

# Business Analytics (108-1)

## Assignment 2

*Due: 9:00 am, Tue 22-Oct-2019*

### 1.

The manager of a commuter rail transportation system was recently asked by her governing board to determine which factors have a significant impact on the demand for rides in the large city served by the transportation network. The system manager has collected data on variables thought to be possibly related to the number of weekly riders on the city's rail system. The file railriders.txt contains these data.

- (a) (*Answer this question before using R.*) What are the expected signs of the coefficients of the explanatory variables in this multiple regression model? Provide reasoning for each of your stated expectations.
- (b) Formulate and estimate a multiple regression model using R with the given data. Perform a test of significance for each of the model's regression coefficients. Interpret each of the estimated regression coefficients. Are the signs of the estimated regression coefficients consistent with your expectation as stated in (a)?
- (c) What proportion of the total variation in the number of weekly riders is *not* explained by this multiple regression model?

### 2.

An antique collector believes that the price received for a particular item increases with its age and with the number of bidders. The file antique.txt contains data for 32 recently auctioned comparable items.

- (a) Formulate and estimate a multiple regression model with R. Interpret each of the estimated regression coefficients. Is the antique collector correct in believing that the price received for the item increases with its age and with the number of bidders?
- (b) Interpret the residual standard error and the coefficient of determination.
- (c) Suppose the antique collector believes that the rate of increase of the auction price with the age of the item will be driven upward by a large number of bidders. How would you revise the multiple regression model?
- (d) Interpret each of the estimated coefficients in the revised model.
- (e) Which model fits the given data better? Give your reasons.

### 3.

When car dealers lease a car, how do they decide what to charge? One answer, if you've got a lot of unpopular cars to move, is to charge whatever it takes to get the cars off the lot. A different answer considers the so-called residual value of the car at the end of the lease. The residual value of a leased car is the value of this car in the used-car market, also known as the previously owned car market.

How should we estimate the residual value of a car? The residual value depends on how much the car was worth originally, such as the manufacturer's list price. Let's take this off the table by limiting our attention to a particular type of car. Let's also assume that we are looking at cars that have not been damaged in an accident.

What else matters? Certainly the age of the car affects its residual value. Customers expect to pay less for older cars, on average. Older cars have smaller residual value. The term of the lease, say 2 or 3 years, has to cover the cost of the ageing of the car. Another factor that affects the residual value is the type of use. An older car that is in the great condition might be worth more than a newer car that has been heavily driven. It seems as though the cost of a lease ought to take both duration and use into account.

The file `used_bmw.txt` lists 218 BMW's popular 3-series.

- (a) Why does a manufacturer need to estimate the residual value of a leased car in order to determine annual payments on the lease?
- (b) Check scatterplots of the variables. Do the relationships appear straight enough to permit using multiple linear regression with these variables?
- (c) Fit the appropriate multiple linear regression model.
- (d) Does this model meet the assumptions for multiple linear models?
- (e) Build confidence intervals for the partial effects of age and mileage.
- (f) Summarize the results of your model. Recommend terms for leases that cover the costs of ageing and mileage.
- (g) Do you have any caveats that should be pointed out with your recommended terms? For example, are there any evident lurking variables?