**ACS 2011 data**

Data file: *acs\_data\_bach.csv*

The data set is a random subsample of individual responses to the ACS survey from the year 2011. The individuals in the sample are wage and salary workers ages 25-54 with a bachelors degree (but no higher), who reported positive wage and salary income.

I recommend that you read a little bit about the data source, the American Community Survey (ACS). A link to their web site is provided on the Canvas site under ECON 41 Data Analysis Project. Run by the U.S. Census Bureau, the ACS surveys about 1% of the U.S. population annually, asking a wide variety of questions about demographics, education, work, earnings, housing, and the like.

Here are the variable definitions in your sample, in alphabetical order:

|  |  |
| --- | --- |
| *Variable* | *Definition* |
| age | Age in years |
| age\_spouse | Age of spouse in years (NA if not married with spouse present) |
| citizen | U.S. citizen (binary) |
| ethnicity | Ethnic group (categorical or “factor” variable)\*: Asian, Black, Hispanic, Other, or White.  \* This variable is based on people’s self-reported race and Hispanic origin. In the ACS, race and Hispanic origin are separate questions, so someone who is Hispanic can be of any race. These responses are combined here, with anyone reporting being Hispanic assigned to the Hispanic category. Thus all the other categories are non-Hispanic. See also the discussion of the *race* variable below. |
| female | Female (binary) |
| forborn | Foreign born (=0 if US born, = 1 if foreign born) |
| ftfy | Full-time, full-year worker (binary): worked at least 50 weeks past year, usual weekly hours 35+ |
| hispanic | Hispanic origin (binary): Hispanic may be of any race |
| income\_spouse | Total income of spouse (NA if not married with spouse present) |
| incwage | Wage and salary income earned over previous year, in $ |
| major | Field of study for bachelors degree (categorical or “factor” variable): Humanities, Natural sciences, Engineering, Social sciences, Public service, Education, Business, Health, Technical, or Other |
| marrspp | Married with spouse present (binary) |
| metro | Lives in metropolitan area (binary) |
| nchild | Number of own children living in household (any age) |
| nchlt5 | Number of own children under 5 years old living in household |
| race | Race (categorical or “factor” variable)\*\*: Asian, Black, Other , or White.  \*\* This variable is based on people’s self-reported race. In the ACS, people are permitted to report more than one racial identity, although a large majority (95%) report just one. For those reporting multiple races, the data source used here (IPUMS) assigns the “most likely” single race response using a statistical procedure. In this variable, Hispanics are included as whatever race they reported. So, for example, *race = White* includes both Hispanic and non-Hispanic whites, whereas *ethnicity = White* includes only non-Hispanic whites (Hispanic whites are *ethnicity = White*). The category Asian includes Pacific Islanders. |
| region | Region of residence (categorical or “factor” variable): Midwest, Northeast, South, or West. |
| single | Never married (binary) |
| south | Resides in South (binary) |
| speakeng | Speaks English well (binary) |
| years\_usa | Years in the United States if foreign-born; age if US-born |
| yrmarr | Year married (0 if not married) |