# Customer Brand Preferences Report

The complete survey shows the numbers 1 and 2 to assign brand. The key showed that 0 is Acer and 1 is Sony. I interpreted that the number 1 was Acer, and the number 2 was Sony.

I split the complete survey in a train set and a test set. The train set had 75% of the data and the test set had 25% of the data.

A K-fold cross validation was used, with number of folds (k) set at 10.

The models tried were Stochastic Gradient Boosting (gbm), Random Forest (rf), and C5.0.

The gbm model had the best metrics in terms of accuracy and Kappa.

It had a 92.4% accuracy and 0.84 Kappa.

Accuracy Kappa

0.9244139 0.8400827

The rf model had the following metrics:

Accuracy Kappa

0.9195635 0.8295368

The C5.0 model had the following metrics:

Accuracy Kappa

0.9126920 0.8136931

The gbm model had the following variable importance:

Chart

Description automatically generated

* Salary, and age were the most important variables in this model.

The predicted brand values were the following:

Chart, bar chart

Description automatically generated

Acer Sony

Observations 1947 3053

Percentage 39% 61%

Chart

Description automatically generated

Acer Sony

Observations 5691 9207

Percentage 38% 62%

It seems that the customers prefer the brand Sony, and the brand that they choose is highly dependent of their salary and age.