1. Names:
   * Art Grichine – ArtGrichine@csu.fullerton.edu – 50%
   * Adam Beck – AdamJBeck@csu.fullerton.edu – 50%
2. A brief description about the software tool you used or program you implemented.
   * Our software consists of a Python script which was created by Chris Strelioff and is found at: <http://chrisstrelioff.ws/sandbox/2015/06/08/decision_trees_in_python_with_scikit_learn_and_pandas.html>

This link leads to a tutorial of how to create decision trees in python with the ‘*scikit-learn’* and ‘*pandas’* libraries. The tutorial uses a data set for flowers. We modified the script to fit our data set and project goal. Our script follows a different pre-processing for the data as well as different settings to the decision tree (Entropy vs. Gini).

The heart of the program uses the ‘pandas’ library to extract the information from the csv dataset and place it into a pandas DataFrame data structure. The class column is defined as the ‘BusinessCategory’ column. Then, the DecisionTreeClassifier module from the SciKit-Learn library is used to train the model based on the defined parameters. Once the model has been trained, it is exported to the graphviz module (which is contained in the SciKit-Learn library) for graphical output. Also, a get\_code() function was written to generate the rules for the given tree. This code is output to the screen as well as a file called ‘code.txt’.

1. The results of your data analysis including the following elements:
   * The name of the class/decision column for your analysis.
   * A brief explanation of the process of your analysis listing a series of steps taken, specifying input and output if there is any, to produce the final decision tree from this data set.
   * A snapshot of the final decision learned. If the tree is too big, show only the important part of the tree based on your judgment (to meet the managers’ expectations).
   * Give two rules you can obtain from the decision tree that can be used to explain the nature of the data. If a rule is too big, you can show only the important part again based on your judgment.
   * A brief explanation of knowledge learned from the data analysis in layman terms so that two managers can understand.