

CenSoc-DMF-WW2-Army-Enlistment File Codebook*

Page	Variable	Label
2	serial_number	Army Serial Number
3	yob	Year of Birth
4	fname	First Name
5	mname	Middle Name
6	lname	Last Name
7	sex	Sex
8	date_of_enlistment_ddmmyyyy	Date of Enlistment
9	place_of_birth	Place of Birth
10	residence_state	State of Residence at Enlistment
11	residence_county	County of Residence at Enlistment
12	place_of_enlistment	Place of Enlistment
13	education	Education
14	grade_code	Army Grade (rank)
15	branch_code	Army Branch
16	term_of_enlistment	Term of Enlistment
17	race	Race
18	citizenship	Citizenship
19	civilian_occupation	Civilian Occupation
20	marital_status	Marital Status at Enlistment
21	height	Height at Enlistment (inches)
22	weight_or_AGCT	Weight (pounds) at Enlistment or AGCT
score		
23	component_enlistment	Army component
24	source	Source of Army Personnel

Summary: The CenSoc-DMF-WW2-Army-Enlistment File (N = 1,541,899) links the full-count 1940 Census to the Social Security Death Master File (DMF) and the National Archives' public release of the United States World War II Army Enlistment Records, circa 1938-1946. Records were linked using the standard variant of the ABE method developed by Abramitzky, Boustan, and Eriksson (2012, 2014, 2017). Variables ending in '_dmf' are from the DMF, and those ending in '_enlistment' are from the Army Enlistment file. To merge 1940 Census data with this data set, researchers can obtain a copy of the 1940 Census from IPUMS-USA and link on the individual-level, unique identifier HISTID variable.

*Last updated: 21 April, 2023

serial_number

Label: Army Serial Number

Description: serial_number is an assigned number.

yob

Label: Year of Birth

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



fname

Label: First Name

Description: fname, utilized ABE cleaning method.

mname

Label: Middle Name

Description: mname reports a person's middle name, utilized ABE cleaning method.

lname

Label: Last Name

Description: lname reports a persons last name cleaned, utilized ABE cleaning method.

SEX

Label: Sex

Description: sex reports a person's sex, as recorded in the Social Security Death Master File.

date_of_enlistment_ddmmyyy

Label: Date of Enlistment

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

```
## [1] "00"
```

```
## [1] "99"
```

```
## # A tibble: 101 x 2
```

```
##   year_of_birth      n
```

```
##   <chr>         <int>
```

```
## 1 00          75666
```

```
## 2 01          37455
```

```
## 3 02          44819
```

```
## 4 03          47044
```

```
## 5 04          54527
```

```
## 6 05          87980
```

```
## 7 06         115443
```

```
## 8 07         139574
```

```
## 9 08         152448
```

```
## 10 09        164656
```

```
## 11 10        192693
```

```
## 12 11        200464
```

```
## 13 12        232993
```

```
## 14 13        258405
```

```
## 15 14        302627
```

```
## 16 15        359775
```

```
## 17 16        415755
```

```
## 18 17        475096
```

```
## 19 18        570516
```

```
## 20 19        614067
```

```
## 21 20        624811
```

```
## 22 21        659665
```

```
## 23 22        623293
```

```
## 24 23        557203
```

```
## 25 24        511462
```

```
## 26 25        365790
```

```
## 27 26        374861
```

```
## 28 27        326087
```

```
## 29 28        193051
```

```
## 30 29         59317
```

```
## 31 30          790
```

```
## 32 31          794
```

```
## 33 32         1861
```

```
## 34 33          306
```

```
## 35 34          249
```

```
## 36 35          371
```

```
## 37 36          176
```

```
## 38 37           34
```

```
## 39 38           79
```

```
## 40 39          223
```

```
## 41 40         2677
```


##	42 41	655
##	43 42	1398
##	44 43	661
##	45 44	2594
##	46 45	3795
##	47 46	811
##	48 47	92
##	49 48	424
##	50 49	1717
##	51 50	4193
##	52 51	1013
##	53 52	1637
##	54 53	912
##	55 54	3426
##	56 55	29699
##	57 56	2151
##	58 57	642
##	59 58	2586
##	60 59	2754
##	61 60	1306
##	62 61	459
##	63 62	586
##	64 63	190
##	65 64	365
##	66 65	2032
##	67 66	1365
##	68 67	75
##	69 68	582
##	70 69	675
##	71 70	77
##	72 71	462
##	73 72	255
##	74 73	41
##	75 74	64
##	76 75	81
##	77 76	40
##	78 77	119
##	79 78	180
##	80 79	150
##	81 80	2402
##	82 81	692
##	83 82	367
##	84 83	216
##	85 84	710
##	86 85	5141
##	87 86	854
##	88 87	650
##	89 88	4558
##	90 89	2987
##	91 90	2684
##	92 91	1146
##	93 92	1686
##	94 93	1427
##	95 94	2145

```

## 96 95          4748
## 97 96          3324
## 98 97          20950
## 99 98          32762
## 100 99         35023
## 101 <NA>       160393

## [1] "00"
## [1] "T"

## # A tibble: 101 x 2
##   date_of_enlistment_yy      n
##   <chr>                <int>
## 1 00                  12480
## 2 01                   993
## 3 02                  3143
## 4 03                   691
## 5 04                  1089
## 6 05                  1208
## 7 06                   294
## 8 07                   156
## 9 08                    73
## 10 09                  394
## 11 10                  211
## 12 11                  390
## 13 12                   46
## 14 13                   37
## 15 14                   94
## 16 15                   41
## 17 16                   17
## 18 17                  109
## 19 18                    5
## 20 19                   21
## 21 20                   68
## 22 21                   10
## 23 22                   65
## 24 23                   92
## 25 24                  176
## 26 25                   67
## 27 26                   15
## 28 27                  151
## 29 28                    9
## 30 29                   26
## 31 30                   78
## 32 31                   40
## 33 32                  180
## 34 33                  185
## 35 34                  375
## 36 35                  307
## 37 36                   82
## 38 37                   86
## 39 38                  2040
## 40 39                 49476
## 41 40                348927

```

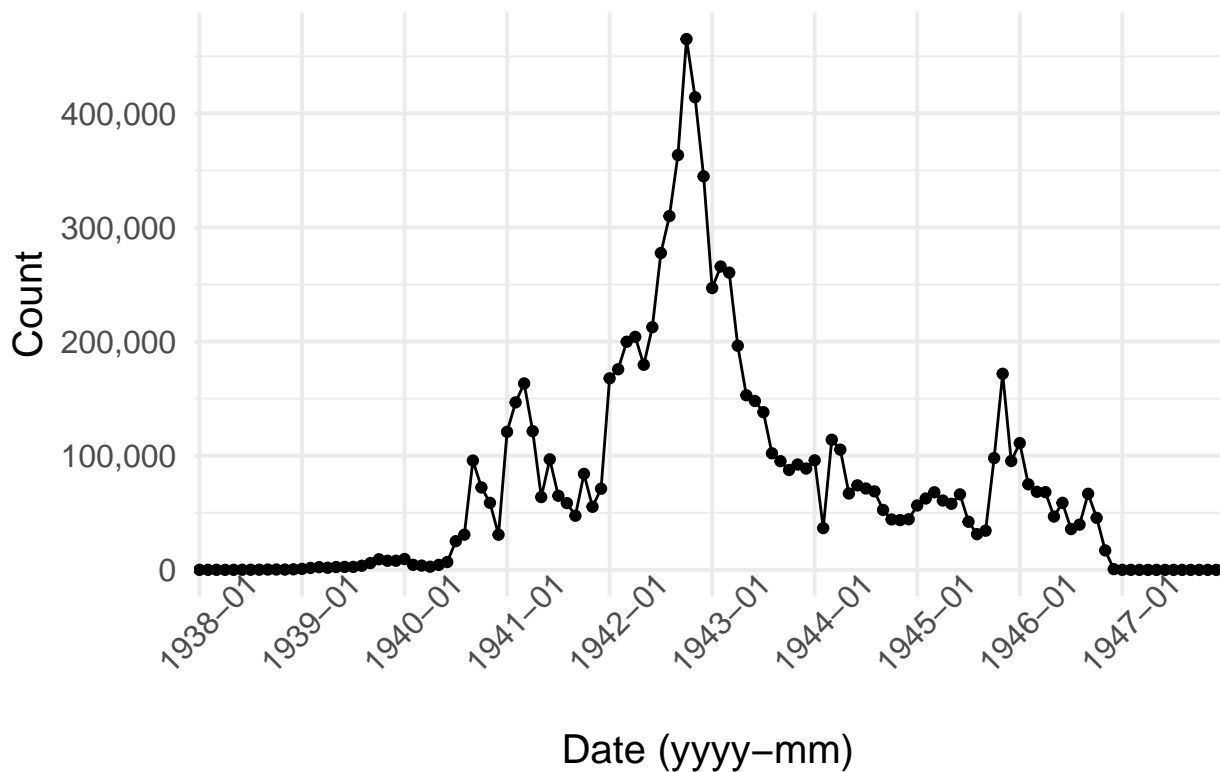
##	42 41	1095603
##	43 42	3320905
##	44 43	1878170
##	45 44	819906
##	46 45	847614
##	47 46	635303
##	48 47	156
##	49 48	82
##	50 49	1039
##	51 50	1059
##	52 51	82
##	53 52	472
##	54 53	1874
##	55 54	1850
##	56 55	2394
##	57 56	552
##	58 57	132
##	59 58	13
##	60 59	812
##	61 60	111
##	62 61	15
##	63 62	124
##	64 63	238
##	65 64	207
##	66 65	239
##	67 66	178
##	68 67	20
##	69 68	25
##	70 69	102
##	71 70	169
##	72 71	44
##	73 72	237
##	74 73	136
##	75 74	414
##	76 75	308
##	77 76	333
##	78 77	36
##	79 78	12
##	80 79	84
##	81 80	92
##	82 81	5
##	83 82	21
##	84 83	2
##	85 84	55
##	86 85	8
##	87 86	19
##	88 87	10
##	89 88	98
##	90 89	173
##	91 90	443
##	92 91	38
##	93 92	375
##	94 93	260
##	95 94	270

```
## 96 95 326
## 97 96 371
## 98 97 3
## 99 98 211
## 100 99 2063
## # ... with 1 more row

## [1] 202288

## # A tibble: 7,145 x 2
##   date_enlistment_recode    n
##   <date>                <int>
## 1 NA                    213886
## 2 1940-09-16             51922
## 3 1940-10-15             35730
## 4 1940-11-25             27318
## 5 1941-02-10             22139
## 6 1941-02-24             20796
## 7 1942-10-30             20649
## 8 1942-11-04             20177
## 9 1942-12-15             20004
## 10 1942-10-27            19986
## # ... with 7,135 more rows
```

Enlistment Date (Month and Year Only)



place__of__birth

Label: Place of Birth

Description: place__of__birth reports a person's place of birth at time of enlistment, 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=2079&dt=893&c_id=24982

residence__state

Label: State of residence

Description: residence__state is a 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=2057&dt=893&c__id=24998

residence__county

Label: County of residence

Description: residence__county is a numeric 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 numeric variable that reports where an person 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. For a complete list of codes, please see: https://usa.ipums.org/usa-action/variables/EDUC#codes_section

grade_code

Label: Army grade ((alphanumeric code)rank)

Description: grade_code is a string variable that records a person's grade (rank), as reported in the World War II Army Enlistment Records.

grade_code	label	n	freq %
1	Master Sergeant	26283	0.264%
2	First Sergeant	11421	0.115%
3	Technical Sergeant	24471	0.246%
4	Staff Sergeant	48140	0.484%
5	Sergeant	71338	0.717%
6	Corporal	89783	0.902%
7	Private First Class	258778	2.6%
8	Private	8215100	82.5%
9	Aviation Cadet	77836	0.782%
NA	Technician 3d Grade	377082	3.79%
NA	Technician 4th Grade	377082	3.79%
NA	Technician 5th Grade	377082	3.79%

```
## # A tibble: 12 x 4
##   grade_code label          n `freq %`
##   <int> <chr>          <int> <chr>
## 1         1 Master Sergeant 26283 0.264%
## 2         2 First Sergeant 11421 0.115%
## 3         3 Technical Sergeant 24471 0.246%
## 4         4 Staff Sergeant 48140 0.484%
## 5         5 Sergeant 71338 0.717%
## 6         6 Corporal 89783 0.902%
## 7         7 Private First Class 258778 2.6%
## 8         8 Private 8215100 82.5%
## 9         9 Aviation Cadet 77836 0.782%
## 10        NA Technician 3d Grade 377082 3.79%
## 11        NA Technician 4th Grade 377082 3.79%
## 12        NA Technician 5th Grade 377082 3.79%
```

Code list also is available at: https://aad.archives.gov/aad/popup-codelist.jsp?cl_id=2072&dt=893&c_id=24999

branch__code

Label: Army branch (alphanumeric code)

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	3976527	43.200000
1	General Staff Corps	1699	0.018500
2	No branch assignment	2829637	30.800000
3	Philippine Scouts	27402	0.298000
4	Professor US Military Academy	2537	0.027600
5	Military Intelligence	7835	0.085200
6	Corps of Military Police	19502	0.212000
7	Specialist Reserve	101	0.001100
8	Detached Enlisted Men's List	13125	0.143000
9	Inactive Reserve	56127	0.610000
10	Infantry	231058	2.510000
18	Women's Army Corps	71592	0.778000
20	Air Corps	955300	10.400000
30	Field Artillery	109267	1.190000
38	Transportation Corps	20736	0.225000
40	Coast Artillery Corps or Army Mine Planter Service	83355	0.906000
50	Quartermaster Corps	90281	0.981000
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	55386	0.602000
71	Medical Corps - For Officers only	74	0.000804
72	Hospital Dietitians	829	0.009010
73	Physical Therapy Aides	112	0.001220
74	Dental Corps - For Officers only	198	0.002150
75	Veterinary Corps or Service - For Officers of the Veterinary Corps and for Enlisted Men of the Veterinary Service	873	0.009490
77	Medical Administrative Corps - For Officers only	669	0.007270
78	Adjutant General's Department	1312	0.014300
79	Pharmacy Corps	267	0.002900
80	Chemical Warfare Service	6094	0.066200
82	Corps of Engineers	46725	0.508000
87	Ordnance Department	57965	0.630000
90	Signal Corps	95326	1.040000
93	Cavalry	28088	0.305000
96	Chaplains	1612	0.017500
97	Finance Department	2439	0.026500
98	Inspector General's Department	613	0.006660
99	Judge Advocate General's Department	11907	0.129000
NA	NA	393662	4.280000

term_of_enlistment

Label: Term of enlistment

Description: term_of_enlistment_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.

term_of_enlistment	label	n	freq %
1	One year enlistment	217859	2.370
2	Enlistment for Alaska	28447	0.309
3	Enlistment for Hawaiian Department	473506	5.150
4	Enlistment for assignment to another corps area	23166	0.252
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	6281919	68.300
6	Enlistment for the Panama Canal Department	339363	3.690
7	Enlistment for the Philippine Department	120956	1.310
8	Enlistment for the Puerto Rican Department	9950	0.108
9	Three year enlistment	100269	1.090
NA	NA	1604797	17.400

Codes also available at: https://aad.archives.gov/aad/popup-codelist.jsp?cl_id=2076&dt=893&c_id=24979

race

Label: Race

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

Codes also available at: https://aad.archives.gov/aad/popup-codelist.jsp?cl_id=2059&dt=893&c_id=24984

citizenship

Label: Citizenship

Description: citizenship is a string variable that reports an individual's citizenship, as reported in the World War II Army Enlistment Records. 1 indicates the individual was a US citizen, 0 indicates a noncitizen.

Codes also available at: https://aad.archives.gov/aad/popup-codelist.jsp?cl_id=2059&dt=893&c_id=24984

civillian__occupation

Label: civilian occupation

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

marital_status	label	n	freq %
1	Single, with dependents	2048974	22.300
3	Separated, with dependents	157700	1.710
4	Divorced, with dependents	215371	2.340
5	Widower or widow, with dependents	68436	0.744
6	Single, without dependents	6471628	70.300
NA	NA	238123	2.590

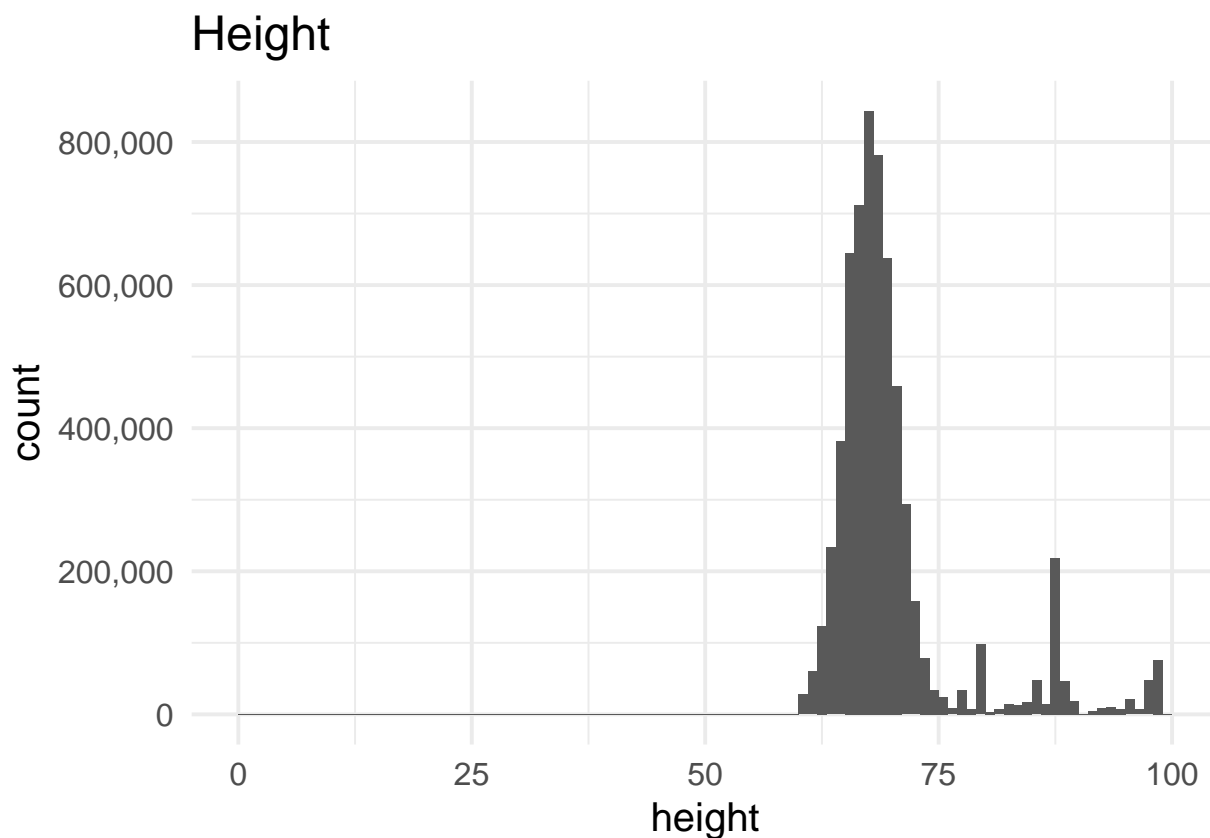
Codes also available at: https://aad.archives.gov/aad/popup-codelist.jsp?cl_id=2081&dt=893&c_id=25001

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.

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. Height may have been recorded through about 1943, but height data recorded after February 1943 are likely unreliable. See the [National Archives' technical documentation on the Army Serial Number Electronic File](#) for further information.

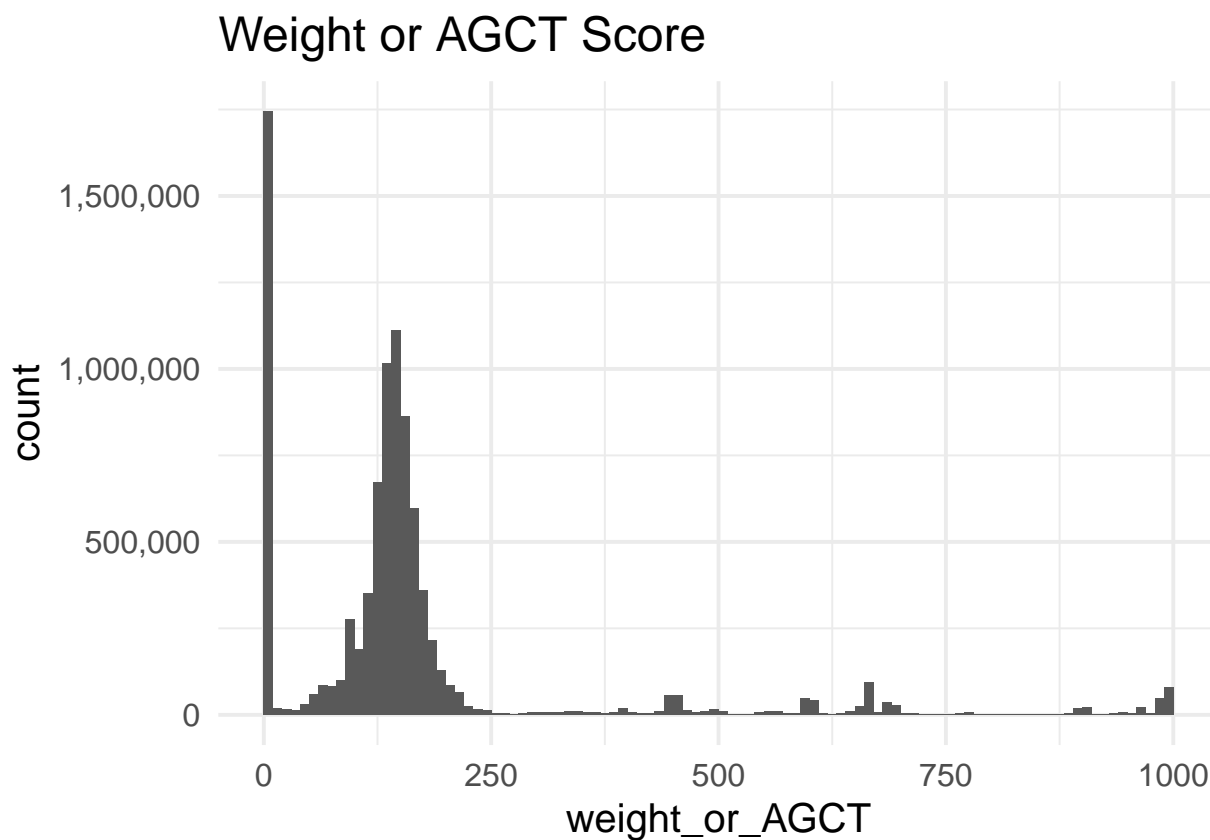


weight_or_AGCT_enlistment

Label: Weight (pounds) at enlistment or AGCT score

Description: weight_or_AGCT_enlistment 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.

Note: Please be aware that instructions on how to use this field on enlistment punch cards varied over time. This field was explicitly used for both weight and AGCT scores, but may also encode military occupation or other information. Weight may have been recorded through about 1943 at some enlistment sites, but the recording of AGCT score instead of weight at some enlistment sites likely began in March 1943. Refer to the [National Archives' technical documentation on the Army Serial Number Electronic File](#) for further information.



component__enlistment

Label: Army Component

Description: component__enlistment 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.

component	label	n	freq %
1	Regular Army (including Officers, Nurses, Warrant Officers, and Enlisted Men)	1294899	14.1000
2	Regular Army Reserve (Enlisted Men Only)	2901	0.0315
3	Reserves - exclusive of Regular Army Reserve and Officers of the Officers Reserve Corps on active duty under the Thomason Act (Officers and Enlisted Men – O.R.C. and E.R.C., and Nurses-Reserve Status)	511487	5.5600
4	National Guard (Officers, Warrant Officers, and Enlisted Men)	251739	2.7400
5	Philippine Scouts (Officers and Enlisted Men)	15377	0.1670
6	Army of the United States - includes the following: Voluntary enlistments effective December 8, 1941 and thereafter; One year enlistments of National Guardsman whose State enlistment expires while in the Federal Service; Officers appointed in the Army of the United States under Army Regulations 605-10	679919	7.3900
7	Selectees (Enlisted Men)	6063465	65.9000
9	Women's Army Corps	140580	1.5300
NA	NA	239865	2.6100

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

source

Label: Source of army personnel

Description: source is a string variable that reports the source of the enlisted person, as reported in the World War II Army Enlistment Records.

Codes also available at: https://aad.archives.gov/aad/popup-codelist.jsp?cl_id=2077&dt=893&c_id=24981