R Notebook

Intro

In the wake of the Great Recession of 2009, there has been a good deal of focus on employment statistics, one of the most important metrics policymakers use to gauge the overall strength of the economy. In the United States, the government measures unemployment using the Current Population Survey (CPS), which collects demographic and employment information from a wide range of Americans each month. In this exercise, we will employ the topics reviewed in the lectures as well as a few new techniques using the September 2013 version of this rich, nationally representative dataset

Dataset used

(Compact version) **PeopleInHousehold**: The number of people in the interviewee's household. **Region**: The census region where the interviewee lives. **State**: The state where the interviewee lives. **MetroAreaCode**: A code that identifies the metropolitan area in which the interviewee lives (missing if the interviewee does not live in a metropolitan area). The mapping from codes to names of metropolitan areas is provided in the file MetroAreaCodes.csv. **Age**: The age, in years, of the interviewee. 80 represents people aged 80-84, and 85 represents people aged 85 and higher. **Married**: The marriage status of the interviewee. **Sex**: The sex of the interviewee. **Education**: The maximum level of education obtained by the interviewee. **Race**: The race of the interviewee. **Hispanic**: Whether the interviewee is of Hispanic ethnicity. **CountryOfBirthCode**: A code identifying the country of birth of the interviewee. The mapping from codes to names of countries is provided in the file CountryCodes.csv. **Citizenship**: The United States citizenship status of the interviewee. **EmploymentStatus**: The status of employment of the interviewee. **Industry**: The industry of employment of the interviewee (only available if they are employed).

1.1) Loading and Summarize the Dataset

```r
library(readr)
library(skimr)
```

```
## Warning: package 'skimr' was built under R version 3.4.4
```

```r
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 3.4.4
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'
```

```
## The following object is masked from 'package:dplyr':
##
##     combine
```

```r
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 3.4.2
## -- Attaching packages -----------------------------------------------------------
## v ggplot2 3.1.0     v purrr   0.2.5
## v tibble  2.0.1     v stringr 1.3.1
## v tidyr   0.8.0     v forcats 0.3.0
## Warning: package 'ggplot2' was built under R version 3.4.4
## Warning: package 'tibble' was built under R version 3.4.4
## Warning: package 'tidyr' was built under R version 3.4.3
## Warning: package 'purrr' was built under R version 3.4.4
## Warning: package 'forcats' was built under R version 3.4.3
## -- Conflicts --------------------------------------------------------------------
## x gridExtra::combine() masks dplyr::combine()
## x dplyr::filter()      masks stats::filter()
## x dplyr::lag()         masks stats::lag()
```

```r
CPS = read_csv("CPSData.csv")
```

```
## Parsed with column specification:
## cols(
##   PeopleInHousehold = col_integer(),
##   Region = col_character(),
##   State = col_character(),
##   MetroAreaCode = col_integer(),
##   Age = col_integer(),
##   Married = col_character(),
##   Sex = col_character(),
##   Education = col_character(),
##   Race = col_character(),
##   Hispanic = col_integer(),
##   CountryOfBirthCode = col_integer(),
##   Citizenship = col_character(),
##   EmploymentStatus = col_character(),
##   Industry = col_character()
## )
```

```r
CPS %>% glimpse()
```

```
## Observations: 131,302
## Variables: 14
## $ PeopleInHousehold  <int> 1, 3, 3, 3, 3, 3, 3, 2, 2, 2, 2, 1, 4, 4, 4...
## $ Region             <chr> "South", "South", "South", "South", "South"...
## $ State              <chr> "Alabama", "Alabama", "Alabama", "Alabama",...
## $ MetroAreaCode      <int> 26620, 13820, 13820, 13820, 26620, 26620, 2...
## $ Age                <int> 85, 21, 37, 18, 52, 24, 26, 71, 43, 52, 29,...
## $ Married            <chr> "Widowed", "Never Married", "Never Married"...
## $ Sex                <chr> "Female", "Male", "Female", "Male", "Female...
```

```
## $ Education          <chr> "Associate degree", "High school", "High sc...
## $ Race               <chr> "White", "Black", "Black", "Black", "White"...
## $ Hispanic           <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ CountryOfBirthCode <int> 57, 57, 57, 57, 57, 57, 57, 57, 57, 57, 57,...
## $ Citizenship        <chr> "Citizen, Native", "Citizen, Native", "Citi...
## $ EmploymentStatus   <chr> "Retired", "Unemployed", "Disabled", "Not i...
## $ Industry           <chr> NA, "Professional and business services", N...
```

Most common industry of employment

```r
CPS %>%
  count(Industry) %>%
  arrange(desc(n))
```

```
## Warning: package 'bindrcpp' was built under R version 3.4.4
```

```
## # A tibble: 15 x 2
##    Industry                                     n
##    <chr>                                    <int>
##  1 <NA>                                     65060
##  2 Educational and health services          15017
##  3 Trade                                     8933
##  4 Professional and business services        7519
##  5 Manufacturing                             6791
##  6 Leisure and hospitality                   6364
##  7 Construction                              4387
##  8 Financial                                 4347
##  9 Transportation and utilities              3260
## 10 Other services                            3224
## 11 Public administration                     3186
## 12 Information                               1328
## 13 Agriculture, forestry, fishing, and hunting  1307
## 14 Mining                                     550
## 15 Armed forces                                29
```

```r
# or
sort(table(CPS$Industry), decreasing = T)
```

```
##
##            Educational and health services
##                                      15017
##                                      Trade
##                                       8933
##         Professional and business services
##                                       7519
##                              Manufacturing
##                                       6791
##                    Leisure and hospitality
##                                       6364
##                               Construction
##                                       4387
##                                  Financial
##                                       4347
##               Transportation and utilities
##                                       3260
```

```
##                             Other services
##                                       3224
##                      Public administration
##                                       3186
##                                Information
##                                       1328
## Agriculture, forestry, fishing, and hunting
##                                       1307
##                                     Mining
##                                        550
##                              Armed forces
##                                         29
```

Which state has the fewest interviewees?

```r
sort(table(CPS$State))
```

```
##
##           New Mexico              Montana          Mississippi
##                 1102                 1214                 1230
##              Alabama        West Virginia             Arkansas
##                 1376                 1409                 1421
##            Louisiana                Idaho             Oklahoma
##                 1450                 1518                 1523
##              Arizona               Alaska              Wyoming
##                 1528                 1590                 1624
##         North Dakota       South Carolina            Tennessee
##                 1645                 1658                 1784
## District of Columbia             Kentucky                 Utah
##                 1791                 1841                 1842
##               Nevada              Vermont               Kansas
##                 1856                 1890                 1935
##               Oregon             Nebraska        Massachusetts
##                 1943                 1949                 1987
##         South Dakota              Indiana               Hawaii
##                 2000                 2004                 2099
##             Missouri         Rhode Island             Delaware
##                 2145                 2209                 2214
##                Maine           Washington                 Iowa
##                 2263                 2366                 2528
##           New Jersey       North Carolina        New Hampshire
##                 2567                 2619                 2662
##            Wisconsin              Georgia          Connecticut
##                 2686                 2807                 2836
##             Colorado             Virginia             Michigan
##                 2925                 2953                 3063
##            Minnesota             Maryland                 Ohio
##                 3139                 3200                 3678
##             Illinois         Pennsylvania              Florida
##                 3912                 3930                 5149
##             New York                Texas           California
##                 5595                 7077                11570
```

```r
# or
```

```r
CPS %>% count(State) %>% arrange(n)
```

```
## # A tibble: 51 x 2
##    State             n
##    <chr>         <int>
##  1 New Mexico     1102
##  2 Montana        1214
##  3 Mississippi    1230
##  4 Alabama        1376
##  5 West Virginia  1409
##  6 Arkansas       1421
##  7 Louisiana      1450
##  8 Idaho          1518
##  9 Oklahoma       1523
## 10 Arizona        1528
## # ... with 41 more rows
```

What proportion of interviewees are citizens of the United States

```r
table(CPS$Citizenship)
```

```
##
##     Citizen, Native Citizen, Naturalized          Non-Citizen
##              116639                 7073                 7590
```

```r
CPS %>%
  group_by(Citizenship) %>%
  summarize(n = n()) %>%
  mutate(freq = n/ sum(n))
```

```
## # A tibble: 3 x 3
##   Citizenship               n   freq
##   <chr>                 <int>  <dbl>
## 1 Citizen, Native      116639 0.888
## 2 Citizen, Naturalized   7073 0.0539
## 3 Non-Citizen            7590 0.0578
```

```r
# https://stackoverflow.com/questions/24576515/relative-frequencies-proportions-with-dplyr
0.88832615 + 0.05386818
```

```
## [1] 0.9421943
```

For which races are there at least 250 interviewees in the CPS dataset of Hispanic ethnicity?

```r
CPS %>%
  count(Race, Hispanic)
```

```
## # A tibble: 12 x 3
##    Race            Hispanic     n
##    <chr>              <int> <int>
##  1 American Indian        0  1129
##  2 American Indian        1   304
##  3 Asian                  0  6407
##  4 Asian                  1   113
```

```
##   5 Black                  0 13292
##   6 Black                  1   621
##   7 Multiracial            0  2449
##   8 Multiracial            1   448
##   9 Pacific Islander       0   541
##  10 Pacific Islander       1    77
##  11 White                  0 89190
##  12 White                  1 16731
```

2.1) Evaluating Missing Values

Which variables have at least one interviewee with a missing (NA) value?

```r
# https://stackoverflow.com/questions/8317231/elegant-way-to-report-missing-values-in-a-data-frame

colSums(is.na(CPS))
```

```
##   PeopleInHousehold             Region              State
##                   0                  0                  0
##       MetroAreaCode                Age            Married
##               34238                  0              25338
##                 Sex          Education               Race
##                   0              25338                  0
##            Hispanic CountryOfBirthCode        Citizenship
##                   0                  0                  0
##    EmploymentStatus           Industry
##               25789              65060
```

```r
# or
apply(is.na(CPS), 2, sum)
```

```
##   PeopleInHousehold             Region              State
##                   0                  0                  0
##       MetroAreaCode                Age            Married
##               34238                  0              25338
##                 Sex          Education               Race
##                   0              25338                  0
##            Hispanic CountryOfBirthCode        Citizenship
##                   0                  0                  0
##    EmploymentStatus           Industry
##               25789              65060
```

```r
table(CPS$Region, is.na(CPS$Married))
```

```
##
##              FALSE  TRUE
##   Midwest    24609  6075
##   Northeast  21432  4507
##   South      33535  7967
##   West       26388  6789
```

```r
table(CPS$Sex, is.na(CPS$Married))
```

```
##
##           FALSE  TRUE
##   Female 55264 12217
```

```
##   Male   50700 13121
```

```r
table(CPS$Age, is.na(CPS$Married))
```

```
##
##       FALSE TRUE
##   0       0 1283
##   1       0 1559
##   2       0 1574
##   3       0 1693
##   4       0 1695
##   5       0 1795
##   6       0 1721
##   7       0 1681
##   8       0 1729
##   9       0 1748
##   10      0 1750
##   11      0 1721
##   12      0 1797
##   13      0 1802
##   14      0 1790
##   15   1795    0
##   16   1751    0
##   17   1764    0
##   18   1596    0
##   19   1517    0
##   20   1398    0
##   21   1525    0
##   22   1536    0
##   23   1638    0
##   24   1627    0
##   25   1604    0
##   26   1643    0
##   27   1657    0
##   28   1736    0
##   29   1645    0
##   30   1854    0
##   31   1762    0
##   32   1790    0
##   33   1804    0
##   34   1653    0
##   35   1716    0
##   36   1663    0
##   37   1531    0
##   38   1530    0
##   39   1542    0
##   40   1571    0
##   41   1673    0
##   42   1711    0
##   43   1819    0
##   44   1764    0
##   45   1749    0
##   46   1665    0
##   47   1647    0
##   48   1791    0
```

```
##   49  1989    0
##   50  1966    0
##   51  1931    0
##   52  1935    0
##   53  1994    0
##   54  1912    0
##   55  1895    0
##   56  1935    0
##   57  1827    0
##   58  1874    0
##   59  1758    0
##   60  1746    0
##   61  1735    0
##   62  1595    0
##   63  1596    0
##   64  1519    0
##   65  1569    0
##   66  1577    0
##   67  1227    0
##   68  1130    0
##   69  1062    0
##   70  1195    0
##   71  1031    0
##   72   941    0
##   73   896    0
##   74   842    0
##   75   763    0
##   76   729    0
##   77   698    0
##   78   659    0
##   79   661    0
##   80  2664    0
##   85  2446    0
```

```r
table(CPS$Citizenship, is.na(CPS$Married))
```

```
##
##                        FALSE  TRUE
##   Citizen, Native      91956 24683
##   Citizen, Naturalized  6910   163
##   Non-Citizen           7098   492
```

How many states had all interviewees living in a non-metropolitan area (aka they have a missing MetroAreaCode value)? For this question, treat the District of Columbia as a state (even though it is not technically a state).

Approach: find states that have 0 for False

How many states had all interviewees living in a metropolitan area? Again, treat the District of Columbia as a state. Approach: find states that have 0 for True

```r
table(CPS$State, is.na(CPS$MetroAreaCode)) # FALSE= living in metropolitan area/ TRUE: not
```

```
##
##                        FALSE  TRUE
##   Alabama               1020   356
```

```
##    Alaska                   0  1590
##    Arizona               1327   201
##    Arkansas               724   697
##    California           11333   237
##    Colorado              2545   380
##    Connecticut           2593   243
##    Delaware              1696   518
##    District of Columbia  1791     0
##    Florida               4947   202
##    Georgia               2250   557
##    Hawaii                1576   523
##    Idaho                  761   757
##    Illinois              3473   439
##    Indiana               1420   584
##    Iowa                  1297  1231
##    Kansas                1234   701
##    Kentucky               908   933
##    Louisiana             1216   234
##    Maine                  909  1354
##    Maryland              2978   222
##    Massachusetts         1858   129
##    Michigan              2517   546
##    Minnesota             2150   989
##    Mississippi            376   854
##    Missouri              1440   705
##    Montana                199  1015
##    Nebraska               816  1133
##    Nevada                1609   247
##    New Hampshire         1148  1514
##    New Jersey            2567     0
##    New Mexico             832   270
##    New York              5144   451
##    North Carolina        1642   977
##    North Dakota           432  1213
##    Ohio                  2754   924
##    Oklahoma              1024   499
##    Oregon                1519   424
##    Pennsylvania          3245   685
##    Rhode Island          2209     0
##    South Carolina        1139   519
##    South Dakota           595  1405
##    Tennessee             1149   635
##    Texas                 6060  1017
##    Utah                  1455   387
##    Vermont                657  1233
##    Virginia              2367   586
##    Washington            1937   429
##    West Virginia          344  1065
##    Wisconsin             1882   804
##    Wyoming                  0  1624
```

Which region of the U.S. has the largest proportion of interviewees living in a non-metropolitan area?

```r
table(CPS$Region, is.na(CPS$MetroAreaCode)) # FALSE= living in metropolitan area/ TRUE: not
```

```
##
##             FALSE  TRUE
##   Midwest   20010 10674
##   Northeast 20330  5609
##   South     31631  9871
##   West      25093  8084
```

```r
# Proportion of interviewees living in a non-metropolitan area (True/Total)
10674 / (10674 + 20010) # Midwest
```

```
## [1] 0.3478686
```

```r
5609 / (5609 + 20330) # Northeast
```

```
## [1] 0.2162381
```

```r
9871 / (9871 + 31631) # South
```

```
## [1] 0.237844
```

```r
8084 / (8084 + 25093) # West
```

```
## [1] 0.2436628
```

```r
CPS %>%
  group_by(Citizenship) %>%
  summarize(n = n()) %>%
  mutate(freq = n/ sum(n))
```

```
## # A tibble: 3 x 3
##   Citizenship               n   freq
##   <chr>                 <int>  <dbl>
## 1 Citizen, Native      116639 0.888
## 2 Citizen, Naturalized   7073 0.0539
## 3 Non-Citizen            7590 0.0578
```

Which state has a proportion of interviewees living in a non-metropolitan area closest to 30%

(Use tapply() with the mean function to answer)

```r
# FALSE= living in metropolitan area/ TRUE: not
sort(tapply(is.na(CPS$MetroAreaCode), CPS$State, mean), decreasing = TRUE) # summary, group, function
```

```
##               Alaska              Wyoming              Montana
##           1.00000000           1.00000000           0.83607908
##        West Virginia         North Dakota         South Dakota
##           0.75585522           0.73738602           0.70250000
##          Mississippi              Vermont                Maine
##           0.69430894           0.65238095           0.59832081
##             Nebraska        New Hampshire             Kentucky
##           0.58132376           0.56874530           0.50678979
##                Idaho             Arkansas                 Iowa
##           0.49868248           0.49049965           0.48694620
##       North Carolina               Kansas            Tennessee
##           0.37304315           0.36227390           0.35594170
```

```
##             Missouri             Oklahoma            Minnesota
##           0.32867133           0.32764281           0.31506849
##       South Carolina            Wisconsin              Indiana
##           0.31302774           0.29932986           0.29141717
##              Alabama                 Ohio               Hawaii
##           0.25872093           0.25122349           0.24916627
##           New Mexico             Delaware               Oregon
##           0.24500907           0.23396567           0.21821925
##                 Utah             Virginia              Georgia
##           0.21009772           0.19844226           0.19843249
##           Washington             Michigan         Pennsylvania
##           0.18131868           0.17825661           0.17430025
##            Louisiana                Texas               Nevada
##           0.16137931           0.14370496           0.13308190
##              Arizona             Colorado             Illinois
##           0.13154450           0.12991453           0.11221881
##          Connecticut             New York             Maryland
##           0.08568406           0.08060769           0.06937500
##        Massachusetts              Florida           California
##           0.06492199           0.03923092           0.02048401
## District of Columbia           New Jersey         Rhode Island
##           0.00000000           0.00000000           0.00000000
```

```
# Answer: Wisconsin
```

Which state has the largest proportion of non-metropolitan interviewees, ignoring states where all interviewees were non-metropolitan?

```
# Answer: Montana
```

3.1) Integrating Metropolitan Area Data

Codes like MetroAreaCode and CountryOfBirthCode are a compact way to encode factor variables with text as their possible values, and they are therefore quite common in survey datasets. In fact, all but one of the variables in this dataset were actually stored by a numeric code in the original CPS datafile.

When analyzing a variable stored by a numeric code, we will often want to convert it into the values the codes represent. To do this, we will use a dictionary, which maps the the code to the actual value of the variable. We have provided dictionaries MetroAreaCodes.csv and CountryCodes.csv, which respectively map MetroAreaCode and CountryOfBirthCode into their true values. Read these two dictionaries into data frames MetroAreaMap and CountryMap.

```
MetroAreaMap = read_csv("MetroAreaCodes.csv")
```

```
## Parsed with column specification:
## cols(
##   Code = col_character(),
##   MetroArea = col_character()
## )
```

```
MetroAreaMap %>% glimpse()
```

```
## Observations: 271
## Variables: 2
## $ Code      <chr> "00460", "03000", "03160", "03610", "03720", "06450"...
## $ MetroArea <chr> "Appleton-Oshkosh-Neenah, WI", "Grand Rapids-Muskego...
```

```
CountryMap = read_csv("CountryCodes.csv")
```

```
## Parsed with column specification:
## cols(
##   Code = col_integer(),
##   Country = col_character()
## )
```

```
CountryMap %>% glimpse()
```

```
## Observations: 149
## Variables: 2
## $ Code    <int> 57, 66, 73, 78, 96, 100, 102, 103, 104, 105, 106, 108,...
## $ Country <chr> "United States", "Guam", "Puerto Rico", "U. S. Virgin ...
```

To merge in the metropolitan areas, we want to connect the field MetroAreaCode from the CPS data frame with the field Code in MetroAreaMap. The following command merges the two data frames on these columns, overwriting the CPS data frame with the result:

CPS = merge(CPS, MetroAreaMap, by.x="MetroAreaCode", by.y="Code", all.x=TRUE)

The first two arguments determine the data frames to be merged (they are called "x" and "y", respectively, in the subsequent parameters to the merge function). by.x="MetroAreaCode" means we're matching on the MetroAreaCode variable from the "x" data frame (CPS), while by.y="Code" means we're matching on the Code variable from the "y" data frame (MetroAreaMap). Finally, all.x=TRUE means we want to keep all rows from the "x" data frame (CPS), even if some of the rows' MetroAreaCode doesn't match any codes in MetroAreaMap (for those familiar with database terminology, this parameter makes the operation a left outer join instead of an inner join).

Review the new version of the CPS data frame with the summary() and str() functions. What is the name of the variable that was added to the data frame by the merge() operation?

```
CPS = merge(CPS, MetroAreaMap, by.x = "MetroAreaCode", by.y = "Code", all.x = TRUE)

summary(CPS)
```

```
##  MetroAreaCode   PeopleInHousehold    Region              State          
##  Min.   :10420   Min.   : 1.000    Length:131302      Length:131302     
##  1st Qu.:21780   1st Qu.: 2.000    Class :character   Class :character  
##  Median :34740   Median : 3.000    Mode  :character   Mode  :character  
##  Mean   :35075   Mean   : 3.284                                         
##  3rd Qu.:41860   3rd Qu.: 4.000                                         
##  Max.   :79600   Max.   :15.000                                         
##  NA's   :34238                                                          
##       Age          Married              Sex             Education        
##  Min.   : 0.00   Length:131302      Length:131302      Length:131302     
##  1st Qu.:19.00   Class :character   Class :character   Class :character  
##  Median :39.00   Mode  :character   Mode  :character   Mode  :character  
##  Mean   :38.83                                                           
##  3rd Qu.:57.00                                                           
##  Max.   :85.00                                                           
##                                                                          
##      Race              Hispanic      CountryOfBirthCode Citizenship       
##  Length:131302      Min.   :0.0000   Min.   : 57.00     Length:131302     
##  Class :character   1st Qu.:0.0000   1st Qu.: 57.00     Class :character  
##  Mode  :character   Median :0.0000   Median : 57.00     Mode  :character  
```

```
##                      Mean   :0.1393   Mean   : 82.68
##                      3rd Qu.:0.0000   3rd Qu.: 57.00
##                      Max.   :1.0000   Max.   :555.00
##
##  EmploymentStatus     Industry          MetroArea
##  Length:131302      Length:131302      Length:131302
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
##
```

```r
str(CPS)
```

```
## 'data.frame':    131302 obs. of  15 variables:
##  $ MetroAreaCode     : int  10420 10420 10420 10420 10420 10420 10420 10420 10420 10420 ...
##  $ PeopleInHousehold : int  4 4 2 4 1 3 4 4 2 3 ...
##  $ Region            : chr  "Midwest" "Midwest" "Midwest" "Midwest" ...
##  $ State             : chr  "Ohio" "Ohio" "Ohio" "Ohio" ...
##  $ Age               : int  2 9 73 40 63 19 30 6 60 32 ...
##  $ Married           : chr  NA NA "Married" "Married" ...
##  $ Sex               : chr  "Male" "Male" "Female" "Female" ...
##  $ Education         : chr  NA NA "Some college, no degree" "High school" ...
##  $ Race              : chr  "White" "White" "White" "White" ...
##  $ Hispanic          : int  0 0 0 0 0 0 0 1 0 0 ...
##  $ CountryOfBirthCode: int  57 57 57 362 57 57 203 57 57 57 ...
##  $ Citizenship       : chr  "Citizen, Native" "Citizen, Native" "Citizen, Native" "Citizen, Naturali
##  $ EmploymentStatus  : chr  NA NA "Retired" "Not in Labor Force" ...
##  $ Industry          : chr  NA NA NA NA ...
##  $ MetroArea         : chr  "Akron, OH" "Akron, OH" "Akron, OH" "Akron, OH" ...
```

How many interviewees have a missing value for the new metropolitan area variable? Note that all of these interviewees would have been removed from the merged data frame if we did not include the all.x=TRUE parameter.

```r
table(is.na(CPS$MetroArea)) # True = have missing values -> live in non-metro area
```

```
##
## FALSE  TRUE
## 97064 34238
```

Which metropolitan area has the highest proportion of interviewees of Hispanic ethnicity? Hint: Use tapply() with mean, as in the previous subproblem. Calling sort() on the output of tapply() could also be helpful here.

```r
sort(tapply(CPS$Hispanic,CPS$MetroArea, mean), decreasing = TRUE)
```

```
##                                     Laredo, TX
##                                    0.966292135
##                      McAllen-Edinburg-Pharr, TX
##                                    0.948717949
##                        Brownsville-Harlingen, TX
##                                    0.797468354
##                                    El Paso, TX
```

```
##                                        0.790983607
##                                     El Centro, CA
##                                        0.686868687
##                                   San Antonio, TX
##                                        0.644151565
##                                        Madera, CA
##                                        0.614035088
##                                Corpus Christi, TX
##                                        0.606060606
##                                        Merced, CA
##                                        0.566037736
##                                       Salinas, CA
##                                        0.557692308
##                                   Las Cruses, NM
##                                        0.542056075
##                                        Tucson, AZ
##                                        0.506622517
##                     Riverside-San Bernardino, CA
##                                        0.502325581
##                                  Bakersfield, CA
##                                        0.489795918
##            Miami-Fort Lauderdale-Miami Beach, FL
##                                        0.467824968
##                                      Victoria, TX
##                                        0.465517241
##                                      Santa Fe, NM
##                                        0.461538462
##             Los Angeles-Long Beach-Santa Ana, CA
##                                        0.460263286
##                                  Albuquerque, NM
##                                        0.441707718
##                         Cape Coral-Fort Myers, FL
##                                        0.438356164
##                          Visalia-Porterville, CA
##                                        0.438016529
##                                        Fresno, CA
##                                        0.409240924
##                Vineland-Millville-Bridgeton, NJ
##                                        0.407407407
##             Santa Barbara-Santa Maria-Goleta, CA
##                                        0.401515152
##                     Killeen-Temple-Fort Hood, TX
##                                        0.386138614
##                Oxnard-Thousand Oaks-Ventura, CA
##                                        0.359550562
##                   Houston-Baytown-Sugar Land, TX
##                                        0.359005458
##                                        Yakima, WA
##                                        0.357142857
##                                       Midland, TX
##                                        0.352941176
##                                       Modesto, CA
##                                        0.341772152
##                                       Danbury, CT
```

```
##                                              0.339285714
##                                                 Waco, TX
##                                              0.329113924
##                                             Stockton, CA
##                                              0.321243523
##                         San Jose-Sunnyvale-Santa Clara, CA
##                                              0.316417910
##                                    Austin-Round Rock, TX
##                                              0.310077519
##                                               Pueblo, CO
##                                              0.307692308
##                                             Longview, TX
##                                              0.292307692
##                                              Lubbock, TX
##                                              0.285714286
##                            Dallas-Fort Worth-Arlington, TX
##                                              0.283950617
##                       Poughkeepsie-Newburgh-Middletown, NY
##                                              0.273631841
##                          San Diego-Carlsbad-San Marcos, CA
##                                              0.269018743
##                      Sacramento-Arden-Arcade-Roseville, CA
##                                              0.263868066
##                                             Amarillo, TX
##                                              0.261363636
##                               Phoenix-Mesa-Scottsdale, AZ
##                                              0.254376931
##                                   Las Vegas-Paradise, NV
##                                              0.251732102
##                                            Waterbury, CT
##                                              0.248407643
##                            San Luis Obispo-Paso Robles, CA
##                                              0.246753247
##                                           Farmington, NM
##                                              0.234375000
##                                    Santa Rosa-Petaluma, CA
##                                              0.232558140
##                                        Denver-Aurora, CO
##                                              0.232047872
##                                                 Napa, CA
##                                              0.229508197
## New York-Northern New Jersey-Long Island, NY-NJ-PA
##                                              0.228508042
##                                  Beaumont-Port Author, TX
##                                              0.227642276
##                                      Springfield, MA-CT
##                                              0.219354839
##                                              Orlando, FL
##                                              0.213114754
##                                                Salem, OR
##                                              0.211764706
##                                              Reading, PA
##                                              0.211267606
##                                    Vallejo-Fairfield, CA
```

```
##                                    0.210526316
##                                Columbus, GA-AL
##                                    0.203389831
##            San Francisco-Oakland-Fremont, CA
##                                    0.199855700
##                                Reno-Sparks, NV
##                                    0.196774194
##                      Naples-Marco Island, FL
##                                    0.182926829
##             Chicago-Naperville-Joliet, IN-IN-WI
##                                    0.167388167
##                                    Greeley, CO
##                                    0.160493827
##             Tampa-St. Petersburg-Clearwater, FL
##                                    0.159144893
##                                      Ocala, FL
##                                    0.157894737
##                               Fayetteville, NC
##                                    0.155844156
##                             Salt Lake City, UT
##                                    0.154910097
##                    Santa-Cruz-Watsonville, CA
##                                    0.151515152
##           Fayetteville-Springdale-Rogers, AR-MO
##                                    0.148837209
##                                    Boulder, CO
##                                    0.146198830
##                          Ogden-Clearfield, UT
##                                    0.144208038
##                      Grand Rapids-Wyoming, MI
##                                    0.138157895
##                     Scranton-Wilkes Barre, PA
##                                    0.136363636
##                     Lakeland-Winter Haven, FL
##                                    0.134228188
##                                    Wichita, KS
##                                    0.133489461
##                              Trenton-Ewing, NJ
##                                    0.131868132
##                                   Prescott, AZ
##                                    0.129629630
##                               Jacksonville, NC
##                                    0.126984127
##                                  Green Bay, WI
##                                    0.125000000
##                                     Lawton, OK
##                                    0.123711340
##                      Athens-Clark County, GA
##                                    0.123076923
##                             Kansas City, MO-KS
##                                    0.121621622
##      Washington-Arlington-Alexandria, DC-VA-MD-WV
##                                    0.121378980
##                     Fort Collins-Loveland, CO
```

```
##                                  0.121359223
##                                  Olympia, WA
##                                  0.121212121
##                         Colorado Springs, CO
##                                  0.120967742
##                             Raleigh-Cary, NC
##                                  0.119047619
##            Charlotte-Gastonia-Concord, NC-SC
##                                  0.117988395
##                                    Chico, CA
##                                  0.116666667
##                         Kankakee-Bradley, IL
##                                  0.114942529
##                                    Tulsa, OK
##                                  0.114551084
##        Providence-Fall River-Warwick, MA-RI
##                                  0.114273205
##      Fort Walton Beach-Crestview-Destin, FL
##                                  0.112500000
##             Bridgeport-Stamford-Norwalk, CT
##                                  0.112328767
##             New Orleans-Metairie-Kenner, LA
##                                  0.111716621
##                                   Durham, NC
##                                  0.111111111
##                    Waterloo-Cedar Falls, IA
##                                  0.108974359
##                           Oklahoma City, OK
##                                  0.107615894
##    Hartford-West Hartford-East Hartford, CT
##                                  0.105084746
##                   Norwich-New London, CT-RI
##                                  0.103448276
##                                Lancaster, PA
##                                  0.102564103
##                               Tuscaloosa, AL
##                                  0.102564103
##              Port St. Lucie-Fort Pierce, FL
##                                  0.100917431
##      Deltona-Daytona Beach-Ormond Beach, FL
##                                  0.100000000
##        Portland-Vancouver-Beaverton, OR-WA
##                                  0.094582185
##                                   Topeka, KS
##                                  0.093406593
##              Augusta-Richmond County, GA-SC
##                                  0.093167702
##                        Boise City-Nampa, ID
##                                  0.093167702
##        Davenport-Moline-Rock Island, IA-IL
##                                  0.091666667
##                             Jacksonville, FL
##                                  0.091603053
##             Leominster-Fitchburg-Gardner, MA
```

```
##                                   0.090909091
##                             Atlantic City, NJ
##                                   0.090090090
##                  Seattle-Tacoma-Bellevue, WA
##                                   0.088446215
##                Hickory-Morgantown-Lenoir, NC
##                                   0.087719298
##            Allentown-Bethlehem-Easton, PA-NJ
##                                   0.086826347
##                            Fort Smith, AR-OK
##                                   0.085714286
##           Atlanta-Sandy Springs-Marietta, GA
##                                   0.085695876
##            Milwaukee-Waukesha-West Allis, WI
##                                   0.085434174
##                                  Medford, OR
##                                   0.085365854
##                       Lansing-East Lansing, MI
##                                   0.084033613
##                             Worcester, MA-CT
##                                   0.083333333
##                          Baltimore-Towson, MD
##                                   0.082265678
##                  Shreveport-Bossier City, LA
##                                   0.082191781
##                                 Syracuse, NY
##                                   0.080717489
##                                 Columbia, SC
##                                   0.079037801
##     Philadelphia-Camden-Wilmington, PA-NJ-DE
##                                   0.078458844
##                         Chattanooga, TN-GA
##                                   0.077844311
##                       Eugene-Springfield, OR
##                                   0.076530612
##                         Canton-Massillon, OH
##                                   0.076271186
##                               Vero Beach, FL
##                                   0.075949367
##                     Greensboro-High Point, NC
##                                   0.075697211
##                               Utica-Rome, NY
##                                   0.075000000
##                               Des Moines, IA
##                                   0.073852295
##                                New Haven, CT
##                                   0.073122530
##                             Indianapolis, IN
##                                   0.071929825
##                  Omaha-Council Bluffs, NE-IA
##                                   0.070010449
##                              Tallahassee, FL
##                                   0.069767442
##               Boston-Cambridge-Quincy, MA-NH
```

```
##                                   0.069537909
##          Nashville-Davidson-Murfreesboro, TN
##                                   0.069306931
##                                  Kingston, NY
##                                   0.068965517
##                   Panama City-Lynn Haven, FL
##                                   0.067796610
##                                Ocean City, NJ
##                                   0.066666667
##                                Provo-Orem, UT
##                                   0.064724919
##                                 Anderson, IN
##                                   0.064516129
##                                   Monroe, MI
##                                   0.063492063
##                                   Peoria, IL
##                                   0.062500000
##                                Lafayette, LA
##                                   0.060773481
##                                Asheville, NC
##                                   0.060344828
##                  Cleveland-Elyria-Mentor, OH
##                                   0.060205580
##                                 Honolulu, HI
##                                   0.059644670
##   Myrtle Beach-Conway-North Myrtle Beach, SC
##                                   0.058823529
##                                   Racine, WI
##                                   0.058823529
##                                Rochester, NY
##                                   0.058631922
##                    Bremerton-Silverdale, WA
##                                   0.057471264
##                                    Dover, DE
##                                   0.057017544
##                            Winston-Salem, NC
##                                   0.055118110
##                         Birmingham-Hoover, AL
##                                   0.053571429
##             Palm Bay-Melbourne-Titusville, FL
##                                   0.053571429
##                                  Decatur, Al
##                                   0.052083333
##        Minneapolis-St Paul-Bloomington, MN-WI
##                                   0.052008239
##   Virginia Beach-Norfolk-Newport News, VA-NC
##                                   0.050251256
##                  South Bend-Mishawaka, IN-MI
##                                   0.049382716
##                         Anniston-Oxford, AL
##                                   0.049180328
##                                 Valdosta, GA
##                                   0.047619048
##                 Sarasota-Bradenton-Venice, FL
```

```
##                           0.046875000
##                            Albany, GA
##                           0.044117647
##                          Rockford, IL
##                           0.043859649
##                          Columbus, OH
##                           0.043557169
##                       Springfield, MO
##                           0.043478261
##                       Gainesville, FL
##                           0.042857143
##                          Richmond, VA
##                           0.042857143
##                      York-Hanover, PA
##                           0.042735043
##                          Columbia, MO
##                           0.042553191
##                       Sioux Falls, SD
##                           0.042016807
##                       Punta Gorda, FL
##                           0.041666667
##                        Binghamton, NY
##                           0.041095890
##           Albany-Schenectady-Troy, NY
##                           0.041044776
##                          Lawrence, KS
##                           0.040816327
##                 Lexington-Fayette, KY
##                           0.040404040
##         Cincinnati-Middletown, OH-KY-IN
##                           0.040333797
##                             Flint, MI
##                           0.039215686
##            Michigan City-La Porte, IN
##                           0.038961039
##                     Louisville, KY-IN
##                           0.038535645
##                      Johnson City, TN
##                           0.038461538
##                       Baton Rouge, LA
##                           0.038167939
##                        Greenville, SC
##                           0.037837838
##            Detroit-Warren-Livonia, MI
##                           0.037666174
##     Little Rock-North Little Rock, AR
##                           0.037128713
##                        Fort Wayne, IN
##                           0.036764706
##                            Toledo, OH
##                           0.034042553
##                 Champaign-Urbana, IL
##                           0.032786885
##         Youngstown-Warren-Boardman, OH
```

```
##                          0.032679739
##                   Kalamazoo-Portage, MI
##                          0.031496063
##                        Iowa City, IA
##                          0.030534351
##                  Rochester-Dover, NH-ME
##                          0.030534351
##                      St. Louis, MO-IL
##                          0.030334728
##                       Janesville, WI
##                          0.030303030
##                          Roanoke, VA
##                          0.030303030
##                         Billings, MT
##                          0.030150754
##                      Springfield, OH
##                          0.029411765
##                     Memphis, TN-MS-AR
##                          0.028735632
##         Pensacola-Ferry Pass-Brent, FL
##                          0.028037383
##                        Lynchburg, VA
##                          0.027397260
##     Saginaw-Saginaw Township North, MI
##                          0.027027027
##                     Coeur d'Alene, ID
##                          0.025641026
##                          Spokane, WA
##                          0.025641026
##                         Fargo, ND-MN
##                          0.025462963
##                      Lake Charles, LA
##                          0.024691358
##                          Madison, WI
##                          0.024647887
##                             Erie, PA
##                          0.022988506
##                Harrisburg-Carlisle, PA
##                          0.022988506
##             Muskegon-Norton Shores, MI
##                          0.022222222
##                             Bend, OR
##                          0.021428571
##                     Evansville, IN-KY
##                          0.020202020
##                      Spartanburg, SC
##                          0.020202020
##                Niles-Benton Harbor, MI
##                          0.019607843
##                        La Crosse, WI
##                          0.017543860
##              Buffalo-Niagara Falls, NY
##                          0.017441860
##        Charleston-North Charleston, SC
```

```
##                                0.017241379
##                                 Joplin, MO
##                                0.016949153
##                             Pittsburgh, PA
##                                0.016393443
##                              Duluth, MN-WI
##                                0.015873016
##                        Gulfport-Biloxi, MS
##                                0.015384615
##                           Cedar Rapids, IA
##                                0.015306122
##                  Kingsport-Bristol, TN-VA
##                                0.014925373
##                                 Bangor, ME
##                                0.014423077
##                             Bellingham, WA
##                                0.014285714
##                            Springfield, IL
##                                0.013157895
##                                  Akron, OH
##                                0.012987013
##                    Holland-Grand Haven, MI
##                                0.012820513
##                                Altoona, PA
##                                0.012195122
##                              St. Cloud, MN
##                                0.012195122
##                         Oshkosh-Neenah, WI
##                                0.011764706
##               Portland-South Portland, ME
##                                0.011412268
##                                 Wausau, WI
##                                0.010416667
##                             Montgomery, AL
##                                0.009708738
##           Burlington-South Burlington, VT
##                                0.009132420
##                                Jackson, MS
##                                0.009009009
##                                Appleton,WI
##                                0.008000000
##                             Charleston, WV
##                                0.007633588
##                              Knoxville, TN
##                                0.005952381
##                                 Monroe, LA
##                                0.005586592
##                                 Dayton, OH
##                                0.003731343
##                               Anderson, SC
##                                0.000000000
##                              Ann Arbor, MI
##                                0.000000000
##                        Barnstable Town, MA
```

```
##                                         0.000000000
##                                Bloomington-Normal IL
##                                         0.000000000
##                                     Bloomington, IN
##                                         0.000000000
##                                   Bowling Green, KY
##                                         0.000000000
##                                         Decatur, IL
##                                         0.000000000
##                                     Eau Claire, WI
##                                         0.000000000
##                                       Florence, AL
##                                         0.000000000
##                      Hagerstown-Martinsburg, MD-WV
##                                         0.000000000
##                                   Harrisonburg, VA
##                                         0.000000000
##                       Huntington-Ashland, WV-KY-OH
##                                         0.000000000
##                                     Huntsville, AL
##                                         0.000000000
##                                        Jackson, MI
##                                         0.000000000
##                                      Johnstown, PA
##                                         0.000000000
##                                          Macon, GA
##                                         0.000000000
##                                         Mobile, AL
##                                         0.000000000
##                                      Salisbury, MD
##                                         0.000000000
##                                       Savannah, GA
##                                         0.000000000
##                                  Warner Robins, GA
##                                         0.000000000
```

Determine the number of metropolitan areas in the United States from which at least 20% of interviewees are Asian.

```
sort(tapply(CPS$Race == "Asian", CPS$MetroArea, mean), decreasing = TRUE)
```

```
##                                       Honolulu, HI
##                                         0.501903553
##                 San Francisco-Oakland-Fremont, CA
##                                         0.246753247
##                San Jose-Sunnyvale-Santa Clara, CA
##                                         0.241791045
##                            Vallejo-Fairfield, CA
##                                         0.203007519
##                                         Fresno, CA
##                                         0.184818482
##                                  Warner Robins, GA
##                                         0.166666667
##                                       Stockton, CA
```

```
##                                           0.155440415
##                                     Atlantic City, NJ
##                                           0.144144144
##              Sacramento-Arden-Arcade-Roseville, CA
##                                           0.142428786
##                     San Diego-Carlsbad-San Marcos, CA
##                                           0.142227122
##                Los Angeles-Long Beach-Santa Ana, CA
##                                           0.135056070
##                                           Olympia, WA
##                                           0.131313131
##                                           Salinas, CA
##                                           0.125000000
## New York-Northern New Jersey-Long Island, NY-NJ-PA
##                                           0.104270660
##                           Seattle-Tacoma-Bellevue, WA
##                                           0.099601594
##                               Visalia-Porterville, CA
##                                           0.090909091
##                                         Green Bay, WI
##                                           0.088235294
##                                        La Crosse, WI
##                                           0.087719298
##                                         Ann Arbor, MI
##                                           0.082352941
##                                       Bakersfield, CA
##                                           0.081632653
##                             Greensboro-High Point, NC
##                                           0.079681275
##                                Las Vegas-Paradise, NV
##                                           0.078521940
##            Minneapolis-St Paul-Bloomington, MN-WI
##                                           0.076725026
##                             Brownsville-Harlingen, TX
##                                           0.075949367
##                                Bloomington-Normal IL
##                                           0.075000000
##                   Oxnard-Thousand Oaks-Ventura, CA
##                                           0.074906367
##                                      Lake Charles, LA
##                                           0.074074074
##                             Norwich-New London, CT-RI
##                                           0.073891626
##                     Atlanta-Sandy Springs-Marietta, GA
##                                           0.072809278
##       Washington-Arlington-Alexandria, DC-VA-MD-WV
##                                           0.070624850
##                Portland-Vancouver-Beaverton, OR-WA
##                                           0.069788797
##          Hartford-West Hartford-East Hartford, CT
##                                           0.066666667
##                                      Cedar Rapids, IA
##                                           0.066326531
##                                         Rochester, NY
```

```
##                                            0.065146580
##                                           Columbia, MO
##                                            0.063829787
##                     Dallas-Fort Worth-Arlington, TX
##                                            0.062801932
##                                            Danbury, CT
##                                            0.062500000
##                        Riverside-San Bernardino, CA
##                                            0.062015504
##                      Houston-Baytown-Sugar Land, TX
##                                            0.061249242
##                                            Boulder, CO
##                                            0.058479532
##                 Chicago-Naperville-Joliet, IN-IN-WI
##                                            0.058441558
##                                        Reno-Sparks, NV
##                                            0.058064516
##                                   Baltimore-Towson, MD
##                                            0.057990560
##                                          Lancaster, PA
##                                            0.057692308
##                 Nashville-Davidson-Murfreesboro, TN
##                                            0.057425743
##                                      Fort Smith, AR-OK
##                                            0.057142857
##                                             Merced, CA
##                                            0.056603774
##                                            Madison, WI
##                                            0.056338028
##                                             Peoria, IL
##                                            0.053571429
##                                          Iowa City, IA
##                                            0.053435115
##                                        Springfield, IL
##                                            0.052631579
##                                  Austin-Round Rock, TX
##                                            0.052325581
##                            Buffalo-Niagara Falls, NY
##                                            0.052325581
##                      Boston-Cambridge-Quincy, MA-NH
##                                            0.052041274
##                     San Luis Obispo-Paso Robles, CA
##                                            0.051948052
##               Fayetteville-Springdale-Rogers, AR-MO
##                                            0.051162791
##                                            Orlando, FL
##                                            0.050819672
##                                       Raleigh-Cary, NC
##                                            0.050595238
##                                              Tulsa, OK
##                                            0.049535604
##                                    Anniston-Oxford, AL
##                                            0.049180328
##                     Burlington-South Burlington, VT
```

```
##                                       0.048706240
##                                  Jacksonville, FL
##                                       0.048346056
##                                  Jacksonville, NC
##                                       0.047619048
##                 Milwaukee-Waukesha-West Allis, WI
##                                       0.047619048
##                                     New Haven, CT
##                                       0.047430830
##                                 Trenton-Ewing, NJ
##                                       0.043956044
##                        Detroit-Warren-Livonia, MI
##                                       0.043574594
##                                   Gainesville, FL
##                                       0.042857143
##                       Portland-South Portland, ME
##                                       0.042796006
##                                       Decatur, Al
##                                       0.041666667
##                                   Albuquerque, NM
##                                       0.041050903
##                                      Syracuse, NY
##                                       0.040358744
##                                     Duluth, MN-WI
##                                       0.039682540
##               Tampa-St. Petersburg-Clearwater, FL
##                                       0.039192399
##              Providence-Fall River-Warwick, MA-RI
##                                       0.038966725
##                   Bridgeport-Stamford-Norwalk, CT
##                                       0.038356164
##                                    Pittsburgh, PA
##                                       0.038251366
##                       Phoenix-Mesa-Scottsdale, AZ
##                                       0.038105046
##                                    Des Moines, IA
##                                       0.037924152
##            Fort Walton Beach-Crestview-Destin, FL
##                                       0.037500000
##                                    Fort Wayne, IN
##                                       0.036764706
##                                      Richmond, VA
##                                       0.036734694
##                      Huntington-Ashland, WV-KY-OH
##                                       0.036585366
##                                        Mobile, AL
##                                       0.036363636
##                                Salt Lake City, UT
##                                       0.035961272
##                 Palm Bay-Melbourne-Titusville, FL
##                                       0.035714286
##             Miami-Fort Lauderdale-Miami Beach, FL
##                                       0.035392535
##                            Lexington-Fayette, KY
```

```
##                                        0.035353535
##                     Hickory-Morgantown-Lenoir, NC
##                                        0.035087719
##                                 Oklahoma City, OK
##                                        0.034768212
##                                  Worcester, MA-CT
##                                        0.034722222
##                                Kansas City, MO-KS
##                                        0.034303534
##                        Cape Coral-Fort Myers, FL
##                                        0.034246575
##                                 Harrisonburg, VA
##                                        0.033333333
##          Philadelphia-Camden-Wilmington, PA-NJ-DE
##                                        0.032924694
##                                   Greenville, SC
##                                        0.032432432
##                                 Denver-Aurora, CO
##                                        0.031914894
##                                     Anderson, SC
##                                        0.031250000
##                          Athens-Clark County, GA
##                                        0.030769231
##                              Gulfport-Biloxi, MS
##                                        0.030769231
##                                      Wichita, KS
##                                        0.030444965
##                                        Akron, OH
##                                        0.030303030
##                      Omaha-Council Bluffs, NE-IA
##                                        0.029258098
##                                   Montgomery, AL
##                                        0.029126214
##                                   Bellingham, WA
##                                        0.028571429
##                                     Fargo, ND-MN
##                                        0.027777778
##                                     Columbia, SC
##                                        0.027491409
##                        Lakeland-Winter Haven, FL
##                                        0.026845638
##        Virginia Beach-Norfolk-Newport News, VA-NC
##                                        0.026800670
##                           Rochester-Dover, NH-ME
##                                        0.026717557
##                             Ogden-Clearfield, UT
##                                        0.026004728
##                                  Fayetteville, NC
##                                        0.025974026
##                          Holland-Grand Haven, MI
##                                        0.025641026
##                    Augusta-Richmond County, GA-SC
##                                        0.024844720
##                                  Indianapolis, IN
```

```
##                                    0.024561404
##                        Naples-Marco Island, FL
##                                    0.024390244
##                                     Bangor, ME
##                                    0.024038462
##                     Bremerton-Silverdale, WA
##                                    0.022988506
##                                Baton Rouge, LA
##                                    0.022900763
##                  Albany-Schenectady-Troy, NY
##                                    0.022388060
##            Little Rock-North Little Rock, AR
##                                    0.022277228
##              Cincinnati-Middletown, OH-KY-IN
##                                    0.022253129
##                                     Topeka, KS
##                                    0.021978022
##       Deltona-Daytona Beach-Ormond Beach, FL
##                                    0.021428571
##          Davenport-Moline-Rock Island, IA-IL
##                                    0.020833333
##                      Eugene-Springfield, OR
##                                    0.020408163
##                                 El Centro, CA
##                                    0.020202020
##                                     Tucson, AZ
##                                    0.019867550
##                                   Savannah, GA
##                                    0.019801980
##                                      Flint, MI
##                                    0.019607843
##                    Fort Collins-Loveland, CO
##                                    0.019417476
##                                    Spokane, WA
##                                    0.019230769
##                                Las Cruses, NM
##                                    0.018691589
##                Pensacola-Ferry Pass-Brent, FL
##                                    0.018691589
##                                   Prescott, AZ
##                                    0.018518519
##                                   Columbus, OH
##                                    0.018148820
##                             Memphis, TN-MS-AR
##                                    0.017241379
##                    Panama City-Lynn Haven, FL
##                                    0.016949153
##                          Champaign-Urbana, IL
##                                    0.016393443
##                                       Napa, CA
##                                    0.016393443
##                          Colorado Springs, CO
##                                    0.016129032
##                                  Johnstown, PA
```

```
##                               0.015873016
##                     Kalamazoo-Portage, MI
##                               0.015748031
##                         Winston-Salem, NC
##                               0.015748031
##             Sarasota-Bradenton-Venice, FL
##                               0.015625000
##         Charlotte-Gastonia-Concord, NC-SC
##                               0.015473888
##                                 Dover, DE
##                               0.015350877
##                        Corpus Christi, TX
##                               0.015151515
##         Allentown-Bethlehem-Easton, PA-NJ
##                               0.014970060
##                                 Ocala, FL
##                               0.013157895
##            Youngstown-Warren-Boardman, OH
##                               0.013071895
##                            Provo-Orem, UT
##                               0.012944984
##                  Waterloo-Cedar Falls, IA
##                               0.012820513
##                     Birmingham-Hoover, AL
##                               0.012755102
##                           Springfield, MO
##                               0.012422360
##                               Greeley, CO
##                               0.012345679
##                               Medford, OR
##                               0.012195122
##                         Louisville, KY-IN
##                               0.011560694
##                   Harrisburg-Carlisle, PA
##                               0.011494253
##                              Kingston, NY
##                               0.011494253
##                     Boise City-Nampa, ID
##                               0.010869565
##                                Lawton, OK
##                               0.010309278
##              Cleveland-Elyria-Mentor, OH
##                               0.010279001
##                              Lawrence, KS
##                               0.010204082
##                         Evansville, IN-KY
##                               0.010101010
##                           Sioux Falls, SD
##                               0.010084034
##                  Grand Rapids-Wyoming, MI
##                               0.009868421
##                                Yakima, WA
##                               0.008928571
##                         Coeur d'Alene, ID
```

```
##                                0.008547009
##                          York-Hanover, PA
##                                0.008547009
##                                Toledo, OH
##                                0.008510638
##                   Santa Rosa-Petaluma, CA
##                                0.007751938
##       Santa Barbara-Santa Maria-Goleta, CA
##                                0.007575758
##                                Dayton, OH
##                                0.007462687
##                                  Bend, OR
##                                0.007142857
##                               Modesto, CA
##                                0.006329114
##                       Chattanooga, TN-GA
##                                0.005988024
##                                Monroe, LA
##                                0.005586592
##          Charleston-North Charleston, SC
##                                0.004310345
##                           San Antonio, TX
##                                0.003294893
##          New Orleans-Metairie-Kenner, LA
##                                0.002724796
##                          St. Louis, MO-IL
##                                0.002092050
##                                Albany, GA
##                                0.000000000
##                               Altoona, PA
##                                0.000000000
##                              Amarillo, TX
##                                0.000000000
##                              Anderson, IN
##                                0.000000000
##                               Appleton,WI
##                                0.000000000
##                             Asheville, NC
##                                0.000000000
##                        Barnstable Town, MA
##                                0.000000000
##                  Beaumont-Port Author, TX
##                                0.000000000
##                              Billings, MT
##                                0.000000000
##                            Binghamton, NY
##                                0.000000000
##                           Bloomington, IN
##                                0.000000000
##                         Bowling Green, KY
##                                0.000000000
##                       Canton-Massillon, OH
##                                0.000000000
##                            Charleston, WV
```

```
##                                           0.000000000
##                                             Chico, CA
##                                           0.000000000
##                                      Columbus, GA-AL
##                                           0.000000000
##                                           Decatur, IL
##                                           0.000000000
##                                            Durham, NC
##                                           0.000000000
##                                        Eau Claire, WI
##                                           0.000000000
##                                           El Paso, TX
##                                           0.000000000
##                                              Erie, PA
##                                           0.000000000
##                                        Farmington, NM
##                                           0.000000000
##                                          Florence, AL
##                                           0.000000000
##                        Hagerstown-Martinsburg, MD-WV
##                                           0.000000000
##                                        Huntsville, AL
##                                           0.000000000
##                                           Jackson, MI
##                                           0.000000000
##                                           Jackson, MS
##                                           0.000000000
##                                        Janesville, WI
##                                           0.000000000
##                                     Johnson City, TN
##                                           0.000000000
##                                            Joplin, MO
##                                           0.000000000
##                                  Kankakee-Bradley, IL
##                                           0.000000000
##                         Killeen-Temple-Fort Hood, TX
##                                           0.000000000
##                             Kingsport-Bristol, TN-VA
##                                           0.000000000
##                                         Knoxville, TN
##                                           0.000000000
##                                         Lafayette, LA
##                                           0.000000000
##                             Lansing-East Lansing, MI
##                                           0.000000000
##                                            Laredo, TX
##                                           0.000000000
##                    Leominster-Fitchburg-Gardner, MA
##                                           0.000000000
##                                          Longview, TX
##                                           0.000000000
##                                           Lubbock, TX
##                                           0.000000000
##                                         Lynchburg, VA
```

```
##                                   0.000000000
##                                   Macon, GA
##                                   0.000000000
##                                   Madera, CA
##                                   0.000000000
##                       McAllen-Edinburg-Pharr, TX
##                                   0.000000000
##                       Michigan City-La Porte, IN
##                                   0.000000000
##                                   Midland, TX
##                                   0.000000000
##                                   Monroe, MI
##                                   0.000000000
##                       Muskegon-Norton Shores, MI
##                                   0.000000000
##        Myrtle Beach-Conway-North Myrtle Beach, SC
##                                   0.000000000
##                          Niles-Benton Harbor, MI
##                                   0.000000000
##                                Ocean City, NJ
##                                   0.000000000
##                              Oshkosh-Neenah, WI
##                                   0.000000000
##                   Port St. Lucie-Fort Pierce, FL
##                                   0.000000000
##             Poughkeepsie-Newburgh-Middletown, NY
##                                   0.000000000
##                                    Pueblo, CO
##                                   0.000000000
##                               Punta Gorda, FL
##                                   0.000000000
##                                    Racine, WI
##                                   0.000000000
##                                   Reading, PA
##                                   0.000000000
##                                   Roanoke, VA
##                                   0.000000000
##                                  Rockford, IL
##                                   0.000000000
##                Saginaw-Saginaw Township North, MI
##                                   0.000000000
##                                     Salem, OR
##                                   0.000000000
##                                 Salisbury, MD
##                                   0.000000000
##                                  Santa Fe, NM
##                                   0.000000000
##                       Santa-Cruz-Watsonville, CA
##                                   0.000000000
##                         Scranton-Wilkes Barre, PA
##                                   0.000000000
##                      Shreveport-Bossier City, LA
##                                   0.000000000
##                      South Bend-Mishawaka, IN-MI
```

```
##                                                   0.000000000
##                                               Spartanburg, SC
##                                                   0.000000000
##                                          Springfield, MA-CT
##                                                   0.000000000
##                                             Springfield, OH
##                                                   0.000000000
##                                               St. Cloud, MN
##                                                   0.000000000
##                                             Tallahassee, FL
##                                                   0.000000000
##                                              Tuscaloosa, AL
##                                                   0.000000000
##                                             Utica-Rome, NY
##                                                   0.000000000
##                                                Valdosta, GA
##                                                   0.000000000
##                                             Vero Beach, FL
##                                                   0.000000000
##                                                Victoria, TX
##                                                   0.000000000
##                          Vineland-Millville-Bridgeton, NJ
##                                                   0.000000000
##                                                    Waco, TX
##                                                   0.000000000
##                                               Waterbury, CT
##                                                   0.000000000
##                                                  Wausau, WI
##                                                   0.000000000
```

```
# answer: 4
```

Determine which metropolitan area has the smallest proportion of interviewees who have received no high school diploma

```
sort(tapply(CPS$Education == "No high school diploma", CPS$MetroArea, mean, na.rm = T))
```

```
##                                               Iowa City, IA
##                                                  0.02912621
##                                           Bowling Green, KY
##                                                  0.03703704
##                                       Kalamazoo-Portage, MI
##                                                  0.05050505
##                                      Champaign-Urbana, IL
##                                                  0.05154639
##                                   Bremerton-Silverdale, WA
##                                                  0.05405405
##                                                Lawrence, KS
##                                                  0.05952381
##                                     Bloomington-Normal IL
##                                                  0.06060606
##                                           Jacksonville, NC
##                                                  0.06122449
##                                             Eau Claire, WI
```

```
##                                   0.06250000
##         Palm Bay-Melbourne-Titusville, FL
##                                   0.06666667
##                            Salisbury, MD
##                                   0.06779661
##                          Gainesville, FL
##                                   0.06896552
##                 Fort Collins-Loveland, CO
##                                   0.06936416
##                              Altoona, PA
##                                   0.07142857
##                              Madison, WI
##                                   0.07423581
##                          Tallahassee, FL
##                                   0.07500000
##                            Fargo, ND-MN
##                                   0.07902736
##              Albany-Schenectady-Troy, NY
##                                   0.07929515
##                           Ocean City, NJ
##                                   0.08000000
##                Lakeland-Winter Haven, FL
##                                   0.08130081
##                             Billings, MT
##                                   0.08280255
##                        Coeur d'Alene, ID
##                                   0.08333333
##       Burlington-South Burlington, VT
##                                   0.08394161
##                                Akron, OH
##                                   0.08421053
##                            Ann Arbor, MI
##                                   0.08695652
##                            Asheville, NC
##                                   0.08695652
##           Pensacola-Ferry Pass-Brent, FL
##                                   0.08695652
##                       Oshkosh-Neenah, WI
##                                   0.08823529
##                  Rochester-Dover, NH-ME
##                                   0.08928571
##                            Knoxville, TN
##                                   0.08965517
##                           Pittsburgh, PA
##                                   0.09060403
##                     Barnstable Town, MA
##                                   0.09090909
##       Bridgeport-Stamford-Norwalk, CT
##                                   0.09563758
##                            Johnstown, PA
##                                   0.09615385
##                    Austin-Round Rock, TX
##                                   0.09629630
##                            La Crosse, WI
```

```
##                                     0.09677419
##                                    Boulder, CO
##                                     0.09701493
##                 Charleston-North Charleston, SC
##                                     0.09890110
##                                  Fort Wayne, IN
##                                     0.09900990
##                                    Roanoke, VA
##                                     0.10169492
##                                   Prescott, AZ
##                                     0.10204082
##                       Santa Rosa-Petaluma, CA
##                                     0.10280374
##                             Evansville, IN-KY
##                                     0.10389610
##                                    Spokane, WA
##                                     0.10434783
##           Poughkeepsie-Newburgh-Middletown, NY
##                                     0.10559006
##             Tampa-St. Petersburg-Clearwater, FL
##                                     0.10579710
##                       Grand Rapids-Wyoming, MI
##                                     0.10612245
##                    Portland-South Portland, ME
##                                     0.10638298
##                                   Honolulu, HI
##                                     0.10739300
##                     Michigan City-La Porte, IN
##                                     0.10769231
##                       Eugene-Springfield, OR
##                                     0.11038961
##                  Boston-Cambridge-Quincy, MA-NH
##                                     0.11080485
##                                       Bend, OR
##                                     0.11111111
##                                 Vero Beach, FL
##                                     0.11428571
##                   Sarasota-Bradenton-Venice, FL
##                                     0.11464968
##         Fort Walton Beach-Crestview-Destin, FL
##                                     0.11475410
##                                      Flint, MI
##                                     0.11538462
##                                Cedar Rapids, IA
##                                     0.11564626
##         Minneapolis-St Paul-Bloomington, MN-WI
##                                     0.11638204
##             Portland-Vancouver-Beaverton, OR-WA
##                                     0.11657143
##      Washington-Arlington-Alexandria, DC-VA-MD-WV
##                                     0.11683748
##                                     Mobile, AL
##                                     0.11702128
##                     Scranton-Wilkes Barre, PA
```

```
##                                             0.11724138
##                                            Topeka, KS
##                                             0.11724138
##                                  Colorado Springs, CO
##                                             0.11764706
##                                           Olympia, WA
##                                             0.11764706
##                                       Reno-Sparks, NV
##                                             0.11764706
##                                           Appleton,WI
##                                             0.11827957
##                                          Santa Fe, NM
##                                             0.11904762
##          Virginia Beach-Norfolk-Newport News, VA-NC
##                                             0.11909651
##                   Allentown-Bethlehem-Easton, PA-NJ
##                                             0.11929825
##                                         Rochester, NY
##                                             0.12132353
##                         Seattle-Tacoma-Bellevue, WA
##                                             0.12168793
##                                   Kansas City, MO-KS
##                                             0.12172775
##                                              Napa, CA
##                                             0.12244898
##                                         Duluth, MN-WI
##                                             0.12264151
##                                         New Haven, CT
##                                             0.12354312
##                                  Canton-Massillon, OH
##                                             0.12371134
##                                      Fayetteville, NC
##                                             0.12500000
##                     San Luis Obispo-Paso Robles, CA
##                                             0.12500000
##                                      Worcester, MA-CT
##                                             0.12605042
##            Philadelphia-Camden-Wilmington, PA-NJ-DE
##                                             0.12717253
##                   Davenport-Moline-Rock Island, IA-IL
##                                             0.12727273
##                            Waterloo-Cedar Falls, IA
##                                             0.12800000
##                                            Pueblo, CO
##                                             0.12844037
##                                       Baton Rouge, LA
##                                             0.12871287
##                                            Racine, WI
##                                             0.12903226
##                                        Des Moines, IA
##                                             0.12944162
##                          Detroit-Warren-Livonia, MI
##                                             0.12964642
##                         Omaha-Council Bluffs, NE-IA
```

```
##                                             0.12972973
##                                          Richmond, VA
##                                             0.12990196
##                                          Savannah, GA
##                                             0.13013699
##                                           Danbury, CT
##                                             0.13043478
##                                       Bloomington, IN
##                                             0.13095238
##                                          Valdosta, GA
##                                             0.13157895
##                                            Wausau, WI
##                                             0.13157895
##                  Deltona-Daytona Beach-Ormond Beach, FL
##                                             0.13178295
##                                             Tulsa, OK
##                                             0.13178295
##                              Harrisburg-Carlisle, PA
##                                             0.13286713
##                             Las Vegas-Paradise, NV
##                                             0.13307985
##           Myrtle Beach-Conway-North Myrtle Beach, SC
##                                             0.13333333
##                                        Provo-Orem, UT
##                                             0.13366337
##                                          Anderson, IN
##                                             0.13461538
##                                             Chico, CA
##                                             0.13461538
##                                      St. Louis, MO-IL
##                                             0.13461538
##                              Niles-Benton Harbor, MI
##                                             0.13513514
##                                 Ogden-Clearfield, UT
##                                             0.13571429
##                                 Baltimore-Towson, MD
##                                             0.13583333
##                            Buffalo-Niagara Falls, NY
##                                             0.13684211
##                      Milwaukee-Waukesha-West Allis, WI
##                                             0.13693694
##                    Chicago-Naperville-Joliet, IN-IN-WI
##                                             0.13737734
##                                   Louisville, KY-IN
##                                             0.13785047
##                                         Lynchburg, VA
##                                             0.13793103
##                                            Peoria, IL
##                                             0.13829787
##                                       Sioux Falls, SD
##                                             0.13832200
##                                             Ocala, FL
##                                             0.13888889
##                       Leominster-Fitchburg-Gardner, MA
```

```
##                                  0.14035088
##                           Oklahoma City, OK
##                                  0.14137214
##            San Diego-Carlsbad-San Marcos, CA
##                                  0.14188267
##                           Jacksonville, FL
##                                  0.14244186
##                          Atlantic City, NJ
##                                  0.14285714
##                    Holland-Grand Haven, MI
##                                  0.14285714
##                                Medford, OR
##                                  0.14285714
##                    Naples-Marco Island, FL
##                                  0.14285714
##                            Punta Gorda, FL
##                                  0.14285714
##                               Victoria, TX
##                                  0.14285714
##                          Winston-Salem, NC
##                                  0.14285714
##                         Salt Lake City, UT
##                                  0.14338235
##           Atlanta-Sandy Springs-Marietta, GA
##                                  0.14421553
##                                Decatur, IL
##                                  0.14516129
##                            Springfield, IL
##                                  0.14516129
##                                 Monroe, MI
##                                  0.14545455
##                          Denver-Aurora, CO
##                                  0.14574558
##     Hartford-West Hartford-East Hartford, CT
##                                  0.14574899
##                                Greeley, CO
##                                  0.14615385
##            San Francisco-Oakland-Fremont, CA
##                                  0.14651368
##                      Boise City-Nampa, ID
##                                  0.14653465
##                             Greenville, SC
##                                  0.14666667
##                     Birmingham-Hoover, AL
##                                  0.14678899
##          Saginaw-Saginaw Township North, MI
##                                  0.14754098
##                Santa-Cruz-Watsonville, CA
##                                  0.14814815
##                          Trenton-Ewing, NJ
##                                  0.14814815
##                     Lexington-Fayette, KY
##                                  0.14838710
##           San Jose-Sunnyvale-Santa Clara, CA
```

```
##                                            0.14922481
##                                       Bellingham, WA
##                                            0.15000000
##                          Norwich-New London, CT-RI
##                                            0.15060241
##                                          Lubbock, TX
##                                            0.15094340
##                       Huntington-Ashland, WV-KY-OH
##                                            0.15151515
##                                        St. Cloud, MN
##                                            0.15151515
##                                          Jackson, MS
##                                            0.15168539
##                                           Dayton, OH
##                                            0.15207373
##                                  Chattanooga, TN-GA
##                                            0.15217391
##                                         Syracuse, NY
##                                            0.15428571
## New York-Northern New Jersey-Long Island, NY-NJ-PA
##                                            0.15573586
##                                         Columbia, SC
##                                            0.15600000
##                                         Columbus, OH
##                                            0.15617716
##                                 Memphis, TN-MS-AR
##                                            0.15714286
##                                          Orlando, FL
##                                            0.16108787
##                                 Warner Robins, GA
##                                            0.16216216
##                        Cleveland-Elyria-Mentor, OH
##                                            0.16250000
##                                         Columbia, MO
##                                            0.16279070
##                                           Durham, NC
##                                            0.16326531
##              Miami-Fort Lauderdale-Miami Beach, FL
##                                            0.16356589
##                                    Indianapolis, IN
##                                            0.16371681
##                                     Albuquerque, NM
##                                            0.16424116
##                           Cape Coral-Fort Myers, FL
##                                            0.16528926
##                                         Amarillo, TX
##                                            0.16666667
##                                Anniston-Oxford, AL
##                                            0.16666667
##                            Athens-Clark County, GA
##                                            0.16666667
##                                       Binghamton, NY
##                                            0.16666667
##                        Phoenix-Mesa-Scottsdale, AZ
```

```
##                                        0.16687737
##                                     Green Bay, WI
##                                        0.16831683
##                                        Bangor, ME
##                                        0.16860465
##               Providence-Fall River-Warwick, MA-RI
##                                        0.16915688
##                      Muskegon-Norton Shores, MI
##                                        0.16923077
##                                    Tuscaloosa, AL
##                                        0.16949153
##                                      Rockford, IL
##                                        0.17021277
##                                    Las Cruses, NM
##                                        0.17283951
##                             Gulfport-Biloxi, MS
##                                        0.17307692
##                                    Huntsville, AL
##                                        0.17391304
##                                    Utica-Rome, NY
##                                        0.17391304
##                                 Fort Smith, AR-OK
##                                        0.17441860
##                Charlotte-Gastonia-Concord, NC-SC
##                                        0.17444717
##                                     El Centro, CA
##                                        0.17567568
##                                          Erie, PA
##                                        0.17567568
##                                       Jackson, MI
##                                        0.17741935
##                 Cincinnati-Middletown, OH-KY-IN
##                                        0.17773788
##                               Springfield, MA-CT
##                                        0.17829457
##                                       Reading, PA
##                                        0.17857143
##                            Vallejo-Fairfield, CA
##                                        0.17924528
##                                         Salem, OR
##                                        0.17985612
##              Nashville-Davidson-Murfreesboro, TN
##                                        0.18112245
##                                  Johnson City, TN
##                                        0.18181818
##                                       Wichita, KS
##                                        0.18181818
##                                  York-Hanover, PA
##                                        0.18181818
##                                    Janesville, WI
##                                        0.18292683
##                         Lansing-East Lansing, MI
##                                        0.18348624
##                        Greensboro-High Point, NC
```

```
##                                  0.18357488
##                                Decatur, Al
##                                  0.18421053
##                                 Albany, GA
##                                  0.18604651
##             Augusta-Richmond County, GA-SC
##                                  0.18796992
##                             Charleston, WV
##                                  0.18834081
##                Shreveport-Bossier City, LA
##                                  0.18918919
##                           Raleigh-Cary, NC
##                                  0.18959108
##                                 Toledo, OH
##                                  0.18965517
##                            Spartanburg, SC
##                                  0.18987342
##            Dallas-Fort Worth-Arlington, TX
##                                  0.19077135
##   Sacramento-Arden-Arcade-Roseville, CA
##                                  0.19136961
##    Santa Barbara-Santa Maria-Goleta, CA
##                                  0.19191919
##                                 Monroe, LA
##                                  0.19205298
##                                  Dover, DE
##                                  0.19220056
##                South Bend-Mishawaka, IN-MI
##                                  0.19354839
##   Fayetteville-Springdale-Rogers, AR-MO
##                                  0.19393939
##                            Columbus, GA-AL
##                                  0.19607843
##                               Kingston, NY
##                                  0.19696970
##             Port St. Lucie-Fort Pierce, FL
##                                  0.19767442
##                              Waterbury, CT
##                                  0.19852941
##          Little Rock-North Little Rock, AR
##                                  0.19939577
##                            Springfield, MO
##                                  0.20000000
##                                Modesto, CA
##                                  0.20325203
##             Houston-Baytown-Sugar Land, TX
##                                  0.20439739
##           Oxnard-Thousand Oaks-Ventura, CA
##                                  0.20657277
##                               Anderson, SC
##                                  0.20689655
##                                Midland, TX
##                                  0.21052632
##            New Orleans-Metairie-Kenner, LA
```

```
##                                          0.21088435
##                                         Fresno, CA
##                                          0.21120690
##                                   Lake Charles, LA
##                                          0.21739130
##                            Visalia-Porterville, CA
##                                          0.21782178
##                                    San Antonio, TX
##                                          0.22004357
##                    Hagerstown-Martinsburg, MD-WV
##                                          0.22222222
##                                         Yakima, WA
##                                          0.22222222
##                    Hickory-Morgantown-Lenoir, NC
##                                          0.22448980
##           Los Angeles-Long Beach-Santa Ana, CA
##                                          0.22882883
##                       Panama City-Lynn Haven, FL
##                                          0.22916667
##                                  Harrisonburg, VA
##                                          0.23287671
##                              Kankakee-Bradley, IL
##                                          0.23437500
##                         Beaumont-Port Author, TX
##                                          0.23469388
##                  Youngstown-Warren-Boardman, OH
##                                          0.23622047
##                      Riverside-San Bernardino, CA
##                                          0.23780488
##                                    Farmington, NM
##                                          0.23913043
##                      Killeen-Temple-Fort Hood, TX
##                                          0.24050633
##                                           Waco, TX
##                                          0.24074074
##                                    Montgomery, AL
##                                          0.24137931
##                                         Tucson, AZ
##                                          0.24603175
##                                      Lafayette, LA
##                                          0.24822695
##                                         Joplin, MO
##                                          0.25000000
##                                       Stockton, CA
##                                          0.25333333
##                         Brownsville-Harlingen, TX
##                                          0.25396825
##                                      Lancaster, PA
##                                          0.26771654
##                                    Bakersfield, CA
##                                          0.27218935
##                 Vineland-Millville-Bridgeton, NJ
##                                          0.27500000
##                                         Lawton, OK
```

```
##                                        0.28000000
##                                        Merced, CA
##                                        0.28358209
##                                Corpus Christi, TX
##                                        0.29702970
##                                       El Paso, TX
##                                        0.30219780
##                                   Springfield, OH
##                                        0.31034483
##                                      Florence, AL
##                                        0.32075472
##                                        Madera, CA
##                                        0.33333333
##                                       Salinas, CA
##                                        0.34090909
##                                        Laredo, TX
##                                        0.34426230
##                          Kingsport-Bristol, TN-VA
##                                        0.36363636
##                                      Longview, TX
##                                        0.38297872
##                        McAllen-Edinburg-Pharr, TX
##                                        0.38297872
##                                         Macon, GA
##                                        0.40816327
```

4.1) Integrating Country of Birth Data

```r
CPS = merge(CPS, CountryMap, by.x = "CountryOfBirthCode", by.y = "Code", all.x = TRUE )

CPS %>% glimpse()
```

```
## Observations: 131,302
## Variables: 16
## $ CountryOfBirthCode <int> 57, 57, 57, 57, 57, 57, 57, 57, 57, 57, 57,...
## $ MetroAreaCode      <int> 10420, 71650, 10420, 10420, 10420, 10420, 1...
## $ PeopleInHousehold  <int> 2, 4, 5, 2, 2, 3, 1, 3, 4, 4, 1, 1, 1, 5, 4...
## $ Region             <chr> "Midwest", "Northeast", "Midwest", "Midwest...
## $ State              <chr> "Ohio", "New Hampshire", "Ohio", "Ohio", "O...
## $ Age                <int> 73, 5, 10, 30, 30, 0, 34, 32, 6, 9, 63, 25,...
## $ Married            <chr> "Married", NA, NA, "Married", "Married", NA...
## $ Sex                <chr> "Female", "Female", "Female", "Female", "Fe...
## $ Education          <chr> "Some college, no degree", NA, NA, "Associa...
## $ Race               <chr> "White", "White", "White", "White", "White"...
## $ Hispanic           <int> 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0...
## $ Citizenship        <chr> "Citizen, Native", "Citizen, Native", "Citi...
## $ EmploymentStatus   <chr> "Retired", NA, NA, "Employed", "Employed", ...
## $ Industry           <chr> NA, NA, NA, "Trade", "Mining", NA, "Constru...
## $ MetroArea          <chr> "Akron, OH", "Boston-Cambridge-Quincy, MA-N...
## $ Country            <chr> "United States", "United States", "United S...
```

How many interviewees have a missing value for the new country of birth variable?

```r
table(is.na(CPS$Country))
```

```
##
##  FALSE   TRUE
## 131126    176
```

Among all interviewees born outside of North America, which country was the most common place of birth?

```r
sort(table(CPS$Country), decreasing = TRUE) # Phillippines
```

```
##
##                      United States                             Mexico
##                             115063                               3921
##                        Philippines                              India
##                                839                                770
##                              China                        Puerto Rico
##                                581                                518
##                        El Salvador                            Vietnam
##                                477                                458
##                            Germany                               Cuba
##                                438                                426
##                             Canada                              Korea
##                                410                                334
##                 Dominican Republic                          Guatemala
##                                330                                309
##                            Jamaica                           Columbia
##                                217                                206
##                           Honduras                              Japan
##                                189                                187
##                            England                             Russia
##                                179                                173
##                              Haiti                             Poland
##                                167                                162
##                             Brazil                              Italy
##                                159                                149
##                               Iran                            Ecuador
##                                144                                136
##                               Peru             Africa, not specified
##                                136                                129
##                           Thailand                     United Kingdom
##                                128                                111
##                             Guyana                           Pakistan
##                                109                                109
##                            Ukraine                             Taiwan
##                                104                                102
##                               Laos                               Iraq
##                                 98                                 97
##                            Nigeria                          Elsewhere
##                                 85                                 81
##                           Ethiopia                              Ghana
##                                 80                                 76
##                          Nicaragua                             France
##                                 76                                 73
```

```
##                  South Korea                  Somalia
##                           73                       72
##                        Egypt                Argentina
##                           65                       64
##                    Hong Kong                 Portugal
##                           64                       64
##         Bosnia & Herzegovina                Venezuela
##                           61                       61
##          Trinidad and Tobago                   Israel
##                           60                       57
##                       Greece                    Kenya
##                           56                       55
##                      Romania                  Liberia
##                           55                       52
##                     Cambodia             South Africa
##                           49                       48
##                       Turkey                  Lebanon
##                           48                       45
##              Myanmar (Burma)                    Nepal
##                           45                       44
##                       Panama                Australia
##                           44                       43
##                   Bangladesh                    Spain
##                           42                       41
##          Asia, not specified                  Ireland
##                           39                       39
##                        Chile                   Jordan
##                           37                       36
##                      Armenia                 Cameroon
##                           35                       32
##                        Syria                     Guam
##                           32                       31
##                     Bulgaria               Costa Rica
##                           29                       29
##                 Saudi Arabia              Netherlands
##                           29                       28
##                       Sweden              Afghanistan
##                           28                       26
##                    Indonesia                  Hungary
##                           26                       25
##                      Belarus                 Scotland
##                           24                       24
##                   Yugoslavia              New Zealand
##                           24                       23
##                  Switzerland                    Yemen
##                           23                       23
##                       Azores                     USSR
##                           22                       22
##                     Malaysia                   Serbia
##                           20                       20
##        Europe, not specified               Uzbekistan
##                           19                       19
##  West Indies, not specified                  Albania
##                           19                       18
```

```
##                         Norway                         Austria
##                             18                              17
##                        Morocco                       Sri Lanka
##                             17                              17
##            U. S. Virgin Islands                         Uruguay
##                             17                              17
##                     Cape Verde                         Eritrea
##                             15                              15
##                   Sierra Leone                          Uganda
##                             15                              15
##             Antigua and Barbuda                         Belgium
##                             13                              13
##                        Bermuda                         Bolivia
##                             13                              13
##                        Grenada                           Sudan
##                             13                              13
##                        Croatia                       Macedonia
##                             12                              12
##                        Moldova                  Czech Republic
##                             12                              11
##                       Dominica                        Paraguay
##                             11                              11
##                        Bahamas                         Finland
##                             10                              10
##                         Kuwait                       Lithuania
##                             10                              10
##                        Algeria         Americas, not specified
##                              9                               9
##                         Belize                            Fiji
##                              9                               9
## St. Vincent and the Grenadines   South America, not specified
##                              9                               7
##                      St. Lucia                        Barbados
##                              7                               6
##                        Denmark                          Latvia
##                              6                               6
##                          Samoa                         Senegal
##                              6                               6
##                      Singapore                        Slovakia
##                              6                               6
##                          Tonga                        Zimbabwe
##                              6                               6
##                        Georgia                      Azerbaijan
##                              5                               3
##                 Czechoslovakia                St. Kitts--Nevis
##                              3                               3
##               Northern Ireland                        Tanzania
##                              2                               2
```

What proportion of the interviewees from the "New York-Northern New Jersey-Long Island, NY-NJ-PA" metropolitan area have a country of birth that is not the United States? For this computation, don't include people from this metropolitan area who have a missing country of birth.

```r
tapply(CPS$Country != "United States", CPS$MetroArea == "New York-Northern New Jersey-Long Island, NY-N
```

```
##     FALSE      TRUE
## 0.1392772 0.3086603
```

```r
# In this case, the group is the interviewees in the Metro Area "NY-NJ-LI", the summary is the intervie
```

Which metropolitan area has the largest number (note – not proportion) of interviewees with a country of birth in India? Hint – remember to include na.rm=TRUE if you are using tapply() to answer this question.

```r
table(CPS$Country == "India")
```

```
##
##  FALSE   TRUE
## 130356    770
```

```r
India =
  CPS %>%
  filter(Country == "India")

India %>%
  count(MetroArea) %>%
  arrange(desc(n))
```

```
## # A tibble: 84 x 2
##    MetroArea                                              n
##    <chr>                                              <int>
##  1 New York-Northern New Jersey-Long Island, NY-NJ-PA    96
##  2 <NA>                                                  60
##  3 Washington-Arlington-Alexandria, DC-VA-MD-WV          50
##  4 Philadelphia-Camden-Wilmington, PA-NJ-DE              32
##  5 Chicago-Naperville-Joliet, IN-IN-WI                   31
##  6 Detroit-Warren-Livonia, MI                            30
##  7 Atlanta-Sandy Springs-Marietta, GA                    27
##  8 San Francisco-Oakland-Fremont, CA                     27
##  9 Hartford-West Hartford-East Hartford, CT              26
## 10 Minneapolis-St Paul-Bloomington, MN-WI                23
## # ... with 74 more rows
```

```r
# Answer: New York-Northern New Jersey-Long Island, NY-NJ-PA
```

In Brazil?

```r
Brazil =
  CPS %>%
  filter(Country == "Brazil")

Brazil %>%
  group_by(!is.na(MetroArea)) %>%
  count(MetroArea) %>%
  arrange(desc(n))
```

```
## # A tibble: 50 x 3
## # Groups:   !is.na(MetroArea) [2]
```

```
##    `!is.na(MetroArea)` MetroArea                                          n
##    <lgl>               <chr>                                          <int>
##  1 FALSE               <NA>                                              20
##  2 TRUE                Boston-Cambridge-Quincy, MA-NH                    18
##  3 TRUE                Miami-Fort Lauderdale-Miami Beach, FL             16
##  4 TRUE                Los Angeles-Long Beach-Santa Ana, CA               9
##  5 TRUE                Washington-Arlington-Alexandria, DC-VA-MD-WV       8
##  6 TRUE                Bridgeport-Stamford-Norwalk, CT                    7
##  7 TRUE                New York-Northern New Jersey-Long Island, NY-~     7
##  8 TRUE                San Francisco-Oakland-Fremont, CA                  6
##  9 TRUE                Danbury, CT                                        5
## 10 TRUE                Davenport-Moline-Rock Island, IA-IL                4
## # ... with 40 more rows
```

```
# answer: Boston-Cambridge-Quincy, MA-NH
```

In Somalia

```
Somalia =
  CPS %>%
  filter(Country == "Somalia")

Somalia %>%
  count(MetroArea) %>%
  arrange(desc(n))
```

```
## # A tibble: 14 x 2
##    MetroArea                                n
##    <chr>                                <int>
##  1 Minneapolis-St Paul-Bloomington, MN-WI    17
##  2 <NA>                                     9
##  3 Phoenix-Mesa-Scottsdale, AZ              7
##  4 Seattle-Tacoma-Bellevue, WA              7
##  5 St. Cloud, MN                            7
##  6 Columbus, OH                             5
##  7 Fargo, ND-MN                             5
##  8 Burlington-South Burlington, VT          3
##  9 Portland-South Portland, ME              3
## 10 Portland-Vancouver-Beaverton, OR-WA      3
## 11 Houston-Baytown-Sugar Land, TX           2
## 12 Sioux Falls, SD                          2
## 13 Dayton, OH                               1
## 14 Richmond, VA                             1
```

```
# Answer: Minneapolis-St Paul-Bloomington, MN-WI
```