

# Standardisierter Mikrodatensatz der Tourismusstatistik (Reisegewohnheiten) für die Jahre 2008 und 2009

Bernhard Meindl, Johanna Ostertag

Letztes Update dieser Dokumentation: Juli 2016

## Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Besonderheiten des Datensatzes</b>	<b>2</b>
2.1	Originaldaten . . . . .	2
2.2	Modifikation der Originaldaten . . . . .	2
<b>3</b>	<b>Geheimhaltung</b>	<b>3</b>
3.1	Software . . . . .	3
3.2	Direkte Identifikationsvariablen . . . . .	4
3.3	Indirekte Identifikationsvariablen . . . . .	4
3.4	Stichprobenziehung . . . . .	4
3.5	Schlüsselvariablen für die Geheimhaltung . . . . .	4
3.6	Lokale Unterdrückung . . . . .	5
3.7	Mikroaggregation . . . . .	5
3.8	Postrandomisierung . . . . .	6
<b>4</b>	<b>Zusammenfassung</b>	<b>6</b>
<b>A</b>	<b>Anhang: Datenbeschreibung des SDS</b>	<b>8</b>

## Wichtiger Hinweis

Diese Dokumentation gilt nur für die anonymisierten Datensätze der Jahre 2008 und 2009, da die Datensätze ab 2012 eine andere Struktur aufweisen und diese ab 2012 in einer jährlich adaptierten Dokumentation beschrieben werden.

## 1 Einleitung

Ein strategisches Ziel der Bundesanstalt STATISTIK AUSTRIA ist es, für den Zweck der wissenschaftlichen Forschung und Lehre ausgewählte Mikrodatsätze der amtlichen

Statistik bereitzustellen. Diese werden in Form von Standardisierten Datensätzen (SDS) über die [Webseite der Statistik Austria](#) zugänglich gemacht. Als SDS werden grundsätzlich Einzeldatenstze bezeichnet, die vor der Veröffentlichung so aufbereitet wurden, dass die gesetzlichen Regelungen hinsichtlich Datenschutz erfüllt sind. Durch die Anwendung statistischer Anonymisierungsverfahren wird das Risiko, dass auf Informationen über einzelne statistische Einheiten rückgeschlossen werden kann, minimiert. Des Weiteren müssen potentielle Datennutzer Nutzungsbestimmungen akzeptieren, bevor der Zugriff auf einen SDS ermöglicht wird.

In dieser Arbeit wird die Erstellung eines anonymisierten Datensatzes aus dem Bereich der Tourismusstatistik beschrieben. Der SDS liefert ab dem 1. Quartal 2008 vierteljährliche Informationen zu den Reisegewohnheiten der Österreicher und Österreicherinnen. Die notwendigen Anonymisierungsschritte wurden so gewählt, dass der resultierende anonymisierte Mikrodatensatz höchstmöglichen Informationsgehalt bei gleichzeitig möglichst niedrigem Identifikationsrisiko aufweist.

Der anonymisierte Datensatz kann sowohl als reine Text-Datei (csv-File zum einfachen Import etwa in Microsoft Excel) als auch als [R-Datensatz](#) bezogen werden.

## 2 Besonderheiten des Datensatzes

### 2.1 Originaldaten

Der SDS über die Reisegewohnheiten der Österreicher und Österreicherinnen basiert auf vierteljährlichen Stichprobenerhebungen, die laufend mit dem Ziel durchgeführt werden, das nationale Reiseverhalten der im Inland wohnhaften Personen ab 15 Jahren abzubilden. Je Quartal werden im Rahmen dieser Erhebung rund 3.500 ausgewählte, im Inland wohnhafte Personen ab 15 Jahren, verteilt über Österreich, telefonisch befragt (Grundgesamtheit von rd. 6.98 Mio.). Die Teilnahme an der Erhebung ist freiwillig. Der Auswahlrahmen für die proportional geschichtete Stichprobe ist das Zentrale Melderegister (ZMR). Die Telefonnummern werden dem öffentlichen Telefonbuch entnommen. Zusätzlich wurden Informationen über das Alter, das Wohnsitzbundesland sowie das Geschlecht der Reisenden und Nicht-Reisenden aus dem Zentralen Melderegister hinzugefügt. Aus den umfangreichen Informationen des Datenbestandes wurden schlussendlich insgesamt 55 Variablen ausgewählt, die - zusammen mit den durchgeführten Umkodierungen - in Anhang (A) beschrieben sind.

Zu erwähnen ist, dass sich die Variablenbezeichnungen mit den Bezeichnungen aus dem Fragebogen decken. Damit soll die Arbeit mit den Mikrodaten erleichtert werden.

### 2.2 Modifikation der Originaldaten

Es wird nun kurz beschrieben, auf welche Art und Weise bestehende Variablen aus dem authentischen Datenbestand für den SDS modifiziert beziehungsweise umkodiert wurden, um der Geheimhaltung Rechnung zu tragen. Nachfolgend werden alle Variablen, die umkodiert wurden, aufgelistet.

- *Frage0\_1* (höchste abgeschlossene Schulbildung)

- *Frage0\_2* (Berufstätigkeit)
- *Frage0\_4* (Anzahl von Kindern unter 15 im Haushalt)
- *Frage0\_5* (Alter)
- *vA2\_2\_1*, *vB2\_2\_1*, *vC2\_2\_1* (hauptsächlicher Reisezweck)
- *vA2\_5* (Reise in welchem ausländischen Staat überwiegend verbracht)
- *vA2\_6*, *vB2\_6* (Reiseorganisation)
- *vA2\_8*, *vB2\_8* (vorwiegende Unterbringung)
- *vB2\_5* (Reise in welchem Bundesland überwiegend verbracht)
- *vC2\_3* (Tagesreise in welchen ausländischen Staat)

Aus Anhang (A) ist die tatsächliche Kodierung aller im SDS enthaltenen Variablen ersichtlich.

Zu erwähnen ist weiters, dass der Datenbestand aus Gründen der Übersichtlichkeit in zwei Dateien aufgesplittet wurde.

- **Personen-File:** dieses File enthält alle persönlichen Informationen über die Reisenden und Nicht-Reisenden sowie alle Variablen, die nicht spezifisch für durchgeführte Auslands-, Inlands- oder Tagesreisen sind, sowie die zur Hochrechnung notwendige GewichtsvARIABLE.

Im letzten Quartal jedes Jahres werden die für die Stichprobe ausgewählten Personen ausserdem nach ihrer Reisetätigkeit im gesamten vergangenen Jahr gefragt. Diese Fragen (Variablenblock *vDx*) ist ebenfalls im Standardisierten Datensatz vorhanden.

- **Reisen-File:** diese Datei enthält Informationen über alle Reisen, die von den Reisenden, die im Personen-File enthalten sind, durchgeführt wurden.

Zu bemerken ist, dass das Personen-File und das Reisen-File über einen eindeutigen Identifier - die Variable *ID* - zusammengeführt werden können.

## 3 Geheimhaltung

Es wird nun die Anonymisierungsprozedur beschrieben, die durchgeführt wurde um aus dem authentischen Datenbestand ab dem 1. Quartal 2008 einen SDS-File zu erzeugen.

### 3.1 Software

Die Anonymisierungsprozedur mit der freien Statistiksoftware *R* ([R Development Core Team, 2006](#)) sowie dem von Statistik Austria entwickelten und frei verfügbaren R-Package *sdcMicro* ([Templ, 2007](#)) (statistical disclosure control for **m**icrodata) durchgeführt. Das Package kann von den Servern des R Comprehensive Archive Network ([CRAN](#)) heruntergeladen werden. *sdcMicro* weist wesentliche Vorteile gegenüber der für Geheimhaltung von Mikrodaten empfohlenen "Standardsoftware"  $\mu$ -Argus auf. Ausserdem wird *sdcMicro* ständig aktualisiert, verbessert und weiterentwickelt.

### 3.2 Direkte Identifikationsvariablen

Direkte Identifikationsvariablen ermöglichen es einem Datenangreifer bestimmte Personen in einem Datensatz eindeutig zu identifizieren. Solche Variablen müssen daher aus einem Standardisierten Datensatz entfernt werden um den Geheimhaltungsanforderungen gerecht werden zu können. Als Beispiel für eine direkte Identifikationsvariable könnte etwa die Sozialversicherungsnummer genannt werden, die von einem Angreifer dazu genutzt werden könnte, eine Person im Standardisierten Datensatz eindeutig zu identifizieren.

Im Datenbestand, der diesem SDS zugrundeliegt, existiert keine direkte Identifikationsvariable.

### 3.3 Indirekte Identifikationsvariablen

Kann durch Kombination mehrerer (meist kategorialer) Variablen eine Person eindeutig im Datensatz identifiziert werden, so werden diese Variablen als indirekte Identifikationsvariablen bezeichnet. Wichtig zu bemerken ist, dass keine indirekte Identifikationsvariable für sich selbst zur eindeutigen Identifizierung einer Person im Datensatz ausreicht.

Als indirekte Identifikationsvariablen in den Daten über die Reisegewohnheiten der Österreicher und Österreicherinnen können beispielsweise die höchste abgeschlossene Schulbildung (*Frage0\_1*), Information über die berufliche Stellung (*Frage0\_2*), das Alter (*Frage0\_5*), Geschlecht (*Frage0\_6*) oder Information über das Wohnsitzbundesland eines Befragten (*Frage0\_7*) herangezogen werden. Kategorielle Variablen können vergrößert oder umkodiert werden um das Risiko einer Reidentifikation einer Person gering zu halten. Letztlich kann es sein, dass in den indirekten Identifikationsvariablen wenige Werte unterdrückt bzw. gelöscht werden müssen um höchstmögliche Anonymität gewährleisten zu können.

### 3.4 Stichprobenziehung

Der erste Schritt bei der Erstellung eines Standardisierten Datensatzes über die Reisegewohnheiten der Österreicher und Österreicherinnen besteht darin, für jedes im SDS vorhandene Quartal eine Zufallsstichprobe mit Zurücklegen aller Befragten zu ziehen. Dabei wird sichergestellt, dass die Anzahl der für die Zufallsstichprobe eines Quartals ausgewählten Einheiten exakt der Anzahl der Befragten in den Originaldaten des entsprechenden Quartals entspricht. In einem zweiten Schritt werden die Reisen einer ausgewählten Person genau so oft reproduziert, wie häufig ein Befragter für die Zufallsstichprobe ausgewählt wurde.

### 3.5 Schlüsselvariablen für die Geheimhaltung

Indirekte Identifikationsvariablen, deren Ausprägungskombinationen ein Angreifer verwenden könnte, um eine eindeutige Identifikation einer Person im Datensatz vorzunehmen, werden als Schlüsselvariable bezeichnet. Für den vorliegenden Datenbestand wurden folgende Variablen als *Schlüsselvariablen* definiert.

- *Frage0\_1* (höchste abgeschlossene Schulbildung, 5 Ausprägungen)
- *Frage0\_2* (Berufstätigkeit, 8 Ausprägungen)
- *Frage0\_5* (Alter, 6 Ausprägungen)
- *Frage0\_6* (Geschlecht, 2 Ausprägungen)
- *Frage0\_7* (Wohnsitzbundesland, 9 Ausprägungen)

Eine Möglichkeit, einen sicheren SDS mit hohem Analysepotential zu erhalten, besteht darin, einzelne Werte in den Schlüsselvariablen zu löschen um schliesslich k-Anonymität gewährleisten zu können.

### 3.6 Lokale Unterdrückung

Im Zuge der Anonymisierungsprozedur wurde durch gezielte Sperrungen von einzelnen Ausprägungen in den Schlüsselvariablen erreicht, dass jeder Ausprägungskombination in den definierten Schlüsselvariablen zumindest 3 Personen zugeordnet werden können. Dies wird auch als *3-Anonymity* bezeichnet.

Es ist zu erwähnen, dass 3-Anonymität für jedes Quartal gewährleistet wird, da die quartalsweisen Daten unabhängig voneinander zu betrachten sind. Tabelle (1) zeigt die Anzahl der notwendigen Sperrungen in den Schlüsselvariablen pro Quartal.

Jahr	Quartal	Frage0_1	Frage0_2	Frage0_5	Frage0_6	Frage0_7
2008	Q1	49	300	194	0	4
2008	Q2	60	262	154	0	4
2008	Q3	55	267	170	0	2
2008	Q4	55	283	168	0	2
2009	Q1	51	258	177	0	0
2009	Q2	64	286	150	0	0
2009	Q3	55	298	132	0	0

Tabelle 1: Sperrungen in den Schlüsselvariablen um 3-Anonymität zu gewährleisten.

### 3.7 Mikroaggregation

Unter Umständen besteht die Möglichkeit dass ein Datenangreifer ihm bekannte Informationen über einen Wert einer numerischen Variable heranzieht, um eine Person im Datensatz erfolgreich zu identifizieren. Insbesondere "Ausreisser" in numerischen Variablen können in Verbindung mit Informationen über andere Schlüsselvariablen dazu verwendet werden, eine positive Identifizierung zu erreichen.

*Mikroaggregation* numerischer Variablen bietet zusätzlichen Schutz gegen Reidentifizierungsversuche. Mikroaggregation bedeutet grundsätzlich, dass möglichst "ähnliche" Objekte in einem ersten Schritt gruppiert werden. In einem zweiten Schritt werden schliesslich

die Ausprägungen einer numerischen Variablen der gewählten Personen durch eine Statistik ersetzt. Bei der verwendeten Statistik handelt es sich oftmals um den Mittelwert. Durch die Mikroaggregation numerischer Variablen wird sichergestellt, dass jede einzelne Ausprägung mehrfach im Datensatz auftritt. Die *Verschmutzung* der Daten selbst durch das Mikroaggregationsverfahren ist gering. Dies ergibt sich aus dem Vergleich univariater bzw. multivariater Statistiken der Originaldaten mit den mikroaggregierten Daten.

Aus Anhang (A) geht hervor, welche Variablen des Standardisierten Datensatzes über die Reisegewohnheiten der Österreicher und Österreicherinnen mikroaggregiert wurden. Bei diesen Variablen wird sichergestellt, dass jeder Wert einer mikroaggregierten Variablen zumindest 3-fach in dieser Variable vorkommt. Als Mikroaggregationsmethode wurde *individual ranking* ausgewählt, da dieses Verfahren auch für Variablen, die fehlende Werte enthalten, verwendet werden kann. Alle Variablen, die mikroaggregiert werden, werden zuerst unabhängig voneinander sortiert. In einem zweiten Schritt findet die Mikroaggregation selbst statt. Den Abschluss des Verfahrens bildet das Zurücksortieren der Ausprägungen nach der ursprünglichen Ordnung.

### 3.8 Postrandomisierung

Um Variablen, die für die Datenanalyse absolut notwendig sind und aus diesem Grunde nicht aus dem Datensatz entfernt oder vergrößert werden können zu schützen, besteht die Möglichkeit, ein Postrandomisierungsverfahren anzuwenden. Die grundlegende Idee bei PRAM ist es, dass grundsätzlich jede Ausprägung in einer Variable entweder unverändert bleibt oder in eine andere Kategorie wechseln kann. Hinter dem Verfahren stehen Übergangsmatrizen, die den Zufallsprozess steuern. Die tatsächlich gewählten Wahrscheinlichkeiten für das Wechseln oder Nicht-Wechseln können publiziert werden. Es ist jedoch zu erwähnen, dass die Wahrscheinlichkeit für einen Verbleib in der jeweiligen Kategorie aus den Originaldaten hoch ist um die innere Struktur des Datensatzes möglichst wenig zu verändern.

Die Variablen des vorliegenden Datensatzes, auf die Postrandomisierungstechniken angewendet wurden, ist aus Anhang (A) ersichtlich.

## 4 Zusammenfassung

Die Aufbereitung und Bereitstellung sensibler Mikrodaten - wie etwa Steuerdaten - für wissenschaftliche Forschung und Lehre ist ein komplexer Prozess. Insbesondere muss das Hauptaugenmerk beim Erstellen des Datensatzes auf die Anonymisierung der Daten gelegt werden um die gegebenen rechtlichen Anforderungen zu erfüllen.

Da allerdings ausschliesslich quartalsweise Zufallsstichproben mit Zurücklegen mit der gleichen Anzahl an Personen wie aus den Originaldaten für die Veröffentlichung des SDS aufbereitet wurden, kann sich ein Datenangreifer auch bei vermeintlicher positiver Identifizierung einer Person aufgrund mehrerer Variablen nicht sicher sein, ob die identifizierte Person überhaupt im standardisierten Datensatz enthalten ist. Durch die weiters angewandten Anonymisierungsverfahren wie die Aggregation beziehungsweise das Umkodieren

kategorieller Variablen, dem Ersetzen kritischer Werte in den Schlüsselvariablen durch *missings*, durch Mikroaggregation numerischer Variablen sowie Postrandomisierung wurde erreicht, dass das Reidentifikationsrisiko aller im SDS verbleibenden Daten sehr gering ist. Allerdings ist anzumerken, dass es 100%-igen Schutz vor Aufdeckung sensibler Information nicht geben kann. Ein Restrisiko bleibt also bestehen.

Beim Erstellen des Standardisierten Datensatzes über die Reisegewohnheiten der Österreicher und Österreicherinnen wurde darauf geachtet, trotz der notwendigen Anonymisierung der Daten das hohe Analysepotential der Daten zu erhalten. Der vorliegende standardisierte Datensatz wird diesem Anspruch gerecht. Des Weiteren ist es möglich, den SDS in Zukunft mit aktualisierten und neuen Daten zu erweitern.

## Literatur

R Development Core Team (2006). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Templ, M. (2007). *scdMicro: A package for statistical disclosure control in R*. In ISI 2007, Lissabon.

## A Datenbeschreibung

In Tabelle (2) werden die Variablen beschrieben, die im SDS Personenfile, in Tabelle (3) jene, die im SDS Reisenfile enthalten sind. In Spalte 3 ist überblicksmässig die für diese Variable angewandte Geheimhaltungaktion beschrieben. Als Unterstützung wurden die Variablennamen verschiedenfärbig markiert, wobei Variablen, die in schwarzer Schrift aufscheinen, nicht verändert wurden. Variablen, die mit **blau** gekennzeichnet sind wurden verändert oder postrandomisiert und **rot** bedeutet, dass diese Variable mikroaggregiert wurde.

Name	Beschreibung	Aktion	Kodierung
ID	eindeutiger Identifier	durchnummeriert	fortlaufend
Jahr	Berichtsjahr		2008, 2009
Quartal	Berichtsquartal		1-4
Frage0_1	höchste abgeschlossene Schulbildung	vergrößert, KeyVar	1: kein Pflichtschulabschluss 2: Pflichtschule, Polytechnische Schule 3: Lehre, Lehrabschluss, BMS 4: AHS 5: Kollegs, Hochschullehrgänge, Uni, FH,..
Frage0_2	Berufstätigkeit	vergrößert,KeyVar	1: selbständig und mithelfend 2: Arbeiter 3: Angestellte 4: Beamte, Vertragsbedienstete 5: Lehrlinge, in Ausbildung befindlich 6: Pensionisten 7: ausschliesslich Haushaltsführend 8: Arbeitslose/Sonstige
Frage0_3	Anzahl der Personen im Haushalt	vergrößert	1-5: tatsächliche Anzahl 6: 6 und mehr Personen
Frage0_4	Anzahl der Personen unter 15 im Haushalt	vergrößert	0: keine Person unter 15 1: mindestens eine Person unter 15
Frage0_5	Alter	vergrößert,KeyVar	1: 15-24 Jahre 2: 25-34 Jahre 3: 35-44 Jahre 4: 45-54 Jahre 5: 55-64 Jahre 6: 65 Jahre und älter
Frage0_6	Geschlecht	KeyVar,PRAM	1: männlich 2: weiblich
Frage0_7	Wohnsitzbundesland	KeyVar,PRAM	1: Burgenland 2: Niederösterreich 3: Kärnten 4: Oberösterreich 5: Salzburg 6: Steiermark 7: Tirol 8: Vorarlberg 9: Wien
vA1	Urlaubsreise und/oder Geschäftsreise ins Ausland		1: Ja 2: Nein
vA2_1	Anzahl der Auslandsreisen		Häufigkeit
vB1	Urlaubsreise und/oder Geschäftsreise ins Inland		1: Ja 2: Nein
vB2_1	Anzahl der Inlandsreisen		Häufigkeit
vC1	Tagesausflug/eintägige Dienstreise		1: Ja 2: Nein
vC1_1	Tagesausflug/eintägige Dienstreise (Inland)		1: Ja 2: Nein
vC1_1_1	Anzahl der Inlandstagesreisen (Freizeit)		Häufigkeit
vC1_1_2	Anzahl der Inlandstagesreisen (Geschäft)		Häufigkeit
vC2_1	Tagesausflug/eintägige Dienstreise (Ausland)		1: Ja 2: Nein
vC2_1_1	Anzahl der Auslandstagesreisen		Häufigkeit
vD1	Urlaubs- und/oder Geschäftsreise ins Ausland (letztes Jahr)		wie vA1



## Anhang: Datenbeschreibung des SDS

vD1_1	Kurzurlaubsreise ins Ausland		wie vD1
vD1_2	Haupturlaubsreise ins Ausland		wie vD1
vD1_3	Geschäftsreise ins Ausland		wie vD1
vD2	Urlaubs- und/oder Geschäftsreise ins Inland (letztes Jahr)		wie vA1
vD2_1	Kurzurlaubsreise ins Inland		wie vD2
vD2_2	Haupturlaubsreise ins Inland		wie vD2
vD2_3	Geschäftsreise ins Inland		wie vD2
Gewicht	Hochrechnungsgewicht		numerisch

Tabelle 2: Beschreibung der Variablen aus dem Personen-File des SDS über Reisegewohnheiten

Name	Beschreibung	Aktion	Kodierung
ID	eindeutiger Identifier	durchnummeriert	fortlaufend
vA2_2	Art der Inlandsreise		1: Urlaubsreise 2: Geschäftsreise
vA2_2_1	Reisezweck	vergrößert	1: Strand/Badeaufenthalt 2: Aktivurlaub 3: Erholungsurlaub 4: Wellness-, Gesundheitsurlaub 5: Verwandten/Bekanntebesuche 6: (nicht-berufliche) Ausbildung 7: Kultur und Besichtigung 8: Shopping, Sonstiges (Urlaub) 9: Kongresse/Messen, Sonstiges (Geschäft)
vA2_2_2	Pauschalreise		1: Ja 2: Nein
vA2_3	Abreisemonat		1: Jänner ... 12: Dezember
vA2_4	Anzahl der Nchtigungen		Häufigkeit
vA2_5	Staat in dem Reise Vuberwiegend verbracht wurde	vergrößert,PRAM	1-30: europ. Länder laut Codebook 31: Bulgarien 32: Russland 33: Afrika 34: Amerika 35: Asien 36: Australien 37: Sonstige
vA2_6	Reiseorganisation	vergrößert	1: selbst, privat 2: ganz/teilweise durch Reisebüro
vA2_7	Verkehrsmittel		1: Flugzeug 2: Schiff 3: Zug 4: Bus/Reisebus 5: PKW 6: Sonstige Verkehrsmittel
vA2_8	vorwiegende Unterbringung	vergrößert	1: Hotels u. ähnliche Betriebe 2: Private Unterkünfte (Bezahlung) 3: Private Unterkünfte (Gratis) 4: sonstige Beherbergungsbetriebe
vA2_9	Anzahl der Personen bei der Reise		Häufigkeit
vA2_10_7	Gesamtausgaben für die Reise	mikroaggregiert	numerisch
vB2_2	Art der Auslandsreise		wie vA2_2
vB2_2_1	Reisezweck	vergrößert	wie vA2_2_1
vB2_3	Abreisemonat		wie vA2_3
vB2_4	Anzahl der Nächtigungen		Häufigkeit wie vA2_4
vB2_5	Bundesland der Reise (überwiegend)		1: Burgenland 2: Niederösterreich 3: Kärnten 4: Oberösterreich 5: Salzburg 6: Steiermark 7: Tirol

## Anhang: Datenbeschreibung des SDS

			8: Vorarlberg 9: Wien 9: nicht zuordenbar
vB2_6	Reiseorganisation	vergrößert	wie vA2_6
vB2_7	Verkehrsmittel		wie vA2_7
vB2_8	vorwiegende Unterbringung	vergrößert	wie vA2_8
vB2_9	Anzahl der Personen bei der Reise		Häufigkeit
vB2_10_1	Gesamtausgaben für die Reise	mikroaggregiert	numerisch
vC2_2	Art der Auslandstagerese		wie vA2_
vC2_2_1	Reisezweck	vergrößert	wie vA2_2_1
vC2_2_2	Abreisemonat		wie vA2_3
vC2_3	Staat in den Tagesreise führte	vergrößert, PRAM	1: Deutschland 2: Italien 3: Schweiz/Liechtenstein 4: Slowakei 5: Slowenien 6: Tschechien 7: Ungarn 8: nicht-Nachbarländer
vC2_4	Verkehrsmittel		wie vA2_7
vC2_5	Anzahl der Personen bei der Reise		Häufigkeit
vC2_6_1	Gesamtausgaben für die Reise	mikroaggregiert	numerisch

Tabelle 3: Beschreibung der Variablen aus dem Reisen-File des SDS über Reisegewohnheiten