Homework 5

1. The following five observations were selected from among the 20 households in the Sahlin's dataset on Mazulu village:

Table 1: Mazulu Village Data

Household	Consumers/ Gardener (X_i)	Acres/ Gardener (Y_i)
1	1.00	1.71
5	1.20	2.21
10	1.46	2.09
15	1.65	2.41
20	2.30	2.36

Do the following by hand (draw the scatterplot with a pencil on paper, calculate all statistics using formulas–no software allowed, and draw the least squares line with a pencil):

- Construct a scatterplot for Y and X.
- Find A and B for the least-squares regression of Y on X, and draw the least-squares line on the scatterplot. Interpret A and B.
- Calculate the standard error of the regression and correlation coefficient. Interpret these statistics.

2. Write a function that emulates the lm function in R for a simple (bivariate) regression. Like the lm function, your function should be able to estimate and report to the screen A and B coefficients, standard errors for these coefficients, and corresponding t-values and p-values. It should also report the residual standard error and R^2. Please email me the R code you write for this problem. That will facilitate my ability to test your code.

3. Examine Angell's data on the moral integration of U.S. cities (Angells is a data file in the car library). Use statistical software to regress moral integration on heterogeneity and geographic mobility for the cities in dataset (multiple regression). Report the finding of the results. Be sure to use a table to report A, B_1, and B_2 and statistics that allow for significance tests to be performed on these three coefficients. Also report the standard error of the regression and the correlation coefficient. Write a paragraph to substantively explain the results of the model.