Stat 6021: Guided Question Set 7

Car drivers like to adjust the seat position for their own comfort. Car designers find it helpful to know where different drivers will position the seat. Researchers at the HuMoSim laboratory at the University of Michigan collected data on 38 drivers. The response variable is *hipcenter*, the horizontal distance of the midpoint of the hips from a fixed location in the car in mm. They measured the following eight predictors:

- x_1: *Age*. Age in years
- x_2: *Weight*. Weight in pounds
- x_3: *HtShoes*. Height with shoes in cm
- x_4: *Ht*. Height without shoes in cm
- x_5: *Seated*. Seated height in cm
- x_6: *Arm*. Arm length in cm
- x_7: *Thigh*. Thigh length in cm
- x_8: *Leg*. Lower leg length in cm

The data are from the `faraway` package in R. After installing the faraway package, load the `seatpos` dataset.

1. Fit the full model with all the predictors. Using the `summary()` function, comment on the results of the t tests and ANOVA F test from the output.
2. Briefly explain why, based on your output from part 1, you suspect the model shows signs of multicollinearity.
3. Provide the output for all the pairwise correlations among the predictors. Comment briefly on the pairwise correlations.
4. Check the variance inflation factors (VIFs). What do these values indicate about multicollinearity?

5. Looking at the data, we may want to look at the correlations for the variables that describe length of body parts: *HtShoes*, *Ht*, *Seated*, *Arm*, *Thigh*, and *Leg*. Comment on the correlations of these six predictors.

6. Since all the six predictors from the previous part are highly correlated, you may decide to just use one of the predictors and remove the other five from the model. Decide which predictor out of the six you want to keep, and briefly explain your choice.

7. Based on your choice in part 6, fit a multiple regression with your choice of predictor to keep, along with the predictors $x_1 = Age$ and $x_2 = Weight$. Check the VIFs for this model. Comment on whether we still have an issue with multicollinearity.

8. Conduct a general linear F test to investigate if the predictors you dropped from the full model were jointly insignificant. Be sure to state a relevant conclusion.

9. Produce the diagnostic plots for your model from part 7. Comment on whether the regression assumptions are met.

10. Based on your results, write your estimated regression equation from part 7. Also report the R^2 of this model, and compare with the R^2 you reported in part 1, for the model with all predictors. Also comment on the adjusted R^2 for both models.