CenSoc-DMF 4.0 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	weight	CenSoc weight

Summary: The CenSoc-DMF Version 4.0 dataset (N = 5,315,385) links the full-count 1940 Census to the Death Master File (DMF) 1975-2023, a collection of death records reported to the Social Security Administration. Records were linked using the conservative variant of the 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.

Researchers should note that only data from years 1975-2020 is weighted. No weights are available for years 2021-2023.
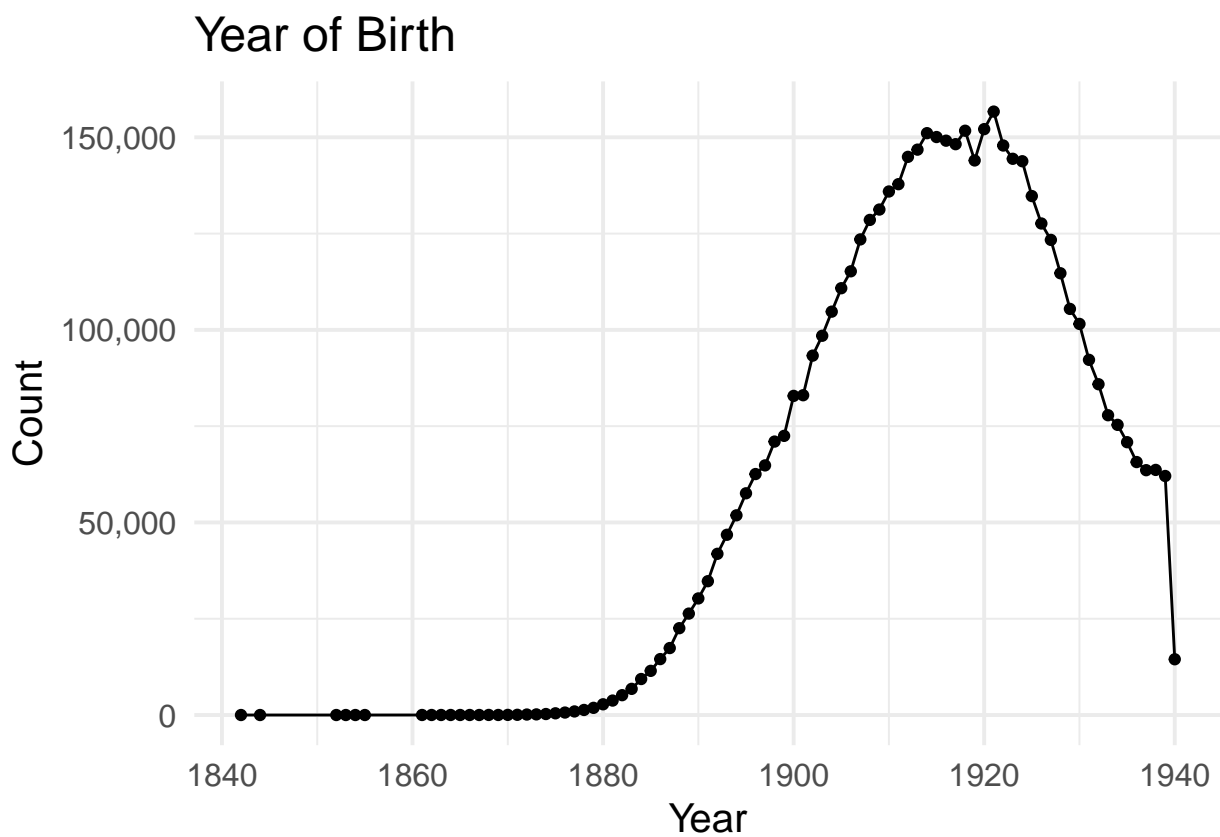

HISTID

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.
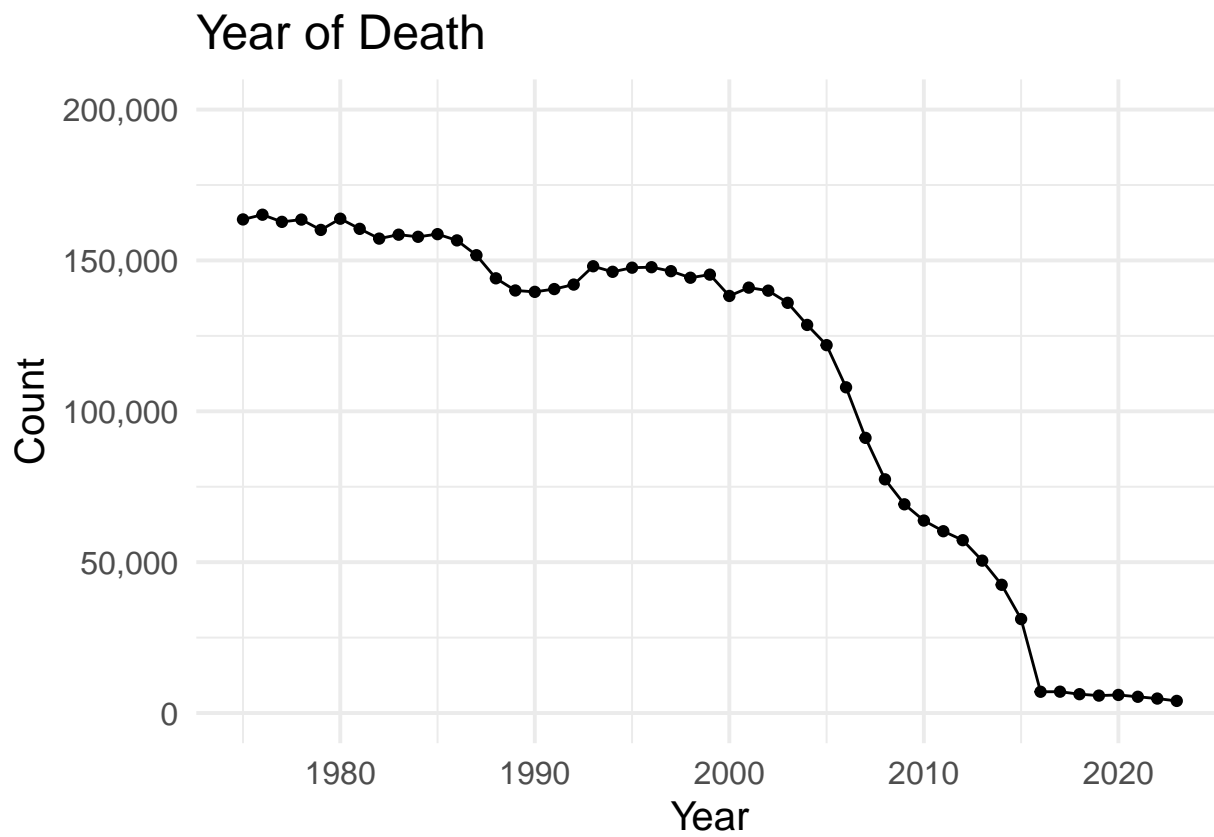
Year of Birth

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	453217	8.5
2	February	427887	8.0
3	March	470375	8.8
4	April	421566	7.9
5	May	432328	8.1
6	June	420742	7.9
7	July	451976	8.5
8	August	469931	8.8
9	September	461348	8.7
10	October	450409	8.5
11	November	420715	7.9
12	December	434890	8.2

dyear

Label: Death Year

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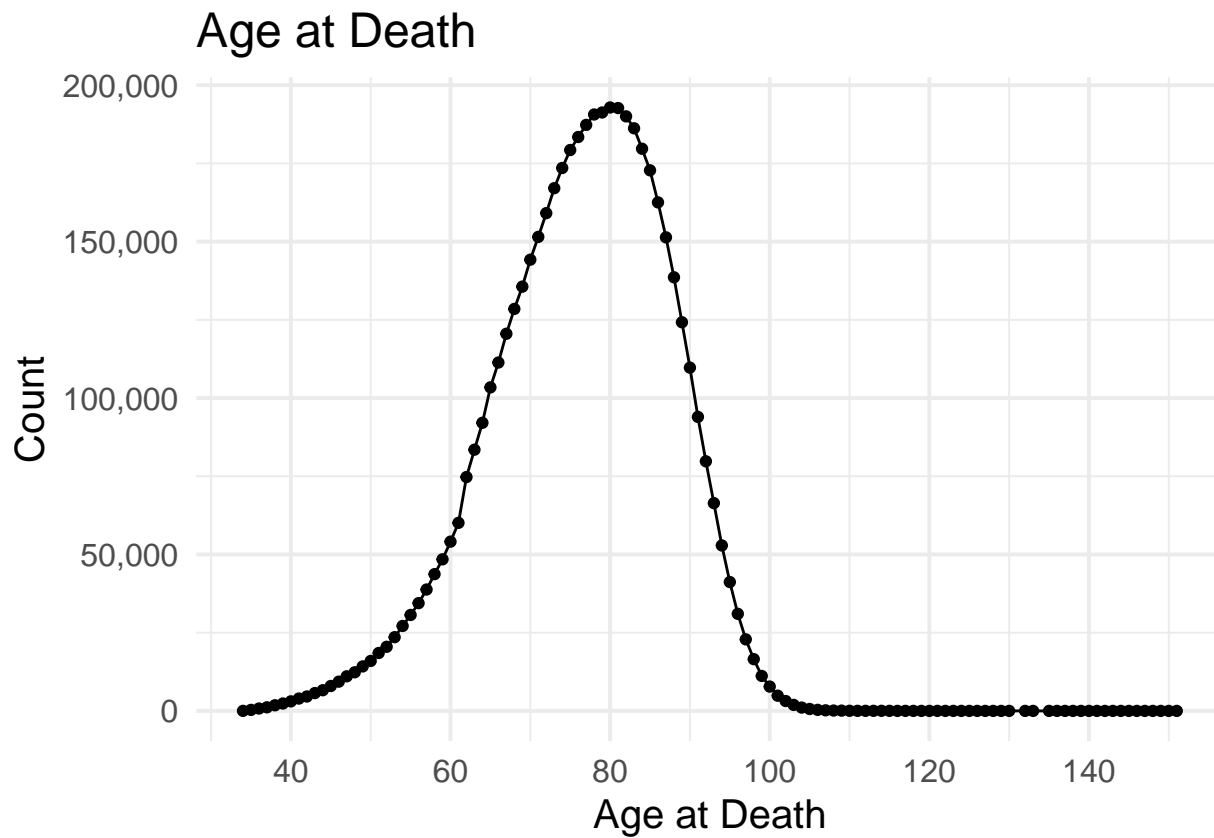
Year of Death

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	513196	9.7
2	February	454799	8.6
3	March	479859	9.0
4	April	442499	8.3
5	May	436956	8.2
6	June	412795	7.8
7	July	416542	7.8
8	August	411236	7.7
9	September	403235	7.6
10	October	435554	8.2
11	November	431308	8.1
12	December	477406	9.0

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.

weight

Label: Sample Weights

Description: A person-weight to National Center for Health Statistics (NCHS) totals for persons (1) dying between 1975-2020 (2) dying between ages 65-100. Weights are based on age at death, year of death, race, and place of birth.

Weights are higher at the end of the this period due to decreased death coverage in the DMF. Please see the technical documentation on our website for more information on how the DMF and how weights are created for this dataset. Users should note that death years after 2020 and cohorts after 1938 are not weighted.

Mean Weight by Year