

1 Mobilität in Tabellen (MiT) - Ziel und Inhalt

Das Tabellen-Tool **MiT** bietet einen schnellen Einstieg zur einfachen Nutzung der Daten der Haushaltsbefragung zum Personenverkehr Mobilität in Deutschland (MiD). Für jede Tabelle werden die Originaldaten auf der Basis ausgewählter aufbereiteter Variablen ausgezählt.

Es können Tabellen mit den Inhalten

- Haushalte,
- Personen,
- Wege am Stichtag,
- zurückgelegte Kilometer am Stichtag, oder
- im Verkehr zugebrachte Zeit am Stichtag

erstellt werden. Auch bezogene Werte, wie z.B. Wege je Person, km je Weg, können berechnet werden. Die Daten können nach maximal drei Brechungen (Variablen) unterteilt ausgezählt werden; z.B. Benutztes Verkehrsmittel in den Zeilen, Wegezweck in den Spalten, dies für jedes Bundesland und Deutschland insgesamt (Datenblöcke). Ausgewiesen werden die mit den jeweils korrekten Gewichten bestimmten hochgerechneten Werte. Zusätzlich können zu diesen Ergebnissen auch die Zeilen- bzw. Spalten-Prozenttabellen angefordert werden. Zur Beurteilung der statistischen Verlässlichkeit der Resultate ist auch der Ausweis der ungewichteten jeweiligen Fallzahlen sowie der relativen Standardfehler möglich.

Die Verkehrserhebung MiD ist aufgrund der zwei Befragungsformen (telefonisch: CATI, schriftlich: PAPI) und der differenzierten Einbeziehung der Kinder sehr komplex: In CATI ist der Umfang der Befragung höher, bei einem Stellvertreter-Interview (Proxy) werden jedoch einige Fragen nicht gestellt, andere Fragen waren nur für bestimmte Personengruppen relevant und wurden nur an diese gerichtet. Für die Auswertung sind daher drei Sachverhalte zu unterscheiden:

- Wenn die Frage für die befragte Person irrelevant ist (z.B. Führerscheinbesitz bei Kindern, regelmäßige berufliche Wege bei Nicht-Erwerbstätigen) ist diese im telefonischen Interview i.a. durch die Filterführung übersprungen worden. In den Daten ist i.a. „trifft nicht zu (tnz)“ codiert. Diese fehlenden Angaben werden als „qualified missing“ bezeichnet und können interpretiert und an der entsprechenden Stelle der Tabelle gezählt werden (z.B. als Person ohne Führerschein, als Person ohne berufliche Wege).
- Wenn aus anderen erhebungstechnischen Gründen einzelne Fragen nicht gestellt worden sind (z.B. Einschätzungsfragen an Kinder), so ist die Grundgesamt kleiner als die Gesamtbevölkerung. Um dennoch den Bezug zur gesamten Wohnbevölkerung in der Eckzahl ausweisen zu können, ist bei den aus solchen Fragen resultierenden Variablen eine Ausprägung „nicht enthalten, weil ...“ zugefügt worden.
- Andere fehlende Angaben (missings) haben sehr unterschiedliche Hintergründe. Neben der Verweigerung ist die proxy-Befragung der Hauptgrund für fehlende Angaben. Wenn ein Stellvertreter-Interview geführt worden ist, sind einige Fragen nicht gestellt worden. Für das Auswertungsprogramm ist hier die generelle Annahme getroffen worden, dass die Angaben der Antwortter auf die Nicht-Antwortter übertragen werden können, diese also der Struktur der Antwortter entsprechen. Die in der Tabelle gewünschten Variablen werden nachgewichtet, so dass –unabhängig von der Anzahl fehlender Angaben– die gleiche hochgerechnete Eckzahl (s. Tabelle) für Haushalte, Personen, bzw. Wege ausgewiesen werden sollte.¹

¹ Die hochgerechneten Einzelwerte können bei unterschiedlichen Tabellen verschieden sein, wenn die auszählenden Variablen nicht im jeweiligen Hochrechnungsrahmen enthalten sind. Im Zweifel ist der Wert, der mit weniger fehlenden Werten entstanden ist, vorzuziehen.

Die Fallzahlen mit fehlenden Angaben werden unterhalb jeder Fallzahlen-Tabelle ausgewiesen, so dass auch eine Kontrolle der gesamten Fallzahlen möglich ist.

Tabelleninhalt können die Zahl der Haushalte, die Zahl der Personen, die Anzahl der Wege, die zurückgelegte Wegstrecke (Personenkilometer Pkm) oder die Wegezeiten sein. Die Eckwerte sind der folgenden Tabelle zu entnehmen. Wenn der Tabelleninhalt eine höhere Hierarchie hat, als die Variablen nach denen ausgezählt werden soll, sind stets alle Elemente Tabelleninhalt. Dies gilt auch für den Tabelleninhalt im Nenner. Zum Beispiel liefert Wege / Person nicht Wege je mobiler Person,² sondern Wege je Person.

Übersicht Eckwerte

Deutschland Basiserhebung	Haushalte	Personen	Wege am Tag *
gewichtet in 1000	37.711	82.441	272.245
Fallzahl	25.848	61.729	169.759
	* einschl. regelm. berufliche Wege		

2 Hinweise zur Bedienung

2.1 Installation

Systemvoraussetzungen: Windows-32-Bit-System, ca. 20 MB freier Plattenplatz.

Eine Installationsroutine ist nicht erforderlich. Kreieren Sie ein Verzeichnis auf ihrer Festplatte und entpacken Sie die Datei MiT.zip in dieses Verzeichnis.

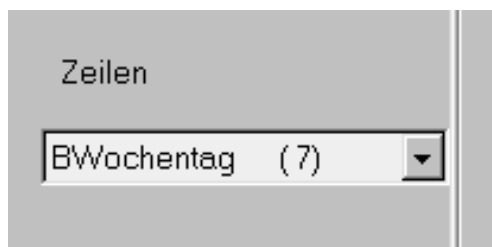
2.2 Bedienung

The screenshot shows a dialog box titled "Tabelle zusammenstellen". It has a blue title bar with a close button on the right. The main area is light gray and contains several sections:


- Blöcke:** A dropdown menu showing "(keine)".
- Haupt-Verkehrsmittel (6):** A label.
- Spalten:** A dropdown menu showing "WVmittel6 (6)".
- Zeilen:** A dropdown menu showing "B-Land (16)".
- Inhalt:** A dropdown menu showing "Wege".
- je:** A dropdown menu showing "-".
- Tabellen:** A group box containing three checked checkboxes: "Hochgerechnete Ergebnisse", "Spalten %", and "Zeilen %".
- Stichprobeninformation:** A group box containing two checked checkboxes: "Rel. Standardfehler" and "Ungew. Fallzahlen".

At the bottom of the dialog are two buttons: "OK" and "Abbruch".

² Um diese Werte auszuzählen, kann als dritte Variable „Person mobil am Stichtag“ z.B. als Blockvariable gewählt werden.



Der Hauptdialog erfolgt über das Dialogfenster zur Tabellenauswahl. Es wird zu Programmbeginn geöffnet, und wenn Sie den Menüeintrag „NeueTabelle“ wählen.

Im Tabellen-Auswahl-Fenster symbolisiert das innere Rechteck die entstehenden Tabellen. Mit einem Klick auf das jeweilige Dreieck  öffnen Sie Auswahlmenüs für Zeilen, Spalten und Blöcke. Zur Erleichterung der Tabellengestaltung ist die Zahl der Variablenausprägungen in Klammern angegeben. Am besten beginnen Sie mit dem Tabelleninhalt, z.B. mit Wege, ggf. wählen Sie auch eine Größe für den Nenner, z.B. je Person. Dann legen Sie die Zeileninhalte fest (z.B. Haushaltseinkommen), und die Spalteninhalte (z.B. HanzPkw –Anzahl Pkw im Haushalt). Wenn Sie eine Kurzbezeichnung anklicken, wird oberhalb des Tabellenrechtecks ein erläuternder Text zu dieser Variablen gezeigt. Die Variablen sind nach Basis-Variablen der regionalen und zeitlichen (Stichtag) Gliederung, nach Haushalts-, Personen-, Wege- und Ziehungs-Variablen geordnet. Nutzen Sie parallel zum Tabellen-Auswahl-Fenster die Suchfunktion der Hilfe, wenn Sie mit den Variablen nicht vertraut sind: Öffnen Sie die Hilfe, unter Suchen geben Sie die Anfangsbuchstaben des gesuchten Begriffs (z.B. Fahrrad) ein und sehen die Kurzbezeichnungen der entsprechenden Variablen.

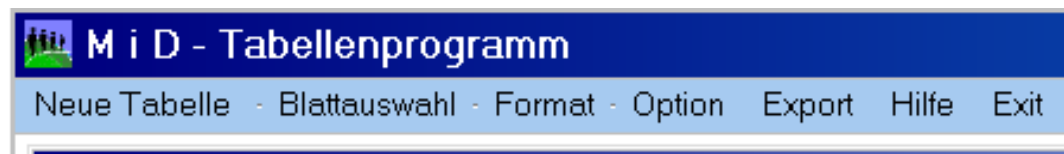
Rechts neben dem Tabellenrechteck können die möglichen Ergebnistabellen gewählt werden. Die ungewichteten Fallzahlen und die relativen Standardfehler können zur Beurteilung der statistischen Verlässlichkeit der Ergebnisse dienen. Wenn Sie den Standardfehler auswählen, steigt die Rechenzeit mit der Zahl der Ergebnisfelder, d.h. Sie müssen ggf. länger auf die Tabellenausgabe warten und sollten das Häkchen nur setzen, wenn Sie die Werte wirklich benötigen.

Wenn Sie mit der Tabellen-Zusammenstellung fertig sind, drücken Sie die OK-Taste des Tabellenauswahl-Dialogs. Sie erhalten nun das Ergebnis in einem neuen Fenster.

Das Tabellen-Auswahl-Fenster erhalten Sie jederzeit, wenn Sie in der Kommandozeile des Programmfensters „**Neue Tabelle**“ wählen.

Die erzeugten Tabellenblätter können Sie verschieben oder schließen, leider nicht in der Größe ändern. Wenn der Überblick verloren gegangen ist können Sie das gewünschte Tabellenblatt mit dem Menü **Blattauswahl** in den Vordergrund holen

Weitere Menüpunkte sind:



Format

Sie können das Aussehen der Ergebnis-Tabelle verändern. Sie können die Zahl der Nachkomma-Stellen im Bereich von 0 bis 3 erhöhen oder verringern, oder die senkrechten Striche ausblenden. Dies ist notwendig, wenn Sie die Tabelle, z.B. in Excel weiterverwenden wollen. Dazu markieren Sie ausgewählte Daten mit gedrückter linker Maustaste oder markieren alle Daten, indem sie auf den Rand des Tabellenfensters klicken, bzw. die Tastenkombination Strg-a drücken. Dann kopieren Sie mit Strg-c alles in die Zwischenablage, wechseln in das Zielprogramm und fügen dort aus der Zwischenablage ein.

Option

Über diesen Menüeintrag können Sie die Sprache wechseln. Wenn Sie Englisch wählen, wird statt des Kommas ein Dezimalpunkt ausgegeben.

Export

Um die Ergebnisse weiterzuverwenden, empfiehlt sich der Weg über die Zwischenablage in ein Tabellenkalkulationsprogramm. Sie können Tabellen(teile) mit der Maus markieren, in die Zwischenablage kopieren (Strg-c); in ihr Zielprogramm wechseln (Quattro Pro, Exel, OpenOffice „,) und dort an der gewünschten Stelle einfügen. Eventuell muss noch die erste Spalte in der Breite angepasst werden.

Sie können auch den Export-Menüeintrag **print.rtf** wählen. Dann wird in dem Verzeichnis, in dem sich MiT befindet eine Datei print.rtf mit den Ergebnissen angelegt. Diese können Sie mit dem Textverarbeitungs-Programm ihrer Wahl öffnen und dann drucken. Es empfiehlt sich, die Vorschau zu nutzen. Bei Tabellen mit bis zu vier Datenspalten ist die Papier-Voreinstellung A4 Hochformat. Ab fünf Spalten ist DIN A3 Quer voreingestellt, um breite Tabellen zu ermöglichen. Prüfen Sie über die Druckvorschau und den Seite-einrichten-Dialog, ob Sie auch anderes Papierformat für Ihre Tabelle reicht. Speichern Sie das Ergebnis unter einem anderen Namen, da MiT immer wieder print.rtf überschreibt, wenn Ausgabe gewählt wird.

Hilfe

Über die Hilfe-Funktion können Sie u.a. nach Inhalten der Erhebung Mobilität in Deutschland suchen und die zugehörige Variablen-Kurzbezeichnung finden.

Mit **Exit** beenden Sie das Programm.

Im Ausgabefenster werden die gewünschten Tabellen angezeigt und können mit einem Mausklick oder über das Menüfeld **Blattauswahl** in den Vordergrund geholt werden. Auch ein Verschieben oder schließen einzelner Tabellenblattfenster ist möglich, leider keine Größenänderung

3 Nützliche Hinweise

- Für die Spalten sollten Sie Variable mit nicht zu vielen Ausprägungen wählen. Manche Variablen (z.B. Alter) stehen deshalb nur als Block- und Zeilenvariable zur Verfügung, aber nicht im Spaltenauswahl-Menü.
- Wenn das Tabellen-Auswahlfenster im Vordergrund ist, ist das Hauptmenü blockiert. Verlassen Sie zuerst den Haupt-Dialog mit OK oder Abbruch.
- Die Berechnung der Standardfehler ist rechenintensiv. Die Rechenzeit steigt linear mit der Zahl der Tabellenfelder. Für die übrigen Zählungen werden im wesentlichen lediglich die Haushalts-, Personen- bzw. Wegedaten nur einmal durchlaufen und auf die richtigen Tabellenfelder verteilt, die Berechnungen sind weitgehend unabhängig vom Umfang der Tabellen. Bei langsamen Rechnern verzichten Sie, wenn möglich, auf die Berechnung der Standardfehler.
- Personenvariable sind personenbezogene Informationen. Wenn Sie z.B. PrbwBranche als Zeilenvariable wählen und Personen als Inhalt, erhalten Sie die Bevölkerung, aufgeteilt nach denen, die regelmäßige berufliche Wege berichtet haben, nach Branchen unterteilt. Wenn Sie als Inhalt Wege wählen, erhalten Sie *alle Wege* dieser Gruppen, nicht nur deren als rbw-Wege berichtete Verkehrsteilnahme. Um die rbw-

Wege nach Branchen zu erhalten, müssen Sie als weitere Brechung (z.B. Spaltenvariable) eine Unterteilung der *Wege* nach Zwecken wählen z.B. *WzwGesDdiff*.

3.1 faq - Tips

- 1) Mein Professor verlangt, dass die Tabellen in meiner Arbeit in der Schrift Arial Narrow 14 Pkte sind, weil er sich keine Lesebrille zulegen will. Muss ich alle Zahlen abtippen?
Nein, nutzen Sie den Menüknopf Export / print.rtf. Die Tabelle, die Sie dann als print.rtf-Datei erhalten, können Sie in eine „echte“ Tabelle ihres Textverarbeitungsprogramms umwandeln, weil Tab-Zeichen als Spaltentrenner eingearbeitet sind. In Word2000 geht es so: Tabelle ohne Striche erzeugen; dann print.rtf mit Word öffnen, die Tabelle (ggf. ohne Fusszeilen) markieren, Menü Tabelle / Umwandeln wählen und Tabstopps markieren. Enter drücken und Tabelle / Autoanpassen Inhalt zur Spaltenformatierung wählen. Die Tabelle noch mal markieren und Sie könne auch beliebige nichtproportionale Schriften wählen, die Spaltenbreiten verändern usf.
- 2) Ich brauche speziell Tabellen zum Thema Fahrrad.
Über Menüeintrag Hilfe die Hilfe starten und Fahrrad suchen lassen, oder die letzte Seite dieser Anleitung neben den Rechner legen.
- 3) Warum erscheinen alle Wege, auch wenn ich nur den Einkaufsverkehr sehen will, und *W ZwEinkaufdiff* als Variable gewählt habe?
Als erste oder letzte Zeile/Spalte wird bei allen Variablen mit den „nicht dazugehörigen Werten“ auf die Gesamtzahl ergänzt. Zum einen können Sie dann den hochgerechneten Eckwert mit obiger Tabelle vergleichen, zum anderen ist aus technischen Gründen (Nachgewichtung) immer der Gesamtbezug notwendig. Wenn Sie z.B. Anteilswerte der Unterzwecke im Einkaufsverkehr benötigen, kopieren Sie die Tabelle mit den hochgerechneten Werten einfach in ein Spreadsheetprogramm und rechnen dort ohne diese ergänzte Reihe weiter.

4 Methodik

4.1 Hochrechnung

Ausgewiesen werden hochgerechnete Werte. Die Hochrechnung setzt auf den Haushalts- und Personengewichtungsfaktoren auf, die dafür sorgen, dass die Stichprobe in der Verteilung wichtiger Strukturen (Haushaltsgröße, Alter und Geschlecht der Einwohner) mit den Verteilungen in der Grundgesamtheit übereinstimmt. Diese Gewichte sind auch in den Datensätzen enthalten, die für Auswertungen ohne das Tabellen-Tool ausgeliefert werden. Für die Erzielung korrekter Ergebnisse ist die Verwendung dieser Gewichtungsfaktoren unerlässlich.

Für die Ableitung von hochgerechneten Niveausergebnissen (z.B. Zahl der Haushalte mit Internetanschluss, Zahl der Einwohner mit Pkw-Fahrerlaubnis, Personenkilometer in Deutschland) sind darüber hinaus Faktoren notwendig, mithilfe derer die Stichprobenergebnisse auf die Grundgesamtheit hochgerechnet werden. Dies ist kein triviales Problem, da die Grundgesamtheit der Stichprobe zwar theoretisch eindeutig definiert durch die deutschsprachigen Einwohner ist. Forschungspraktisch jedoch gibt es angesichts der inhaltlichen und methodischen Komplexität der Erhebung (CATI/PAPI, angepasste Erhebungsinstrumente für Kinder und Erwachsene, Proxyangaben) wechselnde Hochrechnungszielgrößen, die die Grundgesamtheit nicht immer komplett umfassen. Dies

führt dazu, dass es keinen einheitlichen Haushalts- und Personenhochrechnungsfaktor (vergleichbar mit dem o.a. Gewichtungsfaktor) in den ausgelieferten Daten geben kann. Die eigene Ableitung von Niveauwerten erfordert für den Datennutzer gute Kenntnisse der Anlage der Erhebung und Sorgfalt bei der Ableitung der jeweiligen Hochrechnungszielgröße und der Bildung eines entsprechenden Hochrechnungsfaktors.

Im Tabellentool werden für jede gewählte Tabelle abhängig vom Inhalt (Haushalte- Personen- oder Wegeinformationen) Hochrechnungsfaktoren bestimmt und verwendet. Die Hochrechnung erfolgt zweistufig. Im ersten Schritt sind, basierend auf den Gewichtungsfaktoren für Haushalte und Personen, Hochrechnungsfaktoren für Haushalte, Personen, Wege und regelmäßige berufliche Wege zu bestimmen. Diese generelle Hochrechnung wird

- bei Haushalten nach Bundesland und Haushaltsgröße (16 x 5 Kategorien),
- bei Personen und bei normal berichteten Wegen nach Bundesland, Geschlecht und Altersklassen (16 x 2 x 8 Kategorien)

vorgenommen. Die Hochrechnungsfaktoren für die Wege unterscheiden sich von den Personen-Hochrechnungsfaktoren dadurch, dass die Stichprobe (der Nenner der Faktorenberechnung) nicht aus allen antwortenden Personen besteht, sondern nur aus denen, zu denen Auskünfte zu ihrem Mobilitätsverhalten am Stichtag (ob selbst oder als proxy berichtet) vorliegen. Da die Gewichtungsfaktoren auch einen saisonalen Ausgleich umfassen, sind die hochgerechneten Wegezahlen als die Wege der Wohnbevölkerung an einem mittleren Stichtag zu interpretieren.

Um ein vollständiges Bild des Verkehrs am Stichtag auszuweisen, sind auch die Antworten zu den regelmäßigen beruflichen Wegen zu berücksichtigen. Es sind aus diesen Angaben pauschalisierte Wegesätze generiert worden. Stichprobenumfang sind hier alle Auszubildenden und Erwerbstätigen, die nach solchen Wegen gefragt worden sind. Dies ist wiederum nur eine Teilmenge aller erhobenen Auszubildenden und Erwerbstätigen (z.B. ist bei Proxy-Interviews nicht nach den regelmäßigen Wegen gefragt worden).

Im zweiten Schritt der Hochrechnung werden zusätzlich die jeweiligen Block-, Zeilen-, und Spaltenvariablen der gewählten Tabelle berücksichtigt. Die Berechnung der Hochrechnungsfaktoren des zweiten Schritts erfolgt in Anlehnung an die Struktur der generellen Hochrechnung, bei Haushalten nach Bundesland und Haushaltsgröße und bei Personen nach Bundesland, Geschlecht und Altersklasse. Bei den Wegedaten wird in diesem Berechnungsschritt auch berücksichtigt, dass bei Kilometer- und Zeitangaben zusätzliche fehlende Angaben auftreten können: Die Faktoren werden iterativ aus den sich mit fehlenden Werten für die Hochrechnungskategorien berechneten Größen w_{lvz} , bezogen auf die sich ohne fehlende Werte ergebenden Zielgrößen W_{lvz} bestimmt (l, v, z bezeichnen Land, Verkehrsmittel und Zweck). Letztere sind das Ergebnis der mit den Hochrechnungsfaktoren der ersten Stufe gewichteten Wege aller befragten Personen mit Antworten zur Mobilität. Bei Haushalts- und Personen- Hochrechnung wird entsprechend verfahren, die Zielgrößen H_{lg} (l bezeichnet das Bundesland, g die Haushaltsgrößenklasse) bzw. P_{lsa} (mit l für das Land, s für das Geschlecht der Person, und a für die Altersklasse) sind wiederum diejenigen aus dem ersten Hochrechnungsschritt.

4.2 Varianzschätzung und Standardfehler

Neben der Gewichtung, die erst sinnvolle Auswertungen der Befragung ermöglicht, gehört die Schätzung der Verlässlichkeit der so berechenbaren Populationswerte zur methodischen Aufbereitung der Erhebung. Jede Stichprobenerhebung ist mit statistischen Stichprobenfehlern behaftet. Die Berechnung dieses Fehlers ist bei komplexen Erhebungen

nur näherungsweise möglich. Die *MiD*-Erhebung hat aufgrund des hohen Stichprobenumfangs und der hohen methodischen Qualität niedrige Fehlerwerte. Dennoch ist in das Tabellentool die im Folgenden geschilderte Schätzmethode aufgenommen worden.

Der Standardfehler schätzt die Variationsbreite, die für einen Tabellenwert zu erwarten wäre, wenn nicht eine, sondern viele gleiche bundesweite *MiD*-Erhebungen durchgeführt worden wären. Dieser Wert ist ein Schätzer für den statistisch zu erwartenden Stichprobenfehler, also ein Maß für die statistische Variationsbreite der ausgewiesenen Größen, der sich daraus ergibt, dass nicht die gesamte Bevölkerung befragt worden ist, sondern nur eine Stichprobe. Es gibt unterschiedliche Schätzverfahren für die Stichprobenfehler. Bei dem verwendeten komplexen Auswahldesign ist die analytische Bestimmung (die Berechnung über eine Formel) der Varianz des Totals eigentlich nicht möglich. Der hier verwendete Ansatz stellt aber eine gute Approximation nach dem gegenwärtigen wissenschaftlichen Stand dar.

Das gewählte Vorgehen nutzt die Informationen aus der Ziehung. Erster Schritt der Stichprobenbildung war eine geschichtete Ziehung von Gemeinden. Diese sind die primären Ziehungseinheiten (primary sampling units), kurz mit *psu* bezeichnet. Die Siedlungsstruktur hat erheblichen Einfluss auf das Verkehrsverhalten (ÖV-Angebot, Erreichbarkeit der Ziele). Die Differenzierungsmöglichkeiten beim Gemeindetyp werden aber durch den finanziell möglichen Stichprobenumfang begrenzt. In der Basisstichprobe sind 297 Gemeinden mit 340 *psu* enthalten.

Die Auswahl dieser *psu* erfolgte zufällig aus Schichten, die nach

- Bundesländern und
 - Regionstyp (Kombination aus BIK-Typ und BBR-Typisierung)
- gebildet worden sind (vgl. Kapitel 4). In den ausgewählten Gemeinden sind zufällig Personen ab 14 Jahren aus den Einwohnermeldekarteien gezogen worden, befragt wurden alle Personen des so bestimmten Haushalts.

Während die Ziehung aus Gemeinden als Klumpeneffekt im Vergleich zur reinen Zufallsauswahl erhöhend auf den Stichprobenfehler wirkt, wird durch die Schichtung der Gemeinden nach Regionstypen mit ähnlicher Verkehrsstruktur der Schätzfehler verringert. Im benutzten Verfahren werden beide Effekte berücksichtigt, die weitere Klumpung der Personen- und Wegeinformationen durch den Haushaltsbezug konnte jedoch nicht einbezogen werden.

Die Gesamtvarianz ergibt sich als Summe über die Varianzschätzungen für die einzelnen Schichten, da die Ziehungen unabhängig voneinander erfolgt sind. Die Varianz jeder Schicht wird aus der Varianz zwischen den *psu*-Werten in jeder Schicht (between) bestimmt, da unterstellt werden kann, dass dieser Term wesentlich größer ist als die Varianz der Werte in jeder *psu* (within).³ Bezeichnet man mit z_{yhi} die Summe der m_{hi} mit w_{hij} gewichteten Werte y_{hij} der *i*-ten *psu* (Gemeinde) der Schicht *h*

$$z_{yhi} = \sum_{j=1}^{m_{hi}} w_{hij} y_{hij} \quad (1)$$

³ Vgl. Rendtel, Ulrich; Schimpl-Neimanns, Bernhard (2000): Varianzschätzungen für den faktisch anonymisierten Mikrozensus. Jahrbücher für Nationalökonomie und Statistik, Vol. 220, S.759–776.

und mit $\bar{z}_{yh} = \frac{1}{n_h} \sum_{i=1}^{n_h} z_{yhi}$ den entsprechenden Mittelwert über die n_h verschiedenen psu der Schicht, wird die Varianz eines Populationstotals Y , d.h. $V(Y)$ nach der Formel

$$\hat{V}(\hat{Y}) = \sum_h \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} (z_{yhi} - \bar{z}_{yh})^2 \quad (2)$$

geschätzt.

Bei bezogenen Merkmalen $Q=Y/X$, z.B. Wegen je Person, ist die Kovarianz zwischen Y und X zu berücksichtigen. Zur Berechnung der Varianz wird eine Taylorentwicklung der Funktion benutzt (Delta-Methode). Der Varianz-Schätzer der Verteilung von \hat{Q} ist

$$\frac{1}{X^2} [V(\hat{Y}) - 2SCov(\hat{Y}, \hat{X}) + \hat{Q}^2 V(\hat{X})]$$

Definiert und benutzt man $d_{hij} = \frac{1}{\hat{X}} (y_{hij} - \hat{Q}x_{hij})$ anstelle von y_{hij} in (1), unterscheidet sich der Rechenablauf nur durch diese Transformation.

Aus der Varianzschätzung für das Total $\hat{V}(\hat{Y})$ bzw. für den Quotienten $\hat{V}(\hat{Q})$ werden die Standardfehler

$$\hat{s}_y = \sqrt{\hat{V}(\hat{Y})} \quad \text{bzw.} \quad \hat{s}_q = \sqrt{\hat{V}(\hat{Q})} \quad (3)$$

errechnet.

In den Tabellen des Tabellentools wird der relative Standardfehler ausgewiesen:

$$\hat{k} = \frac{\hat{s}}{\bar{z}} * 100 \quad (4)$$

Dieser Quotient von Standardfehler s und hochgerechnetem Wert hat gegenüber der Varianz bzw. der Standardabweichung den Vorteil, dass die Werte unmittelbar in der Größenordnung vergleichbar sind. Wenn die Standardabweichung benötigt wird, z.B. um Konfidenzintervalle zu schätzen, kann leicht über die Formeln (2) bis (4) zurückgerechnet werden.

5 Was nicht geht

Die zur Entwicklung benutzte Software ist eigentlich nicht auf Windows ausgerichtet. Die Zeichenzahl je Tabellenblatt ist auf 65000 Zeichen begrenzt, eine Größenänderung der Ausgabefenster ist nicht möglich.

Bewusst ist auf die Programmierung von Filterung und der Möglichkeit, Variablen neu zu definieren, verzichtet worden. Dies hätte die Bedienung weiter kompliziert.

Die ebenfalls in MiD erhobenen Daten zum Fernverkehr können mit dem Tool nicht ausgewertet werden, ebenso wenig die differenzierten Angaben zu den vorhandenen und genutzten Kraftfahrzeugen.

Kurzbez.	Anz.Auspr.	Langbezeichnung	Kurzbez.	Anz.Auspr.	Langbezeichnung	Kurzbez.	Anz.Auspr.	Langbezeichnung	Kurzbez.	Anz.Auspr.	Langbezeichnung
Basisvariable			Personenvariable			noch: Personenvariable			Verkehrsverhalten am Stichtag		
Räumlich			Personenvariable			Führerscheinbesitz			Zeit und Entfernung		
BBR Regtyp	3	BBR-Regionsgrundtyp	P Alter	98	Alter der Person	PMopedFs	2	Pers.mit Moped-Führersch.	W ab h	24	Wegebeginn (Stunde)
BBR Kreistyp	9	BBR-Kreistyp	P Altkl	9	Altersklasse	PMopedFsJahr	75	JahrFührerschein: Moped	W ab Min	60	Wegebeginn (Minute)
BBR diffRegtyp	7	differenzierter Regionstyp	PMobEinschr	11	Pers.m. Mob.einschränkt	PMoiRadFs	2	Pers.mit Mot.rad-Führersch.	W an h	24	Wegeende (Stunde)
BBR Gemtyp	17	Gemeindetyp	PHauptbesch	12	Hauptbeschäftigung derzeit	PMoiRadFsJahr	75	JahrFührerschein: MotRad	W an Min	60	Wegeende (Minute)
BIK-Typ	5	BIK-Strukturtyp	PLebZyklus	11	Person im Lebenszyklus	PPkwFs	2	Pers.mit Pkw-Führersch.	W Dauerkl	9	Wegedauer(Klassen) in Min
B Wohnumfeld	53	Wohnumfeld (LOCAL-Typ)	PSchulabschl	15	Schulabschluss	PPkwFsJahr	75	JahrFührerschein: Pkw	W Dauer	99	Wegedauer in Min
B Wohn-Status	8	Wohnumfeld(LOCAL-Status)	PSEX	2	Geschlecht der Person	PLkwFs	2	Pers.mit Lkw-Führersch.	W km	99	Wegeänge
B Land	16	Bundesland	PStaatsangD	2	Staatsangeh. Deutsch	PLkwFsJahr	75	JahrFührerschein:erwerb Lkw	W km Kl.	8	Wege-Entfernungsklasse
B WestOst	2	West- Ostdeutschland	Verkehrsmittelnutzung, -verfügbarkeit						W km/h	99	Geschw. Haus zu Haus kmh
			PBahnnutz	6	allg.Nutzg Bahn längere Str				W Geschwkl	7	Geschw.kl. Haus-Haus kmh
BWochentag	7	Wochentag	PBahnHstMin	11	Pers. n. Bahnhof-Entf. Min.						
B Tagtyp	4	Stichtag Ferien- Feiertag	PBahnHstMeter	8	Pers. n. Bahnhof-Entf.Meter	Pmobil	2	Person war mobil			
B Monat	12	Monat des Stichtags	PÖSPVnutz	6	Pers.n. allg. ÖSPV-Nutzung	PNormsTag	8	Pers.n. Normalität Stichtag	Verkehrsmittel		
B Tag	31	Tag im Monat (Stichtag)	PBusHstMeter	8	Fußwegentf. Bushaltest.Meter	PSTKz	4	Kfz-Verfügbarkeit am Stichtag	W AnzVM	6	Anz. gen.Verkehrsm. je Weg
B Woche	52	Kalenderwoche des Stichtags	PBusHstMin	10	Fußwegentf. Bushaltest.Min	PSTWeiter	6	Weiter am Stichtag	W Vmittel17	17	Haupt-Verkehrsmittel diff.
			Haushaltsvariable						W Vmittel6	6	Haupt-Verkehrsmittel (6)
H Groesse	5	Personenzahl im Haushalt	PTicketÖPNV	8	Überwiegende Fahrscheinart				W Vmittel3	3	Haupt-Verkehrsmittel (3)
H Lebensph	12	Haushalt - Lebensphase	PZieleÖPNV	6	Erreichbkt übl. Ziele:ÖPNV	regelm. berufl.Wege					
			PPkwnutz	6	Pers.n. allgem. Pkw-Nutzung	Prbw	2	Pers. mit regelm.berufl. W.			
			PPkwVrfügbar	5	Pkw-Verfügbarkeit allgemein	PrbwAnzWege	47	P.n.Anz. regelm.berufl. W.	Zweck		
H AnzKrad	6	AnzahlKrafträder im Haushalt.	PZielePkw	6	Erreichbkt übl. Ziele: Pkw	PrbwAnzWegekl	6	P.n.Anz. regelm.berufl. W.	W Zweck	7	Haupt-Wegezweck
H AnzLkw	6	AnzahlLkw im Haushalt	PRad	2	Pers. nach Besitz Fahrrad	PrbwBranche	18	P.n. Branche (rbWege)	W Zweckdiff	39	Zweck/Ziel differenziert
H AnzLkwgr	6	AnzahlLkw ab 3.5 t Nutzlast	PRadnutz	6	allgemeine Nutzung:Fahrrad	PrbwKilometer	99	P.n. Entf. (rbWege)	W ZwAusbdiff	3	Wegezweck Ausbildg diff.
H AnzLkwkl	6	AnzahlLkw bis 3.5t Nutzlast	PRadpark	7	Fahrradunterbringung	PrbwKilometerKl	7	P.n. Entf.kl (rbWege)	W ZwBegldiff	4	Wegezw. Begleitg diff.
H AnzPkw	6	AnzahlPkw im Haushalt	PRadzugang	7	Fahrradzugang	PrbwVerkehrsmitt	15	P.n. Verkehrsm. (rbWege)	W ZwBerufdiff	3	Wegezweck Beruf differenz.
H AnzRaeder	6	AnzahlFahrräder im Haushalt				PrbwZweck	5	P.n. Zweck (rbWege)	W ZwEinkaufdiff	6	Wegezw. Einkauf differenz.
			sonstige Ausstattung						W ZwErleddiff	16	Wegezw. Erledigg differenz.
H Wohnlage	4	Wohnlage							W ZwFreiztdiff	22	Wegezw. Freizeit differenz.
H Eink2	2	Haush.-Nettoeink. (2 Kl.)							W ZwGesDdiff	3	Wirtschaftsverk. differenz.
H Eink8	8	Haush.-Nettoeink. (8 Kl.)							Zielungs-Zusatzinformationen		
HhatComputer	2	Haushalt mit Computer							ZHcatipapi	2	Haush.rekrut. tel./schriftl
HhatHandy	2	Haushalt mit Handy							ZPcatipapi	2	Pers.befragg tel./schriftl.
HhatInternet	2	Haushalt mit internet-Ansch							ZP Alter		- Alter Person(Einw.meldeamt)
HhatTelefon	2	Haushalt mit Telefon							ZP Sex	3	m/w Person(Einw.meldeamt)