Berkeley Unified Numident Mortality Database Codebook

ssn

Label: Social Security Number

Description: ssn reports a person's Social Security number, as recorded in the Numident death records. Uniquely identifies all records in the dataset.

fname

Label: First Name

Description: fname reports the first 16 letters of the person's first name, as recorded in the Numident death records.

mname

Label: Middle Name

Description: mname reports the first 16 letters of the person's middle name, as recorded in the Numident death records.

lname

Label: Last Name

Description: lname reports the first 21 letters of the person's last name, as recorded in the Numident death records.

sex

Label: Sex

Description: sex reports a person's sex, as recorded in the Numident death, application, or claims Records.

sex	label	n	freq %
1	Men	25769568	52.20
2	Women	22686047	46.00
NA	Missing	882212	1.79

race_first

Label: race first

Description: race_first reports a person's race, as recorded on their most recent application entry. This variable is only available for persons who submitted 2 or more applications.

Note: Before 1980, the race schema in the Social Security application form contained three categories: White, Black, and Other. In 1980, the SSA added three categories: (1) Asian, Asian American, or Pacific Islander, (2) Hispanic, and (3) North American Indian or Alaskan Native. They also removed the Other category.

race_first	label	n	freq %
1	White	28850917	58.476
2	Black	4066081	8.241
3	Other	689360	1.397
4	Asian, Asian American, or Pacific Islander	207656	0.421
5	Hispanic	286517	0.581
6	North American Indian or Alaskan Native	17485	0.035
NA	Missing	15219811	30.848

race_first_cyear

Label: First Race: Application Year

Description: race_first_cyear is a numeric variable reporting the year of the application on which a person reported their first race. This variable is only available for persons who submitted 2 or more applications.

race_first_cmonth

Label: First Race: Application Month

Description: race_first_cmonth is a numeric variable reporting the month of the application on which a person reported their first race.

race_last

Label: race last

Description: race_last reports a person's race, as recorded on their most recent application entry.

Note: Before 1980, the race schema in the Social Security application form contained three categories: White, Black, and Other. In 1980, the SSA added three categories: (1) Asian, Asian American, or Pacific Islander, (2) Hispanic, and (3) North American Indian or Alaskan Native. They also removed the Other category.

race_last	label	n	freq %
1	White	28417859	57.599
2	Black	4040665	8.190
3	Other	469636	0.952
4	Asian, Asian American, or Pacific Islander	305470	0.619
5	Hispanic	794762	1.611
6	North American Indian or Alaskan Native	89624	0.182
NA	Missing	15219811	30.848

race_last_cyear

Label: First Race: Application Year

Description: race_last_cyear reports the year of the application on which a person reported their last race.

race__cmonth

Label: Last Race: Application Month

Description: race_last_cmonth is a numeric variable reporting the month of the application on which a person reported their last race. This variable is only available for persons who submitted 2 or more applications.

race_change

Label: Change in Race Response Flag

Description: race_change is a numeric dichotomous variable reporting whether an individual changed their race response across their Social Security applications.

race_change	label	n	freq %
0	No Change	33263321	67.40
1	Change	855047	1.73
NA	No Race Information	15219459	30.80

bpl

Label: Birthplace

Description: bpl is a numeric variable reporting the person's place of birth, as recorded in the Numident application or claims records. The numeric codes match the detailed IPUMS-USA Birthplace codes.

For a complete list of IPUMS Birthplace codes, please see: https://usa.ipums.org/usa-action/variables/BPL

byear

Label: Birth Year

Description: byear is a numeric variable reporting the person's year of birth, as recorded in the Numident death records.

Year of Birth
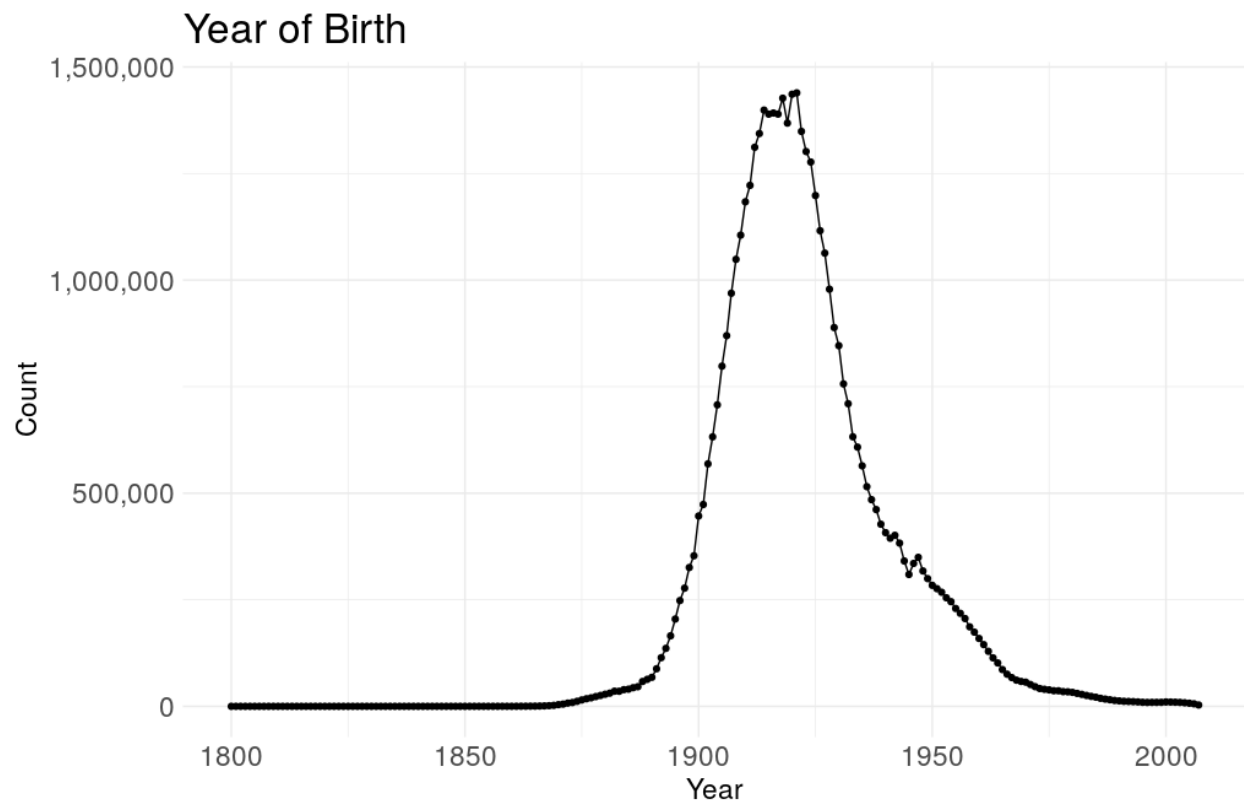
bmonth

Label: Birth Month

Description: bmonth is a numeric variable reporting the person's month of birth, as recorded in the Numident death records.

bmonth	label	n	freq %
1	January	4181375	8.47
2	February	3933555	7.97
3	March	4296940	8.71
4	April	3976307	8.06
5	May	4051432	8.21
6	June	3937147	7.98
7	July	4221224	8.56
8	August	4362872	8.84
9	September	4279353	8.67
10	October	4173868	8.46
11	November	3884566	7.87
12	December	4039185	8.19
NA	Missing	3	0.00

bday

Label: Birth Day

Description: bday is a numeric variable reporting the person's day of birth, as recorded in the Numident death records.

Note: It appears from the surplus of counts on the 1st and the 15th of the month that SSA might have done some imputation of missing values to these dates.

bday	n	freq %
1	1752583	3.600
2	1658244	3.400
3	1609261	3.300
4	1637955	3.300
5	1622040	3.300
6	1626699	3.300
7	1605110	3.300
8	1631598	3.300
9	1590390	3.200
10	1710297	3.500
11	1586088	3.200
12	1670822	3.400
13	1568796	3.200
14	1621064	3.300
15	1782188	3.600
16	1613051	3.300
17	1604846	3.300
18	1616778	3.300
19	1576661	3.200
20	1643349	3.300
21	1546797	3.100
22	1631722	3.300
23	1592459	3.200
24	1591364	3.200
25	1657548	3.400
26	1568447	3.200
27	1577419	3.200
28	1624799	3.300
29	1471946	3.000
30	1414317	2.900
31	901308	1.800
NA	31881	0.065

dyear

Label: Death Year

Description: dyear is a numeric variable reporting the person's year of death, as recorded in the Numident death records.

Year of Death 65+

BUNMD ·· BUNMD Complete Cases — Human Mortality Database
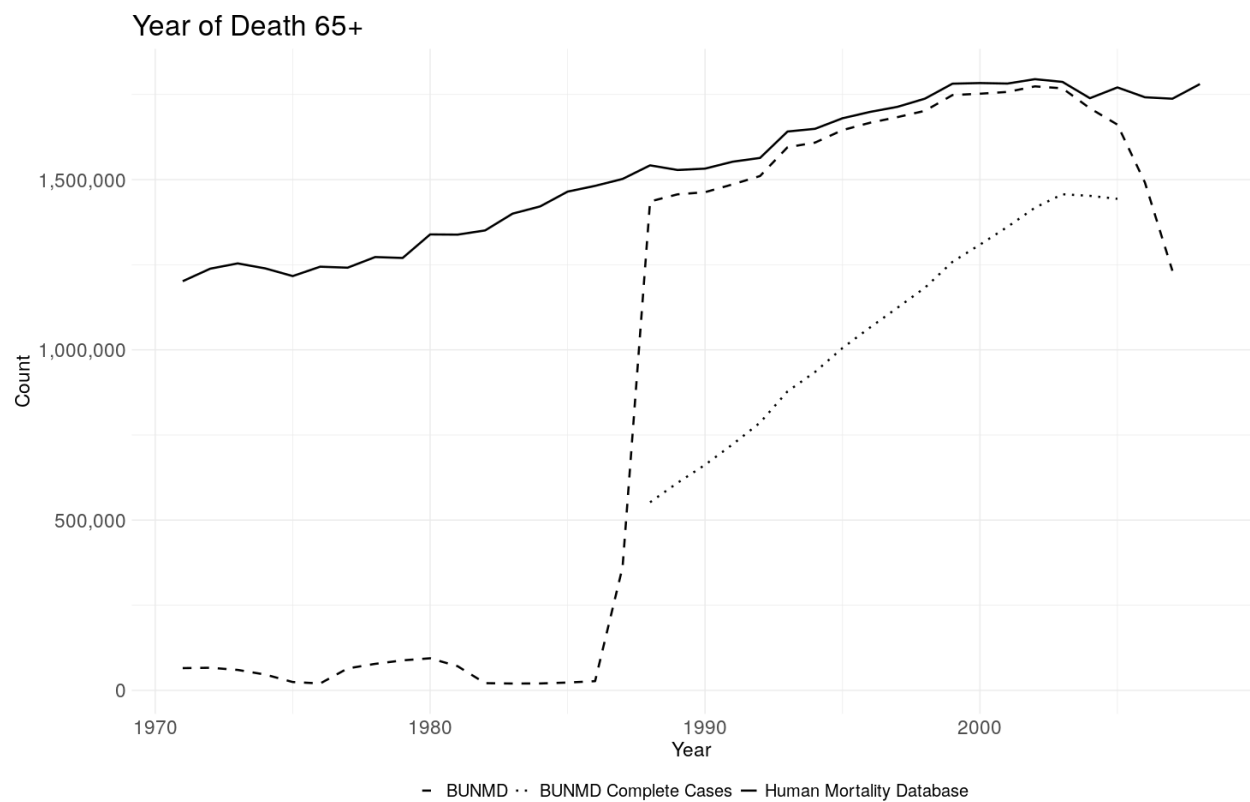
dmonth

Label: Death Month

Description: dmonth is a numeric variable reporting the person's month of death, as recorded in the Numident death records.

dmonth	label	n	freq %
1	January	4610022	9.34
2	February	4119282	8.35
3	March	4353899	8.82
4	April	4036794	8.18
5	May	4023999	8.16
6	June	3828836	7.76
7	July	3924851	7.96
8	August	3887262	7.88
9	September	3835198	7.77
10	October	4125914	8.36
11	November	4102011	8.31
12	December	4489759	9.10

dday

Label: Day of Death

Description: dday is a numeric variable reporting the person's day of death, as recorded in the Numident death records.

Note: It appears from the surplus of counts on the 1st and the 15th of the month that SSA might have done some imputation of missing values to these dates.

dday	n	freq %
1	1432114	2.9
2	1240840	2.5
3	1258505	2.6
4	1281969	2.6
5	1241857	2.5
6	1238121	2.5
7	1234527	2.5
8	1237353	2.5
9	1234152	2.5
10	1236852	2.5
11	1234080	2.5
12	1234552	2.5
13	1230163	2.5
14	1231013	2.5
15	3392922	6.9
16	1227374	2.5
17	1228022	2.5
18	1223495	2.5
19	1223837	2.5
20	1230608	2.5
21	1218728	2.5
22	1217328	2.5
23	1219353	2.5
24	1216433	2.5
25	1214084	2.5
26	1216347	2.5
27	1215884	2.5
28	1217904	2.5
29	1130910	2.3
30	1104857	2.2
31	713927	1.4
NA	9559716	19.0

death_age

Label: Age at Death (Years)

Description: death_age is a numeric variable reporting the person's age at death in years, calculated using birth and death information recorded in the Numident death records.
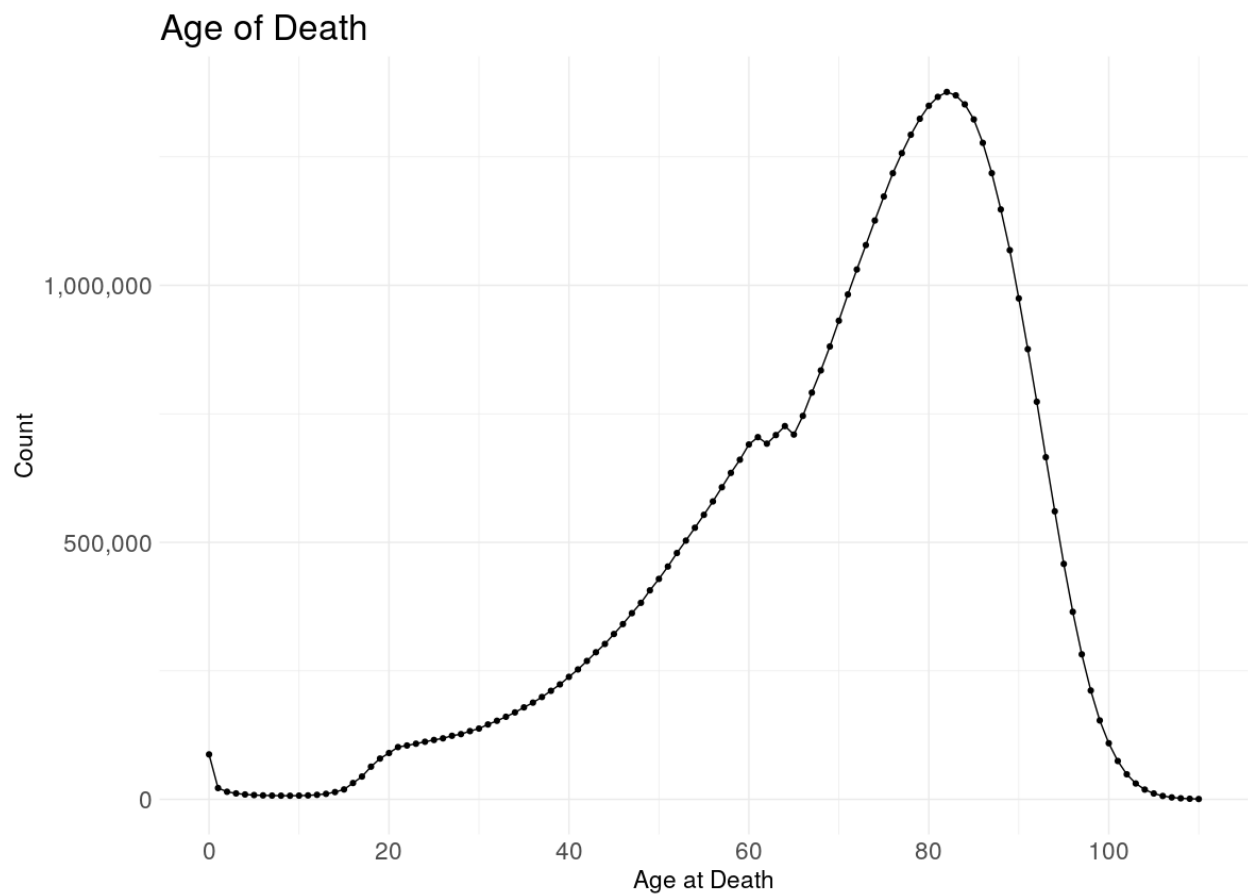
Age of Death

zip_residence

Label: ZIP Code of Residence at Time of Death

Description: zip_residence is a string variable (9-characters) reporting a person's ZIP code of residence at time of death, as in the Numident death records.

socstate

Label: State where Social Security Number Issued

Description: The state in which the person's social security card was issued. Determined by first three (3) digits of Social Security number, as recorded in Numident death records. The numeric codes match the detailed IPUMS-USA Birthplace codes.

Top 25 States by count:

socstate	label	n	freq
3600	New York	4980871	10.0
4200	Pennsylvania	3266642	6.6
600	California	3255785	6.6
1700	Illinois	2855031	5.8
NA	NA	2707652	5.5
4800	Texas	2562712	5.2
3900	Ohio	2460960	5.0
2600	Michigan	2034399	4.1
3400	New Jersey	1510334	3.1
2500	Massachusetts	1465566	3.0
2900	Missouri	1299368	2.6
1800	Indiana	1248048	2.5
1300	Georgia	1171471	2.4
4700	Tennessee	1115600	2.3
1200	Florida	1043773	2.1
5500	Wisconsin	1021002	2.1
5100	Virginia	1014527	2.1
100	Alabama	1010756	2.0
2100	Kentucky	917858	1.9
2200	Louisiana	899017	1.8
2700	Minnesota	817858	1.7
4000	Oklahoma	779894	1.6
1900	Iowa	764643	1.5
2400	Maryland	728497	1.5
5300	Washington	675565	1.4

father_fname

Label: Father's First Name

Description: father_fname is a character variable reporting the first 16 letters of the person's father's first name, as recorded in the Numident death records.

father_mname

Label: Father's Middle Name

Description: father_mname is a character variable reporting the first 16 letters of the person's father's middle name, as recorded in the Numident death records.

father_lname

Label: Father's Last Name

Description: father_lname is a character variable reporting the first 21 letters of the person's father's last name, as recorded in the Numident death records.

mother_fname

Label: Mother's First Name

Description: mother_fname is a character variable reporting the first 16 letters of the person's mother's first name, as recorded in the Numident death records.

mother_mname

Label: Mother's Middle Name

Description: mother_mname is a character variable reporting the first 16 letters of the person's mother's last name, as recorded in the Numident death records.

mother_lname

Label: Mother's Last Name

Description: mother_lname is a character variable reporting the first 21 letters of the person's mother's middle name, as recorded in the Numident death records.

age_first_application

Label: Age at First Social Security Application

Description: age_first_application reports the age at which the person submitted their first Social Security Application.

Age of First Application
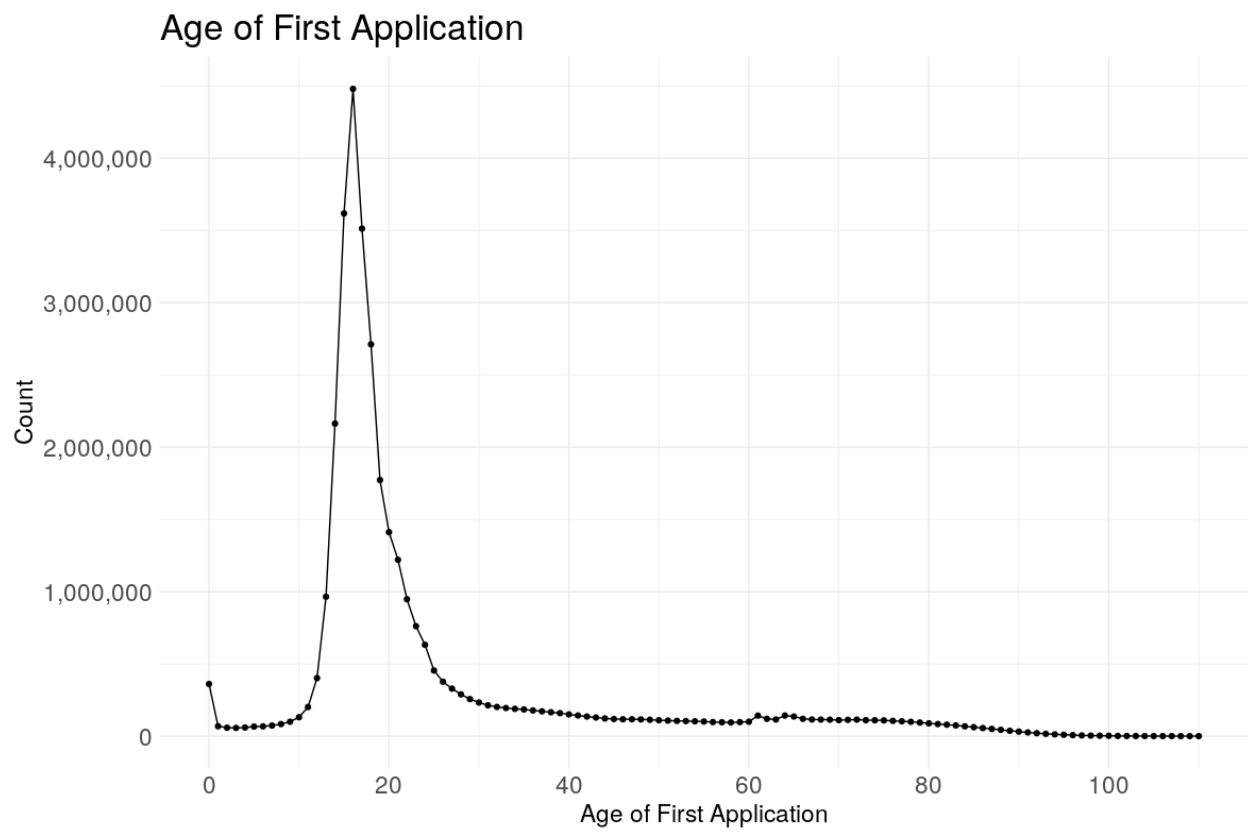
number_apps

Label: Total Number of Applications

Description: number_apps reports the person's total number of Social Security applications, as recorded in the Numident application records. This variable was top-coded at 10.

number_apps	n	freq %
0	14166703	29.000
1	17646617	36.000
2	10107569	20.000
3	4309930	8.700
4	1710582	3.500
5	719300	1.500
6	323112	0.650
7	155220	0.310
8	80363	0.160
9	43690	0.089
10	74741	0.150

number_claims

Label: Total Number of Claims

Description: number_claims reports the person's total number of Social Security claims, as recorded in the Numident claim records. This variable was top-coded at 3.

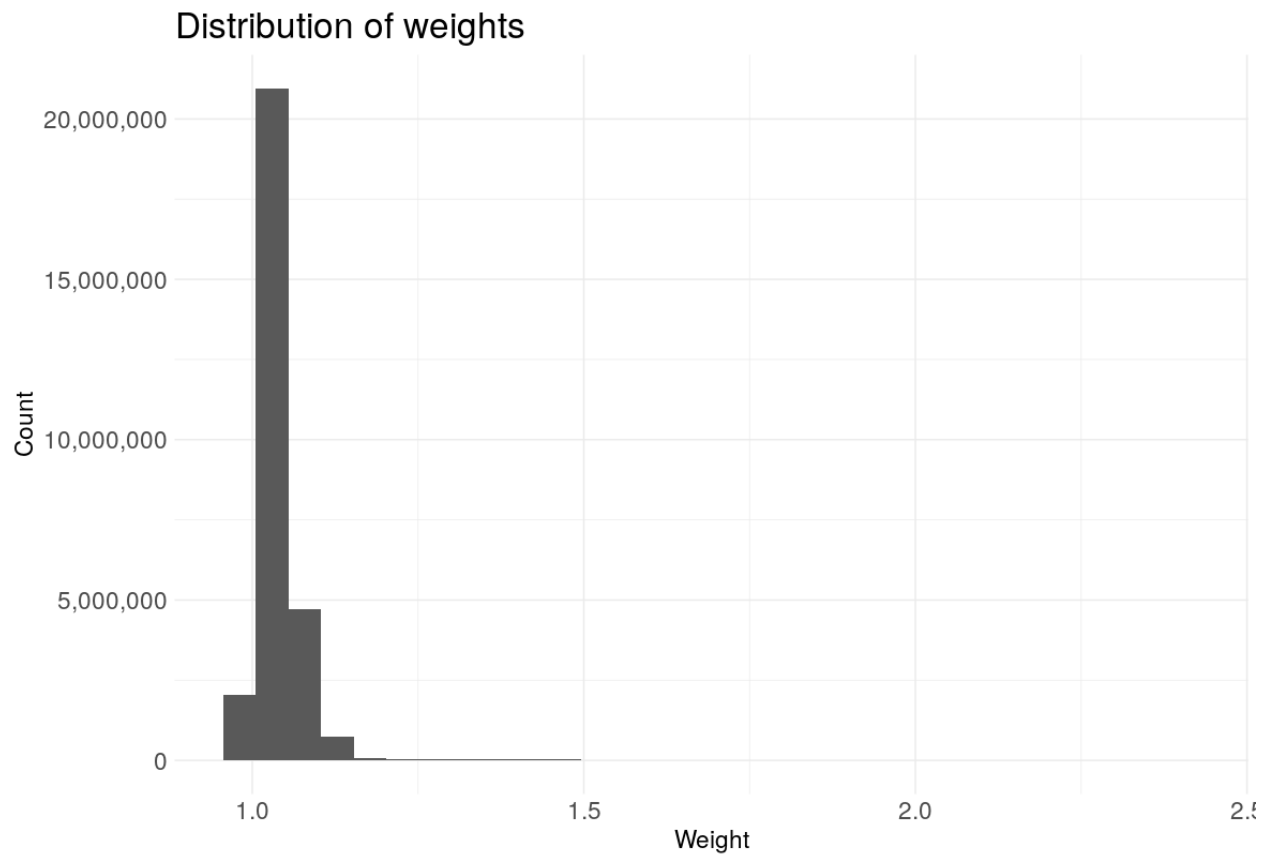| number_claims	n	freq
0	32931311	66.75
1	16360228	33.16
2	45187	0.09
3+	1101	0.00

weight

Label: High Coverage Sample Weight

Description: A post-stratification person-weight to Human Mortality Database (HMD) totals for persons (1) born between 1895-1940 (2) dying between 1988-2005 and (3) dying between ages 65-100 (4) with a non-missing value for sex. Please see the CenSoc methods protocol for more details on weighting procedure.

Value	Label
NA	No weight assigned
Min Weight	0.961
Max Weight	2.383

Distribution of weights

ccweight

Label: High Coverage "Complete Case" Sample Weight

Description: A post-stratification person-weight to Human Mortality Database (HMD) totals for persons (1) born between 1895-1940 (2) dying between 1988-2005 (3) dying between ages 65-100 (4) with a non-missing value for sex (5) with a non-missing value for birth place and (6) with a non-missing value for race. Please see the CenSoc methods protocol for more details on weighting procedure.

Value	Label
NA	Incomplete case (No weight assigned)
Min Weight	1.018
Max Weight	18.71

Distribution of ccweights