

## Die 100 größten Städte der Welt Eine Fallstudie über Ungleichheit

David Drew und Dave Steyne, Sheffield Hallam University

Urtext: The World's 100 largest Cities.- Teaching Statistics 17(1995) Vol.1, p. 17 - 21,  
übersetzt und bearbeitet von Elke Warmuth, Berlin

### Zusammenfassung

Daten für die hundert größten Städte der Welt werden vorgestellt. Sie bilden den Teil einer Fallstudie, in deren Rahmen Studenten in Explorativer Datenanalyse unterrichtet wurden, aber sie sind auch von zusätzlichem Interesse, da sie den Blick auf Armut und Unterentwicklung in der Dritten Welt und den Kontrast zum Wohlstand der Ersten Welt richten.

### Einleitung

Wir können von unseren Studenten kein automatisches Interesse für Datenanalyse oder Statistik als Thema erwarten. In England hat uns die Erfahrung gelehrt, daß nur eine kleine Minderheit von Studenten der Sozialwissenschaften ein ausgeprägtes Interesse für Statistik hat. Ihr Interesse muß gewonnen werden. Diese Studenten sind wahrscheinlich nicht stark in Mathematik und sie werden wahrscheinlich die mehr mathematischen Teile des Themas schwer finden. Ein Ausweg aus dieser Problemlage ist es, den Unterricht in Statistik und Datenanalyse eng mit der Soziologie zu verbinden, so die Rolle der Mathematik herunterzuspielen und die Studenten in die praktische Datenanalyse einzubeziehen. Dieser Zugang wird in dem ausgezeichneten Buch von MARSH (1988) angewendet, und es scheint so, als ob er bei unseren Studenten funktioniert.

Die Fallstudie ermöglicht den Studenten, ihr Wissen über MINITAB (ein statistisches Softwarepaket) in die Praxis umzusetzen und Boxplots, Tabellen, Kenngrößen und Korrelationsmatrizen zu erstellen. Sie können Erfahrungen sammeln bei der Beurteilung von Datenquel-

len, der Zuverlässigkeit und Stichhaltigkeit der Daten, bei der Herstellung von Tabellen und Graphiken, bei deren Aufnahme in Berichte und der Diskussion der Ergebnisse. Wir haben die Möglichkeit, die Bedeutung von gut aufgebauten Tabellen und die Notwendigkeit von soziologischen Interpretationen der Daten zu zeigen; bei der Datenanalyse geht es um das Aufdecken von Strukturen, aber das kann nicht ohne einen soziologischen Rahmen geschehen. Die Fallstudie kann auch als Einführung in die multivariate Datenanalyse mit Hilfe von Clusteranalyse dienen. Sie wurde erfolgreich in einem Kurs über Untersuchungsmethoden für Studenten der Sozialwissenschaften eingesetzt.

### Die Daten

Die Daten beinhalten zehn Merkmale für die 100 größten Städte der Welt aus einer Studie einer in Washington DC ansässigen Umweltorganisation (vgl. CAMP, BARBERIS und HINDS 1990, und Guardian 1990). Die Liste der Gebiete wurde durch einen Experten von Rand McNally and Company aufbereitet, die Daten für alle Merkmale stammen aus dem Jahre 1989.

Die Statistiken wurden für jede *metropolitan area* (in der Geographie feststehender Begriff, entspricht im Deutschen etwa der Bedeutung von Stadtring) mit Hilfe eines dreizehnseitigen Fragebogens im Land erhoben. Es gibt im Originalbericht eine umfangreiche Liste von Danksagungen an die Planer und Statistiker, die die Information für jede Stadt zusammengestellt haben. Das Anliegen der Übersicht war es, Daten zu sammeln "über den Lebensstandard, die relevant über nationale Grenzen hinweg sind und ohne kulturelle Vorurteile" (CAMP, BARBERIS und HINDS 1990). Die zwei städtischen Regionen, für die keine Daten verfügbar waren, sind Yangon, Myanmar (Rangoon, Burma) und Bukarest (Rumänien). Insgesamt 162 Fragebogen wurden ausgefüllt und zurückgeschickt, viele mit mehrfachen Befragten und oft aus offiziellen Quellen.

Die Daten wurden in Themengruppen eingeteilt:

#### Bevölkerung

Um möglichst zuverlässige aktuelle demographische Daten zur Verwendung bei weiteren Merkmalen zu bekommen, wurde im Fragebogen nach den offiziellen Bevölkerungszahlen und Untergliederungen nach Alter, Geschlecht und Haushalten gefragt.

Es wurde auch nach der jüngsten Schätzung der Gesamtbevölkerung, ob offiziell oder nicht, gefragt.

#### *Öffentliche Sicherheit*

Als Maß für das Niveau der persönlichen Sicherheit und für den Grad der Gewaltverbrechen in der *metropolitan area* wurde im Fragebogen nach der jährlichen Anzahl von Totschlägen gefragt.

#### *Lebensmittelkosten*

Der für Lebensmittel ausgegebene Prozentsatz des Haushaltseinkommens diente im Fragebogen als allgemeiner Indikator für Reichtum und Armut.

#### *Lebensraum und Wohnungsstandard*

Um die Lebensdichte zu bestimmen und unzureichenden Wohnungsstandard zu erkennen, sollten die Befragten die Anzahl der Wohnungen und die Anzahl der Räume je Wohnung nennen und, als qualitative Größen, Fragen zum Vorhandensein von Elektrizitäts- und Wasseranschluß beantworten.

#### *Kommunikation*

Um die städtische Infrastruktur, insbesondere die moderne Kommunikation, beurteilen zu können, sollten die Befragten die Anzahl der Telefonanschlüsse angeben. Sie sollten auch schätzen, bei wieviel Prozent der Anrufversuche eine Verbindung zustandekommt.

#### *Bildung*

Um international vergleichbare Daten über Bildungsstandards, insbesondere die Entwicklung der Oberstufe, zusammenzutragen, fragte man nach dem Prozentsatz der 14- bis 17jährigen Kinder (oder einer vergleichbaren Altersgruppe), die die Schule besuchen.

#### *Gesundheitswesen*

Die Kindersterblichkeitsrate diente als allgemeiner Indikator für das Niveau des Gesundheitswesens.

#### *Ruhe und Frieden*

Da es spezifische Daten über die Umweltlärmbelastigung nicht gibt, wurden die Befragten aufgefordert, in einer 10-Punkte-Skala eine subjektive Einschätzung über den Grad des Umgebungslärms in ihrer Stadt zu geben. Die oberen, mittleren und unteren Punkte der Skala wurden nicht-quantitativ beschrieben.

#### *Verkehrsfluß*

Um ein international vergleichbares Maß für den städtischen Verkehrsstau zu entwickeln, fragte man nach der Entfernung zum nächsten Flughafen und nach der Fahrzeit während der morgendlichen Hauptverkehrszeit mit einem privaten PKW vom Flughafen zu einem zentralen Geschäftsbezirk.

Zusätzlich zu diesen Merkmalen wurden die Städte in drei Gruppen eingeteilt: *die Erste Welt, China und die Dritte Welt*. Die Aufnahme von China als eigener Gruppe geschah einfach, weil wir fühlten, daß die chinesischen Städte interessante abweichende Charakteristiken gegenüber denen der anderen beiden Gruppen haben könnten. Aus den Themengruppen wählten wir die folgenden Merkmale aus:

1. Bevölkerung (in Millionen)
2. Öffentliche Sicherheit: Morde je 100 000 Einwohner
3. Lebensmittelkosten: Prozentsatz des Einkommens ausgegeben für Lebensmittel
4. Lebensraum: Personen pro Zimmer
5. Wohnungsstandard: Prozentsatz der Häuser mit Elektrizität/Wasser
6. Kommunikation: Telefone je 100 Einwohner
7. Bildung: Prozentsatz der Kinder in Oberschulen
8. Gesundheitswesen: Gestorbene Kinder je 1 000 Lebendgeborene
9. Ruhe und Frieden: Niveau des Umgebungslärms (1-10)
10. Verkehrsfluß: Meilen pro Stunde in der Hauptverkehrszeit
11. Zuordnung: 1=Erste Welt, 2=China, 3=Dritte Welt
12. Name der *metropolitan area*

#### **Die Analyse**

Bei Daten dieser Art sind viele verschiedene Analysen möglich. Wir hielten es für nützlich,

den Studenten sowohl offene als auch abgegrenzte Aufgabenstellungen zu geben. D. h. sie sollten z. B. Boxplots und Tabellen herstellen, aber auch die Daten in irgendeiner Weise analysieren, die ihnen angemessen erschien. Das ermöglichte ihnen, ihre eigenen Hypothesen aufzustellen und sie mit den Daten zu konfrontieren. Diese Vorgehensweise berücksichtigte die Zielsetzung, daß schwächere Studenten die Aufgabe zufriedenstellend bewältigen konnten, während stärkere Studenten weitergehen und ihre eigenen Ideen entwickeln konnten.

Auf der Sachebene können die Daten mit der jüngeren Geschichte von Kolonialismus und Imperialismus in der Dritten Welt in Verbindung gebracht werden. Als Großbritannien, Frankreich und die anderen Kolonialmächte ihren Kolonien die Unabhängigkeit zugestanden, hinterließen sie in allen Gebieten schlecht entwickelte Infrastrukturen. Vor allem das Gesundheitswesen, das Bildungswesen, das Wohnungswesen und die ökonomische Infrastruktur waren fast vollständig von ausländischen Firmen anhängig (vgl. RODNEY 1972 und MILES 1989). Die Industrie wird weiterhin von multinationalen Firmen beherrscht, und die neueren unabhängigen Länder entwickelten Vertrauen zu ihren früheren Kolonialisatoren und den Banken. Sie nehmen Kredite auf, um die Infrastruktur zu entwickeln, die sie dringend brauchen, und Schulden und weitere Armut sind die Folge davon. Die "Lebensqualitäts"merkmale in der Fallstudie spiegeln die Folgen dieses Bedingungsgefüges wider. In den späten 80er Jahren setzte sich die Schuldenspirale fort, und in vielen Fällen ist die ökonomische Situation der Länder der Dritten Welt sogar schlechter als zuvor.

Die Daten können benutzt werden, um Fragen der Zuverlässigkeit und Berechtigung von Daten zu diskutieren. Beträchtliche Sorgfalt wird in der Originaldatenquelle auf die Beschränktheit der Daten und auf die Approximationsverfahren zur Schätzung von Daten gelegt (vgl. CAMP, BARBERIS, HIND 1990). Um beispielsweise die Mordrate für Johannesburg zu erhalten, wurden die Daten für die (vorwiegend weiße) Stadt und für das (vorwiegend schwarze) Township Soweto kombiniert. Während im allgemeinen sozioökonomische Kenngrößen zu einer großen Zahl von Fehlerquellen neigen, sind die Mittel, die einige Regierungen für das Sammeln der Daten aufwenden, beachtlich, und das spiegelt sich in der Qualität der Daten wider.

Was die Stichhaltigkeit der Daten angeht, so ist es nicht möglich, aus der Originalstudie zu entnehmen, was die genaue Absicht war, mit der die Daten gesammelt wurden. Die Frage, ob ein ausgewähltes Merkmal die Lebensqualität adäquat widerspiegelt oder nicht und was das bedeutet, hätte mit den Studenten diskutiert werden können. Wir haben es vorgezogen,

das Problem auf eine andere Art anzupacken, und zwar haben wir nahegelegt, daß einige der Größen jeweils als Maße für besondere Probleme der Städte der Dritten Welt dienen können. Armut beispielsweise wird in indirekter Weise durch den Prozentsatz am Haushaltseinkommen, der für Lebensmittel ausgegeben wird, gemessen.

Wir haben univariate, bivariate und multivariate Analysen durchgeführt. Nehmen wir als Beispiel den Prozentsatz an Haushaltseinkommen, der für Lebensmittel ausgegeben wird. Er streut von 9% in Washington DC bis 80% in Ho-Chi-Minh-Stadt, Vietnam, wobei der Median bei 35% liegt (vgl. Tabelle 2 und Bild 1).

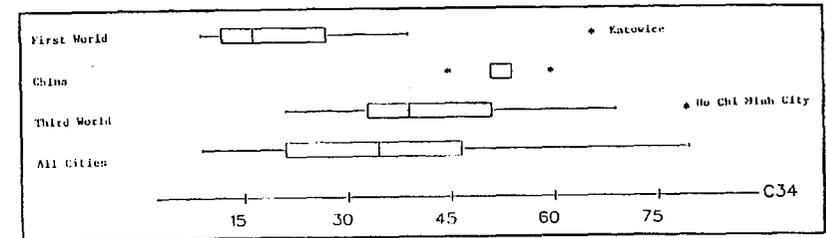


Bild 1: Boxplots der Lebensmittelkosten für die Städtegruppen

In den Ländern der Dritten Welt und in China sind die Prozentsätze wesentlich höher als in den Ländern der Ersten Welt (vgl. Tabelle 3). Die Boxplots zeigen das und lassen uns erkennen, daß Katowice in Polen ein Ausreißer der Erste-Welt-Länder ist. Wenn wir die Länder mit den höchsten Prozentsätzen von Ausgaben für Lebensmittel betrachten, dann sind Städte in Indien, Lateinamerika, Afrika, Südostasien und China gut vertreten (vgl. Tabelle 4). Die Tabelle deutet außerdem darauf hin, daß die ärmsten Städte zugleich diejenigen sind, wo weniger Kinder zur Oberschule geben, obgleich offensichtlich keine starke Korrelation zwischen diesen beiden Größen besteht.

	35		Median: 35
21		47	Halbweite: 26
9		80	Spannweite: 71

Tabelle 2: "Fünfhilfenzusammenfassung" des für Lebensmittel ausgegebenen Anteils am Einkommen (vgl. MARSH 1988)

Städtegruppe	Für Lebensmittel ausgegebener Anteil am Einkommen	Prozentsatz der Kinder in Oberschulen	Morde je 100 000 Einwohner
Erste Welt	17	90	3.1
China	54	76	2.5
Übrige Welt	40	61	6.0
Alle Städte	35	76	4.1

Tabelle 3: Mediane zu ausgewählten Größen, sortiert nach Städtegruppen

Stadt	Für Lebensmittel ausgegebener Anteil am Einkommen	Prozentsatz der Kinder in Oberschulen	Morde je 100 000 Einwohner
Lagos, Nigeria	58	31	*
Calcutta, Indien	60	49	1,1
Istanbul, Türkei	60	67	3,5
Guagzhou, China	60	56	2,5
Bangalore, Indien	62	60	2,8
Dhaka, Bangladesh	63	37	2,4
Kinshasa, Zaire	63	60	*
Katowice, Polen	67	87	2,1
Lima, Peru	70	55	*
Ho-Chi-Minh-Stadt, Vietnam	80	52	2,1
Median aller Städte	35	76	4,1

Tabelle 4: Ausgewählte Merkmale für die 10 Städte, in denen der für Lebensmittel ausgegebene Anteil des Einkommens am größten ist (in Rangfolge)

Wenn wir den Anteil der Wohnungen mit Wasser bzw. Elektrizität als Maß für die Wohnungsqualität und die Kindersterblichkeit als Maß für die Gesundheit nehmen, dann können wir einen Zusammenhang zwischen den Wohnungsbedingungen und der Gesundheit erwarten. Das ist in der Tat der Fall, aber der Zusammenhang ist nicht so einfach (vgl. Bild 2). Alle Länder der Ersten Welt zeichnen sich durch relativ niedrige Kindersterblichkeitsraten und relativ gute Wohnbedingungen aus. Für die Städte der Dritten Welt ist häufig das Gegenteil der Fall, aber es gibt Ausnahmen. Johannesburg beispielsweise hat eine schlechte Wohnqualität, aber eine relativ niedrige Kindersterblichkeitsrate.

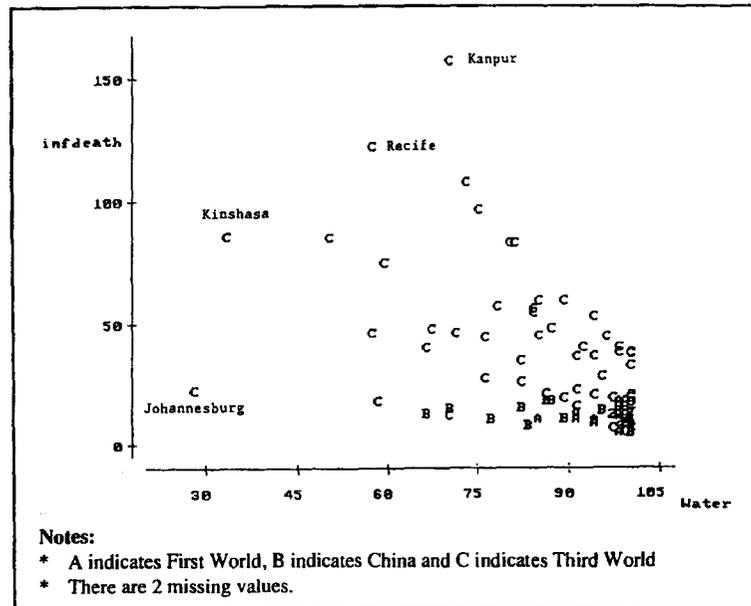


Bild 2: Punktwolke Wasseranschluß (in %) - Kindersterblichkeit je 1000

Die Diskussion der bivariaten Zusammenhänge führt in natürlicher Weise auf die Notwendigkeit multivariater Analysen. Die Beziehungen sind nicht ganz einfach, und die "Lebensqualität", wie immer sie definiert wird, hat mehr als eine Dimension. Der Lebensstandard, gemessen an Wohnung und Bildung, kann in manchen Städten der Ersten Welt hoch sein, aber dieselben Städte zeichnen sich durch hohe Kriminalität aus. Mit einer Clusteranalyse können wir dies untersuchen.

### Schlußfolgerungen

Wir glauben, daß es wichtig ist, den Studenten anspruchsvolle, aber zu bewältigende Probleme zu stellen. Auf diesem Niveau des Unterrichts gibt es eine enorme Schwankungsbreite in den Verhältnissen und Erfahrungen der Studenten. Vielen Studenten fehlt das Selbstver-

trauen. Geeignete Fallstudien sind in der Lage, den schwächeren Studenten Selbstvertrauen zu vermitteln und die fähigen oder fähigeren herauszufordern.

Die Computerentwicklung hat das potentielle Interesse an unseren Kursen erhöht, weil dadurch stumpfsinnige Rechenarbeiten entfallen. Die Studenten erleben, daß Rechnen (in diesem Fall mit MINITAB) interessant, leicht und vergnüglich ist.

Wir glauben, daß es wichtig ist, die Statistik und die sozialwissenschaftlichen Probleme als untrennbar anzusehen. Nur auf diese Weise werden wir es erleben, daß sich Studenten der Sozialwissenschaften für Datenanalyse interessieren und darauf einlassen. Der Unterricht in Datenanalyse in den Sozialwissenschaften macht nun viel mehr Spaß als zu Beginn vor fast 20 Jahren. Unsere Absolventen müssen mit neuen Fähigkeiten ausgerüstet werden, um mit dem Gebiet klarzukommen.

### Anmerkungen

Die Daten und die dazugehörige Dokumentation sind kostenlos erhältlich gegen Einsendung einer unformatierten 3,5-Zoll-Diskette an Dr. David DREW, School of Computing and Management Sciences, Sheffield Hallam University 100 Napier Street, Sheffield, S11 8 HD. Wir danken Elizabeth COATES, die uns auf die Daten aufmerksam gemacht hat, und Tina BEATTY und Praesh PATEL, die die Clusteranalyse übernommen haben.

### Literatur

- CAMP, S.; BARBERIS, M.; HINDS, I. (1990): Cities: Life in the World's 100 Largest Metropolitan Areas.- Population Crisis Committee, Washington DC, Suite 550, 1120 19th Street N.W., Washington DC, 20036-3605.
- Guardian (1990): Manchester 8th in World quality league.- Guardian Dec 29th, 1990.
- MARSH, C. (1988): Exploring Data.- Cambridge: Polity Press.
- MILES, R. (1989).- Racism.- London and New York: Routledge.
- RODNEY, W. (1972): How Europe Underdeveloped Africa.- London: Bogle-L'Ouverture Publications.