Working with the `SubnationalCRVS` R Package

Created by: Jeremy Roth

Last Updated: 22 July 2020

Contents

1 Setup

The `SubnationalCRVS` package, which is still under active development, is hosted on the GitHub page `www.github.com/ConVERGE-UNFPA/SubnationalCRVS` rather than on CRAN. As a result, `SubnationalCRVS` cannot be installed with the usual `install.packages()` function. Instead, `SubnationalCRVS` can be installed with the `install_github()` function from the `devtools` package. The key dependency `DemoTools` (Riffe et al. 2019) is also hosted on GitHub instead of CRAN and can also be installed with `install_github()`.

```
#install.packages("devtools") # install devtools (to be able to use install_github())
library(devtools)
#install_github("timriffe/DemoTools") # install the DemoTools dependency
#install_github("ConVERGE-UNFPA/SubnationalCRVS") # install the SubnationalCRVS package
```

Now we can load `SubnationalCRVS`, specify the name of the local folder in which we will save the plots produced by the package, and create that local folder if it does not already exist.

```r
library(SubnationalCRVS)
my_plots_dir <- "Plots/" # local folder where the plots should be saved
dir.create(my_plots_dir) # create the folder if it does not already exist
```

For this demonstration, we will also load the `dplyr` package to customize the display of some of the tables returned by `SunbationalCRVS`.

```r
library(dplyr)
```

2 Datasets Included with the Package: Tabulations from Ecuador

The `SubnationalCRVS` package comes with tabulations of population and registered deaths in Ecuador – disaggregated by age, sex, and province – created from publicly available datasets from Ecuador (Peralta et al. 2019; INEC 2010) based on the Ecuador's 2001 Census, 2010 Census, and annual counts of registered deaths from 2001 through 2010. The structure of the included example datasets is inspired by the data requirements of the `ddm()` function from the `DDM` package (Riffe, Lima, and Queiroz 2017).

The `ecuador_single_year_ages_combined` dataset reports the estimated populations in Ecuador by single-year ages (0, 1, 2, ... in the `age` column) from both the 2001 Census (`pop1` column) and the 2010 Census (`pop2` column), separately for males and females (`m` and `f` in the `sex` column) and province (in the `province` column).

```r
head(ecuador_single_year_ages)
```

```
##   province_name sex age pop1 pop2 year1 month1 day1 year2 month2 day2
## 1         Azuay   m   0 6086 6750  2001     11   25  2010     11   28
## 2         Azuay   m   1 6555 6984  2001     11   25  2010     11   28
## 3         Azuay   m   2 7232 7090  2001     11   25  2010     11   28
## 4         Azuay   m   3 7101 7095  2001     11   25  2010     11   28
## 5         Azuay   m   4 7083 6961  2001     11   25  2010     11   28
## 6         Azuay   m   5 6583 6895  2001     11   25  2010     11   28
```

The `ecuador_single_year_ages_combined` dataset appends rows to `ecuador_single_year_ages_combined` that report the sex- and single-year-age- disaggregated population estimates for the entire country (as opposed to a single province).

```r
ecuador_single_year_ages_combined %>% filter(province_name == "National") %>%
                                      head()
```

```
##   province_name sex age   pop1   pop2 year1 month1 day1 year2 month2 day2
## 1      National   m   0 117897 123013  2001     11   25  2010     11   28
## 2      National   m   1 133280 141602  2001     11   25  2010     11   28
## 3      National   m   2 141869 141719  2001     11   25  2010     11   28
## 4      National   m   3 134793 143418  2001     11   25  2010     11   28
## 5      National   m   4 139268 143830  2001     11   25  2010     11   28
## 6      National   m   5 131491 136479  2001     11   25  2010     11   28
```

The `ecuador_five_year_ages` and `ecuador_five_year_ages_combined` datasets have the same variables as `ecuador_single_year_ages_combined` and `ecuador_single_year_ages_combined` with two exceptions: (1) the `age` variable now represents five-year age groups (in the `age` column, with 0-4 coded as 0, 5-9 coded as 5, 10-14 coded as 10, etc.) instead of single-year ages; and (2) there is an additional column called `deaths` that reports the registered deaths collected between 2001 and 2010.

```r
head(ecuador_five_year_ages)
```

```
##   province_name sex age  pop1  pop2 deaths year1 month1 day1 year2 month2 day2
## 1         Azuay   m   0 34101 34886    772  2001     11   25  2010     11   28
## 2         Azuay   m   5 34996 36406    232  2001     11   25  2010     11   28
## 3         Azuay   m  10 34946 38125    223  2001     11   25  2010     11   28
## 4         Azuay   m  15 32387 37611    416  2001     11   25  2010     11   28
## 5         Azuay   m  20 25634 33665    480  2001     11   25  2010     11   28
## 6         Azuay   m  25 18606 28376    475  2001     11   25  2010     11   28
```

```r
head(ecuador_five_year_ages_combined)
```

```
##   province_name sex age  pop1  pop2 deaths year1 month1 day1 year2 month2 day2
## 1         Azuay   m   0 34101 34886    772  2001     11   25  2010     11   28
## 2         Azuay   m   5 34996 36406    232  2001     11   25  2010     11   28
## 3         Azuay   m  10 34946 38125    223  2001     11   25  2010     11   28
## 4         Azuay   m  15 32387 37611    416  2001     11   25  2010     11   28
## 5         Azuay   m  20 25634 33665    480  2001     11   25  2010     11   28
## 6         Azuay   m  25 18606 28376    475  2001     11   25  2010     11   28
```

```r
ecuador_five_year_ages_combined %>% filter(province_name == "National") %>%
                                    head()
```

```
##   province_name sex age   pop1   pop2 deaths year1 month1 day1 year2 month2
## 1      National   m   0 678280 743576  22338  2001     11   25  2010     11
## 2      National   m  10 679067 782559   4290  2001     11   25  2010     11
## 3      National   m  15 616725 712878   8293  2001     11   25  2010     11
## 4      National   m  20 569964 638042  13412  2001     11   25  2010     11
## 5      National   m  25 456230 585652  13301  2001     11   25  2010     11
## 6      National   m  30 422307 519493  11811  2001     11   25  2010     11
##   day2
## 1   28
## 2   28
## 3   28
## 4   28
## 5   28
## 6   28
```

3 Example of Subnational Analysis: Provinces in Ecuador

In this section, we describe the key functions of `SubnationalCRVS` in the context of visualizing outputs from a demographic data quality assessment (DDQA) and estimates of death registration completeness (Riffe, Lima, and Queiroz 2017) within provinces of Ecuador. Later in this tutorial, we also show how `SubnationalCRVS` provides corresponding visualizations for national-level results.

3.1 Demographic Data Quality Assessment (DDQA)

3.1.1 Sex ratios: PlotSexRatios()

A key step in this demographic data quality assessment (DDQA) process is to use the `PlotSexRatios()` function to compute and plot sex ratios within each combination of province, sex, and single-year age for the 2001 and 2010 data stored in `ecuador_single_year_ages_combined`.

To use `PlotSexRatios()`, we are required to provide our tabulated data frame in the `data` argument and a few additional required arguments that describe the variable names and values in `data`. These arguments and the expected format of the specified dataset are motivated by the data structure enforced in the `DDM` package (Riffe, Lima, and Queiroz 2017).

- `name.disaggregations` is the name of variable representing the subnational disaggregation (apart from sex, which is required,) in `data`. Here we specify `name.disaggregations="province_name"`.
- `name.sex` is the name of variable representing sex. Here we specify `name.sex="sex"`
- `name.age` is the name of variable representing age. Here we specify `name.age="age"`
- `name.year1` is the name of variable that provides the year of the earlier of the two time periods represented in `data`. Here we specify `name.year1="year1"` (the value of this variable, `"2001"` was the year of Ecuador's 2001 Census)
- `name.month1` is the name of variable that provides the month of the earlier of the two time periods represented in `data`. Here we specify `name.month1="month1"` (the value of this variable, `"11"` was the numerical month, November, of Ecuador's 2001 Census)
- `name.day1` is the name of variable that provides the day of the earlier of the two time periods represented in `data`. Here we specify `name.day1="day1"` (the value of this variable, `"25"` was the day of Ecuador's 2001 Census)
- `name.year2` is the name of variable that provides the year of the later of the two time periods represented in `data`. Here we specify `name.year2="year2"` (the value of this variable, `"2010"` was the year of Ecuador's 2010 Census)
- `name.month2` is the name of variable that provides the month of the later of the two time periods represented in `data`. Here we specify `name.month2="month2"` (the value of this variable, `"11"` was the numerical month, November, of Ecuador's 2010 Census)
- `name.day2` is the name of variable that provides the day of the later of the two time periods represented in `data`. Here we specify `name.day2="day2"` (the value of this variable, `"28"` was the day of Ecuador's 2010 Census)
- `name.population.year1` is the name of variable representing the population in the earlier of the two time periods represented in the dataset. Here we specify `name.population.year2="pop1"`
- `name.population.year2` is the name of variable representing the population in the earlier of the two time periods represented in the dataset. Here we specify `name.population.year2="pop2"`
- `name.male` is the name of value of the `name.sex` variable that represents males. Here we specify `name.males="m"`
- `name.female` is the name of value of the `name.sex` variable that represents females. Here we specify `name.females="f"`

In addition, we specify two optional arguments, `plots.dir` and `name.national`. We set `plots.dir=my_plots_dir` so that the plots will save in the local folder `Plots`; if the `plots.dir` argument is omitted, the plots will be saved in the same working directory of the R script. We also set `name.national="National"` to reflect the fact that national-level datasets are provided in rows where the `name.disaggregations` variable has the value `"National"`. Specifying `name.national` produces visualizations catered specifically to national-level analysis (instead of accommodated further disaggregation) that will be presented later in the tutorial. If `name.national` is not provided, those national-level visualizations are simply not produced.

```r
s <- PlotSexRatios(data=ecuador_single_year_ages_combined,
                   name.disaggregations="province_name",
                   name.males="m",
                   name.females="f",
                   name.age="age",
                   name.sex="sex",
                   name.population.year1="pop1",
                   name.population.year2="pop2",
                   name.year1="year1",
                   name.month1="month1",
                   name.day1="day1",
                   name.year2="year2",
                   name.month2="month2",
                   name.day2="day2",
                   name.national="National",
                   label.subnational.level="Province",
                   plots.dir=my_plots_dir)
```

The plots of sex ratios are saved in the `Plots/` sub-folder we specified with the argument `plots.dir=my_plots_dir`; `plots.dir` is an optional argument and, if we do not specify a value for it, the plots will be saved the working directory. We also specified `label.subnational.level="Province"` so that the disaggregations are labeled `Province` instead of the less clear `province_name`.

The sex ratios for all levels of subnational disaggregation are overlaid in the following "combined" plots separately for each data year.

```r
knitr::include_graphics(path=paste0(my_plots_dir,
                                    "sex_ratios_combined_province_name_",
                                    Sys.Date(),
                                    ".pdf"))
```

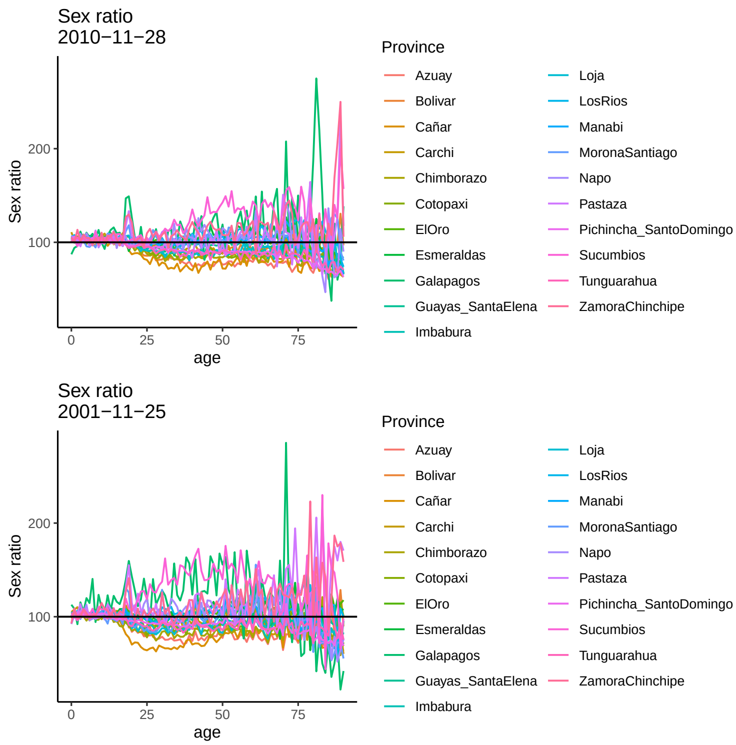

Figure 1: Sex ratios in Ecuador by province, combined plot

Additionally, the sex ratios are plotted in separate figures for each level of subnational disaggregation in the following "disaggregated" plots.

```
knitr::include_graphics(path=paste0(my_plots_dir,
                                    "sex_ratios_by_province_name_",
                                    Sys.Date(),
                                    ".pdf"))
```

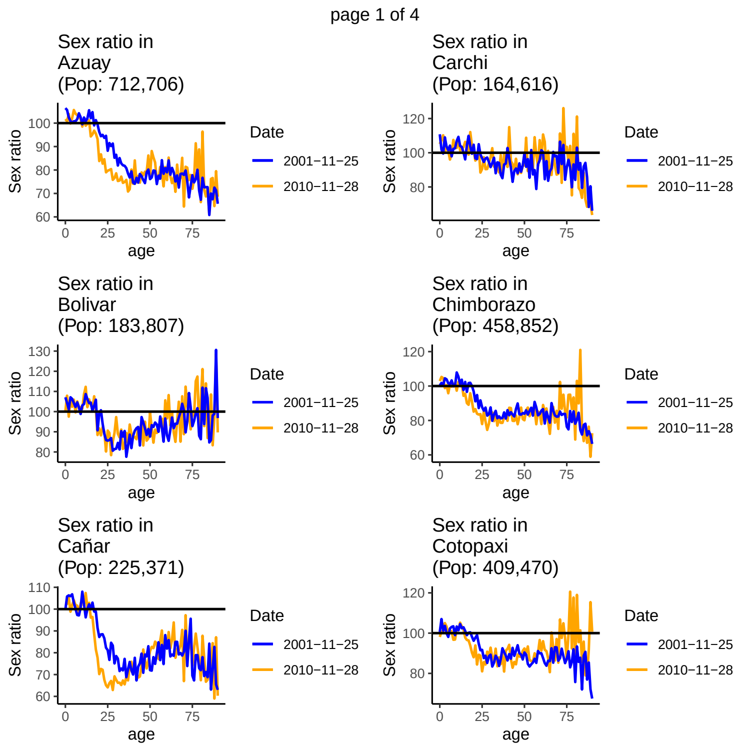

Figure 2: Sex ratios in Ecuador by province, disaggregated plots (only first page shown)

3.1.2 View sex ratios in table

The object returned by `PlotSexRatios()` is a table that shows us the sex ratios for each combination of province, sex, and single-year age in the `sex_ratio_1` column (for the 2001 Census) and the `sex_ratio_2` column (for the 2010 Census).

```
s %>% select(province_name, age, pop1, pop2, sex_ratio_1, sex_ratio_2) %>%
      head()
```

```
##   province_name age  pop1  pop2 sex_ratio_1 sex_ratio_2
## 1         Azuay   0 12073 13092       101.7       106.4
## 2         Azuay   1 13060 13596       100.8       105.6
## 3         Azuay   2 14195 14014       103.9       102.4
## 4         Azuay   3 14217 14124        99.8       100.9
## 5         Azuay   4 14012 13911       102.2       100.2
## 6         Azuay   5 12815 13736       105.6       100.8
```

3.1.3 Age ratios: PlotAgeRatios()

Another step in our demographic data quality assessment is using the `PlotAgeRatios()` function in the `SubnationalCRVS` package to compute and plot age ratios within each combination of province and sex for the 2001 and 2010 data stored in `ecuador_five_year_ages_combined`. The arguments we provide to `PlotAgeRatios()` are actually identical to those we specified for `PlotSexRatios()`, except now we are using the tabulation with five-year age groups (`ecuador_five_year_ages_combined`) instead of the tabulation with single-year ages.

```
a <- PlotAgeRatios(data=ecuador_five_year_ages_combined,
              name.disaggregations="province_name",
              name.males="m",
              name.females="f",
              name.age="age",
              name.sex="sex",
              name.population.year1="pop1",
              name.population.year2="pop2",
              name.year1="year1",
              name.month1="month1",
              name.day1="day1",
              name.year2="year2",
              name.month2="month2",
              name.day2="day2",
              label.subnational.level="Province",
              name.national="National",
              plots.dir=my_plots_dir)
```

3.1.4 View age ratios in combined plot

The following "combined" plots, saved in the `Plots/` folder, show the age ratios for all levels of subnational disaggregation, separately for males and females in each data year.

```
knitr::include_graphics(path=paste0(my_plots_dir,
                                    "age_ratios_combined_province_name_",
                                    Sys.Date(),
                                    ".pdf"))
```

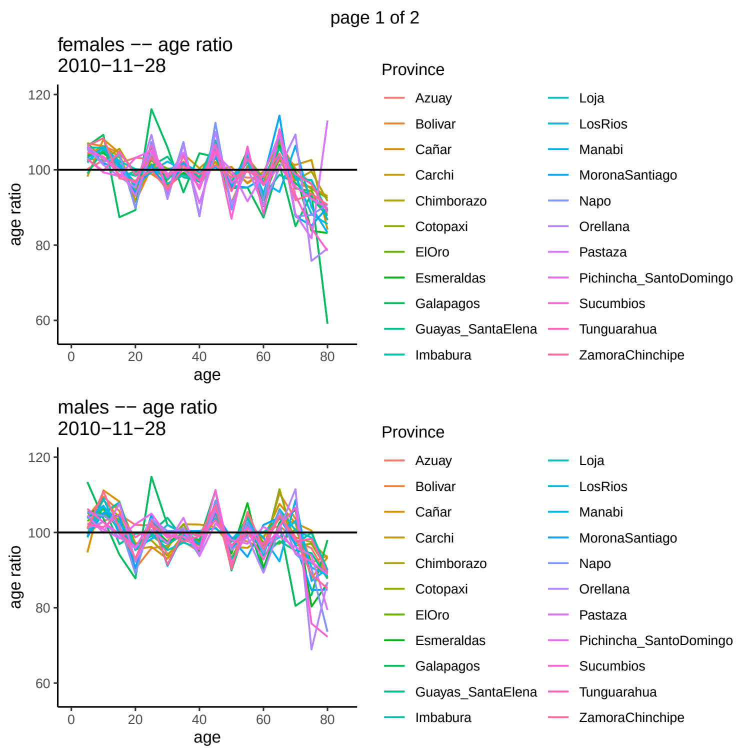

Figure 3: Age ratios in Ecuador by province, combined plot

In addition, `PlotAgeRatios()` also creates the following "disaggregated" plots, saved in the `Plots/` folder, where the age ratios for each level of disaggregation are shown in separate plots, with different sexes and data years overlaid within each plot.

```
knitr::include_graphics(path=paste0(my_plots_dir,
                                    "age_ratios_by_province_name_",
                                    Sys.Date(),
                                    ".pdf"))
```

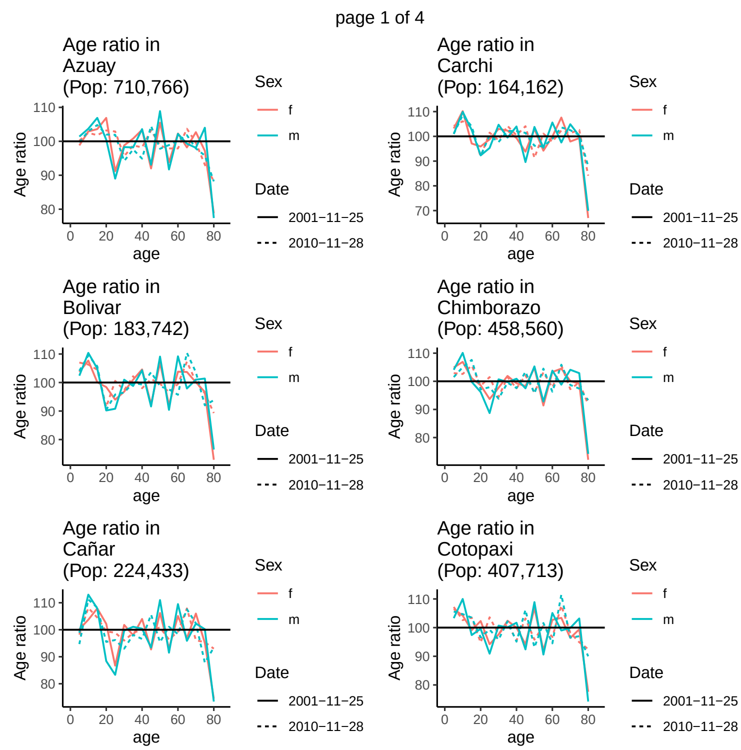

Figure 4: Age ratios in Ecuador by province, disaggregated plots (only first page shown)

Just as the `PlotSexRatios()` function returns a table of disaggregated sex ratios, `PlotAgeRatios()` returns a table of disaggregated age ratios.

```
a %>% select(province_name, age, pop1, pop2, age_ratio_1, age_ratio_2) %>%
      head()
```

```
##   province_name age  pop1  pop2 age_ratio_1 age_ratio_2
## 1         Azuay   0 33491 33876          NA          NA
## 2         Azuay   5 33817 35701        98.8       100.2
## 3         Azuay  10 34975 37366       102.9       102.5
## 4         Azuay  15 34181 37215       103.6       101.8
## 5         Azuay  20 31000 35753       106.9       103.2
## 6         Azuay  25 23844 32054        91.2       102.9
```

3.1.5 Potential age heaping: PlotPotentialAgeHeaping()

To give us a sense of whether "age-heaping" is occurring within the levels of disaggregation present in our `ecuador_single_year_ages_combined` dataset, we turn to the `PlotPotentialAgeHeaping()` function and actually provide the same arguments we used in the `PlotSexRatios()` function.

```
PlotPotentialAgeHeaping(data=ecuador_single_year_ages_combined,
            name.disaggregations="province_name",
              name.males="m",
              name.females="f",
              name.age="age",
              name.sex="sex",
```

```
            name.population.year1="pop1",
            name.population.year2="pop2",
            name.year1="year1",
            name.month1="month1",
            name.day1="day1",
            name.year2="year2",
            name.month2="month2",
            name.day2="day2",
            name.national="National",
            label.subnational.level="Province",
            plots.dir=my_plots_dir)
```

3.1.6 View potential age heaping in combined plot

The following "combined" plots show us estimated population counts by single-year ages with different provinces represented with different overlaid colors, and separate plots for each sex and data year.

```
knitr::include_graphics(path=paste0(my_plots_dir,
                                    "potential_age_heaping_combined_province_name_",
                                    Sys.Date(),
                                    ".pdf"))
```

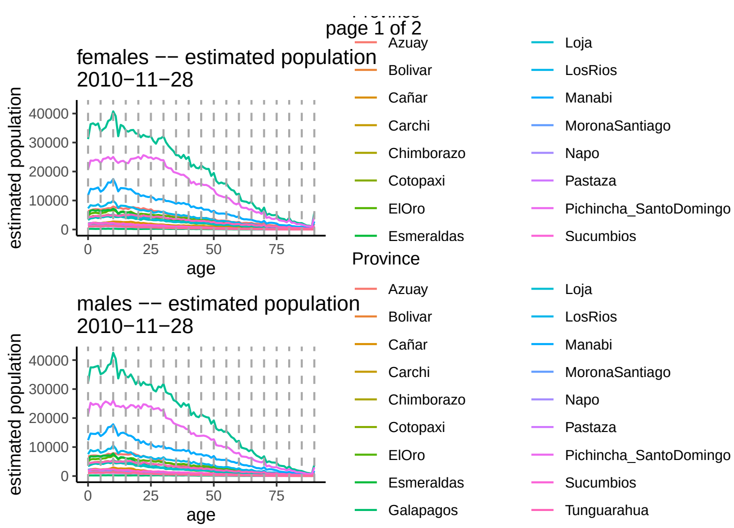

Figure 5: Population counts in Ecuador by single-year age, combined plot

`PlotPotentialAgeHeaping()` present separate plots of population counts for each province in Ecuador, with different sexes and data years overlaid within each plot.

```
knitr::include_graphics(path=paste0(my_plots_dir,
                                    "potential_age_heaping_by_province_name_",
                                    Sys.Date(),
                                    ".pdf"))
```

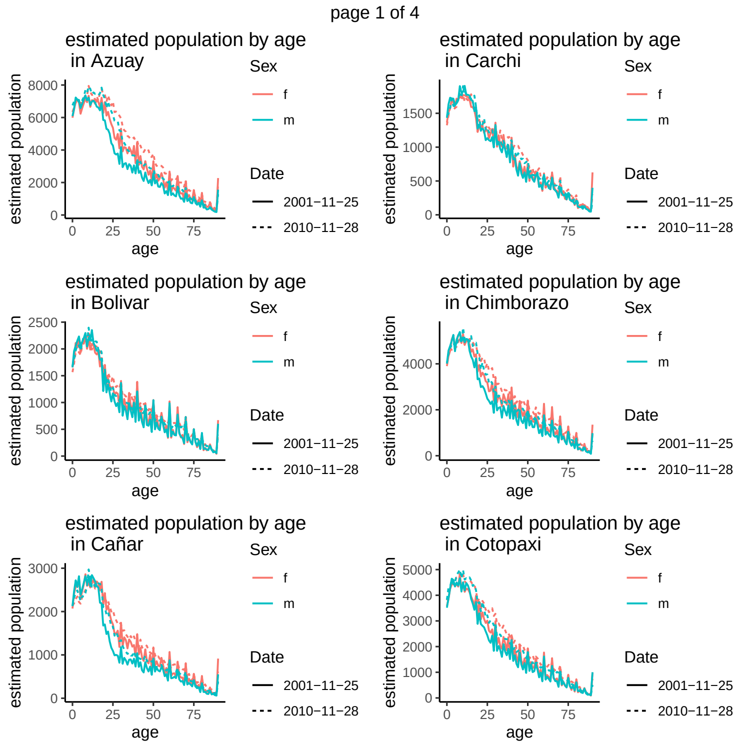

Figure 6: Population counts in Ecuador by single-year age, disaggregated plots

3.1.7 Age heaping indices: PlotAgeHeapingScores()

As a more concise summary of potential age-heaping suggested by the visualizations from `PlotPotentialAgeHeaping`, we now use the `PlotAgeHeapingScores` function with the same arguments we provided to the `PlotAgeRatios` function.

```
ageheaping <- PlotAgeHeapingScores(data=ecuador_single_year_ages_combined,
                                   name.disaggregations="province_name",
                                   name.males="m",
                                   name.females="f",
                                   name.age="age",
                                   name.sex="sex",
                                   name.population.year1="pop1",
                                   name.population.year2="pop2",
                                   name.year1="year1",
                                   name.month1="month1",
                                   name.day1="day1",
                                   name.year2="year2",
                                   name.month2="month2",
                                   name.day2="day2",
                                   name.national="National",
                                   label.subnational.level="Province",
                                   plots.dir=my_plots_dir)
```

One set of plots returned by `PlotAgeHeapingScores()` shows the values of three age-heaping indices within each combination of province, sex, and data year: Roughness, Whipple, and Myers. The indices are computed with the `check_heaping_roughness()`, `check_heaping_Whipple()`, and `check_heaping_myers()` functions, respectively, from the `DemoTools` package (Riffe et al. 2019).

```
knitr::include_graphics(path=paste0(my_plots_dir,
                                    "age_heaping_scores_combined_province_name_",
                                    Sys.Date(),
                                    ".pdf"))
```

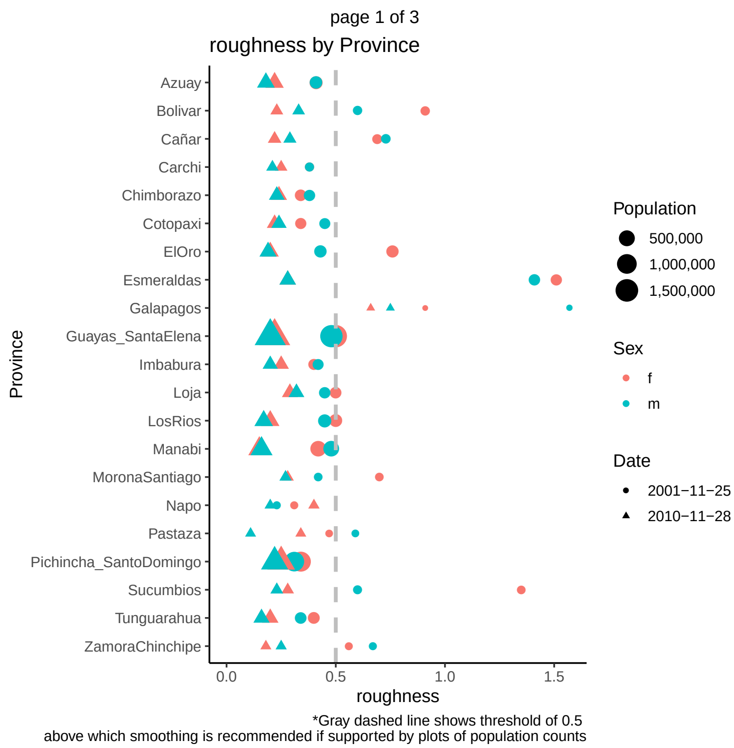

Figure 7: Roughness, Whipple, and Myers indices in Ecuador by province (only first page shown)

Further, `PlotAgeHeapingScores()` also creates plots showing the Noumbissi index in each province computed for each terminal digit from 0-9, using the `check_heaping_Noumbissi()` function from the `DemoTools` package (Riffe et al. 2019).

```
knitr::include_graphics(path=paste0(my_plots_dir,
                                    "Noumbissi_by_province_name_",
                                    Sys.Date(),
                                    ".pdf"))
```

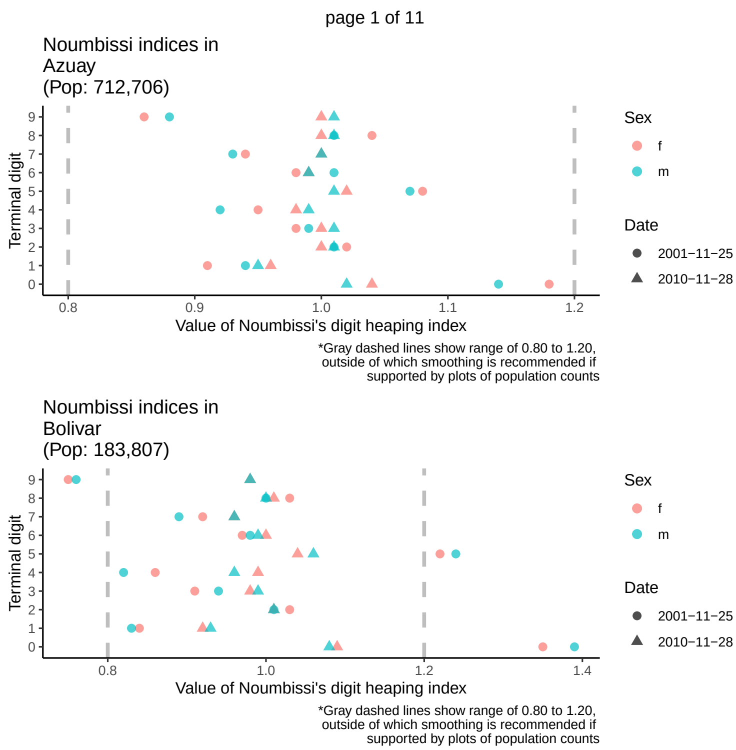

Figure 8: Noumbissi indices in Ecuador by province (only first page shown)

We can also view the age-heaping indices in the table returned by `PlotAgeHeapingScores()`

```
head(ageheaping)
```

```
##   province_name       date total_pop sex roughness Whipple Myers Noumbissi_0
## 1         Azuay 2001-11-25    319983   f      0.41    1.18  4.21        1.18
## 2       Bolivar 2001-11-25     86256   f      0.91    1.37  7.39        1.35
## 3         Cañar 2001-11-25    112041   f      0.69    1.22  4.89        1.22
## 4        Carchi 2001-11-25     77172   f      0.38    1.18  3.75        1.15
## 5    Chimborazo 2001-11-25    213106   f      0.34    1.25  5.44        1.23
## 6      Cotopaxi 2001-11-25    180328   f      0.34    1.27  5.99        1.25
##   Noumbissi_1 Noumbissi_2 Noumbissi_3 Noumbissi_4 Noumbissi_5 Noumbissi_6
## 1        0.91        1.02        0.98        0.95        1.08        0.98
## 2        0.84        1.03        0.91        0.86        1.22        0.97
## 3        0.90        1.02        0.98        0.91        1.11        0.98
## 4        0.91        1.05        0.96        0.92        1.11        0.98
## 5        0.88        1.01        0.97        0.89        1.12        0.98
## 6        0.88        1.02        0.95        0.90        1.18        0.98
##   Noumbissi_7 Noumbissi_8 Noumbissi_9
## 1        0.94        1.04        0.86
## 2        0.92        1.03        0.75
## 3        0.95        1.03        0.80
## 4        0.96        1.04        0.86
## 5        0.94        1.05        0.82
## 6        0.91        1.06        0.80
```

3.2 DDM Estimation of Death Registration Completeness

The structure of the `ecuador_five_year_ages` dataset is inspired by the requirements for the `ddm()` function from the `DDM` package (Riffe, Lima, and Queiroz 2017), which uses established Death Distribution Methods (DDM) to estimate death registration completeness for adults (aged 15+) between two consecutive Censuses (Moultrie et al. 2013). Essentially, the `SubnationalCRVS` package offers its `EstimateDDM()` function only as a convenient wrapper to `DDM::ddm()` to perform DDM estimation of death registration completeness using the same pipeline used to perform the DDQA.

3.2.1 Compute DDM estimates: EstimateDDM()

```
ddm_results <- EstimateDDM(data=ecuador_five_year_ages_combined,
            name.disaggregations="province_name",
            name.age="age",
            name.sex="sex",
            name.males="m",
            name.females="f",
            name.year1="year1",
            name.month1="month1",
            name.day1="day1",
            name.year2="year2",
            name.month2="month2",
            name.day2="day2",
            name.population.year1="pop1",
            name.population.year2="pop2",
            name.national="National",
            name.deaths="deaths",
            deaths.summed=TRUE)
```

```
## [1] "estimating completeness of adult death registration completeness with
GGB, SEG, and GGB-SEG methods within each of 20 possible age ranges within each
of 23 levels of subnational disaggregations..."
```

We note `EstimateDDM()` uses exactly the same arguments as `PlotAgeRatios()`, for example, from the DDQA with two additional required arguments:

- `name.deaths`, which provides the name of the variable representing the count of registered deaths between the two dates represented in `name.date1` and `name.date2`
- `deaths.summed`, which should be set to TRUE when the `name.deaths` variable represents the total number of registered deaths `name.date1` and `name.date2` and set to FALSE when the `name.deaths` variable represents the average number of registered deaths between the two dates.

3.2.2 Plot DDM estimates: PlotDDM()

We can plot the estimated adult death registration completeness (using the "hybrid" GGB-SEG method) with the `PlotDDM` function:

```
PlotDDM(ddm_results=ddm_results,
        label.subnational.level="Province",
        plots.dir=my_plots_dir)
```

```r
knitr::include_graphics(path=paste0(my_plots_dir,
                                    "ggbseg_point_estimates_combined_province_name_",
                                    Sys.Date(),
                                    ".pdf"))
```

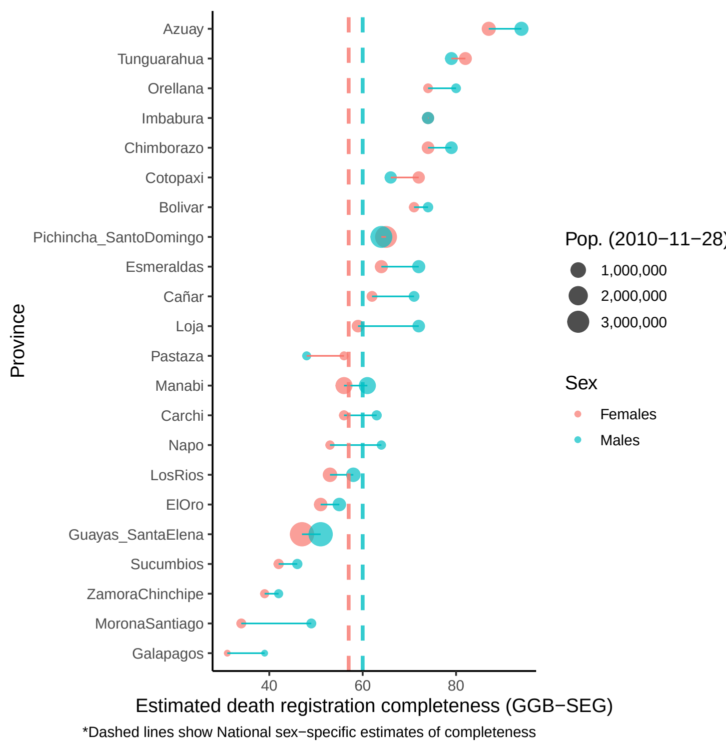

Figure 9: Point estimates of death registration completeness in Ecuador from 2001-2010, using the GGB-SEG method

The `EstimateDDM` function also returns a list, in which the `ddm_estimates` object contains the estimated death registration completeness with the GGB, SEG, and GGB-SEG methods (Moultrie et al. 2013).

```r
head(ddm_results$ddm_estimates)
```

```
##       cod     sex ggbseg  ggb  seg lower_age_range upper_age_range total_pop1
## 1   Azuay Females   0.87 0.99 0.82              15              50     599313
## 2   Azuay   Males   0.94 1.10 0.97              15              50     599313
## 3 Bolivar Females   0.71 0.99 0.72              20              60     170696
## 4 Bolivar   Males   0.74 0.96 0.80              25              60     170696
## 5   Cañar Females   0.62 1.00 0.58              20              55     206346
## 6   Cañar   Males   0.71 0.95 0.79              15              50     206346
##   total_pop2
## 1     710766
## 2     710766
## 3     183742
## 4     183742
## 5     224433
## 6     224433
```

Additionally, `PlotDDM` presents a visualization of the point estimates of completeness and the corresponding root mean squared errors (RMSEs) for each of the permitted values of the age range that is selected `DDM::ddm()` as part of its fitting procedure.

```
knitr::include_graphics(path=paste0(my_plots_dir,
                                    "ggbseg_sensitivity_province_name_",
                                    Sys.Date(),
                                    ".pdf"))
```

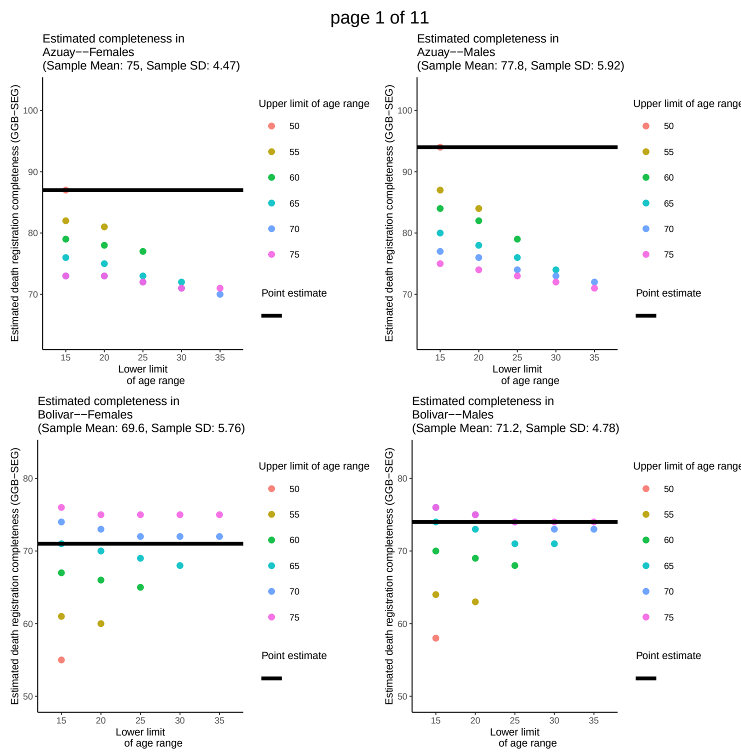

Figure 10: Sensitivity of point estimates of death registration completeness in Ecuador from 2001-2010 to choice of age-range parameter in the GGB-SEG method

```
knitr::include_graphics(path=paste0(my_plots_dir,
                                    "ggbseg_RMSE_ggb_province_name_",
                                    Sys.Date(),
                                    ".pdf"))
```

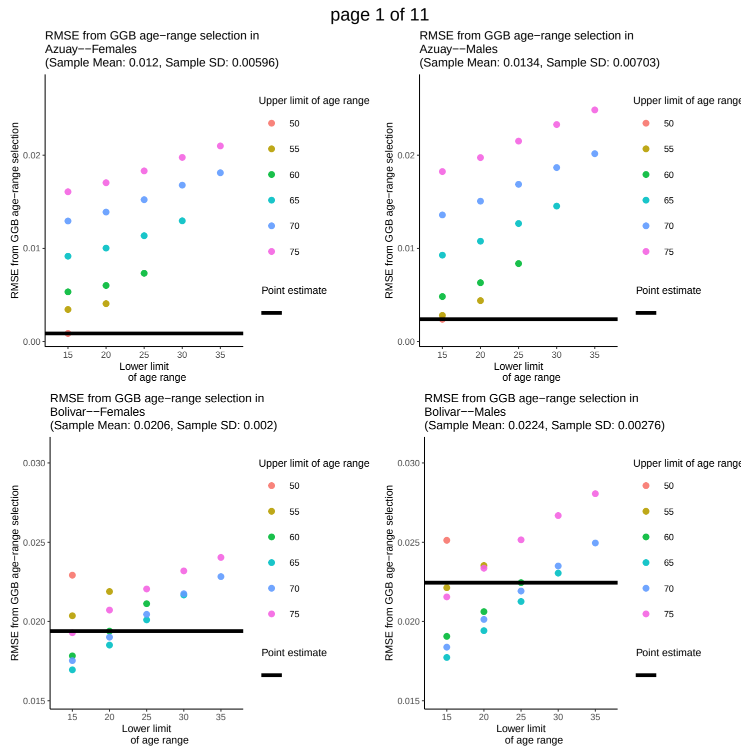

Figure 11: Sensitivity of point estimates of death registration completeness in Ecuador from 2001-2010 to choice of age-range parameter in the GGB-SEG method

The sensitivity estimates are also returned in table by `EstimateDDM`, in the `sensitivity_ddm_estimates` element of its list.

```
head(ddm_results$sensitivity_ddm_estimates, n=5)
```

```
##     cod     sex ggbseg  ggb  seg total_pop1 total_pop2 lower_age_range
## 1 Azuay Females   0.87 0.99 0.82     599313     710766              15
## 2 Azuay Females   0.82 0.87 0.82     599313     710766              15
## 3 Azuay Females   0.81 0.86 0.82     599313     710766              20
## 4 Azuay Females   0.79 0.83 0.82     599313     710766              15
## 5 Azuay Females   0.78 0.82 0.82     599313     710766              20
##   upper_age_range  RMSE_ggb selected_age_range_ggb
## 1              50 0.0008509                   TRUE
## 2              55 0.0034260                  FALSE
## 3              55 0.0040540                  FALSE
## 4              60 0.0053210                  FALSE
## 5              60 0.0060060                  FALSE
```

4 Example of National Analysis: Ecuador

As discussed, specifying the argument `name.national="National"` during the previous DDQA and estimation of death registration completeness permitted `SubnationalCRVS` to save plots specifically designed to display national-level (as opposed to subnational-level) visualizations. These national-level plots are can also be produced by specifying a national-level-only dataset, for example `ecuador_five_year_ages_national` (shown below), as long as the corresponding `name.national` argument is also specified (e.g. `name.national="National"` here)

```
head(ecuador_five_year_ages_national)
```

```
##   province_name sex age   pop1   pop2 deaths year1 month1 day1 year2 month2
## 1      National   m   0 678280 743576  22338  2001     11   25  2010     11
## 2      National   m  10 679067 782559   4290  2001     11   25  2010     11
## 3      National   m  15 616725 712878   8293  2001     11   25  2010     11
## 4      National   m  20 569964 638042  13412  2001     11   25  2010     11
## 5      National   m  25 456230 585652  13301  2001     11   25  2010     11
## 6      National   m  30 422307 519493  11811  2001     11   25  2010     11
##   day2
## 1   28
## 2   28
## 3   28
## 4   28
## 5   28
## 6   28
```

```
tail(ecuador_five_year_ages_national)
```

```
##    province_name sex age   pop1   pop2 deaths year1 month1 day1 year2 month2
## 31      National   f  60 149508 203789  12500  2001     11   25  2010     11
## 32      National   f  65 126310 166559  14635  2001     11   25  2010     11
## 33      National   f  70  99469 123647  18139  2001     11   25  2010     11
## 34      National   f  75  73819  86449  21616  2001     11   25  2010     11
## 35      National   f  80  52335  62304  23629  2001     11   25  2010     11
## 36      National   f  85  72515  55087  52329  2001     11   25  2010     11
##    day2
## 31   28
## 32   28
## 33   28
## 34   28
## 35   28
## 36   28
```

4.1 Plot of sex ratios

```
knitr::include_graphics(path=paste0(my_plots_dir,
                                    "sex_ratios_national_",
                                    Sys.Date(),
                                    ".pdf"))
```

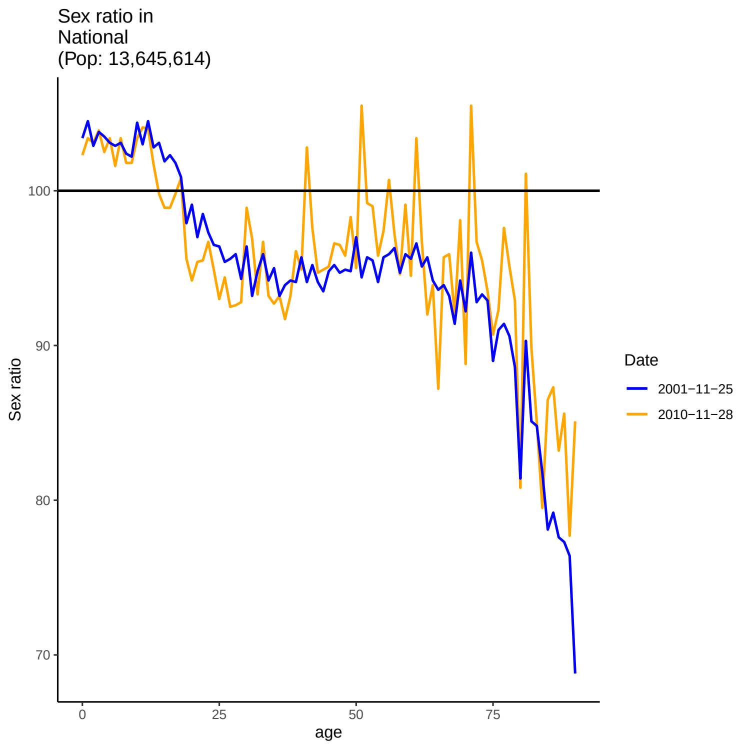

Figure 12: Sex ratios in Ecuador

4.2 Plot of age ratios

```
knitr::include_graphics(path=paste0(my_plots_dir,
                                    "age_ratios_national_",
                                    Sys.Date(),
                                    ".pdf"))
```

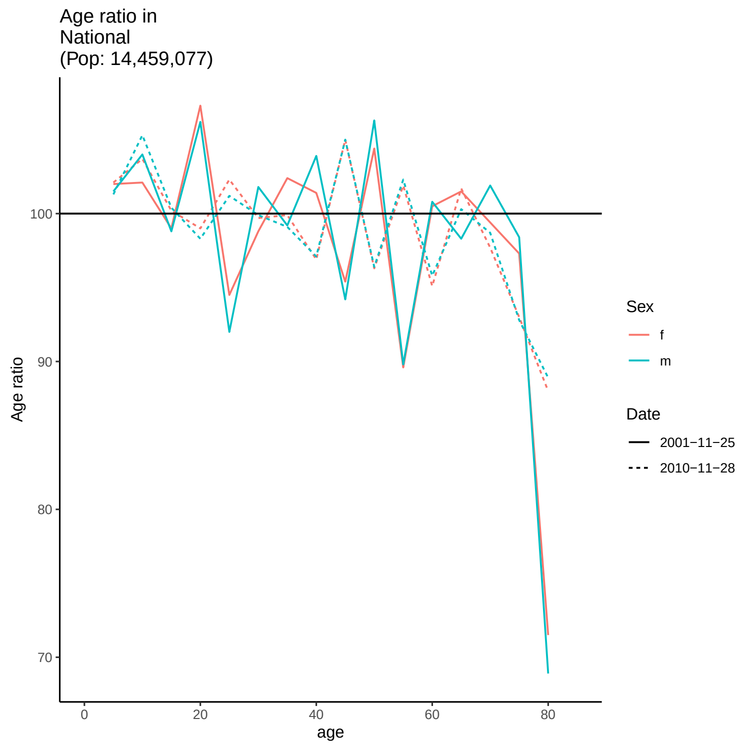

Figure 13: Age ratios in Ecuador

4.3 Plots related to age heaping

4.3.1 Plot of potential age heaping

```
knitr::include_graphics(path=paste0(my_plots_dir,
                                    "potential_age_heaping_National_",
                                    Sys.Date(),
                                    ".pdf"))
```

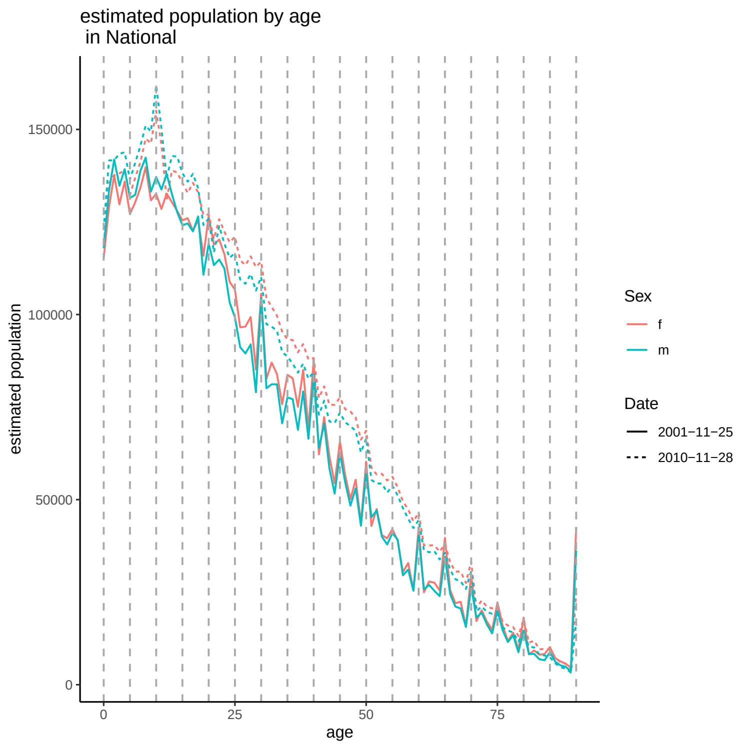

Figure 14: Population counts in Ecuador by single-year age

4.3.2 Plot of age heaping indices

```
knitr::include_graphics(path=paste0(my_plots_dir,
                                    "age_heaping_scores_National_",
                                    Sys.Date(),
                                    ".pdf"))
```

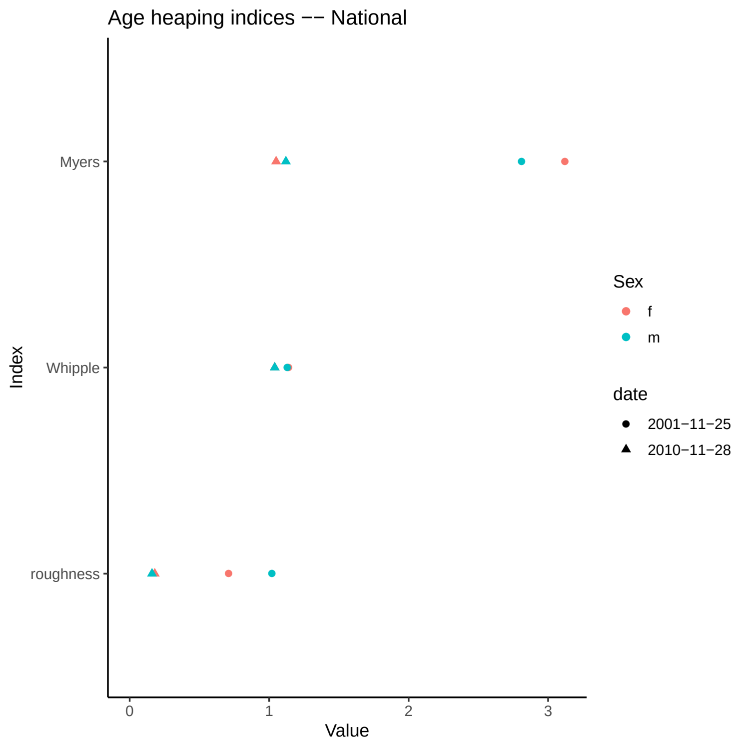

Figure 15: Roughness, Whipple, and Myers indices in Ecuador

4.3.3 Plot of Noumbissi indices

```
knitr::include_graphics(path=paste0(my_plots_dir,
                                    "Noumbissi_National_",
                                    Sys.Date(),
                                    ".pdf"))
```

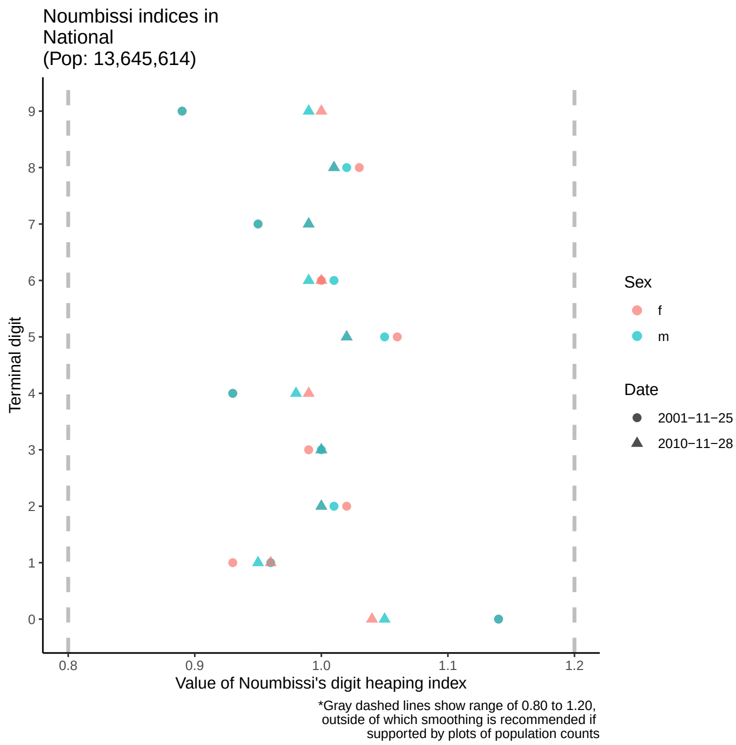

Figure 16: Noumbissi indices in Ecuador

4.4 Plot of DDM estimates

```
knitr::include_graphics(path=paste0(my_plots_dir,
                                    "ggbseg_sensitivity_National_",
                                    Sys.Date(),
                                    ".pdf"))
```

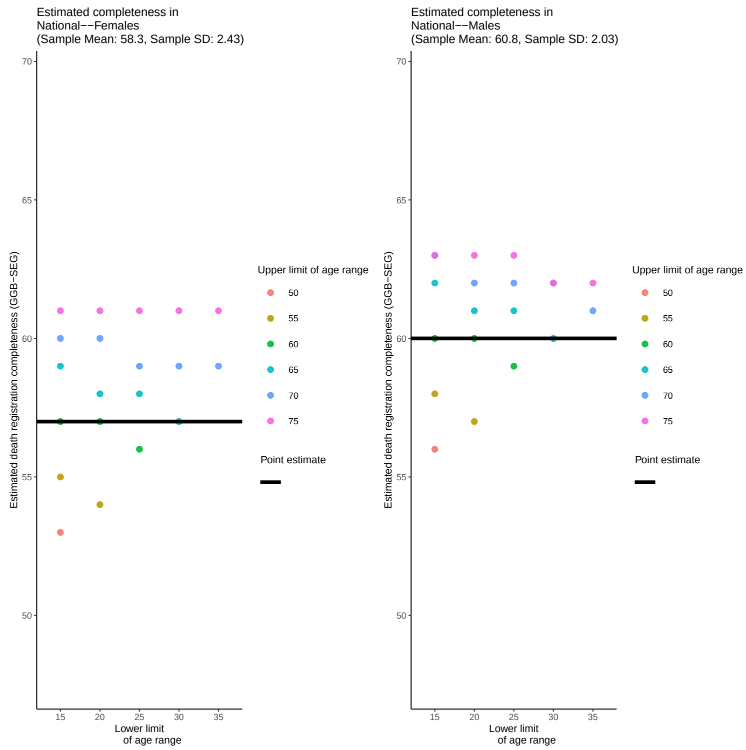

Figure 17: Sensitivity of point estimates of death registration completeness in Ecuador from 2001-2010 to choice of age-range parameter in the GGB-SEG method

```
knitr::include_graphics(path=paste0(my_plots_dir,
                                    "ggbseg_RMSE_ggb_National_",
                                    Sys.Date(),
                                    ".pdf"))
```

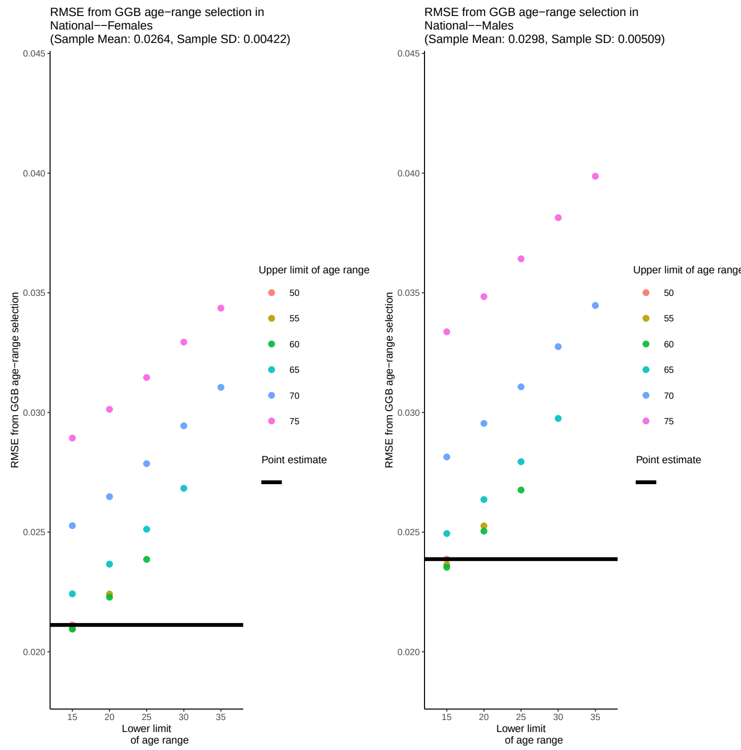

Figure 18: Sensitivity of point estimates of death registration completeness in Ecuador from 2001-2010 to choice of age-range parameter in the GGB-SEG method

References

INEC. 2010. “Censo de Población Y Vivienda.” https://www.ecuadorencifras.gob.ec/censo-de-poblacion-y-vivienda//.

Moultrie, Tom A, RE Dorrington, Allan G Hill, Kenneth Hill, IM Timæus, and Basia Zaba. 2013. *Tools for Demographic Estimation*. International Union for the Scientific Study of Population.

Peralta, Andrés, Joan Benach, Carme Borrell, Verónica Espinel-Flores, Lucinda Cash-Gibson, Bernardo L Queiroz, and Marc Marı’-Dell’Olmo. 2019. “Evaluation of the Mortality Registry in Ecuador (2001–2013)–Social and Geographical Inequalities in Completeness and Quality.” *Population Health Metrics* 17 (1). Springer: 3.

Riffe, T, JM Aburto, M Alexander, S Fennell, I Kashnitsky, M Pascariu, and P Gerland. 2019. “DemoTools: An R Package of Tools for Aggregate Demographic Analysis.”

Riffe, Tim, Everton Lima, and Bernardo Queiroz. 2017. “DDM: Death Registration Coverage Estimation.”