

CenSoc-Numident Demo 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	zip_residence	ZIP Code of residence at time of death
9	weight	CenSoc weight
10	pernum	Person number in sample unit
10	perwt	IPUMS person weight
10	age	Age in 1940
10	sex	Sex in 1940
10	bpl	Place of birth
10	mbpl	Mother's place of birth
10	fbpl	Father's place of birth
10	educd	Educational attainment (detailed)
10	empstatd	Employment status (detailed)
10	hispan	Hispanic/Spanish/Latino origin
10	incnonwg	Had non-wage/salary income over \$50
10	incwage	Wage and salary income
10	marst	Marital status
10	nativity	Foreign birthplace or parentage
10	occ	Occupation
10	occscore	Occupational income score
10	ownership	Ownership of dwelling (tenure)
10	race	Race
10	rent	Monthly contract rent
10	serial	Household serial number
10	statefip	State of residence 1940
10	urban	Urban/rural status

Summary: The CenSoc-Numident Demo dataset was constructed by (i) linking the CenSoc-Numident dataset to the IPUMS 1940 1% census sample and (ii) selecting a set of 20 mortality covariates from the 1940 census. The smaller size of the file — approximately 1% of the records in the full CenSoc-Numident dataset — makes it easier to work with. However, the demo file isn't conducive to high-resolution mortality research, and we recommend working with the complete CenSoc-Numident file for any final analysis.

histid

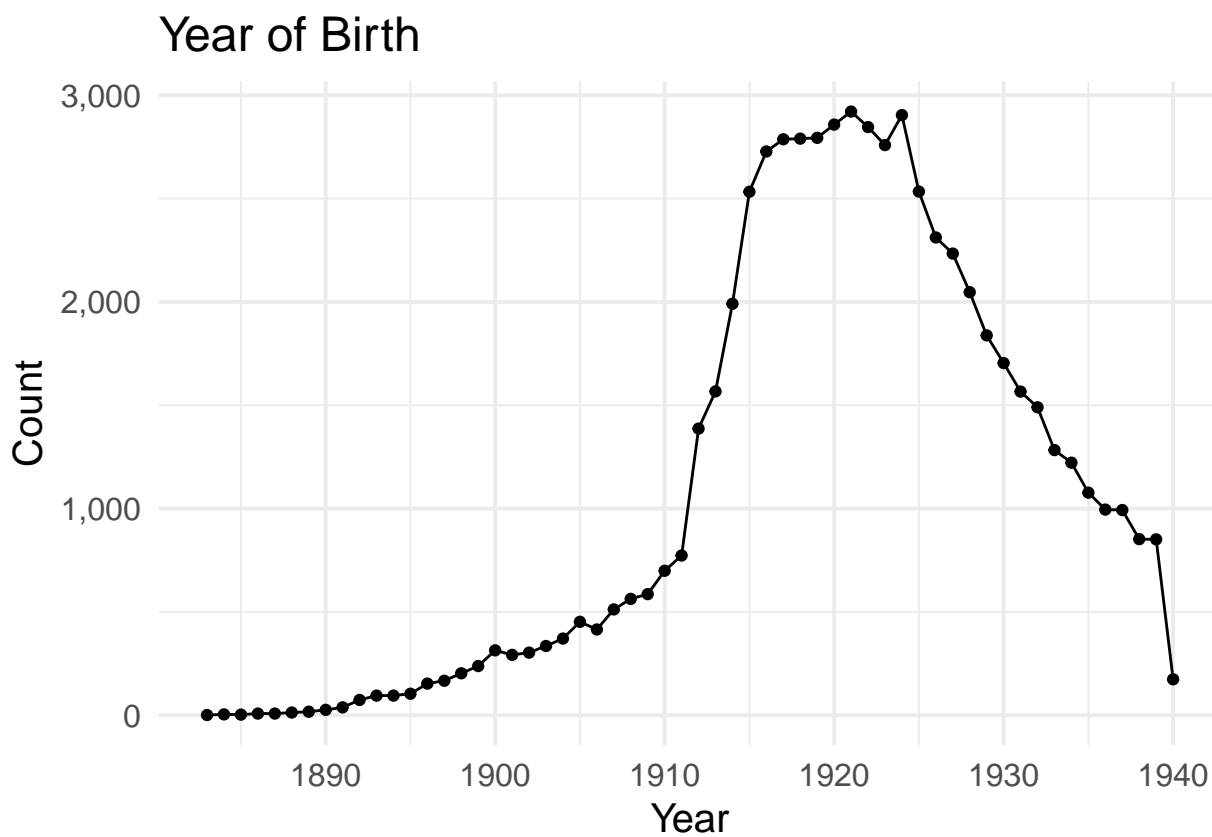
Label: Historical Unique Identifier

Description: histid is a unique individual-level identifier. It can be used to match CenSoc-Numident and IPUMS-USA 1940 Full-Count Census.

byear

Label: Birth Year

Description: byear reports a person's year of birth, as recorded in the Numident death records.



bmonth

Label: Birth Month

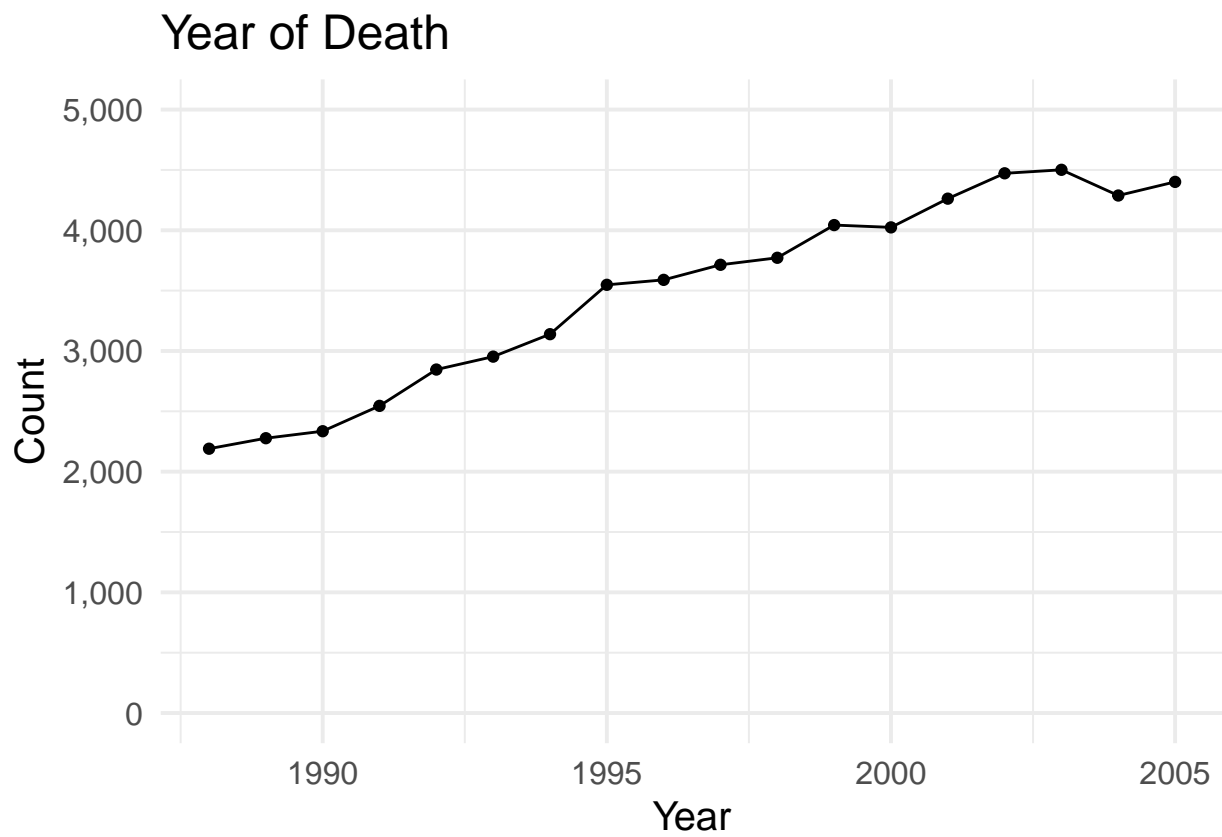
Description: bmonth reports a person's month of birth, as recorded in the Numident death records.

bmonth	label	n	freq %
1	January	5626	8.9
2	February	5350	8.5
3	March	5871	9.3
4	April	3255	5.2
5	May	4791	7.6
6	June	5039	8.0
7	July	5494	8.7
8	August	5868	9.3
9	September	5701	9.1
10	October	5518	8.8
11	November	5129	8.2
12	December	5257	8.4

dyear

Label: Death Year

Description: dyear reports a person's year of death, as recorded in the Numident death records.



dmonth

Label: Death Month

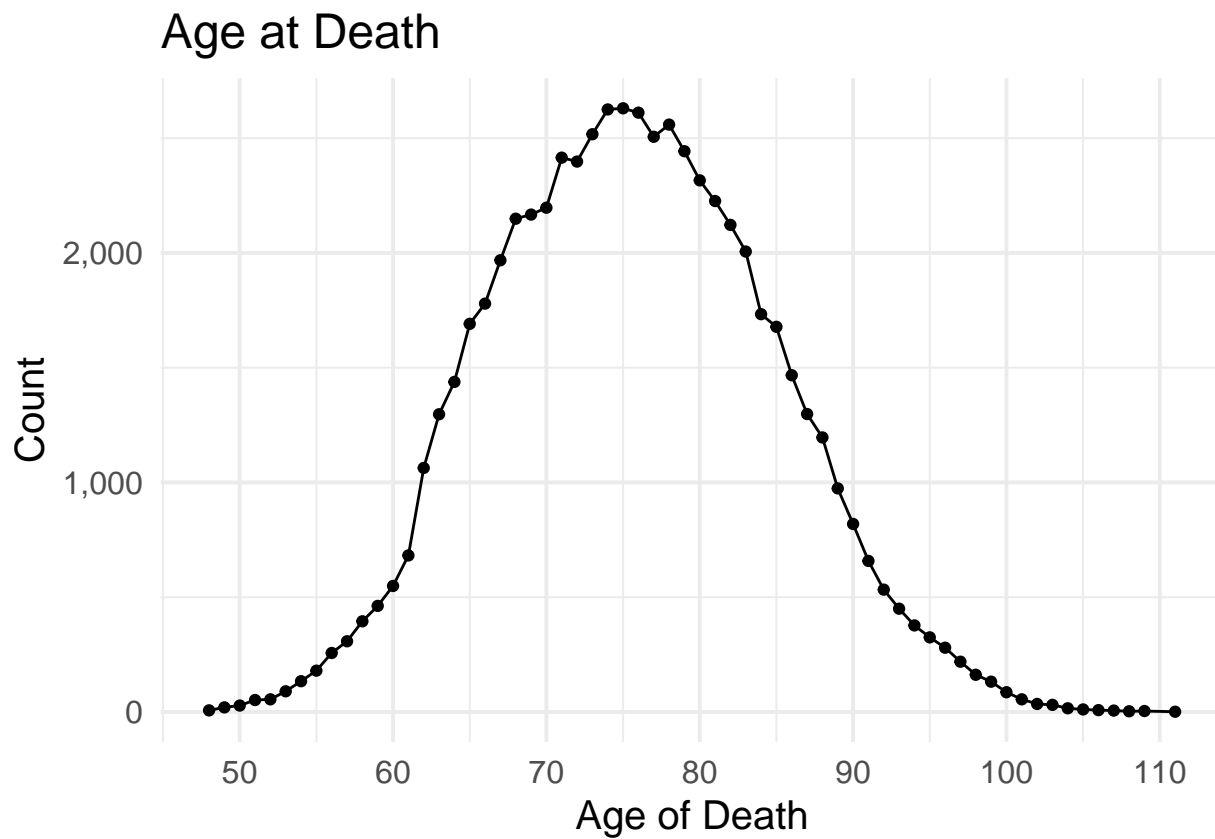
Description: dmonth reports a person's month of death, as recorded in the Numident death records.

dmonth	label	n	freq %
1	January	5782	9.2
2	February	5259	8.4
3	March	5729	9.1
4	April	5104	8.1
5	May	5205	8.3
6	June	4959	7.9
7	July	4884	7.8
8	August	4933	7.8
9	September	4800	7.6
10	October	5274	8.4
11	November	5256	8.4
12	December	5714	9.1

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 Numident death records.



zip_residence

Label: ZIP Code of Residence at Time of Death

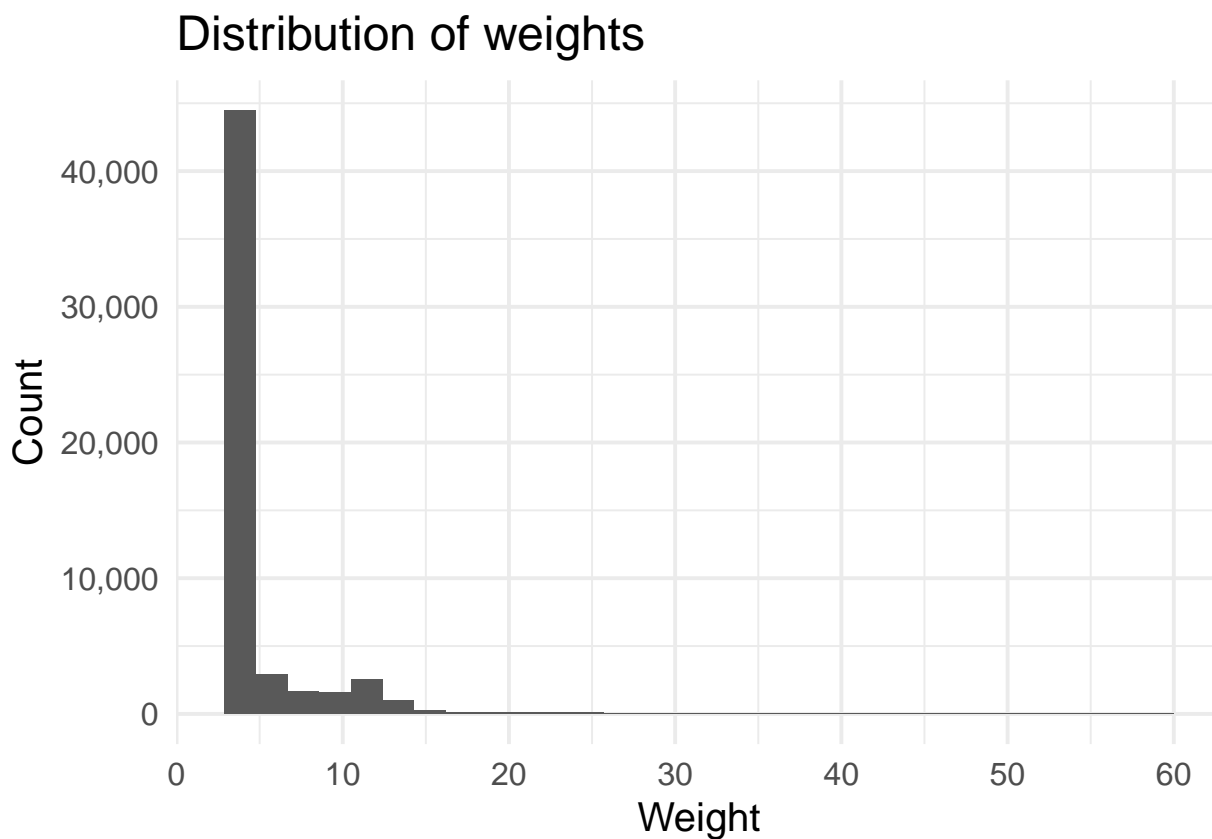
Description: zip_residence is a string variable (9-characters) reporting a person's ZIP Code of residence at time of death, as recorded in the Numident death records.

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
2.97	Min Weight
58.21	Max weight
NA	No Weight Assigned



¹The 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.

IPUMS 1940 Census Variables

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	Description
perwt	Person weight ²
age	Age on April 1st, 1940
sex	Sex
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
incnonwg	Had non-wage/salary income over \$50
incwage	Wage and salary income
marst	Marital status
nativity	Foreign birthplace or parentage
occ	Occupation
occscore	Occupational income score
ownership	Ownership of dwelling (tenure)
pernum	Person number in sample unit
race	Race
rent	Monthly contract rent
serial	Household serial number
statefip	State of residence 1940
urban	Urban/rural status

²The **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.