**Class 3 Activity**

**Frequency Tables and Histograms**

**Part 1**

On a scrap piece of paper, answer the following questions (put your name on it and turn it in! I have scrap paper if you need some).

You ask a sample of UD students how much they have enjoyed their time in college so far on a scale of 1 (not at all) to 5 (very much). You get the following data:

1, 2, 1, 4, 5, 4, 5, 5, 5, 5, 3, 4, 2, 5, 3, 3, 4, 3, 2, 4, 5, 5, 5, 5

1) Is this measure ordinal, scale, or nominal?

2) Make a frequency table of these scores (include the value, frequencies and percentages)

3) Would it make sense to display these data as a grouped frequency table? Why or why not?

4) Make a histogram from the frequency table

5) Describe the general shape of this distribution. Specifically, modality and symmetry/skewness.

**Part 2**

Download Class3.csv and open the data in JASP (note: these are not real data). Complete the following questions on this document and **upload your answers to Part 2 on** **Canvas** by the end of class!

Participants were asked to report their age, gender, levels of depression (total score on a depression inventory ranges from 0 to 29), anxiety (total score on an anxiety inventory ranges from 0 to 19), and a rating of how much they like statistics on a scale of 1 (not at all) to 10 (very much).

1) Did JASP assign the correct level of measurement for each variable? If not, change them (*hint: remember, these are the icons next to the variable names*).

2) How many participants do we have data for?

100 participants

3) Click on “Descriptives>> Descriptive Statistics” at the top of JASP. Select the scale variables (so, except subject ID & gender) and move them into the “variables” box by hitting the arrow ( ). You should now see “Results” to the right of your screen. What information does this give you about the variables?

This gives the valid, missing, mean, standard deviation, minimum and maximum values for each of the variables.

4) Now, expand the “plots” category, and select “distribution plots”. This should create histograms in your results section. Describe the distributions just by the way they look (skewed, normal, how many modes?). *Tip: if you click “display density,” it will give you a trendline that can be helpful in seeing the overall distribution. If you don’t get a trendline for all 4 variables, you may have the level of measurement assigned incorrectly…*

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable** | **Skewed?**  **If so, + or -?** | **Normal?**  **Yes or no?** | **Unimodal, bimodal, etc?** |
| Depression | - | yes | unimodal |
| Anxiety | - | no | multimodal |
| Rating | symmetric | yes | unimodal |
| Age | no | no | multimodal |

5) Now, expand the “Statistics” category, and select “Skewness” and “Kurtosis”. Find these values in the results section. You can get a sense of if a distribution is skewed from looking at a histogram (kurtosis is usually harder to tell by looking at it), but these statistics are more accurate. Fill in the box below with the values (remember, the closer to 0, the more normal/symmetrical the distribution. Skewness > |1| is generally above the cutoff for what we consider skewed, and kurtosis > |2| is generally above the cutoff):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Skewness** | **Above Cutoff?** | **Kurtosis** | **Above Cutoff?** |
| Depression | -0.907 | no | -0.063 | no |
| Anxiety | -1.401 | yes | 1.237 | no |
| Rating | -0.116 | no | -0.376 | no |
| Age | 0.198 | no | -1.740 | no |

6) Gender is a nominal variable. If I asked you to “visualize” gender (AKA make a graph/plot), what type of graph would you use? How would this be different from a histogram?

I would make a bar graph. This is different from a histogram because the bars are all separated not connected.