The level of customer loyalty (Based on OLS Model)

ln this program. we used Ordinary Least Squares regression to do the analysis and prediction. Data is from Professor Jagdip Singh at Weatherhead School of Management from Case Western Reserve University. Programming by Jiaqi Li, Zeyao Yang and Ziming Cui.

**1. Usage Instructions**

If you wish to get a prediction result, change the dictionary paths for the data files containing the training and test data. Next, modify the output dictionary to point to the desired place. The software will output the predicted outcome.

**2. Background**

The level of customer loyalty will greatly affect the profitability of Bancare, so this program intends to collect customer data on LOYALTY and build a predictive model using a range of predictors as described in the “data-description” file.

Factors:

Loyalty consumer loyalty

Satis Satisfaction

Value Value from use

Reputa Reputation

Pricey Price-Value ratio

ATRUST Agent Trust (Log)

CTRUST Company Trust (Quadratic)

**3. Methodology**

To develop a predictive model that:

1. Has high R-square 2. Few predictors 3. Is robust and not biased by violations 4. Includes important interaction and quadratic terms.

First, with the given data, we exclude location and Case ID. Then changed three categorical factors (Age, Sex and Education) to dummy variables. And then, according to the Cook’s Distance Discipline, we deleted the outliers to improve the model. After removing the outliers, we added important interaction (Value of use × reputation) and quadratic terms, improving R Squared from 0.6727 to 0.7779.

**4. Other Concerns**

In our model, it can be seen that the variability (variances) of the residual points decreases with the value of the fitted outcome variable, suggesting non-constant variances in the residual errors (or heteroscedasticity). Then, it’s good, if residuals are spread equally along the ranges of predictors.