CenSoc Enlistment-Census-1940 : Linked WWII Military Enlistment and 1940 Census Data V1 *

Page	Variable	Label
2	HISTID	HISTID
3	byear	Year of Birth
4	sex	Sex
5	date_of_enlistment	Date of Enlistment
6	bpl	Place of Birth
7	residence_state	State of Residence at Enlistment
8	residence_county	County of Residence at Enlistment
9	place_of_enlistment	Place of Enlistment
10	education	Education
11	grade_code	Army Grade (rank)
12	branch_code	Army Branch
13	term_of_enlistment	Term of Enlistment
14	race	Race
15	citizenship	Citizenship
16	civilian_occupation	Civilian Occupation
17	marital_status	Marital Status at Enlistment
18	height	Height at Enlistment (inches)
19	weight_before_march_1943	Weight (pounds)
20	weight_or_AGCT	Weight (pounds) or AGCT score
21	component_enlistment	Army component
22	source	Source of Army Personnel

Summary: The CenSoc Enlistment-Census-1940 file links 2,573,678 men from World War II era (circa 1938-1947) army enlistment records to the 1940 census. We include the HISTID identifier to allow researchers to merge this file with IPUMS 1940 census data. All variables included here, such as marital status and education, were measured at time of enlistment and are not taken from census data. Matches between datasets were established using the conservative variant of the ABE method developed by Abramitzky, Boustan, and Eriksson (2012, 2014, 2017).

HISTID

Label: Historical Unique Identifier

Description: HISTID is a unique individual-level identifier. It can be used to merge this file to full count 1940 census data from IPUMS-USA.

byear

Label: Year of Birth

Description: byear is a numeric variable reporting a person's year of birth, as recorded in the World War II Army Enlistment Records.

Year of Birth

(plot: line chart of Count vs Year, peaking around 1921 near 190,000)
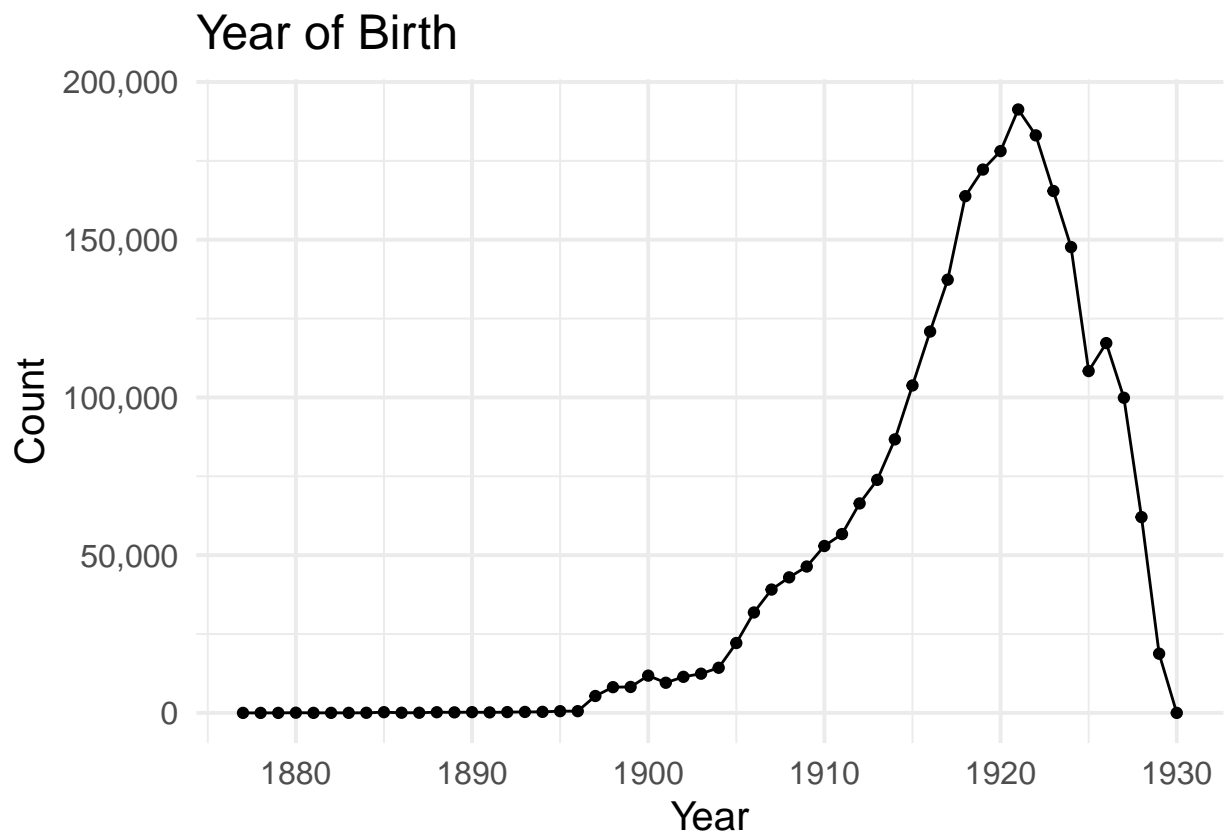
sex

Label: Sex

Description: sex is a numeric variable that reports a person's sex, as inferred by army branch in World War II ensistment records.

Note: Sex was not listed on enlistee punchcards and was inferred from army branch in enlistment records. Only men have been linked due to lower numbers enlisted women.

sex	label	n	freq %
1	Men	2573678	100

date_of_enlistment

Label: Date of Enlistment

Description: date_of_enlistment is a date variable that reports a person's date of enlistment in dd-mm-yyyy format, as reported in the World War II Army Enlistment Records.

Enlistment Date (Month and Year Only)

Count

Date (yyyy–mm)
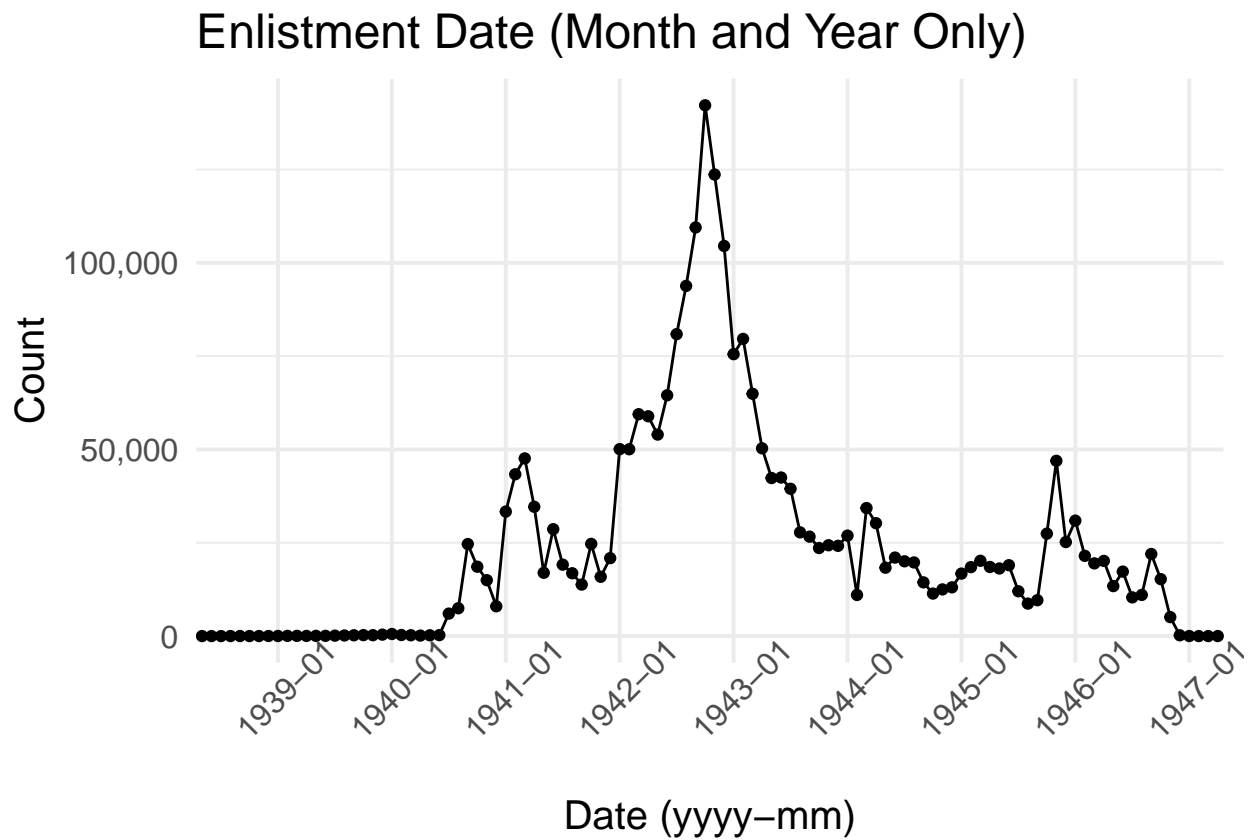
bpl

Label: Place of Birth

Description: bpl is a numeric variable that reports a person's place of birth, as reported in the World War II Army Enlistment Records.

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

residence_state

Label: State of residence

Description: residence_state is a character variable that reports a person's state of residence at time of enlistment, as reported in the World War II Army Enlistment Records. This variable is only available for enlistees residing in the US.

For a complete list of codes, please see: https://aad.archives.gov/aad/popup-codelist.jsp?cl_id=2046&dt=893&c_id=24996

residence_county

Label: County of residence

Description: residence_county is a character variable that reports a person's county of residence at time of enlistment, as reported in the World War II Army Enlistment Records. This variable is only available for enlistees residing in the US.

For a complete list of codes, please see: https://aad.archives.gov/aad/popup-codelist.jsp?cl_id=2057&dt=893&c_id=24998

place_of_enlistment

Label: Place of enlistment

Description: place_of_enlistment_enlistment is a character variable that reports where a person was enlisted, as reported in the World War II Army Enlistment Records.

For a complete list of codes, please see: https://aad.archives.gov/aad/popup-codelist.jsp?cl_id=4123&dt=893&c_id=24997

education

Label: Education

Description: education is a numeric variable that reports an individual's educational attainment at time of enlistment, as reported in the World War II Army Enlistment Records. Codes also available at: https://usa.ipums.org/usa-action/variables/EDUC#codes_section

Note: Exact number of years of schooling is unknown for enlistees who did not complete at least one year of high school. Such enlistees were given originally given a code meaning "grammar school", though it is unknown if this corresponds to the modern equivalent of a 5th, 6th, or 8th grade education. Many enlistees with this code likely had little to no actual education. Before the attack on Pearl Harbor, literacy was required of army inductees, but this requirement relaxed as the war effort increased (Brandt, 2004). By 1942, the Army accepted illiterate men if they were deemed trainable, though academic standards for admission increased again in 1944 and 1945.

education	label	n	freq %
1	Nursery school to grade 4	686708	26.700
3	Grade 9	213721	8.300
4	Grade 10	271080	10.500
5	Grade 11	249967	9.710
6	Grade 12	758652	29.500
7	1 year of college	162463	6.310
8	2 years of college	90552	3.520
9	3 years of college	44965	1.750
10	4 years of college	84032	3.270
11	5+ years of college	11538	0.448

grade_code

Label: Army grade

Description: grade_code is a numeric variable that records a person's grade (rank), as reported in the World War II Army Enlistment Records. Code list also is available at: https://aad.archives.gov/aad/popup-codelist.jsp?cl_id=2072&dt=893&c_id=24999

grade_code	label	n	freq %
1	Master Sergeant	7933	0.308
2	First Sergeant	3216	0.125
3	Technical Sergeant	7255	0.282
4	Staff Sergeant	13420	0.521
5	Sergeant	19435	0.755
6	Corporal	23321	0.906
7	Private First Class	68550	2.660
8	Private	2395450	93.100
9	Aviation Cadet	12364	0.480
NA	NA	22734	0.883

branch_code

Label: Army branch

Description: branch_code is a numeric variable that records the army branch in which the enlistee served, as reported in the World War II Army Enlistment Records.

branch_code	label	n	freq %
0	Branch Immaterial - Warrant Officers, USA	1164986	45.300000
1	General Staff Corps	57	0.002210
2	No branch assignment	819257	31.800000
3	Philippine Scouts	6738	0.262000
4	Professor US Military Academy	315	0.012200
5	Military Intelligence	489	0.019000
6	Corps of Military Police	4779	0.186000
7	Specialist Reserve	9	0.000350
8	Detached Enlisted Men's List	3624	0.141000
9	Inactive Reserve	915	0.035600
10	Infantry	59876	2.330000
20	Air Corps	305442	11.900000
30	Field Artillery	29851	1.160000
38	Transportation Corps	5525	0.215000
40	Coast Artillery Corps or Army Mine Planter Service	21467	0.834000
50	Quartermaster Corps	19679	0.765000
70	Medical Department - For Enlisted Men of the Medical Department (includes Enlisted Men of the Medical Department on Duty with the Dental Corps) or Army Nurse Corps	15027	0.584000
71	Medical Corps - For Officers only	13	0.000505
72	Hospital Dietitians	273	0.010600
73	Physical Therapy Aides	29	0.001130
74	Dental Corps - For Officers only	38	0.001480
75	Veterinary Corps or Service - For Officers of the Veterinary Corps and for Enlisted Men of the Veterinary Service	72	0.002800
77	Medical Administrative Corps - For Officers only	4	0.000155
78	Adjutant General's Department	330	0.012800
79	Pharmacy Corps	29	0.001130
80	Chemical Warfare Service	995	0.038700
82	Corps of Engineers	12808	0.498000
87	Ordnance Department	17544	0.682000
90	Signal Corps	28946	1.120000
93	Cavalry	8130	0.316000
96	Chaplains	74	0.002880
97	Finance Department	803	0.031200
98	Inspector General's Department	81	0.003150
99	Judge Advocate General's Department	2558	0.099400
NA	NA	42915	1.670000

term_of_enlistment

Label: Term of enlistment

Description: term_of_enlistment is a numeric variable that reports the length of a person's enlistment or their department of enlistment, as reported in the World War II Army Enlistment Records. Codes also available at: https://aad.archives.gov/aad/popup-codelist.jsp?cl_id=2076&dt=893&c_id=24979

term_of_enlistment	label	n	freq %
1	One year enlistment	58267	2.2600
2	Enlistment for Alaska	7050	0.2740
3	Enlistment for Hawaiian Department	124647	4.8400
4	Enlistment for assignment to another corps area	4821	0.1870
5	Enlistment for the duration of the War or other emergency, plus six months, subject to the discretion of the President or otherwise according to law	1814917	70.5000
6	Enlistment for the Panama Canal Department	97018	3.7700
7	Enlistment for the Philippine Department	24261	0.9430
8	Enlistment for the Puerto Rican Department	1763	0.0685
9	Three year enlistment	21442	0.8330
NA	NA	419499	16.3000

race

Label: Race

Description: race is a numeric variable that reports an individual's race, as reported in the World War II Army Enlistment Records.

Codes also available at: https://usa.ipums.org/usa-action/variables/RACE#codes_section

race	label	n	freq %
100	White	2390531	92.90000
150	Puerto Rican	293	0.01140
200	Black	150234	5.84000
300	Native American	4630	0.18000
400	Chinese	1506	0.05850
500	Japanese	3219	0.12500
600	Filipino	209	0.00812
630	Native Hawaiian	47	0.00183
991	Other	5716	0.22200
NA	NA	17293	0.67200

citizenship

Label: Citizenship

Description: citizenship is a numeric variable that reports an individual's citizenship, as reported in the World War II Army Enlistment Records.

citizenship	label	n	freq %
0	Non-Citizen	15260	0.593
1	Citizen	2541125	98.700
NA	NA	17293	0.672

civilian_occupation

Label: civilian occupation

Description: civilian_occupation is a numeric variable that reports a person's occupation at time of enlistment, as reported in the World War II Army Enlistment Records.

For complete codes, please see: https://aad.archives.gov/aad/popup-codelist.jsp?cl_id=3323&dt=893&c_id=24986

marital_status

Label: marital status

Description: marital_status is a numeric variable that reports a person's marital status at time of enlistment, as reported in the World War II Army Enlistment Records.

Codes also available at: https://usa.ipums.org/usa-action/variables/MARST#codes_section

marital_status	label	n	freq %
1	Married	589592	22.900
3	Separated	34251	1.330
4	Divorced	48168	1.870
5	Widowed	9135	0.355
6	Single	1887849	73.400
NA	NA	4683	0.182

height

Label: Height

Description: height is a numeric variable that reports a person's height in inches at time of enlistment, as reported in the World War II Army Enlistment Records. Height was filtered based on the range of permissible heights for enlistment in the military, 60 - 78 inches (Foster et al, 1967).

Note: Please be aware that instructions on how to use this field on enlistment punch cards varied over time. This field may encode military occupation or other information depending on time and site of enlistment. We have removed height data for men who enlisted in March of 1943 or later, as the distribution of this variable changes significantly at that point. However, it is possible that other data remains in this field. See the National Archives' technical documentation on the Amy Serial Number Electronic File for further information.
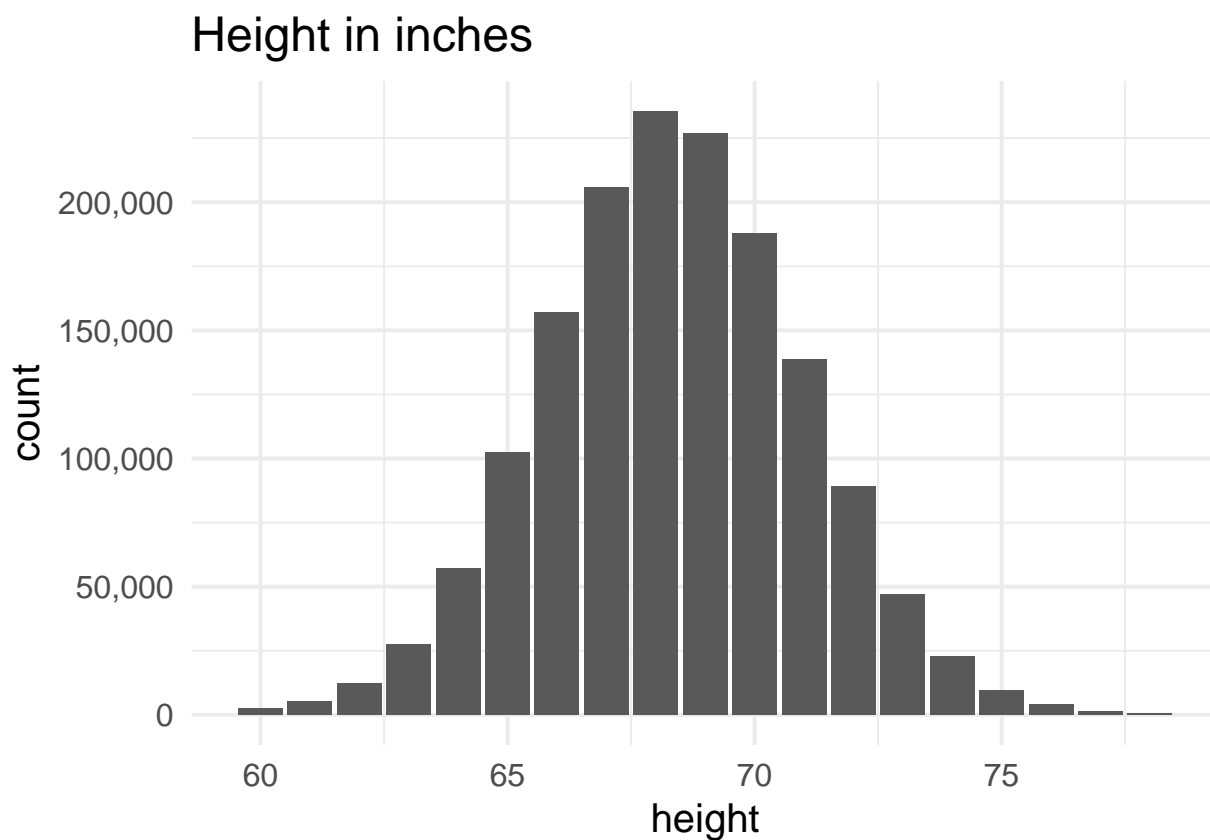
weight_before_march_1943

Description: weight_before_march_1943 is a numeric variable that reports a person's weight in pounds at time of enlistment as reported in the World War II Army Enlistment Records. This variable is only available for enlistees who enlisted before the March of 1943. Weight is filtered to a minimum of 105 pounds, the minimum permissible weight for a 60-inch-tall enlistee (Foster et al, 1967). While there was no clear maximum body weight for army enlistees, we have removed values of over 500.

Note: On enlistment punchcards, the field used to record weight was also used for Army General Classification Test (AGCT) scores, and may also encode military occupation or other information. Weight may have been recorded through 1943 at some enlistment sites, but the recording of AGCT score instead of weight at many enlistment sites likely began in March 1943. Because of this ambiguity, we only make this variable available for enlistees who enlisted before 1943. Refer to the National Archives' technical documentation on the Amy Serial Number Electronic File for further information. For enlistees who enlisted starting in March 1943, we have created an additional variable called **weight_or_AGCT**.
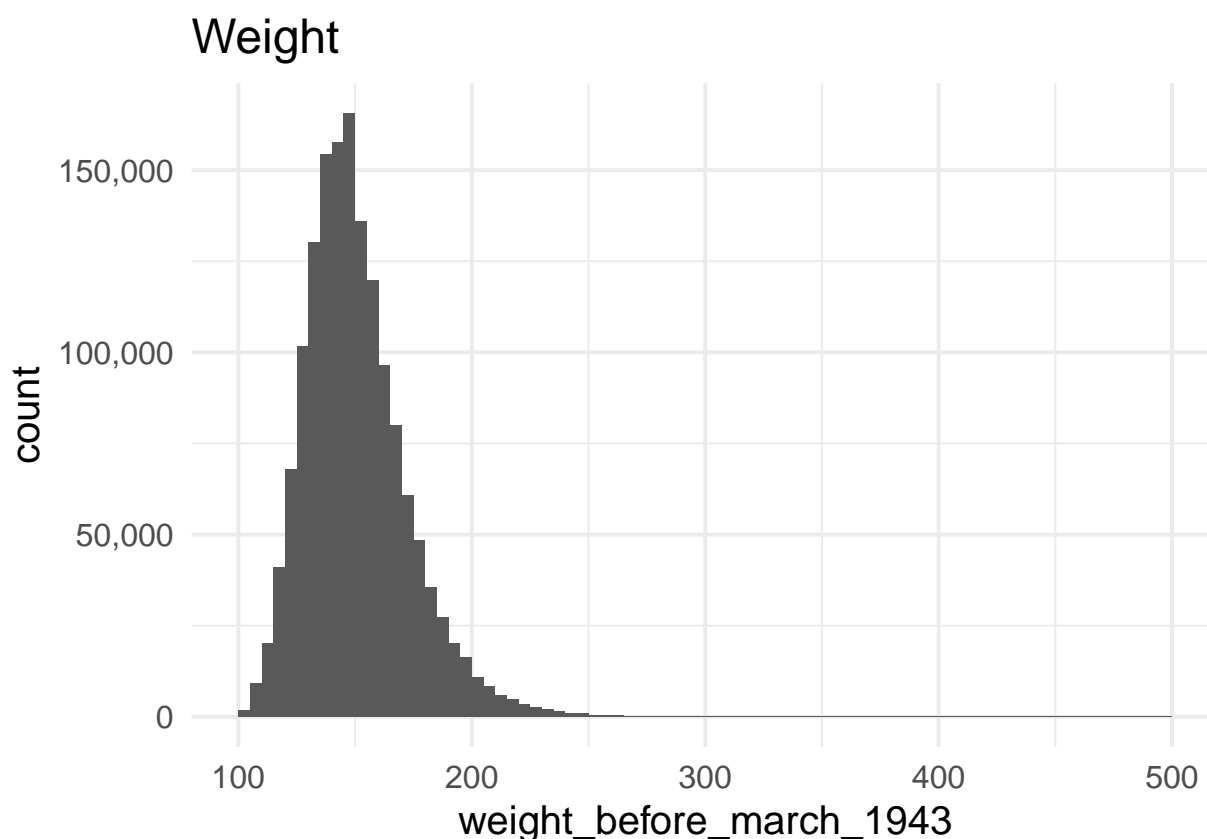
weight_or_AGCT

Label: Weight (pounds) at enlistment or AGCT score

Description: weight_or_AGCT is a numeric variable that reports a person's weight in pounds at time of enlistment or their Army General Classification Test (AGCT) score, as reported in the World War II Army Enlistment Records. This variable is only available for enlistees who enlisted in January 1943 or later. We have removed zeros and values above 500, which presumably exceed the maximum realistic body weights and AGCT scores, but otherwise leave the data in this field unfiltered and unsorted.

Note: On enlistment punchcards, the field used to record weight was also used for Army General Classification Test (AGCT) scores, and may also encode military occupation or other information. Weight may have been recorded through 1943 at some enlistment sites, but the recording of AGCT score instead of weight at many enlistment sites likely began in March 1943. Because of this ambiguity, we created this variable for persons enlisted in March 1943 or later, which may contain weight or AGCT or other information.
Refer to the National Archives' technical documentation on the Amy Serial Number Electronic File for further information. For enlistees prior to January 1943, we used the information recorded in this field to create the **weight_before_march_1943** variable.

Weight or AGCT Score
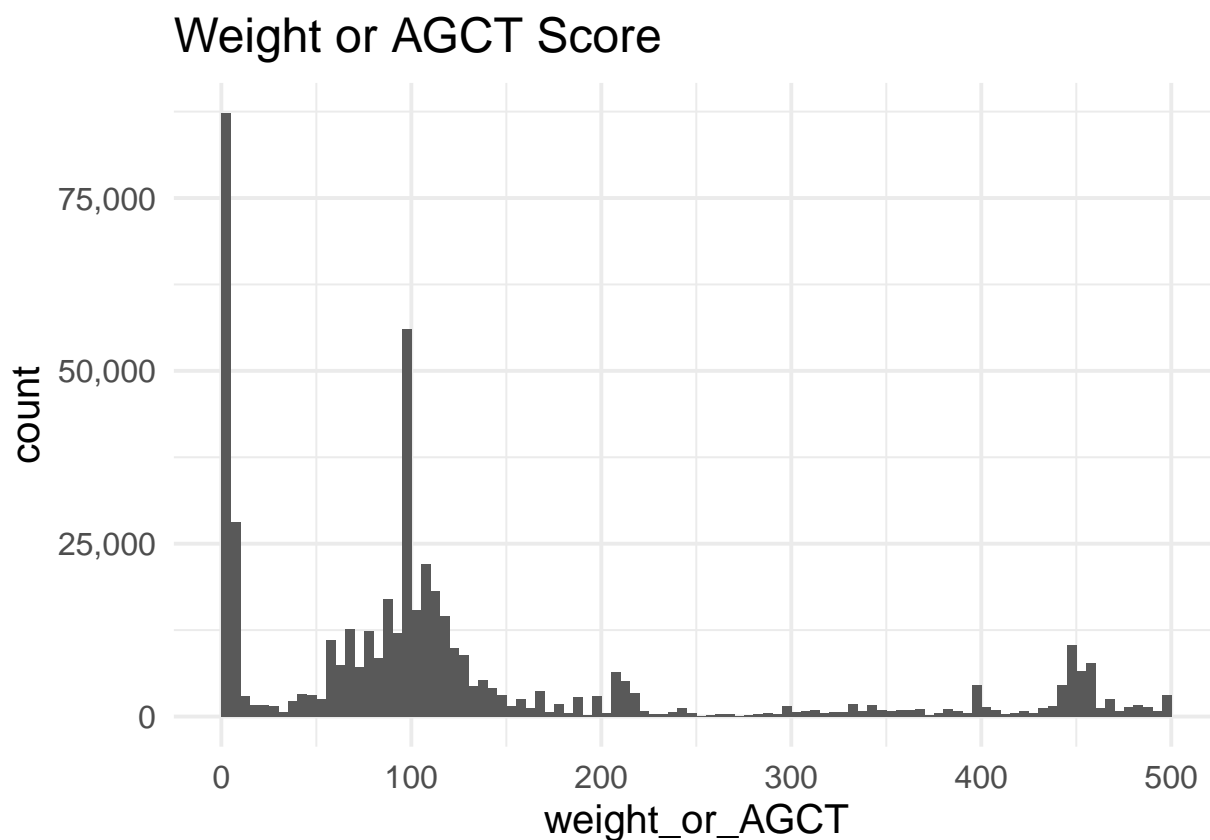
component

Label: Army Component

Description: component is a numeric variable that records the component of the army in which an individual served, as recorded in the World War II Army Enlistment Records.

Codes also available at: https://aad.archives.gov/aad/popup-codelist.jsp?cl_id=2078&dt=929&c_id=25014

component	label	n	freq %
1	Regular Army	348983	13.60000
2	Regular Army Reserve	83	0.00322
3	Reserves	173148	6.73000
4	National Guard	70893	2.75000
5	Philippine Scouts	33	0.00128
6	Army of the United States	211352	8.21000
7	Selectees	1764681	68.60000
9	Women's Army Corps	217	0.00843
NA	NA	4288	0.16700

source

Label: Source of army personnel

Description: source is a numeric variable that reports the source of the enlisted person, as reported in the World War II Army Enlistment Records. This denotes if an enlistee was a civilian prior to enlistment, or if they were already enlisted in the National Guard or another faction of the military. Codes also available at: https://aad.archives.gov/aad/popup-codelist.jsp?cl_id=2077&dt=893&c_id=24981

source	label	n	freq %
0	Civil Life	2130200	82.800000
1	Enlisted Man, Regular Army, within 3 months of Discharge or former WAAC Auxiliary	5676	0.221000
2	National Guard in Federal Service, within 3 months of Discharge	122331	4.750000
3	Enlisted Man, Regular Army, after 3 months of Discharge	14935	0.580000
4	National Guard in Federal Service, after 3 months of Discharge	12436	0.483000
5	Enlisted Reserve or Medical Administrative Corps (MAC) Officer	56790	2.210000
6	Enlisted Man, Philippine Scout or recall to AD of an enlisted man who had been transferred to the ERC	104052	4.040000
7	National Guard	73822	2.870000
8	Returned to Military Control, Surrendered	1381	0.053700
9	Returned to Military Control Apprehended	6	0.000233
11	Selective Service within 3 months of Discharge	9116	0.354000
12	Selective Service after 3 months of Discharge	51	0.001980
NA	NA	42882	1.670000