

Die Anonymisierung der FOBS Daten - Der standardisierte FOBS Datensatz

Bernhard Meindl und Matthias Templ

7. März 2008

Inhaltsverzeichnis

1	Einleitung	1
2	Software	2
3	Anonymisierung des Datensatzes	2
3.1	Direkte Identifikationsvariablen	2
3.2	Indirekte Identifikationsvariable	2
3.3	Schlüsselvariablen für die Geheimhaltung	3
3.4	Umkodierung	4
3.5	Lokale Unterdrückungen	5
4	Folgerungen	6
5	Anhang: Datenbeschreibung FOBS-Daten	8

1 Einleitung

Um die Zusammenarbeit zwischen offizieller Statistik einerseits und wissenschaftlicher Forschung und Lehre andererseits auszubauen und zu fördern sowie dem steigenden Interesse nach Mikrodaten gerecht zu werden, stellt die STATISTIK AUSTRIA interessierten Datennutzern [über ihren Webauftritt](#) ausgewählte Mikrodatensätze zur Verfügung.

Um das Angebot an Mikrodatensätzen für Forschung und Lehre zu erweitern wurde der Datensatz über die [Faktoren des Geschäftserfolgs von neu gegründeten Unternehmen](#) (**F**actors of **B**usiness **S**uccess) oder kurz FOBS anonymisiert. Die FOBS-Erhebung wurde im Oktober 2005 und April 2006 als Stichprobenerhebung durchgeführt und lieferte Meldungen von mehr als 1.700 österreichischen Jungunternehmen. Für Details zur Erhebung selbst sei auf die entsprechende [Publikation](#) verwiesen.

Die FOBS-Daten werden als Standardisierter Datensatz (SDS) angeboten. Der Standardisierte Datensatz kann von Lehrenden und Studierenden nach erfolgter Registrierung und unter Einhaltung der Nutzungsbestimmungen von der [Webseite der ST.AT](#) abgerufen werden.

Um dem österreichischen Bundesstatistikgesetz Rechnung zu tragen, können auch standardisierte Datensätze nur nach einer streng gehandhabten Anonymisierungsstrategie veröffentlicht werden. Durch Anwendung verschiedener Anonymisierungsmethoden wird gewährleistet, dass das Risiko einer korrekten Re-Identifizierung einer statistischen Einheit vernachlässigbar ist. Es wurde jedoch darauf Wert gelegt, das in den Daten vorhandene Analysepotential durch die notwendigen Anonymisierungsschritte nicht zu zerstören und an den Daten selbst möglichst wenig zu verändern.

Der FOBS-Datensatz wird registrierten Nutzern in maschinenlesbarer Form angeboten. Damit können die Nutzerinnen und Nutzer die Daten einfach auf ihren PC's und der Software ihrer Wahl einlesen. Zusätzlich kann der Datensatz auch als speziell für *R* aufbereitetes Datenfile bezogen werden.

2 Software

Die Anonymisierung der FOBS-Daten wurde mittels des hausintern entwickelten, frei verfügbaren R-Package `sdcMicro` (siehe z.B. [Templ \(2007\)](#)) durchgeführt. *R* ([R Development Core Team \(2007\)](#)) ist eine freie, objektorientierte high-level Programmiersprache für Computerstatistik und Graphiken, und `sdcMicro` ist eine Abkürzung für **s**tatistical **d**isclosure **c**ontrol for **m**icrodata und ist über [CRAN](#) (R Comprehensive Archive Network) frei zu beziehen. `sdcMicro` wurde bereits erfolgreich bei der Erstellung von anonymisierten Datensätzen (z.B. SDS über die Lohnsteuerstatistik) verwendet und bietet die Möglichkeit, flexibel und effizient und explorativ anonymisierte Mikrodatensätze zu erzeugen. Weiters bleibt zu erwähnen, dass sowohl ein Handbuch als auch eine online-Hilfe in `sdcMicro` integriert sind.

3 Anonymisierung des Datensatzes

Der nicht-anonymisierte FOBS-Datensatz bestand aus 1761 Beobachtungen und 132 Variablen. 35 Fragebögen wurden jedoch nicht vom Unternehmensgründer selbst ausgefüllt und wurden daher gelöscht. Insgesamt bestand der zu anonymisierende Mikrodatensatz also aus 1726 Beobachtungen und 132 Variablen.

3.1 Direkte Identifikationsvariablen

Variablen, die eine direkte Identifizierung einer statistischen Einheit ermöglichen werden als direkte Identifikationsvariablen bezeichnet. Solche Variablen müssen aus Datenschutzgründen aus dem Datensatz entfernt werden. In den FOBS-Daten existieren allerdings keine direkten Identifikationsvariablen, dieses Problem stellte sich demnach nicht.

3.2 Indirekte Identifikationsvariable

Eine weitere Möglichkeit zu einer Identifikation einer statistischen Einheit bilden indirekte Identifikationsvariablen. Jede der (meist kategoriellen) indirekten Identifikationsvariablen für sich lässt keinen Schluss auf ein bestimmtes Unternehmen zu. Kombiniert man jedoch

mehrere indirekte Identifikationsvariablen, so lassen sich bestimmte statistische Einheiten möglicherweise identifizieren. Entscheidend ist jedoch, dass ein Angreifer *Vorwissen* über indirekte Identifikationsvariablen besitzt. Dieses Vorwissen kann beispielsweise aus anderen, öffentlich zugänglichen Datenquellen wie Registern stammen.

Als Beispiele für mögliche indirekte Identifikationsvariablen in den FOBS-Daten gelten etwa die Wirtschaftsklassifikation NACE, Alter, Geschlecht, Bildung und Nationalität der/s FirmengründerIn oder Information darüber, ob der/die FirmengründerIn bezahlte MitarbeiterInnen beschäftigt. Die Kombination dieser (oder anderer) indirekten Identifikationsvariablen könnte zu einer eindeutigen und korrekten Identifizierung einer statistischen Einheit führen. Ziel der Anonymisierung ist daher, dass das Risiko einer korrekten Identifizierung einer Einheit möglichst gering ausfällt. Dabei ist zu beachten, dass das Re-Identifizierungsrisiko auf einem Wahrscheinlichkeitsansatz basiert.

3.3 Schlüsselvariablen für die Geheimhaltung

Die Anonymisierung der FOBS-Daten erfolgt iterativ in einer explorativen Art und Weise. In einem ersten Schritt wird ein Set indirekter Identifizierungsvariablen ausgewählt. Diese Variablen werden nun als Schlüsselvariable für die Geheimhaltung bezeichnet. Sind die Schlüsselvariablen fixiert, kann nach dem Modell von Benedetti and Franconi ([Benedetti and Franconi, 1998](#)) für jede Beobachtung ein individuelles Risiko der Re-Identifizierung berechnet werden. In die Berechnung des individuellen Risikos der Re-Identifikation fließt auch das Stichprobengewicht einer Beobachtung ein. Je niedriger das Stichprobengewicht einer statistischen Einheit ist, desto größer ist die Chance auf Re-Identifizierung dieser Einheit. Es soll nun einerseits das individuelle Risiko der Re-Identifikation möglichst gering ausfallen und andererseits soll k -Anonymity gewährleistet werden. Das bedeutet, dass jede Ausprägungskombination in den Schlüsselvariablen mit zumindest k Einheiten besetzt ist. Für k wurde in Fall der FOBS-Daten der Wert 3 gewählt.

Erreicht werden kann das Ziel der k -Anonymity, indem einzelne Werte in den Schlüsselvariablen gezielt als *fehlend* gesetzt werden. Weitere Möglichkeiten bestehen darin, Umkodierungen in den Schlüsselvariablen oder verschiedene Ausprägungsmöglichkeiten kategoriemer Variablen zusammenzufassen. So kann beispielsweise das Alter des Firmengründers in Alterskategorien umkodiert werden und Information über die Nationalität in nur zwei unterschiedliche Ausprägungskategorien (etwa Österreich und Nicht-Österreich) vergrößert werden. Für die Anonymisierung der FOBS-Daten wurden folgende Variablen als Schlüsselvariable definiert:

- **NACE:** *Wirtschaftsklassifikation*
- **F2_13code:** *höchste abgeschl. Schulbildung zum Zeitpunkt der U-Gründung*
- **F2_10:** *Geburtsjahr des Unternehmensgründers/der Unternehmensgründerin*
- **F2_6code:** *Haben Sie vor der Gründung des jetzigen Unternehmens bereits ein anderes Unternehmen gegründet?*
- **F2_11code:** *Geschlecht*
- **F2_12code:** *Nationalität des Unternehmensgründers/der Unternehmensgründerin*

- **F3_2b:** *Bezahlte Mitarbeiter?*

3.4 Umkodierung

Die in 3.3 definierten Schlüsselvariablen wurden wie folgt vergrößert bzw. umkodiert.

- **NACE:** *Wirtschaftsklassifikation*
Information über Wirtschaftsklassen war in der Variable NACE auf 6-Steller Ebene vorhanden. Die Variable NACE wurde in insgesamt 7 Kategorien zusammengefasst, wobei jede Kategorie aus mehreren 2-Stellern besteht.
 - 1 ... NACE 2-Steller ≤ 39
 - 2 ... $40 \leq$ NACE 2-Steller < 46
 - 3 ... $46 \leq$ NACE 2-Steller < 55
 - 4 ... $55 \leq$ NACE 2-Steller < 60
 - 5 ... $60 \leq$ NACE 2-Steller < 60
 - 6 ... $65 \leq$ NACE 2-Steller < 70
 - 7 ... NACE 2-Steller ≥ 70
- **F2_13code:** *höchste abgeschl. Schulbildung zum Zeitpunkt der U-Gründung*
In den Originaldaten waren vier mögliche Ausprägungskategorien vorhanden. Im Zuge der Anonymisierung wurde die Variable wie folgt umkodiert.
 - 1 ... Meisterpr., FHS oder UNI
 - 2 ... keine Meisterpr./ keine FHS / keine UNI
- **F2_10:** *Altersgruppe des Unternehmensgründers/der Unternehmensgründerin*
Die Information über das Alter des Unternehmensgründers/der Unternehmensgründerin wurde in insgesamt drei Altersklassen umkodiert.
 - 1 ... bis 39 Jahre
 - 2 ... 40 bis 49 Jahre
 - 3 ... 50 Jahre und älter
- **F2_6code:** *Haben Sie vor der Gründung des jetzigen Unternehmens bereits ein anderes Unternehmen gegründet?*
Die Information, ob ein Unternehmensgründer bereits Gründungserfahrung besitzt, ist in Variable F2_6code abgebildet. In den Originaldaten wird unterschieden, ob ein Unternehmensgründer noch nie, genau einmal oder bereits mehr als einmal ein Unternehmen gründete. F2_6code wurde wie folgt vergrößert.
 - 1 ... Nein
 - 2 ... Ja
- **F2_11code:** *Geschlecht*
Die Variable F2_11code wurde für den Anonymisierungsvorgang nicht geändert werden und ist folgendermaßen kodiert.

- 1 ... weiblich
- 2 ... männlich
- **F2_12code:** *Nationalität des Unternehmensgründers/der Unternehmensgründerin*
Für die Frage welcher Nationalität ein Unternehmensgründer angehört ist, existieren in den Originaldaten drei Antwortmöglichkeiten. Es wird zwischen Österreich, einem anderen EU-Land und Nicht-EU-Ländern unterschieden. Die Variable F2_12code wurde für die nachfolgenden Berechnungen wie folgt umkodiert.
 - 1 ... Österreich
 - 2 ... anderes Land
- **F3_2b:** *Bezahlte Mitarbeiter?*
Die Information ob das Unternehmen bezahlte Mitarbeiter beschäftigt ist in Variable F3_2b abgebildet. Diese Variable wurde aber weder umkodiert noch vergrößert und weist folgende Ausprägungen auf:
 - 1 ... Ja
 - 2 ... Nein

3.5 Lokale Unterdrückungen

Basierend auf den in 3.3 gewählten Schlüsselvariablen wird nun für jede Beobachtung das individuelle Re-Identifikationsrisiko berechnet, das in Histogramm 1 dargestellt ist.

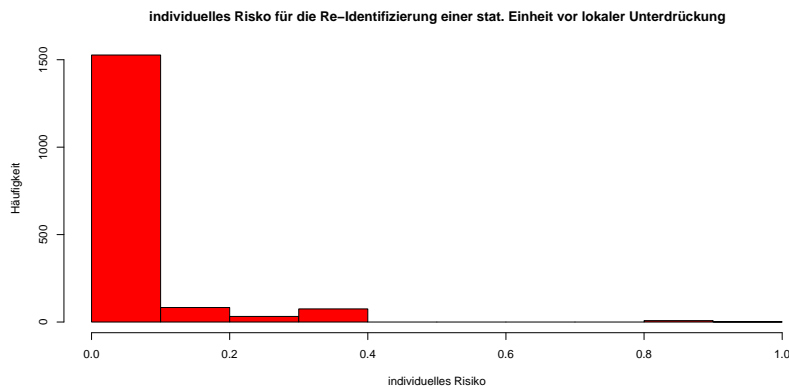


Abbildung 1: Indiv. Risiko der Re-Identifikation vor lokaler Unterdrückung

Man erkennt, aus dem Histogramm 1, dass der überwiegende Teil der Unternehmen ein geringes Re-Identifikationsrisiko aufweist. Allerdings gibt es insgesamt 114 Unternehmen, deren Ausprägungskombinationen in den Schlüsselvariablen *unique* sind. Das bedeutet, dass sich in einer speziellen Schicht¹ nur ein einziges Unternehmen befindet. Außerdem gibt es 102 Unternehmen, die sich in einer Schicht mit genau 2 Unternehmen befinden.

¹eine Schicht entspricht einer möglichen Ausprägungskombination in den Schlüsselvariablen

Ziel ist es nun, bei Unternehmen mit hohem Re-Identifikationsrisiko solange einzelne Werte in den Schlüsselvariablen zu unterdrücken, bis ein ausreichender Schutz gewährleistet ist. Insbesondere soll gewährleistet werden, dass das Re-Identifizierungsrisiko für alle statistischen Einheiten gering ist und dass sich mindestens 3 Elemente in jeder (besetzten) Ausprägungskombination in den Schlüsselvariablen befinden.

sdMicro ermöglicht es, einzelnen Schlüsselvariablen Gewichte zuzuweisen. Damit kann erreicht werden, dass in Schlüsselvariablen, die ein niedriges Gewicht aufweisen, weniger Werte unterdrückt werden als in Schlüsselvariablen mit hohem Gewicht. So wurde für die FOBS-Daten das Gewicht für die Schlüsselvariablen NACE und F2_10 auf 0 gesetzt und damit sichergestellt, dass in diesen Variablen keine Werte unterdrückt werden. Der Grund hierfür ist, dass häufig Auswertungen nach Wirtschaftsklassen sowie Altersgruppen durchgeführt werden.

Nach Durchführung dieser Prozedur ergab sich, dass das Re-Identifikationsrisiko für zwei Unternehmen noch relativ hoch war. Deshalb wurden zwei Werte in der Variable NACE zusätzlich gesperrt. In Tabelle 3.5 ist für jede Schlüsselvariable die Anzahl der notwendigen Unterdrückungen aufgelistet.

Schlüsselvariable:	NACE	F2_6code	F2_10	F2_11code
gesperrte Werte (#):	2	164	0	78
gesperrte Werte (%):	0.12	9.50	0.00	4.52
Schlüsselvariable:	F2_12code	F2_13code	F3_2b	
gesperrte Werte (#):	56	27	31	
gesperrte Werte (%):	3.24	1.56	1.80	

Tabelle 1: unterdrückte Werte in den Schlüsselvariablen

Insgesamt mussten 358 Werte fehlend gesetzt werden. In der Schlüsselvariable F2_10 wird kein Wert unterdrückt, was sich aus der (bewussten) Setzung des Gewichtes auf 0 für diese Variable ergibt. Das Gewicht war auch für die Wirtschaftsklassifikation NACE auf 0 gesetzt, allerdings mussten nachträglich noch zwei Werte in dieser Variable unterdrückt werden um das Re-Identifikationsrisiko zweier Unternehmen zu reduzieren. Durch diese Unterdrückungen wird 3-Anonymity gewährleistet. Es befinden sich also mindestens 3 Unternehmen in jeder besetzten Ausprägungskombination in den Schlüsselvariablen.

Histogramm 2 zeigt das individuelle Re-Identifikationsrisiko nach erfolgter lokaler Unterdrückung, welches im Vergleich zu Abbildung 1 deutlich geringer ausfällt.

4 Folgerungen

Bei der Anonymisierung der FOBS-Daten wurde darauf geachtet, das Risiko einer korrekten Identifikation einzelner statistischer Einheiten möglichst gering zu halten. Durch Umkodierungen und Vergrößerungen in einzelnen Variablen wurde zusätzliche Sicherheit gegenüber möglichen Deanonimisierungsversuchen geschaffen.

Die Auswahl der Schlüsselvariablen erfolgte nach fachstatistischem Wissen in enger Zusammenarbeit mit den für die FOBS-Erhebung zuständigen Fachleuten. Absoluten Schutz vor einer korrekten Re-Identifizierung eines Unternehmens bieten auch die durchgeführten

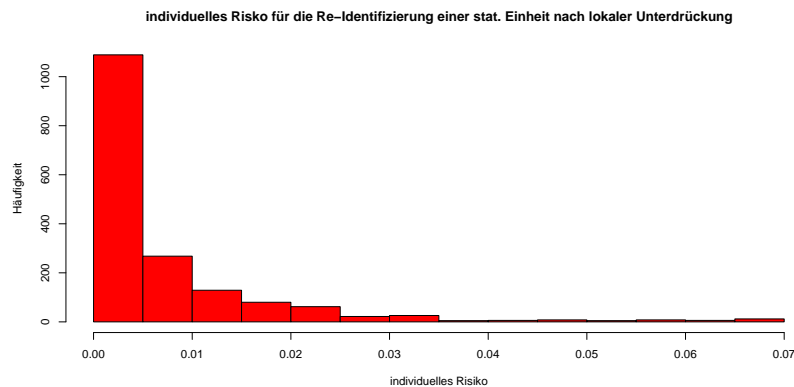


Abbildung 2: Indiv. Risiko der Re-Identifikation nach lokaler Unterdrückung

Anonymisierungsschritten nicht. Die Wahrscheinlichkeit einer korrekten Identifikation ist jedoch gering.

Zusammenfassend kann gesagt werden, dass ein hoher Anonymisierungsgrad erreicht wurde und gleichzeitig das Analysepotential der Daten weitestgehend erhalten blieb. Der Trade-Off zwischen Datenanonymisierung einerseits und Gewährleistung der Nutzbarkeit der Daten ist oftmals schwierig. Es mussten insgesamt 2.95% aller Werte in den gewählten Schlüsselvariablen gesperrt werden um zu gewährleisten, dass jede nicht leere Ausprägungskombination in den Schlüsselvariablen zumindest drei Unternehmen umfasst.

Die Anonymisierung der FOBS-Daten wurde vollständig mit dem hausintern entwickelten R-Package `sdcMicro` durchgeführt. Durch weitere Entwicklungen gegenüber vorherigen Versionen von `sdcMicro` ist es nun möglich, durch Vergabe von Gewichten die Schlüsselvariablen zu "werten" und zu erreichen, dass beispielsweise in Variablen, nach denen häufig ausgewertet wird, nur wenige oder überhaupt keine Werte unterdrückt werden müssen. So wurde etwa gezeigt, dass 3-Anonymität erreicht werden kann, ohne explizit Werte in den Variablen NACE (Wirtschaftsklassen) und F2_10 (Altersklassen) zu sperren. Allerdings war es dennoch notwendig, zwei Werte in der Variable NACE zu sperren. Dies geschah jedoch nicht um 3-Anonymität zu erreichen, sondern um das Re-Identifikationsrisiko einzelner Unternehmen zusätzlich zu reduzieren.

Literatur

Benedetti, R. and Franconi, L. (1998). Statistical and technological solutions for controlled data dissemination. In Pre-proceedings of New Techniques and Technologies for Statistics, volume 1, pages 225–232.

R Development Core Team (2007). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

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

5 Datenbeschreibung FOBS Daten

Es werden nun die Variablen beschrieben, die im SDS enthalten sind. Überblicksmäßig wird die „Aktion“ der Geheimhaltung in Spalte 4 der folgenden Tabelle beschrieben. Zur leichteren Orientierung wurden die Variablennamen mit unterschiedlichen Farben markiert, wobei **blau** bedeutet, dass die Variable verändert oder neu erzeugt wurde. Die Variablen, die in schwarzer Schrift zu sehen sind, wurden nicht verändert.

Variablenname	Skalierung / Identifier	Beschreibung	Aktion	Kodierung/ Spezifizierung
NACE	kat. / indirekt	Wirtschaftsklassifikation	vergrößert durch Zusammenfassung von NACE 2-Stellern	1 = 2-Steller < 40 2 = 40 <= 2-Steller <= 45 3 = 46 <= 2-Steller <= 54 4 = 55 <= 2-Steller <= 59 5 = 60 <= 2-Steller <= 64 6 = 65 <= 2-Steller <= 69 7 = 2-Steller >= 70
F2_1acode	kat. / nein	Neue Herausforderung	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_1bcode	kat. / nein	Der eigene Chef sein	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_1ccode	kat. / nein	Hobby zum Beruf	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_1dcode	kat. / nein	Familientradition	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_1ecode	kat. / nein	Verbindung Arbeit mit Privatleben	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_1fcode	kat. / nein	Kinder sind jetzt groß	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_1gcode	kat. / nein	Aussicht mehr Geld zu verdienen	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_1hcode	kat. / nein	Neues Produkt/neue Dienstleistung	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_1icode	kat. / nein	Internationaler Markt	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_1jcode	kat. / nein	Vermeiden von Arbeitslosigkeit	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_1kcode	kat. / nein	Alte Arbeitssituation war unbefriedigend	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_1lcode	kat. / nein	Exklusiver Subunternehmer für bisherigen Arbeitsgeber	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_1mcode	kat. / nein	Einzige Möglichkeit den Beruf auszuüben	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_2a	kat. / nein	Eigene Finanzmittel oder Ersparnisse	keine	0=Nein, 1=Ja
F2_2b	kat. / nein	Finanzielle Unterstützung durch Familie oder Freunde	keine	0=Nein, 1=Ja
F2_2c	kat. / nein	Bankdarlehen ohne Sicherheiten	keine	0=Nein, 1=Ja
F2_2d	kat. / nein	Bankdarlehen mit Sicherheiten	keine	0=Nein, 1=Ja
F2_2e	kat. / nein	Finanzielle Unterstützung durch andere Unternehmen	keine	0=Nein, 1=Ja
F2_2f	kat. / nein	Beteiligungs- oder Risikokapital	keine	0=Nein, 1=Ja
F2_2g	kat. / nein	Finanzielle Unterstützung durch die öffentliche Hand	keine	0=Nein, 1=Ja

Anhang: Datenbeschreibung FOBS-Daten

F2_3a_acode	kat. / nein	Finanzierung	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_3a_bcode	kat. / nein	Aufbau von Kundenkontakten	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_3a_ccode	kat. / nein	Bezahlung ausstehender Rechnungen	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_3a_dcode	kat. / nein	Preise für eigene Produkte oder Dienstleistungen festlegen	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_3a_ecode	kat. / nein	Finden geeigneter Räumlichkeiten	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_3a_fcode	kat. / nein	Finden von Lieferanten	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_3a_gcode	kat. / nein	Finden von geeignetem Personal	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_3a_hcode	kat. / nein	Effektive Nutzung der EDV	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_3a_icode	kat. / nein	Abwicklung rechtlicher oder administrativer Angelegenheiten	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_3a_jcode	kat. / nein	Allein für alles verantwortlich sein	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_3a_kcode	kat. / nein	Rückhalt vom Ehepartner oder der Familie	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F2_3bcode	kat. / nein	Bezahlte Mitarbeiter im Jahr der Unternehmensgründung	keine	1=Ja, 2=Nein
F2_4code	kat. / nein	Was waren Sie unmittelbar bevor Sie das Unternehmen gründeten?	vergrößert	1=unselbstständig beschäftigt 2=Inhaber(in) eines anderen Unternehmens 3=Student(in) 4=arbeitslos oder ohne Erwerbstätigkeit
F2_5code	kat. / nein	Hatten Sie bereits praktische Erfahrung in der Branche Ihres jetzigen Unternehmens?	keine	1=Ja, 2=Nein
F2_6code	kat. / indirekt	Haben Sie vor der Gründung des jetzigen Unternehmens bereits ein anderes Unternehmen gegründet?	vergrößert	1=Nein, 2=Ja
F2_7code	kat. / nein	Haben Sie ein spezielles Weiterbildungs- oder Kursangebot genutzt?	nein	1=Ja, aus eigener Initiative, 2=Ja, auf Anraten der Behörden, 3=Nein
F2_8acode	kat. / nein	Freunde und Familie	nein	0=Nein, 1=Ja
F2_8bcode	kat. / nein	Spezielle Weiterbildungs- oder Kursangebote für Unternehmer	nein	0=Nein, 1=Ja
F2_8ccode	kat. / nein	Kontakte zu anderen Unternehmen	nein	0=Nein, 1=Ja
F2_8dcode	kat. / nein	Professionelle Berater	nein	0=Nein, 1=Ja
F2_8ecode	kat. / nein	Das AMS	nein	0=Nein, 1=Ja
F2_8fcode	kat. / nein	Organisationen, die auf Unternehmensneugründungen spezialisiert sind	nein	0=Nein, 1=Ja
F2_8gcode	kat. / nein	Ein Finanzinstitut	nein	0=Nein, 1=Ja
F2_8hcode	kat. / nein	Keinen Zugang zu den genannten Hilfestellungen	nein	0=Nein, 1=Ja
F2_8icode	kat. / nein	Keine Hilfe benötigt	nein	0=Nein, 1=Ja
F2_9acode	kat. / nein	Haben Sie derzeit zusätzlich zu Ihrem Unternehmen eine andere Erwerbstätigkeit?	nein	1=Ja, 2=Nein

Anhang: Datenbeschreibung FOBS-Daten

F2_9bcode	kat. / nein	Handelt es sich bei Ihrer anderen Erwerbstätigkeit um eine unselbstständige?	nein	1=Ja, 2=Nein
F2_10	kat. / indirekt	Geburtsjahr des Unternehmensgründers/der Unternehmensgründerin	vergrößert	1=34 Jahre und jünger 2=35-49 Jahre 3=50 Jahre und älter
F2_11code	kat. / indirekt	Geschlecht des Unternehmensgründers/der Unternehmensgründerin	nein	1=weiblich, 2=männlich
F2_12code	kat. / indirekt	Nationalität des Unternehmensgründers/der Unternehmensgründerin	statt 3 nur 2 Klassen	1=Österreich, 2=anderes Land
F2_13code	kat. / indirekt	Höchste abgeschlossene Schulbildung zum Zeitpunkt der Unternehmensgründung	statt 4 nur 2 Klassen	1=Meisterpr., FHS oder UNI, 2=keine Meisterpr., FHS oder UNI
F3_1code	kat. / indirekt	Von wem wird das Unternehmen heute geleitet?	statt 3 nur 2 Klassen	1=Gründer, 2=nicht Gründer
F3_2a	kat. / nein	Unbezahlte oder selbstständige Mitarbeiter	vergrößert	1=höchstens ein unbezahlter oder selbstständiger Mitarbeiter, 2=mehr als ein unbezahlter oder selbstständiger Mitarbeiter
F3_2b	kat. / indirekt	Bezahlte Mitarbeiter	vergrößert	1=keine bezahlten Mitarbeiter, 2=bezahlte Mitarbeiter
F3_3	kat. / nein	Gesamtumsatz 2004	vergrößert	1= 20 Tsd. € und weniger 2=ab 20 Tsd. bis einschließlich 65 Tsd. € 3=ab 65 Tsd. bis einschließlich <=175 Tsd. € 4=über 175 Tsd. €
F3_4	kat. / nein	Gesamtumsatz erste Hälfte 2005	vergrößert	1= 10 Tsd. € und weniger 2=ab 10 Tsd. bis einschließlich 35 Tsd. € 3=ab 35 Tsd. bis einschließlich <=100 Tsd. € 4=über 100 Tsd. €
F3_5code	kat. / indirekt	Von wie vielen Abnehmern/Kunden erzielen Sie den überwiegenden Anteil des Gesamtumsatzes?	nein	1=1 Abneh. 2=2-4 Abneh. 3=5-10 Abneh. 4=mehr als 10 Abnehm.
F3_6a	kat. / nein	Lokaler bzw. regionaler Markt	nein	0=Nein, 1=Ja
F3_6b	kat. / nein	Nationaler Markt	nein	0=Nein, 1=Ja
F3_6c	kat. / nein	Internationaler Markt innerhalb der EU	nein	0=Nein, 1=Ja
F3_6d	kat. / nein	Internationaler Markt außerhalb der EU	nein	0=Nein, 1=Ja
F3_7code	kat. / nein	Die Wirtschaftlichkeit des Unternehmens	nein	1=sehr gut 2=gut 3=gerade noch ausreichend 4=nicht ausreichend
F3_8code	kat. / nein	Kooperation mit anderen Unternehmen	nein	1=Netzwerk/Cluster 2=Subunternehmer 3=Franchise-/Lizenznehmer 4=Joint Venture 5=gar nicht
F3_9code	kat. / nein	Glauben Sie, dass Ihr Unternehmen als innovativ bewertet werden kann?	nein	1=Ja, 2=Nein
F3_9a	kat. / nein	Produktinnovation	nein	0=Nein, 1=Ja
F3_9b	kat. / nein	Prozessinnovation	nein	0=Nein, 1=Ja
F3_9c	kat. / nein	Organisationsinnovation	nein	0=Nein, 1=Ja
F3_9d	kat. / nein	Marketinginnovation	nein	0=Nein, 1=Ja
F3_10acode	kat. / nein	Verkaufshindernis: Wettbewerbssituation	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F3_10bcode	kat. / nein	Verkaufshindernis: Bedarf und Nachfrage	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht

Anhang: Datenbeschreibung FOBS-Daten

F3_10ccode	kat. / nein	Verkaufshindernis: Preisgestaltung	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F3_10dcode	kat. / nein	Verkaufshindernis: Fehlende Marketingkenntnisse	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F3_11acode	kat. / nein	Entwicklungshindernis: Ertragslage/Rentabilität	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F3_11bcode	kat. / nein	Entwicklungshindernis: Verfügbarkeit von Bankkrediten	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F3_11ccode	kat. / nein	Entwicklungshindernis: Verfügbarkeit von Risikokapital	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F3_11dcode	kat. / nein	Entwicklungshindernis: Verfügbarkeit von kurzfristigen Krediten	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F3_11ecode	kat. / nein	Entwicklungshindernis: Finden von qualifiziertem Personal	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F3_11fcode	kat. / nein	Entwicklungshindernis: Personalkosten	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F3_11gcode	kat. / nein	Entwicklungshindernis: Behördliche oder administrative Verpflichtungen	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F3_11hcode	kat. / nein	Entwicklungshindernis: Anpassung an neue technische Standards	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F3_11icode	kat. / nein	Entwicklungshindernis: Führungskompetenz	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F3_11jcode	kat. / nein	Entwicklungshindernis: Vereinbarkeit von Familie und Arbeit	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F3_11kcode	kat. / nein	Entwicklungshindernis: Finden von Geschäftspartnern	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F3_11lcode	kat. / nein	Entwicklungshindernis: Nicht oder spät zahlende Kunden	keine	1=Trifft sehr zu, 2=Trifft zu, 3=Trifft nicht zu, 4=Weiß ich nicht
F4_1code	kat. / nein	Strategische Pläne für die nächsten zwei Jahre	keine	1=wie bisher 2=verschmelzen 3=verkaufen und neues in gleicher Branche gründen 4=verkaufen und neues in anderer Branche gründen 5=verkaufen ohne Neugründung
F4_2acode	kat. / nein	Der Umsatz wird:	keine	1=abnehmen, 2=gleich bleiben, 3=steigen
F4_2bcode	kat. / nein	Die Anzahl der Mitarbeiter wird:	keine	1=abnehmen, 2=gleich bleiben, 3=steigen
F4_2ccode	kat. / nein	Die Auswahl der angebotenen Produkte und Dienstleistungen wird:	keine	1=abnehmen, 2=gleich bleiben, 3=steigen
F4_2dcode	kat. / nein	Die Ertragslage wird:	keine	1=abnehmen, 2=gleich bleiben, 3=steigen
F4_3acode	kat. / nein	Priorität: Mitarbeiter einstellen	keine	0=Nein, 1=Ja
F4_3bcode	kat. / nein	Priorität: Mitarbeitern mehr bezahlen	keine	0=Nein, 1=Ja
F4_3ccode	kat. / nein	Priorität: ins Unternehmen selbst investieren	keine	0=Nein, 1=Ja
F4_3dcode	kat. / nein	Priorität: Kredite und Schulden zurückzahlen	keine	0=Nein, 1=Ja
F4_3ecode	kat. / nein	Priorität: eigenes Gehalt erhöhen	keine	0=Nein, 1=Ja
F4_3fcode	kat. / nein	Priorität: eigene Arbeitszeit reduzieren	keine	0=Nein, 1=Ja
F4_3gcode	kat. / nein	Priorität: nichts von all dem	keine	0=Nein, 1=Ja