

# How to use the Map Reliability Calculator

Addison Larson

2018-12-12

## 1. Download your data.

You can do this using the Census API (as seen below) or by visiting American FactFinder and downloading your data from there. The chunk below downloads the number of Hispanic or Latino residents at the census tract level for PA and NJ and subsets the data for Burlington, Camden, Gloucester, and Mercer Counties in NJ; and Bucks, Chester, Delaware, Montgomery, and Philadelphia Counties in PA.

```
library(tidycensus); library(tidyverse)
trct <- get_acs(state = c("PA", "NJ"), geography = "tract",
               variables = "B03003_003",
               output = "wide") %>%
  mutate(stcty = substr(GEOID, 1, 5)) %>%
  subset(stcty %in% c("34005", "34007", "34015", "34021",
                    "42017", "42029", "42045", "42091", "42101")) %>%
  select(-stcty)
```

Here's what the results look like:

```
head(trct)

## # A tibble: 6 x 4
##   GEOID      NAME                                B03003_003E B03003_003M
##   <chr>      <chr>                                <dbl>      <dbl>
## 1 420171001~ Census Tract 1001.02, Bucks County, ~    118         72
## 2 420171001~ Census Tract 1001.03, Bucks County, ~    161        108
## 3 420171001~ Census Tract 1001.04, Bucks County, ~    666        299
## 4 420171001~ Census Tract 1001.05, Bucks County, ~    248        259
## 5 420171002~ Census Tract 1002.01, Bucks County, ~    173        143
## 6 420171002~ Census Tract 1002.06, Bucks County, ~    360        215
```

## 2. Prep your data.

The calculator will not work if you supply it extra columns, blank cells, or special characters. The chunk below grabs the estimate and the MOE columns and drops any observations with NAs.

```
trct <- trct %>%
  select(B03003_003E, B03003_003M) %>%
  drop_na()
```

Here's what the results look like:

```
head(trct)

## # A tibble: 6 x 2
##   B03003_003E B03003_003M
##   <dbl>      <dbl>
## 1     118         72
## 2     161        108
## 3     666        299
```

|      |     |     |
|------|-----|-----|
| ## 4 | 248 | 259 |
| ## 5 | 173 | 143 |
| ## 6 | 360 | 215 |

### 3. Export your data.

Export data as a .csv. Change the file path below to somewhere on your PC.

```
setwd("D:/alarson")
write.csv(trct, file = "export.csv", row.names = FALSE)
```

### 4. Use the calculator.

Visit the Map Reliability Calculator and drop in your new dataset.

The calculator has a lot of functions. Let's walk through some of them:

1. **Data upload**, located top left. Click the **browse** button to add your dataset.
  - **File has a header**, located top left. Uncheck this if your .csv has no column names.
  - **Include estimates of 0 in error calculations**, located top left. Check this if you have lots of zeroes in your dataset.
2. **Summary of expected errors**, located top center. A quick summary of the classification schemes and number of breaks good for your dataset.
3. **Select number of classes**, located top left. Once you're set on a number of classes for your map, select the number in the drop-down menu. Your selection changes the results of **Error by number of classes**.
4. **Acceptable error percentage**, located top left. Default is 10%.
5. **Custom comma-delimited breaks**, located top left.
6. **Error by number of classes**, located bottom center. When you select a number of classes, this section provides:
  - Overall expected classification error
  - Lower bound error by class
  - Upper bound error by class
  - Number of observations in each class
  - Total class error
7. **Summary of data uploaded**, located bottom left. Shows summary statistics of your data.
8. **Break values**, located bottom left. This can be especially useful when you want to tweak the breaks of an existing classification scheme to make it better.