



Mobilität in Deutschland – MiD

# Nutzerhandbuch

Ausgabe Januar 2019

eine Studie des:



Bundesministerium  
für Verkehr und  
digitale Infrastruktur

durchgeführt von:

**infas**

In Kooperation mit:



DLR



Research  
Daten  
Analysen  
Strategien

**infas 360**





**Ergebnisbericht für/eine Studie des:**

Bundesministerium für Verkehr und digitale Infrastruktur  
Referat G 13 – Prognosen, Statistik und Sondererhebungen  
Robert-Schuman-Platz 1  
53175 Bonn  
FE-Projektnummer 70.904/15



**vorgelegt von:**

infas Institut für angewandte  
Sozialwissenschaft GmbH  
Friedrich-Wilhelm-Straße 18  
53113 Bonn



**in Kooperation mit:**

Deutsches Zentrum für Luft- und Raumfahrt e. v.  
Institut für Verkehrsforschung  
Rutherfordstraße 2  
12489 Berlin



IVT Research GmbH  
M 4, 10  
68161 Mannheim



infas 360 GmbH  
Ollenhauerstraße 1  
53113 Bonn

**Projekt:**

5431  
Bonn, Januar 2019  
Version 1.0

Der Inhalt dieses Berichts darf ganz oder teilweise nur mit unserer schriftlichen Genehmigung veröffentlicht, vervielfältigt, gedruckt oder in Informations- und Dokumentationssystemen (information storage and retrieval systems) gespeichert, verarbeitet oder ausgegeben werden

# Inhaltsverzeichnis

<b>1</b>	<b>Aufgabenstellung .....</b>	<b>9</b>
<b>2</b>	<b>Übersicht zu den Begleitdokumenten des Nutzerhandbuchs.....</b>	<b>11</b>
<b>3</b>	<b>Studiendesign und Begriffsdefinition .....</b>	<b>13</b>
3.1	Konzeption und Inhalte der MiD.....	13
3.2	Definition zentraler Begriffe.....	15
<b>4</b>	<b>Datenaufbereitung.....</b>	<b>21</b>
4.1	Formale Datenaufbereitung .....	21
4.2	Weiterführende Datenaufbereitung.....	22
<b>5</b>	<b>Das Datenangebot der MiD 2017.....</b>	<b>25</b>
5.1	Struktur der Daten .....	25
5.2	Datensatzpakete.....	28
5.3	Bereitstellung der Datensatzpakete .....	30
<b>6</b>	<b>Hinweise zur Datennutzung .....</b>	<b>31</b>
6.1	Gewichtung und Hochrechnung.....	31
6.2	Ermittlung der korrekten Tagesstrecke.....	33
6.3	Stellvertreterinterviews, Missings und Module .....	34
6.4	Die Variable Mobilität am Stichtag .....	36
6.5	Vergleich mit Vorgängererhebungen .....	37



# Tabellenverzeichnis

<b>Tabelle 1</b>	Fallzahl der sechs Teildatensätze der MiD 2017 .....	26
<b>Tabelle 2</b>	Kodierung antwortbedingt fehlender Werte .....	27
<b>Tabelle 3</b>	Kodierung designbedingt fehlender Werte .....	28
<b>Tabelle 4</b>	Namen und erzielte Fallzahl von Gewichtungs- und Hochrechnungsvariable je Teildatensatz der MiD .....	32
<b>Tabelle 5</b>	Häufigkeitsverteilung der Pkw-Verfügbarkeit als Fahrer oder Mitfahrer .....	33
<b>Tabelle 6</b>	Pkw-Verfügbarkeit als Fahrer oder Mitfahrer hochgerechnet auf die Bevölkerung ab 14 Jahren .....	33
<b>Tabelle 7</b>	Häufigkeitsverteilung bei der Frage nach Online-Shopping .....	34
<b>Tabelle 8</b>	Häufigkeitsverteilung bei der Frage nach Home Office .....	35
<b>Tabelle 9</b>	Häufigkeitsverteilung der differenzierten Variable Mobilität am Stichtag .....	36
<b>Tabelle 10</b>	Häufigkeitsverteilung der einfachen Variable Mobilität am Stichtag .....	37



# Abbildungsverzeichnis

Abbildung 1	Übersicht der Fragebogeninhalte der MiD 2017 .....	14
Abbildung 2	Bestimmung des ökonomischen Status über Haushaltsnettoeinkommen und gewichtete Haushaltsgröße .....	18
Abbildung 3	Regionalstatistische Raumtypologie (RegioStaR) des BMVI.....	19
Abbildung 4	Zusammengefasster Regionalstatistische Raumtypologie (RegioStaR 7).....	20
Abbildung 5	Datensatzstruktur: Teildatensätze und Schlüsselvariablen .....	25
Abbildung 6	Übersicht der MiD-Datensatzpakete.....	28
Abbildung 7	Zugelassener Nutzerkreis der MiD-Datensatzpakete.....	30





# 1 Aufgabenstellung

Mit der aktuellen Erhebung Mobilität in Deutschland 2017 (MiD 2017) steht Wissenschaft, Planung und Politik erneut eine umfangreiche Datengrundlage für die Analyse der Alltagsmobilität in Deutschland zur Verfügung. Die Studie ist zu weiten Teilen eine Replikation der Vorgängerstudie. Gegenwärtige Entwicklungen im Verkehr und geänderte Rahmenbedingungen für Erhebungen haben aber auch Anpassungen erfordert. Über die Aufnahme von Modulfragen, die jeweils nur einem Teil der Befragten gestellt wurden, konnte das Fragenprogramm gegenüber den vorherigen MiD-Erhebungen erweitert werden. Um der abnehmenden Teilnahmebereitschaft an Erhebungen entgegenzuwirken, wurde die Methodenvielfalt (Online-, Telefoninterview, schriftlicher Fragebogen) auf alle Phasen der Erhebung ausgeweitet. Darüber hinaus enthalten die über die Clearingstelle Verkehr des Deutschen Zentrums für Luft- und Raumfahrt erhältlichen MiD-Daten nicht nur die Fälle aus der Bundesstichprobe, sondern auch die Fälle von über 60 regionalen Auftraggebern, die sich der Bundeserhebung angeschlossen haben (Aufstocker).

Die MiD weist aufgrund ihrer thematischen Breite (Erfassung der Stichtagsmobilität sowie von regelmäßigen beruflichen Wegen und Reisen) und ihres Erhebungsdesigns (unterschiedliche Erhebungsformen, die Möglichkeit von Stellvertreterinterviews) eine nicht zu unterschätzende Komplexität auf. Ziel des vorliegenden Nutzerhandbuchs mit seiner umfassenden Dokumentation im Anhang ist die Bereitstellung von grundlegenden Informationen zu Studiendesign, Datenaufbereitung und zentralen Begriffen sowie zum praktischen Umgang mit den Daten.

Generell können vier Zugänge zu den Daten der MiD unterschieden werden:

**Ergebnisbericht:** Die zentralen Ergebnisse der MiD sind im Ergebnisbericht zusammengefasst. Der Bericht konzentriert sich auf die Darstellung der Alltagsmobilität in Deutschland im Jahr 2017. Im Sommer 2019 folgt ein weiterer Bericht mit Fokus auf die Entwicklung der Mobilität im Zeitverlauf. Für diesen Bericht wird zunächst ein konsistenter Zeitreihendatensatz hergestellt.

**Tabellenband:** Der Tabellenband enthält eine Grundauswertung der MiD-Daten in tabellarischer Form. Die Tabellen können als Nachschlagewerk oder für den Abgleich eigener Ergebnisse genutzt werden.

**Tabellierungswerkzeug Mobilität in Tabellen (MiT):** Mit dem Tabellierungswerkzeug MiT können ohne umfangreiche Kenntnis von Statistikprogrammen eigene Auswertungen mit den Daten vorgenommen werden. Im Gegensatz zu den Vorgängererhebungen bedarf es dafür keiner Registrierung. Das Tool steht online unter [www.mobilitaet-in-tabellen.de](http://www.mobilitaet-in-tabellen.de) zur Verfügung. Gegenwärtig enthält das Tool nur die Daten der MiD 2017.

**Daten der MiD:** Die Daten der MiD sind in verschiedenen Teildatensätzen abgelegt. Diese stehen den Nutzern in vier verschiedenen Varianten (Datensatzpaketen) zur Verfügung (siehe Kapitel 5.2). Das Standard-Datensatzpaket entspricht dem Public Use File der MiD 2008. Der Datensatz enthält detaillierte soziodemografische und -ökonomische Angaben zu Personen und Haushalten. Der konkrete Raumbezug geht nicht über das Bundesland hinaus. Darüber hinaus werden zwei Datensatzpakete mit detaillierten Rauminformationen angeboten. Das Regional-Datensatzpaket enthält die Gemeindeverbandskennziffer und erlaubt damit räumliche Analysen auf der Ebene von Gemeindeverbänden. Das Lokal-Datensatzpaket stellt für den Startpunkt und das Ziel von Wegen sowie den Wohnort der Befragten die Angabe zur geografischen Gitterzelle bereit. Damit sind Analysen unterhalb der Gemeindeebene möglich. Aus Datenschutzgründen wurde bei beiden Datensatzpaketen der Variablenbestand reduziert und die soziodemografischen Angaben vergrößert. Schließlich gibt es ein absolut anonymisiertes Datensatzpaket, bei dem sowohl die soziodemografischen als auch die räumlichen Merkmale vergrößert und der Variablenbestand reduziert wurden.

Das Nutzerhandbuch widmet sich primär dem Standard-Datensatzpaket der Studie. Zielgruppe ist ein breites Spektrum an Nutzerinnen und Nutzern mit unterschiedlichem Wissenshintergrund. Die Kapitel bzw. Begleitdokumente des Nutzerhandbuchs sind nicht für jeden gleichermaßen von Bedeutung.

Das Nutzerhandbuch ist vielmehr so angelegt, dass es parallel zur Arbeit und in Bezug auf die jeweils aktuell anstehende Fragestellung selektiv gelesen werden kann.

## 2 Übersicht zu den Begleitdokumenten des Nutzerhandbuchs

Das vorliegende Nutzerhandbuch wird durch folgende Begleitdokumente ergänzt:

**Fragenübersicht:** Mit diesem Dokument kann ein schneller Überblick über die Inhalte des Haushaltsinterviews und des Personen-Wege-Interviews gewonnen werden. Es stellt in übersichtlicher Form die folgenden Informationen bereit: den Originallaut jeder Frage, die Antwortkategorien, den Namen der Variable, in der die Angaben abgelegt wurden, sowie den Hinweis, ob es sich um eine Frage aus dem Kernfragebogen oder einem Zusatzmodul handelt.

**Codepläne:** Für alle Teildatensätze stehen detaillierte Codepläne zur Verfügung. Darin enthalten sind folgende Angaben: Variablenname und -label, Werte und Wertelabel, Positionsnummer, Messniveau, Variablentyp, Anfangs- und Endspalte.

**Erhebungsunterlagen:** Hierbei handelt es sich um die Unterlagen, die an die teilnehmenden Haushalte der MiD 2017 verschickt wurden: Anschreiben, Informationsflyer, Datenschutzerklärung sowie ein Wegeblatt für Erwachsene und eines für Kinder, das den teilnehmenden Personen als Gedächtnisstütze für die Wege am Berichtstag dienen sollte.

**Schriftlicher Fragebogen und CAWI- und CATI- Fragebogenmaster:** Bereitgestellt werden sowohl die schriftlichen Fragebogen des Haushalts- und Personen-Wege-Interviews als auch die beiden Fragebogenmaster, die Grundlage für die Programmierung des Online- (CAWI) und des Telefoninterviews (CATI) waren. Diesen Dokumenten kann der Ablauf der Interviews und der originale Wortlaut der Fragen entnommen werden.

**Variablenaufbereitung:** Neben einer formalen Aufbereitung der Daten wurde eine inhaltliche Bereinigung und Plausibilisierung vorgenommen. Für jeden Teildatensatz wird variablenweise dokumentiert, welche Bereinigungs- und Aufbereitungsschritte erfolgt sind. Darüber hinaus wird erläutert, wie analytische, aus mehreren Originärantworten abgeleitete Variablen, gebildet wurden.



## 3 Studiendesign und Begriffsdefinition

### 3.1 Konzeption und Inhalte der MiD

Eine umfassende Darstellung von Erhebungskonzept und Feldverlauf ist dem Methodenbericht zu entnehmen. An dieser Stelle werden nur die wichtigsten Eckpunkte der Erhebung zusammengefasst.

Die Erhebung Mobilität in Deutschland 2017 bestand aus zwei aufeinanderfolgenden Phasen:

- In der ersten Phase wurde eine Haushaltsbefragung durchgeführt, bei der die Haushaltsgröße, Angaben zu den einzelnen Haushaltsmitgliedern und den im Haushalt vorhandenen Verkehrsmitteln sowie einige weitere Merkmale erhoben wurden. Diese Angaben wurden stellvertretend von einem Haushaltsmitglied für alle Haushaltsmitglieder gemacht.
- In einer zweiten Phase wurden alle Haushaltsmitglieder im Rahmen eines Personen-Wege-Interviews einzeln nach persönlichen Merkmalen, nach ihren Wegen an einem fest vorgegebenen Stichtag und nach ihren Reisen mit Übernachtung in den letzten drei Monaten befragt.

Bei der Erhebung gab es keine Alterseinschränkung. Bis zum Alter von neun Jahren wurden die Eltern gebeten, stellvertretend für ihre Kinder zu antworten. Im Alter von zehn bis 13 Jahren konnten entweder die Kinder selbst oder ihre Eltern antworten. Ab 14 Jahren sollten die Personen selbst an der Befragung teilnehmen. War dies nicht möglich, konnte das Interview auch in diesem Fall stellvertretend von einem anderen Haushaltsmitglied geführt werden.

Die Feldarbeit für die MiD 2017 erstreckte sich kontinuierlich über mehr als zwölf Monate. Nach der Stichprobenziehung wurde für jede Haushaltsadresse ein konkreter Stichtag für die Erhebung der Wege festgesetzt, der im weiteren Verlauf den sukzessiven Einsatz einer Adresse für den Versand bzw. die Vorlage im Telefonstudio bestimmte. Durch die Trennung in die beiden oben beschriebenen Phasen und die sich dadurch ergebenden Laufzeiten für Ankündigungen

bzw. Wiedervorlagen erstreckte sich die Feldzeit von Ende Mai 2016 bis Anfang Oktober 2017.

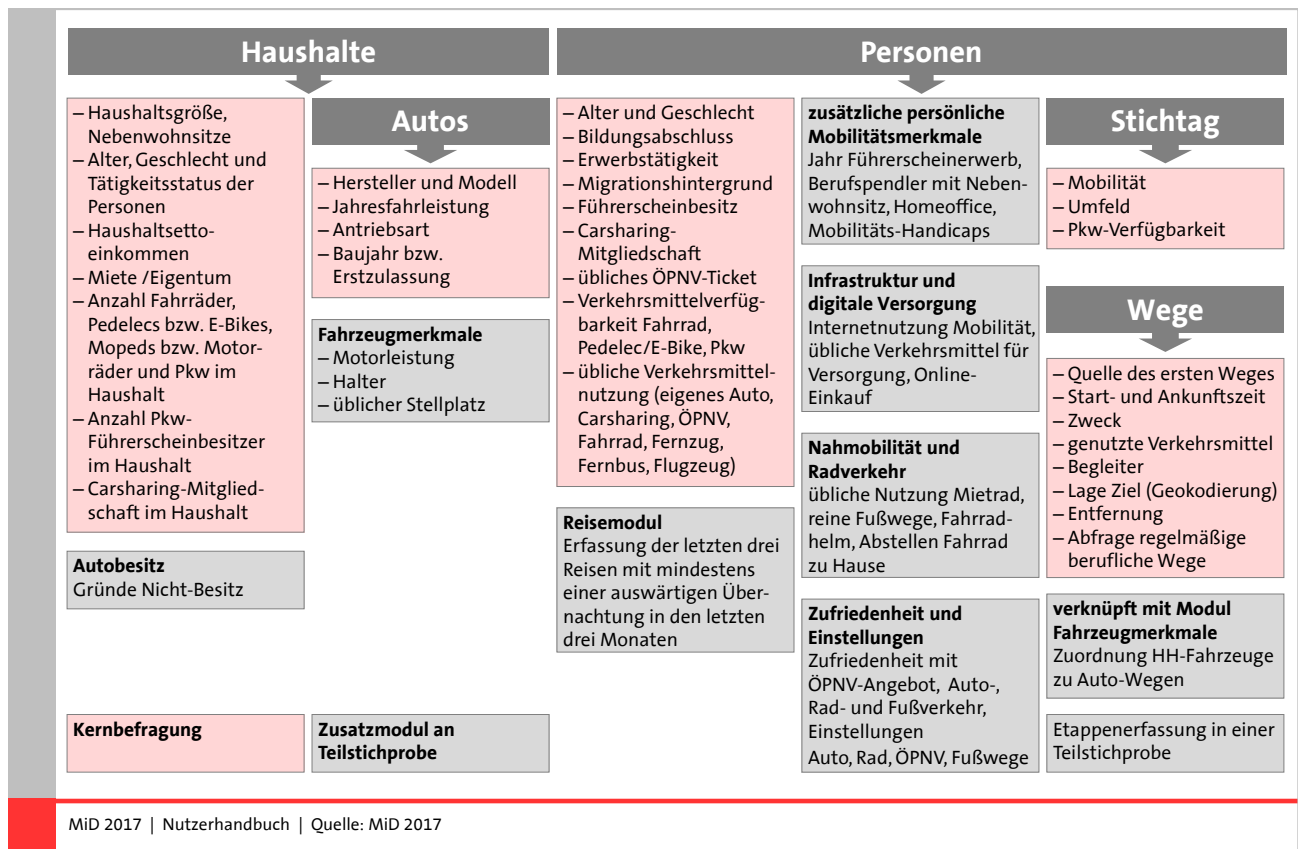
Der so realisierte Datensatz der MiD-Stichprobe umfasst die Angaben von 316.361 Personen aus 156.420 Haushalten. Im Rahmen der MiD wurden fast eine Million Wege erhoben. Von der MiD-Stichprobe entfallen 33.389 Haushalte auf die vom Bundesministerium für Verkehr und digitale Infrastruktur in Auftrag gegebene Bundesstichprobe und 123.031 Haushalte auf regionale Vertiefungsstichproben von über 60 Auftraggebern, die sich der Basiserhebung angeschlossen haben. Im Gegensatz zur Erhebung aus dem Jahr 2008 sind in den Datensätzen der MiD 2017 auch die Fälle der regionalen Vertiefungsstichproben enthalten.

In der ersten MiD 2002 wurde auf Basis einer Methodenstudie der Wechsel von einem rein postalisch-schriftlichen Verfahren zu einer schriftlichen und telefonischen Erhebung vollzogen. Für die darauf folgende MiD 2008 wurde auf das postalisch-schriftliche Verfahren in der Wegeerfassung verzichtet. In der Erhebung 2017 wurden in beiden Erhebungsstufen (Haushalte und Personen- bzw. Wegeerfassung) alle drei Methoden eingesetzt: telefonische, schriftliche und Online-Befragung. Dieses schon in der Ausschreibung festgelegte Verfahren sollte den Befragten möglichst alle Optionen nahe an ihrer individuellen Präferenz bieten und in der Kombination zu einer guten Abdeckung der Grundgesamtheit führen.

Die Befragungsinhalte der beiden Interviews sind *Abbildung 1* zu entnehmen. Insgesamt können sechs Ebenen der Datenauswertung unterschieden werden: Haushalte, Personen, Wege, Autos, Reisen und Etappen.

Das Kernstück der Erhebung bildet die Erfassung der Wege am Stichtag. Im Rahmen des Personen-Wegeinterviews wurden detaillierte Angaben von bis zu zwölf Wegen (Telefon- und Online-Interview) bzw. bis zu acht Wegen (schriftliches Interview) erfasst. Dazu gehören der Wegezweck, die Start- und Zieladresse des Weges, die genutzten Verkehrsmittel, Angaben zum

Abbildung 1 Übersicht der Fragebogeninhalte der MiD 2017



Fahrer bzw. zur Fahrerin bei Wegen mit Verkehrsmitteln des motorisierten Individualverkehrs, die Anzahl der begleitenden Personen und die Entfernung des zurückgelegten Weges. Von Wegen, die über diese zwölf bzw. acht detailliert zu erfassenden Wege hinausgehen, wurde die Anzahl notiert.

Die Angaben zur Stichtagsmobilität bieten ein umfassendes Bild der Mobilität in Deutschland. Von dem sequenziellen Ausschnitt eines einzelnen Tages kann jedoch nicht auf das generelle Mobilitätsverhalten der befragten Person geschlossen werden. Aus diesem Grund wurde neben den Wegen an einem fest vorgegebenen Stichtag auch die allgemeine Nutzungshäufigkeit der Verkehrsmittel erhoben. Über eine fünfstufige Skala konnten die Befragten angeben, wie häufig sie das Auto, das Fahrrad, Bus oder Bahn in der Region, die Bahn auf längeren Strecken (ab 100 Kilometer einfache Entfernung), das Flugzeug sowie neuere Mobilitätsangebote wie Fernbusse, Car- und Bikesharing nutzen.

Um auch längere Reisen in einem ausreichenden Umfang zu erfassen, wurden diese nicht nur in der Stichtagsabfrage, sondern auch rückblickend für die letzten drei Monate vor dem Erhebungsstag erfasst. Zunächst

wurde nach der Anzahl der Reisen mit mindestens einer auswärtigen Übernachtung gefragt. Für die letzten drei dieser Reisen wurden der hauptsächliche Zweck, die Lage des Ziels, die genutzten Verkehrsmittel, die Anzahl der Übernachtungen, die Anzahl der begleitenden Haushaltsmitglieder und die ungefähre Entfernung des Reiseziels erfasst.

Bei Stichtagserhebungen kommt es in der Regel zu einer Untererfassung von Wegen im Wirtschaftsverkehr, da Personen, bei denen während der Arbeitszeit sehr viele Wege anfallen, in der Stichtagsabfrage oft nicht darüber berichten. Aus diesem Grund wurde bereits 2002 ein Wirtschaftsverkehrsmittel entwickelt, das sich speziell an berufliche Vielfahrer wendet (zum Beispiel Vertreter, Lieferanten, Busfahrer, Postboten). Das Wirtschaftsverkehrsmittel liefert Ergebnisse zum berufsbedingten Verkehr eines Tages (Hauptzweck, Anzahl der zurückgelegten Wege, Gesamtentfernung der Wege und überwiegend genutztes Verkehrsmittel). Der insgesamt über die MiD abgebildete Wirtschaftsverkehr setzt sich aus den gesondert erfassten regelmäßigen beruflichen Wegen (rbW) und den über die normale Wegeabfrage erhobenen sporadischen dienstlichen Wegen zusammen.

Zusätzlich zu den bisher dargestellten, in dieser Form auch in der Vorgängerstudie enthaltenen Fragen wurden in 2017 erstmals Fragen im Rahmen von Modulen erhoben. Auf diese Weise konnten neue Fragen in die Erhebung aufgenommen werden. Da den Befragten jeweils nur ein Teil der in den insgesamt sieben Modulen zusammengefassten Fragen gestellt wurde, konnte einer zu hohen Belastung durch einen zu umfangreichen Fragenkatalog entgegengewirkt werden.

Die folgenden sieben Module können unterschieden werden:

- Reisemodul (nur Basisstichprobe),
- Automodul (nur Basisstichprobe),
- Persönliche Mobilitätsmerkmale,
- Infrastruktur und digitale Versorgung,
- Nahmobilität und Radverkehr,
- Zufriedenheit und Einstellungen,
- Etappenmodul (ab der zweiten Feldhälfte).

In der Bundesstichprobe erhielten alle Befragten das Reisemodul und die Hälfte zufallsgesteuert das Automodul. Das Etappenmodul wurde ab der zweiten Feldhälfte eingesetzt und deutschlandweit sowie für die Vertiefungsstichprobe von Baden-Württemberg erhoben. Die übrigen vier Module wurden in der Bundesstichprobe so verteilt, dass jede Person davon zwei in einer zufälligen Kombination erhielt. Die regionalen Auftraggeber konnten aus den vier Modulen (persönliche Mobilitätsmerkmale, Infrastruktur und digitale Versorgung, Nahmobilität und Radverkehr, Zufriedenheit und Einstellungen) zwei Module auswählen oder sich für den zufälligen Einsatz der Module für die von ihnen beauftragten Haushalte entscheiden. Die Steuerung der Module war dabei auf die beiden Erhebungsmethoden Telefon- und Online-Erhebung beschränkt. Das heißt Haushalten, die für ein Modul ausgewählt wurden und schriftlich an der Erhebung teilgenommen haben, konnten die Modulfragen nicht gestellt werden.

### 3.2 Definition zentraler Begriffe

Im Folgenden werden zentrale Begriffe der Studie Mobilität in Deutschland vorgestellt. Die genaue Abgrenzung der Begriffe ist bei der Interpretation der Ergebnisse zu berücksichtigen. Dabei können drei Arten von Begriffen unterschieden werden:

- Begriffe, die Grundlage für die Abgrenzung des Erhebungsgegenstandes waren. Hierunter fallen Begriffe wie ein Weg, eine Etappe, eine Reise.

- Begriffe, die bei der Datenaufbereitung von Relevanz waren, wenn bspw. aus Variablen mit Originäranworten der Befragten neue, analytische Variablen abgeleitet wurden. Hierzu gehören Begriffe wie Hauptverkehrsmittel und ökonomischer Status.
- Begriffe, die durch die Zuspelung externer Rauminformationen für die MiD bedeutend sind, wie der regionalstatistische Raumtyp.

#### Zentrale Begriffe bei der Abgrenzung des Untersuchungsgegenstands

**Weg:** Von einem Weg wird gesprochen, wenn sich eine Person außer Haus zu Fuß oder mit anderen Verkehrsmitteln von einem Ort zu einem anderen Ort bewegt. Hin- und Rückweg werden als zwei Wege gezählt. Erfolgt auf dem Weg zu einem Ziel der Umstieg zwischen verschiedenen Verkehrsmitteln, bleibt es weiterhin ein Weg. Pro Weg werden alle genutzten Verkehrsmittel erhoben, nicht jedoch der Zeitanteil (mit Ausnahme einer kleinen Stichprobe des Etappenmoduls) und der Entfernungsanteil, der pro Weg auf die verschiedenen Verkehrsmittel entfällt.

**Etappe:** Eine Etappe ist die Teilstrecke eines Weges, für die dasselbe Verkehrsmittel genutzt wird. Findet ein Wechsel des Verkehrsmittels statt, beginnt eine neue Etappe des Weges.

**Reise:** Eine Reise definiert sich bei der MiD über das Merkmal der Übernachtung. Im Sinne der MiD handelt es sich um eine Reise, wenn mindestens eine auswärtige Übernachtung angefallen ist. Für das Vorliegen einer Reise muss keine Mindestentfernung zurückgelegt werden.

**Regelmäßiger beruflicher Weg (rbW):** Hierzu gehören alle regelmäßig von Personen durchgeführten Wege, die während der Arbeitszeit aufgrund ihres Berufs anfallen. Typische Berufsgruppen mit regelmäßigen beruflichen Wegen sind Handwerker, Busfahrer, Postboten, Vertreter und Lieferanten. Der Weg zur Arbeit gehört nicht zu den regelmäßigen beruflichen Wegen.

**Wirtschaftsverkehr:** Unter Wirtschaftsverkehr werden alle Wege in Ausübung des Berufs verstanden. Die meisten Wege des in der MiD ermittelten Wirtschaftsverkehrs werden über das separate Befragungsmodul zu den regelmäßigen beruflichen Wegen erfasst, das sich an spezielle Berufsgruppen (Busfahrer, Vertreter etc.) richtet. Doch auch bei Berufen, die nicht per se durch eine Vielzahl an beruflichen Wegen

gekennzeichnet sind, fallen dienstliche Wege an. Wurde am Stichtag zum Beispiel eine Dienstreise unternommen, taucht dieser Weg im Wegeprotokoll unter der Wegezweckkategorie dienstlich auf. Der über die MiD abgebildete Wirtschaftsverkehr setzt sich daher aus den rbW und den im Wegeprotokoll berichteten dienstlichen Wegen zusammen.

### Zentrale Begriffe der Datenaufbereitung

**Computer Assisted Telephone Interview (CATI):** Befragungsmethode, bei der ein Interviewer telefonisch einen Teilnehmer befragt. Im Interview wird ein computergestütztes Fragebogenskript verwendet, das automatisch die Filterführung steuert und die Angaben speichert.

**Computer Assisted Web Interview (CAWI):** Befragungsmethode, bei der Teilnehmer einen schriftlichen Fragebogen selbst ausfüllen, der über das Internet bereitgestellt wird. Der programmierte Fragebogen übernimmt dabei die Filtersteuerung.

**Paper And Pencil Interview (PAPI):** Befragungsmethode, bei der die Teilnehmer einen schriftlichen Papierfragebogen ausfüllen.

**Stellvertreterinterview (Proxy-Interview):** Interview, bei dem die Interviewfragen nicht von der Person selbst, sondern stellvertretend von einem anderen Mitglied im Haushalt beantwortet wurden. Der Umfang des Fragenkatalogs war in diesem Fall eingeschränkt.

**Hauptverkehrsmittel:** Pro Weg wurden alle genutzten Verkehrsmittel erfasst. Wurde nur ein Verkehrsmittel genutzt, so war dieses das Hauptverkehrsmittel. Besteht ein Weg dagegen aus mehreren Etappen, für die unterschiedliche Verkehrsmittel genutzt wurden, wird das Verkehrsmittel zum Hauptverkehrsmittel erklärt, mit dem aller Wahrscheinlichkeit nach die längste Teilstrecke des Weges zurückgelegt wurde. Dazu wurde auf eine Hierarchie der Verkehrsmittel zurückgegriffen.

**Hauptwegezweck:** Im Rahmen der MiD erfolgte bei Online- und Telefoninterviews eine differenzierte Erfassung des Wegezwecks mit mehr als 40 Detailzwecken. Bei der Variable Hauptwegezweck wurde die Anzahl der Zwecke auf die sieben wesentlichen reduziert, die auch im Rahmen der schriftlichen Erhebung erfasst wurden. Zudem wurden den Rückwegen vom vorherigen Weg und den Wegen nach Hause

der Zweck des zuvor genannten Weges bzw. bei einer Wegekette der Zweck mit dem höchsten Rang zugeordnet. Damit hat der Hauptwegezweck folgende Ausprägungen: Arbeit, dienstlich, Ausbildung, Einkauf, Erledigung, Freizeit und Begleitung.

Beim Wegezweck Arbeit handelt es sich um den Weg zur Arbeitsstätte, der Wegezweck dienstlich steht für Wege während der Arbeitszeit. Erledigungen können zum Beispiel ein Arztbesuch oder ein Behördengang sein. Bei Begleitwegen hat die berichtende Person eine andere Person begleitet, zum Beispiel hat eine Mutter ihr Kind zur Schule gebracht. Während in diesem Fall beim Kind der Wegezweck Ausbildung angegeben ist, wird der Weg bei der Mutter dem Wegezweck Begleitung zugeordnet.

**Haushaltstyp (Variante 1):** Die Einteilung des differenzierten Haushaltstyps erfolgte über die Anzahl erwachsener Personen und Kinder im Haushalt. Da in Abhängigkeit vom Alter deutliche Unterschiede im Mobilitätsverhalten bestehen, wurde bei den Ein- und Zweipersonenhaushalten zusätzlich eine Differenzierung nach dem Alter vorgenommen. Unterschieden werden die Altersgruppen unter 30 Jahren, 30 bis 59 Jahre sowie 60 Jahre und mehr. Bei den Zweipersonenhaushalten erfolgt die Zuordnung über das Alter der jüngeren Person im Haushalt. Mehrpersonenhaushalte mit Kindern wurden nach dem Alter des Kindes oder der Kinder ebenfalls weiter unterteilt. Daraus ergeben sich elf Gruppen:

- Einpersonenhaushalte mit Personen im Alter von 18 bis unter 30 Jahren (junge Alleinlebende),
- Einpersonenhaushalte mit Personen im Alter von 30 bis unter 60 Jahren (Alleinlebende mittleren Alters),
- Einpersonenhaushalte mit Personen im Alter von 60 Jahren und mehr (ältere Alleinlebende),
- Haushalte mit zwei Erwachsenen mit mindestens einer Person im Alter von 18 bis unter 30 Jahren (junge Zweipersonenhaushalte),
- Haushalte mit zwei Erwachsenen mit mindestens einer Person im Alter von 30 bis unter 60 Jahren (Zweipersonenhaushalte mittleren Alters),
- Haushalte mit zwei Erwachsenen im Alter von 60 Jahren und mehr (ältere Zweipersonenhaushalte),



- Haushalte mit drei und mehr erwachsenen Personen (Haushalte mit mindestens drei Erwachsenen),
- Haushalte mit zwei Erwachsenen und mindestens einem Kind unter 6 Jahren,
- Haushalte mit zwei Erwachsenen und mindestens einem Kind unter 14 Jahren,
- Haushalte mit zwei Erwachsenen und mindestens einem Kind unter 18 Jahren,
- Alleinerziehende.

**Haushaltstyp (Variante 2):** Die Einteilung des zweiten Haushaltstyps basiert ebenfalls auf dem Alter der Haushaltsmitglieder und der Frage, ob Kinder im Haushalt leben. Eine Unterscheidung von Einperson- und Mehrpersonenhaushalten findet nicht statt. Der Haushaltstyp setzt sich aus den folgenden vier Kategorien zusammen:

- Junge Haushalte: Alle im Haushalt lebenden Personen sind zwischen 18 und 34 Jahre alt.
- Familienhaushalte: Mindestens eine Person im Haushalt ist jünger als 18 Jahre.
- Haushalte mit Erwachsenen: Alle im Haushalt lebenden Personen sind 18 Jahre und älter. Es lebt mindestens eine Person mittleren Alters (zwischen 35 und 64 Jahren) im Haushalt, oder der Haushalt setzt sich sowohl aus jungen als auch älteren Personen zusammen.
- Haushalte mit Personen ab 65 Jahren: Alle im Haushalt lebenden Personen sind 65 Jahre und älter.

**Modalgruppen:** Die Einteilung in Modalgruppen basiert auf den Angaben zur allgemeinen Nutzungshäufigkeit der Verkehrsmittel Auto, Fahrrad und öffentlichen Verkehrsmitteln. Wird üblicherweise nur eines der drei Verkehrsmittel im Wochenverlauf genutzt (ein bis drei Mal pro Woche oder häufiger), handelt es sich um eine monomodale Person, werden zwei oder alle drei der Verkehrsmittel innerhalb einer Woche genutzt um eine multimodale Person. In Summe können drei monomodale und je nach Kombination der Verkehrsmittel vier multimodale Personengruppen unterschieden werden:

- Monomodale Autofahrer,
- Monomodale ÖV-Nutzer,
- Monomodale Fahrradfahrer,
- Multimodale Auto-Fahrrad-Nutzer,
- Multimodale Auto-ÖV-Nutzer,
- Multimodale Fahrrad-ÖV-Nutzer,
- Multimodale Auto-Fahrrad-ÖV-Nutzer.

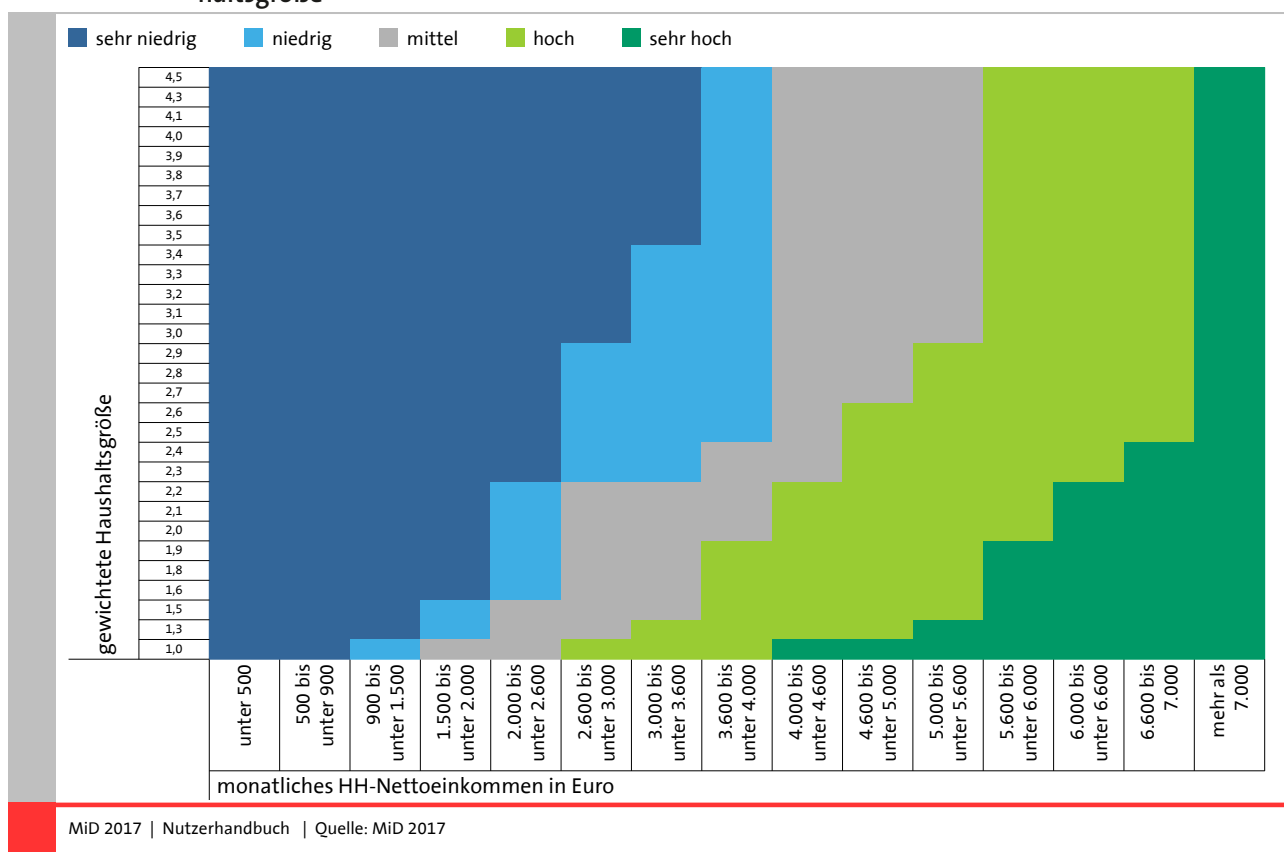
**Verhaltenshomogene Gruppen:** Die Zugehörigkeit einer Person zu einer verhaltenshomogenen Gruppe wird aus den Angaben zu Alter, Erwerbstätigkeitsstatus und Auto-Verfügbarkeit abgeleitet. Es gibt zwei Variablen dieses Typs, eine mit neun und eine mit zwölf Ausprägungen. Bei der 9er Variante werden folgende Gruppen unterschieden:

- Berufstätige mit verfügbarem Pkw,
- Berufstätige ohne verfügbaren Pkw,
- Nicht-Berufstätige mit verfügbarem Pkw,
- Nicht-Berufstätige ohne verfügbaren Pkw,
- Studenten,
- Auszubildende,
- Kinder von 0 bis 6 Jahren,
- Schüler von 7 bis 10 Jahren,
- Schüler ab 11 Jahren.

Bei der 12er Variante wird auch bei Studenten, Auszubildenden und Schülern ab elf Jahren nach der Auto-Verfügbarkeit differenziert.

**Ökonomischer Status:** Der ökonomische Status eines Haushalts leitet sich gemäß dem Prinzip des Äquivalenzeinkommens, das sich in der Sozial- und Armutsforschung für Analysen der Einkommensverteilung etabliert hat, aus dem Haushaltsnettoeinkommen und der gewichteten Haushaltsgröße ab. Ohne Gewichtung der Haushaltsgröße wären Familien überproportional oft in den niedrigen Einkommensgruppen vertreten. Die gewichtete Haushaltsgröße wird aus der Anzahl und dem Alter der Haushaltsmitglieder bestimmt. Kinder unter 14 Jahren gehen mit dem Faktor 0,3 ein. Die erste Person ab 14 Jahren im Haushalt erhält den Gewichtungsfaktor 1, alle weiteren Personen ab 14 Jahren den Faktor 0,5. Für eine vierköpfige Familie mit zwei Kindern unter 14 Jahren ergibt sich so eine gewichtete Haushaltsgröße von  $1 + 0,5 + 0,3 + 0,3 = 2,1$  anstelle einer ungewichteten Haushaltsgröße von 4. Anhand der Matrix aus Haushaltsnettoeinkommen und gewichteter Haushaltsgröße wird jeder Haushalt einer von fünf Kategorien, die von sehr niedrig bis sehr hoch reichen, zugeordnet (Abbildung 2).

**Abbildung 2** Bestimmung des ökonomischen Status über Haushaltsnettoeinkommen und gewichtete Haushaltsgröße



### Zentrale Begriffe zur Rauminformation

**Regionalstatistischer Raumtyp (RegioStaR):** Hierbei handelt es sich um die „Regionalstatistische Raumtypologie für die Mobilitäts- und Verkehrsforschung (RegioStaR)“, die vom Bundesministerium für Verkehr und digitale Infrastruktur mit Unterstützung des Bundesinstituts für Bau-, Stadt- und Raumforschung (BBSR) konzipiert wurde (siehe [www.bmvi.de/regiostar](http://www.bmvi.de/regiostar)). Diese ist ein neues Instrument, um Wirkungszusammenhänge zwischen Verkehr und räumlichen Strukturen analysieren und räumlich unterschiedliches Mobilitätsverhalten sichtbar machen zu können. Die regionalstatistische Raumtypologie ist inzwischen auch im Katalog der räumlichen Gliederungen der statistischen Ämter des Bundes und der Länder hinterlegt und kann dort für Sonderauswertungen der amtlichen Statistik verwendet werden.

Die Typologie basiert auf siedlungsstrukturellen Merkmalen auf der Ebene von Gemeindeverbänden. Sie ist damit einerseits deutlich feinteiliger und räumlich präziser als etablierte Raumtypologien auf der Ebene von Kreisen und kreisfreien Städten. Zudem sind die Ergebnisse dadurch in der zeitlichen Entwicklung besser vergleichbar, weil Änderungen der

Gemeindeverbände nur wenig Auswirkungen auf die Ergebnisse der Typen haben, während die Kreise durch Gebietsreformen immer größer und heterogener werden und zu größeren Brüchen in der Zeitreihe führen (zum Beispiel der Zusammenschluss der Stadtkreise Aachen und Hannover mit ihrem jeweiligen Umlandkreis).

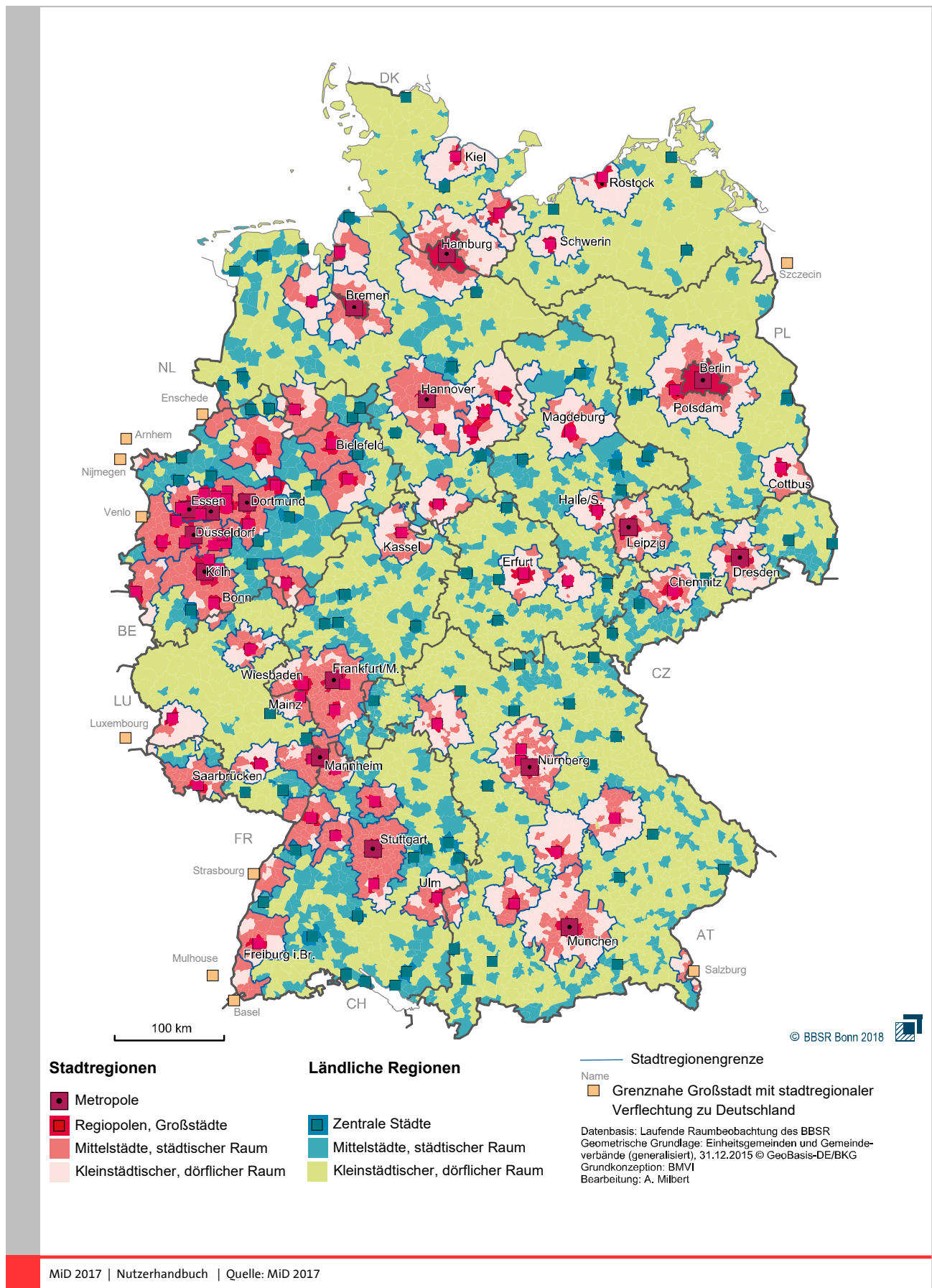
Die regionalstatistische Raumtypologie ist hierarchisch aufgebaut. Für die Zuordnung und Verwendung von Daten auf der Ebene der Gemeindeverbände fächert sich die Typologie von einem groben, in Stadtregionen und ländliche Regionen unterschiedenen Regionstyp, über einen in vier Typen differenzierten Regionstyp letztlich in 17 Raumtypen auf. Diese 17 Raumtypen (RegioStaR 17) bilden den Kern der Raumtypologie und sind Grundlage für verschiedene Zusammenfassungen in Raum- und Gemeindetypen (Abbildung 3). Der regionalstatistische Raumtyp ist in verschiedenen Varianten in den Datensätzen enthalten. Der Karte in Abbildung 4 ist die räumliche Verteilung des zusammengefassten regionalstatistischen Raumtyps (RegioStaR 7) zu entnehmen, der Grundlage für die Erstellung des Ergebnisberichts war.

**Abbildung 3** Regionalstatistische Raumtypologie (RegioStaR) des BMVI

Regionalstatistischer Regionstyp <i>RegioStaR 2</i>	<b>1 Stadtregion</b>		<b>2 Ländliche Region</b>	
Differenzierter regionalstatistischer Regionstyp <i>RegioStaR 4</i>	<b>11 Metropolitane Stadtregion</b>	<b>12 Regiopolitane Stadtregion</b>	<b>21 Stadtregionsnahe ländliche Region</b>	<b>22 Periphere ländliche Region</b>
Regionalstatistischer Raumstyp <i>RegioStaR 17</i>	111 Metropole 112 Großstadt 113 Mittelstadt 114 Städtischer Raum 115 Kleinstädtischer, dörflicher Raum	121 Regiopole  123 Mittelstadt 124 Städtischer Raum 125 Kleinstädtischer, dörflicher Raum	211 Zentrale Stadt  213 Mittelstadt 214 Städtischer Raum 215 Kleinstädtischer, dörflicher Raum	221 Zentrale Stadt  223 Mittelstadt 224 Städtischer Raum 225 Kleinstädtischer, dörflicher Raum
Zusammengefasster regionalstatistischer Raumstyp <i>RegioStaR 7</i>	71 Metropolen (111) 72 Regiopolen und Großstädte (112, 121) 73 Mittelstädte, städtischer Raum (113, 114, 123, 124) 74 Kleinstädtischer, dörflicher Raum (115, 125)		75 Zentrale Städte (211, 221)  76 Mittelstädte, städtischer Raum (213, 214, 223, 224) 77 Kleinstädtischer, dörflicher Raum (215, 225)	

MiD 2017 | Nutzerhandbuch | Quelle: MiD 2017

Abbildung 4 Zusammengefasster Regionalstatistische Raumtypologie (RegioStaR 7)



## 4 Datenaufbereitung

Die Datenaufbereitung erfolgte analog zu 2008 in einem mehrstufigen Verfahren. Generell haben die Komplexität der Datensätze und der Datenbereinigung zugenommen. Dies ist zum einen auf den Methodenmix aus CATI, CAWI und PAPI zurückzuführen, der in 2017 im Gegensatz zur Vorgängerstudie in beiden Erhebungsphasen (Haushaltsinterview und Personen-Wege-Interview) angewendet wurde. Zum anderen hat der Umfang der Datensätze stark zugenommen. Infolgedessen fällt die Anzahl unlogischer Fälle, deren Anteil nach wie vor gering ist, um ein mehrfaches höher aus als 2008. Für die Sicherstellung einer hohen Konsistenz der Daten wurden umfangreiche Prüfroutinen entwickelt, die beim Vorliegen bestimmter Ausgangsbedingungen und bei ausgewählten Variablen trotz Datenumfang auch Einzelfallanalysen beinhaltet haben.

Die Datenaufbereitung umfasste die folgenden Schritte:

- Erfassen der schriftlichen Fragebogen,
- Einfügen regelmäßiger beruflicher Wege in den Wegedatensatz,
- formale Bereinigung und Kodierung fehlender Werte,
- Extremwertbereinigung und Prüfung der inhaltlichen Konsistenz,
- Ergänzen und Löschen von Wegen,
- Imputation von Werten,
- Anpassung der Anzahl immobiler Personen,
- Bildung analytischer Variablen sowie
- Zuspielen zusätzlicher Rauminformationen.

Die Datenbereinigung der aktuellen MiD folgte wie in den Vorgängererhebungen dem generellen Grundsatz, nur dann korrigierend einzugreifen, wenn eine hohe Gewissheit über den „wahren“ Wert vorliegt. Im Zweifel wurden Inkonsistenzen belassen. Es handelt sich

zumeist nur um wenige Fälle, die bei genauer Betrachtung irritieren können, jedoch keine Auswirkungen auf das Ergebnis haben.

Im Folgenden werden die Bereinigungs- und Plausibilisierungsschritte erläutert. Eine ausführliche, variablengenaue Beschreibung des Vorgehens ist dem Begleitdokument „Variablenaufbereitung“ des Nutzerhandbuchs zu entnehmen.

### 4.1 Formale Datenaufbereitung

#### Erfassen der schriftlichen Fragebogen

Die Erfassung der zurückgesendeten Papierfragebogen erfolgte händisch mithilfe einer Eingabemaske. Dieses Verfahren ist zeitintensiver als das Einscannen von Fragebogen, weist allerdings den Vorteil einer genaueren Erfassung auf. Darüber hinaus können Notizen am Rand des Fragebogens erfasst werden. Für das einheitliche Einlesen der Fragebogen wurden allgemeine und variablenspezifische Erfassungsregeln erstellt. So wurden nicht beantwortete Fragen generell auf „keine Angabe“ gesetzt. Wurde bei der Frage zur Carsharing-Mitgliedschaft zum Beispiel sowohl „ja, bei einem Anbieter“ als auch „ja, bei mehreren Anbietern“ angekreuzt, wurde der Code für die Mitgliedschaft bei mehreren Anbietern eingetragen.

#### Einfügen regelmäßiger beruflicher Wege in den Wegedatensatz

Da es bei Stichtagserhebungen in der Regel zu einer Untererfassung des Wirtschaftsverkehrs kommt, enthält die MiD spezifische Fragen zu den Wegen, die von bestimmten Berufsgruppen regelmäßig während ihrer Arbeitszeit durchgeführt werden (siehe Kapitel 3.1). Aus diesen pauschal erhobenen Angaben zu Wegeanzahl, genutztem Verkehrsmittel und insgesamt zurückgelegter Entfernung wurden einzelne Wege erzeugt und in den Wegedatensatz integriert. Die Entfernung wurde gleichmäßig auf die Wege verteilt, das berichtete Verkehrsmittel auf alle rbW übertragen und der Wegezweck auf „dienstlich“ gesetzt. Die rbW einer Person sind somit alle identisch und haben keine Angabe zu Start- und Ankunftszeit.

### Formale Bereinigung und Kodierung fehlender Werte

Die MiD verfügt sowohl beim Telefon- (CATI) als auch beim Online-Interview (CAWI) über eine komplexe Filterführung. Der erste Schritt der formalen Datenaufbereitung war es, den dadurch entstandenen fehlenden Werten (Missings) einen eindeutig zu interpretierenden Code zuzuweisen. Im Ergebnis kann damit bei jeder Variablen für jeden Fall (etwa eine Person oder ein Haushalt) nachvollzogen werden, warum kein Wert im gültigen Bereich oder ein fehlender Wert vorliegt. Um eine ausreichende Antwortqualität sicherzustellen, wurden daran anschließend Fälle identifiziert und gelöscht, bei denen sehr häufig keine Antwort auf Fragen gegeben wurde und daher nur der Wert für „keine Angabe“ vorlag.

Beim schriftlichen Fragebogen (PAPI) sind die Möglichkeiten der Filtersetzung begrenzt. Damit sich die Datenauswertung jeweils auf eine klar umrissene Personengruppe bezieht (zum Beispiel Personen ab einem bestimmten Alter), wurde die Filtersetzung vom CATI und CAWI im Rahmen der Datenaufbereitung nachträglich auf die PAPI-Fälle übertragen.

Ein weiterer Arbeitsschritt der formalen Bereinigung war die Zuordnung der offenen Nennungen zu einem gültigen Code. Allein beim Wegezweck lagen 26.000 solcher Angaben vor, die im Nachhinein kategorisiert wurden.

### Extremwertbereinigung und Prüfung der inhaltlichen Konsistenz

Im Anschluss an die formale Aufbereitung wurde auf Ebene der einzelnen Variablen überprüft, ob die Angaben in einem plausiblen Wertebereich liegen. Werte außerhalb eines festgesetzten Bereichs wurden auf „unplausibel“ gesetzt. Deutlich komplexer gestaltete sich die variablenübergreifende inhaltliche Konsistenzprüfung. So können die Angaben verschiedener Variablen nicht zusammenpassen (zum Beispiel eine Person ist zu jung, um einen Führerschein zu besitzen) oder es liegen für dieselbe, designbedingt sowohl im Haushalts- als auch im Personen-Wege-Interview erhobene Frage unterschiedliche Angaben vor (zum Beispiel Altersangaben im Haushalts- und Personen-Wege-Interview unterscheiden sich). In diesem Fall wurden unter Berücksichtigung mehrerer Angaben soweit möglich Plausibilisierungsregeln aufgestellt. Lagen unterschiedliche Angaben zu demselben Sachverhalt auf Haushalts- und Personenebene vor, wurde der Angabe auf Personenebene die höhere

Priorität eingeräumt. Im Zweifel wurde die Inkonsistenz belassen.

## 4.2 Weiterführende Datenaufbereitung

### Ergänzen von Wegen

Für welche der drei Erhebungsmethoden CATI, CAWI und PAPI sich die Probanden entscheiden, hat nur geringen Einfluss auf das Ergebnis. Lediglich bei der Anzahl der berichteten Wege ist ein Methodeneffekt erkennbar (Gruschwitz et al. 2018). So werden beim PAPI weniger Wege berichtet als bei den Interviewer- und Computer-gestützten Verfahren mit der Möglichkeit der gezielten Nachfrage, ob alle Wege des Tages berichtet wurden. Da viele Befragten beim Papierfragebogen insbesondere wenige Nach-Hause-Wege angeben, wurde im Zuge der Datenaufbereitung ein Verfahren zur Identifikation dieser Fälle und zum Einfügen von auf den Tagesablauf abgestimmten Nach-Hause-Wegen entwickelt.

### Löschen zu viel berichteter Wege

Die Analyse der Daten hat gezeigt, dass ein nicht unerheblicher Anteil an Personen ihre am Stichtag während der Arbeitszeit durchgeführten regelmäßigen beruflichen Wege sowohl bei den allgemeinen Fragen zum Wirtschaftsverkehr als auch im Wegeprotokoll berichten. Da aus den pauschal erhobenen Angaben zum Wirtschaftsverkehr wie beschrieben Einzelwege generiert und in den Wegedatensatz integriert werden, liegt damit eine doppelte Berichterstattung vor. In diesem Fall wurde das Einfügen der regelmäßigen beruflichen Wege in den Wegedatensatz rückgängig gemacht.

Zur Reduktion des Aufwandes für die Befragten wurden Wege, bei denen ein Haushaltsmitglied die Begleitung durch ein anderes Haushaltsmitglied berichtet hat, automatisch auf die Begleitperson übertragen. Somit lagen bei der Beantwortung der Fragen im Online- oder Telefoninterview durch die Begleitperson bereits Angaben vor. Dennoch wurden diese Wege zum Teil neu eingegeben. Diese durch identische Angaben identifizierbaren Wege wurden gelöscht.

### Imputation von Werten

Bei jeder Erhebung entstehen Datenlücken durch fehlende oder unplausible Angaben. Bei der Ausweisung von Prozentwerten und Mittelwerten sind kleine Anteile fehlender Werte kein Problem. Schwieriger wird

es, wenn auf die Gesamtbevölkerung hochgerechnete Absolutwerte ermittelt werden. Das Ergebnis fällt in diesem Fall zu niedrig aus.

Bei der MiD 2008 kamen nur einfache Verfahren zum Einsatz, wie das Einsetzen des Mittelwerts auf Basis bestehender Angaben bei metrischen Variablen. Bei der MiD 2017 wurden dagegen statistische Schätzverfahren genutzt, die je nach betroffener Personengruppe und Situation das Einsetzen spezifischer Werte erlaubt. Die Imputation von fehlenden Werten erfolgte für die drei, für die Berechnung von Mobilitätskennwerten grundlegenden Variablen „Hauptverkehrsmittel“, „Entfernung“ und „Wegedauer“. Darüber hinaus wurden beim Haushalts-Nettoeinkommen fehlende Angaben ersetzt.

### Anpassung der Anzahl immobiler Personen

Im Gegensatz zu 2008 wurden in 2017 alle Personen zu ihrer Mobilität am Stichtag gefragt. Damit ist eine gute Ausgangslage für die Ermittlung der Mobilitätsquote gegeben. Ein Teil der mobilen Personen hat im Weiteren jedoch nicht über ihre Wege am Stichtag berichtet. Dadurch ergeben sich bei der Berechnung von durchschnittlichen Tageswerten für die Gesamtbevölkerung zu niedrige Werte, da alle immobilen Personen an dieser Stelle mit einer Null einfließen, für einen Teil der mobilen Personen jedoch kein Wert im gültigen Bereich vorliegt. Für die Ermittlung korrekter Werte wurde analog zu den mobilen Personen ein Teil der immobilen Personen auf Missing gesetzt. Die Auswahl der Personen erfolgte über die modellbasierte Berechnung der Wahrscheinlichkeit, dass eine immobile Person zur Gruppe der Nichtberichter gehört.

### Bildung analytischer Variablen

Nicht alle in den Datensatzpaketen enthaltenen Variablen stammen direkt aus der Befragung. Bei einem Teil der Variablen handelt es sich um Informationen, die aus den erhobenen Angaben abgeleitet wurden. Beispiele für diese „analytischen Variablen“ sind der Haushaltstyp, der ökonomische Status des Haushalts, die Einteilung in mono- und multimodale Personengruppen oder die Klassifikation der Autos nach Kriterien des Kraftfahrt-Bundesamtes.

### Zuspielen zusätzlicher Rauminformationen

Analog zu 2008 stehen auch bei der aktuellen Erhebung zusätzliche Rauminformationen zur Verfügung. Die Informationen stammen aus drei Quellen: dem

Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR, zum Beispiel regionalstatistischer Raumtyp), dem Statistischen Bundesamt (zum Beispiel politischen Gemeindegrößenklasse) und infas360 (zum Beispiel Qualität des öffentlichen Personennahverkehrs am Wohnort). Über diese Variablen lassen sich der Einfluss von Raum- und Siedlungsstruktur auf die Verkehrsnachfrage näher analysieren. Ein Teil der Raumvariablen wurde den Einzeldatensätzen zugespielt. Ein anderer Teil der Raumvariablen ist in einem separaten, bei der Clearingstelle Verkehr des DLR erhältlichen Datensatz abgelegt und kann bei Bedarf über die Identifikationsnummer der Haushalte zugespielt werden.





# 5 Das Datenangebot der MiD 2017

## 5.1 Struktur der Daten

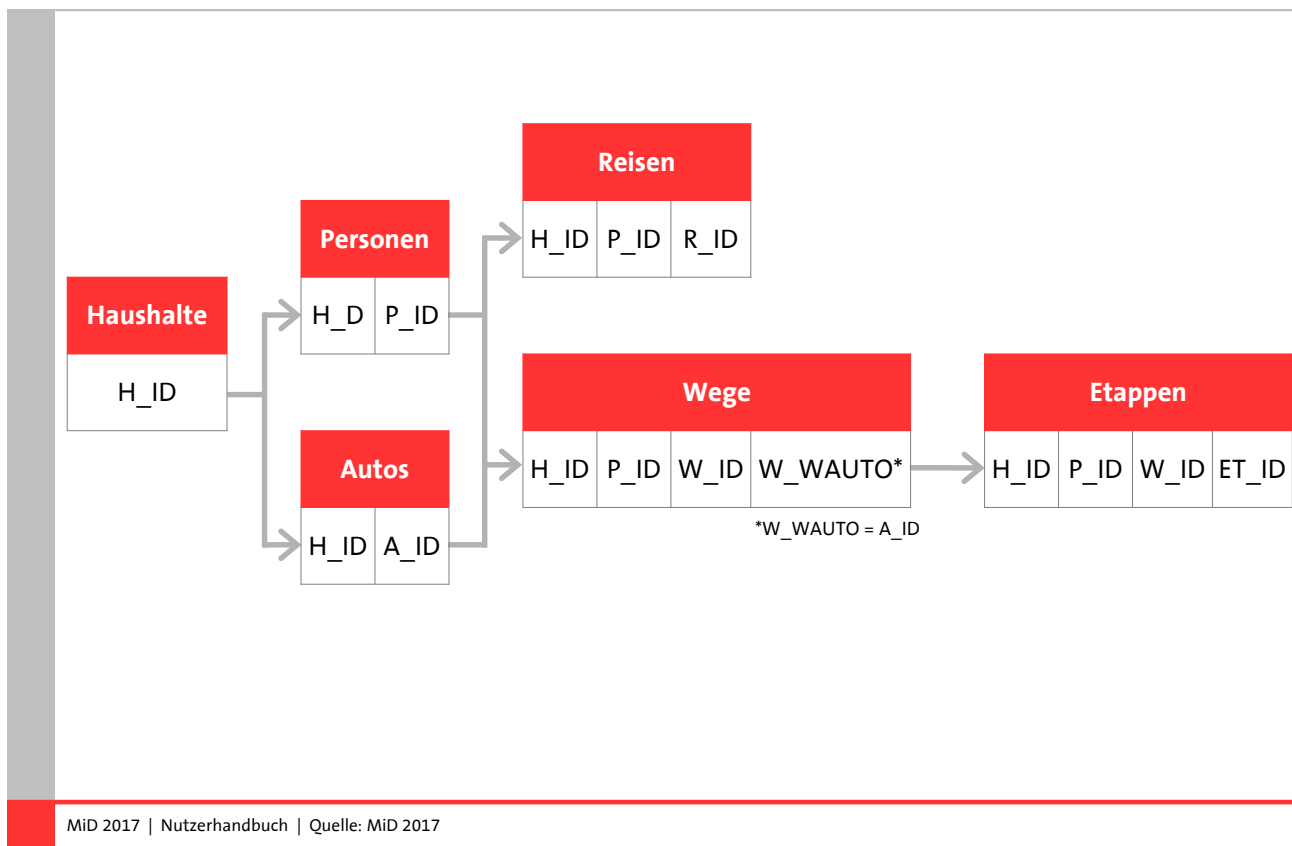
Die in der MiD 2017 im Haushalts- und Personen-Wege-Interview erfassten Daten sind in mehreren Einzeldatensätzen abgelegt. In Summe gibt es sechs Dateien. Jede Datei steht für eine Hauptinformations- und Auswertungsebene. Diese sind: Haushalte, Personen, Wege, Autos, Reisen und Etappen. Die in den Dateien enthaltenen Schlüsselvariablen erlauben die Verknüpfung der Einzeldatensätze und damit bei Bedarf die Übertragung einzelner Variablen zwischen den Datensätzen. *Abbildung 5* liefert eine Übersicht der Datensatzstruktur und der in den einzelnen Teildatensätzen vorhandenen Schlüsselvariablen.

Die sechs Dateien enthalten die nachfolgend beschriebenen Informationen.

**Datei Haushalte:** Die Haushalte bilden die oberste Ebene der Datensatzstruktur. Die Datei enthält Angaben zu: Haushaltsgröße, Alter, Geschlecht und Tätigkeit der Haushaltsmitglieder, Ausstattung mit Pkw und anderen Verkehrsmitteln, Haushaltsnettoeinkommen etc. Die Informationen auf Haushaltsebene können den anderen fünf Dateien über die Schlüsselvariable H\_ID zugespielt werden.

**Datei Personen:** Diese Datei enthält Informationen zu jeder Person eines Haushalts, die ein Personen-Wege-Interview geführt hat. Enthalten sind zum Beispiel Informationen zu Alter, Tätigkeit und Führerscheinbesitz sowie zur allgemeinen Nutzungshäufigkeit der Verkehrsmittel. In Abhängigkeit des für eine Person relevanten Moduls sind Informationen zu

**Abbildung 5** Datensatzstruktur: Teildatensätze und Schlüsselvariablen



Einstellungen, zur Zufriedenheit mit der Verkehrssituation am Wohnort oder zur Nutzung mobiler Endgeräte für digitale Mobilitätsdienste etc. enthalten. Darüber hinaus werden aus der Wegedatei abgeleitete aggregierte Informationen zum Verkehrsverhalten am Stichtag bereitgestellt, wie die Anzahl der am Stichtag zurückgelegten Wege, die Tagesstrecke und die Dauer aller Wege. Jede Person ist über die Haushaltsnummer (H\_ID) und Personennummer (P\_ID) eindeutig identifizierbar.

**Datei Wege:** Diese Datei enthält die von mobilen Personen für den Stichtag berichteten Wege. Für jeden Weg sind u. a. Informationen zu Start- und Ankunftszeit, gewählte Verkehrsmittel, Wegelänge und -dauer sowie Wegezweck und Anzahl der Begleitpersonen enthalten. Da bei der MiD das Inlandsprinzip angewendet wird, das heißt nur Wege innerhalb Deutschlands berücksichtigt werden, wurden die Wege von Personen, die ausschließlich im Ausland unterwegs waren, nicht erhoben.

**Datei Autos:** Bei der MiD wurden für bis zu drei Fahrzeuge eines Haushalts detaillierte Angaben erfasst. Die Gesamtanzahl der Pkw ist im Haushaltsdatensatz abgelegt. Die Detailinformationen zu den Pkw umfassen u. a. Antriebsart, Baujahr und Jahresfahrleistung. Bei Haushalten, die für das Fahrzeug-Modul ausgewählt wurden, liegen zudem Angaben zu Halter, Motorleistung und üblichem Stellplatz vor. Im Wegedatensatz sind darüber hinaus Angaben zur Nutzung des Fahrzeugs auf einzelnen Wegen am Stichtag enthalten. Daraus wurde die Gesamtanzahl der mit dem Pkw am Stichtag zurückgelegten Wege und Kilometer abgeleitet und in den Autodatensatz eingefügt. Über die Haushaltsnummer (H\_ID) und die Pkw-Nummer (A\_ID) ist jeder Pkw eindeutig identifizierbar und einem Haushalt zuzuordnen.

**Datei Reisen:** Die Personen der Basisstichprobe wurden nach ihren Reisen in den letzten drei Monaten befragt. Bis zu drei Reisen wurden detailliert berichtet. Während im Personendatensatz die Gesamtanzahl der Reisen einer Person enthalten sind, stellt der Reisedatensatz Informationen zu jeder detailliert berichteten Reise zur Verfügung. Hierzu gehören die genutzten Verkehrsmittel, die Entfernung, die Anzahl der Übernachtungen und die Anzahl der Begleiter aus dem Haushalt. Jede Reise ist durch eine eindeutige Identifikationsnummer bestehend aus Haushalts-, Personen- und Reisennummer gekennzeichnet.

**Datei Etappen:** Für einen kleinen Teil der berichteten Wege in der MiD stehen Informationen auf der Ebene von Etappen zur Verfügung. Dabei wurde die Erhebung von Etappen im Rahmen einer Wegeerhebung erprobt. Die etappengenaue Erfassung liegt für die Wege von Personen vor, die für das Modul Etappen ausgewählt waren und die bei der Wegeerhebung für einen Weg entweder die Nutzung mehrerer Verkehrsmittel oder die Nutzung eines öffentlichen Verkehrsmittels berichtet haben. In diesem Fall wurden die Personen gebeten, die Reihenfolge der genutzten Verkehrsmittel sowie von Abschnitten zu Fuß anzugeben und für jede Etappe die Etappenlänge und die Etappendauer zu berichten. Der Schwerpunkt des Einsatzes lag auf dem Land Baden-Württemberg und einer zufälligen Verteilung über das restliche Bundesgebiet. Der Einsatz erfolgte ab Dezember 2016 und ausschließlich im Rahmen des telefonischen und des Online-Interviews. Aufgrund des Versuchscharakters des Einsatzes und der regionalen und zeitlichen Disproportionalität, die die Gewichtung nicht ausgleicht, erlaubt der Datensatz keine einfache Auswertung und anschließende Generalisierung der Ergebnisse. Es gibt aber einen Einblick in die Erhebung von Etappen im Rahmen eines Wegekonzepts.

Im Gegensatz zu 2008 sind in den Daten der MiD 2017 auch die Fälle der regionalen Vertiefungsstichproben enthalten. Die Fallzahl der Vertieferstichproben ist in Summe deutlich höher als die der Basisstichprobe. Der Variable „BASISAUF“ ist zu entnehmen, welche Fälle zur Basis- und welche zu einer Vertieferstichprobe gehören. In Tabelle 1 ist für jede der sechs Dateien die Fallzahl nach Basis- und Vertiefungsstichprobe dargestellt.

**Tabelle 1** Fallzahl der sechs Teildatensätze der MiD 2017

Datei	gesamt	Basisstichprobe	Vertiefungsstichproben
Haushalte	156.420	33.389	123.031
Personen	316.361	65.459	250.902
Wege	960.619	195.199	765.420
Autos	216.844	45.253	171.591
Reisen	38.905	38.905	/
Etappen*	11.000	1.842	9.158

\* Im Datensatz sind nur Wege enthalten, bei denen mehr als ein Verkehrsmittel oder ein öffentliches Verkehrsmittel genutzt wurde

MiD 2017 | Nutzerhandbuch | Quelle: MiD 2017

Die Bezeichnung der in den Datensätzen enthaltenen Variablen basiert auf der nachfolgenden Namenskonvention:

- Variablen, deren Inhalt direkt von den Befragten erhoben wurde, beginnen je nach Informationsebene mit einem entsprechenden Großbuchstaben: H steht für Haushalte, P für Personen, W für Wege, R für Reisen, A für Autos, ET für Etappen.
- Die Indexbuchstaben HP kennzeichnen Variablen, bei denen zum selben Sachverhalt vorliegende Angaben auf Haushalts- und Personenebene nach der in Kapitel 4.1 beschriebenen Regel zusammengeführt worden sind.
- Variablen mit Informationen zu dem pro Haushalt fest vorgegebenen Stichtag für die Mobilitätserfassung beginnen mit ST. Der Buchstabe M steht für Modul.
- Mehrfachnennungsvariablen enden auf \_A, \_B etc., wobei der letzte Buchstabe für „keine Angabe“ steht.
- Im Haushaltsdatensatz enden die Namen der Variablen, die Angaben zu den einzelnen Haushaltsmitgliedern enthalten, auf \_1, \_2 bis \_8.
- Die Namen von Variablen mit Inhalten, die direkt aus der Befragung stammen, bestehen aus Großbuchstaben, die Namen von analytischen, das heißt von Originalvariablen abgeleiteten Variablen bestehen aus Kleinbuchstaben.

Bei der Kodierung fehlender Werte besteht folgende Konvention:

Analog zur MiD 2008 wurden ausschließlich positive Werte für die Kodierung fehlender Werte verwendet. Bei der Abgrenzung gegenüber den gültigen Werten kann zwischen den antwortbedingt fehlenden Werten und den designbedingt fehlenden Werten unterschieden werden.

**Antwortbedingt fehlende Werte** sind durch die Werte 94, 95, 9 bzw. 99 gekennzeichnet. Sofern der gültige Wertebereich diese Werte beinhaltet, wird ihnen die Indexziffer 9 vorangestellt. Die Anzahl der Indexziffern richtet sich dabei nach der Feldbreite der jeweiligen Variable. Die während der Erhebung zur Verfügung stehenden Antwortkategorien „weiß nicht“ oder „kann oder möchte ich nicht beantworten“ wurden im Zuge der Datenaufbereitung zu „keine Angabe“ zusammengefasst. Tabelle 2 ist die Bedeutung der Codes für fehlende Werte zu entnehmen.

**Tabelle 2** Kodierung antwortbedingt fehlender Werte

Wert	Wertelabel
94, 994, 9994, ...	unplausibler Wert
95, 995, 9995, ...	„Wert nicht zu berechnen“ oder „nicht zuzuordnen“
9, 99, 999, ...	keine Angabe

MiD 2017 | Nutzerhandbuch | Quelle: MiD 2017

**Designbedingt fehlende Werte** sind durch dreistellige Codes gekennzeichnet. Sofern der gültige Wertebereich diese Werte beinhaltet, sind den Codes entsprechende Indexziffern vorangestellt. Die Anzahl der Indexziffern richtet sich dabei nach der Feldbreite der jeweiligen Variable. Die letzten drei Ziffern beinhalten die spezifische Kodierung. Die erste dieser drei Ziffern gibt dabei die Art des fehlenden Wertes an. Zum Beispiel steht die 1 für Module, die 2 für die Interviewart. In Tabelle 3 sind alle Arten designbedingt fehlender Werte mit je einem Beispiel dargestellt.

Jeder Code hat nur eine Bedeutung. Die Codes wurden zunächst fortlaufend vergeben. Im Zuge der Datenaufbereitung wurden zur Vereinfachung allerdings einige Codes zusammengefasst. Die Codes entsprechen nicht den Codes der MiD 2008.

**Tabelle 3** Kodierung designbedingt fehlender Werte

Ziffer	Art des fehlenden Werts	Beispiel Wert	Wertelabel
1	Module	101, 1101, 10101, ...	Modul Fahrzeugmerkmale und Fahrzeugbesitz nicht erhalten
2	Interviewart	202, 2202, 20202, ...	im PAPI nicht erhoben
3	Haushaltsgröße	308, 3308, 30308, ...	Person 8 im HH nicht vorhanden
4	Soziodemographie	403, 4403, 40403, ...	Person unter 16 Jahren
5	Mobilitätseigenschaften	502, 5502, 50502, ...	HH ohne Auto
6	Stichtag außer Haus	602, 6602, 60602, ...	Person ohne Außer-Haus-Aktivität
7	Wege	701, 7701, 70701, ...	bei rbW nicht erhoben
8	Mobilität am Stichtag	804, 8804, 80804, ...	Person mit unbekannter Mobilität
9	Etappen	904, 9904, 90904, ...	Etappenerfassung nicht vorgesehen

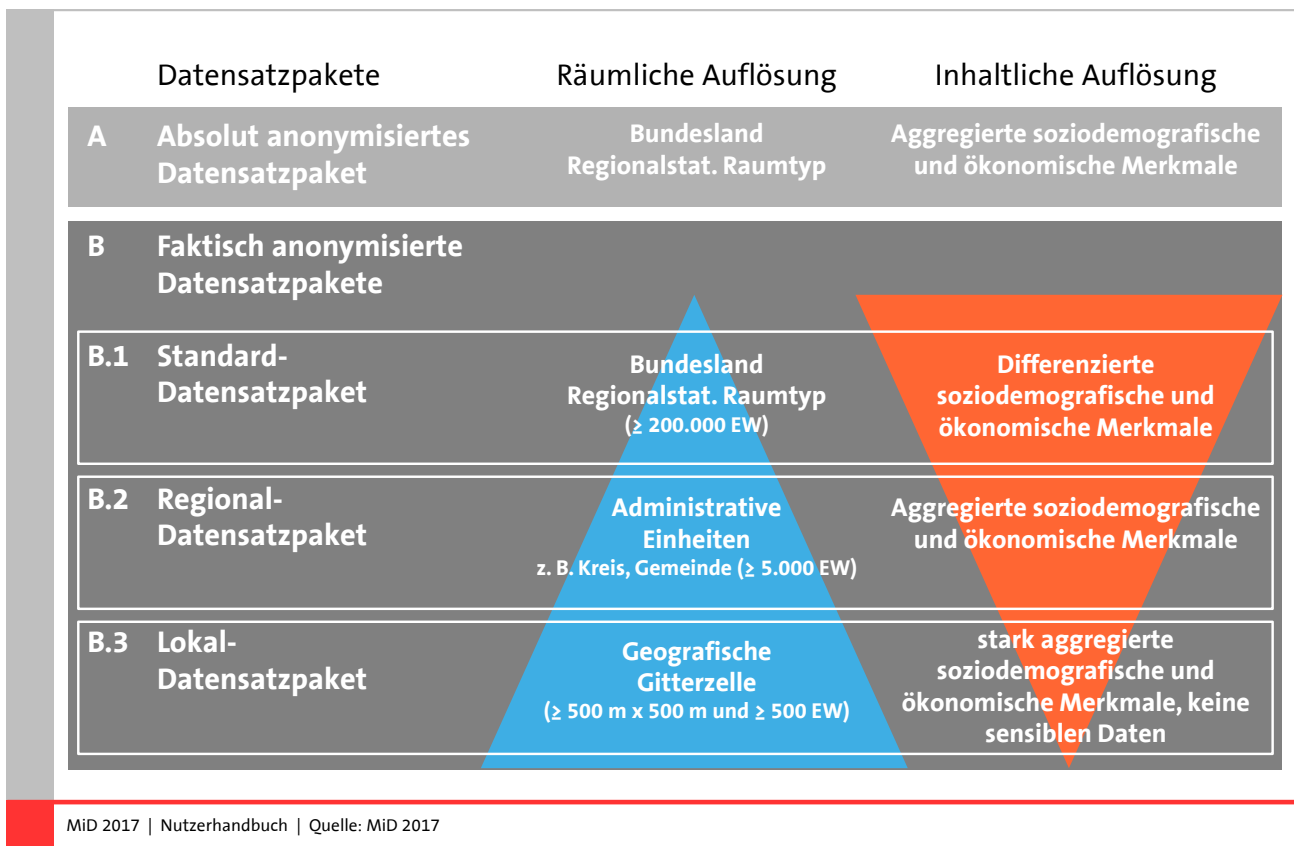
MiD 2017 | Nutzerhandbuch | Quelle: MiD 2017

## 5.2 Datensatzpakete

Die im vorherigen Abschnitt beschriebenen Einzeldatensätze der MiD 2017 stehen in vier Varianten zur Verfügung. Die vier Datensatzpakete unterscheiden sich im Variablenbestand und in der Auflösung der räumlichen und soziodemografischen Angaben. Im Nachfolgenden werden die vier Datensatzpakete und ihre Einsatzfelder beschrieben. **Abbildung 6** enthält eine Übersicht der MiD-Datensatzpakete.

### (1) Absolut anonymisiertes Datensatzpaket (Datensatzpaket A)

Bei diesem Datensatzpaket wurde der Variablenbestand reduziert und die soziodemografischen und räumlichen Angaben vergrößert, sodass eine Reidentifizierung der auskunftsgibenden Personen ausgeschlossen ist. So wird zum Beispiel das Alter nicht als metrischer Wert bereitgestellt, sondern in Gruppen. Neben dem Bundesland stehen lediglich raumbeschreibende Variablen wie zum Beispiel der zusammengefasste regionalstatistische Gemeindetyp, die Luftlinienentfernung zum nächsten Bahnhof und

**Abbildung 6** Übersicht der MiD-Datensatzpakete

die Qualität des Nahverkehrs am Wohnort zur Verfügung. Für alle Analysen, die keinen höheren Differenzierungsgrad der soziodemografischen Angaben und keine konkrete räumliche Verortung von Startpunkt und Ziel der Wege oder des Wohnortes erfordern, stellt dieses Datensatzpaket eine geeignete Grundlage dar. Das Datensatzpaket enthält die Teildatensätze Haushalte, Autos, Personen und Wege.

## **(2) Faktisch anonymisierte Datensatzpakete (Datensatzpakete B1, B2 und B3)**

Neben dem absolut anonymisierten Datensatzpaket gibt es drei faktisch anonymisierte Datensatzpakete. Bei diesen Datensatzpaketen hat ebenfalls eine Informationsreduktion stattgefunden. Dabei gilt folgender Grundsatz: Je tiefer die räumliche Auflösung der Daten, umso gröber werden die Angaben zur Soziodemografie bereitgestellt. Umgekehrt geht eine hohe Auflösung auf soziodemografischer Ebene mit einer starken Vergröberung auf der räumlichen Ebene einher.

### **B1: Standard-Datensatzpaket der MiD**

Das Standard-Datensatzpaket der MiD entspricht dem Public-Use-File der 2008er MiD-Daten. Dieses Datensatzpaket enthält die höchste Auflösung von soziodemografischen und -ökonomischen Angaben zu Haushalten und Personen. Um eine Reidentifizierung der Befragten auszuschließen, werden nur eingeschränkte Rauminformationen bereitgestellt. Für die Auswertung des Zusammenhangs von Mobilität und Raumstruktur stehen im Standard-Datensatzpaket die Angaben zum Bundesland sowie den Wohnort beschreibende Raumvariablen zur Verfügung, die vom BBSR, dem Statistischen Bundesamt und infas360 bereitgestellt wurden. Der Bestand an Raumvariablen ist gegenüber Datensatzpaket A etwas erweitert. Eine konkrete „Verortung“ der teilnehmenden Haushalte ist nicht möglich und für eine Vielzahl an Analysen, wie zum Beispiel der Ergebnisbericht der MiD zeigt, nicht nötig. Dieses Datensatzpaket enthält als einziges alle sechs Teildatensätze.

### **Datensätze mit detaillierten Rauminformationen**

Neben dem Standard-Datensatzpaket stehen zwei Datensatzpakete mit einer höheren räumlichen Tiefe zur Verfügung. In dem einen Fall werden für die Start- und Zielorte der Wege sowie die Wohnorte der Personen die Gemeindeverbandskennziffern bereitgestellt. In dem anderen Fall liegen die räumlichen Angaben

auf Ebene von kleinräumigen Gitterzellen vor. Beide Datensatzpakete richten sich an Nutzer mit spezifischen Auswertungsinteressen. Die zwei Datensatzpakete enthalten die drei Teildatensätze Haushalte, Personen und Wege.

### **B2: Regional-Datensatzpaket – Datensätze mit Angabe der Gemeindeverbandskennziffer**

In diesem Datensatzpaket werden die Gemeindeverbandskennziffer sowie ein deutlich erweitertes Set an raumbeschreibenden Variablen bereitgestellt. Das Regional-Datensatzpaket bietet damit die Möglichkeit zu räumlichen Analysen auf der Ebene von Gemeindeverbänden. Zur Einhaltung der Richtlinien des Datenschutzes sind die Merkmalsausprägungen der soziodemografischen und -ökonomischen Variablen zum Teil vergrößert und der Variablenbestand reduziert. Bei Analysen auf Ebene der Gemeindeverbände sollten jeweils nur solche berücksichtigt werden, die mit einer ausreichenden Fallzahl in den Daten enthalten sind. Bei Gemeindeverbänden mit weniger als 5.000 Einwohnern steht die Kennziffer aus Datenschutzgründen nicht zur Verfügung.

### **B3: Lokal-Datensatzpaket – Datensätze mit Angabe von kleinräumigen Gitterzellen**

Im Unterschied zu 2008 wurden 2017 die Adressen der teilnehmenden Haushalte sowie die Start- und Zieladressen der Wege erfasst. Die so entstandenen Geokoordinaten wurden zur faktischen Anonymisierung der Daten in mindestens 500x500 Meter großen und mindestens 500 Einwohner umfassenden MiD-Gitterzellen abgelegt. Im Lokal-Datensatzpaket sind die räumlichen Informationen auf Gitterzellenebene enthalten. Damit sind Analysen unterhalb der Gemeindeebene möglich. Analog zum Regional-Datensatzpaket ist der Variablenbestand reduziert, die soziodemografischen Angaben sind vergrößert.

Bei allen vier beschriebenen Datensatzpaketen sind die Einzeldatensätze weitgehend bereinigt und plausibilisiert. Bei den drei grundlegenden Variablen zur Mobilität „Hauptverkehrsmittel“, „Entfernung“ und „Wegedauer“ wurden fehlende Werte imputiert. Darüber hinaus stehen analytische Variablen zur Verfügung, deren Werte nicht erhoben sondern aus den Merkmalen von Befragungsvariablen abgeleitet wurden (zum Beispiel Mobilitätstyp oder ökonomischer Status des Haushalts).

**Abbildung 7** Zugelassener Nutzerkreis der MiD-Datensatzpakete

Datensatzpakete	Wer	Wofür?
<b>A</b> Absolut anonymisiertes Datensatzpaket	Alle	Verkehrspolitik, -planung und -forschung
<b>B</b> Faktisch anonymisierte Datensatzpakete	MiD-Auftraggeber	Verkehrspolitik, -planung und -forschung
	Forschungseinrichtungen (Staatliche, staatlich anerkannte Hochschulen; Forschungseinrichtungen, deren Tätigkeit überwiegend aus öffentlichen Mitteln finanziert wird)	Verkehrsforschung
	Weitere Dritte mit einem öffentlichen Auftrag	Verkehrsforschung

MiD 2017 | Nutzerhandbuch | Quelle: MiD 2017

Im Frühjahr 2019 wird ein fünftes Datensatzpaket zur Verfügung gestellt, das sich im Grad der Aufbereitung unterscheidet und dem in 2008 bereitgestellten Zusatzfile entspricht. Dieses Datensatzpaket kennzeichnet eine nur formale Bereinigung. So wurde fehlenden Werten ein eindeutiger Code zugewiesen. Eine inhaltliche Plausibilisierung der Daten hat nur in sehr geringem Maß stattgefunden. Dieses Datensatzpaket steht einem speziellen Nutzerkreis zur Verfügung, der für eigene Aufbereitungen und Analysen bei allen Fällen auf die Originalangaben der Probanden zurückgreifen möchte.

### 5.3 Bereitstellung der Datensatzpakete

Alle Datensatzpakete können über die Clearingstelle Verkehr des Deutschen Zentrums für Luft- und Raumfahrt erworben werden ([www.clearingstelle-verkehr.de](http://www.clearingstelle-verkehr.de)). Während das absolut anonymisierte Datensatzpaket A allen Personen und Institutionen zur Verfügung steht, ist der Nutzerkreis der faktisch anonymisierten Datensatzpakete eingeschränkt. Die drei Datensatzpakete B1, B2 und B3 werden ausschließlich für die Durchführung unabhängiger wissenschaftlicher Forschung sowie von Forschungsprojekten im Bundesinteresse und im öffentlichen Auftrag bereitgestellt. Bei den beiden Datensatzpaketen mit detaillierten Rauminformationen (B2 und B3) muss zudem dargelegt werden, wofür die Informationen auf kleinräumiger Ebene benötigt werden. Die Anträge werden einzeln geprüft. Für die Nutzung der Daten ist die Unterzeichnung eines speziellen Nutzungsvertrags notwendig.

## 6 Hinweise zur Datennutzung

### 6.1 Gewichtung und Hochrechnung

Die Grundgesamtheit der MiD 2017 umfasst die Wohnbevölkerung der Bundesrepublik Deutschland im Alter ab null Jahren. Sollen für diese Grundgesamtheit repräsentative Ergebnisse berechnet werden, sind bei der Auswertung der Variablen Gewichtungsfaktoren zu verwenden. Über die Multiplikation der Variablen mit einem Gewichtungsfaktor werden zum einen die unterschiedlichen Auswahlwahrscheinlichkeiten bei der Stichprobenziehung ausgeglichen. Zum anderen wird die Verteilung relevanter Merkmale (zum Beispiel Alter und Geschlecht) an die der Grundgesamtheit angeglichen. Sollen auf die Grundgesamtheit hochgerechnete Absolutwerte berechnet werden, wird der Gewichtungsfaktor zusätzlich mit einem Faktor multipliziert, um von der Anzahl der Fälle in der Stichprobe auf die Anzahl der Fälle in der Grundgesamtheit zu kommen. Das Ergebnis ist ein Hochrechnungsfaktor.

Vor allem auf Haushaltsebene ergeben sich ohne Gewichtung schiefe Verteilungen. Dies ist auf die Bildung der Stichprobe zurückzuführen, die auf der Ziehung von Personen bzw. Telefonanschlüssen basiert. Auswahlgrundlage für die MiD sind zum einen die Einwohnermelderegister der Gemeinden, zum anderen die Festnetz- und Mobilfunkanschlüsse in Deutschland. So steht zum Beispiel eine aus dem Einwohnermelderegister ausgewählte Person stellvertretend für ihren Haushalt. Da jeweils alle Personen aus dem Haushalt befragt werden, ergibt sich die Haushaltsstichprobe unmittelbar aus der Personenstichprobe. Dies führt zwangsläufig zu einer Unterrepräsentierung kleiner Haushalte: Die Wahrscheinlichkeit in die Stichprobe zu kommen, fällt bei Personen aus Mehrpersonenhaushalten höher aus als bei Personen, die in Einpersonenhaushalten leben. Dies gilt analog für die Festnetz- und die Mobilfunkstichprobe, da die Zahl der Anschlüsse im Haushalt tendenziell mit der Haushaltsgröße zunimmt.

In jedem der sechs Teildatensätze der MiD ist ein Gewichtungsfaktor enthalten. In den Teildatensätzen Haushalte, Personen und Wege steht darüber hinaus ein Hochrechnungsfaktor zur Verfügung. Die genaue

Bildung des Haushalts- und Personengewichts ist dem Methodenbericht zu entnehmen. An dieser Stelle werden lediglich die bei der Berechnung der Haushalts- und Personengewichte einbezogenen Merkmale genannt. Danach werden die auf den Haushalts- bzw. Personengewichten basierenden Gewichte von Auto-, Wege- und Reisedatensatz beschrieben. Am Ende werden Hinweise zur Berechnung von hochgerechneten Werten gegeben.

Bei der Kalibrierung auf Haushaltsebene („Haushaltsgewicht“) wurden die folgenden Merkmale einbezogen:

- Bundesland des Wohnortes,
- Haushaltsgröße (eine Person, zwei, drei, vier, fünf und mehr Personen) und
- Wohnsituation (Miete/ Eigentum)

sowie die zeitlichen Merkmale

- Quartal (Januar bis März etc.) und
- Wochentag (Montag, Dienstag, Mittwoch, ..., Sonntag) des Berichtstags.

Um mit dem Hochrechnungsfaktor die Gesamtzahl der in Deutschland existierenden 41 Mio. Haushalte zu erzielen, wurde das Gewicht jedes Falls im Haushaltsdatensatz mit 261,88 multipliziert. Jeder Haushalt im Datensatz repräsentiert damit rund 262 Haushalte in Deutschland.

Die Kalibrierung auf Personenebene („Personengewicht“) erfolgte anhand der Merkmale:

- Bundesland,
- Erwerbstätigkeit bzw. überwiegendem Lebensunterhalt,
- Schulabschluss,
- Elementargebiet,
- Geschlecht und
- Altersgruppe.

Um mit dem Hochrechnungsfaktor die Gesamtzahl der in Deutschland lebenden 82 Mio. Menschen zu erzielen, wurde das Gewicht eines jeden Falls im Personendatensatz mit 259,75 multipliziert. Jede Person im Datensatz repräsentiert damit rund 260 Personen in Deutschland.

Ausgangsbasis der Gewichte im Wegedatensatz sind die Personengewichte. Da nicht alle Personen, die am Stichtag mobil waren, über ihre Wege berichtet haben, würde das Verkehrsaufkommen ohne Ausgleich dieser fehlenden Wege systematisch zu niedrig ausfallen. Um dies zu vermeiden, wurden alle berichteten Wege bzw. deren Gewichte mit einem Hebefaktor multipliziert, sodass sie in Summe auch die Wege der „mobilen Nicht-Berichter“ beinhalten. Zusätzlich erfolgte ein Ausgleich für Wege, die über die zwölf (Telefon- und Online-Interview) bzw. acht (schriftliches Interview) detailliert berichteten Wege hinausgehen und nur noch als Anzahl notiert wurden. Auch diese weiteren Wege sind nicht im Wegedatensatz enthalten, müssen aber für die korrekte Ermittlung des Verkehrsaufkommens integriert werden. In diesem Fall wurden die Wegegewichte von Personen mit weiteren Wegen um den Faktor angehoben, der sich über die Division von Gesamtwegezahl (berichtete Wege plus weitere Wege) durch die berichteten Wege ergibt. Nur so ergibt sich hochgerechnet das für alle mobilen Personen am Stichtag korrekte Verkehrsaufkommen von 257 Mio. Wegen und bei Summenbildung der Entfernungsangaben je Weg die korrekte Verkehrsleistung von 3.214 Mio. Kilometern.

Die Gewichte im Reisedatensatz basieren ebenfalls auf den Personengewichten. In diesem Fall hat ein Ausgleich stattgefunden, wenn die Person in den letzten drei Monaten mehr Reisen als die drei im Reisedatensatz dokumentierten unternommen hat. Hierfür wurde das Personengewicht mit einem Faktor

multipliziert, der sich aus der Division von Gesamtanzahl der Reisen durch die Anzahl berichteter Reisen ergibt. Bei den Gewichten im Autodatensatz handelt es sich um die Haushaltsgewichte. Für die Gewichtung der Etappen im Etappendatensatz werden die Wegegewichte verwendet.

Die Gewichte sind in jedem Datensatz fallzahlnormiert. Im Mittel ergibt sich daher jeweils der Wert 1. Werden die Hochrechnungsfaktoren verwendet, ergeben sich:

- im Haushaltsdatensatz die Gesamtanzahl der Haushalte in Deutschland,
- im Personendatensatz die Bevölkerungsanzahl Deutschlands,
- im Wegedatensatz die geschätzte Zahl der an einem durchschnittlichen Tag in Deutschland von der deutschen Wohnbevölkerung insgesamt zurückgelegten Wege.

In **Tabelle 4** sind für jeden Teildatensatz die Namen der Gewichtungs- und falls vorhanden Hochrechnungsfaktoren angegeben sowie die mit diesen Faktoren erzielte Anzahl an Fällen.

Bei der Auswertung von Merkmalen mit vielen fehlenden Werten oder wenn sehr spezifische Gruppen mit kleinen Fallzahlen analysiert werden (zum Beispiel die ab 80-Jährigen mit Führerschein in einem Mehrpersonenhaushalt im ländlichen Raum), sollte vor der Interpretation der Ergebnisse über einen Vergleich von gewichteten und ungewichteten Werten der Einfluss des Gewichts analysiert werden.

**Tabelle 4** Namen und erzielte Fallzahl von Gewichtungs- und Hochrechnungsvariable je Teildatensatz der MiD

Ebene	Name Gewichtungsvariable	Normierte Fallzahl	Name Hochrechnungsvariable	Hochgerechnete Fallzahl
Haushalte	H_GEW	156.420	H_HOCH	40.963.002
Personen	P_GEW	316.361	P_HOCH	82.175.684
Wege	W_GEW	960.619	W_HOCH	257.466.940
Autos	A_GEW	216.844	/	/
Reisen	R_GEW	38.905	/	/
Etappen	ET_GEW	11.000	/	/

MiD 2017 | Nutzerhandbuch | Quelle: MiD 2017



**Tabelle 5** Häufigkeitsverteilung der Pkw-Verfügbarkeit als Fahrer oder Mitfahrer

Pkw-Verfügbarkeit als Fahrer oder Mitfahrer (einschließl. Carsharing), unabhängig vom Stichtag	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
		%	%	%
<b>Gültig</b>				
1 jederzeit	43.855.637	53,4	77,0	77,0
2 gelegentlich	7.800.392	9,5	13,7	90,7
3 gar nicht	5.299.287	6,4	9,3	100,0
<b>Gesamt</b>	<b>56.955.316</b>	<b>69,3</b>	<b>100,0</b>	
<b>Fehlend</b>				
9 keine Angabe	3.183.701	3,9		
206 Erwachsener ab 14 Jahre (Proxy)	11.936.889	14,5		
402 Kind unter 14 Jahren	10.099.778	12,3		
<b>Gesamt</b>	<b>25.220.368</b>	<b>30,7</b>		
<b>Gesamt</b>	<b>82.175.684</b>	<b>100,0</b>		

MiD 2017 | Nutzerhandbuch | Quelle: MiD 2017

Bei der Hochrechnung auf Gesamtwerte ist folgendes zu beachten: Viele Variablen der MiD weisen fehlende Werte (Missings) auf. Für die korrekte Ermittlung hochgerechneter Werte muss oft ein Teil dieser Missings bei der Hochrechnung einbezogen werden. Das Vorgehen wird im Folgenden exemplarisch anhand der Pkw-Verfügbarkeit erläutert.

In Tabelle 5 ist die Häufigkeitsverteilung der Variable Pkw-Verfügbarkeit als Fahrer oder Mitfahrer dargestellt. Die Ausgabe ist mit dem Hochrechnungsfaktor gewichtet. In Summe ergibt sich so die Gesamtanzahl der Bevölkerung von 82 Mio. Menschen. Den Missing-Codes ist zu entnehmen, dass die Frage Personen unter 14 Jahren nicht gestellt wurde. 3,9 Prozent der Personen haben die Frage erhalten, diese jedoch nicht beantwortet. Weiteren 14,5 Prozent der Befragten wurde die Fragen nicht gestellt, obwohl sie 14 Jahre

und älter sind, da das Interview stellvertretend von einer anderen Person im Haushalt geführt wurde.

Soll berechnet werden, wie viele Menschen in Deutschland jederzeit, gelegentlich oder gar nicht über einen Pkw verfügen können, müssen die beiden Gruppen mit der Kategorie „keine Angabe“ und „Erwachsene ab 14 Jahre (Proxy)“ einbezogen werden. Vereinfacht wird angenommen, dass sich die Antwortkategorien bei Personen ohne Antwort auf die gleiche Weise verteilen würden wie bei Personen mit Antwort. Die 15 Mio. Menschen ohne Antwort werden danach zu 77 Prozent der Gruppe mit der Antwort „jederzeit“ zugeschlagen, 13,7 Prozent der Gruppe mit der Antwort „gelegentlich“ und 9,3 Prozent der Gruppe mit der Antwort „gar nicht“. Das auf diese Weise erzielte Ergebnis ist Tabelle 6 zu entnehmen. Von den insgesamt 72 Mio. Menschen in Deutschland ab 14 Jahren steht danach 55 Mio. Menschen jederzeit ein Pkw zur Verfügung, zehn Mio. Menschen gelegentlich und sieben Mio. Menschen gar nicht.

**Tabelle 6** Pkw-Verfügbarkeit als Fahrer oder Mitfahrer hochgerechnet auf die Bevölkerung ab 14 Jahren

Pkw-Verfügbarkeit als Fahrer oder Mitfahrer	Häufigkeit
<b>Gültig</b>	
1 jederzeit	55.498.502
2 gelegentlich	9.871.253
3 gar nicht	6.706.151
<b>Gesamt</b>	<b>72.075.906</b>
<b>Fehlend</b>	
Personen unter 14 Jahren	10.099.778
<b>Gesamt</b>	<b>82.175.684</b>

MiD 2017 | Nutzerhandbuch | Quelle: MiD 2017

## 6.2 Ermittlung der korrekten Tagesstrecke

Im Wegedatensatz ist für jeden Weg die Entfernung enthalten. Werden die Entfernungsangaben aller Wege einer Person aggregiert, ergibt sich die Tagesstrecke. Die so ermittelten Tagesstrecken sind bei Personen, die neben den detailliert im Wegedatensatz beschriebenen Wegen weitere Wege berichtet haben, mit einem Hebefaktor zu multiplizieren. Dabei wird angenommen, dass die nicht mehr detailliert berichteten Wege den bereits berichteten entsprechen. Hat eine Person im schriftlichen Datensatz zum Beispiel acht Wege mit einer Gesamtlänge von 20 Kilometern

detailliert beschrieben und angegeben, dass sie zwei weitere Wege durchgeführt hat, dann wird die Tagesstrecke mit dem Faktor 1,25 multipliziert. Der Faktor ergibt sich aus der Division der Gesamtsumme von zehn Wegen durch acht detailliert berichtete Wege. Die Tagesstrecke der Person erhöht sich damit auf 25 Kilometer. Wird die so gebildete Variable Tagesstrecke auf Personenebene mit dem Personenfaktor gewichtet, ergeben sich repräsentative Ergebnisse für die Personen Deutschlands.

Die hier beschriebene Variable Tagesstrecke wurde bereits aus den im Wegedatensatz enthaltenen Wegen abgeleitet und in den Personendatensatz integriert. Soll eine Differenzierung der Tagesstrecke zum Beispiel nach Verkehrsmitteln oder Wegezweck vorgenommen werden, ist das hier beschriebene Verfahren mit der jeweils gewünschten Untergliederung anzuwenden.

### 6.3 Stellvertreterinterviews, Missings und Module

Die MiD ist durch eine hohe Anzahl an Missing-Kategorien gekennzeichnet. Hierbei handelt es sich nur zu einem sehr geringen Anteil um die Nicht-Beantwortung von Fragen (item non response). Der größte Anteil der Missings beschreibt designbedingt fehlende Werte. So ergeben sich bei Modulfragen hohe Ausfälle,

da sie nur Personen, die für das Modul ausgewählt waren, gestellt wurden. Auch im Falle von Stellvertreterinterviews wurden viele Fragen übersprungen. Darüber hinaus kommen aus inhaltlichen Gründen fehlende Werte vor (qualified missings). So wurden Fragen, die für bestimmte Personengruppen irrelevant waren, ebenfalls nicht gestellt. Nicht-berufstätige Personen wurden beispielsweise nicht nach regelmäßigen beruflichen Wegen gefragt.

Bei der Auswertung muss daher jeweils folgendes beachtet werden:

1. Welchem Personenkreis wurde die Frage gestellt?
2. Auf welche Grundgesamtheit bezieht sich das Ergebnis?
3. Welchem Teil der Befragten wurde die Frage ggf. nicht gestellt, obwohl sie aufgrund ihrer Eigenschaften zur Grundgesamtheit gehören?

Tabelle 7 enthält das Beispiel der Häufigkeit von Online-Shopping. Die Kategorie 402 deutet darauf hin, dass die Frage Kindern unter 14 Jahren nicht gestellt wurde. Die Grundgesamtheit der Frage sind damit alle Personen ab 14 Jahren. Der Kategorie 104 ist zu entnehmen, dass es sich bei dieser Frage um eine Modulfrage handelt. Ein Großteil der Befragten hat die Frage zum Online-Shopping damit nicht erhalten, obwohl sie älter als 14 Jahre sind. Zudem waren folgende

**Tabelle 7** Häufigkeitsverteilung bei der Frage nach Online-Shopping

Häufigkeit Online-Shopping (CATI/CAWI)	Häufigkeit	Prozent	Gültige Prozepte	Kumulierte Prozepte
		%	%	%
<b>Gültig</b>				
1 täglich bzw. fast täglich	861	0,3	1,4	1,4
2 an 1-3 Tagen pro Woche	5.035	1,6	8,4	9,8
3 an 1-3 Tagen pro Monat	20.347	6,4	33,9	43,7
4 seltener als monatlich	17.665	5,6	29,4	73,2
5 fast nie	13.818	4,4	23,0	96,2
6 habe kein Internet	2.188	0,7	3,6	99,9
9 keine Angabe	90	0	0,1	100,0
<b>Gesamt</b>	<b>60.004</b>	<b>19,0</b>	<b>100,0</b>	
<b>Fehlend</b>				
104 Modul Infrastruktur und digitale Versorgung nicht erhalten	189.020	59,7		
202 im PAPI nicht erhoben	37.292	11,8		
206 Erwachsener ab 14 Jahre (Proxy)	21.651	6,8		
402 Kind unter 14 Jahren	8.394	2,7		
<b>Gesamt</b>	<b>256.357</b>	<b>81,0</b>		
<b>Gesamt</b>	<b>316.361</b>	<b>100,0</b>		

MiD 2017 | Nutzerhandbuch | Quelle: MiD 2017

Personengruppen für die Beantwortung der Frage ausgewählt, es liegen aber keine Antworten vor: (1) Personen, für die stellvertretend eine andere Person aus dem Haushalt geantwortet hat, (2) Personen, die für das Modul ausgewählt waren, denen im schriftlichen Papierfragebogen die Frage aber nicht gestellt werden konnte sowie (3) wenige Personen, die die Frage erhalten aber nicht beantwortet haben. Bei (1) und (2) handelt es sich um designbedingt fehlende Werte, dies ist die Masse der fehlenden Werte, bei (3) um item non response. Lediglich 90 von 60.004 Befragten haben die Frage erhalten und nicht beantwortet.

Ein anderes Beispiel ist die Frage nach Home Office (Tabelle 8). In diesem Fall wurde die Frage Personen erst ab 16 Jahren gestellt und zwar nur dann, wenn sie zuvor angegeben haben, dass sie berufstätig sind. Die Grundgesamtheit sind damit alle Personen ab 16 Jahren, die berufstätig sind. Auch hier handelt es sich um eine Modulfrage, sodass für einen großen Teil der Befragten (Kategorie 103) designbedingt keine Antwort vorliegt. Folgende Gruppen waren für die Beantwortung der Frage ausgewählt, es liegen aber keine Antworten vor: (1) Personen, die berufstätig sind, für die jedoch stellvertretend eine andere Person aus dem Haushalt geantwortet hat, (2) Personen, die berufstätig sind und für das Modul ausgewählt waren, aber im schriftlichen Papierfragebogen die Frage nicht erhalten haben, (3) und auch hier nur wenige

Personen, die berufstätig sind und die Frage erhalten, aber nicht beantwortet haben (item non response).

Wie ist nun bei der Interpretation der Ergebnisse der Umstand zu berücksichtigen, dass einem großen Teil der Befragten, der zur Grundgesamtheit gehört, aus unterschiedlichen Gründen die jeweilige Frage nicht gestellt wurde? Tatsächlich ist die Auswahl der Erhebungsmethode durch die Probanden, ebenso wie das Führen von Stellvertreterinterviews und der Ausfall bei vorgelagerten Fragen nicht rein zufällig über die Stichprobe verteilt. Bei der Interpretation der Ergebnisse kann dennoch vereinfacht angenommen werden, dass die vorliegenden Antworten für alle Befragten der Grundgesamtheit gelten und dem „wahren“ Wert sehr nahe kommen. Theoretisch ließen sich der Methodeneffekt und der Einfluss des Ausfalls durch Stellvertreterinterviews berechnen. Dabei handelt es sich aber um ein aufwendiges Verfahren, das in den meisten Fällen nur geringen Einfluss auf das Ergebnis haben dürfte. Daher wurde auch beim Ergebnisbericht vereinfacht das erzielte Ergebnis als Wert für die Grundgesamtheit angenommen. Die beiden Beispiele verdeutlichen in jedem Fall den Nutzen, sich in die Vielzahl an Missing-Kategorien der MiD einzuarbeiten, um die Ergebnisse richtig einordnen und interpretieren zu können.

**Tabelle 8** Häufigkeitsverteilung bei der Frage nach Home Office

Home Office ja/nein (CATI/CAWI)	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
		%	%	%
<b>Gültig</b>				
1 ja	5.100	1,6	16,8	16,8
2 nein	25.242	8,0	83,2	100,0
9 keine Angabe	8	0	0	100,0
<b>Gesamt</b>	<b>30.350</b>	<b>9,6</b>	<b>100,0</b>	
<b>Fehlend</b>				
103 Modul persönliche Mobilitätsmerkmale nicht erhalten	197.398	62,4		
202 im PAPI nicht erhoben	34.857	11,0		
206 Erwachsener ab 14 Jahre (Proxy)	20.589	6,5		
403 Person unter 16 Jahren	8.364	2,6		
404 Nicht-Berufstätige(r)	21.232	6,7		
406 Berufstätige(r) als Nebentätigkeit oder Auszubildende(r)	3.025	1,0		
407 Person ohne Info zu Berufstätigkeit oder -umfang	546	0,2		
<b>Gesamt</b>	<b>286.011</b>	<b>90,4</b>		
<b>Gesamt</b>	<b>316.361</b>	<b>100,0</b>		

MiD 2017 | Nutzerhandbuch | Quelle: MiD 2017

## 6.4 Die Variable Mobilität am Stichtag

Die Variable Mobilität am Stichtag gibt an, wie viele Personen an dem vorgegebenen Berichtstag außer Haus waren und Wege unternommen haben und wie viele nicht. Im Datensatz sind zwei Variablen zur Mobilität am Stichtag enthalten. Die differenzierte Variante der Variablen weist eine Besonderheit auf, die im Folgenden beschrieben wird.

Personen mit der Kategorie 1, 2 und 3 der differenzierten Variable Mobilität am Stichtag waren an dem vorgegebenen Tag mobil und haben Wege berichtet (Tabelle 9). Je nachdem, ob sie nur bei der detaillierten Wegeerfassung, nur bei den Fragen zu regelmäßigen beruflichen Wegen oder bei beidem über Wege berichtet haben, sind sie einer der drei Kategorien zugeordnet. Von diesen Personen sind Wege im Wegedatensatz enthalten. In Kategorie 5 sind Personen zusammengefasst, die mobil waren, aber keine Wege berichtet haben. Entsprechend gibt es für diese Personen keine Wege im Wegedatensatz.

Personen mit der Kategorie 4 bei der differenzierten Variable Mobilität am Stichtag waren an dem vorgegebenen Tag mobil, aber nur im Ausland. Da in der MiD das Inlandsprinzip gilt und nur Wege innerhalb von Deutschland betrachtet werden, sind diese Personen in der einfachen Variable Mobilität am Stichtag, die nur zwischen nicht mobil und mobil unterscheidet, der Kategorie nicht mobil zugeordnet (Tabelle 10).

Bei der differenzierten Variable Mobilität am Stichtag gibt es darüber hinaus die Kategorien 0 und 6, beide mit der Bedeutung „nicht mobil“. Die Personen beider

Kategorien haben für den Stichtag angegeben, nicht mobil gewesen zu sein. Die Differenzierung von zwei Kategorien hat folgenden Hintergrund: Wird zum Beispiel die durchschnittliche Anzahl der Wege am Stichtag berechnet, gehen Personen ohne Wege mit dem Wert 0, mobile Personen mit der Anzahl berichteter Wege ein. Nun gibt es jedoch Personen, die mobil waren, aber keine Wege berichtet haben. Ihre Wege fehlen bei der Berechnung des Mittelwerts. Würden nun alle nicht mobilen Personen bei der Berechnung berücksichtigt, obwohl bei den mobilen Personen Angaben fehlen, läge die durchschnittliche Wegeanzahl systematisch zu niedrig. Um dies zu verhindern, wurde mithilfe eines Logit-Modells unter Verwendung der Informationen „Wegeprotokoll fehlt und Stichtag liegt länger als vier Tage zurück“, Vollständigkeit des Interviews, Person war am Stichtag verreist, Erhebungsmethode und Alter die Wahrscheinlichkeit bestimmt, dass eine nicht mobile Person analog zu den „mobilen Nicht-Berichtern“ im Weiteren ebenfalls nicht berichtet hätte. 2.100 Personen mit der höchsten Wahrscheinlichkeit wurden in der Kategorie 6 „nicht mobil (potenziell ohne Wegeangabe analog Kat. 5)“ zusammengefasst. Auf diese Weise kann nun auch bei den nicht mobilen Personen die Gruppe der (potenziellen) „Nicht-Berichter“ bei der Berechnung ausgeschlossen und der korrekte Mittelwert von 3,1 Wegen pro Person und Tag ermittelt werden.

Bei der einfachen Variable Mobilität am Stichtag wurden die Kategorien der differenzierten Mobil-Variable folgendermaßen zusammengefasst: Kategorie 1, 2, 3 und 5 ergeben mobil, Kategorie 0, 4 und 6 ergeben nicht mobil.

**Tabelle 9** Häufigkeitsverteilung der differenzierten Variable Mobilität am Stichtag

Mobilität am Stichtag (Differenzierung nach rbW und Ausland)	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
		%	%	%
<b>Gültig</b>				
0 nicht mobil	38.190	12,1	12,1	12,1
1 mobil (nur detailliert erfasste Wege)	246.224	77,8	77,8	89,9
2 mobil (nur rbW)	2.611	0,8	0,8	90,7
3 mobil (detailliert erfasste Wege und rbW)	10.674	3,4	3,4	94,1
4 mobil (nur entfernteres Ausland)	3.760	1,2	1,2	95,3
5 mobil (ohne Wegeangaben)	11.300	3,6	3,6	98,9
6 nicht mobil (potenziell ohne Angabe, analog Kat 5)	2.100	0,7	0,7	99,5
9 Mobilität unbekannt	1.502	0,5	0,5	100,0
<b>Gesamt</b>	<b>316.361</b>	<b>100,0</b>	<b>100,0</b>	

MiD 2017 | Nutzerhandbuch | Quelle: MiD 2017

**Tabelle 10** Häufigkeitsverteilung der einfachen Variable Mobilität am Stichtag

Mobilität am Stichtag	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
		%	%	%
<b>Gültig</b>				
0 nicht mobil	44.050	13,9	13,9	13,9
1 mobil	270.809	85,6	85,6	99,5
9 Mobilität unbekannt	1.502	0,5	0,5	100,0
<b>Gesamt</b>	<b>316.361</b>	<b>100,0</b>	<b>100,0</b>	

MiD 2017 | Nutzerhandbuch | Quelle: MiD 2017

## 6.5 Vergleich mit Vorgängererhebungen

Die vorliegenden MiD-Daten aus den Jahren 2002, 2008 und 2017 sind nur bedingt für eine Analyse der zeitlichen Veränderungen im Verkehr geeignet. Von einer direkten Gegenüberstellung der Ergebnisse der drei Querschnittserhebungen der MiD wird aus den folgenden Gründen abgeraten.

Durch den Zensus im Jahr 2011 hat sich der Hochrechnungsrahmen für die Bevölkerung Deutschlands geändert. Absolutwerte sind daher nur eingeschränkt vergleichbar. Trotz des Ansatzes der Replikation wurden Änderungen im Erhebungsdesign und Verbesserungen bei der Datenaufbereitung vorgenommen. So wurde, um der immer geringer werdenden Teilnahmebereitschaft an Erhebungen entgegenzuwirken, die Methodenvielfalt (Online-, Telefoninterview, schriftlicher Fragebogen) in allen Phasen der Erhebung beibehalten. Schließlich sind die Ergebnisse der Vorläuferstudien durch eine überproportional hohe Beteiligung Pkw-besitzender Haushalte beeinflusst. All dies erschwert die Vergleichbarkeit. Da sich viele Mobilitätskennwerte nur langsam über die Zeit verändern, kann oft nicht eindeutig interpretiert werden, ob die leichte Zu- oder Abnahme von Werten tatsächliche Veränderungen der Mobilität darstellen oder den Unterschieden zwischen den Erhebungen geschuldet sind. Eine konsistente Datenaufbereitung und analoge Gewichtung und Hochrechnung auf vergleichbare Bevölkerungswerte sind daher unabdingbar.

Aktuell wird vom Projektkonsortium der MiD 2017 eine konsistente Zeitreihe mit den Daten aller drei Erhebungsjahre hergestellt. Für einen ausgewählten Variablenbestand wird für alle drei Erhebungsjahre eine analoge Variablenaufbereitung vorgenommen. Darüber hinaus findet eine Neugewichtung der 2002er und 2008er Daten statt. Der Zeitreihendatensatz ist voraussichtlich über die Clearingstelle erhältlich und stellt eine geeignetere Datenbasis für einen Zeitvergleich dar.



#### **Studiendurchführung und Auswertung:**

Folkert Aust	Michael Herter
Marcus Bäumer	Zafirios Kiatipis
Marco Berg	Katja Köhler
Kerstin von der Burg	Brigitte Krämer
Christian Dickmann	Tobias Kuhnimhof
Johannes Eggs	Karen Marwinski
Markus Eichhorn	Claudia Nobis
Bernd Ermes	Manfred Pfeiffer
Robert Follmer	Christian Prinz
Reiner Gilberg	Martina Roggendorf
Dana Gruschwitz	Menno Smid
Seline Günther	Manuel Trittel
Heinz Hautzinger	Barbara Wawrzyniak

#### **Text:**

Katja Köhler  
Claudia Nobis

#### **Layout und Grafik:**

Astrid Blome  
Sigrid Phiesel

#### **1. Auflage**

Bonn im Januar 2019

