Week 2 – Data exploration & cleaning

| Question | Answer |
| --- | --- |
| 1. What is your name? |  |
| 1. When conducting statistical analyses, we should always keep in mind the larger context in which we are working with data. What are steps in the research process which are completed before we analyze data? (Hint: This was on the week 1 assignment.) | 1. Generate research questions (Field 1.4) 2. Generate research hypotheses (Field 1.5) 3. Determine study design, variables, levels of measurement (Field 1.6, LSWJ 2.2) 4. Collect data |
| 1. Data needs to be examined and any problems fixed before analyses can be done. In statistics, what are 4 main areas to examine as part of pre-analysis checking? (Hint: This was on the week 2 assignment.) | 1. Data accuracy 2. Missing data 3. Outliers 4. Statistical assumptions |
| 1. What are important steps to complete when checking data accuracy? (Hint: This was on the week 2 assignment.) | 1. Make sure the data types are correct. (What is the level of measurement of each variable?) 2. Check the data for typos. (i.e. mailes for males, feemales for females) 3. Check the data for nonsensical values. (i.e. 1000 year old people) 4. Check categories make sense. (i.e. variable is gender but has a category of blue) 5. Correct problems if possible or omit the data from the analyses. 6. Reverse code instrument items if needed. 7. Calculate summary scores for any instruments that need it. 8. Keep track of what you do so you can report it as part of the analysis steps you completed. (Be transparent.) |
| 1. Open the Week3.rds dataset in Jamovi. (Note any difficulties.) | No issues |
| 1. What are the variables in the dataset and the level of measurement for each one? Make sure the level of measurement for each variable is correct in Jamovi. (Notice the first one is done for you.) | 1. sex – nominal (no problems noted) 2. Research-nominal (no problems) 3. Height -continuous no problems 4. Weight- continuous -no problems 5. Lvst 1- changed to ordinal ranking of something 6. Lvst 2- changed to ordinal ranking of something 7. Lvst 3- changed to ordinal ranking of something 8. Lvst 4- changed to ordinal ranking of something 9. First 1- changed to ordinal ranking of something 10. First 2- changed to ordinal ranking of something 11. First 3- changed to ordinal ranking of something |
| 1. Check the dataset for accuracy using visual inspection and descriptive statistics. Note any problems. | Didn’t seem to have any issues |
| 1. Create a new variable in Jamovi for BMI. The formula for BMI is weight (kg)/height (m)^2. Weight is already in kg and height is in cm. | You will need to create a computed variable for this. |
| 1. Create a new variable in Jamovi for lvst4\_recode. Change the values in lvst4 according to the following pattern: 1=5, 2=4, 3=3, 4=2, 5=1. | You will need to create a transformed variable for this. |
| 1. Save your Jamovi (.omv) file. You will turn it in. |  |
| 1. Open the .Rmd file. Note any difficulties | None noted |
| 1. Run the code blocks. Note any difficulties. | This is error I received  You can uncomment the code below to try different ways to examine the dataset in RStudio. (Remove the # sign from one line at a time.)  ```{r, eval=FALSE}  The View() command opens a data tab in RStudio. Notice the eval=FALSE option is set for this block.  Don't evaluate blocks which use the View() command when knitting a file. It could be problematic.  View(dataset)  ``` get error unexpected symbol in the view |
| 1. What level of measurement are the newly created LOSS and FOSS variables? | 1. LOSS\_total – ordinal only using numeric instead of characters 2. FOSS\_total – ordinal only using numeric instead of characters |
| 1. What descriptive statistics and plots should you use to examine the new variables? | 1. LOSS\_total – central tendency and standard deviation, normal distribution using histograms 2. FOSS\_total – central tendency and standard deviation, normal distribution using histograms |
| 1. What are some advantages of using an average to score an instrument over a sum? | * Final scores are in same scale for easier interpretation   -can calculate average with missing data |
| 1. What units are z-scores in? | It is the standard deviation from the mean |
| 1. What is the meaning of the following z-scores? | 1. Negative z-score – below mean average 2. Positive z-score – above mean average 3. Z-score of 0 – it is the mean score 4. Z-score with absolute value > 2 –more than 2 standard deviations from the mean |
| 1. Why are z-scores useful? | Allows for calculation of probability of something happening with in a standard deviation |
| 1. What impact does scoring an instrument using sum vs average have on a z-score? | Z score is based on the standard deviation from the mean so we need the mean or the average in order to obtain the z score |
| 1. Are there any z-scores in the LOSS and FOSS variables with an absolute value > 2? If so, how many in each? | 1. LOSS\_total – 1 2. FOSS\_total – 2 |
| 1. After you finished reading through the .Rmd file. Knit the file to markdown and save the .md file. You will turn the file in as part of your assignment. Describe any difficulties. If you encounter errors close all instances of RStudio and try opening it again. (Don’t save your workspace.) If you encounter errors while creating the file and can’t figure out how to fix them before the assignment is due. You can paste the text of the .Rmd file with the output into a Word or pdf file and turn in that. |  |
| 1. What did you like/dislike about this assignment? | I had some issues getting the files to run  I think the data is much easier to read in jamovi |
| 1. How would you change this assignment to make it better for future students? | Not sure we can |
| 1. Complete the answer submission in Canvas and turn in your RStudio and Jamovi files. (The answer submission in Canvas allows multiple attempts. You may want to save the text you provide in open responses like the two questions above so you can just paste it into the answer field if you make additional attempts to improve your score.) |  |