

## Nutzungskonzept Business-Tax-Panel

Das Business-Tax-Panel führt 6 Ertrags- und Umsatzsteuerstatistiken sowie das Unternehmensregister (siehe Tabelle 1) im Querschnitt und über die Zeit zusammen. Für den Zeitraum 2013 bis 2019 liegen über 66 Mio. Einheiten vor. Das Panel wird jährlich um das aktuellste Statistikjahr erweitert.

Darüber hinaus ist der Merkmalskranz der zugrundeliegenden Steuerstatistiken unterschiedlich breit. Aus den verschiedenen Vordrucken und darin verwendeten Kennzahlen werden über 2.700 Merkmale erfasst.

Tabelle 1: Datengrundlagen des Business-Tax-Panels

Statistiken	EVAS
Gewerbesteuerstatistik	73511
Körperschaftsteuerstatistik	73211
Umsatzsteuerstatistik (Voranmeldungen)	73311
Statistik über die Personengesellschaften und Gemeinschaften	73121
Umsatzsteuerstatistik (Veranlagungen)	73321
Einnahmenüberschussrechnung (aus Lohn- und Einkommensteuerstatistik)	73111
Unternehmensregister	52111

Zu beachten gilt, dass nicht jede Beobachtung in jeder der Statistiken oder in jedem Jahr enthalten ist. Dies liegt an unterschiedlichen Erfassungsgrenzen (beispielsweise die Erfassungsgrenze für die Umsatzsteuer-Voranmeldung von 17.500 Euro an Lieferungen und Leistungen für die Statistikjahre 2007 bis 2019) oder der gewählten Rechtsform (beispielsweise müssen Kapitalgesellschaften immer bilanzieren, Pflicht zur doppelten Buchführung, und dürfen keine Einnahmenüberschussrechnung machen). Wird beispielsweise Angaben aus der Gewerbe- und Körperschaftsteuer benötigt sind lediglich noch 8.435.399 Beobachtungen betroffen (siehe Tabelle 2, sowie weitere Übersicht im Metadatenreport Teil 2).

Tabelle 2: Mögliche Zusammenführungswünsche für den Zeitraum 2013-2019

Gewerbe- steuer	Körper- schaft- steuer	Umsatzsteuer- Voranmeldung	Statistik über Personen- gesellschaften und Gemein- schaften	Umsatzsteuer- Veranlagung	Statistisches Unternehmens- register	Einnahmen- überschuss- rechnung	Anzahl
		x			x		22.091.952
				x		x	21.747.863
x	x						8.435.399
x	x			x			6.332.178
			x	x			4.455.929
x		x	x				1.933.709

Die Datensatzbeschreibungen zu den jeweiligen Steuerstatistiken enthalten einige Informationen um vorab zu prüfen, welche Analysen möglich sind.

Grundlegend kann gesagt werden, dass die Variablen immer mit dem Anfangsbuchstaben der jeweiligen Statistik beginnen, dann eine ein Ordnung der Variablen nach Ordnungsmerkmal oder Kennzahl erfolgt an die sich dann der Sachbereich und die Kennzahl anschließt. Je nach Statistik können Sonderheiten auftreten, weshalb es von Vorteil ist den Hinweistext auf dem ersten Tabellenblatt der Datensatzbeschreibung zu beachten.

Beispiel: e\_c51106

"e" - Variable aus der Einnahmenüberschussrechnung

"c" - Kennzahl

51 - Sachbereich 51 "Anlage SE - Sonderbetriebseinnahmen"

106 - Kennzahl 106 "Private Kfz-Nutzung"

Bei der Eingrenzung nach Statistiken wird auch die Variablenauswahl um die nicht angefragten Statistiken reduziert. Zusätzlich ist zu prüfen, welche der Variablen für die Analyse benötigt werden und inwiefern diese Variablen entsprechend befüllt ist. Aus den Datensatzbeschreibungen kann für jedes Jahr entnommen werden, erstens, ob die Kennzahl in dem Jahr enthalten ist und wie die Befüllung (0 - nicht belegt, 1 -weniger als 1% belegt, 5 - weniger als 5% belegt oder „“(leer) in mehr als 5% belegt) im entsprechenden Jahr ist. Variable die in keinem der angegebenen Berichtsjahre Werte enthalten wurde bereits aus dem Datensatz entfernt (abzulesen an der Spalte „Hinweis“ in den jeweiligen Datensatzbeschreibungen).

Folgende Beispiele können aus Tabelle 3 entnommen werden:

Die Kennzahlen wurden lediglich in den Jahren 2018 und 2019 verwendet.

Kennzahl e\_c65357 enthält in 2018 in weniger als 5% der Beobachtung einen anderen Wert als „0“ oder „.“, in 2019 sogar in weniger als 1% der Beobachtungen.

Kennzahl e\_c55863 ist nicht im Datensatz enthalten, da die Kennzahl entweder nicht verwendet wurde (2017) oder nicht befüllt ist (2018 und 2019).

Anhand dieser Informationen kann entnommen werden inwiefern welche Kennzahlen sich für die Analysen eignen.

Tabelle 3: Beispiel Befüllung von Kennzahlen aus der Einnahmenüberschussrechnung

Variable	Beschreibung	Hinweis	2017	2018	2019	Filled2 017	Filled 2018	Filled 2019
e_c65357	Höchstbetrag der Nichtabziehbaren Schuldzinsen, mindestens 0 (in Cent)			e_c65357	e_c65357		5	1

e_c55863	Gesamtbetrag der bereits berücksichtigten Beträge, für die Teilfreistellungen nach InvStG gelten (in Cent)	Die Variable enthält in keinem der angegebenen Berichtsjahre Werte und wurde daher aus dem Datensatz entfernt.		e_c55863	e_c55863		0	0
----------	--	--	--	----------	----------	--	---	---

Aufgrund der jährlichen Erweiterung, der Anzahl an verschiedenen Steuerstatistiken und der Menge an verfügbaren Merkmalen sowie unterschiedlichen Belegungsquote ist eine Eingrenzung der

- (1) Steuerstatistiken  
und
- (2) Variablen  
zielführend.

Diese Eingrenzungen dient zusätzlich dazu den IT-technischen Gegebenheiten Rechnung zu tragen und erhöht die Laufzeitgeschwindigkeit der Analysen.

Bei Nutzung der kontrollierten Datenfernverarbeitung entstehen für die Nutzung des ersten Gastwissenschaftsarbeitsplatzes keine zusätzlichen Kosten. Die Ergebnisfreigabe erfolgt dabei ausschließlich über die kontrollierte Datenfernverarbeitung. Die Nutzung weiterer Gastwissenschaftsarbeitsplätze oder eines Gastwissenschaftsarbeitsplatzes mit Ergebnisfreigabe ist kostenpflichtig. Am Gastwissenschaftsarbeitsplatz kann, bedingt durch die benötigte Speicherplatzgröße, die Nutzung einer Stichprobe erforderlich sein.