Notions worth repeating before Final (Google Doc)

- Interpretations necessary. Integrate descriptive and inferential results.
- Inferential results (confidence intervals and *P*-values) should be cognizant of parameter being inferred. Examples of parameters:
    - population correlation coefficient (rho); estimator is *r*
    - slope coefficient parameter(s) (beta$_i$); estimator are b_i
    - binomial parameter (*p*); estimator is "p-hat"
    - relative risk parameter (*RR*); estimator is "RR-hat"
    - odds ratio parameter (*OR*); estimator is "OR-hat"
- Interpretation of confidence intervals
    - Intends to capture location of *parameter* (consider entire interval)
    - CI length quantifies precision (half confidence interval length = margin of error)
    - With due caution, and not over-simplifying, can be used to judge significance at various levels, e.g., a 95% CIs can be used to judge statistical significance at alpha = .05 level, a 90% CIs can be used to judge statistical significance at the alpha = .10 level, and so on.
- Interpretation of *P* values - Quantifies evidence *against* the null hypothesis, and nothing else (therefore, you must know and be aware of H_0)
    - "Significance" language without context will surely be misinterpreted; see: Cohen,1994 (link active)
    - *Not* a measure of effect size
    - No sharp boundaries (surely God loves *P* = .05 as nearly much as *P* = .06)
- When using inferential methods, i.e., *P*-values and CIs, be aware of assumptions (e.g., L.I.N.E.; expected values more than 5; etc. etc.)
- Systematic errors in public health research (e.g., information bias, selection bias, confounding) more important than random errors (CIs and *P*-values do *not* address systematic error)
- Be aware of sample types
    - Single (e.g., Ch 16), independent (e.g., Ch 17), paired (e.g., §18.6), case-control (18.5)
    - Experimental vs. observational designs
    - Naturalistic vs. cohort vs. case-control observational samples
- Computer doesn't replace knowledge:
    - Know what you are looking for
    - Computer output can be misleading: Illustration: SPSS output: Despite what output says, these are not risk statistics (RR NOT risk) + *RR* statistics do not apply because this is a case-control study.
- Illustrations
    - Sample size for estimating *p* (exercises 16.19 and 16.20, p. 371)
    - Sample size for comparing proportions: Lab 3B
    - Naturalistic/cohort sample (prison.sav), SPSS output -- Lab 3B
    - Stratified analysis: confounding and interaction -- Lab 5
    - Labs 6: simple regression