Wolfram Fischer

Streuungsfächerkarten und pseudogeografische Anordnungen

Mit Beispielen zum verfügbaren Einkommen und zu Krankenkassenprämien in der Schweiz

Schweizer Tage der öffentlichen Statistik in Vaduz

September 2012 · Letzte Änderung: 25.09.2012

Inhaltsverzeichnis

0	Abstract	2
1	Einleitung	3
2	Methoden 2.1 Streuungsfächer 2.2 Pseudogeografische Anordnung 2.3 Programmierung mit R	5
3	Daten 3.1 Statistik über Einkommen und Lebensbedingungen	7 7 8
	Resultate 4.1 Statistik über Einkommen und Lebensbedingungen	19
5	Diskussion	20
6	Anhang 6.1 Abkürzungen und Internetverweise	21 21 21

0 Abstract

Es wird eine neue Visualisierungs-Methode vorgestellt, welche über die bewährte Darstellung von Quartilen mit Hilfe von Boxplots hinausgeht. Im Weiteren wird gezeigt, wie Grafikfelder geografisch ausgerichtet kompakt positioniert werden können.

EINLEITUNG

Ein *Streuungsfächer* bildet Median, Quartile und zwei weitere wählbare Quantile auf einer Kreislinie ab. – Eine *pseudogeografische Anordnung* kann alternativ zur Abbildung geografisch gebundener Daten nach geografischen Koordinaten verwendet werden. Grafikfelder gleicher Grösse (oder Minitabellen mit Kennzahlen) werden matrixförmig so nebeneinander und übereinander angeordnet, dass die neuen Positionen in etwa den geografischen Positionen entsprechen. – Die Grafiken wurden mit der frei verfügbaren Statistiksoftware R erstellt.

METHODEN

Die neuen Methoden werden beispielhaft vorgestellt: (a) Die Streuung des verfügbaren Einkommens wird in Abhängigkeit von Alter, Geschlecht, Bildungsstand und Grossregion dargestellt. Datengrundlage ist die Statistik der Haushalte zu Einkommen und Lebensbedingungen (SILC) des Bundesamtes für Statistik (BFS) des Jahres 2010. – (b) Es werden die Streuungen der Prämienangebote der Krankenversicherer im Jahre 2010 zur obligatorischen Krankenversicherung für Erwachsene in den kantonalen Prämienregionen gezeigt. Dazu wurden öffentlich abrufbare Daten des Bundesamtes für Gesundheit (BAG) verwendet.

DATENMATERIAL

Die erarbeiteten Grafiken ermöglichen Vergleiche von Streuungen in Abhängigkeit nicht nur von einer, sondern auch von mehreren Variablen.

RESULTATE

Streuungsfächer ermöglichen eine kompaktere Darstellung als Boxplots. – Streuungsfächer können sowohl nebeneinander wie auch übereinander gestellt werden. Dies ermöglicht Vergleiche in zwei Richtungen. – Eine Hintergrundfärbung aufgrund des Medianwertes sowie die unterschiedlichen Öffnungen der Fächer lenken das Auge schnell zu aussergewöhnlichen Messwerten und Streuungen. – Veränderte Messwerte bewirken eine Drehung der Medianlinie. In einer Streuungsfächerkarte kann dies eine faszinierende Gesamtkomposition ergeben, die an Strömungsbilder von Winden erinnert.

DISKUSSION

Eine geografisch ausgerichtete Darstellung ermöglicht eine platzsparende Anordnung von Grafikfeldern unabhängig von der Grösse des dazugehörigen geografischen Territoriums. – Sie ermöglichen auf den ersten Blick eine grobe Einschätzung regionaler Unterschiede. – Sofern der Betrachter die realen geografischen Positionen kennt, findet er auch bei einer Vielzahl von Grafikfeldern relativ rasch ein ihn interessierendes Feld. – Pseudeogeografische Anordnungen sind gegenüber einer Anordnung entsprechend den realen geografischen Positionen und Flächen dann besonders attraktiv, wenn sich die territorialen Ausdehnungen sehr stark unterscheiden oder wenn es für Vergleiche wünschenswert ist, dass die grafischen Elemente gleich gross sind.

Fischer (2010) Neue Grafiken I

REFERENZ

Fischer W. Neue Grafiken zur Datenvisualisierung. Band 1: Speichengrafiken, Streuungsfächerkarten, Differenz-, Sequenz- und Wechseldiagramme. Wolfertswil (ZIM) 2010: 107 S. Internet: http://www.fischer-zim.ch/studien/Neue-Grafiken-I-1003-Info.htm.

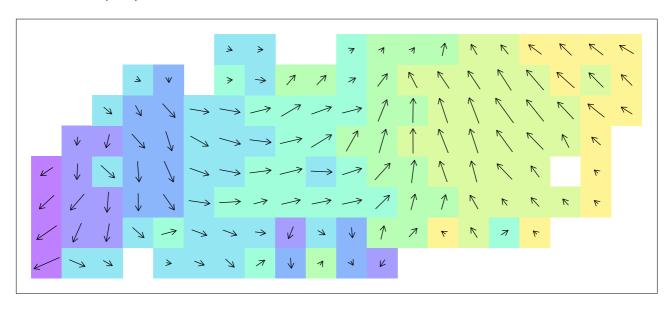
1 Einleitung

Es wird eine neue Visualisierungs-Methode vorgestellt, welche über die bewährte Darstellung von Quartilen mit Hilfe von Boxplots hinausgeht. Im Weiteren wird gezeigt, wie Grafikfelder geografisch ausgerichtet kompakt positioniert werden können.

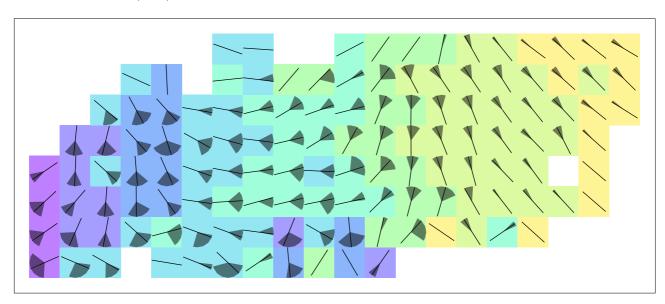
2 Methoden

2.1 Streuungsfächer

Tafel 1: Wind 5 (2012)



Tafel 2: Windfächer 5 (2012)



Die erste Kunstgrafik zeigt ein Strömungsbild, in welchem sich die Strömungsrichtungen von links nach rechts immer weiter im Gegenuhrzeigersinn drehen. Diese Bewegung ist auch von unten nach oben zu sehen, allerdings weniger deutlich.

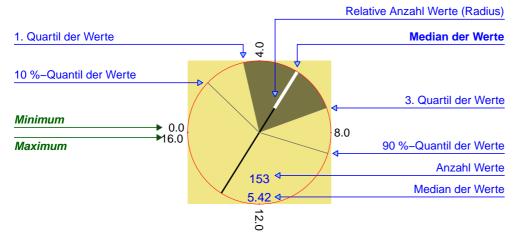
Tafel 1

Wenn man die Strömungsrichtungen von links nach rechts betrachtet, fallen zwei Bereiche mit besonders ausgeprägten Richtungsänderungen auf: In der Mitte und im ersten Viertel.

Gut sichtbar sind auch Regionen mit «ungeordneten Strömungsverhältnissen»: Auf den unteren beiden Zeilen und im linken Bereich.

Die zweite Kunstgrafik zeigt die gleichen Strömungsrichtungswerte. Zusätzlich gibt es aber noch kleine Fächer, welche die Streuungen der Strömungsrichtungswerte sichtbar machen. Die Streuungen im rechten Bereich sind deutlich kleiner. Die Grafikzellen in Regionen mit den «ungeordneten Strömungsverhältnissen» weisen mehrheitlich grosse Streuungen auf.

▶ Tafel 2



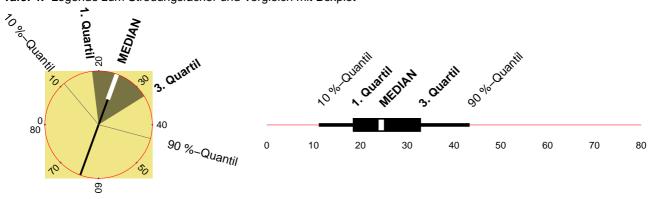
Tafel 3: Legende zum Streuungsfächer (360°)

Die «Streuungsfächerkarte»¹ ist eine Grafik, deren Gestaltung von meteorologischen Windströmungskarten inspiriert wurde. Anstelle eines Pfeiles, welcher die Windrichtung anzeigt, wird ein gerichtetes Fächersymbol verwendet. Es zeigt Anzahl (Länge der vorderen, weissen Mittellinie), Median (Symbolausrichtung) und Streuung (schwarzer Fächer und graue Federn) der Messwerte. Es kann anstelle von Boxplots² verwendet werden.

Die Federn zeigen je nach Bedarf auf die Quantile zu $5\,\%$ und $95\,\%$ oder zu $10\,\%$ und $90\,\%$ der Messwerte.

► Tafeln 3 und 4

Tafel 4: Legende zum Streuungsfächer und Vergleich mit Boxplot



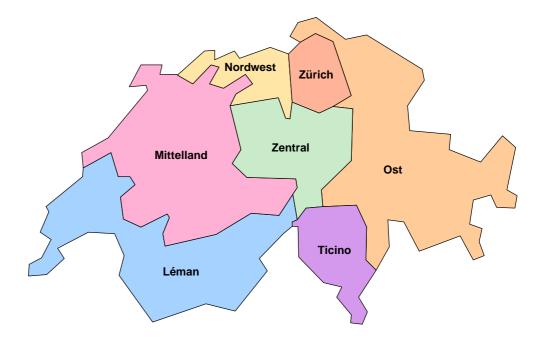
Streuungsfächerkarte

¹ Englisch: «fan chart». – Fischer [Neue Grafiken I, 2010]: 27 ff.

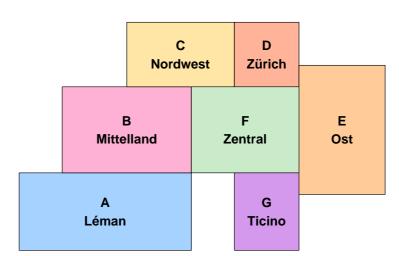
² Tukey [EDA, 1977]. Das Prinzip des Boxplots ist unter dem Namen «range bar» bereits zu finden in: Spear [Charting Statistics, 1952]: 166.

2.2 Pseudogeografische Anordnung

Tafel 5:Die sieben
Grossregionen
der Schweiz



Tafel 6:Schema der sieben
Grossregionen
der Schweiz



Tafel 7:Raster der sieben
Grossregionen
der Schweiz (3x3)



25.09.2012 5

Geografisch definierte Gebiete – wie z.B. Kantone, Regionen, Länder – sind normalerweise unterschiedlich gross. Deshalb dürfen deren Flächen nicht benutzt werden, um Daten abzubilden, die keinen direkten Bezug zur Fläche haben.

In pseudogeografisch angeordneten Grafiken wird pro Gebiet ein Quadrat (oder Rechteck) einheitlicher Grösse verwendet. Diese Felder werden matrixförmig neben- und übereinander gestellt. So ist einerseits die Zeichungsfläche möglichst gut belegt, andererseits bleiben die relativen geografischen Positionen einigermassen gewahrt. Geografisch westlich liegende Gebiete sind in solchen Rastern also links zu suchen, geografisch südlich liegende Gebiete unten, usw. Im Idealfall sind die direkten Nachbarn in dieser pseudogeografischen Anordnung auch in Wirklichkeit Nachbarn.

Pseudogeografisch angeordnete Grafiken

► Tafeln 5 bis 9

2.3 Programmierung mit R

Die statistischen Auswertungen und deren grafische Aufarbeitung wurden in R, einer Implementation von S, geschrieben.³

Grafik-Programmierung

³ Vgl. http://www.r-project.org/; Chambers [R, 2008]; Dalgaard [R, 2008]; Sarkar [Lattice, 2008]; Murrell [R Graphics, 2005].

СН	BS		SH	TG	AR
JU	BL	AG	ZH	SG	AI
NE	so	LU	ZG	SZ	GL
FR	BE	OW	NW	UR	GR
GE	VD	VS		TI	

Tafel 8: Kantonsraster, mit Gruppierung nach Grossregionen (5x6)

Tafel 9: Raster der Prämienregionen, mit Gruppierung nach Grossregionen (4x11)

СН	BS	BL	.2	ZH	.2	.3	SH	.2	TG	AR
NE	JU	SO	AG	LU	.2	.3	ZG	SZ	SG	AI
FR	.2	BE	.2	.3	NW	OW	UR	GL	.2	.3
GE	VD	.2	.3	VS	.2	TI	.2	GR	.2	.3

3 Daten

3.1 Statistik über Einkommen und Lebensbedingungen

SILC

In der "Statistik über Einkommen und Lebensbedingungen" (SILC) werden – parallel zu Erhebungen in 25 europäischen Ländern – Einkommenstruktur, Armut und soziale Ausgrenzung gemssen.⁴ Nach Piloterhebungen in den Jahren 2004 und 2005 wurden 2009 Daten aus rund 7 000 Haushalten und etwas über 17 000 Personen erhoben.⁵

Im ersten Auswertungsteil dieser Arbeit wurden Daten aus der Erhebung 2010 zur "Statistik über Einkommen und Lebensbedingungen" (SILC) des Bundesamtes für Statistik (BFS) verwendet. Die Daten beziehen sich auf das Jahr 2009.⁶

Als abhängige Variable wurde das verfügbare Äquivalenzeinkommen pro Person nach Schweizer Berechnungsart verwendet.⁷ Es basiert auf dem «verfügbaren Einkommen».

Das verfügbare Einkommen eines Haushalts wurde berechnet, indem man die Summe der Bruttoeinkommen⁸ sämtlicher Haushaltsmitglieder um «die obligatorischen Transferausgaben, d. h. Sozialversicherungsbeiträge, Steuern, Krankenkassenprämien für die Grundversicherung und zu leistende Unterhaltsbeiträge» reduzierte.⁹

Das verfügbare monatliche Äquivalenzeinkommen ist für alle Personen eines Haushalts – unabhängig vom Alter – identisch. Es wurde berechnet, indem man das verfügbare Haushaltseinkommen durch die gewichtete Anzahl der Personen im Haushalt dividierte. Das älteste Haushaltsmitglied wurde dabei mit 1, jede weitere Person ab 14 Jahren mit 0.5 und jedes Kind unter 14 Jahren mit 0.3 gewichtet. (Bei dieser Berechnungsweise darf das Äquivalenzeinkommen aller Personen nicht addiert werden. Um das insgesamt verfügbare Einkommen zu berechnen, müssten die verfügbaren Haushaltseinkommen aller Haushalte addiert werden.)

Um die Daten, die bis zum Ende der Erhebung effektiv erhoben werden konnten, der geplanten, gemäss der Volkszählung statifizierten Stichprobe anzupassen, wurden für jede Person Gewichte berechnet: Ein haushaltsabhängiges Gewicht und ein Gewicht für die Person als Individuum. 11 Das haushaltsabhängige Gewicht ist für alle Personen im gleichen Haushalt identisch.

Entsprechend der Empfehlung des BFS wurde für Auswertungen zum Äquivalenzeinkommen das haushaltsabhängige Gewicht verwendet.¹²

In den SILC-Daten der Erhebung 2010 wurde das Alter berechnet als <2009 – Geburts-jahr). Kinder, die 2010, aber noch vor dem Interview, zur Welt kamen, wurden auch gezählt und erhielten – gleich wie die Kinder mit Jahrgang 2009 – ein Alter von null Jahren. Da sich das Äquivalenzeinkommen auf das Jahr 2009 bezieht, müssten eigentlich die im Jahr 2010 geborenen Kinder weggelassen werden. Das BFS empfahl, sie jedoch aus Gründen der internationalen Vergleichbarkeit nicht aus dem Datenbestand zu entfernen (und das Äquivalenzeinkommen nicht neu zu berechnen). – Entsprechend sind Personen, die im Datenjahr 2009 noch lebten und im Erhebungsjahr 2010 gestorben sind, *nicht* im Datenbestand enthalten.

Die Namen der in den Grafiken verwendeten Variablen beginnen mit *P_*. Dies bedeutet, dass es sich um personenbezogene Werte handelt. Variablen, die haushaltsbezogene Werte

Verfügbares Einkommen

Äquivalenzeinkommen

Gewichtung der Haushalte und der Personen

Im Erhebungsjahr geborene und gestorbene Personen

Variablenbezeichnungen

25.09.2012 7

⁴ BFS-CH [Einführung SILC Medienmitteilung, 2007]. – http:// www.silc.bfs.admin.ch /.

⁵ Vgl. BFS-CH [Resultate SILC 2009, 2010].

⁶ Vgl. BFS-CH [Resultate SILC 2010 Medienmitteilung, 2011].

⁷ Variable *CPD010* aus der Personenzusatzdatei der SILC-Befragung 2010.

⁸ Das Buttoeinkommen umfasst «die Einkommen aus unselbstständiger und selbstständiger Erwerbstätigkeit, Renten und Sozialtransfers, Vermögenserträge, Unterhaltszahlungen sowie andere regelmässige Transfereinkommen von anderen Haushalten usw. Um den finanziellen Vorteilen bei selbst genutztem Wohneigentum oder Mietobjekten, deren Mietzinsen unter dem marktüblichen Preis liegen, Rechnung zu tragen, wird zum Bruttoeinkommen der betreffenden Haushalte ein um die effektiven Wohnkosten reduzierter Betrag addiert (fiktive Miete), der dem Nutzungswert des Objekts entspricht.» BFS-CH [Resultate SILC 2009, 2010]: 8.

⁹ BFS-CH [Resultate SILC 2009, 2010]: 8.

¹⁰ BFS-CH [Resultate SILC 2009, 2010]: 8f. Vgl. auch: Förster [Equivalences Scales, 2009] und Förster et al. [OECD Income Measuring, 2009]: 7f,

¹¹ Pea [SILC Pondération Guide, 2012]. Vgl. auch Graf [SILC Pondération Méthode, 2008]: 36 ff.

¹² Wie in Pea [SILC Pondération Guide, 2012]: 2 und auch in der Diskussion mit dem BFS vorgeschlagen, wurde die Variable *RB050* verwendet.

 $^{^{13}}$ Vgl. Eurostat [Algorithms, 2009]: 18. – Verwendet wurde die Variable $\it RX020$ aus der Registerdatei.

bezeichnen, beginnen mit H.

Die Variable zum Äquivalenzeinkommen nach Schweizer Berechnungsart heisst in dieser Studie *P_Äquivalenzeinkommen_CH*. Sie gibt das Äquivalenzeinkommen in 1 000 Fr. pro Jahr an

Altersgruppen sind beschriftet als [0,18) oder [18,25) usw. Die eckige Klammer bezeichnet einen Wert, der eingeschlossen ist, die runde Klammer einen ausgeschlossenen Wert. Somit steht [18,25) für 18 bis 24 Jahre.

Beschriftung Altersgruppen

3.2 Prämienangebote der Krankenversicherer

In der Schweiz publizieren die gut 60 Krankenversicherer ihre Angebote mit den Prämien für das Folgejahr jeweils im Herbst. Jeder Krankenversicherer kann die Prämien nach Alter, «Prämienregion», Franchise und Versicherungsmodell differenzieren.

Pro Kanton wurden vom Bundesamt für Sozialversicherung (BSV) pro Kanton ein bis drei Prämienregionen definiert. Dazu wurde jede Gemeinde aufgrund des damaligen Kostenniveaus einer Prämienregion zugeordnet.¹⁴

Für den zweiten Auswertungsteil dieser Arbeit wurden die beim Bundesamt für Gesundheit (BAG) verfügbaren Daten zu den Prämienangeboten der Krankenversicherer im Jahre 2012 zur obligatorischen Krankenversicherung ab Alter 26 in den kantonalen Prämienregionen verwendet.¹⁵

Prämienangebote

Prämienregionen

Daten des BAG

4 Resultate

4.1 Statistik über Einkommen und Lebensbedingungen

Vorbemerkung: Damit die Präzision der Daten abgeschätzt werden kann, verlangt das BFS – sinnvollerweise – die Angabe der Konfidenzintervalle von Resultaten. Die Darstellung von Konfidenzintervallen bei Streuungsfächern wurde bislang noch nicht entwickelt. Die Konfidenzintervalle werden deshalb nur in den Boxplots angezeigt. Sie entsprechen der Breite der Kerben.¹⁶

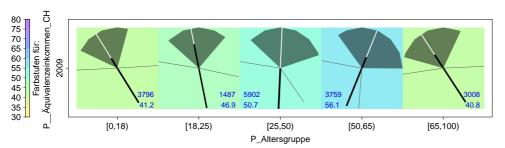
Anzeige des Konfidenzintervalls

Die erste Grafik mit einer Unterteilung nach fünf Altersgruppen (mit etwa 20-Jahres-Klassen, so wie sie auch im SILC-Projekt verwendet wurden¹⁷) zeigt, dass Personen im Erwerbsalter

Fünf Altersgruppen

▶ Tafel 10

¹⁷ Vgl. z. B. BFS-CH [Resultate SILC 2010 Medienmitteilung, 2011]: 6.



Tafel 10: Streuung des Äquivalenzeinkommens 2009 nach fünf Altersgruppen



Die schwarzen Fächer zeigen die Werte der Variablen "P__Äquivalenzeinkommen_CH" (gemessen in 1000 Fr. pro Jahr, gewichtet) vom ersten bis zum dritten Quartil auf einer Kreisskala von 0 bis 100. Die grauen Federn zeigen die Positionen des 10 %– und des 90 %–Quantils. Mindestanzahl Personen pro Zelle: 200. Die Zahlen zeigen: Anzahl (Personen), Median.

BFS, Erhebung über Einkommen und Lebensbedingunger SILC-2010 Version 26.6.12, inkl. fiktive Miete SILC 2010 : 17952 Personen

Z I M - Test / 102 SILC.fanchart:fan] 25.09.2012

¹⁴ Die einheitlichen Prämienregionen werden seit 2004 angewendet. – Vgl. auch http://www.bsv.admin.ch / dokumentation / medieninformationen / archiv / ? lang = de & msgsrc = / 2002 / d / 02053001.htm [2012-08].

¹⁵ Exceldateien zu den «Krankenversicherungsprämien OKP 2012 CH – EU ... 04.10.2011» unter http://www.bag.admin.ch > Themen > Krankenversicherung > Statistiken > Statistikpublikationen [2012-08]. – Vgl. auch http://www.priminfo.ch.

 $^{^{16}}$ In der Funktion *boxplot.stats* von R wird das Konfidenzintervall berechnet als \pm 1.58 \times Interquartilsabstand \div Quadratwurzel von n. Diese Faustregel ist auch in Sachs/Hedderich [Statistik, 2006]: 274 zu finden.

in Haushalten mit tendenziell höherem Äquivalenzeinkommen leben. Personen im Pensionsalter jedoch leben in Haushalten mit deutlich verringertem mittleren Äquivalenzeinkommen.

Aufgrund der Öffnung des Fächers und besonders an der Spreizung der Federn erkennt man, dass die Streuungen in der Altersgruppe der 50- bis 64-Jährigen am Grössten sind.

Acht Altersgruppen

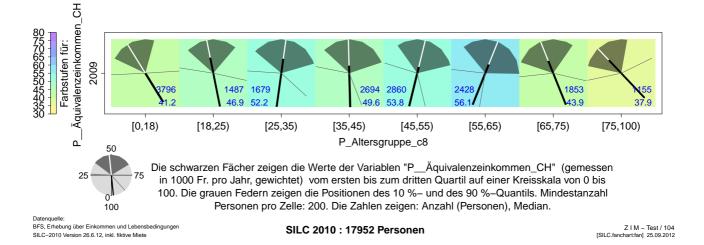
► Tafel 11

Tafel 12

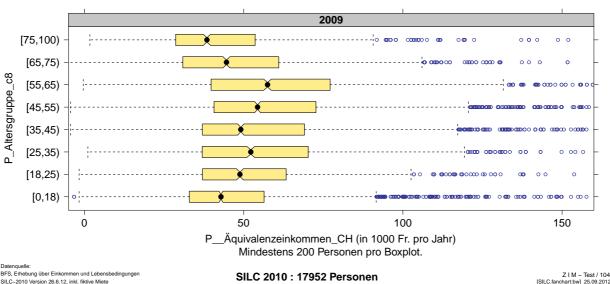
Bei einer differenzierteren Unterteilung nach acht Altersgruppen (mit etwa 10-Jahres-Klassen) zeigt sich, dass das mittlere Äquivalenzeinkommen nicht einfach mit zunehmendem Alter steigt, sondern dass es einen Rückgang gibt bei den 35- bis 44-Jährigen. Ausserdem leben Über-75-Jährige in Haushalten mit deutlich niedrigerem mittleren Äguivalenzeinkommen als 65- bis 75-Jährige.

Der Vergleich der Streuungsfächerkarte mit dem Boxplots zeigt, dass beim Boxplot die Balken mit der Streuungsangabe auf den ersten Blick dominanter sind als die Punkte, welche die Mediane anzeigen. Die Streubreiten zum Äquivalenzeinkommen im Erwerbsalter entwickeln sich mehr oder weniger kontinuierlich. Erst bei einem bewussteren Blick auf den Verlauf der Medianpunkte fällt auch hier auf, dass der Median zum Äquivalenzeinkommen in der Altersgruppe der 35- bis 45-Jährigen deutlich nach links verschoben ist, also niedriger liegt.

Tafel 11: Streuung des Äguivalenzeinkommens 2009 nach acht Altersgruppen



Tafel 12: Streuung des Äquivalenzeinkommens 2009 nach acht Altersgruppen



SILC-2010 Version 26.6.12, inkl. fiktive Miete

SILC 2010: 17952 Personen

9 25 09 2012

Wenn die Auswertung nicht nur nach Alter, sondern gleichzeitig auch nach Geschlecht differenziert wird, zeigt sich erstmals ein wesentlicher Pluspunkt der Streuungsfächerkarte. Mit ihr ist nicht nur ein direkter Vergleich der Äquivalenzeinkommen von einer Altersgruppe zur nächsten möglich (horizontal), sondern auch ein direkter Vergleich zwischen den Äquivalenzeinkommen bei Männern und Frauen in der gleichen Altersgruppe (vertikal).

Die Streuungsfächerkarte zeigt, dass die erwachsenen Männer in allen Altersgruppen in Haushalten mit einem höheren mittleren Äquivalenzeinkommen leben als die Frauen. (Die Fächer der Männer sind etwas mehr nach rechts gedreht als jene der Frauen.) Am ausgeprägtesten ist dieser Unterschied bei den 55- bis 64-Jährigen und im Pensionsalter.

Dass das mittlere Äquivalenzeinkommen bei Männern in der Altersgruppe der 55- bis 64-Jährigen deutlich höher ist als bei den Altersgruppen der unmittelbar Jüngeren und Älteren ist sowohl in der Streuungsfächerkarte auch auch im Boxplot-Diagramm gut zu erkennen.

Dass die mittleren Äquivalenzeinkommen bei den Frauen dieser Altersgruppe der 55- bis 64-Jährigen mehr streuen als in den daneben liegenden Altersgruppen, erscheint vielleicht in der Boxplot-Darstellung etwas augenfälliger, kann aber anhand der Spreizung der Federn und der Öffnung des Fächers auch gut aus der Streuungsfächerkarte gelesen werden.

Die Auswertung nach Altersgruppen und Bildungsstand zeigt, dass der Rückgang des mittleren Äquivalenzeinkommens bei den 35- bis 44-Jährigen im Wesentlichen bei Personen mit Sekundar-II-Ausbildung auftritt. Während das mittlere Äquivalenzeinkommen bei jüngeren Personen mit Tertiär-Ausbildung in dieser Altersgruppe noch ein wenig tiefer ist, scheint bei den Personen mit Primarschulausbilung das mittlere Äquivalenzeinkommen hauptsächlich in der Altersstufe der 25- bis 34-Jährigen zurückzugehen. Es könnte die Hypothese überprüft werden, ob dies z. B. mit einer früheren Familiengründung zusammenhängt. (Es könnte auch geprüft werden, ob das reduzierte mittlere Äquivalenzeinkommen einfach deswegen entsteht, weil das Haushaltseinkommen während der Familienzeit rechnerisch auf mehr Personen verteilt wird, oder ob die Reduktion besonders auch damit zusammenhängt, dass während der Familienzeit in manchen Familien das summierte Erwerbs-Arbeitspensum reduziert wird.)

Die Streuungen des Äquivalenzeinkommens nehmen pro Altersgruppe mit zunehmendem Bildungsstand zu. Dies ist in der Streuungsfächerkarte klar ersichtlich: Die Fächer auf der obersten Zeile sind deutlich grösser als die Fächer auf der mittleren Zeile. Im Boxplot-Diagramm sieht man beim Vergleich der drei Bildungsstandsgruppen zumindest gut, dass die Streuungsbalken grundsätzlich bei höherem Bildungsstand grösser sind.

Altersgruppen und Geschlecht

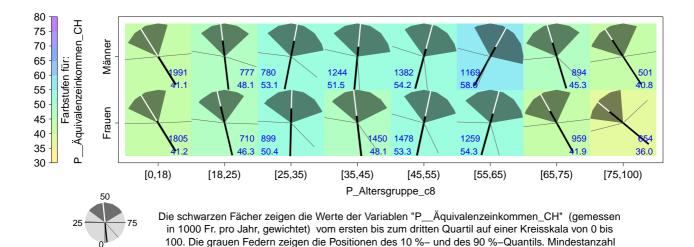
► Tafeln 13 und 14 (S. 11)

Altersgruppen und Bildungsstand

► Tafel 15 (S. 12)

► Tafel 16 (S. 12)

Tafel 13: Streuung des Äquivalenzeinkommens 2009 nach acht Altersgruppen und Geschlecht



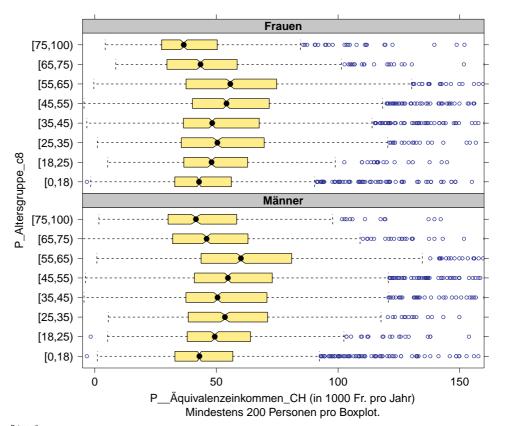
100
Datenquelle:
BFS, Erhebung über Einkommen und Lebensbedingungen
SILC-2010 Version 26.6.12, inkl. fiktive Miete

SILC 2010: 17952 Personen

Personen pro Zelle: 200. Die Zahlen zeigen: Anzahl (Personen), Median.

Z I M - Test / 109 [SILC.fanchart:fan] 25.09.2012

Tafel 14: Streuung des Äquivalenzeinkommens 2009 nach acht Altersgruppen und Geschlecht

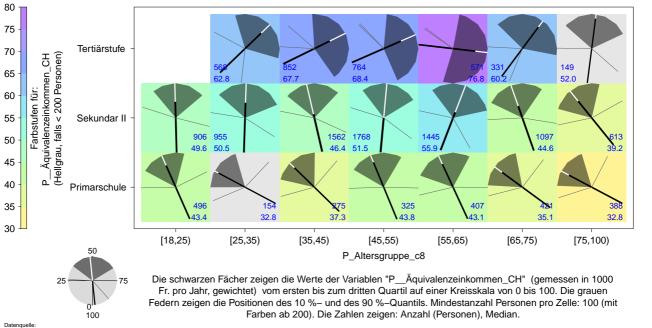


Datenquelle: BFS, Erhebung über Einkommen und Lebensbedingungen SILC–2010 Version 26.6.12, inkl. fiktive Miete

SILC 2010: 17952 Personen

Z I M – Test / 109 [SILC.fanchart:bw] 25.09.2012

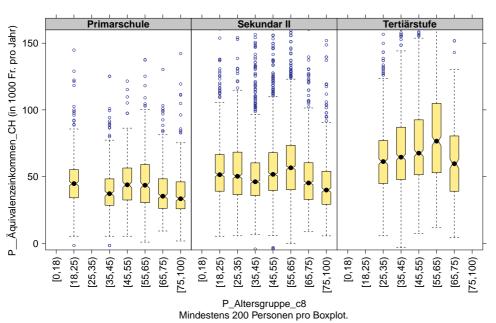
Tafel 15: Streuung des Äquivalenzeinkommens 2009 nach acht Altersgruppen und Bildungsstand



Datenquelle: BFS, Erhebung über Einkommen und Lebensbedingungen SILC–2010 Version 26.6.12, inkl. fiktive Miete

SILC 2010 : 17952 Personen

Z I M - Test / 112 [SILC.fanchart:fan] 25.09.2012



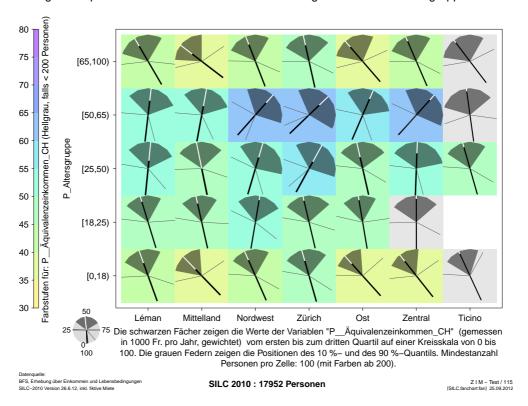
Tafel 16: Streuung des Äquivalenzeinkommens 2009 nach acht Altersgruppen und Bildungsstand

Datenquelle: BFS, Erhebung über Einkommen und Lebensbedingungen SII C-2010 Version 26.6.12 inkl. fiktive Miete

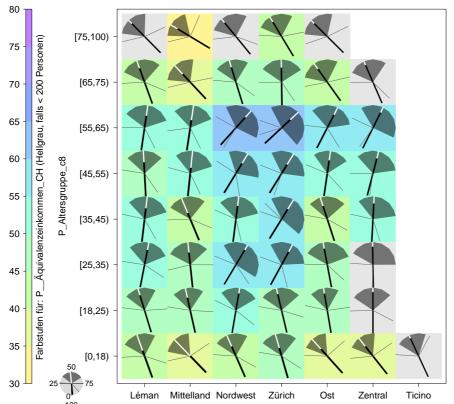
SILC 2010 : 17952 Personen

Z I M - Test / 112 [SILC.fanchart:bw] 25.09.2012

Tafel 17: Streuung des Äquivalenzeinkommens 2009 nach Grossregionen und fünf Altersgruppen



Tafel 18: Streuung des Äquivalenzeinkommens 2009 nach Grossregionen und acht Altersgruppen



Die schwarzen Fächer zeigen die Werte der Variablen "P__Äquivalenzeinkommen_CH" (gemessen in 1000 Fr. pro Jahr, gewichtet) vom ersten bis zum dritten Quartil auf einer Kreisskala von 0 bis 100. Die grauen Federn zeigen die Positionen des 10 %- und des 90 %-Quantils. Mindestanzahl Personen pro Zelle: 100 (mit Farben ab 200).

BFS, Erhebung über Einkommen und Lebensbedingungen SILC 2010 : 17952 Personen

Z I M - Test / 116 [SILC.fanchart:fan] 25.09.2012

Streuungsfächerkarten haben ihre Stärke dort, wo sowohl in x- wie auch in y-Richtung viele Kategorien abgebildet werden. Deshalb wurden in den nächsten beiden Grafiken waagrecht die sieben Grossregionen und senkrecht fünf bzw. acht Altersstufen abgebildet.

Da es sich bei den Grossregionen um nominale Kategorien handelt (und nicht um ordinale), ist eine Analyse der Entwicklung von links nach rechts nur beschränkt sinnvoll. Problemlos verglichen werden dürfen die Werte von jeweils zwei Regionen. Am einfachsten ist dies bei nebeneinander liegenden Regionen.

Man sieht auf diese Weise z.B., dass das mittlere Äquivalenzeinkommen in der Grossregion «Espace Mittelland» niedriger ist als in der Region «Nordwestschweiz»: Alle Fächer der Nordwestschweiz sind im Vergleich zu den Fächern des Mittellandes etwas mehr nach rechts gedreht. Anhand der Zahlen sieht man, dass der Unterschied durchgehend mehrere tausend Franken (pro Jahr) beträgt.

Beim Zeichnen eines Boxplot-Diagramms muss man sich entscheiden, welche Kategorien man für den unmittelbaren Vergleich benutzen will. Da der Vergleich zwischen den ordinalen Altersgruppen sinnvoller ist als jener zwischen den nominalen Grossregionen, wurden Grafikfelder pro Grossregion gezeichnet. Sie sind pseudogeografisch angeordnet.

Beim einfachen Vergleich der Äquivalenzeinkommen in den Grossregionen zeigt sich, dass der Quervergleich zwischen den verschiedensten Grossregionen bei pseudogeografisch angeordneten Streuungsfächern einfacher ist als im Boxplot-Diagramm. Ein Streuungsfächer hat bis zu vier direkte und damit gut vergleichbare Nachbarn. Ziemlich gut vergleichbar sind darüberhinaus auch die vier diagonal positionierten Nachbarn.

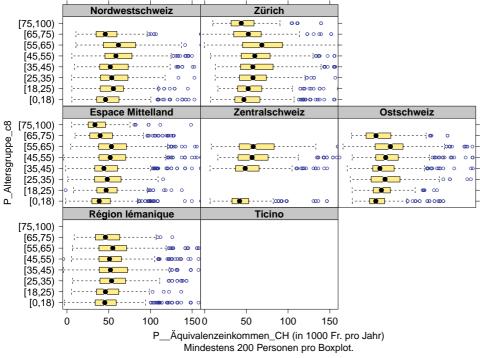
Altersgruppen und Grossregionen

▶ Tafel 17 und 18 (S. 13)

► Tafel 19

Grossregionen

► Tafel 20 und 21 (S. 15)



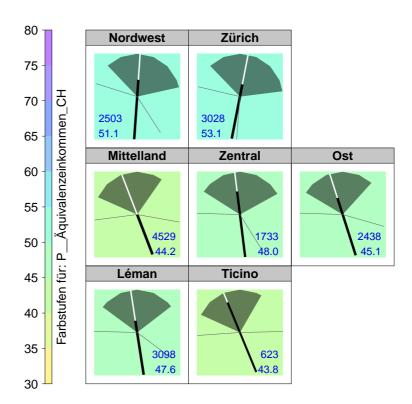
Tafel 19: Streuung des Äquivalenzeinkommens 2009 nach Grossregionen und acht Altersgruppen

Datenquelle: BFS, Erhebung über Einkommen und Lebensbedingunger SILC-2010 Version 26.6.12, inkl. fiktive Miete

SILC 2010 : 17952 Personen

Z I M - Test / 116 [SILC.fanchart:bw] 25.09.2012

Tafel 20: Streuung des Äquivalenzeinkommens 2009 (nach Grossregionen)



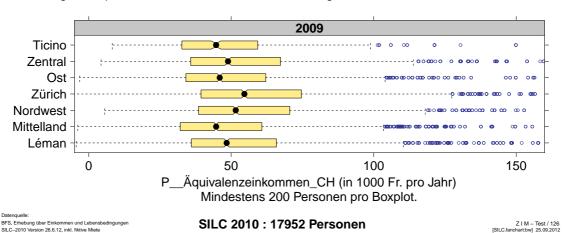


Die schwarzen Fächer zeigen die Werte der Variablen "P__Äquivalenzeinkommen_CH" (gemessen in 1000 Fr. pro Jahr, gewichtet) vom ersten bis zum dritten Quartil auf einer Kreisskala von 0 bis 100. Die grauen Federn zeigen die Positionen des 10 %– und des 90 %–Quantils. Mindestanzahl Personen pro Zelle: 200. Die Zahlen zeigen: Anzahl (Personen), Median.

Datenquelle: SILC 2010 : 17952 Personen BFS, Erhebung über Einkommen und Lebensbedingungen SILC-2010 Version 26.6.12, inkl. fiktive Miete

Z I M - Test / 126 [SILC.fanchart:fan] 25.09.2012

Tafel 21: Streuung des Äquivalenzeinkommens 2009 nach Grossregionen



Die Auswertung des Äguivalenzeinkommens nach Haushaltstypen und Verstädterungsgrad wirkt auf den ersten Blick unübersichtlich. Um eine bessere Interpretation zu ermöglichen, arbeitete ich in zwei Richtungen weiter: Einerseits durch Veränderung der Reihenfolge der Haushalte und andererseits durch Auswertung des gesamten verfügbaren Einkommen pro Haushalt.

Wenn die Haushaltstypen nach dem Median der Äquivalenzeinkommen geordnet werden, ergibt sich eine schöne Fächerkarte. Die äussersten Spalten zeigen sogleich, dass das mittlere Äquivalenzeinkommen in Haushalten mit zwei Erwachsenen unter 65 Jahren am höchsten ist und in Haushalten von Alleinerziehenden oder von Familien mit mehr als zwei Kindern am kleinsten. Insgesamt aber ist die Entwicklung des zunehmenden mittleren Äguivalenzeinkommen von links nach rechts nicht unmittelbar interpretierbar.

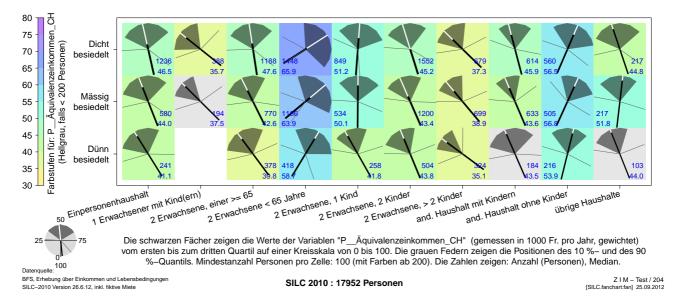
Die Grafik zeigt auch, dass die Unterschiede des mittleren verfügbaren Äquivalenzeinkommens zwischen unterschiedlich dicht besiedelten Gebieten klein sind.

Haushaltstypen und Verstädterungsgrad

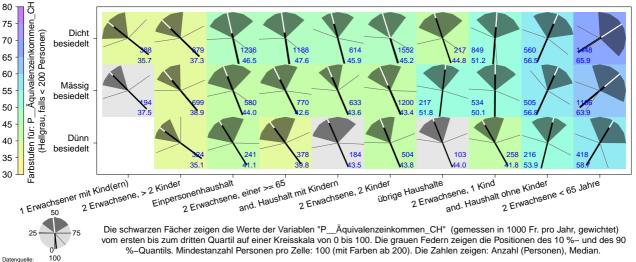
► Tafeln 22

► Tafel 23

Tafel 22: Streuung des Äquivalenzeinkommens 2009 nach Haushaltstyp und Verstädterungsgrad



Tafel 23: Streuung des Äquivalenzeinkommens 2009 nach Haushalttyp (geordnet) und Verstädterungsgrad



100 BFS, Erhebung über Einkommen und SILC-2010 Version 26.6.12, inkl. fiktiv SILC 2010: 17952 Personen

16 25 09 2012 Haushaltseinkommen nach Haushaltstypen und Verstädterungsgrad

▶ Tafeln 24 und 25

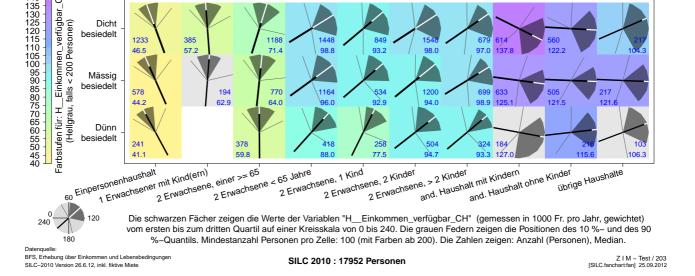
Versuchsweise machte ich anschliessend auch Auswertungen zum gesamten verfügbaren Einkommen des Haushalts, in dem die Personen leben (Schweizer Berechnungsart).

Als ich die Haushaltstypen einigermassen logisch und zugleich mehr oder weniger nach aufsteigender Personenzahl ordnete, erschien eine besonders schöne und interessante Streuungsfächerkarte.

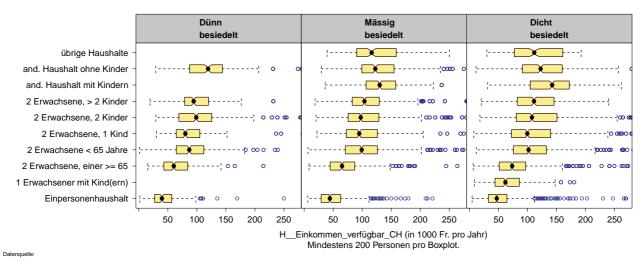
Insgesamt drehen sich die Fächer von links nach rechts im Uhrzeigersinn, d. h. das mittlere verfügbare Haushaltseinkommen wird von links nach rechts (meist) höher. Von unten nach oben kann vielfach nur eine leichte Drehung im Uhrzeigersinn ausgemacht werden; manchmal gibt es keine Drehung, vereinzelt eine Drehung im Gegenuhrzeigersinn. Die Unterschiede des mittleren verfügbaren Haushaltseinkommens zwischen dünn und mässig besiedelten Gebieten sind grösser als zwischen mässig und dicht besiedelten Regionen.

Was bei der Gesamtbetrachtung auffällt, sind einerseits die hohen Streuungen, auf welche die Fächer hinweisen, die trotz grosszügig ausgelegter Skala recht breit sind. Andererseits gibt es drei Farbblöcke, in denen deutlich unterschiedliche mittlere verfügbare Haushaltsein-

Tafel 24: Streuung des verfügbaren Haushaltseinkommens 2009 nach Haushaltstyp und Verstädterungsgrad



Tafel 25: Streuung des verfügbaren Haushaltseinkommens 2009 nach Haushaltstyp und Verstädterungsgrad



BFS, Erhebung über Einkommen und Lebensbedingungen SILC-2010 Version 26.6.12, inkl. fiktive Miete

SILC 2010 : 17952 Personen

Z I M - Test / 203 [SILC.fanchart:bw] 25.09.2012

kommen vorgefunden werden. Haushalte mit einem einzigen Erwachsenen im Erwerbsalter verfügen im Mittel über ein deutlich niedrigeres Haushaltseinkommen als Haushalte mit mehr als einem Erwachsenen im Erwerbsalter. (Es könnte überprüft werden, ob dieser Unterschied hauptsächlich auf doppelverdienende Paare zurückgeführt werden kann oder auf ein höheres Einkommensniveau der erwerbstätigen Person[en] in diesen Haushalten.)

Es zeigt sich auch, dass das mittlere verfügbare Haushaltseinkommen in Haushalten mit zwei Erwachsenen und einem oder mehr Kindern im Mittel deutlich weniger unterscheidet als gegenüber Haushalten mit nur einem oder mehr als zwei Erwachsenen. Das mittlere Haushaltseinkommen scheint stark abhängig von der Anzahl Erwachsenen und weniger von der Anzahl Kinder. (Überprüft werden könnte, ob die Haushalte mit mehr Kindern grundsätzlich weniger Geld zur Verfügung haben oder ob das Verhältnis der Gewichtungsfaktoren von 0.3 pro Kind und 0.5 pro zusätzlichem Erwachsenen zur Diskussion gestellt werden müsste.)

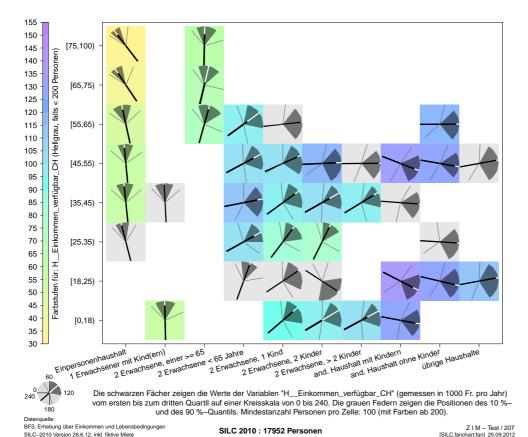
In einer weiteren Auswertung habe ich das Haushaltseinkommen nach Haushaltstypen und Altersgruppen differenziert. Von den drei Blöcken unterscheiden sich hier der linke und der mittlere recht deutlich, d. h. das mittlere Haushaltseinkommen scheint hauptsächlich von der Anzahl Erwachsener abhängig zu sein.

Bei der Betrachtung der untersten Zeile mit den Kindern im Alter von null bis 17 Jahren zeigt sich, dass Kinder mit mehr Geschwistern (und zwei Erwachsenen) tendenziell in Haushalten mit einem höheren Haushaltseinkommen leben.

Auffällig ist im Weiteren der unerwartet heftig nach unten abgedrehte Fächer der Haushalte mit zwei Erwachsenen und zwei Kindern in der Altersstufe der 18- bis 24-Jährigen. Auch stehen hier die Federn extrem weit ab, ganz besonders jene, welche das 90 %-Quantil anzeigt. Derartige «optische Ausreisser» verdienen einen Kontrollblick in die Daten: Gibt es von dort her eine Erklärung oder handelt es sich bei den gezeigten, unerwartet hohen verfügbaren Haushaltseinkommen tatsächlich um repräsentative Werte? (Die kleine weisse Spitze der Medianlinie zeigt, dass in dieser Zelle Daten von verhältnismässig wenig Personen dargestellt sind.)

Haushaltstypen und Altersgruppen

▶ Tafel 26



Tafel 26: Streuung des verfügbaren Haushaltseinkommens 2009 nach Haushaltstyp und acht Altersgruppen

4.2 Prämienangebote der Krankenversicherer

Prämienangebote pro Kanton

► Tafel 27

Die Streuungsfächer zeigen die Prämienangebote zur obligatorischen Grundversicherung pro Kanton auf einer Kreisskala von 250 bis 550 Franken pro Monat. (Ein Viertelkreis umfasst also 75 Franken.) Die Federn zeigen hier auf das 5%- und das 95%- Quantil. Die Zellen der Kantone sind pseudogeografisch angeordnet. (Der Kanton Genf ist deshalb links unten, der Kanton St. Gallen rechts oben zu finden.)

Die mehr grüne bis gelbe Färbung bei vielen Kantonen in der rechten Hälfte zeigt, dass das die mittleren Prämienangebote in der Zentral- und Ostschweiz oft niedriger sind als in meisten übrigen Kantonen.

Mit Abstand am höchsten sind die – violett gefärbten – Prämienangebote in den Kantonen Genf (GE) und Baselstadt (BS).

In einigen Kantonen fallen die schmalen Fächer auf. Dort variieren die mittleren 50 % der Prämienangebote nur wenig. Beispiele sind die beiden Appenzell (Al und AR), Zug (ZG), Obund Nidwalden (OW, NW), Uri (UR).

In einigen Kantonen fallen die weit ausgespannten grauen Federn auf. Hier streuen die mittleren 90% der Prämienangebote am meisten. Die Beispiele mit der grössten 90%-Spanne sind Baselstadt (BS) und Bern (BE). Bei diesen Kantonen bilden die Federn etwa eine Gerade, d. h. sie markieren einen Unterschied von ca. 150 Franken.

Prämienangebote pro Prämienregion

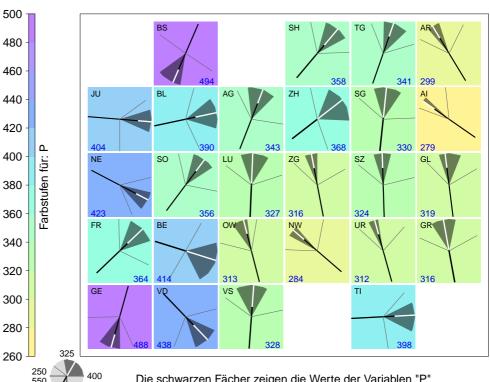
► Tafel 28

Wenn die Streuungsfächer nicht pro Kanton, sondern pro Prämienregion berechnet werden, werden sie in den Kantonen mit mehreren Prämienregionen erwartungsgemäss kleiner.

In dieser Darstellung wird das West-Ost-Gefälle der Prämienangebote sehr augenfällig. Auch die Ausnahmen werden gut erkannt: Die Prämienangebote in Freiburg Region 2 (FR2), in Solothurn (SO), im Aargau (AG) und im Wallis (VS1 und besonders VS2) sind klar niedriger als die Angebote in den meisten übrigen Prämienregionen der westlichen Landeshälfte.

Solche Analysemöglichkeiten zeigen, wie die pseudogeografische Anordnung dem Interpreten dieser Grafik so sehr entgegenkommt, dass er sie kaum mehr speziell in die Überlegungen einbringen muss.

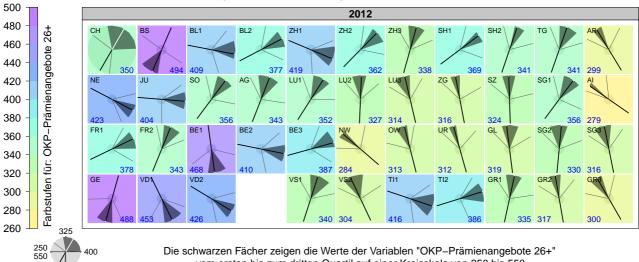
Tafel 27: CH: Krankenkassenprämienangebote pro Kanton



Die schwarzen Fächer zeigen die Werte der Variablen "P"
 vom ersten bis zum dritten Quartil auf einer Kreisskala von 250 bis 550.
 Die grauen Federn zeigen die Positionen des 5 %– und des 95 %–Quantils.
 Die Zahlen zeigen: Median.

BAG OFSP UFSP SFOPH

Z I M - v22 [PRE.128.fanchart.P..C] 02.08.2012



Tafel 28: CH: Krankenkassenprämienangebote pro Prämienregion

Die schwarzen Fächer zeigen die Werte der Variablen "OKP-Prämienangebote 26+
vom ersten bis zum dritten Quartil auf einer Kreisskala von 250 bis 550.
Die grauen Federn zeigen die Positionen des 5 %– und des 95 %–Quantils.
Die Fläche des abgedunkelten Kreises ist proportional zur Anzahl Messwerte.
Die Zahlen zeigen: Median.

Z I M [PRE.09Z.fanchart.P–12802]

5 Diskussion

www.bag.admin.ch

Datenguelle

Streuungsfächer ermöglichen eine kompaktere Darstellung als Boxplots. – Streuungsfächer können sowohl nebeneinander wie auch übereinander gestellt werden. Dies ermöglicht Vergleiche in zwei Richtungen. – Eine Hintergrundfärbung aufgrund des Medianwertes sowie die unterschiedlichen Öffnungen der Fächer lenken das Auge schnell zu aussergewöhnlichen Messwerten und Streuungen. – Veränderte Messwerte bewirken eine Drehung der Medianlinie. In einer Streuungsfächerkarte kann dies eine faszinierende Gesamtkomposition ergeben, die an Strömungsbilder von Winden erinnert.

Streuungsfächer

Eine geografisch ausgerichtete Darstellung ermöglicht eine platzsparende Anordnung von Grafikfeldern unabhängig von der Grösse des dazugehörigen geografischen Territoriums. – Sie ermöglichen auf den ersten Blick eine grobe Einschätzung regionaler Unterschiede. – Sofern der Betrachter die realen geografischen Positionen kennt, findet er auch bei einer Vielzahl von Grafikfeldern relativ rasch ein ihn interessierendes Feld. – Pseudeogeografische Anordnungen sind gegenüber einer Anordnung entsprechend den realen geografischen Positionen und Flächen dann besonders attraktiv, wenn sich die territorialen Ausdehnungen sehr stark unterscheiden oder wenn es für Vergleiche wünschenswert ist, dass die grafischen Elemente gleich gross sind.

Pseudogeografische Anordnungen

6 Anhang

6.1 Abkürzungen und Internetverweise

Tafel 29: Abkürzungen

Abkürzung	Text	Internet
BAG	Bundesamt für Gesundheit	http://www.bag.admin.ch/
BFS	Bundesamt für Statistik	http://www.bfs.admin.ch/
BSV	Bundesamt für Sozialversicherung	http://www.bsv.admin.ch/
CH	Schweiz	http://www.admin.ch/
EU	Europäische Union	http:// europa.eu.int /
EU-SILC	Statistics on Income and Living Conditions	http://epp.eurostat.ec.europa.eu/portal/page/portal/microdata/eu_silc
OKP	Obligatorische Krankenpflege-Grundversicherung	http://www.praemien.admin.ch/
R	The R Project for Statistical Computing	http:// www.r-project.org /
SILC	Statistik über Einkommen und Lebensbedingungen	http://www.silc.bfs.admin.ch/

6.2 Literaturverzeichnis

BFS-CH (2007) Einführung SILC Medienmitteilung

Bundesamt für Statistik. Einführung einer neuen jährlichen Erhebung über Einkommen und Lebensbedingungen (S/LC). Medienmitteilung. Neuchâtel 2007: 3 S. Internet: http://www.bfs.admin.ch/bfs/portal/de/index/infothek/erhebungen_quellen/blank/blank/silc/04.Document.88883.pdf.

BFS-CH (2010) Resultate SILC 2009

Bundesamt für Statistik. *Lebensbedingungen in der Schweiz 2009.* Resultate der Erhebung über die Einkommen und die Lebensbedingungen (SILC). Neuchâtel 2010: 15 S. Internet: http://www.bfs.admin.ch/bfs/portal/de/index/news/publikationen.html?publicationID = 4213.

BFS-CH (2011) Resultate SILC 2010 Medienmitteilung

Bundesamt für Statistik. *Haushalte mit Kindern haben mehr Schwierigkeiten, eine unerwartete Ausgabe zu tätigen*. Erhebung über die Einkommen und die Lebensbedingungen (SILC), Medienmitteilung. Neuchâtel 2011: 8 S. Internet: http:// www.bfs.admin.ch / bfs / portal / de / index / infothek / erhebungen _ _ quellen / blank / blank / silc / 04.Document.88883.pdf.

Chambers (2008) R

Chambers JM. Software for Data Analysis. Programming with R. New York (Springer) 2008: 498 S.

Dalgaard (2008) R

Dalgaard P. Introductory Statistics with R. 2nd Edition, New York (Springer) 2008: 363 S.

Eurostat (2009) Algorithms

European Commission Eurostat, Working Group Living Conditions. *Algorithms to Compute Pensions Indicators Based on EU-SILC and Adopted Under the Open Method of Coordination (OMC)*. 10–12 June 2009. Doc LC-ILC/40/09/EN, Luxemburg 2009: 58 S. Internet: http:// epp.eurostat.ec.europa.eu / portal / page / portal / income _ social _ inclusion _ living _ conditions / documents / tab / Tab / DOC%2520 LC _ ILC _ 40 _ 09 _ EN _ Pensions%2520 Portfolio%2520 June%2520 09.pdf.

Fischer (2010) Neue Grafiken I

Fischer W. Neue Grafiken zur Datenvisualisierung. Band 1: Speichengrafiken, Streuungsfächerkarten, Differenz-, Sequenz- und Wechseldiagramme. Wolfertswil (ZIM) 2010: 107 S. Internet: http://www.fischer-zim.ch/studien/Neue-Grafiken-I-1003-Info.htm.

Förster (2009) Equivalences Scales

Förster MF. What Are Equivalence Scales?. (OECD) 2009: 2 S. Internet: http:// www.oecd.org / social / familiesandchildren / 35411111.pdf.

Förster et al. (2009) OECD Income Measuring

Förster MF, d'Ercole MM. The OECD Approach to Measuring Income Distribution and Poverty. Strengths, Limits and Statistical Issues. (OECD) 2009: 30 S. Internet: http://umdcipe.org/conferences/oecdumd/conf_papers/Papers/The%2520 OECD%2520 Approach%2520 to%2520 Measuring%2520 Income%2520 Distribution%2520 and%2520 Poverty.pdf.

Graf (2008) SILC Pondération Méthode

Graf E. Pondérations du SILC pilote. SILC_I vague 2, SILC_II vague 1, SILC_I et SILC_II combinés. Neuchâtel (BFS) 2008: 77 S. Internet: http:// www.bfs.admin.ch / bfs / portal / de / index / infothek / erhebungen _ _ quellen / methodenberichte.Document.105401.pdf.

Murrell (2005) R Graphics

Murrell P. R Graphics. (Chapman & Hall/CRC) 2005: 301 S.

Pea (2012) SILC Pondération Guide

Pea J. Guide pour utilisateurs des poids. Aus der zu den Daten der SILC-Erhebung 2010 gelieferten Dokumentation, Neuchâtel (BFS) 2012: 3 S.

Sachs/Hedderich (2006) Statistik

Sachs L, Hedderich J. *Angewandte Statistik*. Methodensammlung mit R. 12. vollständig neu bearbeitete Auflage, Berlin Heidelberg (Springer) 2006: 702 S.

Sarkar (2008) Lattice

Sarkar D. Lattice. Multivariate Data Visualization with R. New York (Springer) 2008: 265 S.

Spear (1952) Charting Statistics

Spear ME. Charting Statistics. New York Toronto London (McGraw-Hill) 1952: 253 S. Internet: http://archive.org/details/ChartingStatistics.

Tukey (1977) EDA

Tukey JW. Exporatory Data Analysis. Reading MA (Addison Wesley) 1977: 688 S.