

Autor: Rainer Gallus

Methodische Bewertung aktueller Städterankings: Kriterien bei der Erstellung von Ranglisten

Ranglisten sind beliebt

Einleitung

Die Beliebtheit von Städte- und Regionalvergleichen lässt sich angesichts der Regelmäßigkeit von Beiträgen in der Tages- und Fachpresse kaum leugnen. Mag ein großer Teil lediglich der Befriedigung der Neugier dienen besteht doch die Gefahr, dass sich Urteile über Regionen verfestigen und Standortentscheidungen beeinflusst werden. So ist es auch erklärtes Ziel der Vergleiche, die Transparenz von Standortbedingungen zu steigern und damit wirtschaftlichen und privaten Entscheidungsträgern eine Hilfestellung bei der Wahl von Standort bzw. Arbeits- und Wohnstätte zu bieten. Gleichzeitig bezwecken die Autoren durch die Belebung des interkommunalen Wettbewerbs das Forcieren von Qualitätsverbesserungen, in dem interkommunale Vergleiche den Städten zeigen sollen, ob die eigene Entwicklung derjenigen in anderen Städten entspricht oder ob (und gegebenenfalls in welcher Weise) sie davon abweicht. In dieser Funktion als aussagekräftiges Entscheidungsinstrument müssen Rankings die Entscheidungsstrukturen des Nutzers bestmöglich nachbilden. Daher muss sich der Leser intensiv mit den methodischen Schritten einer Vergleichsstudie auseinandersetzen, um deren Aussagekraft beurteilen zu können, und sich schließlich mit den in diesem Entstehungsprozess getroffenen Entscheidungen identifizieren können. Im Folgenden sollen daher die wichtigsten methodischen Grundlagen aktueller Rankings beleuchtet und diskutiert werden. Hierfür wird eine kurze Klärung zum Ranking-Begriff vorangestellt, bevor zunächst auf die Ziele und Indikatoren, im Schwerpunkt dann auf die Zusammenfassung von Daten eingegangen wird.

Abgrenzung des Rankingbegriffs

Gegenstand von Städtevergleichen ist die Reduktion einer Fülle von Daten und Indikatoren auf (eine bzw. wenige) eindimensionale Kennziffern, um im Ergebnis eine Rangfolge präsentieren zu können. Dabei werden in den Studien die Begriffe „Ranking“, „Rating“ und „Benchmarking“ nicht immer klar abgegrenzt oder synonym verwendet. Ein Blick auf die methodischen Grundlagen dieser Bewertungsinstrumente soll deren Grundverständnis erleichtern.

*Ranking: Zusammenfassung
mehrerer Eigenschaften*

Ein **Ranking** vergleicht Städte, in dem diese direkt hinsichtlich ihrer Leistungsausprägung in eine Reihenfolge gebracht werden. Mehrere Eigenschaften einer Stadt werden in Gesamtwerten zusammengefasst. Im Ergebnis soll jeder Stadt genau ein Rangplatz zugewiesen werden. Ein **Rating** weist einer Stadt keinen Rangplatz sondern eine Wertung zu. Die Beurteilung erfolgt mit Hilfe einer vordefinierten Messskala, nach der die Städte definierten Leistungsklassen zugeteilt werden. Verschiedene Städte können so in ihrer Leistung gleich eingestuft werden. Die Städte werden also nicht direkt verglichen sondern erst durch das Ratingverfahren vergleichbar gemacht. Bei kommunalen Vergleichen im Sinne des **Benchmarkings** steht der Lerneffekt im Vordergrund. In erster Linie entspringt das Benchmarking dem betriebswirtschaftlichen Marketingbereich und hat das Ziel, aus dem Vergleich mit den Besten der Branche Anhaltspunkte für die Verbesserung der eigenen unternehmerischen Leistung zu gewinnen. Im Vergleich zum einfachen Ranking wird der Vergleichsmaßstab also bereits im Vorhinein durch den stärksten Konkurrenten (bzgl. Umsatz, Marktanteil o.ä.) festgelegt.

Zudem geht das Benchmarking weit über den im Grunde deskriptiven Charakter des Rankings hinaus, in dem der Leistungsvergleich in der Art und Weise durchgeführt werden muss, dass ein Lernen vom Besten durch die Ermittlung von „Good-Practice“ ermöglicht wird. Ranking und Benchmarking setzen also beide an einem direkten Vergleich der Städte an, in dem die Städte selbst die Maßstäbe setzen, während das Rating seine Meßlatte unabhängig setzt.

Wie sind nun aktuelle Städtevergleiche in diese Begrifflichkeiten einzuordnen? Tatsächlich stellen sie einen Mix aus diesen drei Grundformen dar: Mit Hilfe von Ranking- und Ratingverfahren verdichten die Vergleichsstudien ihre Datenbasis, gleichzeitig werden die Ergebnisse als Rangfolge präsentiert. Daneben erheben die Studien den Anspruch des Benchmarkings, „Good-Practice“ zu präsentieren und Städte aufzuzeigen, die als Vorbild dienen. Da in der Praxis die Ergebnisrangfolge im Vordergrund steht, wird im Weiteren der Rankingbegriff bevorzugt. Die Autoren von Städterankings sind in der Mehrheit private Forschungsinstitute, die Vergleichsanalysen aus PR-Gründen veröffentlichen. Teilweise werden hierbei umfassende Studien veröffentlicht, die in den Medien aufgegriffen und/oder von diesen gesponsert werden (z. B. von IW Consult (Köln)¹⁾ und Prognos (Köln)). Oft liegen den Rangfolgen aber auch keine separaten Studien zugrunde, wie beispielsweise beim alle zwei Jahre im Wirtschaftsmagazin „Capital“ veröffentlichten Ranking von Feri (Bad Homburg) mit Prognosen zur wirtschaftlichen Entwicklung²⁾. Vielmehr sollen diese Ergebnisse für Analysen werben, die Feri im Rahmen von Auftragsarbeiten oder Abonnements für Immobilieninvestoren anfertigt. Zuletzt erstellen auch verschiedene öffentliche Einrichtungen wie Industrie- und Handelskammern³⁾, sowie Wissenschaftler universitärer Forschungsinstitute Vergleichsstudien, die zumindest im weiteren Sinne auf Rankingmethoden zurückgreifen.

Auf folgende aktuelle Studien wird im Weiteren beispielhaft besonders eingegangen:

- ◆ *Berlin Institut für Weltbevölkerung und globale Entwicklung (2004): Deutschland 2020. Die demographische Zukunft der Nation.* Thematisch umfassender Standortvergleich mit Schwerpunkt „Demographie“ (22 Indikatoren).
- ◆ *IW Consult (2004): Großstadtvergleich. Deutsche Großstädte im Vergleich.* Thematisch umfassender Standortvergleich mit Schwerpunkt „Wirtschaft“ (64 Indikatoren).
- ◆ *Institut für Demoskopie Allensbach (2004): Unternehmerfreundliche Großstadt. Bericht über eine Telefonumfrage unter mittelständischen Unternehmern in den 25 größten Städten Deutschlands.* Befragung regionaler Unternehmer zum Service von sechs wirtschaftsnahen Einrichtungen (IHK, Wirtschaftsförderung, Arbeitsamt, etc.).

Alle Studien stehen im Internet auf den Seiten der jeweiligen Herausgeber als Download zur Verfügung.

Zur Erstellung von Ranglisten

Folgende Themen sind eine Auswahl von Entscheidungen, die bei der Erstellung eines Rankings getroffen werden müssen, und die eine entscheidende Bedeutung auf das Ergebnis des Rankings haben. Deshalb müssen diese sorgfältig durchdacht, vor allem aber auch ausführlich beschrieben werden, um dem Leser die tatsächliche Interpretation der Ergebnisse zu ermöglichen.

1) Der „Großstadtvergleich“ von IW Consult erfolgte in Zusammenarbeit mit der Wirtschaftswoche.

2) Zuletzt wurde ein Ranking im Heft 2/2005 veröffentlicht.

3) So z. B. der Deutsche Industrie- und Handelskammertag (2002): „Standorte in Deutschland“.

1.1 Rankingziel, Zielgruppe, Zielobjekte

Während das **Ziel** eines Rankings in der Regel erkennbar ist, bleibt die konkrete **Zielgruppe** in der Studie oft unspezifiziert. Dabei werden Unternehmer, kommunale Entscheider oder Privathaushalte denkbar verschiedene Maßstäbe an Standorte stellen. Beispielsweise ist das Feri-Ranking an Immobilieninvestoren gerichtet, die vermutlich ein großes Interesse an hohen Mietpreisen haben. Unternehmensgründer und eine Familie mit drei Kindern werden dieses Interesse kaum teilen, so dass die Bewertung der Mietpreise konzeptionell anders ausfallen müsste. Für die Ersteller besteht der Konflikt damit im Grunde darin, einerseits eine möglichst breite Interessengruppe anzusprechen, andererseits aber deren spezifische Entscheidungsstrukturen möglichst genau zu repräsentieren. Die in den Studien behandelten Städte stellen die **Zielobjekte** der Untersuchung dar. Meist werden „die x größten“ oder „alle Großstädte“ (über 100 000 Einwohner) einbezogen. Neben einer inhaltlichen Abgrenzung muss auch auf die regionale Abgrenzung eingegangen werden, die durchaus von Indikator zu Indikator unterschiedlich definiert werden kann, um die Wirkung des Indikators bestmöglich zu erfassen.

Reduktion der Komplexität

1.2 Indikatoren – die Basis des Rankings

Zur Messung von Leistungen steht eine Vielzahl von ökonomischen, gesellschaftlichen und ökologischen Indikatoren zur Verfügung. Aus diesen sind möglichst aussagekräftige Kennziffern auszuwählen, um durch die Reduktion der Komplexität möglichst wenig Information zu verlieren. Neben inhaltlichen Kriterien sollte auch auf ein Mix verschiedener Indikatorenarten geachtet werden. So lassen sich gewonnene Ergebnisse durch Wanderungsbewegungen oder Tourismuszahlen (als indirekte Attraktivitätskennzahlen) überprüfen. Mit Hilfe von subjektiven Indikatoren (Befragungsdaten) kann direkt nach der Zufriedenheit von Personen gefragt werden, während objektive Indikatoren dem Leser die Möglichkeit überlassen, sich auf der Basis der Daten eine subjektive Meinung zu bilden. Letztlich entscheidend ist aber die Validität. „Misst der Indikator tatsächlich auch das, was er messen möchte?“ Besonders ist daher darauf zu achten, dass Indikatoren eindeutig interpretierbar sind.

Transformation und Gewichtung

1.3 Die Zusammenfassung der Daten

Um allen Städten einen vermeintlich repräsentativen Gesamtwert zuordnen zu können, müssen Indikatoren und Dimensionen zusammengefasst, also aggregiert werden. Dazu werden alle Daten auf eine einheitliche Skala gebracht (Transformation) und anschließend über die Bedeutung der Indikatoren und Variablen (Gewichtung) entschieden.

1.3.1 Transformation

Die Transformation der Indikatorenwerte in Leistungsbewertungen bildet den Kern von Städterangfolgen. In der Praxis sind im Grunde drei verschiedene Bewertungstypen erkennbar: Rangskalen, normative Skalen und Benchmarking-Skalen.

Rangskalen

Rangskalen stellen die einfachste Art der Transformation dar: Geordnet nach ihren Indikatorenwerten werden die Städte in eine Reihenfolge gebracht und entsprechend ihrem Rang bepunktet (Rangwerte) – eine Systematik, wie sie eigentlich für ordinal-skalierte Indikatoren vorgesehen ist. Werden Indikatorenwerte mit höherem Skalenniveau durch eine solche Rangskala transformiert, geht in der Folge Information verloren: Die Rangwerte lassen keinerlei Rückschlüsse mehr auf vorhandene Leistungsabstände zu, d.h., die Struktur der Leistungsdaten geht verloren. Der absolute Wert und die Wertabstände treten vollkommen in den Hintergrund. So lassen große Differenzen in Rangfolgen nicht unbedingt auch auf große Unterschiede in den Indikatorenwerten schließen.

In der Praxis findet diese Transformationsform beispielsweise Anwendung in den Studien des Forschungsinstituts empirica Delasasse (Köln). Ein Beispiel wird später noch mögliche Auswirkungen auf ein Gesamtranking verdeutlichen.

Normative Skalen

Bei Anwendung von normativen Skalen werden im Rahmen der Datenzusammenfassung normative Zielwerte definiert, an deren Erreichen Leistungen gemessen werden. Anhand einer so festgelegten Notenskala ordnet beispielsweise die Studie des Berlin Instituts für Weltbevölkerung und globale Entwicklung den Städten einen Notenwert zu (Tabelle 1). Dies kann dazu führen, dass Höchstnoten gar nicht erreicht werden, wie beim Indikator „Kinderzahl pro Frau“: Die für die Bestnote 1 benötigte Kinderzahl wird in keinem Kreis erreicht. Der erzielte Höchstwert liegt bei 1,92 (Note 2). Das herangezogene Maß ist also nicht eine empirische sondern eine normative Vorgabe ⁴⁾.

Tabelle 1

Normative Skala

„Kinderzahl pro Frau“		„Wanderung“ (Saldo je 1 000 Ew.)	
Indikator	Note	Indikator	Note
2 und mehr	1	5 und mehr	1
1,70 bis 1,99	2	0 bis 4,99	2
1,50 bis 1,69	3	-3 bis -0,01	3
1,40 bis 1,49	4	-6 bis -3,01	4
1,30 bis 1,39	5	-9 bis -6,01	5
weniger als 1,30	6	weniger als -9	6
Spanne der empirischen Werte: 0,88 bis 1,92		Spanne der empirischen Werte: -35,04 bis 29,78	

Quelle: Berlin Institut für Weltbevölkerung und globale Entwicklung (2004).

Weiterhin fällt bei der hier vorliegenden Transformationskala auf, dass die Wertebereiche der Noten unterschiedlich groß sind und die empirischen Werte mitunter weit außerhalb der Grenzen liegen: Im Indikator „Wanderung“ wird bereits ab einem Saldo von fünf die Bestnote 1 vergeben. Ein Kreis erreicht jedoch einen Wert von knapp 30. Diese Leistung kommt in der zuzuordnenden Note 1 kaum zur Geltung. Die tatsächlich breite Wertespanne wird damit reduziert, Städte mit stark abweichender Leistung erhalten in den hohen und niedrigen Wertebereichen keine besseren bzw. schlechteren Noten mehr. Die Transformation ist damit nicht linear und bildet die Struktur der Daten nicht hinreichend ab. Die Notenintervalle bilden die Wertebereiche nicht mehr strukturtreu ab, wodurch die aus diesen Noten gebildeten Mittelwerte nicht mehr repräsentativ für die ursprünglichen Leistungswerte sind.

Eine generelle Problematik bei normativen Skalen liegt schließlich in der Festlegung und Begründung der Zielwerte, also in der Frage nach „dem Maß der Dinge“. In der Praxis erscheint eine Anwendung normativer Wertungsskalen daher besonders für Indikatoren mit gesetzlich vorgeschriebenen Grenzwerten geeignet (Umwelt- und Gesundheitsbelastungen).

Benchmarking-Skalen

Benchmarking-Skalen umgehen die Problematik normativer Zielwerte, in dem der tatsächlich empirisch erzielte Maximalwert als Zielwert herangezogen wird, was dem oben dargestellten Grundgedanken des Benchmarkings entspricht: der Vergleich mit den Besten. In der Praxis wird dafür die Wertespanne zwischen höchstem und geringstem Indikatorenwert in eine Punkteskala umgesetzt.

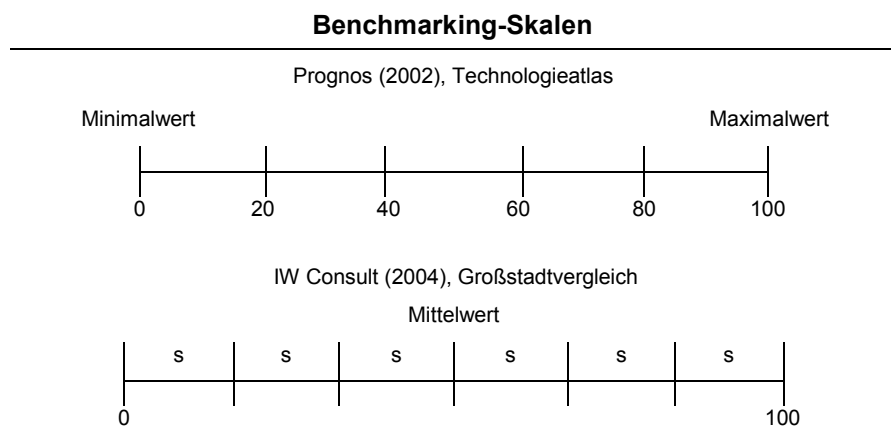
4) In diesem Falle hatten die Autoren vermutlich die Kinderzahl 2,1 im Visier, bei der eine Bevölkerung stabil bleibt.

Im einfachsten Fall erfolgt eine simple Unterteilung in 10 Bereiche und die Vergabe der Punkte von 0 – 10. Nach diesem Prinzip wurden die Daten beispielsweise im „Technologieatlas“ von Prognos für eine Aggregation transformiert: Der Minimalwert wird mit 0 Punkten, der Maximalwert mit 100 Punkten bewertet, dazwischen wird linear interpoliert (siehe Abbildung 1). So entsteht ein Abstandsmaß, das die Position der Stadt auf der Spanne zwischen Maximal- und Minimalwert angibt. Erhält eine Stadt z. B. beim BIP je Einwohner eine Bewertung von 70, erreicht diese Stadt 70 % des BIP-Vorsprungs, den der Erstplatzierte vor dem Minimalwert hat. Dabei ist zu beachten, dass durch diese Vorgehensweise bereits eine einzige extrem überdurchschnittliche Leistung alle anderen vergleichsweise schlecht erscheinen lässt⁵⁾. Weiterhin ist für eine Interpretation der Ergebnisse zu berücksichtigen, dass die Normierung auf eine einheitliche Skala von 10 bzw. 100 Punkten unterschiedliche Wertespannweiten ausgleicht, d.h., stark streuende Wertebereiche werden gestaucht, sehr geringe Wertespannen hingegen durch die Indexierung gedehnt und überzeichnet.

IW Consult mit eleganter Lösung

Eine sehr elegante Lösung, die z. B. auch die Problematik von Ausreißern abschwächt, bietet das Transformationsschema im sog. „Großstadtvergleich“ des Wirtschaftsforschungsinstituts IW Consult. In der Studie wird die Bewertungsskala nicht direkt durch Minimal- und Maximalwert bestimmt, sondern mittels der Abweichung vom Mittelwert (siehe Abbildung 1): Die (theoretische) Höchstpunktzahl erreichen Städte mit einem Indikatorenwert, der um mehr als das dreifache der Standardabweichung über dem Mittelwert liegt. Entsprechend werden Städte, die um mehr als das dreifache der Standardabweichung hinter dem Durchschnitt herinken mit 0 Punkten bewertet. Dazwischen wird linear interpoliert. So bilden die tatsächlich erzielten empirischen Werte über den Mittelwert die Basis der Skala, und auch die Verteilung der Werte wird durch die Standardabweichung berücksichtigt.

Abbildung 1



Quelle: Eigene Darstellung.

Für eine Interpretation ist bei Benchmarking-Skalen im Gegensatz zu normativen Skalen zu beachten, dass hohe Punktzahlen nicht unbedingt mit absolut hohen Indikatorenwerten korrespondieren. Im Extremfall erhält eine Stadt sehr gute Wertungen, wenn die absolut betrachteten schlechten Ergebnisse von allen anderen noch untertroffen werden, so dass im Ergebnis bereits ein schlechter Wert der beste ist.

Der entscheidende Vorteil von normativen Skalen und Benchmarking-Skalen liegt in ihrer Möglichkeit, die Datenstruktur durch die Bewertungen getreu wiederzugeben. Wie oben bereits angesprochen, bilden Rankingskalen die Leistungsabstände nicht mehr repräsentativ ab, wohingegen normative Skalen und Benchmarking-Skalen, wenn entsprechend linear transformiert, strukturtreu sind.

5) So misst die Studie der Handwerkskammer Hamburg (2000), *Metropolregionen im Wettbewerb*, die Entscheidungskonzentration von Städten im Bankensektor durch die Anzahl der ansässigen Banken. Es überrascht wenig, dass Frankfurt/Main hier eine klare Vorrangstellung einnimmt. In Folge dessen erhalten andere Metropolen nur schlechte Wertungen.

Die Abstände der Punktwerte repräsentieren dann also in gewissem Maße die tatsächlichen Leistungsdifferenzen. Ein Beispiel soll mögliche Auswirkungen auf das Gesamtranking verdeutlichen. Angenommen werden folgende Werte der Städte Heidelberg und Kassel innerhalb der Dimension „Wachstumsdynamik“, die zum einen durch Rangwerte, zum anderen durch die Vergabe von Punkten transformiert und anschließend (additiv) aggregiert werden:

Tabelle 2

Aggregation von Rang- und Punktwerten

Stadt	BIP-Wachstum			Prognose des BIP-Wachstums			aggregiert	
	Indikatorwert	Rangwert	Punkteskala 1)	Indikatorwert	Rangwert	Punkteskala 1)	Rangsumme	Punkte-summe
Heidelberg	3,0 %	4	3	1,0 %	33	1	37	4
Kassel	1,0 %	69	1	4,0 %	15	4	84	5

Quelle: Eigene Darstellung.

1) Formel: (Indikatorwert/Maximum) · 10 [Angenommen wird jeweils ein Maximalwert von 10 %.]

Das Beispiel zeigt, dass je nach Transformationsverfahren entweder Heidelberg oder Kassel besser abschneiden. Werden die Rangwerte zusammengefasst, kommt Heidelberg auf 37, Kassel auf 84 Rangpunkte, womit Heidelberg klar vorne liegt. Bei Aggregation der Punkteskala kann Kassel 5 Punkte erzielen, während Heidelberg auf insgesamt nur 4 Punkte kommt, also hinter Kassel platziert wird. Die Ursache für dieses unterschiedliche Ergebnis liegt darin begründet, dass Rangwerte die Abstände der Leistungen nicht strukturgetreu abbilden. Durch die Rangstellungen der Städte erhält Kassel für die hohe Wachstumsprognose nur eine mäßig bessere Platzierung. Dagegen kann Heidelberg bei geringerem Vorsprung im Wachstum einen höheren Rangabstand erzielen. In der Folge sind die Summen und implizit auch die Mittelwerte durch diese unzureichende Abbildung der Leistungsabstände nicht mehr repräsentativ. Damit gilt es festzuhalten: Um ein Ranking auf Repräsentativität und Aussagekraft gemittelter Endwerte hin zu überprüfen, muss der Leser innerhalb des Aggregationsprozesses die Strukturtreue der transformierten Punkteskala überprüfen.

1.3.2 Gewichtung

Gewichtung, eine zentrale Aufgabe

Um eine möglicherweise unterschiedliche Bedeutung einzelner Aspekte zu berücksichtigen, werden Indikatoren und Dimensionen gewichtet. Stark gewichtete Indikatoren haben größeren Einfluss auf das Endergebnis, weshalb der Gewichtung große Beachtung geschenkt werden muss. In der Praxis werden die Punktwertungen entweder *nach* der Transformation mit unterschiedlich hohen Gewichtungsfaktoren multipliziert, oder aber, die Gewichtung wird durch unterschiedlich erzielbare Höchstpunktzahlen direkt in die Transformation *integriert*. Erhält also eine Stadt für die niedrigste Arbeitslosenquote 20 Punkte, während die Stadt mit den größten Grünflächen nur 10 Punkte erhält, so geht die Arbeitslosigkeit mit doppeltem Gewicht in die Gesamtwertung ein.

Mangels genauer Einschätzung verzichten einige Studien auf die Gewichtung von Indikatoren, so dass alle Daten mit gleichem Gewicht in das Endergebnis eingehen. Fließen Gewichtungen ein, sichern sich Studien oftmals auf Basis sogenannter Expertenmeinungen ab, die im Rahmen von Befragungen ermittelt werden. Im „Städtetest Unternehmensgründer“ (empirica Delasasse 1999) resultiert daraus folgendes Gewichtungsschema:

Wie Tabelle 3 (auf Seite 34) zeigt; wird die durchschnittliche Einschätzung der Gewichtung bei dieser Studie direkt in Gewichtungsfaktoren umgesetzt, wodurch diese mit zwei Stellen hinter dem Komma zum Einsatz kommen. Dies impliziert eine methodisch kaum erreichbare Genauigkeit und erschwert eine einfache Interpretation der Ergebnisse.

Tabelle 3

Gewichtung durch Expertenbefragung

Dimensionen	Gewichtung (in %)
Administrative Rahmenbedingungen	21,23
Regionales Umfeld	30,99
Kunden und Marktpotenziale	22,79
Gründungsklima/Erfolgreiche Gründungen	24,99

Datenquelle: Empirica Delasasse 1999; eigene Darstellung.

Auch im „Großstadtvergleich“ von IW Consult (2004) werden Expertenmeinungen herangezogen, die jedoch zur Absicherung von ökonomischen Berechnungen dienen. Das Ergebnis ist recht drastisch:

Tabelle 4

Gewichtung auf Basis ökonomischer Berechnung und Expertenmeinung

		Indikatoren (Anzahl)	Gewichtung in Prozent
Zielgrößen	Wohlstand	Arbeitseinkommen Einkommensteuer	25
	Arbeitsmarkt	Arbeitslosenquote Arbeitsplatzversorgung	25
Einflussgrößen	Standort	-37	15
	Wirtschaft	-8	15
	Struktur	-14	15
	Staat	-5	5
Summe		-64	100

Quelle: IW Consult (2004); eigene Darstellung.

Anmerkung: In der Studie wurde getrennt voneinander ein Niveau- und ein Dynamikranking erstellt. Das hier gezeigte Gewichtungsschema wird auf beide Rankings einzeln angewendet. Für das Gesamtergebnis trifft dieses Schema gleichfalls zu, jedoch geht die entsprechend doppelte Anzahl an Indikatorenwerten ein.

Wie in Tabelle 4 erkennbar entfallen auf vier von insgesamt 64 Indikatoren 50 % des Gewichts. Ein einzelner Indikator der Bereiche „Wohlstand“ und „Arbeitsmarkt“ trägt damit zu einem Achtel zum Gesamtergebnis bei. Indikatoren aus dem Bereich „Standort“⁶⁾ beeinflussen das Endergebnis dagegen im Durchschnitt mit einem Faktor von rund 0,4. Diese unterschiedliche Bedeutung von Indikatoren muss sich der Leser eines Rankings bei der Interpretation von Ergebnissen stets vor Augen halten.

Eine Variante zur Bestimmung von Gewichtungsfaktoren im Falle von Befragungsdaten zeigt die Studie des Instituts für Demoskopie Allensbach in ihrer Ermittlung der „unternehmerfreundlichsten Stadt“ (2004). Ziel der Studie ist es, die Servicefreundlichkeit von sechs wirtschaftsnahen Einrichtungen in den jeweiligen Städten zu beurteilen⁷⁾. Im Rahmen ihrer Unternehmerumfrage wurde ermittelt, wie viele Unternehmen bereits Kontakt zu den Einrichtungen aufgenommen haben, die sogenannte *Nutzungsrate*. Aus diesen Nutzungsraten werden Gewichtungsfaktoren abgeleitet, und zwar individuell für jede Stadt. Zur Ermittlung der Gesamtnote werden die einzelnen Durchschnittsnote mit den so ermittelten Gewichten multipliziert. Je stärker also eine Institution genutzt wird, desto stärker geht ihre Durchschnittsnote in den Gesamtdurchschnitt ein.

6) Zum Bereich „Standort“ zählen u.a. Indikatoren zu Arbeitskosten, Produktivität, Kostenfaktoren und Infrastrukturdaten.

7) Die Unternehmen wurden zur allgemeinen Zufriedenheit und zu einzelnen Servicequalitäten von Wirtschaftsförderungseinrichtungen, Stadtverwaltung, Industrie- und Handelskammern, Handwerkskammern/Kreishandwerkerschaften (HWK), Arbeitsämtern, „anderen städtischen Dienststellen“ und ansässigen Kreditinstituten befragt.

Tabelle 5

Gewichtung mit Hilfe von Nutzungsraten

Institution	Nutzungsrate (in Prozent)	implizite Gewichtung (in Prozent) 1)
Industrie- und Handelskammer	52	17
Handwerkskammern	25	8
andere städtische Dienststellen	63	20
Wirtschaftsförderungseinrichtungen	23	7
Arbeitsamt	72	23
Kreditinstitute	74	24

Quelle: Institut für Demoskopie Allensbach (2004); eigene Darstellung.
1) Eigene Berechnung.

Allgemein gilt: Werden keine expliziten Gewichtungen vorgenommen, wird eine Gleichgewichtung vorausgesetzt. Verschiedene methodische Vorgehensweisen führen jedoch oft zu einer impliziten Gewichtung, besonders, wenn den verschiedenen Dimensionen eine unterschiedliche Anzahl von Indikatoren zugrunde liegt, derselbe Indikator mehrfach verwendet wird oder verwendete Indikatoren eines Index hoch intern korrelieren.

1.3.3 Verknüpfung

Substituierbarkeit von Indikatoren

Voraussetzung für eine Aggregation ist die Substituierbarkeit von Indikatoren, d.h., es wird davon ausgegangen, dass Leistungen „verrechnet“ werden können. Überdurchschnittliche Indikatorenwerte, gleich welcher Art/Dimension können dann unterdurchschnittliche Werte kompensieren. Die Verknüpfung der Indikatoren durch die Aggregation erfolgt in den meisten Studien additiv. Eine interessante Alternative stellt hierbei die multiplikative Verknüpfung dar. In Form des geometrischen Mittels findet diese beispielsweise im Technologieatlas (Prognos 2002) ihre Anwendung. Im Vergleich zur additiven Verknüpfung haben bei multiplikativer Aggregationsformel Extremwerte einen größeren Einfluss auf das Gesamtergebnis und können so weniger einfach ausgeglichen werden, wie folgende Tabelle 6 verdeutlicht.

Tabelle 6

Additive und multiplikative Verknüpfung

Indikatorenwerte	additiv	multiplikativ
5;5	10	25
4;6	10	24
3;7	10	21
2;8	10	16
1;9	10	9
0;10	10	0

Quelle: Eigene Darstellung.

Vergleicht man das Ergebnis von additiver und multiplikativer Verknüpfung von Indikatorenpaaren zeigt sich, dass die Substituierbarkeit durch die Multiplikation erschwert wird. Große Diskrepanzen der Werte reduzieren die Gesamtwertung, bei vollkommener Abwesenheit eines Indikators ist der Gesamtwert gleich Null. Durch geeignete Transformation der Indikatorenwerte könnte so ein Mindestwert (beispielsweise gesetzliche Grenzwerte) einbezogen werden, bei dessen Nichterreichen der Gesamtwert Null ergibt. Die Abwesenheit eines Merkmals oder aber die Bewertung mit Null führt dann direkt zu einem letzten Platz.

Fazit

Um aussagekräftige Informationen zu liefern, müssen Städterankings innerhalb ihres Entstehungsprozesses bestimmte methodische Voraussetzungen bei der Auswahl von Städten und ihren Leistungsindikatoren sowie bei deren Zusammenführung in eine Rangfolge erfüllen.

Das Verfahren der Aggregation der Daten nimmt dabei besonderen Einfluss auf das Ergebnis eines Städtevergleichs, so dass diesem erhöhte Aufmerksamkeit geschenkt werden muss. Nur wenn der Nutzer eines Rankings seine eigenen Entscheidungskriterien innerhalb der Studie wiedererkennt, besitzen die Ergebnisse für ihn Aussagekraft.

Dabei sollte sich der Leser stets vor Augen halten, dass die Aggregation zwar die Komplexität der Darstellung reduziert, nicht aber diejenige des Sachverhalts.

Eine Interpretation von Rankingergebnissen muss in der Folge immer die dahinter verborgene Vorgehensweise berücksichtigen und die dem Ergebnis zugrunde liegende Methodik beleuchten. Somit machen erst das vollständige explizite Offenlegen der Kriterien und eine hinreichende Begründung der Bildung und Auswahl von Indikatoren und deren Zusammenfassung die Ergebnisse diskutierbar und nachvollziehbar.

Anmerkung der Redaktion

Der Autor ist Referent beim Einzelhandelsverband Nordrhein-Westfalen und dort tätig im Bereich Stadtentwicklung, Bauleitplanung und Stadtmarketing.

Dieser Artikel erschien erstmals auch in der Reihe „Stadtforschung und Statistik“, Ausgabe 2 – 2006, eine Zeitschrift des Verbandes Deutscher Städtestatistiker.