Accountants and Auditors from IPUMS

Spencer Graves

9/1/2018

Accountants and auditors from IPUMS

I created an account wit IPUMS and selected and requested historical data on the number of accountants and auditors in the US labor force dating back to the nineteenth century. After a while I got an email saying my extract was ready for download.

However, the obvious thing to download produced a *.dat file, but the obvious instructions for reading it required a* .xml file.

From “help(pac=ipumsr)” I found that the packages included six vignettes. One of those is titled “Introduction to ipumsr - IPUMS Data in R” (https://cran.r-project.org/web/packages/ipumsr/vignettes/ipums.html). From this, I learned that I needed to right click (ctrl-click on a Mac) on “DDI” under “Codebook” and then select and then select “Save link as…”. Moreover, I should NOT do this in Safari. Google Chrome worked for me for this on 2018-09-01.

Following this process, I downloaded “usa_00001.dat” with 4.91 GB of data plus “usa_00001.xml” being a codebook file of size 67 KB.

The “Command File” for R recommends using the following:

```
readAndCompute <- TRUE
# "usa_0001.dat" is huge.
# It takes a long time to read
# (3.63 seconds on a reasonably fast notebook on 2018-09-02)
# and shorter but still long times with other operations
# on "data" (roughly 30 seconds with each computation).
# Therefore, I'm wrapping each computation in a condition,
# so it will only be run if I actually want it.
if(readAndCompute){
#***IF readAndCompute
#***change setwd as needed to
#***the directory containing 'usa_00001.xml' and .dat
  setwd("../../../..")
  library(ipumsr)
  start.time <- Sys.time()
  ddi <- read_ipums_ddi("usa_00001.xml")
  data <- read_ipums_micro(ddi)
  (et <- Sys.time() - start.time)
}
```

```
## Use of data from IPUMS-USA is subject to conditions including that users should
## cite the data appropriately. Use command `ipums_conditions()` for more details.
```

```
## Time difference of 1.984066 mins
```

I timed this, because the first time it seemed to take a long while. Obviously, I extracted a lot moredata than I need. But conveniently, when I'm running this manually, it displays both percent completion and number of MB read so far: 3.63 minutes.

The "data" is an object with a huge number of rows and 8 colums:

```
if(readAndCompute){
  str(data)
  nrow(data)/1e6
}
```

```
## Classes 'tbl_df', 'tbl' and 'data.frame':    114278279 obs. of  8 variables:
##  $ YEAR   : int  1850 1850 1850 1850 1850 1850 1850 1850 1850 1850 ...
##   ..- attr(*, "label")= chr "Census year"
##   ..- attr(*, "var_desc")= chr "YEAR reports the four-digit year when the household was enumerated or included in the census, the ACS, and the "| __truncated__
##  $ DATANUM: num  1 1 1 1 1 1 1 1 1 1 ...
##   ..- attr(*, "label")= chr "Data set number"
##   ..- attr(*, "var_desc")= chr "DATANUM identifies the particular sample from which the case is drawn in a given year. For most censuses, the I"| __truncated__
##  $ SERIAL : num  101 101 101 101 101 101 101 101 101 201 ...
##   ..- attr(*, "label")= chr "Household serial number"
##   ..- attr(*, "var_desc")= chr "SERIAL is an identifying number unique to each household record in a given sample. All person records are assig"| __truncated__
##  $ HHWT   : num  97 97 97 97 97 97 97 97 97 97 ...
##   ..- attr(*, "label")= chr "Household weight"
##   ..- attr(*, "var_desc")= chr "HHWT indicates how many households in the U.S. population are represented by a given household in an IPUMS samp"| __truncated__
##  $ GQ     : 'labelled' int  1 1 1 1 1 1 1 1 1 1 ...
##   ..- attr(*, "label")= chr "Group quarters status"
##   ..- attr(*, "var_desc")= chr "GQ classifies all housing units as falling into one of three main categories: households, group quarters, or va"| __truncated__
##   ..- attr(*, "labels")= Named num  0 1 2 3 4 5 6
##   .. ..- attr(*, "names")= chr  "Vacant unit" "Households under 1970 definition" "Additional households under 1990 definition" "Group quarters--Institutions" ...
##  $ PERNUM : num  1 2 3 4 5 6 7 8 9 1 ...
##   ..- attr(*, "label")= chr "Person number in sample unit"
##   ..- attr(*, "var_desc")= chr "PERNUM numbers all persons within each household consecutively in the order in which they appear on the origina"| __truncated__
##  $ PERWT  : num  97 97 97 97 97 97 97 97 97 97 ...
```

```
##   ..- attr(*, "label")= chr "Person weight"
##   ..- attr(*, "var_desc")= chr "PERWT indicates how many persons in the U.S. popul
ation are represented by a given person in an IPUMS sample. \"| __truncated__
##  $ OCC1950: 'labelled' int  100 999 999 999 999 999 999 999 999 690 ...
##   ..- attr(*, "label")= chr "Occupation, 1950 basis"
##   ..- attr(*, "var_desc")= chr "Universe Note: \"New Workers\" are persons seeking
employment for the first time, who had not yet secured their"| __truncated__
##   ..- attr(*, "labels")= Named num  0 1 2 3 4 5 6 7 8 9 ...
##   .. ..- attr(*, "names")= chr  "Accountants and auditors" "Actors and actresses"
"Airplane pilots and navigators" "Architects" ...
##  - attr(*, "spec")=List of 2
##   ..$ cols   :List of 8
##   .. ..$ YEAR   : list()
##   .. .. ..- attr(*, "class")= chr  "collector_integer" "collector"
##   .. ..$ DATANUM: list()
##   .. .. ..- attr(*, "class")= chr  "collector_double" "collector"
##   .. ..$ SERIAL : list()
##   .. .. ..- attr(*, "class")= chr  "collector_double" "collector"
##   .. ..$ HHWT   : list()
##   .. .. ..- attr(*, "class")= chr  "collector_double" "collector"
##   .. ..$ GQ     : list()
##   .. .. ..- attr(*, "class")= chr  "collector_integer" "collector"
##   .. ..$ PERNUM : list()
##   .. .. ..- attr(*, "class")= chr  "collector_double" "collector"
##   .. ..$ PERWT  : list()
##   .. .. ..- attr(*, "class")= chr  "collector_double" "collector"
##   .. ..$ OCC1950: list()
##   .. .. ..- attr(*, "class")= chr  "collector_integer" "collector"
##   ..$ default: list()
##   .. ..- attr(*, "class")= chr  "collector_skip" "collector"
##   ..- attr(*, "class")= chr "col_spec"
```

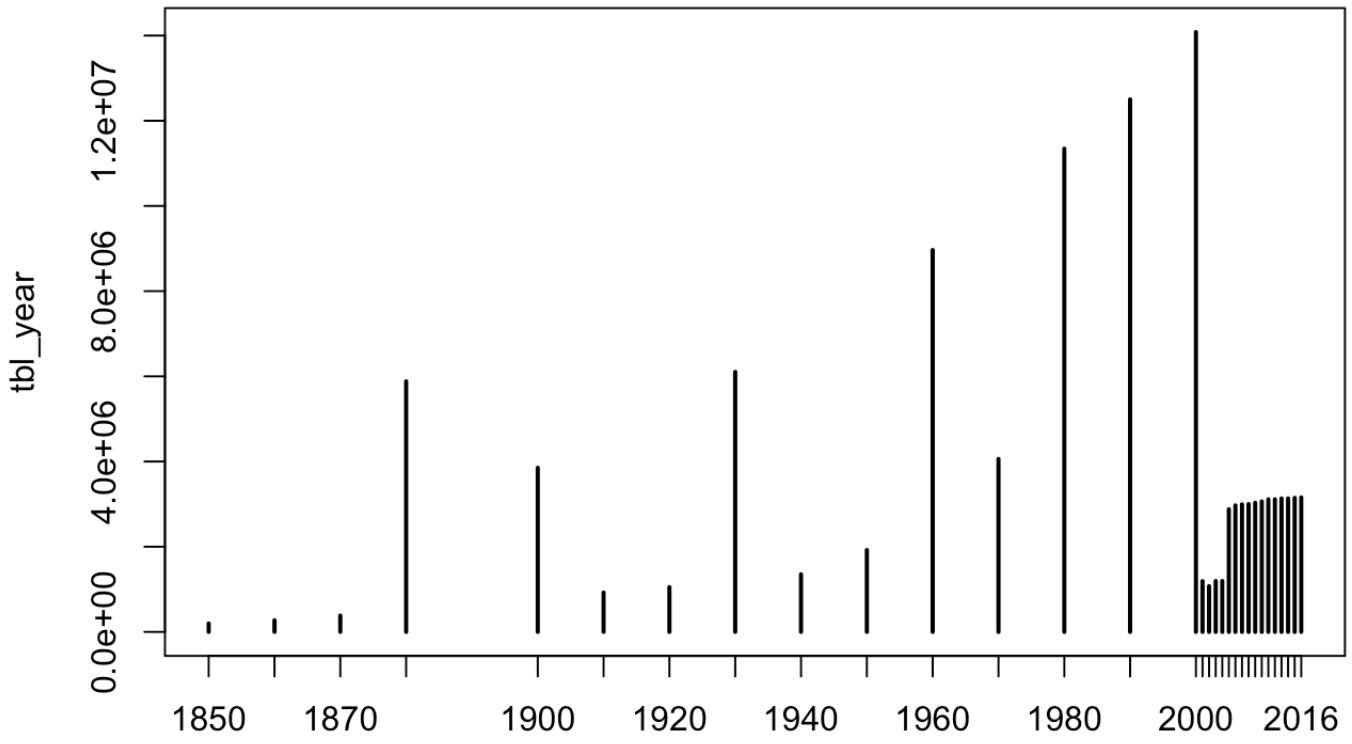
```
## [1] 114.2783
```

“data” is an object of classes “tbl_df”,“tbl” and “data.frame” with over 114 million rows.

That’s too few rows to have one row for each person in the most recent census – or even one row for each household in all the census since 1850:

```
if(readAndCompute){
  startYr <- Sys.time()
  str(tbl_year <- table(data$YEAR))
  (etYr <- Sys.time()-startYr)
  plot(tbl_year)
  tbl_year
}
```

```
##  'table' int [1:31(1d)] 197796 273596 383358 5882038 3852852 923153 1050634 610382
2 1351732 1922198 ...
##  - attr(*, "dimnames")=List of 1
##   ..$ : chr [1:31] "1850" "1860" "1870" "1880" ...
```




```
##
##     1850     1860     1870     1880     1900     1910     1920     1930
##   197796   273596   383358  5882038  3852852   923153  1050634  6103822
##     1940     1950     1960     1970     1980     1990     2000     2001
##  1351732  1922198  8965606  4059942 11343120 12501046 14081466  1192206
##     2002     2003     2004     2005     2006     2007     2008     2009
##  1074628  1194928  1194354  2878380  2969741  2994662  3000657  3030728
##     2010     2011     2012     2013     2014     2015     2016
##  3061692  3112017  3113030  3132795  3132610  3147005  3156487
```

The plot looks funny but shows that we have data from every census except 1890, and that with the listing shows that we also have data for each year between 2000 and 2016.

Let's look at "var_desc" for HHWT:

```
if(readAndCompute){
  attributes(data$HHWT)
}
```

```
## $label
## [1] "Household weight"
##
## $var_desc
## [1] "HHWT indicates how many households in the U.S. population are represented by
a given household in an IPUMS sample. \n\nIt is generally a good idea to use HHWT whe
n conducting a household-level analysis of any IPUMS sample. The use of HHWT is optio
nal when analyzing one of the \"flat\" or unweighted IPUMS samples. Flat IPUMS sample
s include the 1% samples from 1850-1930, all samples from 1960, 1970, and 1980, the 1
% unweighted samples from 1990 and 2000, the 10% 2010 sample, and any of the full cou
nt 100% census datasets. HHWT must be used to obtain nationally representative statis
tics for household-level analyses of any sample other than those.\n\nUsers should als
o be sure to select one person (e.g., PERNUM = 1) to represent the entire household.\
n\nFor further explanation of the sample weights, see \"Sample Designs\" [URL omitted
from DDI.] and \"Sample Weights\" [URL omitted from DDI.]. See also PERWT for a corre
sponding variable at the person level, and SLWT for a weight variable used with sampl
e-line records in 1940 1% and 1950."
```

Let's look at the distribution of HHWT:

```
if(readAndCompute){
  quantile(data$HHWT)
}
```

```
##   0%  25%  50%  75% 100%
##    0   20   24   85 4331
```

Let's also examine the the attributes of OCC1950:

```
if(readAndCompute){
  stOCC <- Sys.time()
  str(OCCcodes <- attributes(data$OCC1950))
  (etOCC <- Sys.time()-stOCC)
}
```

```
## List of 4
##  $ label   : chr "Occupation, 1950 basis"
##  $ var_desc: chr "Universe Note: \"New Workers\" are persons seeking employment fo
r the first time, who had not yet secured their"| __truncated__
##  $ class   : chr "labelled"
##  $ labels  : Named num [1:283] 0 1 2 3 4 5 6 7 8 9 ...
##   ..- attr(*, "names")= chr [1:283] "Accountants and auditors" "Actors and actress
es" "Airplane pilots and navigators" "Architects" ...
```

```
## Time difference of 0.003813028 secs
```

We're especially interested in "labels":

```
if(readAndCompute){
  head(OCCcodes$labels)
  tail(OCCcodes$labels)
}
```

```
##                      Inmate                  New Worker
##                         987                         990
##  Gentleman/lady/at leisure       Other non-occupation
##                         991                         995
## Occupation missing/unknown                N/A (blank)
##                         997                         999
```

"Accountants and auditors" are coded 0, "Gentleman/lady/at liesure" is 991, and there are two different missing value codes, which we should probably examine after we do some tabulations.

Let's sum HHWT within YEAR and OCC1950:

```
if(readAndCompute){
  stYrOcc <- proc.time()
  str(YrOcc <- tapply(data$HHWT, data[c("OCC1950", "YEAR")], sum))
  (etYrOcc <- proc.time()-stYrOcc)
}
```

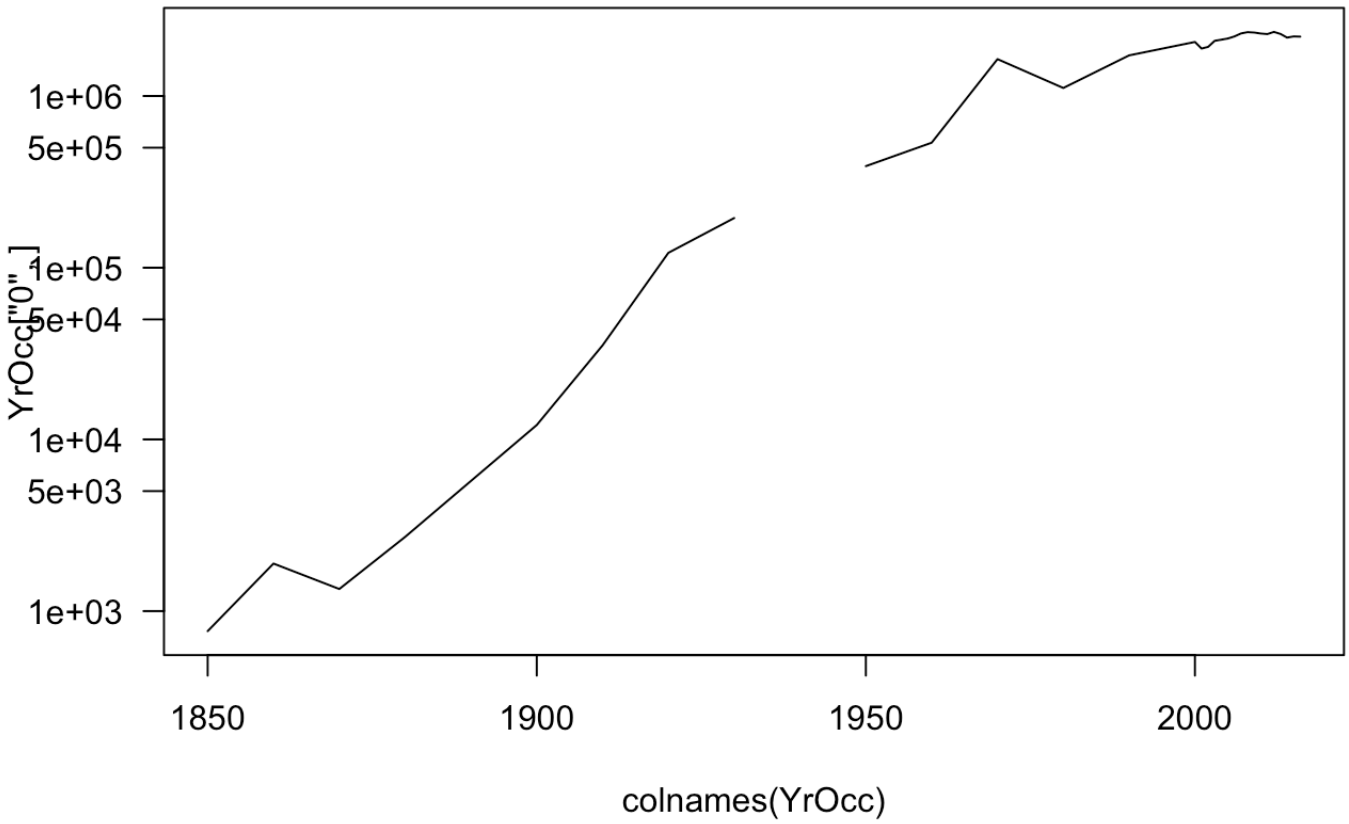
```
##  num [1:281, 1:31] 764 400 NA 397 2130 ...
##  - attr(*, "dimnames")=List of 2
##   ..$ OCC1950: chr [1:281] "0" "1" "2" "3" ...
##   ..$ YEAR   : chr [1:31] "1850" "1860" "1870" "1880" ...
```

```
##    user  system elapsed
##  24.836   4.405  29.608
```

This is an array of OCC1950 by YEAR. The first column should estimate the number of Accountants and Auditors by YEAR.

Let's plot

```
if(readAndCompute){
  plot(colnames(YrOcc), YrOcc['0', ], type='l', log='y', las=1)
}
```



What about the break in this line?

```
if(readAndCompute){
  YrOcc['0',]
}
```

```
##       1850       1860       1870       1880       1900       1910
##     764.49    1891.97    1345.02    2705.24   12129.25   35307.00
##       1920       1930       1940       1950       1960       1970
##  122104.25  194663.56         NA  390396.00  534680.00 1639800.00
##       1980       1990       2000       2001       2002       2003
## 1113580.00 1724241.00 2064061.00 1889156.00 1931920.00 2096069.00
##       2004       2005       2006       2007       2008       2009
## 2127959.00 2161838.00 2223026.00 2314146.00 2354861.00 2342316.00
##       2010       2011       2012       2013       2014       2015
## 2311519.00 2294479.00 2362162.00 2298363.00 2188585.00 2223301.00
##       2016
## 2217376.00
```

1940 is NA. Is this consistent across all OCC1950 codes?

```
if(readAndCompute){
  (yrNA <- colSums(is.na(YrOcc)))
}
```

```
## 1850 1860 1870 1880 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000
##  116   84   73   34   18   16   10    4   66   10   10   22   60   61   95
## 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
##   96   96   96   96   96   96   96   96   96   97   97  102  102  102  102
## 2016
##  102
```

Different occupation codes are missing for different years, ranging from 4 OCC1850 codes not used in 1930 to 116 in 1850 and 102 in 2016.

For the purpose of computing the size of the labor force, I think we should treat those NAs as 0, because people nominally with those occupations would probably have been counted in other categories.

```
if(readAndCompute){
  stNA <- proc.time()
  NA.yr <- colSums(is.na(data))
  (etNA <- proc.time()-stNA)
}
```

```
##    user  system elapsed
##   5.643   8.378  15.473
```

```
if(readAndCompute){
  NA.yr
}
```

```
##    YEAR DATANUM  SERIAL    HHWT      GQ  PERNUM   PERWT OCC1950
##       0       0       0       0       0       0       0       0
```

No NAs in “data”.

Note also that there are ony 281 rows in YrOcc, while OCCcodes $labels has length 283. Let's find which OCCcodes labels were not used:

```
if(readAndCompute){
  st01 <- proc.time()
  str(OCClbls <- table(data$OCC1950))
  (et01 <- proc.time()-st01)
}
```

```
##  'table' int [1:281(1d)] 625027 16950 41091 59521 89510 1273 58142 44603 20513 178
320 ...
##  - attr(*, "dimnames")=List of 1
##   ..$ : chr [1:281] "0" "1" "2" "3" ...
```

```
##    user  system elapsed
##  19.830   4.118  24.281
```

```
if(readAndCompute){
  OCCcodes$labels[!(OCCcodes$labels %in% names(OCClbls))]
}
```

```
## Not yet classified         New Worker
##                979                990
```

“Not yet classified” and “New Worker”.

That makes some sense: These codes may have been generated and may even have been used prior to data cleaning operations. If used, they’ve been eliminated from the data I received.

Let’s delete these two and create a logical variable of length 281 indicating which codes are in the labor force (TRUE / FALSE). To start, let’s look at the list of occupational names to see which would not have been counted in the labor force:

```
if(readAndCompute){
  OCCcodes$labels
}
```

```
##                                     Accountants and auditors
##                                                            0
##                                        Actors and actresses
##                                                            1
##                                  Airplane pilots and navigators
##                                                            2
##                                                    Architects
##                                                            3
##                                     Artists and art teachers
##                                                            4
##                                                      Athletes
##                                                            5
##                                                       Authors
##                                                            6
##                                                     Chemists
##                                                            7
##                                                  Chiropractors
##                                                            8
##                                                    Clergymen
##                                                            9
##                                   College presidents and deans
##                                                           10
##                 Agricultural sciences-Professors and instructors
##                                                           12
##                   Biological sciences-Professors and instructors
##                                                           13
##                           Chemistry-Professors and instructors
##                                                           14
##                           Economics-Professors and instructors
##                                                           15
##                         Engineering-Professors and instructors
##                                                           16
##                Geology and geophysics-Professors and instructors
##                                                           17
##                         Mathematics-Professors and instructors
##                                                           18
##                      Medical Sciences-Professors and instructors
##                                                           19
##                            Physics-Professors and instructors
##                                                           23
##                          Psychology-Professors and instructors
##                                                           24
```

```
##                         Statistics-Professors and instructors
##                                                            25
##              Natural science (nec)-Professors and instructors
##                                                            26
##              Social sciences (nec)-Professors and instructors
##                                                            27
##            Non-scientific subjects-Professors and instructors
##                                                            28
##              Subject not specified-Professors and instructors
##                                                            29
##                                 Dancers and dancing teachers
##                                                            31
##                                                       Dentists
##                                                            32
##                                                      Designers
##                                                            33
##                                 Dietitians and nutritionists
##                                                            34
##                                                     Draftsmen
##                                                            35
##                                       Editors and reporters
##                                                            36
##                                     Aeronautical-Engineers
##                                                            41
##                                        Chemical-Engineers
##                                                            42
##                                           Civil-Engineers
##                                                            43
##                                     Electrical-Engineers
##                                                            44
##                                     Industrial-Engineers
##                                                            45
##                                     Mechanical-Engineers
##                                                            46
##                         Metallurgical, metallurgists-Engineers
##                                                            47
##                                         Mining-Engineers
##                                                            48
##                                          Engineers (nec)
##                                                            49
##                                       Entertainers (nec)
##                                                            51
##                               Farm and home management advisors
##                                                            52
##                                  Foresters and conservationists
##                                                            53
##                                 Funeral directors and embalmers
```

```
##                                                 54
##                                  Lawyers and judges
##                                                 55
##                                          Librarians
##                                                 56
##                         Musicians and music teachers
##                                                 57
##                                 Nurses, professional
##                                                 58
##                         Nurses, student professional
##                                                 59
##                              Agricultural scientists
##                                                 61
##                                Biological scientists
##                                                 62
##                         Geologists and geophysicists
##                                                 63
##                                       Mathematicians
##                                                 67
##                                          Physicists
##                                                 68
##                             Misc. natural scientists
##                                                 69
##                                       Optometrists
##                                                 70
##                                         Osteopaths
##                                                 71
##                   Personnel and labor relations workers
##                                                 72
##                                        Pharmacists
##                                                 73
##                                       Photographers
##                                                 74
##                                Physicians and surgeons
##                                                 75
##                                     Radio operators
##                                                 76
##                         Recreation and group workers
##                                                 77
##                                     Religious workers
##                                                 78
##           Social and welfare workers, except group
##                                                 79
##                                         Economists
##                                                 81
##                                       Psychologists
##                                                 82
```

```
##                                  Statisticians and actuaries
##                                                            83
##                                       Misc social scientists
##                                                            84
##                                Sports instructors and officials
##                                                            91
##                                                     Surveyors
##                                                            92
##                                           Teachers (n.e.c.)
##                                                            93
##                                  Medical and dental-technicians
##                                                            94
##                                         Testing-technicians
##                                                            95
##                                           Technicians (nec)
##                                                            96
##                                  Therapists and healers (nec)
##                                                            97
##                                                Veterinarians
##                                                            98
##            Professional, technical and kindred workers (nec)
##                                                            99
##                                  Farmers (owners and tenants)
##                                                           100
##                                                Farm managers
##                                                           123
##                                  Buyers and dept heads, store
##                                                           200
##                               Buyers and shippers, farm products
##                                                           201
##                                         Conductors, railroad
##                                                           203
##                                                   Credit men
##                                                           204
##                               Floormen and floor managers, store
##                                                           205
##                                Inspectors, public administration
##                                                           210
##                           Managers and superintendants, building
##                                                           230
##                   Officers, pilots, pursers and engineers, ship
##                                                           240
##  Officials and administratators (nec), public administration
##                                                           250
##                           Officials, lodge, society, union, etc.
##                                                           260
##                                                    Postmasters
```

```
##                                                 270
##                       Purchasing agents and buyers (nec)
##                                                 280
##                  Managers, officials, and proprietors (nec)
##                                                 290
##                                       Agents (nec)
##                                                 300
##                       Attendants and assistants, library
##                                                 301
##                  Attendants, physicians and dentists office
##                                                 302
##                                Baggagemen, transportation
##                                                 304
##                                       Bank tellers
##                                                 305
##                                        Bookkeepers
##                                                 310
##                                           Cashiers
##                                                 320
##                               Collectors, bill and account
##                                                 321
##                         Dispatchers and starters, vehicle
##                                                 322
##                  Express messengers and railway mail clerks
##                                                 325
##                                      Mail carriers
##                                                 335
##                               Messengers and office boys
##                                                 340
##                                 Office machine operators
##                                                 341
##                              Shipping and receiving clerks
##                                                 342
##                    Stenographers, typists, and secretaries
##                                                 350
##                                    Telegraph messengers
##                                                 360
##                                     Telegraph operators
##                                                 365
##                                     Telephone operators
##                                                 370
##                          Ticket, station, and express agents
##                                                 380
##                       Clerical and kindred workers (n.e.c.)
##                                                 390
##                             Advertising agents and salesmen
##                                                 400
```

```
##                                                Auctioneers
##                                                        410
##                                              Demonstrators
##                                                        420
##                                     Hucksters and peddlers
##                                                        430
##                                  Insurance agents and brokers
##                                                        450
##                                                  Newsboys
##                                                        460
##                                Real estate agents and brokers
##                                                        470
##                                       Stock and bond salesmen
##                                                        480
##                                Salesmen and sales clerks (nec)
##                                                        490
##                                                   Bakers
##                                                        500
##                                                Blacksmiths
##                                                        501
##                                                Bookbinders
##                                                        502
##                                               Boilermakers
##                                                        503
##                         Brickmasons,stonemasons, and tile setters
##                                                        504
##                                              Cabinetmakers
##                                                        505
##                                                 Carpenters
##                                                        510
##                                  Cement and concrete finishers
##                                                        511
##                                    Compositors and typesetters
##                                                        512
##                               Cranemen,derrickmen, and hoistmen
##                                                        513
##                                  Decorators and window dressers
##                                                        514
##                                               Electricians
##                                                        515
##                                  Electrotypers and stereotypers
##                                                        520
##                                Engravers, except photoengravers
##                                                        521
##                   Excavating, grading, and road machinery operators
##                                                        522
##                                               Foremen (nec)
```

```
##                                                    523
##                                 Forgemen and hammermen
##                                                    524
##                                               Furriers
##                                                    525
##                                               Glaziers
##                                                    530
##                       Heat treaters, annealers, temperers
##                                                    531
##            Inspectors, scalers, and graders log and lumber
##                                                    532
##                                         Inspectors (nec)
##                                                    533
##         Jewelers, watchmakers, goldsmiths, and silversmiths
##                                                    534
##                                       Job setters, metal
##                                                    535
##      Linemen and servicemen, telegraph, telephone, and power
##                                                    540
##                                     Locomotive engineers
##                                                    541
##                                       Locomotive firemen
##                                                    542
##                                              Loom fixers
##                                                    543
##                                               Machinists
##                                                    544
##                           Airplane-mechanics and repairmen
##                                                    545
##                          Automobile-mechanics and repairmen
##                                                    550
##                      Office machine-mechanics and repairmen
##                                                    551
##                 Radio and television-mechanics and repairmen
##                                                    552
##                Railroad and car shop-mechanics and repairmen
##                                                    553
##                                Mechanics and repairmen (nec)
##                                                    554
##                           Millers, grain, flour, feed, etc
##                                                    555
##                                             Millwrights
##                                                    560
##                                           Molders, metal
##                                                    561
##                                Motion picture projectionists
##                                                    562
```

```
##                     Opticians and lens grinders and polishers
##                                                            563
##                        Painters, construction and maintenance
##                                                            564
##                                                    Paperhangers
##                                                            565
##                        Pattern and model makers, except paper
##                                                            570
##                              Photoengravers and lithographers
##                                                            571
##                         Piano and organ tuners and repairmen
##                                                            572
##                                                     Plasterers
##                                                            573
##                                    Plumbers and pipe fitters
##                                                            574
##                         Pressmen and plate printers, printing
##                                                            575
##                                  Rollers and roll hands, metal
##                                                            580
##                                       Roofers and slaters
##                                                            581
##                     Shoemakers and repairers, except factory
##                                                            582
##                                     Stationary engineers
##                                                            583
##                                Stone cutters and stone carvers
##                                                            584
##                                     Structural metal workers
##                                                            585
##                                      Tailors and tailoresses
##                                                            590
##             Tinsmiths, coppersmiths, and sheet metal workers
##                                                            591
##                      Tool makers, and die makers and setters
##                                                            592
##                                                   Upholsterers
##                                                            593
##                         Craftsmen and kindred workers (nec)
##                                                            594
##                                Members of the armed services
##                                                            595
##                                   Auto mechanics apprentice
##                                                            600
##                           Bricklayers and masons apprentice
##                                                            601
##                                       Carpenters apprentice
```

```
##                                                    602
##                                 Electricians apprentice
##                                                    603
##                       Machinists and toolmakers apprentice
##                                                    604
##                          Mechanics, except auto apprentice
##                                                    605
##                       Plumbers and pipe fitters apprentice
##                                                    610
##                         Apprentices, building trades (nec)
##                                                    611
##                     Apprentices, metalworking trades (nec)
##                                                    612
##                              Apprentices, printing  trades
##                                                    613
##                        Apprentices, other specified trades
##                                                    614
##                           Apprentices, trade not specified
##                                                    615
##                            Asbestos and insulation workers
##                                                    620
##                       Attendants, auto service and parking
##                                                    621
##                                  Blasters and powdermen
##                                                    622
##                        Boatmen, canalmen, and lock keepers
##                                                    623
##                                     Brakemen, railroad
##                                                    624
##                                           Bus drivers
##                                                    625
##                     Chainmen, rodmen, and axmen, surveying
##                                                    630
##                         Conductors, bus and street railway
##                                                    631
##                                  Deliverymen and routemen
##                                                    632
##                Dressmakers and seamstresses, except factory
##                                                    633
##                                                 Dyers
##                                                    634
##                      Filers, grinders, and polishers, metal
##                                                    635
## Fruit, nut, and vegetable graders, and packers, except facto
##                                                    640
##                           Furnacemen, smeltermen and pourers
##                                                    641
```

```
##                                          Heaters, metal
##                                                    642
##                          Laundry and dry cleaning Operatives
##                                                    643
##              Meat cutters, except slaughter and packing house
##                                                    644
##                                               Milliners
##                                                    645
##                                Mine operatives and laborers
##                                                    650
##                    Motormen, mine, factory, logging camp, etc
##                                                    660
##                  Motormen, street, subway, and elevated railway
##                                                    661
##                               Oilers and greaser, except auto
##                                                    662
##                   Painters, except construction or maintenance
##                                                    670
##                                Photographic process workers
##                                                    671
##                                     Power station operators
##                                                    672
##                                      Sailors and deck hands
##                                                    673
##                                                  Sawyers
##                                                    674
##                                           Spinners, textile
##                                                    675
##                                         Stationary firemen
##                                                    680
##                                       Switchmen, railroad
##                                                    681
##                                Taxicab drivers and chauffeurs
##                                                    682
##                                     Truck and tractor drivers
##                                                    683
##                                           Weavers, textile
##                                                    684
##                                Welders and flame cutters
##                                                    685
##                          Operative and kindred workers (nec)
##                                                    690
##                               Housekeepers, private household
##                                                    700
##                                Laundresses, private household
##                                                    710
##                               Private household workers (nec)
```

```
##                                                    720
##                  Attendants, hospital and other institution
##                                                    730
##          Attendants, professional and personal service (nec)
##                                                    731
##                        Attendants, recreation and amusement
##                                                    732
##                       Barbers, beauticians, and manicurists
##                                                    740
##                                                Bartenders
##                                                    750
##                                                Bootblacks
##                                                    751
##                          Boarding and lodging house keepers
##                                                    752
##                                     Charwomen and cleaners
##                                                    753
##                             Cooks, except private household
##                                                    754
##                                Counter and fountain workers
##                                                    760
##                                          Elevator operators
##                                                    761
##                                    Firemen, fire protection
##                                                    762
##                             Guards, watchmen, and doorkeepers
##                                                    763
##          Housekeepers and stewards, except private household
##                                                    764
##                                       Janitors and sextons
##                                                    770
##                                     Marshals and constables
##                                                    771
##                                                 Midwives
##                                                    772
##                                    Policemen and detectives
##                                                    773
##                                                  Porters
##                                                    780
##                                          Practical nurses
##                                                    781
##                                       Sheriffs and bailiffs
##                                                    782
##                               Ushers, recreation and amusement
##                                                    783
##                                     Waiters and waitresses
##                                                    784
```

```
##                       Watchmen (crossing) and bridge tenders
##                                                          785
##                Service workers, except private household (nec)
##                                                          790
##                                                  Farm foremen
##                                                          810
##                                Farm laborers, wage workers
##                                                          820
##                         Farm laborers, unpaid family workers
##                                                          830
##                         Farm service laborers, self-employed
##                                                          840
##                                     Fishermen and oystermen
##                                                          910
##                   Garage laborers and car washers and greasers
##                                                          920
##                      Gardeners, except farm and groundskeepers
##                                                          930
##                                   Longshoremen and stevedores
##                                                          940
##                        Lumbermen, raftsmen, and woodchoppers
##                                                          950
##                                                    Teamsters
##                                                          960
##                                               Laborers (nec)
##                                                          970
##                                          Not yet classified
##                                                          979
##                     Keeps house/housekeeping at home/housewife
##                                                          980
##                              Imputed keeping house (1850-1900)
##                                                          981
##                        Helping at home/helps parents/housework
##                                                          982
##                                           At school/student
##                                                          983
##                                                      Retired
##                                                          984
##                                 Unemployed/without occupation
##                                                          985
##                     Invalid/disabled w/ no occupation reported
##                                                          986
##                                                       Inmate
##                                                          987
##                                                   New Worker
##                                                          990
##                                       Gentleman/lady/at leisure
```

```
##                                                  991
##                                Other non-occupation
##                                                  995
##                             Occupation missing/unknown
##                                                  997
##                                          N/A (blank)
##                                                  999
```

“Keeps house/housekeeping at home/housewife, 980” has traditionally not been considered part of the labor force. “Imputed keeping house (1850-1900), 981” should probably be included with that. Similarly, “Helping at home/helps parents/housework, 982” has probably not been traditionally considered part of the labor force. Same with “At school/student, 983” and “Retired, 984”, “Unemployed/without occupation, 985”, “Invalid/disabled w no occupation reported, 986”, “Inmate, 987”, “Gentleman/lady/at leisure, 991”, “Other non-occupaton, 995”, “Occupation missing/unknown, 997”, and “N/A (blank), 999”.
Since “Keeps house … 980” exists, “Housekeepers, private household, 700” must be considered part of the labor force, as with “Laundresses, private household, 710” and “Private household workers (nec), 720”, I think.

What about “Farm laborers, unpaid family workers, 830”? Are these part of the labor force? Probably, being between “Farm laborers, wage workers, 820” and “Farm service laborers, self-employed, 840”.

```
if(readAndCompute){
  LaborForce <- rep(TRUE, length=nrow(YrOcc))
  names(LaborForce) <- rownames(YrOcc)
  LaborForce[as.character(c(980:987, 991, 995, 997, 999))] <- FALSE

  LaborForce[!LaborForce]
  table(LaborForce)

  OccNms <- OCCcodes$labels[OCCcodes$labels %in% names(LaborForce)]
  OccNames <- names(OccNms)
  names(OccNames) <- OccNms
  OccNames[!LaborForce]
}
```

```
##                                          980
## "Keeps house/housekeeping at home/housewife"
##                                          981
##          "Imputed keeping house (1850-1900)"
##                                          982
##    "Helping at home/helps parents/housework"
##                                          983
##                           "At school/student"
##                                          984
##                                  "Retired"
##                                          985
##              "Unemployed/without occupation"
##                                          986
## "Invalid/disabled w/ no occupation reported"
##                                          987
##                                   "Inmate"
##                                          991
##                 "Gentleman/lady/at leisure"
##                                          995
##                      "Other non-occupation"
##                                          997
##                "Occupation missing/unknown"
##                                          999
##                               "N/A (blank)"
```

That all looks good.

```
if(readAndCompute){
  str(TotalLaborForce <- colSums(YrOcc[LaborForce, ], na.rm=TRUE))
  plot(names(TotalLaborForce), TotalLaborForce, type='l', log='y',
     las=1)
  TotalLaborForce
}
```

```
##  Named num [1:31] 5337210 8378962 12935933 18033591 29271906 ...
##  - attr(*, "names")= chr [1:31] "1850" "1860" "1870" "1880" ...
```


```
##      1850      1860      1870      1880      1900      1910      1920
##   5337210   8378962  12935933  18033591  29271906  38878405  41794569
##      1930      1940      1950      1960      1970      1980      1990
##  49023147  51508164  61353379  86905200 219049800 128412700 143783301
##      2000      2001      2002      2003      2004      2005      2006
## 166544358 159068575 161437410 162253758 164259225 166918520 171604050
##      2007      2008      2009      2010      2011      2012      2013
## 172724909 172853102 172977129 173049142 171466219 172587199 173424328
##      2014      2015      2016
## 174497753 175731168 177135176
```

The number for 1970 seems suspect at 219 million, but the other numbers look plausible and roughly consistent with other sources. In particular, they seem more consistent with the Bicentennial Edition: Historical Statistics of the United States, Colonial Times to 1970 (https://www.census.gov/library/publications/1975/compendia/hist_stats_colonial-1970.html) than the Labor Force numbers (items Ba1033 and Ba1159) in the more recent Historical Statistics of the United States

(https://en.wikipedia.org/wiki/Historical_Statistics_of_the_United_States), whose numbers for “Accountants and auditors” (item Ba1161) seem to contain some fairly blatant errors, e.g., 0 for 1940 and 1700 and 1200 for 1860 and 1870, respectively, while the labor force grew by over 40% in that decade.

Let’s look at the ratio, being “Accountants and auditors” as a percent of the labor force:

```
if(readAndCompute){
  str(AApct <- (YrOcc["0", ] / TotalLaborForce))
  plot(names(AApct), 100*AApct, type='l', log='y', las=1)
  AApct
}
```

```
##  Named num [1:31] 0.000143 0.000226 0.000104 0.00015 0.000414 ...
##  - attr(*, "names")= chr [1:31] "1850" "1860" "1870" "1880" ...
```

```
##          1850          1860          1870          1880          1900
## 0.0001432378 0.0002258001 0.0001039755 0.0001500112 0.0004143649
##          1910          1920          1930          1940          1950
## 0.0009081391 0.0029215339 0.0039708499            NA 0.0063630725
##          1960          1970          1980          1990          2000
## 0.0061524512 0.0074859689 0.0086718837 0.0119919420 0.0123934610
##          2001          2002          2003          2004          2005
## 0.0118763621 0.0119669908 0.0129184620 0.0129548827 0.0129514568
##          2006          2007          2008          2009          2010
## 0.0129543912 0.0133978707 0.0136234813 0.0135411890 0.0133575872
##          2011          2012          2013          2014          2015
## 0.0133815221 0.0136867741 0.0132528292 0.0125421959 0.0126517170
##          2016
## 0.0125179880
```

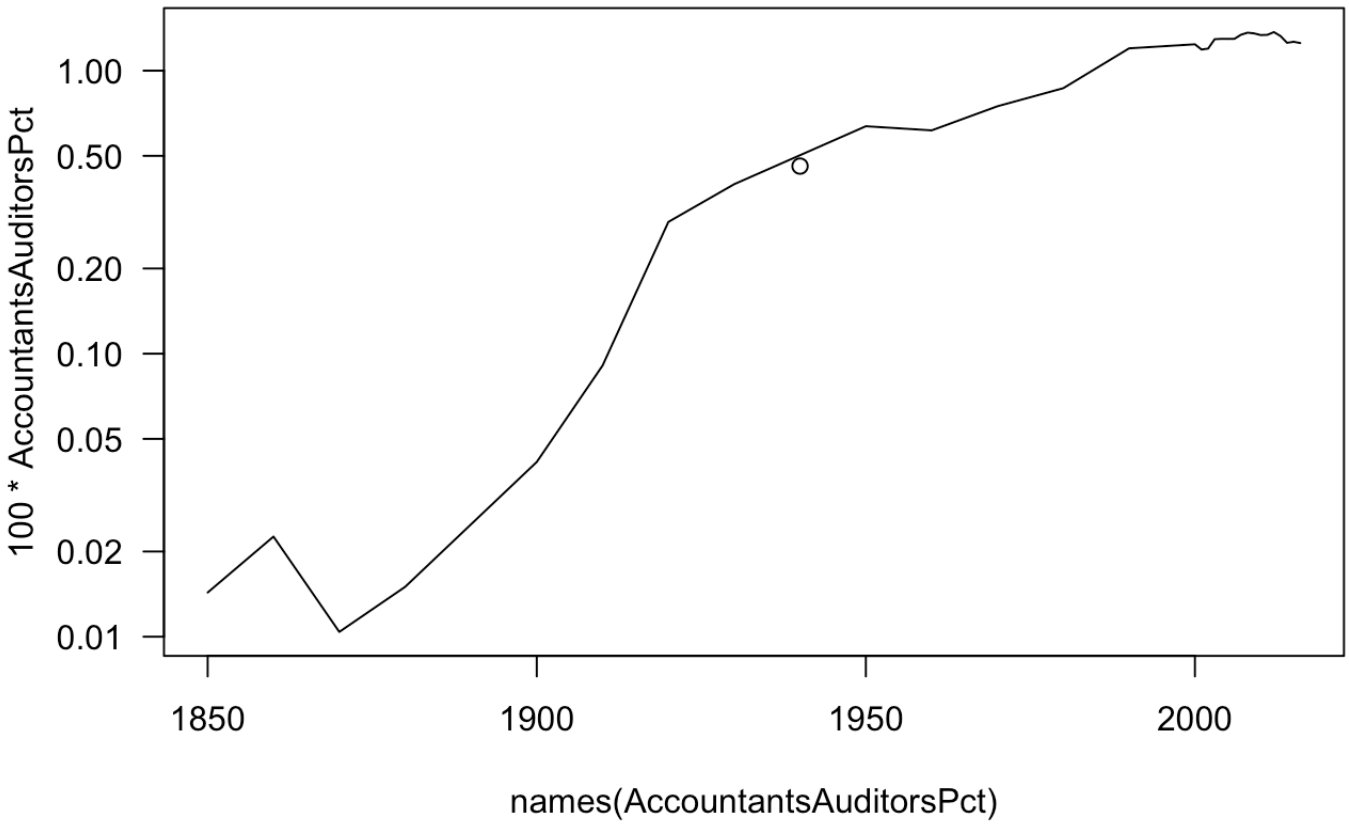
These numbers all look reasonably plausible, both internally consistent and moderately consistent with other sources.

Let's compare this with the 0.46% number from the Bicentennial Edition: Historical Statistics of the United States, Colonial Times to 1970 (https://www.census.gov/library/publications/1975/compendia/hist_stats_colonial-1970.html), which was used by Wyatt and Hecker (2006) "Occupational changes during the 20th century" (https://www.bls.gov/opub/mlr/2006/03/art3full.pdf):

```
if(readAndCompute){
  plot(names(AApct), 100*AApct, type='l', log='y', las=1)
  points(1940, 0.46)
}
```

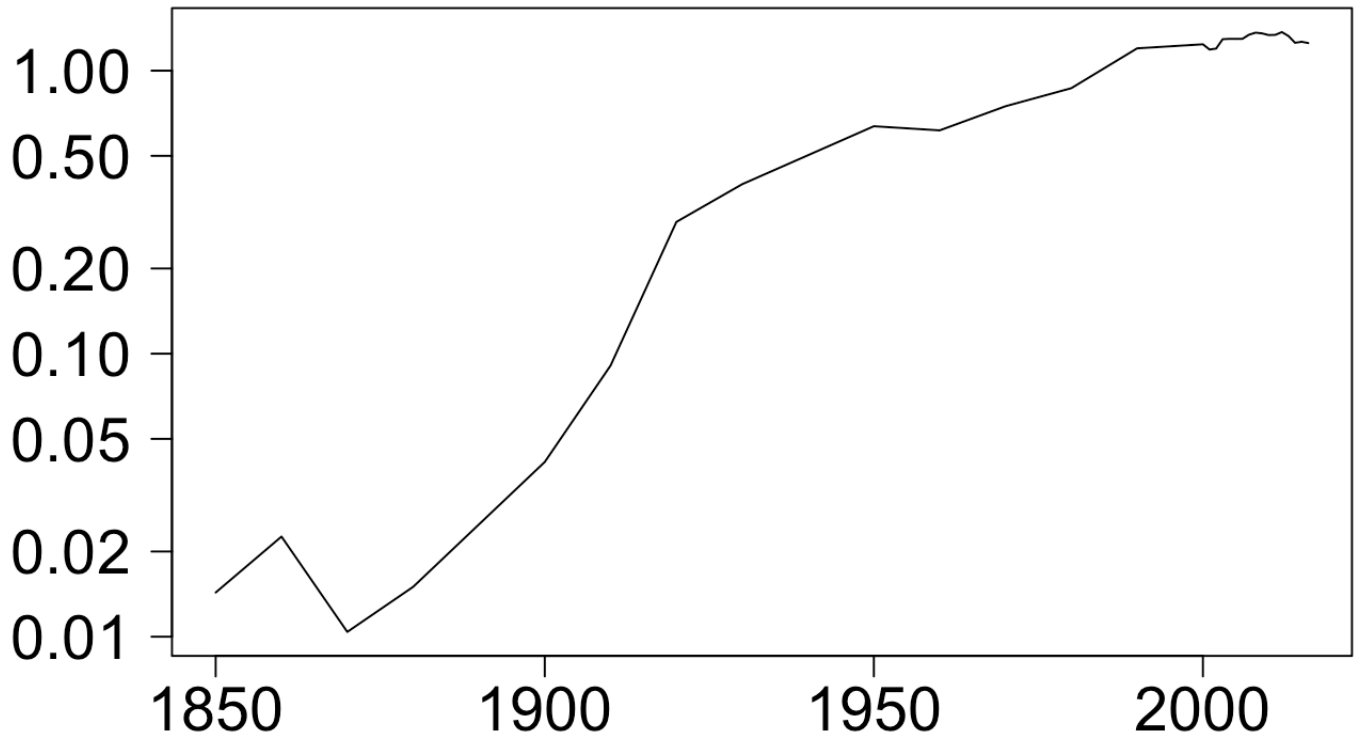
Let's just ignore that NA:

```
if(readAndCompute){
  AccountantsAuditorsPct <- AApct[!is.na(AApct)]
} else {
  library(Ecdat)
  data(AccountantsAuditorsPct)
}
plot(names(AccountantsAuditorsPct), 100*AccountantsAuditorsPct,
       type='l', log='y', las=1)
points(1940, 0.46)
```



Let's drop the 1940 point from Bicentennial Edition: Historical Statistics of the United States, Colonial Times to 1970 (https://www.census.gov/library/publications/1975/compendia/hist_stats_colonial-1970.html) and create an svg file suitable for Wikimedia Commons:

```
plot(names(AccountantsAuditorsPct), 100*AccountantsAuditorsPct,
     type='l', log='y', las=1,
     xlab='', ylab='', cex.axis=1.8)
```




```
svg('AccountantsAuditorsUS.svg')
plot(names(AccountantsAuditorsPct), 100*AccountantsAuditorsPct,
     type='l', log='y', las=1,
      xlab='', ylab='', cex.axis=1.8)
dev.off()
```

```
## quartz_off_screen
##                 2
```

Let's save "AccountantsAuditorsPct" and port to another platform, where "svg" works as advertised. [On 2018-08-31 "svg" failed to use "cex.axis=1.8" using R 3.5.1 on macOS 10.13.6, but R 3.2.1 under Windows 7 worked as expected.]

```
if(readAndCompute){
  save(AccountantsAuditorsPct, file='AccountantsAuditorsPct.rda')
}
```