

Data Analysis Module I Part 1: Introductory Exercise

This assignment is worth five points and is due September 17.

Instructions:

1. Go to the website http://www.ssdan.net/ida_resources.shtml. Scroll down the page until you find WebChip Launcher. Click on the Launcher.
2. In the box, after the line “datasets/pc” enter “cen1990/educimm9.dat” (without quotation marks, all lower case). Click “Submit Query.” You should get a screen with gray buttons at the top and the output listed below.

CHIP

1990 Persons Age 25+, wgted 5% PUMS, Frey/U-Mich for SSDAN - not for commercial purposes.

RaceLat	Gender	Age	Imm	Educ
6	2	5	4	7

N = 158511248

The output provides information about the data set you have opened. This data set (educimm9.dat) contains 1990 data on the U.S. population. The data set includes five variables – RaceLat (racial and ethnic status in six categories), Gender, Age (five categories), Imm (immigration status over time in four categories), Educ (in seven categories).

3. Click the gray Marginals button and you will get the marginal frequencies for all the variables contained in the file – i.e., the percentage of the U.S. population that fall into each of the five variables.
4. Answer the following questions. In 1990, what percentage of the U.S. population
 - a. . . . was Black? _____
 - b. . . . was **NOT** non-Latino White (NLWhite)? _____
 - c. . . . had less than a high school degree? _____
 - d. . . . was a college graduate? _____
 - e. . . . was not born (FB = foreign born) in the United States? _____