**Note to self: transfer this to a markdown document for the sake of the repo**

# Variable Name Conventions

## Intro.

American Community Survey (ACS) tables each have a number—sometimes a very large number—of columns, each of which with a detailed description of what count or calculation it represents. The Census gives each field a unique ID for handling by programmers, and a long, plain English description for the sake of interpretation. Neither this ID or description are especially suitable for programmers since the former (e.g. B17034\_012E for …) and the latter--“” in the case of table element B17034\_0122E—is too long to be useful in writing programs.

A key component off this project is establishing a short-hand for naming table elements that:

1. has a predictable, systematic structure for use in programming, and
2. allows for relatively short names for practical use. In particular, a number of very highly-used program packages—such as STATA and SAS—in the social sciences have limits of 32 characters for variable names.

With these two requirements, it is also hoped that variable names will be as readable as possible. It is hoped that community members contributing to this project will offer ideas for how to balance readability with concision in naming.

## Conventions.

### Convention for table field elements

*<Table name>\_<field/value pair 1 {\_field/value pair 2 […etc]}>*

### Convention for constructed variables.

## Notes

* On case sensitivity: these conventions will be properly preserved by STATA and R, which are case sensitive languages. In languages that are not case sensitive such as SAS, users would need to use <specific options/means for reading/writing in ways that preserve case> and to make use of the convention’s