Calories and Sodium in Hot Dogs

Name of Group Members:

Note: For this computer lab, everything should be done in RStudio. You will create a PDF document at the end that can be printed at the end of class.

Use the data on my website recording the calorie and sodium in different meat hot dogs. Test to see if there is a significant relationship between calories and sodium in meat hot dogs.

1. What are the null and alternative hypothesis?
2. Read in the data using the following command
    a. hotdogs<-read.csv("http://mathcs.muhlenberg.edu/~davidson/Hotdogs.csv")
3. Create a scatterplot of the data. Describe the general relationship including form and direction.
4. Fit a linear model (add trendline and run analysis under REGRESSION section)
5. Calculate the correlation between calories and sodium. Is this strong moderate or weak?
6. Based on the correlation, what is the R^2 value? What does this mean in terms of the problem?
7. Based on the output of the model, what is the R^2 value? Do your answers match?
8. Test the assumptions of the model:
    a. Are there any patterns in the residual plot?
    b. Are residuals normally distributed (qqplot)?
    c. Is scatterplot linear?
9. What is the equation of a line?
10. If there is a new brand of hot dog with 150 calories, how many milligrams of sodium do you estimate that one of these hotdogs contains?
11. There is a point (107, 144) of (calories, sodium). What is the residual for this point?
    a. Residual = y(obs) – y(pred)
12. What are your conclusions based on the analysis you've run?
13. There seems to be an outlier in the data (in fact, the same point mentioned in the residual question). The analysis was run again without this outlier present and the correlation was found to be 0.834 (with an R^2 value of 69.5%). Do you think this is an "influential" point (i.e. does it drastically change the results?)

Note: If you want to add the confidence bands to your scatterplot, in your xyplot command, simply add

```
panel=panel.lmbands
```