunter teilweise mit Längen von neun oder mehr Jahren. Die Programmbereiche Politik-Gesellschaft-Umwelt und Gesundheit als Schwerpunktbereiche sind nicht prägend für langlebige Typen. Nur der Programmbereich Grundbildung-Schulabschlüsse zeigt für den Typ mit niedrigen Gewichten eine gewisse Langlebigkeit. Es fällt auf, dass unter den Einrichtungen mit einem konstanten Typ zwei durchgehend den Sprachenbereich betonen. Es handelt sich um Frankfurt und Offenbach, den beiden Gemeinden über 200 Tsd. Einwohner mit der höchsten Ausländerquote. Die dritte Einrichtung betont durchgehend den Programmbereich Arbeit-Beruf. Es handelt sich um eine Einrichtung in einem Kreis mit einer vergleichsweise hohen Arbeitslosenquote. Die vierte Einrichtung ohne Typveränderung hat einen langlebigen Schwerpunkt im Programmbereich Kultur-Gestalten.

Insgesamt zeigen sich Einrichtungen in ihrem Veranstaltungsprofil durchaus wandlungsfähig. Das zeigen die häufigen Wechsel des Typs. Gleichzeitig wird deutlich, dass sich die Nachfragestruktur nicht so schnell zu ändern scheint oder dass möglicherweise eine gewisse Beharrungstendenz im Planungsbereich die Entwicklung einer Programmstruktur prägt. Dafür sprechen die teilweise über längere Zeiträume unveränderten Einrichtungstypen des Beispiels. Jedenfalls ist die „Typgeschichte" eine jeweils einrichtungsspezifische Eigenschaft. Gleiche Verläufe sind selten. Das belegt bereits das voranstehende Beispiel.

Tabelle 17 Typwechsel beim Veranstaltungsprofil von 33 hessischen Volkshochschulen zwischen 1987 und 2003

Übergänge			Nach (1988 – 2003)													insgesamt
			""	A	B	C	D	E	F	a	b	c	d	e	f	
			101	22	84	38	87	59	25	0	1	5	3	13	90	
Von (1987 – 2002)	""	103	*70*	1	1	2	7	5	3					3	11	**33**
	A	22	1	*15*	1		1	2							2	**7**
	B	87	1	1	*78*		1		1						5	**9**
	C	37				*33*		1					1		2	**4**
	D	85	5		2		*70*					1		1	6	**15**
	E	56	2	1		1		*47*	3						2	**9**
	F	24	2					2	*18*		1			1		**6**
	a	0								*0*						**0**
	b	1	1								*0*					**1**
	c	5					1					*4*				**1**
	d	3											*2*		1	**1**
	e	11	2				1							*8*		**3**
	f	94	17	4	2	3	6	2							*60*	**34**
insgesamt			**31**	**7**	**6**	**6**	**17**	**12**	**7**	**0**	**1**	**1**	**1**	**5**	**29**	**123**

Bei den Typbezeichnungen entsprechen A/a „Politik-Gesellschaft-Umwelt“, B/b „Kultur-Gestalten“, C/c „Gesundheit“; D/d „Sprachen“, E/e „Arbeit-Beruf“, F/f „Grundbildung-Schulabschlüsse“.

8 Schlussbemerkung

Die Beschäftigung gerade mit im letzten Abschnitt diskutierten Fragestellungen wird vermutlich besonders über Langzeitanalysen deutlicher erkennen lassen, welche Faktoren in der Entwicklung markante Rollen spielen und damit Theorie bildend sind. Denn zum einen werden über Profile und ihre Typisierung hochkomplexe Sachverhalte berücksichtigt und es wird vermieden, unergiebig isolierten Einzelbefunden nachzugehen. Zum andern ermöglicht die Reduzierung der Komplexität nach den vorgeschlagenen Methoden erst den Zugang für Analysen.

Ich habe mich in diesem Beitrag darauf beschränkt, ein Ergebnis für die erste Frage am Beispiel der Volkshochschulstatistik zu geben, würde mich aber außerordentlich darüber freuen, wenn er Anregung für weitergehende wissenschaftliche Analysen gibt. Sie wären anwendungsorientiert, weil der Nutzen für die Praxis der Weiterbildung in Einrichtungen und für ihre Supportstrukturen wie Verbände unmittelbar erkennbar ist.

Stefan Hummelsheim danke ich für die kritische Durchsicht einer Entwurfsfassung dieses Beitrags.

9 Tabellen- und Abbildungsverzeichnis

Tabelle 1 Finanzierung zweier Volkshochschulen 2003 nach Finanzierungsquellen 3
Tabelle 2 Verteilung der Unterrichtsstunden für zwei Volkshochschulen 2003 auf Programmbereiche 4
Tabelle 3 Tabellarische Darstellung einer Häufigkeitsverteilung/eines Profils 6
Tabelle 4 Goodness of fit nach Lienert (1962) 8
Tabelle 5 Kritische Werte für den χ^2-Anpassungstest 8
Tabelle 6 Finanzierung zweier Volkshochschulen 2003 nach Finanzierungsquellen („Finanzierungsprofil“) 9
Tabelle 7 Verteilung der Unterrichtsstunden für Nordrhein-Westfalen 2003 auf Programmbereiche ("Programmprofile") 10
Tabelle 8 Beispielergebnisse 11

Tabelle 9 Optimierung des Toleranzintervalls für Profilanalysen 17
Tabelle 10 Übersicht über festgestellte Typen bei der Finanzierungsprofilanalyse (980 Volkshochschulen 2003) – Variante (1): Nur die größte Kontingenzkomponente ist berücksichtigt..... 19
Tabelle 11 Übersicht über festgestellte Typen bei der Programmprofilanalyse (980 Volkshochschulen 2003) – Variante 1: Nur die größte Kontingenzkomponente ist berücksichtigt..... 19
Tabelle 12 Zusammenhang zwischen Programmprofiltypen und Gemeindegrößenklasse ... 21
Tabelle 13 Zusammenhang zwischen Veranstaltungsprofil und Finanzierungsprofil 2003.... 22
Tabelle 14 Gebietsspezifische Typenbildung nach der Variante (2) 24
Tabelle 15 Veranstaltungsprofile einer Einrichtung nach Variante (1) wischen 1987 und 2003 26
Tabelle 16 Die Entwicklung des „Bundesveranstaltungsprofils“ im Vergleich zum Basisjahr 1987..... 27
Tabelle 17 Typwechsel beim Veranstaltungsprofil von 33 hessischen Volkshochschulen zwischen 1987 und 2003..... 29

Abbildung 1 Finanzierungsprofile 4
Abbildung 2 Programmprofile 5
Abbildung 3 Vergleich von Finanzierungsprofilen: Kontingenzkomponenten 13
Abbildung 4 Vergleich von Programmprofilen: Kontingenzkomponenten 14
Abbildung 5 Vergleich von Finanzierungsprofilen: Auffällige Kontingenzkomponenten..... 15
Abbildung 6 Vergleich von Programmprofilen: Auffällige Kontingenzkomponenten 16
Abbildung 7 Abhängigkeit des Anteil nicht erkennbarer Typen bei Profilanalysen von der Toleranzbreite (980 Volkshochschulen 2003) 17
Abbildung 8 Affinitäten zwischen Programmbereichen Volkshochschulen 2003 auf der Basis der beiden auffälligsten Kontingenzkomponenten..... 20

10 Literatur

- BBR Bundesamt für Bauordnung und Raumordnung (2003): INKAR Indikatoren und Karten zur Raumentwicklung. Ausgabe 2003. CD-ROM zu Berichte, Band 17.
- Bortz, Jürgen (1989): Statistik für Sozialwissenschaftler. Dritte, neu bearbeitete Auflage. Springer Verlag, Berlin u. a.
- Clauß, Günter/Finze, Falk-Rüdiger/Partsch, Lothar (2002): Statistik für Soziologen, Pädagogen, Psychologen und Mediziner. Grundlagen. Verlag Harri Deutsch, Frankfurt am Main
- Pehl, Klaus (1987): VHS-Profile im Spiegel der Statistik. Arbeitspapier 110 der Pädagogischen Arbeitsstelle des Deutschen Volkshochschul-Verbandes. Frankfurt am Main
- Pehl, Klaus (2001): Informationen zu Regionen als raumstrukturellem Begriff. Online im Internet: URL http://www.die-bonn.de/service/statistik/dokumente/2001-10-16%20Raumstruktureller%20Begriff.pdf
- Pfanzagl, Johann (1966): Allgemeine Methodenlehre der Statistik II. Höhere Methoden unter besonderer Berücksichtigung der Anwendungen in Naturwissenschaften, Medizin und Technik. Walter de Gruyter & Co, Berlin
- Sachs, Lothar (2003): Angewandte Statistik. Anwendung statistischer Methoden. 11., überarbeitete und aktualisierte Auflage. Springer Verlag Berlin u. a.
- Upton, Graham/Cook, Ian (2004): A Dictionary of Statistics. OXFORD University Press Oxford New York