GPH 483/598 Geographic Information Analysis **Spring 2010**
Exercise 1

Instructions: Using `R` you are to carry out a statistical analysis of the problem below. Your work must be turned in on the blackboard site and include: [1] Your script; [2] Your short interpretation and findings.
Due: February 10, 13:00 mst.

Research Question: You are an economic geographer interested in the impact of accessibility on residential rental markets. More specifically, you want to investigate the hypothesis that apartment rents per square foot are significantly higher for locations within 1,000 feet of a trolley stop, compared to apartments that are further than 1,000 feet away from the nearest trolley stop.

Data: The file `trolley.dat` has the sample data for your analysis. It has N rows, with the first row being a header. Each subsequent row is either an apartment or a trolley record. If the record is for an apartment, the value in the first column is the total monthly rent, R for the apartment, the second column is the square footage of the apartment s, while the third and forth column are the x and y locational coordinates for the apartment. If the record is for a trolley station, the values are -9 for the first two columns, while the third and forth columns are the locational coordinates for the trolley station.

All jobs in this city are in the central business district (CBD), and all commuters must go to one of the trolley stations to commute to the CBD each day which is located at $(x = 0, y = 0)$. You are interested in testing if the apartments within 1,000 feet of a trolley station command a significantly higher rent (per sq foot) than those beyond 1,000 feet of a station.

Suppose you also know that the apartment rent gradient (rent per square foot at different distances) is as follows:

$$r_i = r_0 - 0.019(d_i) + \beta T_i \tag{1}$$

where r_i is the rent (in cents) per square foot of apartment space at d_i feet from the CBD, r_0 is the rent at the CBD, β is the amount by which apartment rents per square foot increase if the apartment is within 1,000 feet of a trolley station, and $T_i = 1$ if i is within 1,000 feet of a trolley station otherwise $T_i = 0$.

You also know that the apartment size gradient is:

$$s_i = 500 + 0.20(d_i) \tag{2}$$

where s_i is the square footage of an apartment at d_i feet from the CBD. Finally, the total monthly rent for an apartment is given as:

$$R_i = s_i \times r_i \ . \tag{3}$$

Tasks:

1. Process your data:
   (a) What is the value of N?
   (b) Determine how many trolley stations are in the sample
   (c) Determine how many apartments are in the sample
   (d) What is the size of the bounding rectangle that contains all the apartments?

2. Visualize your data:
   (a) Plot the locations of the apartments
   (b) On the same plot, add the locations of the trolley stations and a 1,000 foot circular buffer around each station
   (c) On a separate graph, plot the size gradient, s
   (d) Plot the total monthly rent gradient, R
   (e) Plot the rent per square foot gradient, r
   (f) Provide an interpretation of each of your plots
3. Computationally, determine which apartments are within 1,000 feet of a trolley stop and which are not (assume distance can be measured as the crow flies).
4. What is the value of r_0?
5. Clearly formulate your null and alternative hypotheses regarding your research question
6. Based on your sample, provide an estimate of β
7. Select a test statistic
8. Calculate the statistic and its p-value
9. Interpret your findings with regard to the hypotheses