

Table of Contents

[The Ancestry.com Online Census Images Data Base](#)

[Ancestry.com Indexes Bugs & Features](#)

[My Rules for Unambiguous Citation to the USCensus Page](#)

[The USCensus Page/Folio MetaNumber](#)

[The “Census Day” and the “Census Year”](#)

[My USCensus Abstracting Conventions](#)

[...1790 – 1840 USCensuses](#)

[...1850 – 1930 USCensuses](#)

[...The Abstract Header](#)

[...The Abstract Person-Lines](#)

[...Abbreviations](#)

[...Abbreviations: Property Ownership](#)

[...Abbreviations: Education and Occupational Status](#)

[...Abbreviations: Vital Statistics \(birth and marriage\)](#)

[...Abbreviations: Health/Disabilities](#)

[...Abbreviations: Origins, Citizenship, Veteran Status](#)

[Example Abstracts](#)

[Deriving Birth Dates from Ages in the USCensus](#)

The Ancestry.com Online Census Images Data Base

My USCensus citations are generally to the Ancestry.com Online Images USCensus Database, a digitized facsimile of the pages of the NARA USCensus film series. Where the actual source is instead a NARA or an FHL film, the film# will be cited instead of the Ancestry.com image#.

Although the Ancestry.com database is technically one step further removed from the original documents filmed by NARA, it is in every way a “better” source than the NARA films. The images have been digitally cleaned up and enhanced, their viewing is not subject to the usual dirt and scratches interposed by the microfilm reader, and filming defects in the NARA series can usually be overcome by blowing up the Ancestry.com images.

What’s more, the Ancestry.com data are better (though still imperfectly) organized and titled than the NARA film series, they are newly indexed on every name for all years except 1910, and these indexes are searchable not only on name (by Soundex or wildcard), but on any combination of name, census location, birthplace of subject, etc. And unlike the NARA films, where the film frame#s are not usually apparent, the Ancestry.com database can be cited to the individual image#, eliminating all backing and filling, and overcoming in this way, the frequent ambiguity of the original census numbering.

The only weakness of the Ancestry database is that the subdivision and titling of the local jurisdictional records clusters, or *enumeration sections*, has not been carried out consistently, probably because of a misplaced inclination to copy the structure of the even more inconsistently organized and often mistitled NARA films. However, my citations obviate these weaknesses by including all place name fields found on the original document (as well as the less-useful census administrative fields, which I relegate to parentheses). From these fields, the exact enumeration section can be determined in nearly every case, whether the source be the NARA film or the Ancestry DB, and if the latter, the image# I cite completes the non-ambiguous reference to a specific page.

Of course it is possible that MyFamily.com, the company behind Ancestry, will someday make the mistake of “perfecting” the organization of its data, which may in some cases, render my cited image#s incorrect. For that reason, and to ensure that my citations continue to point back to the original NARA documents, I have made an extensive study of the numbering schemes applied to the originals, and have developed a fully-defined, and nearly comprehensive, set of rules for identifying USCensus *metanumbers*, and citing unambiguously to them, which is more than any of the scholarly “authorities” in this field have done.

Ancestry.com Indexes–Bugs & Features

The Ancestry.com indexes for 1850-1930 now index every name—all except 1910 that indexes only the first person of each surname in the household. The fields that can be searched on, besides name and place of residence vary widely between indexes.

Soundex searches of surname, and wildcard searches of either given or surnames is possible, with some limitations. The chief of these is that a search for names beginning with certain text requires a minimum prefix of three characters; thus one can search on all names beginning with “Pri*”, but not all names beginning with “Pr*”. Ancestry.com theoretically also implements the single-character-substitution wildcard character (“?”), but I have not been able to get this to work. Some indexes also provide certain “fuzzy” as opposed to exact, and with it ranked matching, but the rules for this are inadequately defined, and I haven’t had time yet to try to deduce them by trial and error experiment.

The Soundex implementation of the Ancestry.com indexes from 1850-1900 & 1920-1930 (and probably those from 1790-1840) erroneously index names like “Mc Mahon” as MC_, rather than MCMN. The 1910 index avoids this problem by concatenating these two name parts *before* indexing them. The Soundex rules say to concatenate multi-part names as part of the indexing process, while

preserving the spaces in the name itself. The result of this error is that a search for MCMN (which brings up mostly non-McMAHONs to begin with) is still not comprehensive: it must be supplemented with any number of variant exact name searches of the form “Mc Mahon” etc. etc. Since there are so many of these, the purpose of Soundex searching (to bring up all or most sound alike names, at the cost of many “false positives”) is largely defeated.

Names like “St John” are supposed to be (1) expanded to “Saint”, then (2) concatenated into a single string. Ancestry.com’s indexing does neither of these, so searching for “St”, “Ste”, “Saint” names is a real nightmare.

Given names are not Soundexed, and in fact, only the first name is indexed at all, although if multiple initials/names are used as argument, all names that match on any of the parts are brought up. One good feature though, is that standard abbreviations of first names are first expanded, eg. “Wm” to “William”, and also the old-fashioned “Jno” to “John”. This is true for all of the years from 1850-1930 (I haven’t checked for earlier years).

My Rules for Unambiguous Citation to the USCensus Page

My primary page# citation is to the Ancestry.com image#, which designates a particular image within the sequence of pages of a geographic or jurisdictional records cluster—which I shall term, for purposes of this discussion, an *enumeration section* or just *section*.

My secondary page# citation is to the document page or folio *metanumber*, which will be more fully explicated below.

Since, to provide the necessary interpretive context, my abstracts are generally of whole households, the citation itself is usually to the complete household, even though only one of its members is of present interest; for example, “161A(05-10)” tells us that the household occupies lines 5-10 on (metanumbered) page 161A. A household that spills over onto the next page might be designated “161A(36-40)-161B(01)”. The only exception to this rule is when an individual member is a boarder or inmate in a largish institution, like a hotel or hospital; in this case only the individual’s line# will be cited, and the entry as a whole will be identified as, eg. “John Robb, boarder (Alma Murphy boarding house)”.

I do not bother to report the uninformative “dwelling#” of later censuses, nor even the corresponding, rather arbitrary (and usually ignored) “family#”, though in those exceptional cases where the enumerator actually attempted to differentiate multiple families within a household, that information will be noted in my abstracts, by the separator line “**–next family–**”.

The USCensus Page/Folio MetaNumber

When the reports of all the census takers were gathered and collated, the original page numbering was superceded by an assigned metanumber series spanning multiple enumeration sections. Usually, the metanumber is the one stamped (or, before 1830, written and often underlined), on every other page. The 1830 & 1840 censuses were gathered on pre-printed forms in *folio* format, and for these years, the metanumber is a folio number. For the other years, I have refined the metanumber into a page number by appending a suffix (from 1880 on, the metanumber suffix is actually pre-printed on the census form).

In most cases, the first page of a section is a right-side page of the bound book, and bears the first (usually stamped) metanumber of the section. This number is, in effect, a *leaf* number since there is one unique number per two-sided leaf; I refine this leaf number into a true *page* number by appending “A” to indicate the front side of the leaf, and “B” the back side. For 1880 only, since the pages are pre-printed successively “A”, “B”, “C”, “D”, then “A” and so forth, I simply use the pre-printed number as the suffix. For 1830 and 1840 the pre-printed two-page-spread forms were actually folio numbered, and are cited accordingly.

In more uncommon cases (usually confined to 1820 and before), the first page of a section is an unnumbered left-side page of the original bound book, with the metanumber appearing on the right-side page. This is, in effect, a folio numbering scheme applied to page-formatted material. I distinguish this folio numbering variant from the normal leaf numbering scheme by suffixing its pages “L” = left-side, or “R” = right-side.

Finally, some of the earliest filmed images were of loose leaves joined together either left-to-right, or top-to-bottom, and these latter I suffix “T” = top or “B” = bottom. in my citations. There are a few numbering schemes that fall outside these categories (too rare to be worth specifying), which I have been able to handle by extended analogy with the above rules. There are even (more rarely) sections with no clearly defined metanumbers at all; in some cases I have interpolated numbers, and in others left them blank.

The “Census Day” and the “Census Year”

Although the census data might be gathered on any day over a span of months (and I show this actual enumeration date in my abstract), for each decennial USCensus an official Census Day was fixed, and a corresponding Census Year defined as the 12 month period preceding the Census Day. Irrespective of when the enumerator’s interview took place, and the date he gathered set down, all data was supposed to be “as of” the official Census Day. Thus if the Census Day was 1Jun, a new baby born 15Jun, and the interview on 21Jun, the child was supposed to be disregarded; similarly ages, # years married etc. etc. were supposed to be strictly “as of” the Census Day and Year.

I have abstracted the actual date of gathering because this formal rule was often ignored or misinterpreted. It might be difficult for an enumerator not to make note of a squalling baby in the arms of its mother, the informant for the household, and since many informants had only a tenuous familiarity with the English language, or otherwise lacked the capacity to grasp such nice formal distinctions as an “as of” date, there are many deviations from these rules, and the probability that there may have been such deviations can best be judged from the overall context, of which the actual date of gathering is an element.

The Census Day varied over the course of time:

For...	the “census day” was...
1790–1820	1st Monday in August
1790	2Aug
1800	4Aug
1810	6Aug
1820	7Aug
1830–1900	1Jun
1910	15Apr
1920	1Jan
1930–1980	1Apr

My USCensus Abstracting Conventions

Names of persons are rendered literally, though if abbreviated in the original, they will be expanded, within square brackets, where there is no doubt from the abbreviation itself that name is meant. I shall occasionally, in my editorial notes, refer to surnames generically (using a single standardized spelling); in such cases the generic surname will be rendered in all-caps.

Given names, when referred to in editorial notes with relation to earlier generations, will be qualified with superscripts, following the usual convention: the first immigrant ancestor is generation 1, his/her children are generation 2 (even when they too are immigrants), their children are generation 3, etc.; meanwhile, the first immigrant’s parents (who did not emigrate) are generation A, their parents are generation B, etc.

Place names will generally be normalized to their correct form and spelling, and standard abbreviations will be used for US States, and for foreign countries, and in some cases foreign provinces and principalities.

Dates are normalized to standard genealogical date format, eg. “23Jul1870”.

Transcribed text appears in normal type, with italics reserved for the editorial voice. Text that can’t actually be read, but which can be reasonably inferred, will appear in normal type within square brackets, along with any italicized editorial commentary.

Normal type within square brackets thus represents an editorial filling out of what is actually readable in the original—by expansion of an abbreviation, or by interpolation of missing or illegible characters or even a whole word, which can be inferred with confidence from context, or perhaps from an associated or companion document. Where such inference is problematic, the bracketed text will be supplemented by an editorial “?”, or where alternate readings are possible, they will be separated by editorial “|”(OR sign)s. Editorial notation and commentary within the abstract itself is confined to italicized text within square brackets, with the exception of certain editorial symbols that may be interpolated into the text of the abstract in *unbracketed normal type*:

- asterisks (“*”) are substituted for illegible characters in the original, one for each
- dashes (“---”) represents a normally required field left blank in the original
- ellipsis dots (“...”) are used in the normal way to indicate skipping of part of a data field, or of one or more whole person-lines of a household.

In addition, each abstract may be followed by a brief italicized, but *unbracketed*, editorial note.

Interlineations will be silently merged into the text at the intended point, but overwritten or struck-out text will generally be discussed in the editorial note.

1790 – 1840 USCensuses

For these censuses in which only the head of household is named, I generally use a tabular format with columnar headings to align the number of household inhabitants in each sex/age category in a single column. For example, here’s my 1830 USCensus abstract template:

1830 USCensus	Free Whites													Sla Fre		
	M: 0	5	10	15	20	30	40	50	60	70	80	90	100	ves	Col	TOT
	F: 4	-9	-14	-19	-29	-39	-49	-59	-69	-79	-89	-99	+			
*****	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ii+ccccSS-tttt:fppp-LL	0	0	0	0	0	0	0	0	0	0	0	0	0			

The first row represents the males, the second the females. The asterisks (*****) are replaced by the name; the data below it on the 2nd line is the citation information— *ii* is the Ancestry image#; *ccccSS* is the county-state abbreviation; *tttt* the township (if any); *fppp* the folio/page#; and *LL* the line#.

1850 – 1930 USCensuses

With the great amount of information for these modern, every-name, USCensuses, I have resorted to many abbreviations and conventionalizations, although where odd spellings or abbreviations in the original text leave room for doubt about what is meant, this will be noted either in the abstract itself, or in its supplementary editorial note. The remainder of this document is concerned with defining these conventions.

The Abstract Header

The date supplied for the household entry is the full date on which the cited data was actually solicited by the enumerator, eg. "23Jul1870", *and not the official "as of" date for that Census Year.*

The header place names appear in the format: <county name>"Co"<state ZIP code abbr>," <town(ship)> which may be followed by "PO " <rural post office>, or by "Wd" <ward#>, <street address> when the household lies in a town or city.

The census administrative fields, like Enumeration District, as well as any pagination local to the enumeration section, appear in a standard format within parentheses, eg. "(SD5,ED23,s13A)", followed by the unique citeable page metanumber, and the parenthesized range of lines covered by the household, eg. "page 72A(42-47)".

The Abstract Person-Lines

Names are abstracted in the format <surname>"," <given name(s)> regardless of the order encountered in the original, although the actual text of the name is rendered literally; dittoed surnames are always represented by ", the standard ditto mark. I have adopted these conventions to facilitate eyeball scanning of whole pages of abstracts.

Age is in years (as found), unless it is given in months (eg. 6/12), when it will be suffixed, as, eg. "6mon".

Sex is indicated only when not determinate from name or relationship, by prefixing Age with "[M]" or "[F]".

Race is indicated only where it is other than white by prefixing Age with "[b1]" for "black", or "[mu]" for "mulatto".

Married status is abbreviated to "sing", "marr", "wid", "div".

Relationships are abbreviated to: "son", "grson", "dau", "husb", "wife", "fath", "moth", "mothIL", "neph", "niece", "stpson", "brdr"=boarder, "inmt"=inmate, "srvt"=servant etc..

Occupation is generally abstracted literally, or, if conventionalized, will be encased in angle brackets (eg. "works in mine" or "<miner>"; "laborer", "day labor", or "<lab>"). The default woman's occupation will generally be conventionalized, ie. "housekeeper", "keeping house" and the like to "<KH>".

Birthplace (of self, or of parents) will be standardized to my usual abbreviations: 2char state ZIP codes, 3char country codes, or 4char province/independent principality codes, eg. "PA", "Ire", or "Hano" for the German state of Hanover. Tables are available for the codes I use, but most are created by simple truncation of the place name word, or by abbreviating with the first consonants of each of several words.

All originally dittoed fields besides name are dittoed in the abstract by editorial construction of the dittoed text, in square brackets, eg. if the head of household is born in "Minn", and his wife's birthplace field is dittoed, the abstracted representations would be "MN", and "[MN]", respectively. Where there is no overt dittoing in the original (ie. the field is left blank) no dittoing will be extrapolated, ie. the field will be left blank in my abstract or marked with a "-" to indicate that an expected field is missing.

Abbreviations

Values of the many other fields, which vary from census to census, are abstracted in abbreviated form, viz.:

Property Ownership

\$nnn	= Real Property (1850 USCensus only)
nnn/nnn	= Real Property/Personal Property (1860-1870)
ownsHF	= Owns Home Free (of mortgage) (1900-1920)
ownsHM	= Owns Home with Mortgage (1900-1920)
ownsFF	= Owns Farm Free (of mortgage) (1900-1910)
ownsFM	= Owns Farm with Mortgage (1900-1910)
rents	= rents (1900-1920)
ownsH\$nnn	= Owns Farm of value \$nnn (1930)
ownsH\$nnn	= Owns Home of value \$nnn (1930)
rents\$nn	= Rents @ \$nn/month (1930)
radio	= owns Radio (1930)

Education and Occupational Status

AS	= attended school within the year (1850-1930)
xRW21	= can't Read or Write, and aged 21 or over (1850-1860)
xR,xW	= can't Read, and/or can't Write (any age) (1870-1880)
R,W	= can Read, and/or can Write (1900-1930)

Emp	= employed (1910,1930)
UnEmp	= unemployed (1910,1930)
UnEmp#m	= unemployed # months (1880,1900)
UnEmp#w	= unemployed # weeks (1910)

Vital Statistics (birth and marriage)

b<Mmm>	= "If born within the [Census] year, state/give the month" (1870,1880)
b<MmmYYYY>	= "Date of Birth" (1900)
mCY	= "Married within/during the [Census] year" (1850,1860,1880)
m<Mmm>	= "If married within the [Census] year, state the month" (1870)

m<#>y	= "Number of years of present marriage" (1900,1910)
m@<#>	= "Age at first marriage [in years]" (1930)

Sometimes, in 1900-1930, the ordinal number of the marriage will be indicated, eg. "m1"; in such cases, this number will be prepended, eg. "2m3y" means 2nd marriage, 3 yrs duration.

Health/Disabilities

Dis: <typ> = Disabled "-" specific literal text in Disabled box (1850-1880)
 (for 1850-1870, there's a single box, headed with the categories: deaf, dumb, blind, insane, idiotic, pauper, convict; usually one of these words will be written in the box by the census taker, where the condition applies. In 1880, check-columns are provided for blind, deaf & dumb, idiotic, insane, and a

catchall column headed “Maimed, crippled, bedridden, or otherwise disabled”. In 1900-1930, no attempt is made to gather such information.)

Origins, Citizenship, Veteran Status

FathFor = father foreign born (1870)

MothFor = father foreign born (1870)

ParFor = parents foreign born (if both foreign born)

Aln = alien (1910-1930)

xmCit21 = age 21 or over and NOT a male citizen (1870)

I've inverted this indicator so that I can normally leave this field blank.

mCit21xV = male citizen aged 21 & over where right to vote denied xcept for rebellion/crime (1870)

ImmYYYY = immigrated in year indicated (1900-1930)

Nat = naturalized (1900-1930)

NatYYYY = naturalized in year indicated (1920)

#YinUS = # years in the US (1900)—abstracted only when imm yr. is doubtfully legible or contradictory

Vet-Union = veteran (“survivor”) of the Union Army (1910)

Vet-Confed = veteran (“survivor”) of the Confederate Army (1910)

Vet-< war > = veteran of the indicated war (1930)

Examples of my USCensus Abstract Format

John Orth household, 11Jul1870 USCensus, HennepinCoMN, St Anthony, wd1
(p113), page 799(01-10), Ancestry.com image 1.

	Aged	Occupation	Property Real/Pers	Born	Miscellaneous
Orth, John	49	brewer	20000/5000	Ger	ParFor, mCit21
" , Mary	39	<KH>		Ger	ParFor,
" , John	20	clerk in brewery		MN	ParFor
" , Sophia	16			MN	ParFor
" , Edward	14			MN	ParFor
" , Virginia	12			MN	ParFor
" , Alfred	10			MN	ParFor
Wernel, Philip	81	at home		Ger	ParFor, mCit21
" , Catherine	71	at home		Ger	ParFor
---new family---					
Gilman, Fanny	78	keeps house		NY	

Although, according to my conventions, I could have omitted "mCit21" for the two foreign born males here, I've chosen to specify their citizenship status precisely because they were foreign born.

James A Robb household, 29Jun1880 USCensus, CrawfordCoWI, Utica
(SD2,ED54,p29), page 161A(09-16), Ancestry.com image 25.

	Aged	Rel2	Marr	Occupation	Self	Fath	Moth	Miscellaneous
Robb, James A	46	[head]	marr	miller	OH	PA	Ire	
" , Fanny G	37	wife	marr		WI	VA	OH	
" , John G	17	son		works on farm	WI	OH	WI	AS
" , Sadie J	15	dau			WI	OH	WI	AS
" , Carrie L	12	dau			WI	OH	WI	AS
" , Lora P	1	dau			WI	OH	WI	
Robb, Joel	76	fath	wid		PA	PA	PA	
Gay, Sarah	73	moth-il	wid		OH	VA	VA	

Henry Conklin household, 1Jun1880 USCensus, HardinCoOH, Mt Victory
(SD2,ED99,p3), page 85C(30-33), Ancestry.com image 3.

	Aged	Rel2	Marr	Occupation	Self	Fath	Moth	Misc
Conklin, Henry	23	Head	marr	laborer	IA	[OH]	[OH]	
" , Ella	23	wife	marr	<KH>	OH	MS	OH	
" , Charles	4mon	son	---		[OH]	IA	[OH]	b.JAN
Pever, Sarah	62	mothIL	wid		MD	OH	[OH]	

Charles G Fuller household, 23Jun1900 USCensus, CookCoIL, Evanston, wd1, 1305 Forest Ave
(SD1,ED1155,s30A-30B), pages 35A(49-50)-35B(51-55), Ancestry.com images 59-60.

	Aged	Born	Rel2	Marr	Chldrn Tot/Lv	Occupation	Self	Fath	Moth	Misc
Fuller, Charles G	44	Apr1856	head	m15y		occulist & **ist	NY	VT	VT	ownsSHF
" Isabella	38	Jul1861	wife	m15y	2 2		NY	NY	NY	
---page 35B---										
Fuller, Dorothy	13	Aug1886	dau	sing		at school	IL	NY	NY	
" , Elizabeth	6	Nov1893	dau	sing		at school	IL	NY	NY	
White, Lucy E	28	Sep1871	brdr	sing		school teacher	NY	NY	NY	unemp4m
Johnson, Ella	24	Sep1875	srvt	sing		domestic	Swe	Swe	Swe	imm1894
Olson, Mary	28	Jan1872	srvt	sing		domestic	Nor	Nor	Nor	imm1891?

Deriving Birth Dates from Ages in the USCensus

For a couple of reasons, ages in the USCensus tend to be off by a year, or sometimes two, either way; thus birth years calculated from such ages should always be qualified by “abt”.

For one thing, the age recorded was supposed to be as of a particular Census Day, which, for most of the 19th century was June 1st. However, the census enumerator might actually collect the data anywhere from June to December of the year, and if he were to ask an informant in August about the age of a person who had had a birthday in July, but failed to make it clear that he wanted the latter’s age as of the 1st of June, he might well get his current age, especially if the informant was hard of hearing, or aged, or a person for whom English was his second language.

As another contributor to inaccuracy, if you were to ask a person (even today) his age just a few days, weeks, or even months before his birth day, and if you didn’t stress that it was important to be strictly accurate, he might well tell you the age that he was about to be—the age he was closest to.

The accuracy of answers to such questions as these depended both on the enumerator and the informant, and on the degree of rapport and clarity of communication between them.

There are certain general rules, or guidelines, that may be applied in resolving conflicts in estimating the year of birth from age in the census:

(1) Ages of children or young adults within their natal family groupings are generally the most accurate, and the younger the children the better. This is both because the ages reported are closest to being contemporary with the event, and because in most case the parents, who have first hand knowledge of the births of their children, are the informants.

(2) A closely related reason is that the variability of the ages of children listed in order of age within family groupings is constrained by the ages of their siblings. And this constraint is reinforced when the same family (with at least some of the same children) appears in successive censuses. Where (as in most cases) the children come along about every two years, overlaps with siblings preclude more than a year or so of inaccuracy, and ideally all the children should be followed into adulthood in later censuses to establish as many points of age reference as possible.

(3) Some census takers were obviously sloppy and lazy, and others were clearly careful and conscientious. Examining the surrounding records on the page or adjacent pages of a particular census section can provide a pretty good idea about the reliability of the enumerator. Sometimes, for example, one finds an oldest child with about the right age, but then all the younger ones have ages exactly 2 years less than the preceding one, suggesting that the census taker, or the informants, balked at providing, or recording the specific details on the family.

(4) Some subjects (or informants) were given to fudging their ages for various reasons, and multiple censuses can turn up patterns that suggest this.

(5) The 1900 USCensus is usually the most accurate because subjects were asked to provide the month as well as the year of their births; the enumerator then calculated their ages from their birth year and month, thus resolving the ambiguities that can arise when the age question is asked at various times after the official Census Day. Also, this additional probing for details about the subject’s birthday tended to encourage people to make their answers more accurate. Errors might still be made in the year because the subject might only know off the top of his head how old he was, not the exact year he was born, and he might make a mistake in calculating this. If it seems farfetched that a person wouldn’t be quite sure of his birth year, consider that in the pre-modern period people were almost never asked such questions, or required to fill out forms, like we moderns have to waste so much of our time doing. However, while there thus remains some potential for getting the birth year wrong in the 1900 census, people generally did know their birth days, and the informant of the household typically did too, so the month of birth can reasonably be presumed to be accurate.