**MGS 8040 Data Mining**

**Assignment 2: Business Application and Data**

1.      Write a few sentences describing how data-mining might be used in a particular industry.

Answer the following pertaining to the industry application you discussed above.

2.      Identify one or more dependent variables (y), at least one of which is categorical.

3.      Identify some independent variables that might be used to predict the categorical y.

4.      Create a spreadsheet with 10 sample records (you can make up the data) that show reasonable values for any five Xs and the Y. Copy and paste as a table in the Word document that you will turn in. Clearly describe what each variable is.

5.      What would be a reasonable outcome period for this model to predict y?

6.      What would be a reasonable sample time frame for the prediction/classification problem? Give a date range, and explain in a sentence or two what must happen during that time frame for a record to qualify to be in your sample.