CenSoc-DMF Codebook*

Page	Variable	Label
2	HISTID	Historical unique identifier
3	byear	Year of birth
4	bmonth	Month of birth
5	dyear	Year of death
6	dmonth	Month of death
7	death_age	Age at death (years)
8	link_abe_exact_conservative	Flag for conservative ABE match
9	weight	CenSoc weight
10	weight_conservative	CenSoc weight (Conservative Sample)
11	pernum	Person number in sample weight
11	perwt	IPUMS person weight
11	age	Age in 1940
11	sex	Sex in 1940
11	bpl	Place of birth
11	mbpl	Mother's place of birth
11	fbpl	Father's place of birth
11	educd	Educational attainment (detailed)
11	empstatd	Employment status (detailed)
11	hispan	Hispanic/Spanish/Latino origin
11	inconwg	Had non-wage/salary income over $50
11	marst	Marital status
11	nativity	Foreign birthplace or parentage
11	occ	Occupation
11	occscore	Occupational income score
11	ownership	Ownership of dwelling (tenure)
11	race	Race
11	rent	Montly contract rent
11	serial	Household serial number
11	statefip	State of residence 1940
11	urban	Urban/rural status

Summary: The CenSoc-DMF Version 2.1 dataset (N = 7,762,221) links the full-count 1940 Census to the Death Master File (DMF), a collection of death records reported to the Social Security Administration. Records were linked using the standard and conservative ABE method developed by Abramitzky, Boustan, and Eriksson (2012, 2014, 2017). To merge 1940 Census variables with this dataset, researchers can obtain a copy of the 1940 Census from IPUMS-USA and link on the individual-level, unique identifier HISTID variable.

*Last updated: 05 January, 2023

HISTID

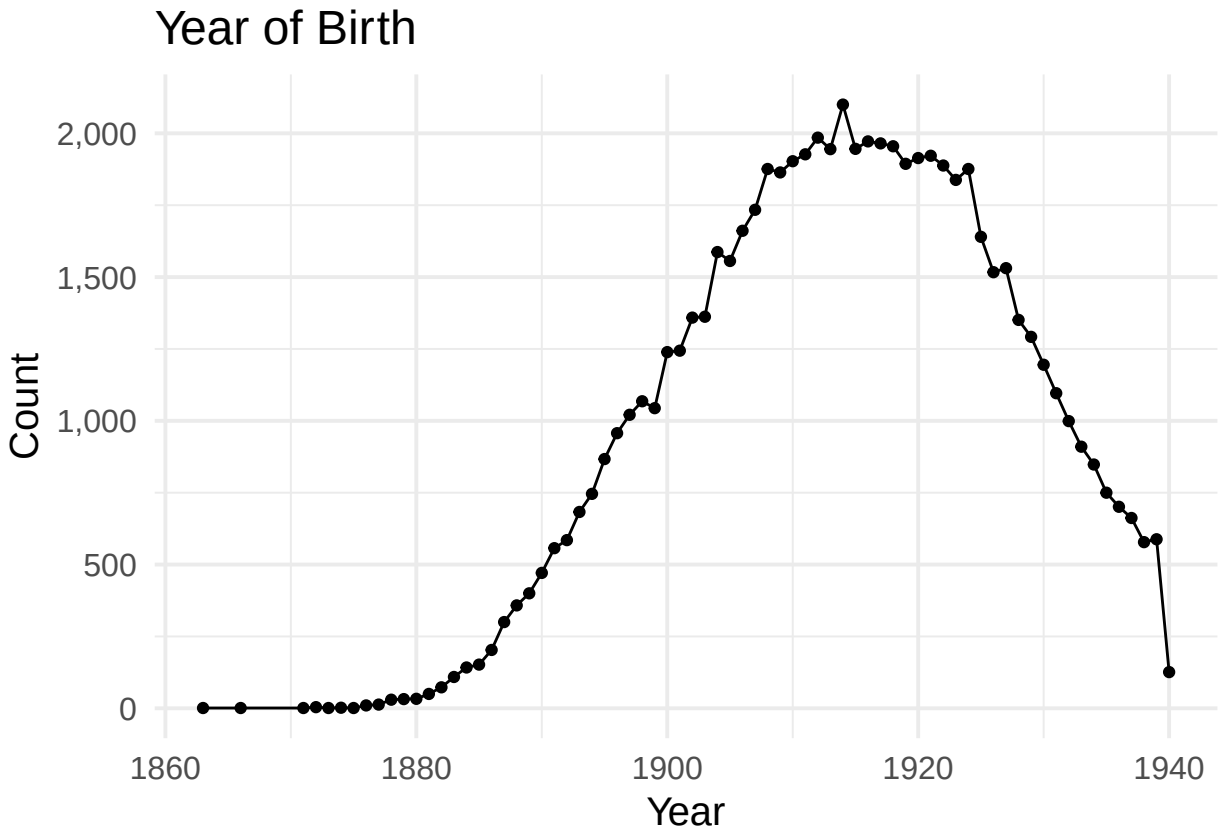
Label: Historical Unique Identifier

Description: HISTID is a unique individual-level identifier. It can be used to merge the CenSoc-DMF file with the 1940 Full-Count Census from IPUMS.

byear

Label: Birth Year

Description: byear reports a person's year of birth, as recorded in the Social Security Death Master File.



bmonth

Label: Birth Month

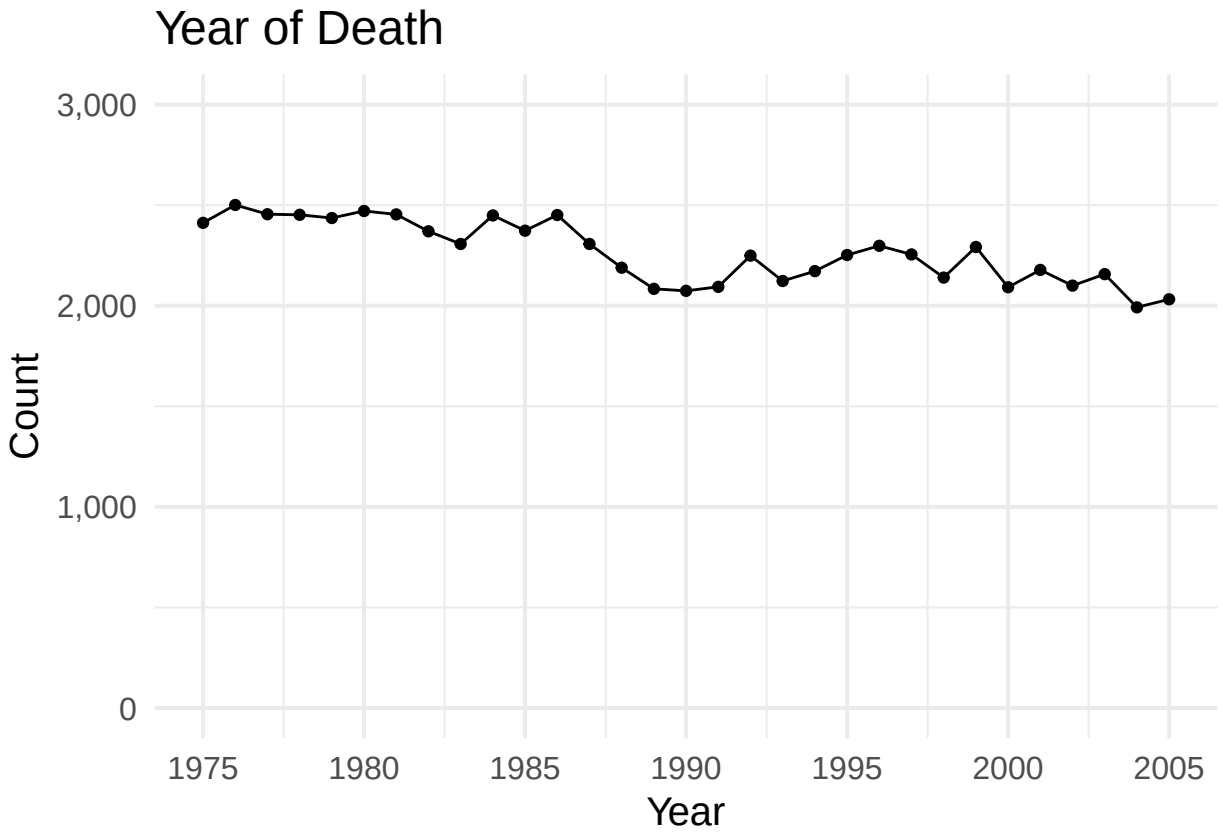
Description: bmonth reports a person's month of birth, as recorded in the Social Security Death Master File.

bmonth	label	n	freq %
1	January	6061	8.6
2	February	5642	8.0
3	March	6311	9.0
4	April	5558	7.9
5	May	5719	8.1
6	June	5488	7.8
7	July	5918	8.4
8	August	6231	8.9
9	September	6024	8.6
10	October	5952	8.5
11	November	5575	7.9
12	December	5732	8.2

dyear

Label: Death Year

Description: dyear reports a person's year of death, as recorded in the Social Security Death Master File.



dmonth

Label: Death Month

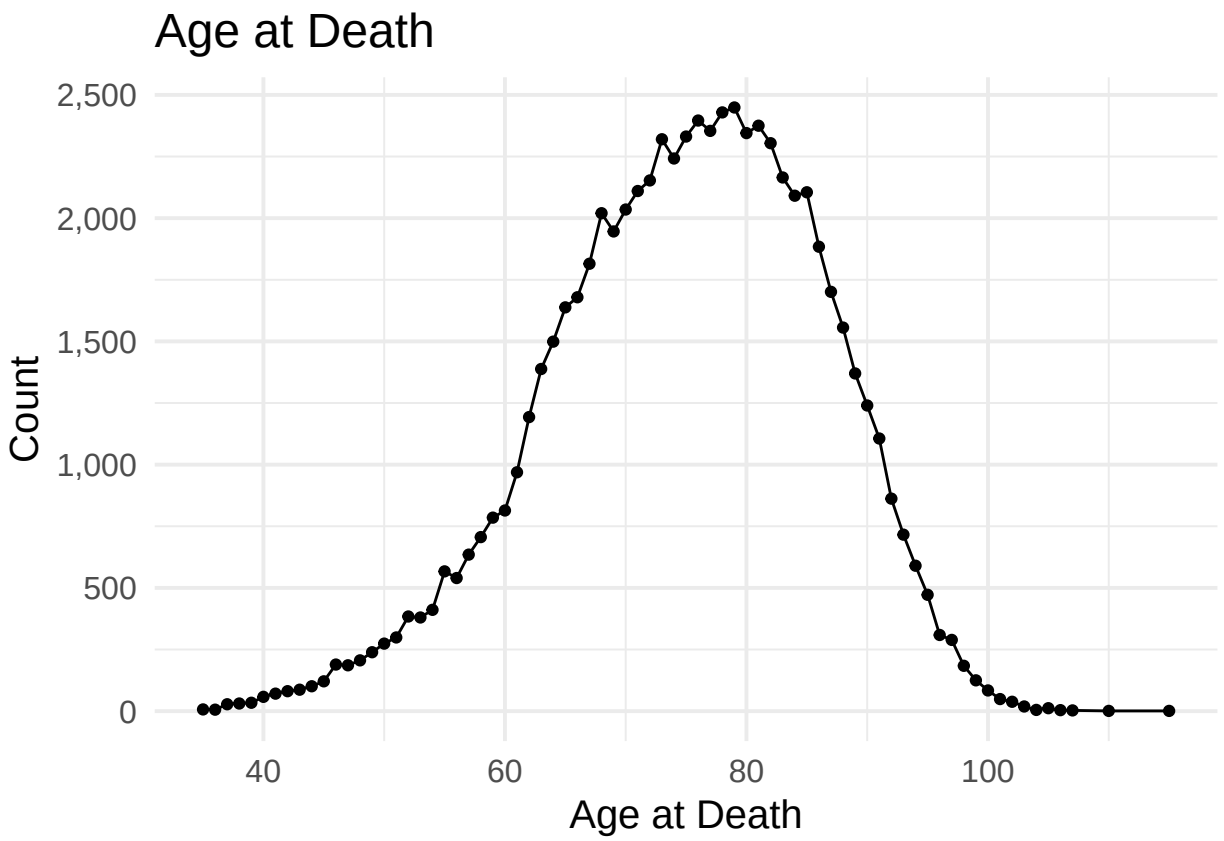
Description: dmonth reports a person's month of death, as recorded in the Social Security Death Master File.

dmonth	label	n	freq %
1	January	6743	9.6
2	February	5924	8.4
3	March	6312	9.0
4	April	5959	8.5
5	May	5810	8.3
6	June	5454	7.8
7	July	5541	7.9
8	August	5547	7.9
9	September	5303	7.6
10	October	5751	8.2
11	November	5670	8.1
12	December	6197	8.8

death_age

Label: Age at Death (Years)

Description: death_age reports a person's age at death in years, calculated using the birth and death information recorded in the Social Security Death Master File.



link_abe_exact_conservative

Label: Flag for conservative ABE match

Description: A flag variable reporting whether a match was established with the ABE conservative match with exact names.

link_abe_exact_conservative	label	n	freq %
1	Conservative and Standard ABE Link	42300	60.2
0	Standard ABE Link Only	27911	39.8

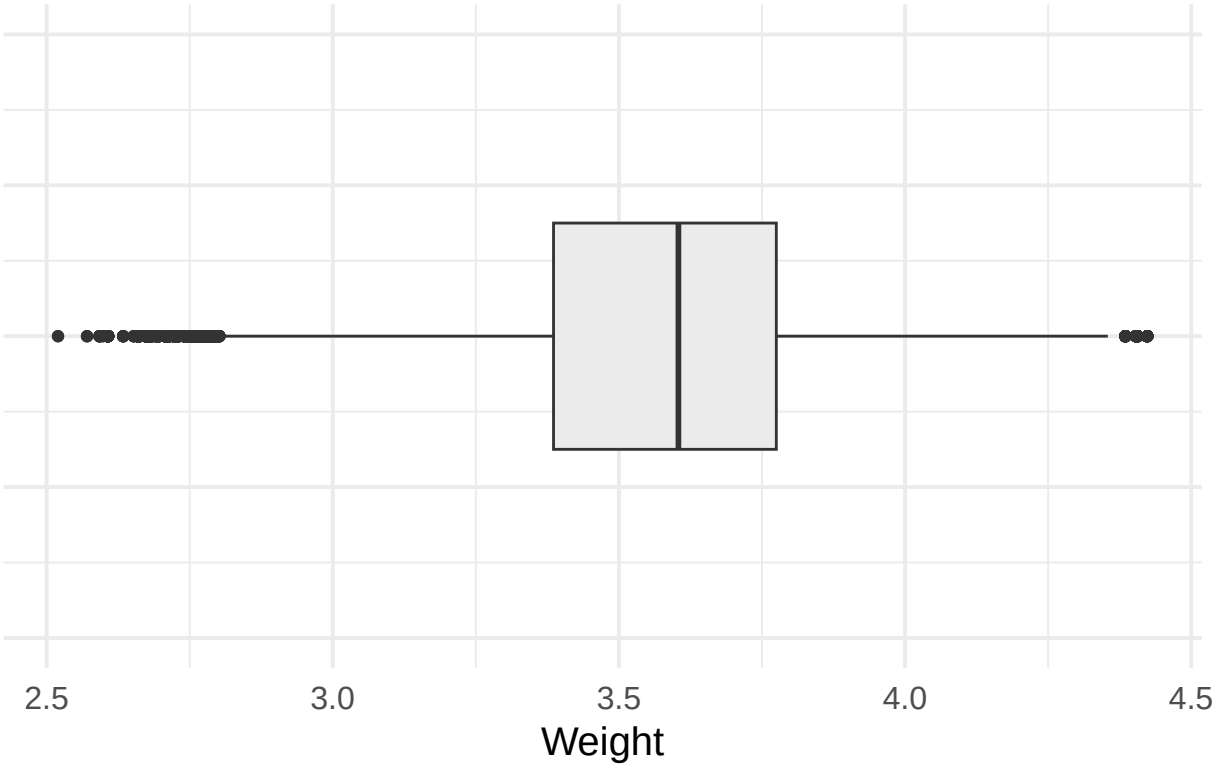
weight

Label: Sample Weights[1]

Description: A post-stratification person-weight to Human Mortality Database (HMD) totals for persons (1) born between 1895-1939 (2) dying between 1975-2005 (3) dying between ages 65-100. Please see the CenSoc Methods Protocol for more details on weighting procedure.

Value	Label
2.52	Min Weight
4.42	Max Weight
NA	No Weight Assigned

Distribution of Weights



[1]The IPUMS-USA 1940 1% sample also includes a weight (`perweight`) to account for the 1940 sampling procedure (thus no weights for the 100% complete count 1940 census). For analysis, we recommend using both sets of weights. A final weight can be constructed by multiplying the two weights together.

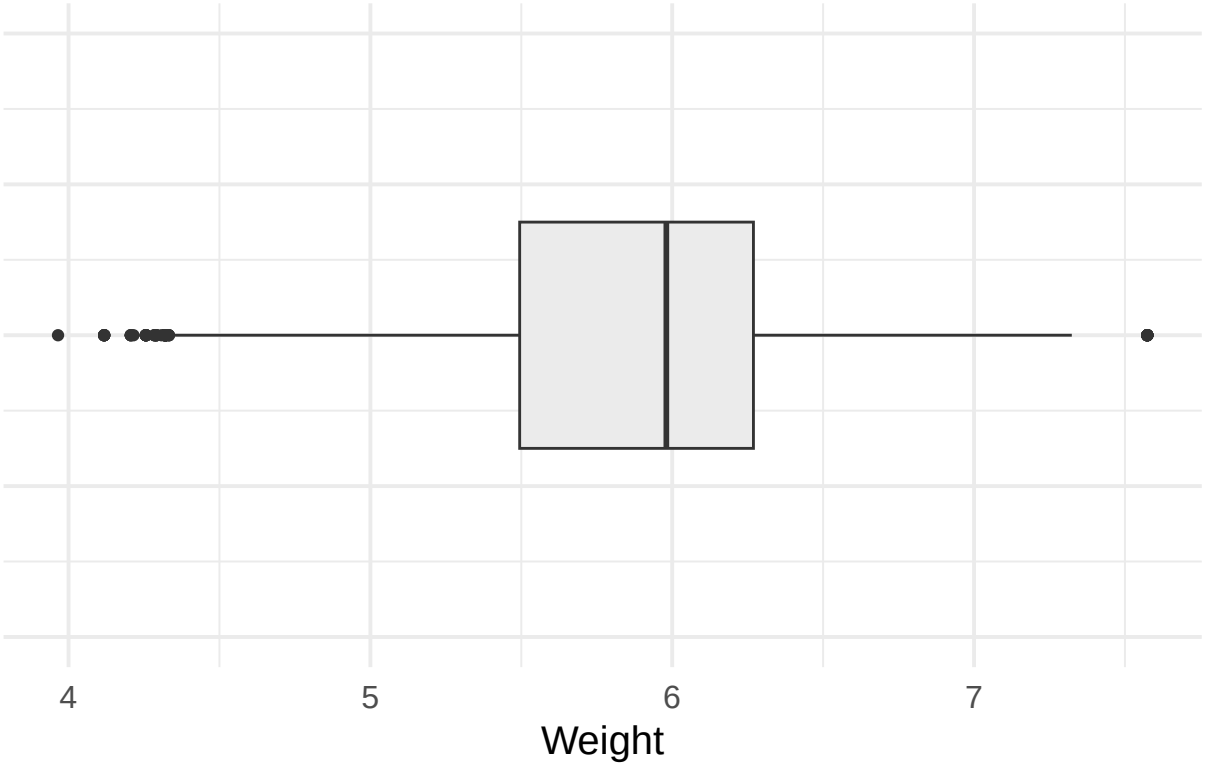
weight_conservative

Label: Sample Weights (Conservative Sample)

Description: A post-stratification person-weight to Human Mortality Database (HMD) totals (only for matches established via the conservative ABE algorithm) for persons (1) born between 1895-1939 (2) dying between 1975-2005 (3) dying between ages 65-100. Please see the CenSoc Methods Protocol for more details on weighting procedure.

Value	Label
3.97	Min Weight
7.57	Max Weight
NA	No Weight Assigned

Distribution of Weights



IPUMS 1940 Census Variable

The variables below are from the IPUMS-USA 1940 1% census sample. We recommend looking at the terrific documentation on the IPUMS-USA website: https://usa.ipums.org/usa/index.shtml.

Variable	Label
pernum	Person number in sample weight
perwt	IPUMS person weight[2]
age	Age in 1940
sex	Sex in 1940
bpl	Place of birth
mbpl	Mother's place of birth
fbpl	Father's place of birth
educd	Educational attainment (detailed)
empstatd	Employment status (detailed)
hispan	Hispanic/Spanish/Latino origin
inconwg	Had non-wage/salary income over $50
marst	Marital status
nativity	Foreign birthplace or parentage
occ	Occupation
occscore	Occupational income score
ownership	Ownership of dwelling (tenure)
race	Race
rent	Montly contract rent
serial	Household serial number
statefip	State of residence 1940
urban	Urban/rural status

[2]The `perweight` accounts for the 1940 sampling procedure to construct the 1% sample, and thus is only available in the 1940 1% sample. For analysis, we recommend using both the IPUMS `perweight` and the CenSoc `weight.` A final weight can be constructed by multiplying the two weights together