Review for Test 1

Chapters 1, 2, 3, and 10

YOU NEED TO KNOW

1. The different sampling techniques: SRS, cluster sample, stratified sample, systematic sample, multistage sample, convenience sample, volunteer sample (you need to know which ones are based on chance, thus good sampling techniques)
2. The issues we need to deal with when sampling: undercoverage, nonresponse, response bias, and bias in general
3. The difference between observational studies and experiments (and why we prefer experiments)
4. Surveys—wording effect
5. Experiments: treatment, treatment group, placebo, control group
6. The best possible experiment to avoid bias: randomized controlled experiment (double-blind, if possible)
7. Blind, double-blind experiments, blocking, matched-pair design
8. Issues with experimentation: lack of realism, unethical (illegal) issues
9. How to outline the design of the experiment
10. Where and why we use random assignment
11. How to identify categorical and quantitative variables
12. How to construct a frequency, and relative frequency tables and histograms
13. What type of graphical representations we use to display categorical variables—pie chart, bar graph. You need to know how to construct and interpret them.
14. What type of graphical representations we use to display quantitative variables—dotplot, histogram (frequency, and relative frequency), stem-and-leaf (also split)
   How to describe distributions—look for the overall pattern, center, spread and outliers
15. Asses from a graphical display whether the shape of a distribution is roughly symmetric, skewed, or neither, also asses whether the distribution has one or more peaks.
16. Describe the distribution by giving numerical measures of center (mean, median, and mode) and spread (range, IQR, and variance/standard deviation)
17. The basic properties of the standard deviation
18. How to find the five-number summary and construct a boxplot.
19. How to indentify outliers using the 1.5IQR rule, and construct a modified boxplot
20. How to “read” boxplots (like the True or False questions in Quiz 2)
21. Which numerical summary is more appropriate for the distribution (mean and standard deviation, or median and IQR)
22. How the skewness and outliers of a distribution affect the mean and the median
23. How to find z-scores, and decide if the value is usual or unusual.
24. How to use the Empirical rule.
25. How make scatterplots to display the relationship between two quantitative variables, and describe the scatterplots
26. Whether or not it is appropriate to use correlation to describe the relationship between two quantitative variables
27. How to find the correlation coefficient
28. The basic properties of the correlation coefficients
29. Be able to match scatterplots with their correlation coefficients.
30. How to find the equation of the least squares regression line either by using your calculator or the two formulas.
31. The basic properties of the least squares regression line
32. Sketch the regression line on your plot
33. How to use a regression line to predict values of the response variable
34. What extrapolation means.
35. Understand that correlation (association) does not imply causation. Be able to think of lurking variables in a situation.

You should review the slides, the class worksheets (you can try to do those we didn’t do in class; that would be a good practice for you), the summary pages from the website, and the book. You are responsible to know everything (and only that) I talked about in class.

Email me if you have any questions: andrea.nemeth.898@csun.edu

Have a good weekend, and study hard!