**Module 2 Review**

**You are not required to respond to the following list of statements and questions. However, the list may help you to summarize Module 2 and, thus, prepare for Exam 1. Moreover, being able to address these statements/questions will help to further your understanding of the key concepts covered in this class.**

(1) What is the difference between the frequency distribution and the relative frequency?

(2) What types of data can be summarized using the frequency distribution and the relative frequency?

(3) Why is the relative frequency more insightful than the raw frequency?

(4) What does a t-test do?

(5) A mean (i.e., average) comparison can be conducted across levels of and

variable types.

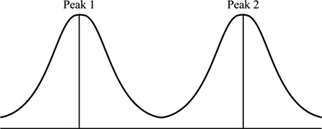
(6) Examining mean differences across levels of sex (i.e., male vs. female) is relatively easy because sex is a categorical variable and the dataset can be stratified by the respective dummy codes (i.e., 1 vs. 2). However, examining mean difference across levels of a continuous variable (e.g., conscientiousness) is more different. A may used to stratify the continuous variable and, thus, help to assess if a mean difference exists across levels of a continuous variable.

(7) Know the difference between mean, median, and mode.

(8) Descriptive statistics can be referred to as the to more sophisticated, in-depth analyses. What does this statement mean?

(9) What is standard deviation mean and why is it a useful descriptive statistic?

(10) Imagine that you are an HR Analyst who is interested in measuring the level of stress among your organization’s employees. You plot their workplace stress scores and observe the following distribution:



a. What type of distribution is being observed?

b. Would you trust the mean workplace stress score? Explain your answer.

(11) Measures of central tendency like the mean are not very informative unless the corresponding measures of , like are also reported.

(12) What can threaten descriptive statistic estimates (e.g., mean)?

(13) There are several “types” of missing data. What type of missingness cannot be ignored and why?

(14) Imagine that you are an HR Analyst who wants to know if whites perform better on a job applicant test than blacks. Specifically, you want to know if disparate treatment or disparate impact in occurring.

a. According to the NHST framework, what does the null hypothesis propose when examining this type of question?

b. You perform a t-test and observe a statistically *non-significant* *p*-value. What do you conclude?

c. You perform a t-test and observe a statistically *significant* *p*-value. What do you conclude?

(15) Should the t-test result (i.e., the p-value that tells us if a mean difference does [does not] exist) always be trusted? Explain your answer.

(16) What is the difference between a t-test and ANOVA?