Math 124 Spring 2005
Assignment 3
Due: May 17, 2005

Dr Ben Bolstad
bolstad_math124@bmbolstad.com
http://math124sfsu.bmbolstad.com

Introduction and important notices

This assignment is intended to give you practice using Excel to perform regression analysis. You will be able to use some of the techniques you used on earlier reports, along with some new methods for fitting simple linear regression lines. You should submit your solution to this assignment as a written **typed** report. **Do not** submit print outs of raw Excel spreadsheets (ie I do not want to see a print out of all the data). You should **copy and paste** your graphs and other analysis results into your report. **Do not** wait until the last moment to start this assignment, rushing on the last day will result in a poor report (along with creating more end of semester pressure for yourself). **Do** ask questions about the assignment if you need help. **Do** feel free to work with your friend to learn Excel, but **do not** try to hand in the same report. Your final report should be your own work. **Remember** to look at the example reports on the website.

Regression and Exploratory Data Analysis for New 2004 Cars

On the website you will find a data file containing 21 variables, measured for 428 models of cars produced in the year 2004. In addition you will find a text file explaining each of these variables in more detail.

Data Sources: Kiplinger's Personal Finance, http://www.amstat.org/publications/jse/jse_data_archive.html

Your task

Explore this dataset using the techniques discussed in class. Where appropriate fit linear regression lines. The suggested analysis below should be used as a initial guide, but you are free to carry out additional analyzes. It is expected you will follow appropriate conventions by accurately labeling your plots, tables etc. Don't forget to identify outliers.
A satisfactory report (ie at the "B" level) will, at the least, include the following:

- A histogram of Horsepower values.
- A histogram of Weights.
- Give a boxplot of City MPG broken down by Vehicle type.
- Give a boxplot of Weights broken down by Vehicle type.
- A scatterplot of Dealer cost against Horsepower.
- A linear regression to predict Dealer cost based on Horsepower.
- Use your linear regression line to predict the Dealers cost when a vehicle horsepower is 150.

A stellar report (ie at the "A" level) will include all of the above plus

- Fit a linear regression where you predict Highway MPG using Vehicle Weight. Show the corresponding scatterplot.
- Fit a linear regression where you predict Highway MPG using 1/(Vehicle Weight). Show the corresponding scatterplot.
- Show residuals plots for each of the previous situations. What do you notice?
- Which of the two linear regressions do you prefer and why?
- Using your preferred model, predict the Highway MPG if a vehicle weighs 2500 lbs.
- Examine how approximate profit margin (ie difference between Suggested Retail Price and Dealer cost) relates to Vehicle type using a boxplot.
- Brief introduction and conclusion sections for your report.