STAT 425 **Applied Regression and Design** **Spring 2020**

Homework No. 3

1 Instructions

Due Date: Friday Mar 13th, in Class
Homework presentation should be neat. You can submit the homework on loose leaf paper or submit through Compass. R codes should be submitted through Compass. **Please submit your digital files in .pdf or .html format**, so it is easier for the graders to open your files. If you submit in paper and more than one sheet of paper is used, the assignment should be stapled together. You must show all work for full credit. If you feel it would help, you are encouraged to work together on Homework, but you have to present assignments individually using your own words. The aim of the Homework is to learn the material and practice for the exams. **Late assignments will not be accepted**. Graduate students should attempt **all** problems. Undergraduate students can skip problems marked as GR.

2 Problems

1. **Problem 1**: Using the *sat* dataset, fit a model with the *total SAT score* as the response and *expend, salary, ratio and takers* as predictors. Perform regression diagnostics on this model to answer the following questions. Display any plots that are relevant. Do not provide any plots about which you have nothing to say. Suggest possible improvements or corrections to the model where appropriate.

   (a) Check the constant variance assumption for the errors.
   (b) Check the normality assumption.
   (c) Check for large leverage points.
   (d) Check for outliers.
   (e) Check for influential points.
   (f) Check the structure of the relationship between the predictors and the response.

2. **Problem 2**: Using the *teengamb* dataset, fit a model with *gamble* as a response and the other variables as predictors. Answer the questions posed in the previous question.

3. **Problem 3 (GR)**: Using the *prostate* dataset, fit a model with *lpsa* as a response and the other variables as predictors. Answer the questions posed in the previous question.

Note: All datasets are from the faraway library in R