

BLACKWELL

ELECTRONICS

*Customer Brand
Preferences
Report*

 **ubiquum**
By
Michail
Alexakis

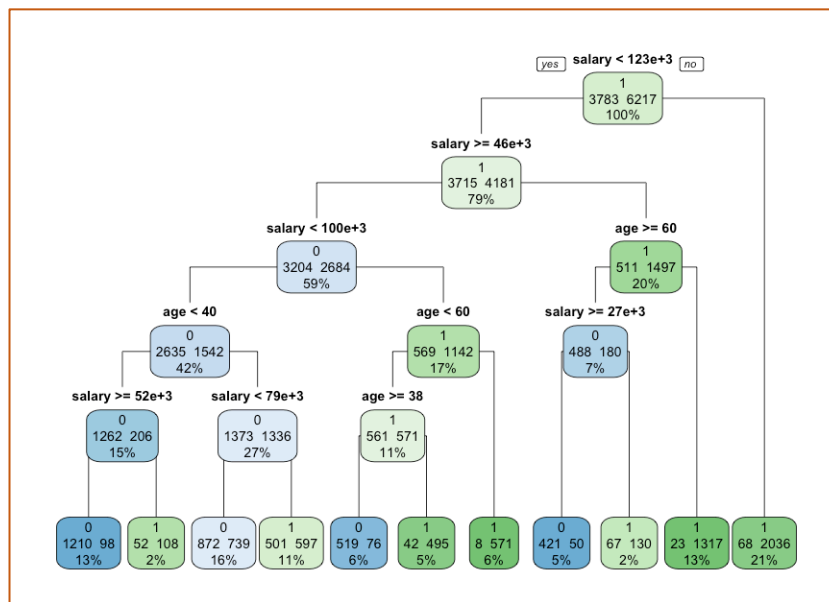
Introduction

This report is about to show **the preferences of Blackwell's customers in two different electronic Brands** so that the sales team to decide with which manufacturer they should pursue a deeper strategic relationship. These are Sony and Acer based to the existing data of 10.000 customers. The prediction was places in an incomplete survey response of 5.000 customers so that to show what they would prefer too.

Data visualization

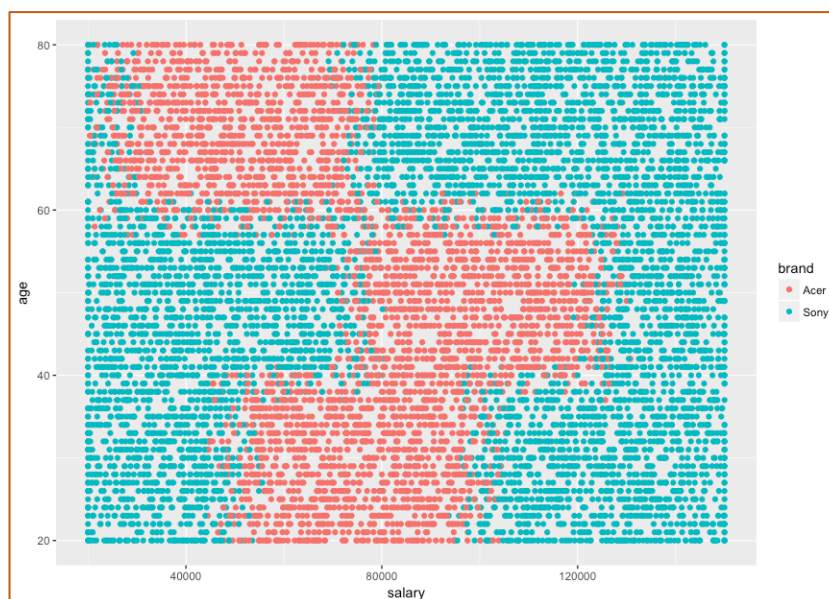
For the first phase of the investigation, and using the Complete survey response supported from the survey key sheet, the data was observed so that to convert them into the preference types.

Yet, they were visualized so that to show relationship between the variables. **Using a 5-levels decision**



tree, was obvious that the brand is related with the age and the salary of the customers. Thus, the strongest attributes for the prediction are these two.

Chart 1: Classifiers testing and prediction results.



It was, also, useful to use bivariate plot which shows the same attributes to be related with the brand preferences.

Chart 2: Bivariate plot between the Brand and the independent variables age and salary.

Building Models with KNN and Random Forest classifiers

For the next phase, a **10-fold Cross-Validation** was created for the process for training and testing the data. It was necessary, a **pre-process** to take place so that the data to be centred and scaled.

Using the `train()` function, the **Knn and Random Forest model** was created. Applying the models to the test package of data, it was easy to assess on which classifiers performs better. Also a `confusionMatrix()` function was used to show clearly the results of each classifiers.

Knn Matrix	Acer	Sony
Acer	852	82
Sony	93	1472

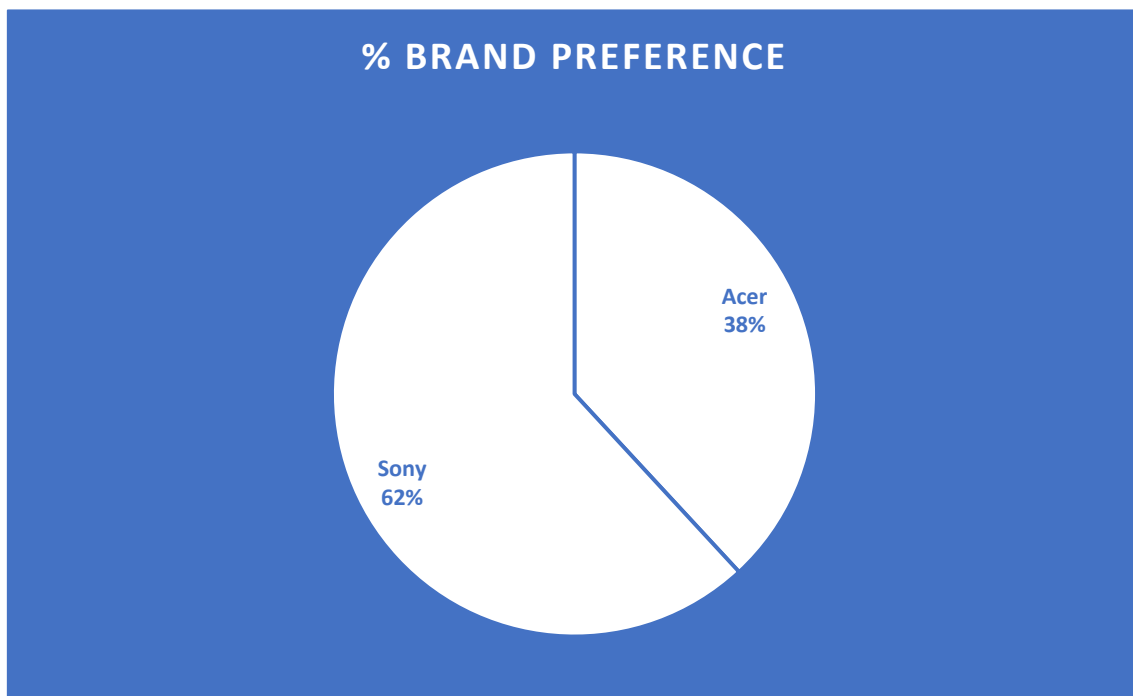
RF Matrix	Acer	Sony
Acer	844	95
Sony	101	1459

Classifier	Accuracy	Kappa	Sony Preference	Acer Preference
KNN	92,90%	0,851	1565	934
Random Forest	92,16%	0,833	1560	939

Table 3: Classifiers testing and prediction sum results

Final Prediction

The final prediction happened **applying the Knn model into the incomplete survey** response of the 5.000 customers. Applying and visualizing the results it is shown that **Sony is leading the customers' preferences**. The **pie chart** below shows the percentage of the brand's preference.



Pie Chart 4: Percentage of the customers' preferences

Conclusions

The investigation was supported from a complete Survey of 10.000 costumers and an incomplete survey that the brand preference is missing. It is remarkable that, the 62% of the customers in the complete survey prefer Sony while the rest 38% Acer. The results of the prediction give the same percentage for the 5.000 customers of the incomplete survey. This is happening because the customers characteristics in both Complete Survey and Incomplete Survey are similar as it is shown in the histograms below.

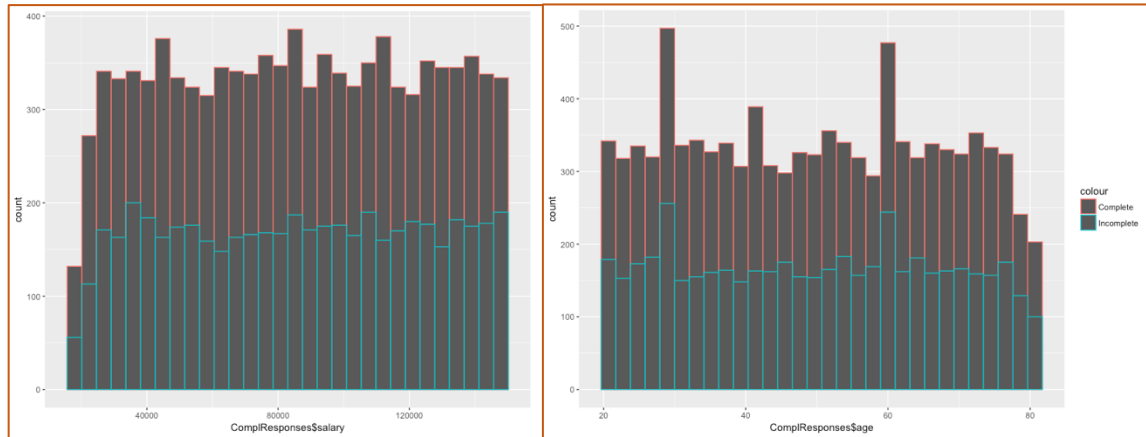


Chart 5: Histogram that shows the similarity between the surveys.

The incomplete survey includes data from 5.000 customers. So, the frequency is lower but still very similar to the complete survey data.