REVIEW SHEET FOR SESSION 2

1. Know how to interpret the values in a **frequency distribution** and a **cumulative frequency distribution** expressed as counts or percentages.

2. Know how to interpret a **frequency histogram** and a **cumulative frequency histogram**.

3. Know how to interpret a **frequency polygon** and a **cumulative frequency polygon**.

4. Know how to convert a **stem-and-leaf plot** back to a list of scores.

5. Be able to recognize positive and negative skewness, and identify outliers using a **box-and-whisker plot**.

6. Know that the mean, median and mode are all measures of **central tendency**.

7. Know that the arithmetic mean, which is usually simply called the **mean**, is the same as the average.

8. Know that the median is the value that divides the distribution such that half of the scores are above the median and half are below the median.

9. Know that the **mode** is the most frequently occurring value.

10. Know that the mode should be used as the measure of central tendency when the data are categorical.

11. Know that the median should be used to measure central tendency when a numeric variable is highly skewed.

12. Know that the mean should be used to measure central tendency when a numeric variable is more or less symmetric.

13. Know that measures of **dispersion** indicate how spread out the scores are.

14. Know that the **range**, **mean absolute deviation**, **variance**, and **standard deviation** are all measures of dispersion.

15. Know that standard deviation is the square root of the variance.
16. Be able to identify \( \frac{\sum_{i=1}^{N} (X_i - \mu)^2}{N} \) as the formula for variance of the population and know that it is represented by the symbol, \( \sigma^2 \).

17. Be able to identify \( \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1} \) as the formula for calculating the sample variance and that it is represented by the symbol, \( S^2 \).

18. Know that \( S^2 \) is an unbiased estimate of \( \sigma^2 \), and that the sum of squares is divided by \( n-1 \) to correct the bias that would have occurred if we had simply divided through by \( n \).

19. Know that each row of an SPSS data matrix contains the values for a case (e.g., a person), and each column contains the values for a particular variable (e.g., an item on a questionnaire).

20. Know that statistical software packages must address the fact that data are not generally available for every case on every variable. When the respondent has not selected an option or no option is applicable, the value is considered missing, and data that are not missing are called valid values or valid responses.

21. Know that statistical software packages, such as SPSS, generally restrict the analyses to valid responses.

22. Know that frequency distributions generated by SPSS produce columns labeled percent and valid percent. Percent is the percentage of total responses for each response category. The total number of responses is the number of valid responses and responses indicating that the data were missing. The valid percent is the percentage of valid responses for each response category. For example, let’s say a survey is administered to 100 elementary school students. Fifty of these students attend an after school program, and 50 do not. The first question on the survey was “Do you like the activities of the after school program?”. Of the 50 students who attended, 40 indicated yes, and 10 indicated no. In this example, 40 percent of the students said they like the activities [i.e., \( (40/100) \times 100 \)], but the valid percent was 80 [i.e., \( (40/50) \times 100 \)].