**RUTGERS UNIVERSITY**

**Bloustein School of Planning and Public Policy**

**Applied Multivariate Methods**

**Fall 2020**

**Dawne Mouzon, Ph.D.**

**Problem Set #4:**

**INTER-ITEM RELIABILITY AND SCALE CLEANING**

**NAME: \_\_\_\_\_\_\_\_\_\_HASSAN KHURSHID\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**SUBMISSION INSTRUCTIONS: Please upload to Canvas by 11:59 pm next Thursday night.**

**INTER-ITEM RELIABILITY AND SCALE CLEANING**

**(100 points total)**

**Please be sure to follow all instructions for creating recodes and constructing scales (including prep work). The requested output can be found after each question. Please copy and paste your code and output after each question/subquestion.**

**Your output must be copied/pasted. Do not use screenshots or pictures. Use the Courier New font, size 10 to make your output line up. Do not use bold font for output. Minimize the font if necessary.**

**Copy and paste your code from your do-file only (not the log); copy and paste your output from the results window. Exclude the “casewise” option throughout.**

**Complete the steps and provide the output in the order asked.**

1. **(40 pts) Please run inter-item reliability using the 2018 GSS support of suicide rights we created in Lab 3.** 
   1. **Initial inter-item reliability output**
   2. **Output for any changes required to the inter-item reliability analysis (hint: this may be a two-step process). If you drop any variables from the scale, provide output for inter-item reliability after each item dropped.**
   3. **Provide a one-sentence description of how the overall composite scale should be revised.**
   4. **Paste:** 
      1. **frequencies of the (edited) overall scale (after dropping any variable(s) you chose to drop in parts a, b, and c)**
      2. **simple descriptives of the scale and final input items (exclude “detail” option)**
      3. **consistency checks using “list” showing ~5 cases for all three potential outcomes (use “nolabel” option to see values only)**
   5. **Now limit the edited scale to include only those with valid data on the final input items included in the scale. Then paste:**
      1. **Output from mvpatterns (only including final set of input scale items)**
      2. **Frequencies from new “nmiss” variable (“*nmiss\_suicide\_scale*”)**
      3. **Frequencies for *suicide\_scale*, both with and without the “miss” option**
   6. **Your code for the entire problem**

a. alpha suicide1r-suicide4r, item

Test scale = mean(unstandardized items)

average

item-test item-rest interitem

Item | Obs Sign correlation correlation covariance alpha

-------------+-----------------------------------------------------------------

suicide1r | 1497 + 0.6526 0.3239 .091818 0.8742

suicide2r | 1532 + 0.8276 0.6946 .0669943 0.6644

suicide3r | 1539 + 0.8255 0.6936 .0678485 0.6657

suicide4r | 1522 + 0.8411 0.6818 .0591733 0.6616

-------------+-----------------------------------------------------------------

Test scale | .0715285 0.7742

b. alpha suicide2r-suicide4r, item

Test scale = mean(unstandardized items)

average

item-test item-rest interitem

Item | Obs Sign correlation correlation covariance alpha

-------------+-----------------------------------------------------------------

suicide2r | 1532 + 0.9052 0.7915 .0918691 0.7927

suicide3r | 1539 + 0.9124 0.8062 .0906305 0.7787

suicide4r | 1522 + 0.8806 0.6851 .0929385 0.9053

-------------+-----------------------------------------------------------------

Test scale | .091818 0.8742

-----------------------------------------------------------------------------

. alpha suicide2r-suicide3r, item

Test scale = mean(unstandardized items)

Average interitem covariance: .0929385

Number of items in the scale: 2

Scale reliability coefficient: 0.9053

c. Since the overall composite scale’s Cronbach Alpha increases from 0.7742 to 0.9053, so we will drop suicide1r and suicide4r from the scale.

\*simple descriptive and consistency checks for final output scale and its input items\*

.

sum suicide\_scale, detail

2-item scale for suicide rights

-------------------------------------------------------------

Percentiles Smallest

1% 0 0

5% 0 0

10% 0 0 Obs 1,547

25% 0 0 Sum of Wgt. 1,547

50% 0 Mean .2559793

Largest Std. Dev. .6376664

75% 0 2

90% 2 2 Variance .4066185

95% 2 2 Skewness 2.219214

99% 2 2 Kurtosis 6.144644

. tab suicide\_scale, miss

2-item scale for |

suicide rights | Freq. Percent Cum.

-----------------+-----------------------------------

0 | 1,318 56.13 56.13

1 | 62 2.64 58.77

2 | 167 7.11 65.89

.d. DK | 9 0.38 66.27

.i. IAP | 789 33.60 99.87

.n. NA | 3 0.13 100.00

-----------------+-----------------------------------

Total | 2,348 100.00

. sum suicide\_scale suicide2r-suicide3r

Variable | Obs Mean Std. Dev. Min Max

-------------+---------------------------------------------------------

suicide\_sc~e | 1,547 .2559793 .6376664 0 2

suicide2r | 1,532 .1318538 .3384422 0 1

suicide3r | 1,539 .1260559 .3320202 0 1

list suicide\_scale suicide2r-suicide3r if !missing(suicide2r-suicide3r) in 1/150, nolabel

+--------------------------------+

| suicid~e suici~2r suici~3r |

|--------------------------------|

1. | 2 1 1 |

3. | 0 0 0 |

4. | 0 0 0 |

6. | 0 0 0 |

9. | 0 0 0 |

|--------------------------------|

list suicide\_scale suicide2r-suicide3r if missing(suicide2r) & !missing(suicide3r) in 1/1000, nolabel

+--------------------------------+

| suicid~e suici~2r suici~3r |

|--------------------------------|

133. | 0 .d 0 |

415. | 0 .d 0 |

435. | 0 .d 0 |

744. | 1 .d 1 |

828. | 0 .n 0 |

+--------------------------------+

. list suicide\_scale suicide2r-suicide3r if missing(suicide2r-suicide3r) in 1/150, nolabel

+--------------------------------+

| suicid~e suici~2r suici~3r |

|--------------------------------|

2. | .i .i .i |

5. | .i .i .i |

7. | .i .i .i |

8. | .i .i .i |

14. | .i .i .i |

|--------------------------------|

. mvpatterns suicide2r-suicide3r, sort

Variable | type obs mv variable label

-------------+----------------------------------------------------

suicide2r | float 1532 816 suicide OK if bankrupt

suicide3r | float 1539 809 suicide OK if dishonored family

------------------------------------------------------------------

Patterns of missing values

+------------------------+

| \_pattern \_mv \_freq |

|------------------------|

| ++ 0 1524 |

| .. 2 801 |

| .+ 1 15 |

| +. 1 8 |

+------------------------+

. tab year

gss year |

for this |

respondent | Freq. Percent Cum.

------------+-----------------------------------

2018 | 2,348 100.00 100.00

------------+-----------------------------------

Total | 2,348 100.00

. tab nmiss\_suicide\_scale, miss

Count for |

different |

number of |

missing |

variables |

for 2-item |

suicide |

scale | Freq. Percent Cum.

------------+-----------------------------------

0 | 1,524 64.91 64.91

1 | 23 0.98 65.89

2 | 801 34.11 100.00

------------+-----------------------------------

Total | 2,348 100.00

tab suicide\_scale, miss

2-item scale for |

suicide rights | Freq. Percent Cum.

-----------------+-----------------------------------

0 | 1,297 55.24 55.24

1 | 60 2.56 57.79

2 | 167 7.11 64.91

.a | 824 35.09 100.00

-----------------+-----------------------------------

Total | 2,348 100.00

. tab suicide\_scale

2-item scale for |

suicide rights | Freq. Percent Cum.

-----------------+-----------------------------------

0 | 1,297 85.10 85.10

1 | 60 3.94 89.04

2 | 167 10.96 100.00

-----------------+-----------------------------------

Total | 1,524 100.00

. sum suicide\_scale, detail

2-item scale for suicide rights

-------------------------------------------------------------

Percentiles Smallest

1% 0 0

5% 0 0

10% 0 0 Obs 1,524

25% 0 0 Sum of Wgt. 1,524

50% 0 Mean .2585302

Largest Std. Dev. .6411881

75% 0 2

90% 2 2 Variance .4111222

95% 2 2 Skewness 2.20271

99% 2 2 Kurtosis 6.063697

**/\*\*\*\*\* code for 4-item suicide scale**

**egen suicide\_scale = rowtotal(suicide1r-suicide4r), miss**

**label variable suicide\_scale "4 item scale for suicide rights"**

**replace suicide\_scale = .d if suicide1r == .d & suicide2r == .d & suicide3r ==.d & suicide4r ==.d**

**replace suicide\_scale =.i if suicide1r == .i & suicide2r == .i & suicide3r ==.i & suicide4r ==.i**

**replace suicide\_scale =.n if suicide1r == .n & suicide2r == .n & suicide3r ==.n & suicide4r ==.n**

**label define em .d "DK" .i "IAP" .n "NA" .c "can't choose"**

**label values suicide\_scale em**

**numlabel em, add force**

**sum suicide\_scale, detail**

**tab suicide\_scale, miss**

**list suicide\_scale suicide1r-suicide4r if !missing(suicide1r-suicide4r) in 1/150, nolabel**

**list suicide\_scale suicide1r-suicide4r if missing(suicide2r) & !missing(suicide1r) in 1/150, nolabel**

**list suicide\_scale suicide1r-suicide4r if missing(suicide1r-suicide4r) in 1/150, nolabel\*\*\*\*\*\*/**

**alpha suicide1r-suicide4r, item**

**\*since suicide1r has a cronbach alpha (0.8742) much higher than overall (0.7742), so it is dropped\***

**/\*\*\*\*code for 3-item suicide scale**

**drop suicide1r**

**egen suicide\_scale = rowtotal(suicide2r-suicide4r), miss**

**label variable suicide\_scale "3-item scale for suicide rights"**

**replace suicide\_scale = .d if suicide2r == .d & suicide3r ==.d & suicide4r ==.d**

**replace suicide\_scale =.i if suicide2r == .i & suicide3r ==.i & suicide4r ==.i**

**replace suicide\_scale =.n if suicide2r == .n & suicide3r ==.n & suicide4r ==.n**

**label define em .d "DK" .i "IAP" .n "NA" .c "can't choose"**

**label values suicide\_scale em**

**numlabel em, add force**

**sum suicide\_scale, detail**

**tab suicide\_scale, miss**

**sum suicide\_scale suicide2r-suicide4r**

**list suicide\_scale suicide2r-suicide4r if !missing(suicide2r-suicide4r) in 1/150, nolabel**

**list suicide\_scale suicide2r-suicide4r if missing(suicide2r) & !missing(suicide4r) in 1/150, nolabel**

**list suicide\_scale suicide2r-suicide4r if missing(suicide2r-suicide4r) in 1/150, nolabel**

**alpha suicide2r-suicide4r, item\*\*\*\*\*/**

**alpha suicide2r-suicide4r**

**\*now suicide4r has a cronbach alpha (0.9053)much higher than the new overall (0.8742), so it will be dropped\***

**drop suicide4r**

**egen suicide\_scale = rowtotal(suicide2r-suicide3r), miss**

**label variable suicide\_scale "2-item scale for suicide rights"**

**replace suicide\_scale = .d if suicide2r == .d & suicide3r ==.d**

**replace suicide\_scale =.i if suicide2r == .i & suicide3r ==.i**

**replace suicide\_scale =.n if suicide2r == .n & suicide3r ==.n**

**label define em .d "DK" .i "IAP" .n "NA" .c "can't choose"**

**label values suicide\_scale em**

**numlabel em, add force**

**sum suicide\_scale, detail**

**tab suicide\_scale, miss**

**sum suicide\_scale suicide2r-suicide3r**

**list suicide\_scale suicide2r-suicide3r if !missing(suicide2r-suicide3r) in 1/150, nolabel**

**list suicide\_scale suicide2r-suicide3r if missing(suicide2r) & !missing(suicide3r) in 1/1000, nolabel**

**list suicide\_scale suicide2r-suicide3r if missing(suicide2r-suicide3r) in 1/150, nolabel**

**alpha suicide2r-suicide3r, item**

**\*Now limiting the edited scale to include only valid data**

**mvpatterns suicide2r-suicide3r, sort**

**egen nmiss\_suicide\_scale = rmiss2(suicide2r-suicide3r)**

**label variable nmiss\_suicide\_scale "Count for different number of missing variables for 2-item suicide scale"**

**tab nmiss\_suicide\_scale, miss**

**tab suicide\_scale, miss**

**replace suicide\_scale =.a if nmiss\_suicide\_scale >0 & nmiss\_suicide\_scale <=2**

**tab suicide\_scale**

**tab suicide\_scale, miss**

**sum suicide\_scale, detail**

1. **(20 pts) Please create a 1996 GSS subset and corresponding do-file. Then run inter-item reliability on the “support of abortion rights” scale you created in PS #3.**
   1. **Initial inter-item reliability output**
   2. **Output for any changes required to the inter-item reliability analysis (hint: this may be a two-step process)**
   3. **Provide a one-sentence description of how the overall composite scale should be revised.**
   4. **Paste:** 
      1. **frequencies of the (edited) overall scale (hints: copy/paste/edit code from the 1972-2018 master do-file into your 1996 do-file; add “drop abscale” to your 1996 do-file)**
      2. **simple descriptives of the scale and final input items (exclude “detail” option)**
      3. **consistency checks using “list” showing ~5 cases for all three potential outcomes (use “nolabel” to see values only).**

**Note: if you use the “missing(varlist)” option, you will need to list each individual item with commas interspersed – e.g.,**

**list abscale abscale inputvar1 inputvar2 inputvar3…. if missing(inputvar1, inputvar2, inputvar3…) in 1/100, nolabel**

* 1. **Your code for the entire problem**

. alpha abdefect\_rs-abany\_rs, item

Test scale = mean(unstandardized items)

average

item-test item-rest interitem

Item | Obs Sign correlation correlation covariance alpha

-------------+-----------------------------------------------------------------

abdefect\_rs | 1850 + 0.6760 0.5658 .11048 0.8808

abnomore\_rs | 1825 + 0.8591 0.7812 .08936 0.8507

abhlth\_rs | 1850 + 0.5051 0.4026 .1240499 0.8941

abpoor\_rs | 1830 + 0.8623 0.7871 .0891985 0.8503

abrape\_rs | 1839 + 0.6453 0.5366 .1131676 0.8834

absingle\_rs | 1832 + 0.8722 0.8011 .0884549 0.8479

abany\_rs | 1821 + 0.8448 0.7587 .0904143 0.8533

-------------+-----------------------------------------------------------------

Test scale | .1007188 0.8845

-------------------------------------------------------------------------------

. alpha abdefect\_rs abnomore\_rs abpoor\_rs-abany\_rs, item

Test scale = mean(unstandardized items)

average

item-test item-rest interitem

Item | Obs Sign correlation correlation covariance alpha

-------------+-----------------------------------------------------------------

abdefect\_rs | 1850 + 0.6399 0.5110 .1462238 0.9027

abnomore\_rs | 1825 + 0.8801 0.8064 .1121098 0.8581

abpoor\_rs | 1830 + 0.8823 0.8103 .1119803 0.8577

abrape\_rs | 1839 + 0.6091 0.4837 .1499732 0.9056

absingle\_rs | 1832 + 0.8942 0.8287 .1107246 0.8543

abany\_rs | 1821 + 0.8687 0.7866 .1133632 0.8606

-------------+-----------------------------------------------------------------

Test scale | .1240499 0.8941

-------------------------------------------------------------------------------

. alpha abnomore\_rs abpoor\_rs absingle\_rs abany\_rs, item

Test scale = mean(unstandardized items)

average

item-test item-rest interitem

Item | Obs Sign correlation correlation covariance alpha

-------------+-----------------------------------------------------------------

abnomore\_rs | 1825 + 0.9145 0.8397 .1988312 0.9238

abpoor\_rs | 1830 + 0.9191 0.8461 .1975337 0.9214

absingle\_rs | 1832 + 0.9319 0.8688 .1940779 0.9144

abany\_rs | 1821 + 0.9116 0.8318 .2000792 0.9253

-------------+-----------------------------------------------------------------

Test scale | .1976335 0.9398

-------------------------------------------------------------------------------

Since composite scale achieves a highest cronbach’s alpha of 0.938 from initial overall alpha of 0.8845, so it is best to drop abhlth\_rs, abdefect\_rs, and abrape\_rs from the composite scale.

. sum abscale abnomore\_rs abpoor\_rs absingle\_rs abany\_rs

Variable | Obs Mean Std. Dev. Min Max

-------------+---------------------------------------------------------

abscale | 1,894 1.767159 1.805128 0 4

abnomore\_rs | 1,825 .4673973 .4990727 0 1

abpoor\_rs | 1,830 .4661202 .4989872 0 1

absingle\_rs | 1,832 .44869 .4974961 0 1

abany\_rs | 1,821 .4497529 .4976055 0 1

. sum abscale, detail

4 item abortion right scale

-------------------------------------------------------------

Percentiles Smallest

1% 0 0

5% 0 0

10% 0 0 Obs 1,894

25% 0 0 Sum of Wgt. 1,894

50% 1 Mean 1.767159

Largest Std. Dev. 1.805128

75% 4 4

90% 4 4 Variance 3.258488

95% 4 4 Skewness .2498829

99% 4 4 Kurtosis 1.236394

. tab abscale, miss

4 item |

abortion |

right scale | Freq. Percent Cum.

------------+-----------------------------------

0 | 842 28.99 28.99

1 | 173 5.96 34.95

2 | 120 4.13 39.08

3 | 102 3.51 42.60

4 | 657 22.62 65.22

. | 1 0.03 65.25

.d. DK | 22 0.76 66.01

.i. IAP | 981 33.78 99.79

.n. NA | 6 0.21 100.00

------------+-----------------------------------

Total | 2,904 100.00

. list id abscale abnomore\_rs abpoor\_rs absingle\_rs abany\_rs if missing(abnomore\_rs, abpoor\_rs, absingle\_rs, a

> bany\_rs) in 1/50, nolabel

+----------------------------------------------------------+

| id abscale abnomo~s abpoor~s absing~s abany\_rs |

|----------------------------------------------------------|

2. | 2 .i .i .i .i .i |

4. | 4 .i .i .i .i .i |

8. | 8 .i .i .i .i .i |

9. | 9 .i .i .i .i .i |

14. | 14 .d .d .d .d .d |

|----------------------------------------------------------|

. list id abscale abnomore\_rs abpoor\_rs absingle\_rs abany\_rs if missing(absingle\_rs) & !missing(abpoor\_rs) in

> 1000/1500, nolabel

+------------------------------------------------------------+

| id abscale abnomo~s abpoor~s absing~s abany\_rs |

|------------------------------------------------------------|

1000. | 1000 0 0 0 .d 0 |

1002. | 1002 2 .d 1 .d 1 |

1155. | 1155 1 .d 1 .d 0 |

1308. | 1308 3 1 1 .d 1 |

1325. | 1325 0 0 0 .d .d |

|------------------------------------------------------------|

. list id abscale abnomore\_rs abpoor\_rs absingle\_rs abany\_rs if !missing(abnomore\_rs, abpoor\_rs, absingle\_rs,

> abany\_rs) in 1000/1050, nolabel

+------------------------------------------------------------+

| id abscale abnomo~s abpoor~s absing~s abany\_rs |

|------------------------------------------------------------|

1005. | 1005 0 0 0 0 0 |

1006. | 1006 4 1 1 1 1 |

1007. | 1007 4 1 1 1 1 |

1009. | 1009 0 0 0 0 0 |

1010. | 1010 3 1 1 1 0 |

|------------------------------------------------------------|

**tab abscale, miss**

**sum abdefect\_rs-abany\_rs**

**alpha abdefect\_rs-abany\_rs, item**

**\*since abhlth\_rs has cornbach alpha (0.8941) greater than overall alpha (0.8845) so this is dropped to increase overall alpha\***

**/\*\*\*\* code for 6 item abortion scale**

**drop abscale**

**drop abhlth\_rs**

**egen abscale = rowtotal(abdefect\_rs abnomore\_rs abpoor\_rs-abany\_rs), miss**

**label variable abscale "6 item abortion right scale"**

**tab abscale, miss**

**replace abscale = .i if abdefect\_rs == .i & abnomore\_rs == .i & abpoor\_rs == .i & absingle\_rs ==.i & abrape\_rs ==.i & abany\_rs ==.i**

**replace abscale = .d if abdefect\_rs == .d & abnomore\_rs == .d & abpoor\_rs == .d & absingle\_rs ==.d & abrape\_rs ==.d & abany\_rs ==.d**

**replace abscale = .n if abdefect\_rs == .n & abnomore\_rs == .n & abpoor\_rs == .n & absingle\_rs ==.n & abrape\_rs ==.n & abany\_rs ==.n**

**\*label define yngss\_miss .i "IAP" .d "DK" .n "NA" \*this is already defined in master file (just for reference sake)**

**label values abscale yngss\_miss**

**numlabel yngss\_miss, add**

**sum abscale abdefect\_rs abnomore\_rs abpoor\_rs-abany\_rs**

**sum abscale, detail**

**tab abscale, miss**

**list id abscale abdefect\_rs abnomore\_rs abpoor\_rs-abany\_rs if missing(abdefect\_rs, abnomore\_rs, abpoor\_rs-abany\_rs) in 1/50, nolabel**

**list id abscale abdefect\_rs abnomore\_rs abpoor\_rs-abany\_rs if missing(abrape\_rs) & !missing(absingle\_rs) in 60000/61000, nolabel**

**list id abscale abdefect\_rs abnomore\_rs abpoor\_rs-abany\_rs if !missing(abdefect\_rs, abnomore\_rs, abpoor\_rs-abany\_rs) in 64800/64810, nolabel \*\*\*\*\*/**

**alpha abdefect\_rs abnomore\_rs abpoor\_rs-abany\_rs, item**

**\*since abrape\_rs alpha (0.9056)and abdefect\_rs alpha (0.9027)are greater than overall alpha (0.8941), so both of these will be dropped to increase overall alpha\***

**drop abscale**

**drop abdefect\_rs abrape\_rs**

**egen abscale = rowtotal(abnomore\_rs abpoor\_rs absingle\_rs abany\_rs), miss**

**label variable abscale "4 item abortion right scale"**

**tab abscale, miss**

**replace abscale = .i if abnomore\_rs == .i & abpoor\_rs == .i & absingle\_rs ==.i & abany\_rs ==.i**

**replace abscale = .d if abnomore\_rs == .d & abpoor\_rs == .d & absingle\_rs ==.d & abany\_rs ==.d**

**replace abscale = .n if abnomore\_rs == .n & abpoor\_rs == .n & absingle\_rs ==.n & abany\_rs ==.n**

**\*label define yngss\_miss .i "IAP" .d "DK" .n "NA" \*this is already defined in master file (just for reference sake)**

**label values abscale yngss\_miss**

**numlabel yngss\_miss, add**

**sum abscale abnomore\_rs abpoor\_rs absingle\_rs abany\_rs**

**sum abscale, detail**

**tab abscale, miss**

**list id abscale abnomore\_rs abpoor\_rs absingle\_rs abany\_rs if missing(abnomore\_rs, abpoor\_rs, absingle\_rs, abany\_rs) in 1/50, nolabel**

**list id abscale abnomore\_rs abpoor\_rs absingle\_rs abany\_rs if missing(absingle\_rs) & !missing(abpoor\_rs) in 1000/1500, nolabel**

**list id abscale abnomore\_rs abpoor\_rs absingle\_rs abany\_rs if !missing(abnomore\_rs, abpoor\_rs, absingle\_rs, abany\_rs) in 1000/1050, nolabel**

**alpha abnomore\_rs abpoor\_rs absingle\_rs abany\_rs, item**

**3. (40 points) Using 1972-2018 General Social Survey, recode cesd1-cesd5 in order to create a scale of depressive symptoms (*cesd\_scale*) using recoded variables (new variables should be named cesd1r-cesd5r). Remember to use the principles of good scale construction (please revisit Lab #3 and Principles of Good Scale Construction Worksheet). Carry over the extended missing values to the recoded variables and overall scale and use the “em” value label set throughout.**

* + - * 1. **Initial inter-item reliability output**
        2. **Output for any changes required to the inter-item reliability analysis (after dropping any variable(s) you chose to drop in parts a, b, and c)**
        3. **Provide a one-sentence description of how the overall composite scale should be constructed, based on the results from the inter-item reliability analysis.**
        4. **Paste:**

**frequencies of the overall scale**

**simple descriptives of the scale and final input items (exclude “detail” option)**

**consistency checks using “list” showing ~5 cases for all three potential outcomes (use “nolabel” option to see values only)**

**Note: if you use the “missing(varlist)” option, you will need to list each individual item with commas interspersed – e.g.,**

**list abscale abscale inputvar1 inputvar2 inputvar3…. if missing(inputvar1, inputvar2, inputvar3…) in 1/100, nolabel**

* 1. **Now limit the edited scale to include only those with valid data on the final input items included in the scale. Then paste:**
     1. **Output from mvpatterns (only including final set of input scale items)**
     2. **Frequencies for new “nmiss” variable (“*nmiss\_cesd\_scale*”)**
     3. **Frequencies for *cesd\_scale*, both with and without the “miss” option.**
  2. **Your code for the entire problem**

. tab cesd\_scale, miss

5-item scale of |

depressive |

symptoms | Freq. Percent Cum.

-----------------+-----------------------------------

0 | 80 0.12 0.12

1 | 151 0.23 0.36

2 | 167 0.26 0.61

3 | 160 0.25 0.86

4 | 133 0.21 1.07

5 | 79 0.12 1.19

6 | 63 0.10 1.29

7 | 44 0.07 1.35

8 | 28 0.04 1.40

9 | 23 0.04 1.43

10 | 15 0.02 1.45

11 | 9 0.01 1.47

12 | 8 0.01 1.48

13 | 2 0.00 1.48

14 | 2 0.00 1.49

15 | 3 0.00 1.49

.d. DK | 1 0.00 1.49

.i. IAP | 63,835 98.49 99.98

.n. NA | 11 0.02 100.00

-----------------+-----------------------------------

Total | 64,814 100.00

sum cesd1r-cesd5r cesd\_scale

Variable | Obs Mean Std. Dev. Min Max

-------------+---------------------------------------------------------

cesd1r | 963 .4060228 .6625135 0 3

cesd2r | 965 1.063212 .9307696 0 3

cesd3r | 966 1.026915 .7787995 0 3

cesd4r | 964 .5165975 .7479529 0 3

cesd5r | 964 .56639 .6826366 0 3

-------------+---------------------------------------------------------

cesd\_scale | 967 3.570838 2.740861 0 15

sum cesd\_scale, detail

5-item cesd scale

-------------------------------------------------------------

Percentiles Smallest

1% 0 0

5% 0 0

10% 1 0 Obs 967

25% 2 0 Sum of Wgt. 967

50% 3 Mean 3.570838

Largest Std. Dev. 2.740861

75% 5 14

90% 7 15 Variance 7.512316

95% 9 15 Skewness 1.158891

99% 12 15 Kurtosis 4.482136

alpha cesd1r-cesd5r, item

Test scale = mean(unstandardized items)

average

item-test item-rest interitem

Item | Obs Sign correlation correlation covariance alpha

-------------+-----------------------------------------------------------------

cesd1r | 963 + 0.7913 0.6683 .2148237 0.6771

cesd2r | 965 + 0.6433 0.3685 .2519515 0.7913

cesd3r | 966 + 0.7091 0.5152 .2303547 0.7233

cesd4r | 964 + 0.7144 0.5338 .230267 0.7166

cesd5r | 964 + 0.7784 0.6444 .2161011 0.6827

-------------+-----------------------------------------------------------------

Test scale | .2286951 0.7611

Since cronbach alpha of cesd2r (0.7913) is greater than overall alpha of 5 items (0.7611), we will drop this item to increase the overall alpha.

. list id cesd\_scale cesd1r cesd3r-cesd5r if missing(cesd1r) & !missing(cesd3r), nolabel

+-----------------------------------------------------+

| id cesd\_s~e cesd1r cesd3r cesd4r cesd5r |

|-----------------------------------------------------|

59689. | 90 3 .n 2 1 .n |

59718. | 119 3 .d 1 1 1 |

60386. | 787 0 .d 0 0 0 |

61200. | 1601 0 .n 0 .n .n |

+-----------------------------------------------------+

. list id cesd\_scale cesd1r cesd3r-cesd5r if missing(cesd1r, cesd3r-cesd5r) in 1/10, nolabel

+---------------------------------------------------+

| id cesd\_s~e cesd1r cesd3r cesd4r cesd5r |

|---------------------------------------------------|

1. | 1 .i .i .i .i .i |

2. | 2 .i .i .i .i .i |

3. | 3 .i .i .i .i .i |

4. | 4 .i .i .i .i .i |

5. | 5 .i .i .i .i .i |

|---------------------------------------------------|

. list id cesd\_scale cesd1r cesd3r-cesd5r if missing(cesd1r) & (!missing(cesd3r) | !missing(cesd4r)), nolabel

+-----------------------------------------------------+

| id cesd\_s~e cesd1r cesd3r cesd4r cesd5r |

|-----------------------------------------------------|

59689. | 90 3 .n 2 1 .n |

59718. | 119 3 .d 1 1 1 |

60386. | 787 0 .d 0 0 0 |

61200. | 1601 0 .n 0 .n .n |

+-----------------------------------------------------+

tab cesd\_scale, miss

5-item cesd |

scale | Freq. Percent Cum.

----------------+-----------------------------------

0 | 154 0.24 0.24

1 | 250 0.39 0.62

2 | 169 0.26 0.88

3 | 148 0.23 1.11

4 | 79 0.12 1.23

5 | 70 0.11 1.34

6 | 36 0.06 1.40

7 | 21 0.03 1.43

8 | 15 0.02 1.45

9 | 13 0.02 1.47

10 | 6 0.01 1.48

11 | 3 0.00 1.49

12 | 3 0.00 1.49

.d. DK | 1 0.00 1.49

.i. no IAP Data | 63,835 98.49 99.98

.n. NA | 11 0.02 100.00

----------------+-----------------------------------

Total | 64,814 100.00

sum cesd1r-cesd5r cesd\_scale

Variable | Obs Mean Std. Dev. Min Max

-------------+---------------------------------------------------------

cesd1r | 963 .4060228 .6625135 0 3

cesd3r | 966 1.026915 .7787995 0 3

cesd4r | 964 .5165975 .7479529 0 3

cesd5r | 964 .56639 .6826366 0 3

cesd\_scale | 967 2.509824 2.255349 0 12

sum cesd\_scale, detail

4-item cesd scale

-------------------------------------------------------------

Percentiles Smallest

1% 0 0

5% 0 0

10% 0 0 Obs 967

25% 1 0 Sum of Wgt. 967

50% 2 Mean 2.509824

Largest Std. Dev. 2.255349

75% 4 11

90% 6 12 Variance 5.086601

95% 7 12 Skewness 1.302924

99% 10 12 Kurtosis 4.801772

alpha cesd1r-cesd5r, item

Test scale = mean(unstandardized items)

average

item-test item-rest interitem

Item | Obs Sign correlation correlation covariance alpha

-------------+-----------------------------------------------------------------

cesd1r | 963 + 0.8157 0.6699 .243607 0.7087

cesd3r | 966 + 0.7612 0.5391 .2604666 0.7744

cesd4r | 964 + 0.7616 0.5530 .2616248 0.7641

cesd5r | 964 + 0.8123 0.6580 .2421189 0.7126

-------------+-----------------------------------------------------------------

Test scale | .2519515 0.7913

. mvpatterns cesd1r cesd3r-cesd5r, sort

Variable | type obs mv variable label

-------------+---------------------------------------------------------------

cesd1r | float 96363851 how many times felt depressed in past week

cesd4r | float 96463850 how many times felt lonely in past week

cesd5r | float 96463850 how many times felt sad in past week

cesd3r | float 96663848 how many times felt happy in past week

-----------------------------------------------------------------------------

Patterns of missing values

+------------------------+

| \_pattern \_mv \_freq |

|------------------------|

| .... 4 63847 |

| ++++ 0 961 |

| .+++ 1 2 |

| +.++ 1 1 |

| .+.+ 2 1 |

|------------------------|

| +... 3 1 |

| ...+ 3 1 |

+------------------------+

. tab nmiss\_cesd\_scale

Count for |

missing |

variable |

frequencies |

for 4 item |

scale of |

depression |

symptoms | Freq. Percent Cum.

------------+-----------------------------------

0 | 961 1.48 1.48

1 | 3 0.00 1.49

2 | 1 0.00 1.49

3 | 2 0.00 1.49

4 | 63,847 98.51 100.00

------------+-----------------------------------

Total | 64,814 100.00

. tab cesd\_scale

4-item scale of |

depressive |

symptoms | Freq. Percent Cum.

-----------------+-----------------------------------

0 | 151 15.71 15.71

1 | 250 26.01 41.73

2 | 168 17.48 59.21

3 | 146 15.19 74.40

4 | 79 8.22 82.62

5 | 70 7.28 89.91

6 | 36 3.75 93.65

7 | 21 2.19 95.84

8 | 15 1.56 97.40

9 | 13 1.35 98.75

10 | 6 0.62 99.38

11 | 3 0.31 99.69

12 | 3 0.31 100.00

-----------------+-----------------------------------

Total | 961 100.00

. tab cesd\_scale, miss

4-item scale of |

depressive |

symptoms | Freq. Percent Cum.

-----------------+-----------------------------------

0 | 151 0.23 0.23

1 | 250 0.39 0.62

2 | 168 0.26 0.88

3 | 146 0.23 1.10

4 | 79 0.12 1.23

5 | 70 0.11 1.33

6 | 36 0.06 1.39

7 | 21 0.03 1.42

8 | 15 0.02 1.44

9 | 13 0.02 1.46

10 | 6 0.01 1.47

11 | 3 0.00 1.48

12 | 3 0.00 1.48

.a | 63,853 98.52 100.00

-----------------+-----------------------------------

Total | 64,814 100.00

. sum cesd\_scale, detail

4-item scale of depressive symptoms

-------------------------------------------------------------

Percentiles Smallest

1% 0 0

5% 0 0

10% 0 0 Obs 961

25% 1 0 Sum of Wgt. 961

50% 2 Mean 2.51717

Largest Std. Dev. 2.257849

75% 4 11

90% 6 12 Variance 5.097882

95% 7 12 Skewness 1.301243

99% 10 12 Kurtosis 4.788665

sum cesd1r-cesd5r cesd\_scale

Variable | Obs Mean Std. Dev. Min Max

-------------+---------------------------------------------------------

cesd1r | 963 .4060228 .6625135 0 3

cesd3r | 966 1.026915 .7787995 0 3

cesd4r | 964 .5165975 .7479529 0 3

cesd5r | 964 .56639 .6826366 0 3

cesd\_scale | 961 2.51717 2.257849 0 12

**sum cesd1-cesd5**

**tab1 cesd1-cesd5, miss**

**\*creating recodes for each variable**

**gen cesd1r = cesd1**

**gen cesd2r = cesd2**

**gen cesd3r = cesd3**

**gen cesd4r = cesd4**

**gen cesd5r = cesd5**

**label define ced 0 "none or almost none of the time" 1 "some of the time" 2 "most of the time" 3 "all or almost all of the time".d "DK" .i " no IAP" .n "NA"**

**label define cedrc 0 "all or almost all of the time" 1 "most of the time" 2 "some of the time" 3 "none or almost none of the time".d "DK" .i " no IAP" .n "NA"**

**numlabel ced cedrc, add**

**recode cesd1r cesd2r cesd4r cesd5r (1=0) (2=1) (3=2) (4=3) (.d=.d) (.i=.i) (.n=.n)**

**recode cesd3r (4=0) (3=1) (2=2) (1=3) (.d=.d) (.i=.i) (.n=.n)**

**label variable cesd1r "how many times felt depressed in past week"**

**label variable cesd2r "how many times had restless sleep in past week"**

**label variable cesd3r "how many times felt happy in past week"**

**label variable cesd4r "how many times felt lonely in past week"**

**label variable cesd5r "how many times felt sad in past week"**

**label values cesd1r cesd2r cesd4r cesd5r ced**

**label values cesd3r cedrc**

**tab cesd1 cesd1r, miss**

**tab cesd2 cesd2r, miss**

**tab cesd3 cesd3r, miss**

**tab cesd4 cesd4r, miss**

**tab cesd5 cesd5r, miss**

**tab1 cesd1 cesd1r, miss**

**tab1 cesd2 cesd2r, miss**

**tab1 cesd3 cesd3r, miss**

**tab1 cesd4 cesd4r, miss**

**tab1 cesd5 cesd5r, miss**

**\*creating a depression scale\***

**/\*\*\*\* code for 5-item scale**

**egen cesd\_scale = rowtotal(cesd1r-cesd5r), miss**

**label variable cesd\_scale "5-item scale of depressive symptoms"**

**tab cesd\_scale, miss**

**replace cesd\_scale = .d if cesd1r == .d & cesd2r ==.d & cesd3r ==.d & cesd4r ==.d & cesd5r ==.d**

**replace cesd\_scale =.n if cesd1r == .n & cesd2r ==.n & cesd3r ==.n & cesd4r ==.n & cesd5r ==.n**

**replace cesd\_scale =.i if cesd1r == .i & cesd2r ==.i & cesd3r ==.i & cesd4r ==.i & cesd5r ==.i**

**label values cesd\_scale em**

**tab cesd\_scale, miss**

**sum cesd1r-cesd5r**

**sum cesd\_scale cesd1r-cesd5r \*\*\*\*/**

**alpha cesd1r-cesd5r, item**

**\*since cronbach alpha of cesd2r (0.7913) is greater than overall alpha of 5 items (0.7611), we will drop this item\***

**drop cesd2r**

**egen cesd\_scale = rowtotal(cesd1r cesd3r-cesd5r), miss**

**label variable cesd\_scale "4-item scale of depressive symptoms"**

**tab cesd\_scale, miss**

**replace cesd\_scale = .d if cesd1r == .d & cesd3r ==.d & cesd4r ==.d & cesd5r ==.d**

**replace cesd\_scale =.n if cesd1r == .n & cesd3r ==.n & cesd4r ==.n & cesd5r ==.n**

**replace cesd\_scale =.i if cesd1r == .i & cesd3r ==.i & cesd4r ==.i & cesd5r ==.i**

**label values cesd\_scale em**

**tab cesd\_scale, miss**

**sum cesd1r cesd3r-cesd5r**

**sum cesd\_scale cesd1r cesd3r-cesd5r**

**list id cesd\_scale cesd1r cesd3r-cesd5r if !missing(cesd1r, cesd3r-cesd5r) in 60000/60020, nolabel**

**list id cesd\_scale cesd1r cesd3r-cesd5r if missing(cesd1r, cesd3r-cesd5r) in 1/10, nolabel**

**list id cesd\_scale cesd1r cesd3r-cesd5r if missing(cesd1r) & (!missing(cesd3r) | !missing(cesd4r)), nolabel**

**alpha cesd1r cesd3r-cesd5r , item**

**\*now limiting the output scale to those with complete data on all four input scale items\***

**mvpatterns cesd1r-cesd5r, sort**

**egen nmiss\_cesd\_scale = rmiss2(cesd1r-cesd5r)**

**label variable nmiss\_cesd\_scale "Count for missing variable frequencies for 4 item scale of depression symptoms"**

**tab nmiss\_cesd\_scale, miss**

**replace cesd\_scale =.a if nmiss\_cesd\_scale >0 & nmiss\_cesd\_scale <=4**

**tab cesd\_scale**

**tab cesd\_scale, miss**

**sum cesd\_scale, detail**