Assignment 03 – Replicating descriptive results

Quantitative Methods for the Social Sciences | Winter 2022/23

Please submit via ILIAS the latest by **Tuesday, 1 November 2022**, 12:00 CET.

Complete the following tasks and generate a do-file. Please upload the **do-file** to ILIAS. Include your solution to **all** tasks in this do-file.

Introduction

This assignment prepares you for next week's graded take-home Stata Assignment 1, in which you will partly replicate the analyses of a highly cited article written by Pescosolido et al. (2010). You will work with the General Social Survey (**GSS**), an annual cross-sectional study that has been monitoring trends in attitudes, behaviours, and attributes of the American population since 1972. You use data from two rounds of the GSS: 1996 and 2006 (available on ILIAS).

General preparations

Download "template_a03.zip" from ILIAS. Find the article by Pescosolido et al. 2010 and the GSS codebook in the main folder. The GSS data for 1996 and 2006 are already stored in the data folder.

Then read the **article** Pescosolido et al. 2010. Pay particular attention to **Table 1** and the associated text (pp. 1324-1325). The text describes what the authors did. You will follow their footsteps.

Look at the **codebook** of the GSS to understand how the vignette variables are coded (the documentation that you need starts on p. 1078 of the codebook). In the Pescosolido et al. 2010 article, you can find the vignette variables under the heading "Outcome measures" in the far-left column of Table 1.

Open the **do-file** *template_a03* (located in the zip archive), rename it into assignment03_YOURLASTNAME and save it in the folder *1_dofiles*. From now on, write every command and solution to the tasks into this do-file.

Note: You will use your do-file from this assignment as a **basis** to the graded Take-Home Stata Assignment 1 next week.

1. Combining two data sets

T1 Combine the two datasets *GSS1996.dta* and *GSS2006.dta* using the `append` command and `keep` only the relevant variables: VIGVERSN MENTLILL UPSDOWNS GENETICS IMBALNCE CHARACTR WAYRAISE MEDDOC MENTLDOC MENTLHOS RXMED VIGWORK VIGNEI VIGSOC VIGMAR VIGFRND HURTOTH HURTSELF YEAR WTSSALL. Use the code provided in the template do-file. Now the file contains only data on the relevant variables measured in 1996 and in 2006.

Note: Stata is case-sensitive. This means you need to type capital letters if you want Stata to recognize these variables. If you want to convert all variable into lowercase, use the `rename` command: `rename *, lower`. In this command, the * is a shorthand for "all variables".

2. Data exploration and preparation

T2 Look at the GSS codebook on pp.1078-1080. Check how the variable VIGVERSN was coded. *Hint*: The values 1 – 18 correspond to the Alcohol Dependence vignette, and so on. Do a crosstab [`tab`] of VIGVERSN and YEAR. What does this crosstab tell you about what was asked in the surveys 1996 and 2006?

T3 Only Schizophrenia, Major Depression, and Alcohol Dependence are needed for your partial replication. `recode` the VIGVERSN variable so that
Schizophrenia   = «1»
Major Depression   = «2»
Alcohol Dependence   = «3»
Other values   = missing values
Remember to always recode the original variables into new variables.
Use `help missing` if needed. Next, `label` the values of the recoded variable according to three categories "Schizophrenia", "Major Depression" and "Alcohol Dependence" (use `help label` if needed). Then `tabulate` the recoded version of VIGVERSN by YEAR to check the recoded variable. What are the total sample sizes for Schizophrenia, Major Depression, and Alcohol Dependence?

T4 `tabulate` the Neurobiological and Sociomoral variables from Table 1 (MENTLILL IMBALNCE GENETICS UPSDOWNS CHARACTR WAYRAISE).
Hint: You can use the `tab1` command to tabulate all variables at once. Check what these variables mean and how they are coded. Look at the GSS codebook to check the original coding and write a note about the original scale in your do-file as a comment. Then `recode` the Neurobiological and Sociomoral variables into new variables so that
"Very likely/Somewhat likely"   = «1»
" Not very likely/Not at all likely /Don't know"   = «0»
Other   = missing values.
`label` the values. `tab` the variables again, showing the missings.

T5 `tabulate` the Treatment Endorsement variables from Table 1 (MEDDOC MENTLDOC MENTLHOS RXMED). Check what these variables mean and how they

are coded. Look at the codebook to check the original coding and add a note about the original scale in you do-file as a comment. Then `recode` the Treatment Endorsement variables into new variables so that

"Yes"   = «1»
"No/Don't know"   = «0»
Other   = missing values.

`label` the values. `tabulate` the variables again, showing the missings.

T6 `tabulate` the Social Distance variables from Table 1 (VIGWORK VIGNEI VIGSOC VIGMAR VIGFRND). Check what these variables mean and how they are scaled. Look at the codebook to check the original coding and add the original scale in you do-file as a comment. Then `recode` the Social Distance variables into new variables so that

"Unwilling"   = «1»
"Willing/Don't know "   = «0»
Other   = missing values.

`label` the values. `tabulate` the variables again, showing the missings.

T7 `tabulate` the Dangerousness variables from Table 1 (HURTSELF HURTOTH). Check what these variables mean and how they are coded. Look at the codebook to check the original coding and add the original scale in your do-file as a comment. Then `recode` the Dangerousness variables into new variables so that

"Very likely/Somewhat likely"   = «1»
"Not likely/Don't know"   = «0»
Other   = missing values.

`label` the values. `tabulate` the variables again, showing the missings.