# Data Analytics for Business

## Homework 3 - Part 2

### **Discrete Choice Models**

Suppose a local department of transportation is considering investing to improve the rail system in the city. They want to figure out how reduced commuting time by rail will increase people's choice of rail as their primary commuting mode. They have conducted a survey of 453 local residents about their choice of primary commuting mode and the cost information associated with each possible alternative. They want you to help them with the analysis.

The data file is "Commute\_Mode.csv". It has the following variables. Complete the tasks outlined below.

#### **Data Description**

Column Name	Variable Description
id	The id of individual survey participant
mode	The possible alternatives of commuting mode
choice	Each individual's choice of primary commuting mode: choice=1 means chosen
cost	A measure of the overall traveling cost by each mode
time	The commuting time by each mode

### Tasks:

- 1. Properly import the data using the **mlogit.data()** function. Specify the right data format (i.e., long) and the related variables.
- 2. Estimate the discrete choice model using the **mlogit()** function. Show the R output for the estimation results.
- 3. Examine the estimated coefficients and make sense of them. Are the signs of the coefficients reasonable? How many intercept coefficients are included, and why the number?

- 4. Compute the predicted choice probability for each alternative. Use the **fitted()** function and show the first few individuals' predicted choice probability for all alternatives. Interpret the results: for example, what does the value at the <u>first row</u> and the <u>last column</u> mean?
- 5. Calculate the marginal effects of the covariate "time" of the rail alternative on the choice probability of all alternatives. Evaluate the marginal effects at the mean values of the explanatory variables in the data sample. Interpret the results: for example, what do the values at (i) the <u>last row</u> and <u>first column</u> and (ii) the <u>last row</u> and the <u>last column</u> mean?