Homework 2: Due Thursday, February 20th, 2018

This homework assignment will use data from the UMass Vitamin D Study, a study of 288 young women (ages 18-30) who completed questionnaires, provided a fasting blood sample, and received a bone density scan with the primary purpose of assessing predictors of Vitamin D levels in this population.

18. a. Have you ever used oral contraceptives? ○ yes   ○ no → if no, skip to question 19

   b. If yes, why did you first start using them?

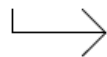   ○ birth control   ○ acne control   ○ PMS
     ○ menstrual cycle regulation   ○ anemia   ○ PCOS
     ○ other

          → If other: [   ]

   c. If yes, at what age did you first use them? [   ]

   d. If yes, how many years, in total, have you used them? [   ]

   e. Are you currently using oral contraceptives? ○ yes   ○ no → if no, skip to question 19

   f. If yes, what type of oral contraceptives? Please name the brand or formula below

[   ]

Below are six questions to assess oral contraceptive use from the survey:

1. This assignment uses the same dataset that was used in Homework #1. Use your working copy of this dataset that includes variable and value labels. Or, download the file "VitDhw1.dta" from Moodle and run your do file from Homework #1 to label the variables and their values.

1. For the variables "age at first use of OCs" (age_oc) and the "number of years used OCs" (yrs_oc):

   1.a. Inspect each variable and describe any issues you notice and how you will address these issues.

   1.b. Fix the issues following this plan.

   1.c. Check your work and verify that you have resolved these issues.

Pubhtlh 490kr
Spring 2018

1. Create a new variable, based on age_oc, that categorizes participants into the following groups for age at first OC use: ≤15, 16, 17, 18, 19, ≥20.

2. For the variables "ever used OCs" (ever_oc) and "current use of OCs" (curr_oc):

    2.a. Cross-tabulate ever_oc with curr_oc.

    2.b. Describe any problems you notice.

    2.c. Create a new variable, using ever_oc and curr_oc, that will code people as never, past, or current OC users. Be sure to address the inconsistencies you noted when creating this new variable.

    2.d. Check your work for errors in creating the new variable.

    2.e. Label this new variable and its values.

    2.f. Produce a tabulation of the new variable you created.

1. For the variable "type of OCs currently used" (type_oc):

    1.a. Produce a tabulation of the values. You will notice that this text variable records many similar drug names as different because of misspellings, capitalizations, etc.

    1.b. Group all variations of "ortho tri cyclen" into a single category. (Hint: use the "encode" command to make this easier; don't worry about regrouping other types).

    1.c. 50

    1.d. 57

    1.e. 58

    1.f. 62

    1.g. 67

    1.h. 94

jhjh

1.

    1.a. Produce a tabulation of the type variable after completing (5.b) above.

Codebook

Variable	Label	Values
id	Study ID number	Continuous
ever_oc	Ever used oral	0=No
why_oc	Why did you first start	1=birth control
oc_why_other	Reason for starting OC use	Text entry
age_oc	Age started using OCs	Continuous (range 12-27)
yrs_oc	How many years used OCs	Continuous (range 0-11)
curr_oc	Currently using OCs	0=No
type_oc	Type of OC currently using	Text entry