MKT-6323: Database Marketing
Group Assignment I

Instructor: N. Kumar, Ph.D.
http://www.utdallas.edu/~nkumar

There are two parts in this assignment. The data file (realestate.xls) is located at course page on my WebCT.

1. **Real Estate Business:** You have just started an e-venture, a website that provides consulting services and assists sellers in selling their houses. The seller of a house, who uses the services of a Real Estate Agent, would ordinarily have to pay a commission of 6% of the selling price (3% to his/her agent and 3% to the buyer's agent). You are attempting to provide this matching service to the sellers for a very nominal fee, say 1%.

   One of the services that you provide is a **recommended selling price**. The recommendation is based on data that you have acquired from a third-party vendor. The data has the following information:

   (i) Price at which the house was sold (in $1000's),
   (ii) Section of the town
   (iii) Lot Size (in 1000's of sq. ft)
   (iv) Number of Bedrooms
   (v) Number of Bathrooms
   (vi) Number of other rooms
   (vii) Number of stories
   (viii) The number of fireplaces
   (ix) Car Garages
   (x) Swimming Pool (1=Yes, 0=No)
   (xi) Whether or not the lot is fenced (1=Yes, 0=No)
   (xii) Age of the home in years

Part I

SAS part of the Assignment (This all in the SAS code – the only change you will need to make to the code is the location of the data set).

1. Run and estimate the regression between Selling Price (i) and the other variables ((ii)-(xii)).
2. Run factor analysis on variables (ii)-(xii) (rotate the factors)
3. Estimate the relationship between selling price and the factors (regression)

Part II (Due at the beginning of the 3rd class)

Analysis and Interpretation of the Output

4. Based on the results of the estimation in #1 (above), answer the following questions:
    a. What is the interpretation of the intercept, the coefficient of *lot size*, *beds* and other variables?
    b. What does the sign of the coefficient tell you?
    c. Based on the relationship you have estimated, how do you determine if a variable such as *lot size* has a significant (statistically significant) impact on sale price?
    d. Are the signs, magnitude and significance of the coefficients consistent with intuition – why/why not?
    e. Does the model fit the data well – what criteria do you use to assess goodness of fit?

5. Based on the results of the estimation in #2, answer the following questions:
    a. What do the correlations between variables reveal?
    b. Assign labels to the factors and provide a justification for these labels – justification should be based on results in #2.
    c. Based on the results how many factors would you consider in summarizing the information in the explanatory variables – why? What percentage of the total information in the explanatory variables is summarized in the factors that you have decided to retain?

6. Based on the results of the estimation in #3, answer the following questions:
    a. How do you interpret the intercept, and the coefficient of the factors?
    b. Are the signs and the magnitude of the coefficients of the factors consistent with intuition and the labels you have assigned in 6c – why/why not?
    c. Does this model fit the data better than the model in step 1 - why/why not?

7. Suppose a seller wanted to spend $10,000 on home improvement before selling the house. Assume that the lot size cannot be changed.
    a. If $10,000 were barely enough to make one update – add a bedroom, add a pool, etc., where would you recommend that the money be spent? Why?
    b. If $10,000 were enough to make two updates – add a bedroom and a pool, etc., which two updates would you recommend and why?

    Base you responses to this question on the results obtained from #1.