

**Konzept zur absoluten Anonymisierung statistischer  
Einzeldaten aus dem AFiD-Panel Agrarstruktur 1999, 2003,  
2007 zur Erstellung eines CAMPUS-Files (CF)**

**CF AFiD-Panel Agrarstruktur**

Version 1.0

## **Vorbemerkung**

Das CAMPUS-File AFiD-Panel Agrarstruktur ist ein absolut anonymes Public-Use-File, das speziell für Lehrende und Studierende erstellt wird. Seine Funktion besteht darin, die praktische Statistikausbildung (insbesondere im Bereich der Arbeit mit Paneldaten) mit amtlichen Einzeldaten anzureichern und damit den Hochschulen ein effektives Werkzeug für eine qualitativ hochwertige Lehre zu liefern.

## **Basismaterial**

Das Basismaterial des vorliegenden CAMPUS-Files bildet das AFiD-Panel Agrarstruktur 1999, 2003, 2007. Dieses umfasst Einzeldaten mit 502 Merkmalen von 535.902 land- und forstwirtschaftlichen Betrieben.

Ausgangspunkt für das AFiD-Panel Agrarstruktur 1999, 2003, 2007 sind die Einzeldaten des Materials Agra02 der Erhebungsjahre 1999, 2003 und 2007 für alle 16 Bundesländer der Bundesrepublik Deutschland.

## **Anonymisierungsmaßnahmen**

Im Folgenden werden die Maßnahmen beschrieben, die zur Einhaltung der Maßgabe der absoluten Anonymität der Einzeldaten eingesetzt werden. Die Anonymisierungsmaßnahmen werden, entsprechend der genannten Reihenfolge auf den Datenbestand angewandt:

- Die Regionalangaben werden auf „Ost“ und „West“ vergrößert

Als regionale Gliederungsebene wird lediglich die Unterscheidung „alte“ und „neue“ Bundesländer (Merkmal Region) aufgenommen. Die Berechnung erfolgt aus dem Merkmal EF2U1 (Bundesland). Dieses Merkmal wird nicht in das CAMPUS-File übernommen.

Die Regionen „West“ und „Ost“ umfassen dabei:

<b>West</b>	<b>Ost</b>
Schleswig-Holstein	Berlin
Hamburg	Brandenburg
Niedersachsen	Mecklenburg-Vorpommern
Bremen	Sachsen
Nordrhein-Westfalen	Sachsen-Anhalt
Hessen	Thüringen
Rheinland-Pfalz	
Baden-Württemberg	
Bayern	
Saarland	

- Das Datenmaterial hat einen stark reduzierten Merkmalsumfang.

Von den im Basismaterial enthaltenen 502 Merkmalen werden nur 48 Merkmale für das CAMPUS File ausgewählt. Die ausgewählten Merkmale können der Merkmalsliste in Anlage 1 entnommen werden.

- Sensible und identifizierende Merkmale werden nur klassifiziert übermittelt

Das Standardbetriebseinkommen (nur 1999) sowie der Standarddeckungsbeitrag (nur 2003 und 2007) werden nur als kategoriale Variable zur Verfügung gestellt. Gleiches gilt für das möglicherweise identifizierende Merkmal des Alters des Betriebsleiters. Die Größenklassen in die die Variablen EF7\_alt, EF7 und EF8 eingeteilt werden, können der Merkmalsliste in Anlage 1 entnommen werden.

- Forstbetriebe werden entfernt

Alle im AFiD-Panel Agrarstruktur 1999, 2003, 2007 enthaltenen Forstbetriebe (EF9 $\geq$ 20) werden aus dem Datensatz entfernt. Nach diesem Schritt sind noch Informationen für 506.150 landwirtschaftliche Betriebe im Datensatz enthalten.

- Besonders auffällige landwirtschaftliche Betriebe werden entfernt

Ein landwirtschaftlicher Betrieb wird in allen Jahren aus dem Datensatz entfernt, wenn er in mindestens einem Erhebungsjahr bei mindestens einer Bodennutzungs- oder Viehbestandsvariable (EF106-EF143, EF201-EF258, EF506-EF543) mit seinem Wert zu den drei größten Betrieben zählt. Die Prüfung wird getrennt nach den beiden Regionen „West“ und „Ost“ und getrennt nach Ökobetrieben und konventionellen Betrieben (EF750) durchgeführt. Insgesamt werden in diesem Schritt 444 landwirtschaftliche Betriebe aus dem Datensatz entfernt.

- Das Datenmaterial ist nur eine Stichprobe des Originalmaterials

Für das CAMPUS-File wird aus den verbliebenen 505.706 landwirtschaftlichen Betrieben eine 5%-Zufallsstichprobe gezogen = 25.285 landwirtschaftliche Betriebe. (Angaben über die Anzahl der Betriebe in den jeweiligen Erhebungsjahren finden sich in Anlage 2.)

- Metrische Variablen werden mit einem Zufallsfaktor multipliziert (Zufallsüberlagerung)

Für alle landwirtschaftlichen Betriebe werden die metrischen Merkmale mit einem Zufallsfaktor multipliziert. Der einem landwirtschaftlichen Betrieb zugewiesene Faktor ist für alle metrischen Merkmale über alle Jahre gleich. Die Hälfte der landwirtschaftlichen Betriebe bekommt einen Faktor zwischen 0,6 und 0,8 und die andere Hälfte einen Faktor zwischen 1,2 und 1,4 zugewiesen.

Nach der Zufallsüberlagerung werden Stückzahlen auf ganze Zahlen gerundet. Die Angaben zur Bodennutzung sowie zu den Arbeitskräfteeinheiten werden auf die zweite Nachkommastelle gerundet.

Die Größenklasse der landwirtschaftlich genutzten Flächen (EF4) wird an den durch die Zufallsüberlagerung veränderten Wert der landwirtschaftlich genutzten Fläche (EF258) angepasst.

Die Summenpositionen der Viehbestände (Pferde insgesamt, Rinder insgesamt, Schafe insgesamt, Schweine insgesamt, Hühner insgesamt und Sonstiges Geflügel insgesamt) werden auf Basis der zufallsüberlagerten dahinterstehenden Einzelpositionen neu berechnet. (Bis auf die Rinder 2 Jahre und älter Milchkühe (EF116) und die Ferkel (EF125) sind die Einzelpositionen jedoch nicht im CAMPUS-File enthalten – siehe Anlage 1).

Die Werte der Großvieheinheiten (EF506 bis EF543) werden auf Basis der zufallsüberlagerten Stückzahlen der jeweils zugrundeliegenden Einzelpositionen neu berechnet.

- Die Datensätze sind systemfrei angeordnet / die Betriebsnummern werden durch systemfreie Nummern ersetzt

Aus der Anordnung der Datensätze im Originalmaterial wäre es möglich, indirekt Regionalinformationen abzuleiten. Daher werden alle Betriebsnummern durch eine zufällig vergebene systemfreie Nummer ersetzt. Das CAMPUS-File wird nach diesen systemfreien Nummern sortiert. Die systemfreien Nummern sind über alle Erhebungsjahre für den gleichen Betrieb identisch.

- Das Datenmaterial weist ein bestimmtes Mindestalter auf

Das Alter der Erhebung (1999, 2003, 2007) bietet eine weitere Schutzwirkung für die enthaltenen Einzeldaten. Die Werte sind durch die nachfolgende Landwirtschaftszählung 2010 sowie die Agrarstrukturerhebung 2013 bereits überholt.

## **Fazit**

Die durchgeführten Maßnahmen stellen die absolute Anonymität der im CAMPUS-File enthaltenen Einzeldaten sicher. Zum Einsatz in der Methodenlehre bietet das vorliegende CAMPUS-File trotz der durchgeführten Maßnahmen vielfältige Analysemöglichkeiten.

## Anlage 1: Liste der enthaltenen Merkmale

Region: 1 = West (alte Bundesländer)  
2 = Ost (neue Bundesländer)

nr: systemfreie Kenn-Nummer der Betriebseinheit.  
Die Kenn-Nummer ermöglicht eine eindeutige Identifizierung der einzelnen Betriebseinheit und strukturiert damit zusammen mit der Zeitpunktinformation aus dem Merkmal EF31 (Berichtsjahr) den Paneldatensatz.

EF4: Größenklasse der landwirtschaftlich genutzten Flächen (LF, aus EF258)

00	=	0 ha		
01	=	0,01	bis unter	1,00 ha
10	=	1,00	bis unter	2,00 ha
20	=	2,00	bis unter	5,00 ha
30	=	5,00	bis unter	10,00 ha
40	=	10,00	bis unter	15,00 ha
50	=	15,00	bis unter	20,00 ha
61	=	20,00	bis unter	25,00 ha
62	=	25,00	bis unter	30,00 ha
71	=	30,00	bis unter	40,00 ha
72	=	40,00	bis unter	50,00 ha
81	=	50,00	bis unter	75,00 ha
82	=	75,00	bis unter	100,00 ha
91	=	100,00	bis unter	200,00 ha
92	=	200,00	bis unter	500,00 ha
93	=	500,00	bis unter	1.000,00 ha
94	=		größer/gleich	1.000,00 ha

EF7\_alt: Größenklasse des Standardbetriebseinkommens in DM (nur 1999)

10	=		unter	2.500 DM
11	=	2.500	bis unter	5.000 DM
20	=	5.000	bis unter	10.000 DM
31	=	10.000	bis unter	15.000 DM
32	=	15.000	bis unter	20.000 DM
40	=	20.000	bis unter	30.000 DM
51	=	30.000	bis unter	40.000 DM
52	=	40.000	bis unter	50.000 DM
60	=	50.000	bis unter	75.000 DM
70	=	75.000	bis unter	100.000 DM
81	=	100.000	bis unter	200.000 DM
82	=	200.000	bis unter	500.000 DM
83	=		größer/gleich	500.000 DM

EF7: Größenklasse des Standarddeckungsbeitrages in EUR (nur 2003 und 2007)

10	=		unter	2.400 EUR
11	=	2.400	bis unter	4.800 EUR
20	=	4.800	bis unter	9.600 EUR
31	=	9.600	bis unter	14.400 EUR
32	=	14.400	bis unter	19.200 EUR
40	=	19.200	bis unter	28.800 EUR
51	=	28.800	bis unter	38.400 EUR
52	=	38.400	bis unter	48.000 EUR
60	=	48.000	bis unter	72.000 EUR
70	=	72.000	bis unter	96.000 EUR
81	=	96.000	bis unter	120.000 EUR
82	=	120.000	bis unter	300.000 EUR
83	=		größer/gleich	300.000 EUR

EF8: Größenklasse nach Alter des Betriebsinhabers (nur S-Betriebe)

15	= 15	bis	19 Jahre
20	= 20	bis	24 Jahre
25	= 25	bis	29 Jahre
30	= 30	bis	34 Jahre
35	= 35	bis	39 Jahre
40	= 40	bis	44 Jahre
45	= 45	bis	49 Jahre
50	= 50	bis	54 Jahre
55	= 55	bis	59 Jahre
60	= 60	bis	64 Jahre
65	= 65	bis	69 Jahre
70	= 70 Jahre und älter		

EF12: Betriebsformen des Betriebsbereichs Landwirtschaft<sup>1</sup> (nur 1999)

- 1 =Marktfruchtbetriebe
- 2 =Futterbaubetriebe
- 3 =Veredlungsbetriebe
- 4 =Dauerkulturbetriebe
- 5 =Landw. Gemischtbetriebe

EF13: Darstellungsbereich der Rechtsform

Betriebe in der Hand von:

- 1 = Einzelpersonen
- 2 = Personengesellschaften } natürliche  
einschließlich Personengemeinschaften } Personen
- 3 = juristische Personen

<sup>1</sup> Ab 2003 Klassifizierung im Merkmal EF72 (Betriebswirtschaftliche Ausrichtung). Diese ist durch unterschiedliche Methodik der Ermittlung nicht direkt vergleichbar.

- EF16: Erhebungsteil  
S = Stichprobenbetrieb  
N = Nicht-Stichprobenbetrieb
- EF31: Berichtsjahr (= 1999, 2003, 2007)  
Das Berichtsjahr bezeichnet das Erhebungsjahr und strukturiert damit zusammen mit der Kenn-Nummer des Betriebs (nr) den Paneldatensatz.
- EF72UG3: Allgemeine Betriebswirtschaftliche Ausrichtung (nur 2003 und 2007)
- |   |   |
|---|---|
| 1 | = Ackerbaubetriebe                      |
| 2 | = Gartenbaubetriebe                     |
| 3 | = Dauerkulturbetriebe                   |
| 4 | = Weideviehbetriebe                     |
| 5 | = Veredlungsbetriebe                    |
| 6 | = Pflanzenbauverbundbetriebe            |
| 7 | = Viehhaltungsverbundbetriebe           |
| 8 | = Pflanzenbau- und Viehhaltungsbetriebe |

### **Viehhaltung: (Angaben in Stück)**

- EF106: Pferde insgesamt  
EF116: Rinder 2 Jahre und älter Milchkühe  
EF119: Rinder insgesamt  
EF124: Schafe insgesamt  
EF125: Ferkel  
EF135: Schweine insgesamt  
EF139: Hühner insgesamt<sup>2</sup>  
EF143: Sonstiges Geflügel insgesamt

### **Bodennutzung (Angaben in ha, soweit nicht anders angegeben)**

- EF201: Winterweizen (ohne Durum)  
EF204: Triticale  
EF205: Roggen  
EF206: Wintergerste  
EF207: Sommergerste  
EF208: Hafer  
EF220: Zuckerrüben ohne Samenbau  
EF229: Winterraps zur Körnergewinnung  
EF241: Feldgras, Grasanbau auf dem Ackerland zum Abmähen und Abweiden  
EF242: Silomais (einschließlich Grünmais und Lieschkolbenschrot)

---

<sup>2</sup> ohne Trut-, Perl- und Zwerghühner

- EF243: Alle anderen Futterpflanzen, auch als Gemenge zur Grünfütter-, Silage- oder Heugewinnung (z.B. Futtererbsen, Wicken, Süßlupinen, Sonnenblumen)
- EF244: Brache (Rotations- und Dauerbrache einschließlich stillgelegter Flächen zur Erlangung der Ausgleichszahlung); 2007: Stillgelegtes / aus der landwirtschaftlichen Erzeugung genommenes Ackerland, Brache
- EF245: Ackerland insgesamt
- EF249: Dauerwiesen (Nutzung durch Schnitt, eventuell Nachweide im Herbst)
- EF250: Mähweiden (Nutzung durch Schnitt und Beweidung wechseln regelmäßig)
- EF251: Dauerweiden (Nutzung durch Beweidung, Schnitt eventuell als Nachmahd)
- EF258: Landwirtschaftlich genutzte Fläche (LF)

### **Großvieheinheiten (Angaben in GV)**

- EF506: Pferde insgesamt
- EF516: Rinder 2 Jahre und älter Milchkühe
- EF519: Rinder insgesamt
- EF524: Schafe insgesamt
- EF525: Ferkel
- EF535: Schweine insgesamt
- EF539: Hühner insgesamt<sup>3</sup>
- EF543: Sonstiges Geflügel insgesamt
- EF544: Vieheinheiten insgesamt

### **Ökologischer Landbau**

- EF750: Unterliegt Ihr Betrieb dem Kontrollverfahren nach der EWG-Verordnung Nr. 2092/91 zum ökologischen Landbau (EWG-Öko-Verordnung)?
- 1 = ja
- 2 = nein

### **Beschäftigte Arbeitskräfte nach Personengruppen im Zeitraum Mai ... - April**

Familienarbeitskräfte (einschl. BI) in Betrieben der Rechtsform Einzelunternehmen

- EF770: insgesamt (AKE)

Ständig und nicht ständig Beschäftigte (einschließlich Saisonarbeitskräfte) in Betrieben aller Rechtsformen (ohne Familien-AK)

- EF775: Insgesamt (AKE)

---

<sup>3</sup> ohne Trut-, Perl- und Zwerghühner

## Anlage 2: Für wie viele Betriebe liegen Informationen im CAMPUS-File vor?

Insgesamt liegen im CAMPUS-File AFiD-Panel Agrarstruktur Informationen für 25.285 landwirtschaftliche Betriebe vor. Die Anzahl der Betriebe in den jeweiligen Jahren ist der folgenden Tabelle zu entnehmen.

	Landwirtschaftliche Betriebe	Darunter: Ökologisch wirtschaftende Betriebe (EF750=1)
	Anzahl	Anzahl
1999	23.627	524
2003	21.005	745
2007	18.663	761
In allen Jahren vorhanden	17.172	341