**RUTGERS UNIVERSITY**

**Bloustein School of Planning and Public Policy**

**Applied Multivariate Methods**

**Fall 2020**

**Dawne Mouzon, Ph.D.**

**Problem Set #6**

**NAME: HASSAN KHURSHID\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

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

**BIVARIATE STATISTICS II**

**(100 points total)**

**Please be sure to follow all instructions for the problems and submit the specified code and output. 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.**

**Reminder: the output we ask for does not include every consistency check you should be running to check your own work. Generally, what we need to grade you is less than what you should do to check your code throughout the assignment.**

**Please use the Courier New font, size 10 to make your output line up.**

***Please use the 2016 General Social Survey for the following question.***

1. **(10 pts) Create a categorical variable (gov\_resp\_scale\_cat) based on the interquartile range for the government responsibility scale (gov\_resp\_scale). The categories on the variable should be: 1=low (<25th percentile); 2=medium (25th-74th percentile); 3=high (75th percentile or higher). All original extended missing values should be carried over to the new variable. Please submit:**
   1. **Frequencies on the new variable**
   2. **Consistency checks**
   3. **Code for the whole problem**

**sum gov\_resp\_scale, detail**

**recode gov\_resp\_scale (min/15=1) (16/23=2) (24/max = 3), gen(gov\_resp\_scale\_cat)**

**label variable gov\_resp\_scale\_cat "3-cat government responsibility scale (IQR)"**

**label define scalecat 1 "low (<25th percentile)" 2 "medium (25th-74th percentile)" 3 "high (75th percentile or higher)" .c "can’t choose" .d "DK" .i "IAP" .n "NA" .a "more than 2 missing", modify**

**label values gov\_resp\_scale\_cat scalecat**

**numlabel scalecat, add**

**tab gov\_resp\_scale\_cat, miss**

**tab gov\_resp\_scale gov\_resp\_scale\_cat, miss**

**.** tab gov\_resp\_scale\_cat, miss

3-cat government responsibility |

scale (IQR) | Freq. Percent Cum.

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

1. low (<25th percentile) | 330 11.51 11.51

2. medium (25th-74th percentile) | 654 22.81 34.32

3. high (75th percentile or higher) | 371 12.94 47.26

.a. more than 2 missing | 1,512 52.74 100.00

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

Total | 2,867 100.00

. tab gov\_resp\_scale gov\_resp\_scale\_cat, miss

10 Item Scale of |

Government's | 3-cat government responsibility scale (IQR)

Responsibilities | 1. low (< 2. medium 3. high ( .a. more | Total

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

0 | 3 0 0 0 | 3

1 | 2 0 0 0 | 2

2 | 4 0 0 0 | 4

3 | 4 0 0 0 | 4

4 | 6 0 0 0 | 6

5 | 1 0 0 0 | 1

6 | 6 0 0 0 | 6

7 | 6 0 0 0 | 6

8 | 14 0 0 0 | 14

9 | 21 0 0 0 | 21

10 | 27 0 0 0 | 27

11 | 31 0 0 0 | 31

12 | 33 0 0 0 | 33

13 | 39 0 0 0 | 39

14 | 57 0 0 0 | 57

15 | 76 0 0 0 | 76

16 | 0 69 0 0 | 69

17 | 0 78 0 0 | 78

18 | 0 87 0 0 | 87

19 | 0 85 0 0 | 85

20 | 0 87 0 0 | 87

21 | 0 90 0 0 | 90

22 | 0 84 0 0 | 84

23 | 0 74 0 0 | 74

24 | 0 0 78 0 | 78

25 | 0 0 69 0 | 69

26 | 0 0 56 0 | 56

27 | 0 0 52 0 | 52

28 | 0 0 47 0 | 47

29 | 0 0 35 0 | 35

30 | 0 0 34 0 | 34

.a. More than two mis | 0 0 0 1,512 | 1,512

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

Total | 330 654 371 1,512 | 2,867

1. **(10 pts) Create a categorical variable (cesd\_scale\_cat) based on the interquartile range for the CES-D scale (cesd\_scale). The categories on the variable should be: 1=low (<25th percentile); 2=medium (25th-74th percentile); 3=high (75th percentile or higher). All original extended missing values should be carried over to the new variable.**
   1. **Frequencies on the new variable**
   2. **Consistency checks**
   3. **Code for the whole problem**

**sum cesd\_scale, detail**

**tab cesd\_scale, miss**

**gen cesd\_scale\_cat = cesd\_scale**

**recode cesd\_scale\_cat (0=1) (1/3=2) (4/12=3)**

**label variable cesd\_scale\_cat "3-cat depression symptoms scale (IQR)"**

**label define scalecat 1 "low (<25th percentile)" 2 "medium (25th-74th percentile)" 3 "high (75th percentile or higher)".c "can’t choose" .d "DK" .i "IAP" .n "NA" .a "all missing", modify**

**label values cesd\_scale\_cat scalecat**

**numlabel scalecat, add**

**tab cesd\_scale\_cat, miss**

**tab cesd\_scale cesd\_scale\_cat, miss**

**tab1 cesd\_scale cesd\_scale\_cat, miss**

. tab cesd\_scale cesd\_scale\_cat, miss

4-item scale of |

depressive | 3-cat depression symptoms scale (IQR)

symptoms | 1. low (< 2. medium 3. high ( .a. all m | Total

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

0 | 151 0 0 0 | 151

1 | 0 250 0 0 | 250

2 | 0 168 0 0 | 168

3 | 0 146 0 0 | 146

4 | 0 0 79 0 | 79

5 | 0 0 70 0 | 70

6 | 0 0 36 0 | 36

7 | 0 0 21 0 | 21

8 | 0 0 15 0 | 15

9 | 0 0 13 0 | 13

10 | 0 0 6 0 | 6

11 | 0 0 3 0 | 3

12 | 0 0 3 0 | 3

.a | 0 0 0 1,906 | 1,906

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

Total | 151 564 246 1,906 | 2,867

. tab1 cesd\_scale cesd\_scale\_cat, miss

-> tabulation of cesd\_scale

4-item scale of |

depressive |

symptoms | Freq. Percent Cum.

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

0 | 151 5.27 5.27

1 | 250 8.72 13.99

2 | 168 5.86 19.85

3 | 146 5.09 24.94

4 | 79 2.76 27.69

5 | 70 2.44 30.14

6 | 36 1.26 31.39

7 | 21 0.73 32.12

8 | 15 0.52 32.65

9 | 13 0.45 33.10

10 | 6 0.21 33.31

11 | 3 0.10 33.41

12 | 3 0.10 33.52

.a | 1,906 66.48 100.00

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

Total | 2,867 100.00

-> tabulation of cesd\_scale\_cat

3-cat depression symptoms scale |

(IQR) | Freq. Percent Cum.

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

1. low (<25th percentile) | 151 5.27 5.27

2. medium (25th-74th percentile) | 564 19.67 24.94

3. high (75th percentile or higher) | 246 8.58 33.52

.a. all missing | 1,906 66.48 100.00

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

Total | 2,867 100.00

1. **(10 pts each, total of 80 pts) Choose the appropriate bivariate test to assess the relationships below. Remember, garbage in, garbage out. For each test, please submit your code and output from the test, as well as a standard write-up. In the write-ups, report the full *p*-value and round point estimates to two decimal places.**
   1. Government responsibility scale (**gov\_resp\_scale**) and education in years (**educ**)

**tab1 gov\_resp\_scale educ, miss**

**pwcorr gov\_resp\_scale educ, obs sig**

**scatter gov\_resp\_scale educ || lfit gov\_resp\_scale educ**

. pwcorr gov\_resp\_scale educ, obs sig

| gov\_re~e educ

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

gov\_resp\_s~e | 1.0000

|

| 1355

|

educ | -0.1418 1.0000

| 0.0000

| 1353 2858

|

A correlation was conducted to determine whether there was a significant correlation between highest number of education in years and government responsibility (10-item scale ranging from 0/least confidence to 30/most responsibility). The correlation was significant (*p* <0.001), however, value of Pearson’s r is only -0.14 indicating a “small” effect/weak or no correlation between education in years and government responsibility. The direction of this small – weak correlation is negative indicating that upon an increase in years of education, the opinions about increasing the government’s responsibility become less vocal.



* 1. Government responsibility scale in categories (**gov\_resp\_scale\_cat**) and education in categories (**educ4cat**)

**tab1 gov\_resp\_scale\_cat educ4cat, miss**

**tab gov\_resp\_scale\_cat educ4cat, chi2 col**

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

| Key |

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

| frequency |

| column percentage |

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

3-cat government |

responsibility scale | highest level of school completed

(IQR) | 1. less t 2. high s 3. some c 4. bachel | Total

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

1. low (<25th percent | 16 94 85 135 | 330

| 8.51 24.87 24.15 31.03 | 24.39

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

2. medium (25th-74th | 82 192 172 207 | 653

| 43.62 50.79 48.86 47.59 | 48.26

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

3. high (75th percent | 90 92 95 93 | 370

| 47.87 24.34 26.99 21.38 | 27.35

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

Total | 188 378 352 435 | 1,353

| 100.00 100.00 100.00 100.00 | 100.00

Pearson chi2(6) = 64.6588 Pr = 0.000

A chi square test was conducted to determine whether there was a significant association between education by category (less than high school high school diploma, some college, bachelor’s degree or higher) and government responsibility scale (low, medium, high). The chi square test was significant (p- value < 0.001) showing that in general, those who had less school education believed that government should take more responsibility than the higher-educated groups. For example, 47.9% of respondents with less than high school education believed the government should take a high level of responsibility, compared to 24.34, 26.99 and 21.38 percent for the higher-educated groups. Between the rest of the three groups who had at least high school education, there was one exception to the diminishing support trend. That was between ‘high school’ and ‘some college’ groups. So, according to our conclusion, the cohort with ‘some college’ must have had lesser percentage of folks siding with high level of responsibility than people with just high school education. But it was contrary to this. (24.44 % vs 26.99 %)

* 1. Government responsibility scale in categories (**gov\_resp\_scale\_cat**) and political party affiliation (**partyid\_rev**).

**tab1 gov\_resp\_scale\_cat partyid\_rev, miss**

**tab gov\_resp\_scale\_cat partyid\_rev, chi2 col**

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

| Key |

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

| frequency |

| column percentage |

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

3-cat government |

responsibility scale | political party affiliation

(IQR) | 1. Democr 2. Indepe 3. Republ 4. Other |Total

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

1. low (<25th percent | 43 121 154 9 | 327

| 9.73 21.80 48.58 33.33 | 24.38

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

2. medium (25th-74th | 234 266 135 12 | 647

| 52.94 47.93 42.59 44.44 | 48.25

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

3. high (75th percent | 165 168 28 6 | 367

| 37.33 30.27 8.83 22.22 | 27.37

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

Total | 442 555 317 27 | 1,341

| 100.00 100.00 100.00 100.00 | 100.00

Pearson chi2(6) = 179.4550 Pr = 0.000

A chi square test was conducted to determine whether there was an association between political party affiliation (Democrat, Republican, Independent, Other) and government responsibility scale (low, medium, high). The chi square test was significant (p- value < 0.001) showing that in general, that the Republican cohort believes government should have lower responsibility than all other groups who apparently have a high percentage of folks who believe that the government should have high responsibility. For instance, 48.6% of Republicans believed in a low level of government responsibility, compared to 9.7% of Democrats, 21.8% of Independents and 33.3% of Others. 37% of Democrats also believed in a high level of government responsibility, compared to 30.27% of Independents, 22.22% of Others, and an 8.8% of Republicans.

* 1. Institutional confidence scale (**inst\_conf\_scale**) and income (**conrinc**)

**tab1 inst\_conf\_scale conrinc, miss**

**pwcorr inst\_conf\_scale conrinc, obs sig**

**scatter inst\_conf\_scale conrinc|| lfit inst\_conf\_scale conrinc**

. pwcorr inst\_conf\_scale conrinc, obs sig

| inst\_c~e conrinc

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

inst\_conf\_~e | 1.0000

|

| 1955

|

conrinc | 0.0151 1.0000

| 0.6139

| 1116 1632

|

A correlation was conducted to determine whether there was a significant linear association between income in dollars and confidence in institutions (13-item scale ranging from 0/least confidence to 26/most confidence). The correlation was not significant (r = 0.0151; p = 0.6139), indicating no significant correlation between income in dollars and confidence in institutions.



* 1. CES-D scale in categories (**cesd\_scale\_cat**) and race/ethnicity (**race\_eth**)

**tab1 cesd\_scale\_cat race\_eth, miss**

**tab cesd\_scale\_cat race\_eth, chi2 col**

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

| Key |

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

| frequency |

| column percentage |

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

3-cat depression | recode of race variable

symptoms scale (IQR) | 1. NH Whi 2. NH Bla 3. Hispan 4. Other | Total

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

1. low (<25th percent | 109 17 21 4 | 151

| 17.19 10.83 17.21 8.33 | 15.71

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

2. medium (25th-74th | 374 90 69 31 | 564

| 58.99 57.32 56.56 64.58 | 58.69

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

3. high (75th percent | 151 50 32 13 | 246

| 23.82 31.85 26.23 27.08 | 25.60

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

Total | 634 157 122 48 | 961

| 100.00 100.00 100.00 100.00 | 100.00

Pearson chi2(6) = 8.7851 Pr = 0.186

A chi square test was conducted to determine whether there was a significant association between race/ethnicity (NH White, NH Black, NH Other, Hispanic) and the CESD-scale of depressive symptoms (low, medium, high). The chi square test was not significant (p = 0.186). For example, 23.8% of NH Whites had a high level of depressive symptoms, compared to 31.2% of NH Blacks, 27.1% of NH Others, and 26.2% of Hispanics.

* 1. Government responsibility scale (**gov\_resp\_scale**) and income (**conrinc**)

**tab1 gov\_resp\_scale conrinc, miss**

**pwcorr gov\_resp\_scale conrinc, obs sig**

**scatter gov\_resp\_scale conrinc|| lfit gov\_resp\_scale conrinc**

. pwcorr gov\_resp\_scale conrinc, obs sig

| gov\_re~e conrinc

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

gov\_resp\_s~e | 1.0000

|

| 1355

|

conrinc | -0.1523 1.0000

| 0.0000

| 770 1632

|

A correlation was conducted to determine whether there was a significant linear association between income in dollars and government responsibility (10-item scale ranging from 0/least confidence to 30/most responsibility). The correlation was significant (r = -0.1523; p <0.001), indicating a small correlation between income in dollars and government responsibility

.

* 1. Government responsibility scale in categories (**gov\_resp\_scale\_cat**) and age in categories (**age4cat**)

**tab1 gov\_resp\_scale\_cat age4cat, miss**

**tab gov\_resp\_scale\_cat age4cat, chi2 col**

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

| Key |

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

| frequency |

| column percentage |

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

3-cat government |

responsibility scale | dummy age

(IQR) | 1. 18-24 2. 25-55 3. 45-64 4. 65 or | Total

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

1. low (<25th percent | 13 94 131 92 | 330

| 11.50 20.30 27.52 30.56 | 24.39

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

2. medium (25th-74th | 56 228 231 138 | 653

| 49.56 49.24 48.53 45.85 | 48.26

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

3. high (75th percent | 44 141 114 71 | 370

| 38.94 30.45 23.95 23.59 | 27.35

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

Total | 113 463 476 301 | 1,353

| 100.00 100.00 100.00 100.00 | 100.00

Pearson chi2(6) = 28.7354 Pr = 0.000

A chi square test was conducted to determine whether there was an association between age (18-24, 25-55, 45-63, 65+) and government responsibility scale (low, medium, high). The chi square test was significant (p- value <0.001) showing that in general younger people believed that the government should take more responsibility. For example, 11.5% of those 18-24 and 20.30% of 25-44 year old folks believe in a low level of government responsibility as opposed to 27.52% of 45-64 year-olds and 30.56% of 65+ year-olds. Whereas a high percentage of youth opined for government to have greater responsibilities (38.94 %) compared to rest of the groups and their percentage of people who believed the same which were appreciably lesser (30.45 %, 23.95 %, and 23.59 % for ’25-55’, ’45-64’, and ’65 or older’, respectively).

* 1. CES-D scale (**cesd\_scale**) and age in years (**age**)

**tab1 cesd\_scale age, miss**

**pwcorr cesd\_scale age, obs sig**

**scatter cesd\_scale age|| lfit cesd\_scale age**

pwcorr cesd\_scale age, obs sig

| cesd\_s~e age

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

cesd\_scale | 1.0000

|

| 961

|

age | 0.0059 1.0000

| 0.8555

| 956 2857

|

A correlation was conducted to determine whether there was a significant link between age in years and the CESD scale (4-item scale ranging from 0/least depressive symptoms to 26/most depressive symptoms). The correlation was not significant(r = 0.0059; p = 0.8555), indicating no significant correlation between age in years and depressive symptoms