

BUS 735: Business Decision Making and Research

Instructor: Dr. James Murray

Homework: Regression Analysis

Due on Wednesday, October 5. Please upload to the appropriate D2L dropbox.

The owner of a moving company needs to estimate how many hours of labor will be required for each upcoming move. Being able to accurately predict this will allow the owner to schedule the right number of employees for each move. If he sends too many employees, he wastes his resources. If he sends too few, his customers are likely to get upset. He collects data from 100 past moves and records the following information:

- **Hours:** How many labor hours the move required
- **Area:** How large the residence was (in square feet)
- **Large:** How many exceptionally large items needed to be carried
- **Bedrooms:** How many bedrooms the residence had
- **Apartment:** Whether or not the residence was an apartment (=1 for apartment, =0 for house).

The data can be downloaded from one of the following:

<http://murraylax.org/datasets/moving.csv>

<http://murraylax.org/datasets/moving.RData>

1. Estimate the regression equation and write down the estimated equation.
2. Suppose the moving company's next customer has a 3 bedroom house (not an apartment) that is 1800 square feet, has two large items that need to be moved. What is your prediction for how many labor hours will be required?
3. What percentage of the variability in labor hours is explained by your explanatory variables?
4. Does it matter whether or not the residence is part of an apartment building or not when determining labor hours for moving? Test the appropriate hypothesis and clearly state your conclusion.
5. Test the hypothesis that at least one of your explanatory variables in your regression model helps explain labor hours for moving.
6. Think about this example. Is there any reason why any of the explanatory variables might be correlated? Which ones? For these variables, compute the Pearson Correlation Coefficient and test whether the correlation is different from zero.
7. Examine whether the relationship is in fact linear. Create scatter plots for each of the pairs of variables below. Comment on each, does a linear relationship look appropriate?
 - (a) X=area of house, Y=labor hours
 - (b) X=number of bedrooms, Y=labor hours
 - (c) X=number of large items, Y=labor hours
8. Save the residuals from your regression. Is there evidence the residuals are normally distributed? Show the appropriate evidence.