

1176²⁰²²

SOEP Survey Papers
Series D – Variable Descriptions and Coding

SOEP-Core v37 – BIOIMMIG: Gene- rated Information on Immigration History

Miriam Gauer, Selin Kara, and SOEP Group

Running since 1984, the German Socio-Economic Panel (SOEP) is a wide-ranging representative longitudinal study of private households, located at the German Institute for Economic Research, DIW Berlin.

The aim of the SOEP Survey Papers Series is to thoroughly document the survey's data collection and data processing. The SOEP Survey Papers is comprised of the following series:

Series A – Survey Instruments (Erhebungsinstrumente)
Series B – Survey Reports (Methodenberichte)
Series C – Data Documentation (Datendokumentationen)
Series D – Variable Descriptions and Coding
Series E – SOEPmonitors
Series F – SOEP Newsletters
Series G – General Issues and Teaching Materials

The SOEP Survey Papers are available at <http://www.diw.de/soepsurveyspapers>

Editors:

Dr. Jan Goebel, DIW Berlin
Prof. Dr. Stefan Liebig, DIW Berlin and Freie Universität Berlin
Prof. Dr. David Richter, DIW Berlin and Freie Universität Berlin
Prof. Dr. Carsten Schröder, DIW Berlin and Freie Universität Berlin
Prof. Dr. Jürgen Schupp, DIW Berlin and Freie Universität Berlin
Prof. Dr. Sabine Zinn, DIW Berlin and Humboldt-Universität zu Berlin

Please cite this paper as follows:

Miriam Gauer, Selin Kara, and SOEP Group. 2022. SOEP-Core v37 – BIOIMMIG: Generated Information on Immigration History. SOEP Survey Papers 1176: Series D – Variable Descriptions and Coding. Berlin: DIW Berlin/SOEP



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

© 2022 by SOEP

ISSN: 2193-5580 (online)

DIW Berlin
German Socio-Economic Panel (SOEP)
Mohrenstr. 58
10117 Berlin
Germany

soeppapers@diw.de

SOEP-Core v37 – BIOIMMIG: Generated Information on Immigration History

Miriam Gauer, Selin Kara, and SOEP Group

2022

Contents

1	Introducing the BIOIMMIG dataset	4
	Sources used to generate the BIOIMMIG dataset	5
2	Data structure and imputating data	6
	Missing values	7
	Using the BIOIMMIG dataset	8
3	Variables, Values and Sources in File BIOIMMIG	9
	biimgrp – BI: Immigration Group	10
	biresper – BI: Residence Status	11
	bicamp – BI: Refugee Residence Y,N	11
	bicampw – BI: Refugee Residence: Weeks	12
	bicampm – BI: Refugee Residence: Months	13
	biwfam – BI: Already Had Family In Country	13
	bifamc – BI: Contacts With Family In Germany	14
	bifamcl – BI: Moved To Same City,Town As Family	15
	birbetr – BI: Reason Migrate: Better	16
	birmoney – BI: Reason Migrate: Money	16
	birfree – BI: Reason Migrate: Freedom	17
	birfam – BI: Reason Migrate: Family	17
	birpoor – BI: Reason Migrate: Poor	17
	birwar – BI: Reason Migrate: War	18
	birjust – BI: Reason Migrate: Just So	18
	birothr – BI: Reason Migrate: Other	19
	biexpr – BI: Expectations In Germany	19
	biexprlv – BI: Expectations: Find Apt	20
	biexprac – BI: Expectations: Accepted by Coworker	20
	biexpran – BI: Expectations: Accepted by Neighbor	21
	birelh – BI: Family in Country of Origin	21
	birelhpc – BI: Family Country of Origin: Parents	22
	birelhgp – BI: Family Country of Origin: Grandparents	22
	birelhcc – BI: Family Country of Origin: Children	23
	birelhbs – BI: Family Country of Origin: Brother,Sister	23
	birelhdr – BI: Family Country of Origin: Distant Relatives	24
	birelhsp – BI: Family Country of Origin: Spouse, Partner	24
	birelhfr – BI: Friends: Country of Origin	25
	birelhmi – BI: Persons Abroad Bring To Germany	25
	birelhs2 – BI: Spouse In Germany	26
	birelhcc2 – BI: Underage Children Not In Germany	26
	bigoback – BI: Go Back Home	27
	bistay – BI: Desire To Stay In Germany	27
	bistayy – BI: Years Desired To Stay In Germany	28
	biscger – BI: Attended School In Germany	29
	biscgrad – BI: Which Grade, Year In German School	30
	biscgerc – BI: Attended Special Foreigner Prep Class	30
	biscgc – BI: Also German Pupils In Class	31
	biscgcf – BI: How Many Pupils Foreign	31
	biscgcfn – BI: Mix Of Nationalities In Class	32
	bireason – BI: Main Reason Immigration To Ger. (since 2014)	32

year – Year, Survey Year	33
hid – Current Wave HH Number	34
cid – Original Household Number, Case ID	34
pid – Never Changing Person ID	35

1 Introducing the BIOIMMIG dataset

The 41 variables contained in the BIOIMMIG dataset are a small pool of selected, and mostly generated, variables that deal with questions related to the situation of immigrants and their descendants in Germany. These questions concern, for example, their reasons for immigrating to Germany, their living conditions upon arrival, their plans to return to their country of origin, and whether they have relatives in the country of origin.

BIOIMMIG was first generated in 1996 and aims to provide the data user with some pre-generated variables that might be useful when conducting migration and integration research. While some variables in the BIOIMMIG dataset are exact copies of migration specific variables recorded in the BIOL or PL dataset, other variables are generated by pulling together a multitude of variables from various datasets and therefore aim to alleviate the coding burden on the data user.

The data available in the BIOIMMIG dataset are longitudinal. This means that the same variable name is used for every year. The different years for which data in the BIOIMMIG file is available can be distinguished by using the variable SYEAR, ranging from 1984 to 2020. Thus, for every respondent included in the dataset, there are as many person-years as the number of years in which the respondent completed interviews. The main unit of observation is therefore person-years.

The sample population of the dataset is defined using the following criteria:

- First, the dataset includes only person-year combinations for those years in which SOEP respondents completed a questionnaire (visible via the variable \$NETTO in the dataset PPATHL). This refers not only to the personal and youth questionnaires but also to the biography questionnaires. In other words, at least one of these interviews had to be realized in a given year so that the person-year combination was considered for the BIOIMMIG dataset. At the same time, the dataset excludes person-year-observations from children and individuals who were not part of a SOEP household in a given year (for more information, see the PPATHL dataset documentation on the current wave's survey status of the variable \$NETTO).
- Second, all available information related to the BIOIMMIG variables is collected and assigned to the person-year observations. Thereafter, all respondents who do not have valid information **on any** variable that is included in the BIOIMMIG dataset are excluded from the dataset for this year.
- Third, all respondents whose only valid information available is the percentage of foreigners in their school class (variable BISCGCF “How Many Pupils Foreign”) were not included in the BIOIMMIG dataset. Most information in the BIOIMMIG dataset comes from migrant specific questions or questionnaires, while information on the variable BISCGCF comes from a question which is asked to all SOEP participants aged 16 to 17 (dataset JUGENDL) and those over 18 who completed a general biography questionnaire (dataset BIOL). Therefore, if the only valid information in the BIOIMMIG dataset was on the variable BISCGCF, it can be assumed that the person does not have a migration background.
- Fourth, all respondents who mentioned to be east German were excluded. This corresponds to value 1 on the variable BIIMGRP which is part of the BIOIMMIG dataset.

Ultimately, for v37 this leaves a sample size of **516,063** person-years and **76,530** individuals in the BIOIMMIG dataset. Please note that the BIOIMMIG sample population cannot be equated with the overall population of immigrants and their descendants in the SOEP,

since the BIOIMMIG dataset only includes respondents who gave information on at least one of the variables that are included in the BIOIMMIG dataset. Researchers who would like to include all respondents with an immigrant origin should therefore cross-check the MIGBACK variable in the PPATHL dataset to determine whether and to what extent their target population is covered in the BIOIMMIG dataset.

Table 1: Development of the sample size of the BIOIMMIG dataset

	Persons absolut	Persons in %	Person-years absolut	Persons-years in %
PPAHTL: Original sample size	157,785	100 %	1,227,349	100%
Interview years: Sample after keeping only those years in which a questionnaire was completed based on \$NETTO	105,521	67 %	753,015	61%
BIOIMMIG information: Sample after excluding all those without valid information in any year of participation	84,641	54 %	561,456	46%
BISCGCF: Sample after excluding all respondents whose only valid BIOIMMIG information is the share of foreigners in class	76,782	49 %	519,729	42%
BIIMGRP: Sample after excluding all east Germans	76,530	49 %	516,063	42%

Source: SOEP v37

Sources used to generate the BIOIMMIG dataset

The BIOIMMIG dataset is build based on some migration specific variables that are collected using individual and biography questionnaires in different samples. As data from new samples and corresponding variables are collected year on year, their information is being fed into the BIOIMMIG dataset. In general, the variables used are both from biographical data stored in the BIOL dataset and individual data stored in the PL dataset.

The following bullet points provides a concise overview on which data sources were used from when on to build the BIOIMMIG dataset.

- In 1994 and 1995 (wave K and L) the SOEP's first immigrant samples (Samples D1 & D2) have been created and information from those specific questionnaires have been used to create the first BIOIMMIG dataset in 1996.

- In 2000 (wave Q) the SOEP started sampling youth. Information from the youth dataset JUGEND, which contains similar information to the LELA dataset but for 16- and 17-year-olds, was incorporated into the BIOIMMIG dataset. In any given year, a person could complete either a youth or a biography questionnaire depending on their age. Thus when you look for the original information, it is either available in the JUGEND dataset or the LELA dataset but never in both.
- In 2013 (wave BD), the IAB-SOEP migration samples (subsamples M1 & M2) have been launched in order to collect a large number of interviews with migrants in Germany. Information from those samples' questionnaires have been added to BIOIMMIG in subsequent years.
- In 2016 and 2017 (wave BG and BH), three IAB-BAMF-SOEP refugee samples (subsamples M3, M4, M5) were launched and added to the overall SOEP sample. The questionnaires of these samples varied slightly from the previous individual and biography questionnaires as they were tailored to capture the specific circumstances of refugees. The information from subsamples M3-M5 have also been added to the BIOIMMIG dataset in subsequent years.
- In 2017, all the cross-sectional data collected through migrant and refugee samples M1-M5 were integrated into the long-format datasets PL and BIOL. From then on those longitudinal datasets have been used to extract the information that is needed to generate the variables of the BIOIMMIG dataset.
- In 2022 (wave BK), data from the newly released migration samples M6, M7 and M8 have been fed into the BIOIMMIG dataset for the first time.

2 Data structure and imputating data

Some of the information included in the BIOIMMIG dataset, especially information taken from the biography dataset BIOL, are assumed to be constant and valid over time. This could be, for example, a respondent's main reason for coming to Germany. In accordance, the respondent's answer is filled into all years going forward in which the respondent takes part in the SOEP survey, even if the respondent doesn't answer that specific question again in subsequent years. We proceeded in the same manner with questions that were posed to the respondent once in the youth questionnaire and then again in the adult questionnaire if the answers remained consistent. For example, when a respondent mentioned that he or she wanted to stay in Germany permanently, first in the youth questionnaire, and later in the biography questionnaire.

It is important to note that the practice of filling in seemingly time-invariant information for all person-years has some disadvantages that researchers should be aware of when conducting analysis with the BIOIMMIG dataset: Firstly, data users have no way of telling in which year the data was actually collected. Secondly, the person-year observations in any given year are artificially inflated which distorts the sample size in any given year. The practice of filling in seemingly time-invariant answers that have only been given once by the respondent is currently reviewed by the SOEP team and might change for v38.

However, some variables are time-variant and it can be assumed that answers and thus values may indeed vary over time. Examples of variables that are based on time-variant answers are the following: BISCGER (school in Germany), BISCGCF (number of foreign students), BIEXPR (expectations from Germany), BIEXPRAC (acceptance by coworkers), BIRELH (family abroad), BIRELHC2 (children living abroad), BIRELHS2 (spouse in Germany), BISTAY

(desire to stay in Germany), BISTAYY (planned duration of stay). For instance, the number of children living abroad might change over time. Thus, both (or more) pieces of information are considered valid for the subsequent years until a different answer is given, and this answer is again considered valid until the next different answer is given. The logic of this procedure is demonstrated for the variable BIRELH (family in home country) in Table 2. By filling in information for person-years in which the respondent hasn't actually answered the question at hand bears (to some extent) the same issues as described in the section above and is likewise reviewed by the SOEP team for v38.

Table 2: Data structure after filling in time-variant information over the years (fictitious data)

PERSNR	HHNR	HHNRAKT	SYEAR	BIRELH	BISTAYY
101	19	19	2000	1	10
101	19	19	2001	1	9
101	19	19	2002	1	8
101	19	19	2003	2	7
101	19	19	2004	2	6
101	19	19	2005	2	5

In this fictitious example, the person replied in 2000 that she had relatives in her home country, but in 2003 her answer changed to *no relatives in home country*. Therefore, the subsequent years are coded accordingly. The number of years she wanted to stay in Germany (BISTAYY) was also adjusted accordingly. The procedure was also adapted for the variable BISTAYY (planned length of stay in Germany in years) which varies per se across time. Therefore, the number of years indicated in one year is logically decremented in the subsequent years. For instance, a planned stay of 10 more years in Germany in 2000 would have been adjusted to nine more years in 2001 (also see BISTAYY in Table 2).

Missing values

Four missing values were assigned to the variables in the BIOIMMIG dataset as shown in Table 3.

- -1 (no answer) is assigned to respondents who received a question that is relevant for generating BIOIMMIG variables but did not answer it.
- -2 (does not apply) is assigned if no information is available for a respondent.
- -5 (not included in this version of questionnaire) is assigned to respondents who did not receive the sample-specific question. This applies mainly to respondents whose samples were included retrospectively in a dataset, such as respondents of the project specific sample FID ("Families in Germany") or respondents given variations of the questions relevant for generating BIOIMMIG variables, such as those in the migration samples in 2016 and 2017.
- -8 (question not part of the survey program this year) is only applicable to datasets in long format. This missing value applies when a question isn't part of the survey program in a given year.

The missing values were filled in across person-years according to the procedure described in the section about data structure and imputing data.

More information on the SOEP's general missing conventions can be found in the SOEP Companion

Table 3: Missing values

Value	Value label (English)
-8	Question not part of the survey program this year
-5	Not included in this version of questionnaire
-2	Does not apply
-1	No answer

Source: SOEP v37

Addaptions had to be made to the variables BISTAY (desire to stay in Germany) and BISTAYY (desired duration of stay). If a respondent replied in a given year that he or she wanted to stay (forever) or (for less than a year) on the respective BISTAY question, then the BISTAYY variable was coded -2 (does not apply). However, if a respondent mentioned in a subsequent year that he or she only wanted to stay in Germany for a specific period of time, the respective information was used to code the values of BISTAY and BISTAYY. Table 4 shows a fictitious example of a respondent changing his or her intention to stay in Germany.

Table 4: Example of inconsistent responses over time to question about desires to stay in Germany (fictitious data)

PERSNR	HHNR	HHNRAKT	SYEAR	BISTAY	BISTAYY
103	21	21	2000	2	10
103	21	21	2001	2	9
103	21	21	2002	1	-2
103	21	21	2003	1	-2
103	21	21	2004	2	15
103	21	21	2005	2	14

Using the BIOIMMIG dataset

The BIOIMMIG file can be used for cross-sectional or panel data analysis. The dataset is provided in *long* format but can easily be reshaped to *wide* format or, if necessary, be combined with cross-sectional data by simply selecting the relevant year via the variable SYEAR in the BIOIMMIG dataset. Relevant variables for merging other SOEP datasets provided in long format to the BIOIMMIG dataset are PID, HID, CID and SYEAR. For merging datasets provided in wide format first use SYEAR to select the cases of the relevant year in the BIOIMMIG dataset and then use PID (PERSNR) and HID (HHNRAKT) as merge variables. For example, merging the BIOIMMIG dataset for the year 1984 with another dataset could be achieved by proceeding in the following manner in Stata:

1. **Open** the BIOIMMIG dataset for 1984:

```
use bioimmig if syear==1984
```

2. **Rename** variables with wave-specific information:

```
rename bicamp camp1984
rename bicampw campw1984
rename (hid pid) (hhnrakt persnr)
```

3. **Save** IDs & renamed variables in a temporary file:

```
sort hhnrakt persnr
keep hhnrakt persnr syear camp1984 campw1984
save bioimmig1984.dta, replace
```

3 Variables, Values and Sources in File BIOIMMIG

In this section you will find more detailed information on all 41 variables of the BIOIMMIG dataset. The information that you will find is the following:

- Firstly, the name and label of the variable. For example *biimgrp* - *BI: Immigration Group* or *biresper* - *BI: Residence Status*.
- Secondly, a question which relates to the variable. Please note that the question has been added in order to give users a better idea about the content of the variable. The question is **not** the exact question from which the variable is derived. The reason for that is quite simple: as you will see, many variables in the BIOIMMIG dataset have been generated from a multitude of variables, all of which might have slightly different question texts but speak to the same topic. Therefore it is not possible to list the exact question text of all variables that have been used to generate the variable in the BIOIMMIG dataset.
- Thirdly, the values of the variables and the person-year count for each of the values.
- Lastly, a table that lists all source variables that have been used to generate the BIOIMMIG variable in question. When there are a lot of variables in one table, this means that all those variables speak to the same concept (for example the concept of which type of immigrant group someone belongs to for variable BIIMGRP) and therefore have been brought together in this one BIOIMMIG variable. In the table the source variables are sorted according to the SOEP dataset in which they can originally be found (for example, BIOL, PL, JUGENDL). From year to year the variables that are used to generate one BIOIMMIG variable might change. This would be recorded in this table. The reason why the variables might change from year to year is mostly due to new variables being added as described in the section on *sources used to generate the BIOIMMIG dataset*. In rare cases it might be that some variables are not deemed appropriate anymore to be included in the generation of a BIOIMMIG variable or that the variable hasn't been asked in the year at question and can therefore not be included. Based on the amount of variables in the table it is also easy to deduct which variables require a lot of effort to be generated (like for example BIIMGRP) and which BIOIMMIG variables are mere copies of variables that are already available in the BIOL, PL or JUGENDL datasets (like for example BIRBETR).

biimgrp – BI: Immigration Group

? There are different judicial principles for moving in to Germany. Which of the following immigrant categories did you belong to when you moved to Germany? (from: soep-core/soep-core-2016-11/6:lzug02[4317])

2	Person Of German Descent From Eastern Europe	31545
3	German Who Lived Abroad	1425
4	Citizen Of EU Country (up to 2009 EC)	9014
5	Asylum seeker, refugee	30301
6	Other Foreigner	25592
-1	No Answer	458
-2	Does not apply	351897
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	65831
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	0

Table 5: Variables used to generate biimgrp

File	Variable
BIOL	lb0020_v1, lb0020_v2, lb0020_v3
	lb0013_h
	lm0020
	lm0020i01
	lm0020i04
	lm0020i05_v1, lm0020i05_v2, lm0020i05_v3
	lm0020i06
	lm0031i01l01 - lm0031i01l10
	lm0031i03l01 - lm0031i03l07
	lm0031i04l01 - lm0031i04l08
	lm0031i07l01_v1, lm0031i07l01_v2
	lm0031i07l02_v1, lm0031i07l02_v2
	lm0031i07l03_v1, lm0031i07l03_v2
	lm0031i07l04 - lm0031i07l09
	lm0031i08l01_v1 - lm0031i08l01_v2, lm0031i08l01_v3
	lm0031i08l02_v1 - lm0031i08l02_v2, lm0031i08l02_v3
	lm0031i08l03_v1 - lm0031i08l03_v2, lm0031i08l03_v3
	lm0031i08l04 - lm0031i08l07
	lm0031i10l01 - lm0031i10l09
	JUGENDL
PL	plj0024_h
	plj0175
	plj0006
	plj0003

biresper – BI: Residence Status

? What is your current residence status? (from: soep-core/soep-core-2016-ll/7:lstat[4319])

? What is your residence status in Germany? (from: soep-core/soep-core-2016-pb-m12-erst/50:lstat[3949])

1	Unlimited	29795
2	Temporary	30877
-1	No Answer	302
-2	Does not apply	398705
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	53471
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	2913

Table 6: Variables used to generate Biresper

File	Variable
BIOL	lb0021
	lb1240_v1, lb1240_v2
PL	plj0680_v1, plj0680_v2
	plj0722
	plm0365
JUGENDL	jl0246
	jl0445

bicamp – BI: Refugee Residence Y,N

? Which of the following would best describe your living situation immediately after moving to Germany? (from: soep-core/soep-core-2014-ll/8;lzugw[3801])

1	Yes	29261
2	No	31185
-1	No Answer	296
-2	Does not apply	303721
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	146598
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	5002

Table 7: Variables used to generate Bicamp

File	Variable
BIOL	lb0022
	lb1249
	lr3236
	lr3440-lr3452
PL	plm0679
HL	hlj0005_h
JUGENDL	jl0253

bicampw – BI: Refugee Residence: Weeks

0	10
1	997
2	1535
3	680
4	669
5	163
6	391
7	4
8	142
10	53
12	44
14	14
15	6
16	3
26	3
32	16
36	13
40	12
41	13
-1	17
-2	358844
-5	57393
-8	95041

Table 10: Variables used to generate *Bicampw*

File	Variable
BIOL	lb0023
JUGENDL	jl0254

bicampm – BI: Refugee Residence: Months

0	10
1	1119
2	848
3	588
4	268
5	378
6	1002
7	358
8	340
9	284
10	286
11	184
12	1176
13	63
14	185
... (21 rows omitted)	2563
38	13
40	1
42	54
48	72
54	1
60	22
66	2
72	29
77	3
90	9
99	27
-1	158
-2	353804
-5	57393
-8	94823

Table 11: Variables used to generate Bicampm

File	Variable
BIOL	lb0024
JUGENDL	jl0255

biwfam – BI: Already Had Family In Country

? There are different legal frameworks for immigrating to Germany. What about you: What was your immigration status at the time of your arrival in Germany? (*from: soep-core/soep-core-2019-113/10:1z01s1[29143]*)

? There are different legal frameworks for immigrating to Germany. What about you: What was your immigration status at the time of your arrival in Germany? (from: [soep-core/soep-core-2019-ll3/44:l3z01s1\[29179\]](#))

? What were your reasons for specifically moving to Germany and not to another country? // Because family members live here (from: [soep-core/soep-core-2019-pb-m345-erst/Q079:lgrd1\[15078\]](#))

1	Yes	31399
-1	No Answer	341
-2	Does not apply	227832
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	118849
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	137642

Table 12: Variables used to generate Biwfam

File	Variable
BIOL	lb0020_v3
	lb0025
	lb0034
	lm0020
	lm0020i01
	lm0020i05_v1, lm0020i05_v2
	lm0020i06
	lm0031i01l01 - lm0031i01l09
	lm0031i03l01 - lm0031i03l09
	lm0031i04l01 - lm0031i04l09
	lm0031i08l01_v1, lm0031i08l01_v2, lm0031i08l01_v3
	lm0031i08l02_v1, lm0031i08l02_v2, lm0031i08l02_v3
	lm0031i08l03_v1, lm0031i08l03_v2, lm0031i08l03_v3
	lm0031i08l04 - lm0031i08l09
	lm0031i10l01 - lm0031i10l109
	lr3168
	JUGENDL

bifamc – BI: Contacts With Family In Germany

? At your arrival, did you have any support of relatives or acquaintances who already lived in Germany? // Yes, by relatives (from: [soep-core/soep-core-2016-ll/10:lzugunt1\[4321\]](#))

? When you moved to Germany, did you have the help of any relatives or friends who already lived in Germany? // Yes, from relatives (from: [soep-core/soep-core-2016-pb-m12-erst/26:lzugunt1\[8460\]](#))

? Did any relatives or acquaintances who already lived in Germany help you move to Germany? // Yes, relatives (from: [soep-core/soep-core-2016-pb-m34/38:lzugunt1\[12237\]](#))

1	Yes	16133
2	No	7152
-1	No Answer	55
-2	Does not apply	378772
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	113951
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	0

Table 13: Variables used to generate *Bifamc*

File	Variable
BIOL	lb0026
	lb1246
	lb1247_v3
	lb0865
	lb0866
	lm0034
	lb1247_v1, lb1247_v2, lb1247_v4
JUGENDL	jl0257

bifamcl – BI: Moved To Same City,Town As Family

1	Yes	21249
2	No	5502
-1	No Answer	385
-2	Does not apply	363900
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	116063
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	8964

Table 14: Variables used to generate *Bifamcl*

File	Variable
BIOL	lb0027
	lm0495i01_v1, lm0495i01_v2
	lr3470-3480
JUGENDL	jl0258

birbetr – BI: Reason Migrate: Better

? Which of the following reasons was the main reason for you to come to Germany? (from: [soep-core/soep-core-2014-ll/9;lzuggr\[4320\]](#))

1	Better Life	6893
-1	No Answer	0
-2	Does not apply	356578
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	57393
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	95199

Table 15: Variables used to generate Birbetr

File	Variable
BIOL	lb0031

birmoney – BI: Reason Migrate: Money

? Which of the following reasons was the main reason for you to come to Germany? (from: [soep-core/soep-core-2014-ll/9;lzuggr\[4320\]](#))

1	Earn Money	3903
-1	No Answer	0
-2	Does not apply	359461
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	57393
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	95306

Table 16: Variables used to generate Birmoney

File	Variable
BIOL	lb0032

birfree – BI: Reason Migrate: Freedom

1	Live In Freedom	5951
-1	No Answer	0
-2	Does not apply	357423
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	57393
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	95296

Table 17: Variables used to generate Birfree

File	Variable
BIOL	lb0033

birfam – BI: Reason Migrate: Family

? Which of the following reasons was the main reason for you to come to Germany? (from: [soep-core/soep-core-2014-ll/9;lzuggr\[4320\]](#))

1	Live With Family	10920
-1	No Answer	0
-2	Does not apply	352899
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	57393
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	94851

Table 18: Variables used to generate Birfam

File	Variable
BIOL	lb0034

birpoor – BI: Reason Migrate: Poor

? Which of the following reasons was the main reason for you to come to Germany? (from: [soep-core/soep-core-2014-ll/9;lzuggr\[4320\]](#))

1	Poverty In Country Of Origin	3370
---	------------------------------	------

-1	No Answer	0
-2	Does not apply	359932
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	57393
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	95368

Table 19: Variables used to generate Birpoor

File	Variable
BIOL	Ib0035

birwar – BI: Reason Migrate: War

? Which of the following reasons was the main reason for you to come to Germany? (from: [soep-core/soep-core-2014-ll/9;zuggr\[4320\]](#))

1	War In Country Of Origin	3465
-1	No Answer	0
-2	Does not apply	359909
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	57393
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	95296

Table 20: Variables used to generate Birwar

File	Variable
BIOL	Ib0036

birjust – BI: Reason Migrate: Just So

1	Just Live In Germany	6413
-1	No Answer	0
-2	Does not apply	219281
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	57393
-6	Version of questionnaire with modified filtering	0

-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	232976

Table 21: Variables used to generate Birjust

File	Variable
BIOL	lb0037_h

birothr – BI: Reason Migrate: Other

? Which of the following reasons was the main reason for you to come to Germany? (from: [soep-core/soep-core-2014-ll/9;lzuggr\[4320\]](#))

1	Other Reasons	5808
-1	No Answer	378
-2	Does not apply	357379
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	57393
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	95105

Table 22: Variables used to generate Birothr

File	Variable
BIOL	lb0038

biexpr – BI: Expectations In Germany

1	Yes	14571
2	Partly	7950
3	No, Not At All	790
-1	No Answer	1275
-2	Does not apply	339861
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	57393
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	94223

Table 23: Variables used to generate Biexpr

File	Variable
BIOL	Ib0040

biexprlv – BI: Expectations: Find Apt

1	Harder	5829
2	As Expected	8877
3	Easier	6485
4	Does Not Apply	4393
-1	No Answer	1132
-2	Does not apply	200226
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	57393
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	231728

Table 24: Variables used to generate Biexprlv

File	Variable
BIOL	Ib0041_h

biexprac – BI: Expectations: Accepted by Coworker

1	Harder	5424
2	As Expected	8471
3	Easier	3294
4	Does Not Apply	8138
-1	No Answer	1285
-2	Does not apply	200299
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	57393
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	231759

Table 26: Variables used to generate Biexprac

File	Variable
BIOL	lb0042_h

biexpran – BI: Expectations: Accepted by Neighbor

1	Harder	6027
2	As Expected	11172
3	Easier	3688
4	Does Not Apply	2912
-1	No Answer	1129
-2	Does not apply	202000
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	57393
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	231742

Table 27: Variables used to generate Biexpran

File	Variable
BIOL	lb0043_h

birelh – BI: Family in Country of Origin

1	Yes	52260
2	No	1862
-1	No Answer	1746
-2	Does not apply	131689
-3	Answer improbable	20
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	48130
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	280356

Table 28: Variables used to generate Birelh

File	Variable
BIOL	lb0044
JUGENDL	jl0262

birelhp – BI: Family Country of Origin: Parents

? Where does your mother currently live? (*from: soep-core/soep-core-2019-pb-m345-erst/Q495:lmleb [15490]*)

? Where does your father currently live? (*from: soep-core/soep-core-2019-pb-m345-erst/Q503:lvleb [15499]*)

1	Parents	19577
-1	No Answer	231
-2	Does not apply	167219
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	52305
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	276731

Table 29: Variables used to generate Birelhp

File	Variable
BIOL	lb0045
	lb0078_v1
	lb0079_v1
	lr3213
	lr3215
JUGENDL	jl0263
PL	plj0118_v1
	plj0119_v1

birelhgp – BI: Family Country of Origin: Grandparents

1	Grandparents	9092
-1	No Answer	360
-2	Does not apply	160806
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	345805

Table 30: Variables used to generate Birelhgp

File	Variable
BIOL	lb0046

File	Variable
JUGENDL	jl0264
PL	pld0017_v1

birelhc – BI: Family Country of Origin: Children

1	Children	1967
-1	No Answer	54
-2	Does not apply	26349
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	117343
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	370350

Table 31: Variables used to generate *Birelhc*

File	Variable
BIOL	lb0047
	lr3193
	lr3195
	lr3198
	lr3200
	lr3204
	lr3206
	lr3208
	lr3211
PL	plj0641

birelhbs – BI: Family Country of Origin: Brother,Sister

1	Brother,Sister	28310
-1	No Answer	124
-2	Does not apply	162370
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	52371
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	272888

Table 32: Variables used to generate *Birelhbs*

File	Variable
BIOL	lb0048 lr3186
JUGENDL	jl0266
PL	plj0124_v1 plj0125_v1

birelhdr – BI: Family Country of Origin: Distant Relatives

1	Distant Relatives	29931
-1	No Answer	1092
-2	Does not apply	155273
-3	Answer improbable	20
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	49702
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	280045

Table 33: Variables used to generate *Birelhdr*

File	Variable
BOL	lb0049 lr3224
JUGENDL	jl0267
PL	plj0130_v1 plj0645

birelhsp – BI: Family Country of Origin: Spouse, Partner

1	Spouse/Fiancee	192
-1	No Answer	4
-2	Does not apply	179478
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	2515
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	333874

Table 34: Variables used to generate *Birelhsp*

File	Variable
BIOL	lb0050
	lm0224_v1
PL	plm0224
	plj0627_v1
	plj0630
	plj0643

birelhfr – BI: Friends: Country of Origin

1	Personal Acquaintance	3812
-1	No Answer	0
-2	Does not apply	166389
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	345862

Table 35: Variables used to generate *Birelhfr*

File	Variable
BIOL	lb0051
JUGENDL	jl0269

birelhmi – BI: Persons Abroad Bring To Germany

1	Yes	3373
2	No	4807
3		28
-1	No Answer	86
-2	Does not apply	38554
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	58579
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	410636

Table 36: Variables used to generate *Birelhmi*

File	Variable
BIOL	lb0052_h
	lr3194
	lr3197
	lr3199
	lr3201
	lr3205
	lr3207
	lr3210
	lr3212
	JUGENDL
PL	plj0628_v1, plj0628_v2
	plj0631_v1, plj0631_v2

birelhs2 – BI: Spouse In Germany

1	Germany, Here In HH	50926
2	Germany, Not In HH	3002
3	Not In Germany	3471
-1	No Answer	49
-2	Does not apply	11258
-3	Answer improbable	1
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	141666
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	305690

Table 37: Variables used to generate Birelhs2

File	Variable
PL	plj0093
	plj0627
	plj0630

birelhc2 – BI: Underage Children Not In Germany

1	Yes	9424
2	No	224001
-1	No Answer	26
-2	Does not apply	254867
-3	Answer improbable	1
-4	Inadmissible multiple response	0

-5	Not included in this version of the questionnaire	4156
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	23588

Table 38: Variables used to generate Birelhc2

File	Variable
PL	plj0097
	plj0159
	plj0099
	plj0642
	plj0137_h

bigoback – BI: Go Back Home

1	Yes, Definitely	419
2	Yes, Probably	955
3	Probably Not	2042
4	No, Definitely Not	6891
-1	No Answer	122
-2	Does not apply	159816
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	345818

Table 39: Variables used to generate Bigoback

File	Variable
BIOL	lb0053
JUGENDL	jl0271

bistay – BI: Desire To Stay In Germany

? Do you want to stay in Germany forever? (from: [soep-core/soep-core-2019-pb-m345-erst/Q106:pnat10\[15113\]](#))

? How long would you like to stay in Germany? (from: [soep-core/soep-core-2019-pb-m345-erst/Q107:pnat11\[15114\]](#))

1	One Year Maximum	1635
---	------------------	------

2	Several Years, Specifically...	37514
3	Stay In Germany Permanently	130819
-1	No Answer	1003
-2	Does not apply	203458
-3	Answer improbable	1
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	106407
-6	Version of questionnaire with modified filtering	25
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	35201

Table 40: Variables used to generate Bistay

File	Variable
PL	plj0084
	plj0085_v1, plj0085_v2
	plj0086_v1, plj0086_v2

bistay – BI: Years Desired To Stay In Germany

? How many years would you like to stay in Germany? // years (from: soep-core/soep-core-2019-pb-m345-erst/Q108:pnat12[15115])

1	2106
2	2535
3	2363
4	2104
5	2276
6	1334
7	1276
8	1362
9	1469
10	2074
11	366
12	414
13	437
14	468
15	664
...	(36 rows omitted)
58	1
59	1
60	3
61	1
65	1
66	1
85	1
89	2

98	1
99	1
-1	1650
-2	486910
-3	2
-5	1845
-8	1038

Table 41: Variables used to generate Bistayy

File	Variable
PL	plj0087

biscger – BI: Attended School In Germany

- ? Where did you last attend school? (from: soep-core/soep-core-2019-ll3/144:lsab3[29276])
- ? What federal state is your last school located in today? (from: soep-core/soep-core-2019-ll3/148:lschbund[29279])
- ? Was the last school you attended in Germany or in another country? (from: soep-core/soep-core-2019-pb-m345-erst/Q282:lsab3_r[15291])

1	Yes	299030
2	No	94660
-1	No Answer	781
-2	Does not apply	115453
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	1310
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	4829

Table 42: Variables used to generate Biscger

File	Variable
BIOL	lb0028
	lb0186_v1 - lb0186_v3
	lb0405
	lb1375
JUGENDL	jl0259

biscgrad – BI: Which Grade, Year In German School

1	7010
2	619
3	636
4	600
5	706
6	693
7	596
8	749
9	619
10	381
11	142
12	52
13	77
-1	431
-2	350568
-5	57393
-8	94791

Table 43: Variables used to generate Biscgrad

File	Variable
BIOL	lb0029
	lb0406
JUGENDL	jl0260

biscgerc – BI: Attended Special Foreigner Prep Class

1	Yes	4664
2	No	12931
-1	No Answer	104
-2	Does not apply	346232
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	57393
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	94739

Table 44: Variables used to generate Biscgerc

File	Variable
BIOL	lb0030
	lb0542

File	Variable
PL	plj0030
JUGENDL	jl0261

biscgc – BI: Also German Pupils In Class

1	Yes	6093
2	No	294
-1	No Answer	17
-2	Does not apply	84037
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	78853
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	346769

Table 45: Variables used to generate Biscgc

File	Variable
BIOL	lb0054
PL	plj0049

biscgcf – BI: How Many Pupils Foreign

? (from: [soep-core/soep-core-2019-ju-lgb/41:jausl\[28931\]](#))

? If you think back to the last grade of school you attended: How many of your fellow students or their parents were not from Germany? (from: [soep-core/soep-core-2019-ju3/41:jausl\[28912\]](#))

? Thinking about your classmates in the last year of school you attended, or your current classmates if you're still in school: How many of your classmates were/are not from Germany or have parents who were/are not from Germany—in other words, how many have a family history of migration? (from: [soep-core/soep-core-2019-ll3/151:lsab10\[29282\]](#))

1	All	1949
2	Most	7412
3	Around 1/2	9724
4	Around 1/4	16294
5	Less Than 1/4	91086
6	No One	163051
-1	No Answer	3992
-2	Does not apply	137046
-3	Answer improbable	0

-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	85424
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	85

Table 46: Variables used to generate *Biscgcf*

File	Variable
BIOL	lb0055
	lb0194_h
PL	plj0050
JUGENDL	jl0176_h

biscgfn – BI: Mix Of Nationalities In Class

1	Only My Nationality	197
2	Mixed	92
-1	No Answer	3
-2	Does not apply	168951
-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	0
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	346820

Table 47: Variables used to generate *Biscgfn*

File	Variable
BIOL	lb0056

bireason – BI: Main Reason Immigration To Ger. (since 2014)

? Which of the following was your main reason for moving to Germany? (*from: soep-core/soep-core-2019-ll3/53:lzuggr[29187]*)

1	Family Reasons/Partnership	11030
2	Economic Reasons	6884
3	Political Reasons	2655
4	Other Reasons	1766
-1	No Answer	145
-2	Does not apply	25474

-3	Answer improbable	0
-4	Inadmissible multiple response	0
-5	Not included in this version of the questionnaire	48997
-6	Version of questionnaire with modified filtering	0
-7	Only available in less restricted edition	0
-8	Question not part of the survey program this year	419112

Table 48: Variables used to generate Bireason

File	Variable
BIOL	lb1245_v1, lb1245_v2

syear – Year, Survey Year

1984	4876
1985	4460
1986	4419
1987	4457
1988	4367
1989	4347
1990	5748
1991	5765
1992	5791
1993	5852
1994	6096
1995	6488
1996	6402
1997	6290
1998	6648
... (7 rows omitted)	97240
2006	15679
2007	14929
2008	14037
2009	14870
2010	21594
2011	23783
2012	23632
2013	26736
2014	23502
2015	23274
2016	25705
2017	28919
2018	26759
2019	26494
2020	26904

Survey Year

hid – Current Wave HH Number

60	1
167	15
175	47
213	13
230	40
396	42
418	15
477	27
523	37
531	17
566	29
574	28
604	49
647	13
728	14
... (47245 rows omitted)	515661
5538388	1
5538418	1
5538523	1
5538566	1
5538590	1
5538744	1
5538779	1
5539180	1
5539198	1
5539414	1
5539546	1
5539600	1
5539660	1
5540200	1
5540404	1

ID variable Households

cid – Original Household Number, Case ID

60	85
167	15
175	47
213	13
230	73
280	4
396	42
418	15
477	27
523	37
531	36

566	29
574	28
582	29
590	30
... (41028 rows omitted)	515538
5538388	1
5538418	1
5538523	1
5538566	1
5538590	1
5538744	1
5538779	1
5539180	1
5539198	1
5539414	1
5539546	1
5539600	1
5539660	1
5540200	1
5540404	1

ID variable Households

pid – Never Changing Person ID

601	19
1602	12
1603	3
1701	24
1704	23
2101	13
2301	37
2302	35
2305	1
3901	21
3902	21
4101	15
4701	27
5201	37
5302	22
... (76500 rows omitted)	515738
55383801	1
55384101	1
55385201	1
55385601	1
55385901	1
55387401	1
55387701	1
55391801	1
55391901	1

55394101	1
55395401	1
55396001	1
55396601	1
55402001	1
55404001	1

ID variable Persons