**Predict response in ACME email campaigns.**

files:

*ACME.str*

*mailing.csv*

Project covers building a CHAID model to predict response to the test mailing and applying it to other group(non test) of customers. The second part covers application of a segmentation model to find clusters of customers in the ACME database based on 3 predictors.

## Supervised learning: Classification.

Customer response for emails prediction.

* Predicted response for customers in high frequency, high recency and high monetary value categories.

(Model Nugget ”Responce\_to\_test”/View/node14)

F 37.452% Customer who don’t respond

T 62.548% Customer who respond

Based on out test email responses, we expect to receive the response from customers in about 62.548 cases out of 100 for high frequency high recency and high monetary value emails

### Assess the model's accuracy.

(Model Analysis node)

The algorithm predicted correctly 97.04% of all responses in testing email dataset.

162 out of 361(162+199) customers who responded positively were identified as such.

### Model application for the rest of the customers (not in the test mailing).

We assume that the model is satisfactory.

There are 20000 who are not included in the test mailing. Of them, 254 customers predicted to respond. (Predicted responses distribution node)

Data about those 254 customers were exported to mailing.csv(attached)(Export ToContact node)

## Unsupervised learning: Segmentation

Cluster customers based on RECENCY, FREQUENCY and MONETARY\_VALUE using TwoStep segmentation model.

Summary of found clusters and formal acceptability of solution.

The TwoStep algorithm found 3 clusters of customers (sizes 33.7%,33%,33.3%). The solution is acceptable ("fair").

### Profile the clusters in terms of the input fields.

Cluster 1 is comprised of customers which received emails with medium Frequency, medium Monetary Value, medium Recency .

Cluster 1 is comprised of customers which received emails with high Frequency, low Monetary Value, low Recency.

Cluster 3 is comprised of customers which received emails with low Frequency, hugh Monetary Value, high Recency.

We can try to predict customer response based on this Segmentation.