1- Suppose that a group of 100 men aged 60-64 received a new flu vaccine in 1986 and that 5 of them died within the next year. Assume that a 1986 mortality table tells us that men aged 60-64 have an approximate probability of death in the next year of .020.

(a) How unusal is it for the 5 men to have died. [Hint: Be careful to review discussions on assessing extreme event. Also, use the binomial model.]

(b) Suppose 10 men had died. Should we be concerned?

2- A Number of studies have shown lichens (certain plants composed of an algae and a fungus) to be excellant bioindicators of air pollution. We are given the following data (read from a graph) on x=NO; wt deposition (g N/m^2) and y = lichen N (%dry weight) x = {.05,.10,.11,.12,.31,.37,.42,.58,.68,.68,.73,.85,.92} y = {.48,.55,.48,.50,.58,.52,1.02,.86,.86,1.00,.88,1.04,1.70}

(a) Does the model seem to be statistically significant based on the F statistic?

(b) What percentage of the variation of y is explained by the variation in x? [Hint - Do not use the adjusted R-squared.]

3- Chronological trend in blood lead gives the following data on y = average blood lead level of white children age 6 months to 5 years and x = amount of lead used in gasoline production (in 1000 tons) for ten 6-month periods.x = {48,59,79,80,95,95,97,102,102,107} y={9.3,11.0,12.8,14.1,13.6,13.8,14.6,14.6,16.0,18.2}.

(a)Construct normal probability plots for x and y. Might it be reasonable to assume that they come from a bivariate normal distribution?

(b)Examine a scatterplot of y versus x. Does it look like there may be some correlation structure?

(c)Does the data provide statistically significant information that there is a linear relationship between blood lead level and the amount of lead used in gasoline production? [Hint: Use an level of .05 and the R function cor.test]

4- Consider the famous Fisher Iris data.

1- Make a pairs plot of this data coloring the observations based on the species designation using a red, green, blue color scheme and a plot symbol of pch = 21.[Hint: You will need to adjust the figure margin using oma=c(3,3,3,15) when you call pairs and also call par(xpd=TRUE) prior to calling legend. Position your legend in the bottomright of the plot]

2- Repeat the exercise of part (a) except produce a parallel coordinates plot.[Hint - It may be impossible to place the legend so it does not obscure the plot a little.]

3- Use the summary function to explore this data.

4- Create a 2 x 2 matrix of box plots of the Fisher iris data with the sepal width and sepal length in the first row and the petal width and petal length in the second row.Make sure that the boxplots are conditioned by species.