PSYC 5143
Spring '24
Exam 1

In completing this exam, you may use any resource except other people. Each problem's point value is listed before the problem. There are **48** possible points; your grade will be the percentage of points you earn. Point values do not necessarily indicate relative ease or difficulty. Please submit your answers in an R file along with *all* R commands you use to arrive at your answers (and nothing superfluous). Avoid theorizing about putative causes of results; really. This is due by 2pm on Monday, March 11, via Blackboard. Good luck!

1) **(16 points)** The data in diabetes.csv has (real) data in it (details here), including several variables; of interest in this question are two of these variables: age and BMI. Significance is of no importance in any of the questions below.

    a. **(2)** Does there appear to be a nonlinear relationship between age (the predictor) and BMI (the outcome)? Generate a scatterplot with a LOESS curve added to it to answer this question (and please do answer it; yes, pictures are worth many words, but I shouldn't be supplying the words here).
    b. **(2)** Model BMI (the outcome) as a function of age (the predictor). Interpret the slope briefly.
    c. **(4)** Without centering age, model BMI on age and its square. Interpret both slopes (but not the intercept) briefly; informatively say what the numbers tell you (e.g., don't simply say there's a positive relationship).
    d. **(4)** Mean-center age and then model BMI on mean-centered age and its square. Informatively interpret the intercept and the slopes briefly.
    e. **(2)** What is the linear slope of age in the quadratic model when age = 60? Find this any way you can.
    f. **(2)** In the plainest language you can use, what is the difference between the linear slope of age in part b and the linear slope of age in part d? (I'm not asking you to do subtraction here to find a difference; I'm asking what the difference is in the *interpretation* of these slopes.)

2) **(10 points)** Briñol et al. (2013)[1] reported an experiment in which they asked students (who were taking a course designed to prevent eating disorders) to write down on paper either positive ***or*** negative thoughts (this is the first factor in this experiment) about their own body. After listing these thoughts, all subjects looked at their thoughts and then either threw them in the trash ***or*** were asked to continue looking at their thoughts and check for grammar & spelling errors (this is the second factor in the experiment); the latter condition is the control. After these manipulations, subjects were given a questionnaire asking about their body attitude, scored so that higher scores indicate a more-positive body attitude. Data that closely match the actual outcome of the study appear in trash.csv. The researchers were interested in the interaction being significant and establishing that negative thoughts would lead to a poorer body attitude only in the control condition, not in the treatment condition.

    Fit a model or models that allow you to (i) test the interaction, (ii) test the (simple) effect of thought-types in the treatment condition, and (iii) the (simple) effect of thought-types in the control conditions. Do not worry about FWER or FDR. Draw informative conclusions for each of the three hypotheses of interest, citing inferential and descriptive statistics in support.

[1] Treating thoughts as material objects can increase or decrease their impact on evaluation. *Psychological Science*.

3) **(22 points)** Assume that the data in rewards.csv represent the effects of food and/or water deprivation on rat behavior in a learning task. The variables are *cond* (for condition/treatment) and trials (the DV)! Treatments 1 and 2 represent control conditions in which the animal received either ad lib food and water (1) or else food and water twice per day (2). In Treatment 3 animals were food deprived, in Treatment 4 they were water deprived, and in Treatment 5 they were deprived of both. The DV, learning, is measured by the number of trials needed to reach a predetermined criterion (i.e., lower numbers indicate better performance). Assume that before running the experiment, the following comparisons were planned:

   i. The two control groups against one another
   ii. The two control groups combined versus the three experimental groups combined
   iii. The two singly-deprived treatments combined versus the doubly-deprived treatment
   iv. The singly deprived treatments against one another

   a. **(2)** Briefly argue whether or not you should do the omnibus ANOVA (i.e., where Model C is an intercept-only model and Model A is one with all predictors in it). (There is no wrong answer to the yes-no aspect of this question; it's your arguments that will be evaluated.)
   b. **(1)** Verify *yes* or *no* that the contrasts implied by these comparisons are orthogonal.
   c. **(1)** Given the comparisons of interest, what method of controlling FWER or FDR would you use here?
   d. **(1)** How many predictors are necessary to fully code group membership in this design?
   e. **(6)** Create contrasts to test the comparisons listed above and fit the model. Report the results, drawing brief conclusions for each contrast, citing descriptive and inferential statistics to support your claims.
   f. **(2)** You *could* use the `TukeyHSD` or the `pairwise.t.test` function (the latter implementing something like the Bonferroni or Benjamini-Hochberg procedure to keep Type I errors under control) in R to compare *all* pairs of groups against one another rather than testing the four comparisons above. Given that both of these functions assume that you'll conduct ten hypothesis tests rather than only the four specified in *i* through *iv* above, what are some strengths and/or weaknesses of using these in place of testing the four *a priori* comparisons of interest.
   g. **(3)** For the third hypothesis and its associated parameter estimate, find *PRE* (do this the easiest way you can) – do not worry about bias! – convert it to f^2, and specify what sample size would be needed to replicate this finding with power = .90. Out of respect for dearly departed Jacob Cohen, if you use G*Power, report *all* of the details, analogous to the footnote below:

> [1] To determine the target sample size, the reported effect size for the disfluency-memory boost (η^2 = 0.031) was converted to Cohen's d = 0.3577; to achieve 95% power with a dependent-samples t test at alpha = .05 would require N = 86 (one-tailed) and N = 104 (two-tailed). Note that

   h. **(2)** Add residuals for your model from part e to the data alongside the original DV. Fit the same model you did in part e with the residuals rather than the DV as the outcome. Using whatever function you like to get *SSR* for each contrast and *SSE* for the whole model, how do the *SSR*s & *SSE* compare to the model from part e? Say why things that changed or stayed the same ... changed or stayed the same.
   i. **(2)** Fit the same model from part e but without the contrast associated with comparison ii. Focusing only on the slopes and their SEs, what changed from the part e model to this model and why? (The *why* part may be hard. Ask for a hint if needed.)
   j. **(2)** Finally, by whatever means necessary, explain why using the original *cond* variable (which is numeric!) as a predictor rather than the contrasts you created (or a factor version of *cond*) will give you results that are difficult to interpret.