DSDR DATA SHARING FOR DEMOGRAPHIC RESEARCH

ICPSR 27063

China Multi-Generational Panel Dataset, Liaoning (CMGPD-LN), 1749-1909

James Z. Lee Hong Kong University of Science and Technology. School of Humanities and Social Science

Cameron D. Campbell University of California-Los Angeles. Department of Sociology, and California Center for Population Research

Liaoning Basic File Codebook

Inter-university Consortium for Political and Social Research P.O. Box 1248 Ann Arbor, Michigan 48106 www.icpsr.umich.edu

Terms of Use

The terms of use for this study can be found at: http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/27063/terms

Information about Copyrighted Content

Some instruments administered as part of this study may contain in whole or substantially in part contents from copyrighted instruments. Reproductions of the instruments are provided as documentation for the analysis of the data associated with this collection. Restrictions on "fair use" apply to all copyrighted content. More information about the reproduction of copyrighted works by educators and librarians is available from the United States Copyright Office.

NOTICE WARNING CONCERNING COPYRIGHT RESTRICTIONS

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use," that user may be liable for copyright infringement.

ICPSR 27063-0001 PROCESSING NOTES

- 1) Because value labels cannot be assigned to character variables in Stata, the following variables will not contain value labels within Stata data files: FATHER_ID, GRANDFATHER_ID, HUSBAND_ID, WIFE_1_ID, and WIFE_2_ID.
- 2) Variable labels were added to the following variables based on correspondence with the Principal Investigators' research staff: RECORD_NUMBER, PERSON_ID, MOTHER_ID, FATHER_ID, FATHER_ID_IMPUTED, GRANDFATHER_ID, GRANDFATHER_ID_IMPUTED, WIFE_1_ID, WIFE_2_ID, HUSBAND_ID, YEAR, REGISTER_SEQ, REGION, DISTRICT, UNIQUE_VILLAGE_ID, RELATIONSHIP, GENERATION, BIRTHYEAR, BROTHER_COUNT, SISTER_COUNT, MALE_COUSIN_COUNT, FEMALE_COUSIN_COUNT, UNCLE_COUNT, AUNT_COUNT, FATHER_ALIVE, MOTHER_ALIVE, SON_COUNT, DAUGHTER_COUNT, DIED, MARRIED_OUT, REMARRIED_OUT, NEXT_BOYS, NEXT_GIRLS, AT_RISK_DIE, AT_RISK_MARRY, AT_RISK_REMARRY, ABSCONDED, NO_STATUS, RETIRED, OLD, ZU_ZHANG, ARTISAN, and SOLDIER.
- 3) Value labels were added to the following variables based on correspondence with the Principal Investigators' research staff: FATHER_ID, FATHER_ID_IMPUTED, GRANDFATHER_ID, GRANDFATHER_ID_IMPUTED, WIFE_1_ID, WIFE_2_ID, HUSBAND_ID, HOUSEHOLD_SEQ, DISTRICT, UNIQUE_VILLAGE_ID, SEX, MARITAL_STATUS, AGE_IN_SUI, BIRTHYEAR, BROTHER_COUNT, SISTER_COUNT, MALE_COUSIN_COUNT, FEMALE_COUSIN_COUNT, UNCLE_COUNT, AUNT_COUNT, FATHER_ALIVE, MOTHER_ALIVE, SON_COUNT, DAUGHTER_COUNT, PRESENT, DIED, MARRIED_OUT, REMARRIED_OUT, NEXT_DIE, NEXT_MARRY, NEXT_REMARRY, NEXT_ABSCONDED, NEXT_BOYS, NEXT_GIRLS, NEXT_3, NEXT_6, AT_RISK_DIE, AT_RISK_MARRY, AT_RISK_REMARRY, ABSCONDED, NO_STATUS, RETIRED, OLD, POSITION, ZU_ZHANG, ARTISAN, BAZI_YEAR, BAZI_MONTH, BAZI_DAY, and BAZI_HOUR.
- 4) Missing values were specified for all variables based on correspondence with the Principal Investigators' research staff. For all variables, the values of -98 and -99 were coded as missing and given the value labels of "Not Applicable" and "Missing" respectively.

ICPSR 27063

China Multi-Generational Panel Dataset, Liaoning (CMGPD-LN), 1749-1909

Variable Description and Frequencies

Note: Frequencies displayed for the variables are not weighted. They are purely descriptive and may not be representative of the study population. Please review any sampling or weighting information available with the study.

Summary statistics (minimum, maximum, mean, median, and standard deviation) may not be available for every variable in the codebook. Conversely, a listing of frequencies in table format may not be present for every variable in the codebook either. However, all variables in the dataset are present and display sufficient information about each variable. These decisions are made intentionally and are at the discretion of the archive producing this codebook.

ICPSR has an FAQ on copyright and survey instruments.

Liaoning Basic File

RECORD_NUMBER: RECORD I.D.

Sequential record identifier. Value generated by specific computer program.

Based upon 1,513,357 valid cases out of 1,513,357 total cases.

Location: 1-9 (width: 9; decimal: 0) Variable Type: character

PERSON_ID: UNIQUE I.D. OF INDIVIDUAL

Unique ID identifying all the records of a specific individual. Value generated by specific computer program by concatenating links between registers assigned by coders.

Based upon 1,513,357 valid cases out of 1,513,357 total cases.

Location: 10-17 (width: 8; decimal: 0) *Variable Type:* character

MOTHER_ID: MOTHER'S I.D.

PERSON_ID of mother for males and never-married females. For married or widowed women, refers to husband's mother. Value generated by specific computer program.

Based upon 1,513,357 valid cases out of 1,513,357 total cases.

Location: 18-25 (width: 8; decimal: 0) *Variable Type:* character

FATHER_ID: FATHER'S I.D.

PERSON_ID of father for males and never-married females. For married or widowed women, refers to husband's father. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
	Missing Data		
-99	Missing	-	-
	Total	1,513,357	100%

Based upon 1,470,511 valid cases out of 1,513,357 total cases.

- Mean: 154492.06
- Minimum: 3
- Maximum: 439000
- Standard Deviation: 108754.58

Location: 26-33 (width: 8; decimal: 0) Variable Type: numeric (Range of) Missing Values: -99

FATHER_ID_IMPUTED: FATHER'S I.D. WAS IMPUTED

Indicates that FATHER_ID in this observation refers to a person not in the dataset inferred from family relationships in the raw data. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	1376064	90.9 %
1	Yes	137293	9.1 %
	Total	1,513,357	100%

Based upon 1,513,357 valid cases out of 1,513,357 total cases.

• Minimum: 0

• Maximum: 1

Location: 34-34 (width: 1; decimal: 0) *Variable Type:* numeric

GRANDFATHER_ID: GRANDFATHER'S I.D.

PERSON_ID of paternal grandfather for males and never-married females. For married or widowed women, refers to husband's grandfather. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
	Missing Data		
-99	Missing	-	-
	Total	1,513,357	100%

Based upon 1,401,105 valid cases out of 1,513,357 total cases.

- Mean: 190311.11
- Minimum: 3
- Maximum: 439018
- Standard Deviation: 133322.68

Location: 35-42 (width: 8; decimal: 0) Variable Type: numeric (Range of) Missing Values: -99

GRANDFATHER_ID_IMPUTED: GRANDFATHER'S I.D. WAS IMPUTED

Indicates that GRANDFATHER_ID in this observation refers to a person not in the dataset inferred from family relationships in the raw data. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	1201028	79.4 %
1	Yes	312329	20.6 %
	Total	1,513,357	100%

Based upon 1,513,357 valid cases out of 1,513,357 total cases.

• Minimum: 0

• Maximum: 1

Location: 43-43 (width: 1; decimal: 0) *Variable Type:* numeric

WIFE_1_ID: WIFE'S I.D.

PERSON_ID of wife. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
	Missing Data		
-99	Missing	-	-
-98	Not applicable	-	-
	Total	1,513,357	100%

Based upon 418,277 valid cases out of 1,513,357 total cases.

- Mean: 130974.08
- Minimum: 2
- Maximum: 266089
- Standard Deviation: 77053.35

Location: 44-51 (width: 8; decimal: 0) Variable Type: numeric (Range of) Missing Values: -99 , -98

WIFE_2_ID: SECOND WIFE'S I.D.

PERSON_ID of second wife. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
15009	-	1	0.0 %
21116	_	1	0.0 %
23179	-	1	0.0 %
27796	-	2	0.0 %
28237	-	1	0.0 %
29275	-	1	0.0 %
36505	-	2	0.0 %
37983	-	1	0.0 %
59801	-	1	0.0 %
61722	-	1	0.0 %
65157	-	2	0.0 %
76548	-	4	0.0 %
81591	-	1	0.0 %
94789	_	1	0.0 %
95628	-	1	0.0 %
110080	-	2	0.0 %
113976	-	1	0.0 %
129798	-	1	0.0 %
134315	-	1	0.0 %
142669	-	2	0.0 %
160490	-	1	0.0 %

Value	Label	Unweighted Frequency	%
161084	-	1	0.0 %
163736	-	8	0.0 %
165222	-	1	0.0 %
173822	-	1	0.0 %
190616	-	1	0.0 %
190800	-	1	0.0 %
197895	-	2	0.0 %
200148	-	2	0.0 %
201780	-	5	0.0 %
202006	-	2	0.0 %
210914	-	1	0.0 %
216580	-	1	0.0 %
216845	-	2	0.0 %
223559	-	10	0.0 %
236689	-	2	0.0 %
236851	-	1	0.0 %
242921	-	2	0.0 %
244843	-	8	0.0 %
247289	-	1	0.0 %
247406	-	2	0.0 %
247497	-	2	0.0 %
247789	-	2	0.0 %
251454	-	2	0.0 %
251457	-	2	0.0 %
260049	-	1	0.0 %
261180	-	1	0.0 %
261819	-	1	0.0 %
265090	-	1	0.0 %
	Missing Data		
-99	Missing	448823	29.7 %
-98	Not applicable	1064439	70.3 %
	Total	1,513,357	100%

Please note that only the first 50 response categories are displayed in the PDF codebook. To view all response categories, please analyze the data file in the statistical package of your choice (SAS, SPSS, Stata, R).

Based upon 95 valid cases out of 1,513,357 total cases.

- Mean: 174696.49
- Median: 201780.00
- Mode: 223559.00
- Minimum: 15009
- Maximum: 265090
- Standard Deviation: 75346.24

HUSBAND_ID: HUSBAND'S I.D.

PERSON_ID of a woman's husband. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
	Missing Data		
-99	Missing	-	-
-98	Not applicable	-	-
	Total	1,513,357	100%

Based upon 400,437 valid cases out of 1,513,357 total cases.

- Mean: 130084.10
- Minimum: 1
- Maximum: 266010
- Standard Deviation: 76801.49

Location: 60-67 (width: 8; decimal: 0) Variable Type: numeric (Range of) Missing Values: -99, -98

HOUSEHOLD_ID: HOUSEHOLD I.D.

Identifies the records of a household in a specific year, for grouping purposes. Based on ZU_SEQ and HOUSEHOLD_SEQ. Value generated by specific computer program.

Based upon 1,513,357 valid cases out of 1,513,357 total cases.

Location: 68-80 (width: 13; decimal: 0) *Variable Type:* character

DATASET: DATASET

Name of the specific state farm population dataset recorded on the title page of all dataset registers. LMGPD comprises 29 such state farm population datasets. Data transcribed from original population registers.

Value	Label	Unweighted Frequency	%
1	Daoyi	120747	8.0 %
2	Gaizhou	42819	2.8 %
3	Dami	32106	2.1 %
4	Chengnei	55678	3.7 %
5	Mianding	56048	3.7 %
6	Niuzhuang	54870	3.6 %
7	Feicheng	71365	4.7 %
8	Manhan	53284	3.5 %
9	Dadianzi	76986	5.1 %
10	Guosantun	35081	2.3 %

Value	Label	Unweighted Frequency	%
11	Bakeshu	48760	3.2 %
12	Daxingtun	86935	5.7 %
13	Nianma dahai	58224	3.8 %
14	Changzhaizi	48800	3.2 %
15	Zhaohuatun	50862	3.4 %
16	Diaopitun	80512	5.3 %
17	Langjiabao	47331	3.1 %
18	Wangzhihuitun	60334	4.0 %
19	Aerjishan	15098	1.0 %
20	Haizhou	119186	7.9 %
21	Wdls shengding	9494	0.6 %
22	Wdls rending	23233	1.5 %
23	Waziyu	75042	5.0 %
24	Wuhu	44790	3.0 %
25	Mianyanding	76738	5.1 %
26	Suba	4162	0.3 %
27	Kaidang	8835	0.6 %
28	Kaidang toucong baoyang	30137	2.0 %
29	Mianhua yanding xiaomen rending	25900	1.7 %
101	Td zhhu	0	0.0 %
102	Td xihu	0	0.0 %
103	Td zhba	0	0.0 %
104	Td xiba	0	0.0 %
105	Td zhho	0	0.0 %
106	Td xiho	0	0.0 %
107	Td zhla	0	0.0 %
108	Td xila	0	0.0 %
111	Fd zhba	0	0.0 %
112	Fd xiba	0	0.0 %
113	Fd zhho	0	0.0 %
114	Fd xiho	0	0.0 %
115	Fd zhla	0	0.0 %
116	Fd xila	0	0.0 %
	Total	1,513,357	100%

Based upon 1,513,357 valid cases out of 1,513,357 total cases.

• Minimum: 1

• Maximum: 29

Location: 81-83 (width: 3; decimal: 0) *Variable Type:* numeric

Gregorian (CE) calendar year in which the register was compiled. Original registers specified imperial reign year. Value assigned by transcribers based on interpretation of original data.

Value	Label	Unweighted Frequency	%
1749	-	3910	0.3 %
1753	-	1158	0.1 %
1756	-	8048	0.5 %
1759	-	8906	0.6 %
1762	-	6190	0.4 %
1765	-	2854	0.2 %
1768	-	11312	0.7 %
1771	-	2823	0.2 %
1774	-	9475	0.6 %
1777	-	16890	1.1 %
1780	-	11813	0.8 %
1783	-	12900	0.9 %
1786	-	6724	0.4 %
1789	-	19462	1.3 %
1792	-	29328	1.9 %
1795	-	23563	1.6 %
1798	-	24474	1.6 %
1801	-	26428	1.7 %
1804	-	32752	2.2 %
1807	-	18996	1.3 %
1810	-	9773	0.6 %
1813	-	26364	1.7 %
1816	-	21431	1.4 %
1819	-	32240	2.1 %
1822	-	41579	2.7 %
1825	-	38267	2.5 %
1828	-	39455	2.6 %
1831	-	36894	2.4 %
1834	-	40271	2.7 %
1837	-	36606	2.4 %
1840	-	38538	2.5 %
1843	-	31997	2.1 %
1846	-	38636	2.6 %
1849	-	33308	2.2 %
1852	-	36196	2.4 %
1855	-	35894	2.4 %
1858	-	38266	2.5 %
1861	-	41204	2.7 %
1864	-	36520	2.4 %
1867	-	39692	2.6 %

Value	Label	Unweighted Frequency	%
1870	-	36359	2.4 %
1873	-	44279	2.9 %
1876	-	36690	2.4 %
1879	-	34164	2.3 %
1882	-	60389	4.0 %
1885	-	53747	3.6 %
1888	-	33182	2.2 %
1903	-	77138	5.1 %
1906	-	83416	5.5 %
1909	-	82856	5.5 %
	Total	1,513,357	100%

Based upon 1,513,357 valid cases out of 1,513,357 total cases.

- Mean: 1849.17
- Median: 1852.00
- Mode: 1906.00
- Minimum: 1749
- Maximum: 1909
- Standard Deviation: 39.83

Location: 84-87 (width: 4; decimal: 0) *Variable Type:* numeric

REGISTER_SEQ: REGISTER SEQUENCE

Sequential identifier for the records in a register. Concatenated with DATASET and YEAR, it also uniquely identifies records. Value assigned by transcribers based on interpretation of original data.

Based upon 1,513,357 valid cases out of 1,513,357 total cases.

Location: 88-94 (width: 7; decimal: 0) *Variable Type:* character

YIHU_SEQ: ZU SEQUENCE NUMBER, WITHIN REGISTER

Sequential number that identifies the records of a specific household group (yihu) in the register for a given year. Ylhu could contain any number of households, and were an intermediate level between banner and household. Value assigned by transcribers based on interpretation of original data.

Value	Label	Unweighted Frequency	%
0	-	8	0.0 %
1	-	37444	2.5 %
2	-	18180	1.2 %
3	-	15140	1.0 %
4	-	16658	1.1 %
5	-	14475	1.0 %
6	-	14385	1.0 %
7	-	17450	1.2 %

Value	Label	Unweighted Frequency	%
8	-	16195	1.1 %
9	-	17604	1.2 %
10	-	14761	1.0 %
11	-	13698	0.9 %
12	-	13986	0.9 %
13	-	13390	0.9 %
14	-	16944	1.1 %
15	-	13453	0.9 %
16	-	13736	0.9 %
17	-	17198	1.1 %
18	-	14395	1.0 %
19	-	15293	1.0 %
20	-	14520	1.0 %
21	-	13169	0.9 %
22	-	13371	0.9 %
23	-	14399	1.0 %
24	-	12424	0.8 %
25	-	14082	0.9 %
26	-	12717	0.8 %
27	-	13186	0.9 %
28	-	15095	1.0 %
29	-	17213	1.1 %
30	-	15371	1.0 %
31	-	15528	1.0 %
32	-	14536	1.0 %
33	-	12992	0.9 %
34	-	13606	0.9 %
35	-	14285	0.9 %
36	-	14252	0.9 %
37	-	15013	1.0 %
38	-	14313	0.9 %
39	-	15899	1.1 %
40	-	15261	1.0 %
41	-	12571	0.8 %
42	-	13569	0.9 %
43	-	14313	0.9 %
44	-	13722	0.9 %
45	-	11475	0.8 %
46	-	12694	0.8 %
47	-	13436	0.9 %
48	-	13012	0.9 %
49	-	15369	1.0 %

Value	Label	Unweighted Frequency	%
	Total	1,513,357	100%

Please note that only the first 50 response categories are displayed in the PDF codebook. To view all response categories, please analyze the data file in the statistical package of your choice (SAS, SPSS, Stata, R).

Based upon 1,513,357 valid cases out of 1,513,357 total cases.

- Mean: 60.12
- Median: 51.00
- Mode: 1.00
- Minimum: 0
- Maximum: 775
- Standard Deviation: 44.53

Location: 95-97 (width: 3; decimal: 0) *Variable Type:* numeric

HOUSEHOLD_SEQ: HOUSEHOLD SEQUENCE NUMBER, WITHIN ZU

Sequential number that identifies the households within a specific household group (zu) in a register. Missing before 1789. Value assigned by transcribers based on interpretation of original data.

Value	Label	Unweighted Frequency	%
0	-	15114	1.0 %
1	-	476410	31.5 %
2	-	280130	18.5 %
3	-	197291	13.0 %
4	-	133467	8.8 %
5	-	92601	6.1 %
6	-	64001	4.2 %
7	-	41777	2.8 %
8	-	28015	1.9 %
9	-	19531	1.3 %
10	-	14520	1.0 %
11	-	10104	0.7 %
12	-	7426	0.5 %
13	-	5419	0.4 %
14	-	4154	0.3 %
15	-	2888	0.2 %
16	-	2451	0.2 %
17	-	2070	0.1 %
18	-	1514	0.1 %
19	-	1163	0.1 %
20	-	985	0.1 %
21	-	652	0.0 %
22	-	509	0.0 %

Value	Label	Unweighted Frequency	%
23	-	370	0.0 %
24	-	259	0.0 %
25	-	294	0.0 %
26	-	226	0.0 %
27	-	222	0.0 %
28	-	125	0.0 %
29	-	203	0.0 %
30	-	264	0.0 %
31	-	177	0.0 %
32	-	90	0.0 %
33	-	230	0.0 %
34	-	128	0.0 %
35	-	296	0.0 %
36	-	278	0.0 %
37	-	130	0.0 %
38	-	81	0.0 %
39	-	155	0.0 %
40	-	174	0.0 %
41	-	220	0.0 %
42	-	145	0.0 %
43	-	130	0.0 %
44	-	146	0.0 %
45	-	98	0.0 %
46	-	91	0.0 %
47	-	71	0.0 %
48	-	106	0.0 %
49	-	49	0.0 %
	Missing Data		
-98	Not Applicable	103003	6.8 %
	Total	1,513,357	100%

Please note that only the first 50 response categories are displayed in the PDF codebook. To view all response categories, please analyze the data file in the statistical package of your choice (SAS, SPSS, Stata, R).

Based upon 1,410,354 valid cases out of 1,513,357 total cases.

- Mean: 3.40
- Median: 2.00
- Mode: 1.00
- Minimum: 0
- Maximum: 123
- Standard Deviation: 4.87

Location: 98-100 (width: 3; decimal: 0) *Variable Type:* numeric

REGION: REGION

Geographical region within Liaoning Province. Value generated by specific computer program based on DATASET and the village information recorded on the original population registers.

Value	Label	Unweighted Frequency	%
1	North In	624003	41.2 %
2	Central In	406869	26.9 %
3	South central In	299957	19.8 %
4	South In	182528	12.1 %
5	Scp	0	0.0 %
99	Missing	0	0.0 %
	Total	1,513,357	100%

Based upon 1,513,357 valid cases out of 1,513,357 total cases.

- Minimum: 1
- Maximum: 4

Location: 101-102 (width: 2; decimal: 0) *Variable Type:* numeric

DISTRICT: ADMINISTRATIVE DISTRICT

Administrative district. Value assigned by machine according to the coder's interpretation of village information recorded on the original population registers.

Value	Label	Unweighted Frequency	%
1	Fu shun	8597	0.6 %
2	Gai zhou	172985	11.4 %
3	Guang ning	17418	1.2 %
4	Jin zhou	8299	0.5 %
5	Ju liu he	21621	1.4 %
6	Kai yuan	360293	23.8 %
7	Liao yang	83287	5.5 %
8	Niu zhuang	167806	11.1 %
9	Shen yang	366212	24.2 %
10	Tie ling	179085	11.8 %
11	Xing jing	10439	0.7 %
12	Xiong yue	31446	2.1 %
13	Xiu yan	1244	0.1 %
	Missing Data		
-99	Missing	84625	5.6 %
	Total	1,513,357	100%

Based upon 1,428,732 valid cases out of 1,513,357 total cases.

• Minimum: 1

• Maximum: 13

Location: 103-105 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -99

UNIQUE_VILLAGE_ID: UNIQUE VILLAGE

Identifies the records for each specific village for grouping purposes. Based on village name recorded in original register. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
2	-	75	0.0 %
3	-	15	0.0 %
4	-	59	0.0 %
5	-	64	0.0 %
6	-	134	0.0 %
7	-	15	0.0 %
8	-	718	0.0 %
9	-	193	0.0 %
10	-	7	0.0 %
11	-	8	0.0 %
12	-	316	0.0 %
13	-	891	0.1 %
14	-	30	0.0 %
15	-	17	0.0 %
16	-	93	0.0 %
17	-	36	0.0 %
18	-	4	0.0 %
19	-	761	0.1 %
20	-	14	0.0 %
21	-	18	0.0 %
22	-	68	0.0 %
23	-	46	0.0 %
24	-	1744	0.1 %
25	-	557	0.0 %
26	-	7	0.0 %
27	-	51	0.0 %
28	-	106	0.0 %
29	-	1331	0.1 %
30	-	347	0.0 %
31	-	14	0.0 %
32	-	32	0.0 %
33	-	249	0.0 %
34	-	3650	0.2 %

Value	Label	Unweighted Frequency	%
35	-	913	0.1 %
36	-	80	0.0 %
37	-	473	0.0 %
38	-	36	0.0 %
39	-	895	0.1 %
40	-	358	0.0 %
41	-	98	0.0 %
42	-	1359	0.1 %
43	-	93	0.0 %
44	-	2170	0.1 %
45	-	3663	0.2 %
46	-	1994	0.1 %
47	-	36	0.0 %
48	-	809	0.1 %
49	-	30	0.0 %
50	-	5089	0.3 %
51	-	1563	0.1 %
	Missing Data		
-99	Missing	106702	7.1 %
	Total	1,513,357	100%

Please note that only the first 50 response categories are displayed in the PDF codebook. To view all response categories, please analyze the data file in the statistical package of your choice (SAS, SPSS, Stata, R).

Based upon 1,406,655 valid cases out of 1,513,357 total cases.

- Mean: 365.38
- Minimum: 2
- Maximum: 699
- Standard Deviation: 170.33

Location: 106-108 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -99

RELATIONSHIP: RELATIONSHIP TO HEAD

Before 1789, relationship to head of household group (zu zhang). Beginning in 1789, relationship to household head. Each number stands for birth order recorded in the register, and each letter stand for a relationship to head (e=head, w=wife, q=concubine, m=mother, f=father, b=brother, z=sister, s=son, d=daughter; o=older, y=younger). For example, f1yb1s means that the person is the head's father's first younger brother's first son, which refers to the head's cousin. 2yb1sw means the head's second younger brother's first son's wife, which refers to wife of the head's nephew. Value assigned by transcribers based on interpretation of original data.

Based upon 1,513,357 valid cases out of 1,513,357 total cases.

Location: 109-126 (width: 18; decimal: 0) *Variable Type:* character

Data transcribed from original population registers.

Value	Label	Unweighted Frequency	%
1	Female	596130	39.4 %
2	Male	917204	60.6 %
	Missing Data		
-99	Missing	23	0.0 %
	Total	1,513,357	100%

Based upon 1,513,334 valid cases out of 1,513,357 total cases.

- Minimum: 1
- Maximum: 2

Location: 127-129 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -99

GENERATION: GENERATION

Generation relative to household head. Head's generation is 3, generation above is 2 (e.g. head's father), generation below is 4 (e.g., head's son), and so on. Value assigned by transcribers based on interpretation of original data.

Value	Label	Unweighted Frequency	%
0	-	251	0.0 %
1	-	10428	0.7 %
2	-	131948	8.7 %
3	-	660448	43.6 %
4	-	545640	36.1 %
5	-	153503	10.1 %
6	-	10784	0.7 %
7	-	343	0.0 %
8	-	11	0.0 %
9	-	1	0.0 %
	Total	1,513,357	100%

Based upon 1,513,357 valid cases out of 1,513,357 total cases.

- Mean: 3.48
- Median: 3.00
- Mode: 3.00
- Minimum: 0
- Maximum: 9
- Standard Deviation: 0.85

Location: 130-130 (width: 1; decimal: 0) *Variable Type:* numeric

Marital status inferred from presence of spouse and children in current and previous registers. Value assigned by transcribers based on interpretation of original data.

Value	Label	Unweighted Frequency	%
1	Married	775260	51.2 %
2	Unmarried	447824	29.6 %
3	Widowed	164534	10.9 %
4	Remarried	42804	2.8 %
	Missing Data		
-99	Missing	82935	5.5 %
	Total	1,513,357	100%

Based upon 1,430,422 valid cases out of 1,513,357 total cases.

- Minimum: 1
- Maximum: 4

Location: 131-133 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -99

AGE_IN_SUI: AGE

Age in sui as recorded in the registers. Individuals were one sui at birth, and aged by one sui on every Lunar New Year. As a result, Western age is on average 1.5 years younger than Chinese age in sui. Data transcribed from original population registers.

Value	Label	Unweighted Frequency	%
1	-	1769	0.1 %
2	-	11461	0.8 %
3	-	18663	1.2 %
4	-	18127	1.2 %
5	-	21650	1.4 %
6	-	22819	1.5 %
7	-	19072	1.3 %
8	-	21549	1.4 %
9	-	21943	1.4 %
10	-	18529	1.2 %
11	-	20425	1.3 %
12	-	21308	1.4 %
13	-	17837	1.2 %
14	-	19859	1.3 %
15	-	21035	1.4 %
16	-	17681	1.2 %
17	-	22371	1.5 %
18	-	24700	1.6 %
19	-	22070	1.5 %
20	-	28193	1.9 %

Value	Label	Unweighted Frequency	%
21	-	28290	1.9 %
22	-	25549	1.7 %
23	-	28439	1.9 %
24	-	28592	1.9 %
25	-	26056	1.7 %
26	-	27801	1.8 %
27	-	27784	1.8 %
28	-	25414	1.7 %
29	-	26833	1.8 %
30	-	27612	1.8 %
31	-	24301	1.6 %
32	-	25852	1.7 %
33	-	25965	1.7 %
34	-	23461	1.6 %
35	-	24636	1.6 %
36	-	25148	1.7 %
37	-	22495	1.5 %
38	-	23119	1.5 %
39	-	23764	1.6 %
40	-	21764	1.4 %
41	-	22325	1.5 %
42	-	22670	1.5 %
43	-	20883	1.4 %
44	-	21063	1.4 %
45	-	21886	1.4 %
46	-	19304	1.3 %
47	-	20022	1.3 %
48	-	20243	1.3 %
49	-	18262	1.2 %
50	-	18670	1.2 %
	Missing Data		
-99	Missing	55842	3.7 %
	Total	1,513,357	100%

Please note that only the first 50 response categories are displayed in the PDF codebook. To view all response categories, please analyze the data file in the statistical package of your choice (SAS, SPSS, Stata, R).

Based upon 1,457,515 valid cases out of 1,513,357 total cases.

- Mean: 35.39
- Median: 33.00
- Mode: 24.00
- Minimum: 1
- Maximum: 742

Standard Deviation: 20.70

Location: 134-136 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -99

BIRTHYEAR: BIRTH YEAR

Gregorian (CE) year of birth, calculated as (YEAR - AGE_IN_SUI + 1). Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
1147	-	1	0.0 %
1188	-	1	0.0 %
1664	-	3	0.0 %
1665	-	2	0.0 %
1666	-	2	0.0 %
1667	-	1	0.0 %
1668	-	5	0.0 %
1669	-	5	0.0 %
1670	-	13	0.0 %
1671	-	9	0.0 %
1672	-	10	0.0 %
1673	-	14	0.0 %
1674	-	12	0.0 %
1675	-	24	0.0 %
1676	-	44	0.0 %
1677	-	25	0.0 %
1678	-	59	0.0 %
1679	-	49	0.0 %
1680	-	68	0.0 %
1681	-	89	0.0 %
1682	-	73	0.0 %
1683	-	111	0.0 %
1684	-	123	0.0 %
1685	-	103	0.0 %
1686	-	136	0.0 %
1687	-	176	0.0 %
1688	-	168	0.0 %
1689	-	200	0.0 %
1690	-	212	0.0 %
1691	-	289	0.0 %
1692	-	281	0.0 %
1693	-	289	0.0 %
1694	-	342	0.0 %
1695	-	333	0.0 %

Value	Label	Unweighted Frequency	%
1696	-	346	0.0 %
1697	-	445	0.0 %
1698	-	384	0.0 %
1699	-	330	0.0 %
1700	-	522	0.0 %
1701	-	583	0.0 %
1702	-	568	0.0 %
1703	-	610	0.0 %
1704	-	584	0.0 %
1705	-	758	0.1 %
1706	-	916	0.1 %
1707	-	856	0.1 %
1708	-	906	0.1 %
1709	-	921	0.1 %
1710	-	949	0.1 %
1711	-	1000	0.1 %
	Missing Data		
-98	Not applicable	55842	3.7 %
	Total	1,513,357	100%

Please note that only the first 50 response categories are displayed in the PDF codebook. To view all response categories, please analyze the data file in the statistical package of your choice (SAS, SPSS, Stata, R).

Based upon 1,457,515 valid cases out of 1,513,357 total cases.

- Mean: 1814.46
- Median: 1816.00
- Mode: 1821.00
- Minimum: 1147
- Maximum: 1909
- Standard Deviation: 45.11

Location: 137-140 (width: 4; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

BROTHER_COUNT: BROTHERS IN HOUSEHOLD

Number of male siblings living in household. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	-	565820	37.4 %
1	-	404095	26.7 %
2	-	220201	14.6 %
3	-	99156	6.6 %
4	-	40338	2.7 %

Value	Label	Unweighted Frequency	%
5	-	16482	1.1 %
6	-	6395	0.4 %
7	-	2912	0.2 %
8	-	1715	0.1 %
9	-	1281	0.1 %
10	-	557	0.0 %
11	-	253	0.0 %
12	-	26	0.0 %
	Missing Data		
-98	Not applicable	154126	10.2 %
	Total	1,513,357	100%

Based upon 1,359,231 valid cases out of 1,513,357 total cases.

- Mean: 1.09
- Median: 1.00
- Mode: 0.00
- Minimum: 0
- Maximum: 12
- Standard Deviation: 1.30

Location: 141-143 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

SISTER_COUNT: SISTERS IN HOUSEHOLD

Number of female siblings living in household. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	-	1284862	84.9 %
1	-	55167	3.6 %
2	-	14256	0.9 %
3	-	3788	0.3 %
4	-	883	0.1 %
5	-	245	0.0 %
6	-	25	0.0 %
7	-	5	0.0 %
	Missing Data		
-98	Not applicable	154126	10.2 %
	Total	1,513,357	100%

Based upon 1,359,231 valid cases out of 1,513,357 total cases.

- Mean: 0.07
- Median: 0.00
- Mode: 0.00

- Minimum: 0
- Maximum: 7
- Standard Deviation: 0.34

Location: 144-146 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

MALE_COUSIN_COUNT: MALE COUSINS IN HOUSEHOLD

Number of male paternal cousins living in household. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	-	931313	61.5 %
1	-	151305	10.0 %
2	-	107939	7.1 %
3	-	70758	4.7 %
4	-	46295	3.1 %
5	-	29447	1.9 %
6	-	19556	1.3 %
7	-	12174	0.8 %
8	-	8310	0.5 %
9	-	5772	0.4 %
10	-	4289	0.3 %
11	-	3166	0.2 %
12	-	2276	0.2 %
13	-	1824	0.1 %
14	-	1144	0.1 %
15	-	748	0.0 %
16	-	604	0.0 %
17	-	422	0.0 %
18	-	351	0.0 %
19	-	358	0.0 %
20	-	220	0.0 %
21	-	166	0.0 %
22	-	157	0.0 %
23	-	57	0.0 %
24	-	39	0.0 %
25	-	12	0.0 %
26	-	16	0.0 %
27	-	2	0.0 %
	Missing Data		
-98	Not applicable	114637	7.6 %
	Total	1,513,357	100%

Based upon 1,398,720 valid cases out of 1,513,357 total cases.

- Mean: 1.02
- Median: 0.00
- Mode: 0.00
- Minimum: 0
- Maximum: 27
- Standard Deviation: 2.11

Location: 147-149 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

FEMALE_COUSIN_COUNT: FEMALE COUSINS IN HOUSEHOLD

Number of unmarried female paternal cousins living in household. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	-	1341338	88.6 %
1	-	35118	2.3 %
2	-	12571	0.8 %
3	-	5449	0.4 %
4	-	2357	0.2 %
5	-	951	0.1 %
6	-	541	0.0 %
7	-	245	0.0 %
8	-	110	0.0 %
9	-	30	0.0 %
10	-	10	0.0 %
	Missing Data		
-98	Not applicable	114637	7.6 %
	Total	1,513,357	100%

Based upon 1,398,720 valid cases out of 1,513,357 total cases.

- Mean: 0.07
- Median: 0.00
- Mode: 0.00
- Minimum: 0
- Maximum: 10
- Standard Deviation: 0.41

Location: 150-152 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

UNCLE_COUNT: PATERNAL UNCLES IN HOUSEHOLD

Number of father's brothers living in household. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	-	960836	63.5 %

Value	Label	Unweighted Frequency	%
1	-	240475	15.9 %
2	-	115150	7.6 %
3	-	47863	3.2 %
4	-	19700	1.3 %
5	-	8205	0.5 %
6	-	3223	0.2 %
7	-	1523	0.1 %
8	-	817	0.1 %
9	-	654	0.0 %
10	-	195	0.0 %
11	-	74	0.0 %
12	-	5	0.0 %
	Missing Data		
-98	Not applicable	114637	7.6 %
	Total	1,513,357	100%

Based upon 1,398,720 valid cases out of 1,513,357 total cases.

- Mean: 0.56
- Median: 0.00
- Mode: 0.00
- Minimum: 0
- Maximum: 12
- Standard Deviation: 1.05

Location: 153-155 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

AUNT_COUNT: PATERNAL AUNTS IN HOUSEHOLD

Number of father's sisters living in household. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	-	1388131	91.7 %
1	-	8154	0.5 %
2	-	1798	0.1 %
3	-	510	0.0 %
4	-	98	0.0 %
5	-	25	0.0 %
7	-	4	0.0 %
	Missing Data		
-98	Not applicable	114637	7.6 %
	Total	1,513,357	100%

Based upon 1,398,720 valid cases out of 1,513,357 total cases.

- Mean: 0.01
- Median: 0.00
- Mode: 0.00
- Minimum: 0
- Maximum: 7
- Standard Deviation: 0.13

Location: 156-158 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

FATHER_ALIVE: FATHER IS ALIVE

Father was alive in the household in the current register. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	759111	50.2 %
1	Yes	600120	39.7 %
	Missing Data		
-98	Not applicable	154126	10.2 %
	Total	1,513,357	100%

Based upon 1,359,231 valid cases out of 1,513,357 total cases.

- Minimum: 0
- Maximum: 1

Location: 159-161 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

MOTHER_ALIVE: MOTHER IS ALIVE

Mother was alive in the current register. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	680513	45.0 %
1	Yes	718207	47.5 %
	Missing Data		
-98	Not applicable	114637	7.6 %
	Total	1,513,357	100%

Based upon 1,398,720 valid cases out of 1,513,357 total cases.

- Minimum: 0
- Maximum: 1

Location: 162-164 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

SON_COUNT: NUMBER OF SONS

Number of sons born to this individual up to this year, inferred from record linkage. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	-	736722	48.7 %
1	-	351327	23.2 %
2	-	181456	12.0 %
3	-	80427	5.3 %
4	-	32200	2.1 %
5	-	10941	0.7 %
6	-	3774	0.2 %
7	-	1165	0.1 %
8	-	397	0.0 %
9	-	182	0.0 %
10	-	60	0.0 %
11	-	42	0.0 %
12	-	9	0.0 %
13	-	18	0.0 %
	Missing Data		
-98	Not applicable	114637	7.6 %
	Total	1,513,357	100%

Based upon 1,398,720 valid cases out of 1,513,357 total cases.

- Mean: 0.84
- Median: 0.00
- Mode: 0.00
- Minimum: 0
- Maximum: 13
- Standard Deviation: 1.15

Location: 165-167 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

DAUGHTER_COUNT: NUMBER OF DAUGHTERS

Number of daughters born to this individual up to this year, inferred from record linkage. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	-	1247385	82.4 %
1	-	107610	7.1 %
2	-	29358	1.9 %
3	-	9540	0.6 %
4	-	3477	0.2 %
5	-	955	0.1 %
6	-	225	0.0 %
7	-	116	0.0 %

Value	Label	Unweighted Frequency	%
8	-	29	0.0 %
9	-	25	0.0 %
	Missing Data		
-98	Not applicable	114637	7.6 %
	Total	1,513,357	100%

Based upon 1,398,720 valid cases out of 1,513,357 total cases.

- Mean: 0.15
- Median: 0.00
- Mode: 0.00
- Minimum: 0
- Maximum: 9
- Standard Deviation: 0.52

Location: 168-170 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

PRESENT: ALIVE AND PRESENT IN THIS REGISTER (VITAL1 == 0 | VITAL1 >= 6)

Alive and present in this register. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	114637	7.6 %
1	Yes	1398720	92.4 %
	Total	1,513,357	100%

Based upon 1,513,357 valid cases out of 1,513,357 total cases.

• Minimum: 0

• Maximum: 1

Location: 171-171 (width: 1; decimal: 0) Variable Type: numeric

DIED: DIED

Annotated in current register as having died in the three years covered by the register. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	1430491	94.5 %
1	Yes	82866	5.5 %
	Total	1,513,357	100%

Based upon 1,513,357 valid cases out of 1,513,357 total cases.

• Minimum: 0

• Maximum: 1

MARRIED_OUT: MARRIED OUT

Annotated in current register as having married out in the three years covered by the register. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	1505038	99.5 %
1	Yes	8319	0.5 %
	Total	1,513,357	100%

Based upon 1,513,357 valid cases out of 1,513,357 total cases.

- Minimum: 0
- Maximum: 1

Location: 173-173 (width: 1; decimal: 0) *Variable Type:* numeric

REMARRIED_OUT: REMARRIED OUT

Annotated in current register as having remarried out in the three years covered by the register. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	1511958	99.9 %
1	Yes	1399	0.1 %
	Total	1,513,357	100%

Based upon 1,513,357 valid cases out of 1,513,357 total cases.

• Minimum: 0

• Maximum: 1

Location: 174-174 (width: 1; decimal: 0) *Variable Type:* numeric

NEXT_DIE: DIE BY NEXT REGISTER (VITAL1 == 1 IN NEXT REGISTER)

Annotated in next available register as having died in the three years covered by that register. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	1237243	81.8 %
1	Yes	74477	4.9 %
	Missing Data		
-98	Not applicable	201637	13.3 %
	Total	1,513,357	100%

Based upon 1,311,720 valid cases out of 1,513,357 total cases.

- Minimum: 0
- Maximum: 1

Location: 175-177 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

NEXT_MARRY: MARRY OUT BY NEXT REGISTER (VITAL1 == 2 IN NEXT REGISTER)

Annotated in next available register as having married in the three years covered by that register. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	1240586	82.0 %
1	Yes	71134	4.7 %
	Missing Data		
-98	Not applicable	201637	13.3 %
	Total	1,513,357	100%

Based upon 1,311,720 valid cases out of 1,513,357 total cases.

- Minimum: 0
- Maximum: 1

Location: 178-180 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

NEXT_REMARRY: REMARRY OUT BY NEXT REGISTER (VITAL1 == 3 IN NEXT REGISTER)

Annotated in next available register as having married out in the three years covered by that register. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	1307941	86.4 %
1	Yes	3779	0.2 %
	Missing Data		
-98	Not applicable	201637	13.3 %
	Total	1,513,357	100%

Based upon 1,311,720 valid cases out of 1,513,357 total cases.

- Minimum: 0
- Maximum: 1

Location: 181-183 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

NEXT_ABSCONDED: WILL BECOME TAO BY NEXT REGISTER (VITAL1 == 4)

Annotated in next available register as having absconded in the three years covered by that register. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	1309516	86.5 %
1	Yes	2204	0.1 %
	Missing Data		
-98	Not applicable	201637	13.3 %
	Total	1,513,357	100%

Based upon 1,311,720 valid cases out of 1,513,357 total cases.

- Minimum: 0
- Maximum: 1

Location: 184-186 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

NEXT_BOYS: BOYS BORN BY NEXT REGISTER

Number of boys born between current and next available register. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	-	1165435	77.0 %
1	-	124930	8.3 %
2	-	17200	1.1 %
3	-	3124	0.2 %
4	-	795	0.1 %
5	-	151	0.0 %
6	-	44	0.0 %
7	-	23	0.0 %
8	-	14	0.0 %
9	-	4	0.0 %
	Missing Data		
-98	Not applicable	201637	13.3 %
	Total	1,513,357	100%

Based upon 1,311,720 valid cases out of 1,513,357 total cases.

- Mean: 0.13
- Median: 0.00
- Mode: 0.00
- Minimum: 0
- Maximum: 9
- Standard Deviation: 0.41

Location: 187-189 (width: 3; decimal: 0) *Variable Type:* numeric

NEXT_GIRLS: GIRLS BORN BY NEXT REGISTER

Number of girls born between current and next available register. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	-	1291000	85.3 %
1	-	18820	1.2 %
2	-	1696	0.1 %
3	-	161	0.0 %
4	-	39	0.0 %
5	-	1	0.0 %
6	-	3	0.0 %
	Missing Data		
-98	Not applicable	201637	13.3 %
	Total	1,513,357	100%

Based upon 1,311,720 valid cases out of 1,513,357 total cases.

- Mean: 0.02
- Median: 0.00
- Mode: 0.00
- Minimum: 0
- Maximum: 6
- Standard Deviation: 0.14

Location: 190-192 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

NEXT_3: NEXT OBSERVATION IS 3 YEARS AWAY

Next triennial observation of this person is in the dataset. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	539039	35.6 %
1	Yes	859681	56.8 %
	Missing Data		
-98	Not applicable	114637	7.6 %
	Total	1,513,357	100%

Based upon 1,398,720 valid cases out of 1,513,357 total cases.

• Minimum: 0

• Maximum: 1

Location: 193-195 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

NEXT_6: NEXT OBSERVATION IS 6 YEARS AWAY

Next available observation is 6 years away. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	1193309	78.9 %
1	Yes	205411	13.6 %
	Missing Data		
-98	Not applicable	114637	7.6 %
	Total	1,513,357	100%

Based upon 1,398,720 valid cases out of 1,513,357 total cases.

• Minimum: 0

• Maximum: 1

Location: 196-198 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

AT_RISK_DIE: AT RISK OF DYING

Identifies observations to include in mortality analysis. Individual is present and an observation is available in next triennial register: Set to 1 if PRESENT and NEXT_3 are both 1. Otherwise set to 0. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	539039	35.6 %
1	Yes	859681	56.8 %
	Missing Data		
-98	Not applicable	114637	7.6 %
	Total	1,513,357	100%

Based upon 1,398,720 valid cases out of 1,513,357 total cases.

- Minimum: 0
- Maximum: 1

Location: 199-201 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

AT_RISK_MARRY: AT RISK OF MARRYING

Identifies observations to include in analysis of first marriage. Individual is present and unmarried and an observation is available in next triennial register: Set to 1 if PRESENT and NEXT_3 are both 1 and MARITAL is 2. Otherwise set to 0. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	1133073	74.9 %
1	Yes	265647	17.6 %

Value	Label	Unweighted Frequency	%
	Missing Data		
-98	Not applicable	114637	7.6 %
	Total	1,513,357	100%

Based upon 1,398,720 valid cases out of 1,513,357 total cases.

• Minimum: 0

• Maximum: 1

Location: 202-204 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

AT_RISK_REMARRY: AT RISK OF REMARRYING

Identifies observations to include in remarriage analysis. Individual is present and widowed and an observation is available in next triennial register: Set to 1 if PRESENT and NEXT_3 are both 1 and MARITAL is 3. Otherwise set to 0. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	1301127	86.0 %
1	Yes	97593	6.4 %
	Missing Data		
-98	Not applicable	114637	7.6 %
	Total	1,513,357	100%

Based upon 1,398,720 valid cases out of 1,513,357 total cases.

- Minimum: 0
- Maximum: 1

Location: 205-207 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

ABSCONDED: ABSCONDED

Annotated in register as absconded (tao). Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	890906	58.9 %
1	Yes	26298	1.7 %
	Missing Data		
-98	Not applicable	596153	39.4 %
	Total	1,513,357	100%

Based upon 917,204 valid cases out of 1,513,357 total cases.

• Minimum: 0

• Maximum: 1

Location: 208-210 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

NO_STATUS: NO STATUS

No administrative status recorded in register for this male. Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	574396	38.0 %
1	Yes	342808	22.7 %
	Missing Data		
-98	Not applicable	596153	39.4 %
	Total	1,513,357	100%

Based upon 917,204 valid cases out of 1,513,357 total cases.

- Minimum: 0
- Maximum: 1

Location: 211-213 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

RETIRED: RETIRED

Administrative status recorded in the register indicates that individual is retired (tui). Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	880207	58.2 %
1	Yes	36997	2.4 %
	Missing Data		
-98	Not applicable	596153	39.4 %
	Total	1,513,357	100%

Based upon 917,204 valid cases out of 1,513,357 total cases.

• Minimum: 0

• Maximum: 1

Location: 214-216 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

OLD: OLD

Administrative status recorded in the register indicates that individual is old (lao). Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	899138	59.4 %
1	Yes	18066	1.2 %
	Missing Data		
-98	Not applicable	596153	39.4 %
	Total	1,513,357	100%

Based upon 917,204 valid cases out of 1,513,357 total cases.

- Minimum: 0
- Maximum: 1

Location: 217-219 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

POSITION: SALARIED POSITION

Value	Label	Unweighted Frequency	%
0	No	902354	59.6 %
1	Yes	14850	1.0 %
	Missing Data		
-98	Not applicable	596153	39.4 %
	Total	1,513,357	100%

Based upon 917,204 valid cases out of 1,513,357 total cases.

- Minimum: 0
- Maximum: 1

Location: 220-222 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

ZU_ZHANG: HEAD OF HOUSEHOLD GROUP (ZU)

Administrative status recorded in the register includes annotation that individual was head of a household group (zu). Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	859213	56.8 %
1	Yes	57991	3.8 %
	Missing Data		
-98	Not applicable	596153	39.4 %
	Total	1,513,357	100%

Based upon 917,204 valid cases out of 1,513,357 total cases.

• Minimum: 0

• Maximum: 1

Location: 223-225 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

ARTISAN: ARTISAN

Administrative status recorded in the register indicates that individual is an artisan (jiang). Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	914551	60.4 %
1	Yes	2653	0.2 %
	Missing Data		
-98	Not applicable	596153	39.4 %
	Total	1,513,357	100%

Based upon 917,204 valid cases out of 1,513,357 total cases.

- Minimum: 0
- Maximum: 1

Location: 226-228 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

SOLDIER: SOLDIER

Administrative status recorded in the register indicates that individual is a soldier (jia or bing). Value generated by specific computer program.

Value	Label	Unweighted Frequency	%
0	No	908191	60.0 %
1	Yes	9013	0.6 %
	Missing Data		
-98	Not applicable	596153	39.4 %
	Total	1,513,357	100%

Based upon 917,204 valid cases out of 1,513,357 total cases.

- Minimum: 0
- Maximum: 1

Location: 229-231 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -98

BAZI_YEAR: BIRTHYR

Numeric code for the animal associated with year of birth in the Chinese zodiac. There are a total of twelve values, following the order of the animals in the traditional calendar.

Value	Label	Unweighted Frequency	%
1	rat	57919	3.8 %
2	ox	298908	19.8 %
3	tiger	72354	4.8 %
4	rabbit	111592	7.4 %
5	dragon	59859	4.0 %
6	snake	72802	4.8 %
7	horse	107257	7.1 %
8	goat	238488	15.8 %
9	monkey	61783	4.1 %
10	rooster	62170	4.1 %
11	dog	118425	7.8 %
12	pig	66836	4.4 %
	Missing Data		
-99	Missing	184964	12.2 %
	Total	1,513,357	100%

Based upon 1,328,393 valid cases out of 1,513,357 total cases.

- Minimum: 1
- Maximum: 12

Location: 232-234 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -99

BAZI_MONTH: MONTH

The lunar month of birth recorded in the original register.

Value	Label	Unweighted Frequency	%
1	-	39056	2.6 %
2	-	231099	15.3 %
3	-	83722	5.5 %
4	-	65978	4.4 %
5	-	76403	5.0 %
6	-	117905	7.8 %
7	-	127386	8.4 %
8	-	179397	11.9 %
9	-	106941	7.1 %
10	-	132112	8.7 %
11	-	22102	1.5 %
12	-	16646	1.1 %
	Missing Data		
-99	Missing	314610	20.8 %

Value	Label	Unweighted Frequency	%
	Total	1,513,357	100%

Based upon 1,198,747 valid cases out of 1,513,357 total cases.

- Mean: 5.97
- Median: 6.00
- Mode: 2.00
- Minimum: 1
- Maximum: 12
- Standard Deviation: 3.01

Location: 235-237 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -99

BAZI_DAY: DAY

The lunar day of birth recorded in the original register.

Value	Label	Unweighted Frequency	%
1	-	55965	3.7 %
2	-	18746	1.2 %
3	-	8890	0.6 %
4	-	8457	0.6 %
5	-	11120	0.7 %
6	-	12751	0.8 %
7	-	13940	0.9 %
8	-	18536	1.2 %
9	-	12842	0.8 %
10	-	11095	0.7 %
11	-	276827	18.3 %
12	-	73960	4.9 %
13	-	35940	2.4 %
14	-	38854	2.6 %
15	-	34613	2.3 %
16	-	55406	3.7 %
17	-	101809	6.7 %
18	-	165601	10.9 %
19	-	83169	5.5 %
20	-	89244	5.9 %
21	-	15305	1.0 %
22	-	5333	0.4 %
23	-	4667	0.3 %
24	-	6126	0.4 %
25	-	6441	0.4 %

Value	Label	Unweighted Frequency	%
26	-	6000	0.4 %
27	-	6204	0.4 %
28	-	9418	0.6 %
29	-	6160	0.4 %
30	-	5338	0.4 %
	Missing Data		
-99	Missing	314600	20.8 %
	Total	1,513,357	100%

Based upon 1,198,757 valid cases out of 1,513,357 total cases.

- Mean: 14.15
- Median: 15.00
- Mode: 11.00
- Minimum: 1
- Maximum: 30
- Standard Deviation: 5.70

Location: 238-240 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -99

BAZI_HOUR: HOUR

The hour of birth recorded in the original register. The numeric codes correspond to "branches," of which there were 12 in a 24 hour cycle.

Value	Label	Unweighted Frequency	%
1	Zi	242371	16.0 %
2	Niu	140237	9.3 %
3	Yin	21781	1.4 %
4	Мао	85357	5.6 %
5	Chen	45294	3.0 %
6	Si	198407	13.1 %
7	Wu	167243	11.1 %
8	Wei	98092	6.5 %
9	Shen	69430	4.6 %
10	You	23573	1.6 %
11	Xu	39200	2.6 %
12	Hai	40461	2.7 %
	Missing Data		
-99	Missing	341911	22.6 %
	Total	1,513,357	100%

Based upon 1,171,446 valid cases out of 1,513,357 total cases.

• Minimum: 1

• Maximum: 12

Location: 241-243 (width: 3; decimal: 0) Variable Type: numeric (Range of) Missing Values: -99