

# Anonymisierungsbeschreibung des Scientific-Use-File der Umsatzsteuerstatistik 2000

## 1 Vorbemerkung

Das Scientific-Use-File der Umsatzsteuerstatistik 2000 basiert als Sekundärstatistik auf Angaben, die im Rahmen der Umsatzsteuervoranmeldung bei den Finanzbehörden anfallen. Die Durchführung der Umsatzsteuerstatistik als Bundesstatistik ist im Gesetz über Steuerstatistiken (StStatG) vom Oktober 1995 geregelt. Die Umsatzsteuerstatistik liefert für fast alle Wirtschaftsbereiche Daten und erfasst mit wenigen Ausnahmen sämtliche Unternehmen.

Die faktische Anonymisierung der Umsatzsteuerstatistik 2000 wurde in erster Linie durch Vergrößerung bei den qualitativen Merkmalen durchgeführt. Der Einsatz von so genannten datenverändernden Verfahren wurde so gewählt, dass es nur dann zu Veränderungen in den Merkmalsausprägungen kommt, wenn die Anonymität der Merkmalsträger auf konventionellem Wege nicht zu gewährleisten war.

## 2 Anonymisierung der qualitativen Merkmale

Aufgrund der notwendigen Anonymisierungsmaßnahmen wurden folgende qualitative Merkmale in ihren Ausprägungen wie folgt reduziert, wobei die Maßnahmen unabhängig von der Unternehmensgröße gelten:

- Die *Regionalangabe* wurde von der Gemeindeebene auf Neue Bundesländer inkl. Berlin und alte Bundesländer ohne Berlin vergrößert.
- Die *Wirtschaftsklassifizierung* (WZ93) ist in unterschiedlicher Tiefengliederung abhängig von den Besetzungszahlen ausgewiesen (zur Einteilung vgl. Datensatzbeschreibung). Dabei werden z.T. auch neue zusammengefasste Positionen gebildet.
- Aus dem Merkmal *Beginn der Steuerpflicht* wurde das Merkmal *Neugründung* mit den Ausprägung 1 = ja und 0 = nein gebildet. Bei Unternehmen mit mehr als 100 Mio. € Umsatz, wurde das Merkmal generell auf 0 gesetzt. Von über 150.000 als Neugründungen gekennzeichneten Unternehmen haben 118 Unternehmen einen Umsatz von über 100 Mio. €, so dass bei diesen dieses Merkmal auf 0 gesetzt und damit die Information unterdrückt ist.
- Das Merkmal *Rechtsform* besitzt nach einer Vergrößerung noch folgende Ausprägungen:
  - Personengesellschaften,
  - Kapitalgesellschaften,
  - Erwerbs- und Wirtschaftsgenossenschaften sowie Betriebe gewerblicher Art von Körperschaften des öffentlichen Rechts und
  - sonstige Rechtsformen.

## 3 Anonymisierung der quantitativen Merkmale

Die quantitativen Merkmale der Umsatzsteuerstatistik wurden mikroaggregiert. Bei einer Mikroaggregation werden drei benachbarte Werte durch ihre Mittelwerte ersetzt. Liegen die benachbarten Merkmalsausprägungen dicht beieinander, so werden sich die im Scientific-Use-File veröffentlichten Mittelwerte von den Ursprungswerten nicht oder nur sehr gering unterscheiden. Dies ist in der Regel bei den Merkmalsausprägungen der Umsatzsteuerstatistik der Fall.

Bei der Umsatzsteuerstatistik besteht das Konzept aus zwei Maßnahmen der Mikroaggregation. Zum einen eine für jedes Merkmal getrennte Mikroaggregation für alle Unternehmen (Mikro21G).

Dies gilt als die schwächste Form der Mikroaggregation und verändert die Merkmalsausprägungen nur geringfügig. Da durch diese getroffenen Maßnahmen die regionalen Branchenmarktführer<sup>1</sup> aber noch nicht genügend geschützt werden, wurde zum anderen eine punktuelle Mikroaggregation durchgeführt. Dabei wurden nur speziell die jeweiligen drei regionalen Marktführer einer Branche mikroaggregiert, wobei das Merkmal *Lieferungen und Leistungen* (Umsatz) das bestimmende Merkmal ist (punktuelle Mikro1G). Da dieses Verfahren getrennt nach den beiden Regionen angewandt wurde, sind insgesamt 408 Merkmalsträger davon betroffen. Diese sind als solche mit einem zusätzlichen Hinweis kenntlich gemacht

Die folgende Tabelle fasst die getroffenen Maßnahmen sowohl für die qualitativen als auch für die quantitativen Merkmale kurz zusammen

**Tabelle: Anonymisierungsmaßnahmen bei der Umsatzsteuerstatistik 2000**

Qualitative Merkmale		
Merkmal	Original	Anonymisiert
Regional-kennung	AGS-Gemeindeebene	1 = Alte Bundesländer 2 = Neue Bundesländer inkl. Berlin
Wirtschafts-klassifikation	Klassifikation der Wirtschaftszweige - Ausgabe 1993 (WZ 1993)	Unterschiedliche Tiefengliederung, abhängig von den Besetzungszahlen
Beginn der Steuerpflicht	1=Beginn vor dem 1.1. des Vorj.; Ende nicht vor d. 1.1. d. Folgejahres 2=Beginn im Vorjahr; Ende nicht vor dem 1.1. des Folgejahres 3=Beginn im Berichtsjahr; Ende nach dem 31.12. des Berichtsjahres 4=Beginn vor dem 1.1. d. Berichtsjahres; Ende im Berichtsjahr 7=Beginn im Berichtsjahr; Ende im Berichtsjahr 8=wie 1 ohne Vorjahresumsätze	0 = keine Neugründung 1 = Neugründung
Rechtsform	1 = Einzelunternehmen 2 = OHG 3 = KG 4 = AG 5 = GmbH 6 = Erwerbs- und Wirtschaftsgenossenschaften 7 = Betriebe gewerblicher Art von Körperschaften des öff. Rechts 8 = Sonstige Rechtsformen	1 = Personengesellschaften 2 = Kapitalgesellschaften 3 = Erwerbs- und Wirtschaftsgenossenschaften sowie Betriebe gewerblicher Art von Körperschaften des öff. Rechts 4 = sonstige
Quantitative Merkmale		
Unternehmen	Original	Anonymisiert
Alle Unternehmen	Werte aus der Umsatzsteuervoranmeldung	Mikroaggregation der Werte für jedes Merkmal getrennt
408 regionale Branchenmarktführer		Zusätzliche punktuelle Mikroaggregation

<sup>1</sup> Mit den regionalen Branchenmarktführern sind die jeweils drei Unternehmen gemeint, die in den jeweiligen Branchen innerhalb ihrer Region die höchsten Umsätze erzielen.