Math 158, Spring 2009
Jo Hardin
semester project (part 2)

Simple Linear Regression

Your task for this project is to apply the tools we've learned in chapters 1, 2 & 3 (Simple Linear Regression) to answer questions about the relationship between two continuous (quantitative) variables. The report should include:

- Introduction (Briefly refresh the reader's mind as to the variables of interest)
- The hypotheses that you'll be addressing. It will probably be that the two variables are linearly related. (Positively? Negatively? Remember, R gives a two-sided p-value, but you can just as easily test that $\beta_1 > 0$ or $\beta_1 < 0$.)
- Check the assumptions for linear regression. Look at the plots of explanatory vs. response and residual vs. predicted (include only what is interesting, which might mean that you don't include a residual plot.) Comment on whether you think the data are linear with constant variability. If not, try transforming the data. Remember, transforming X gives a different relationship between X & Y (might make the relationship more linear); however, transforming Y changes the variability around the line (might make the st. dev. more constant & the relationship different.)
- Compute the test of β_1 (or other test from above) or find a CI for β_1. Remember that if you have transformed data, you should be careful about your interpretations. Your test or CI should have an interpretation in the words of your variables.
- Plot your (transformed?) variables, try to think of one as explanatory and the other as response. Give the reader a CI for both the mean and individual response at some interesting value of the explanatory variable. (That is, at some x-value that is interesting to you.) Interpret these intervals.
- Asses the fit of your model. Discuss the R^2 value and the residual plot(s).
- A Conclusion (Summarize your results. Comment on anything of interest that occurred in doing the project. Were the data approximately what you expected or did some of the results surprise you? What other questions would you like to ask about the data?)

Notes:

- Summarize any output from R. Do not turn in the print outs, but make new tables and summarize so that it flows nicely in the text. I don't need to see the technical calculations.
- I've asked you to do a series of things above, make sure the sections flow nicely into one another. This is a report on the data not a homework assignment. (Try to tell a good story.)
- Do not be tempted to turn in everything you do. Only turn in the interesting parts of the analysis. One of the hardest parts of being a consultant is figuring out what to tell the researcher.
- Computer output that is attached and not described will be ignored.
- Please double space your work.
- Turn in the previous projects with this project. (But let this assignment stand alone, that is, don't expect me to remember your variables.)
- Remember to label all graphs, email me if you are having trouble in R with labels (or really, any troubles in R!)