

macros and loops

version: Friday 26th February, 2016 19:13

due in 2 weeks

1. put everything you have done so far together (can and should reuse past ps), the more lines of code the better, but also the better quality (simplicity, efficiency, etc) the better, must be 300-1k lines without comments and blank lines, or 300-3k lines incl blanks and comments
2. rewrite your earlier code using macros and loops and possibly add new code using macros and loops: use at least 5 macros and 3 loops; at least one nested loop, and at least one branching with **if else**

general directions (always the same):

- i will show your code in class and possibly post some of your code or link to it—again, as per our core values—opensource, transparency, sharing; but if you'd like to keep your code private, that's fine—just let me know, and i will keep your code secret (no penalty, except that you may get little less feedback—usually if we discuss your code in the class, you will benefit from it!)
- you must submit all the code that was executed from the very beginning starting with the very raw data as per replication principle
- all ps are cumulative—you can, and should, include previous code
- use your own dataset; again if you do not have a dataset, email listserv, stop by our offices, etc
- because you are only submitting code, it must load data from Internet—just put your data onto your own website, wordpress, google drive, etc; (when you put data into any public space, try not to violate data copyrights... I haven't heard of anyone having problems with that, but be careful—for instance you may subset dataset to few vars and smaller sample using **sample**); and it is also easier to experiment on small datasets
- keep it simple! drop unnecessary vars; and even retain only certain, say most important, observations; keep it manageable; it is much easier to learn using simple data; can always complicate later!; much better to do right using simple data than do it wrong using complex data!
- have nice structure in your file: sections, subsections, etc; may also have multiple files
- great idea to submit ps as early as possible—we will probably give you some comments; if not, email us and ask for comments!
- it is great to copy code from others; again, one of the rules for this class is 'be lazy': don't reinvent the wheel, whatever you are coding, it has already been done, google things often; but of course you cannot submit 100% code by someone's else.
- if you do something extra/fancy that is relevant and closely related to the assignment questions, it will be extra credit
- use coding rules that we've learned so far
- submit (only) the code into the sakai's dropbox, or email instructor the link to your GIT repo (once is enough, say for 1st ps); ps are due by the beginning of the next class unless indicated otherwise, eg “due in 2 weeks”; late ps are not accepted
- we are on the way to developing the final project with these ps: as we progress, your ps should start resembling a coherent and logical project where you use data management techniques to build new a dataset that can be used to answer interesting questions— say in few sentences (as a comment) why are you doing what you are doing—that is, answer the “so what question”: “ok, you're gonna run all that code, and so what?” what's the goal of all that, why are you doing this? you need a compelling justification for what you are doing; typically: to develop a new dataset (that has not existed before) that can be used to answer some exciting questions: say what are those questions you want to answer; be brief, say couple sentences, and definitely not more than say 100 lines, typically 10-50 lines is enough