STA 2201/442 Assignment 8

Please bring your R printouts for the parts to the quiz. Some may be handed in. *Do not write anything on your printouts in advance except maybe your name (and student number, if you wish).*

The file mcars2.data (see link from course website). has four variables: country of origin (presumably the location of the head office), fuel efficiency in kilometers per litre, length in centimeters and weight in kilograms. Your client gives you these data, and asks two questions.

- Which country produces the most fuel efficient cars?
- Once you allow for size of car, which country produces the most fuel efficient cars?

The typical request from a client is much less clear than this, so please think a bit about what you'd do. Then read the following for some guidance you'd never get in practice.

1. Of course you're going to give the client information about differences in *average* (mean) fuel efficiency, even though that was not requested explicitly. Furthermore, though an unsophisticated client might be content with sample means for the first question, you're going to give her conclusions based on hypothesis tests — because there's random noise in any data set, and hypothesis tests help to filter it out.

2. You'll test for differences among the three means, and then if you reject H_0 at $\alpha = 0.05$, follow up with pairwise comparisons that limit the joint probability of Type I error to 0.05. Do it the easiest way you can. Be able to state the conclusions in plain, non-statistical language. This means that you need to say something about automobiles.

3. Since the client asked for it, you'll start by doing the analysis without the covariates.

4. Then put the covariates in. Start with a model that assumes parallel planes. This is probably what the client expects. If you reject the null hypothesis at $\alpha = 0.05$ (what is H_0?), follow up with pairwise multiple comparisons. I believe the only way you know how to do this is with a Bonferroni correction, so please do that. Calculate the Bonferroni-adjusted p-values with R, so they are on your printout. Again, be able to state the conclusions in plain, non-statistical language. Are your conclusions different from what they were without the covariates?

5. Now, rather than *assuming* parallel planes, you'd better test it.

6. With the covariates set to their sample mean values, is there evidence that average fuel efficiency depends on country of origin?

   (a) Write the null hypothesis in the form $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$

   (b) Do the test and answer Yes or No.

   (c) If the answer is yes, proceed with Bonferroni-corrected multiple comparisons. Again, calculate the Bonferroni-adjusted p-values with `R`, so they are on your printout. Do your conclusions change when the slopes are not parallel?

I could be more precise, but I want you to think of the *intent* of the questions, and do the most reasonable thing.