Bank Data Report

To Alpha OMEGA Bank

As discussed, we have reviewed the data sample you have provided and we in turn have applied some statistical transformation and the Apriori algorithm to help identify scenarios in which your clientele is significantly more likely purchase the Personal Equity Plan (pep).

# Preprocessing

The first steps to executing the task involved the data preprocessing. The data included two continuous variables and one count variable that R reads as being continuous. Continuous data is that which the values are generally considered infinite in scope. Age, which was one of these in your data set, can be measured in whole number years, but can be broken down into a nearly endlessly factional components to measure the value. A client maybe 31 years old, or more accurately 31 years, 12 days, 14 hours, 47 minutes, 15 seconds, 24 milliseconds, and so on.

We changed age, income, and number of children to be discrete variables. For Age we bucketed your customers in the sample data set into three near equal sets. Those of 34 years or younger. Those between 35 and 49, and those aged 49 or older. Similar we broke income into three buckets as well. Those making 24 thousand or less, those making between 25 and 43 thousand, and those more than 43 thousand.

Now that the data has been transformed to discrete variables, the Apriori algorithm can review the discrete items sets and establish patterns of items that are in common. In our case because we wanted to know what factors are influencing the purchase of the pep product, we requested the algorithm specifically give us the outputs that show the predicted peps. We have looked at both Affirmative and Negative responses so that we can not only figure out where the pep product does well and people are likely to take it, but also looked at where you can save resources among those who are less likely to take the product.

# Resource savings rules

We will start with where to save resource. The Apriori has identified that customers with three children, but who do not have a savings account are not likely to subscribe to the pep product.

Not only does the algorithm give out a rule, but it provides additional information by which we help determine the value of this rule. The additional information is three factors.

1. Support - The number of times across the data set that the combination of factors is found. For the rule mentioned the support is 3.6%
2. Confidence – the measure that tells us how often amount the combination of items that the rule holds true. In this case the percentage of the time that this rule holds true is 100% which means in every case of the customer having 3 children and no savings account, they also have not taken the pep product when offered.
3. Lift – helps indicate if the items are correlated together. This is important because negatively correlated products tells that items being together is more by chance than actually impacting each other. A value less than 1 is considered negatively correlated and items greater 1 are positively correlated meaning the rule is more indictive of future outcomes. A good score usually ranges between 1.3 to 1.5, but other values are possible. In the case of aforementioned rule we have a value of 1.8 meaning there is strong relationship between these items.

It is worth noting that of the top 50 rules based on life value, many of them test this output of 3 children and no savings account with other factors. Therefore, we reduced review to just a few rules to help you focus on where the biggest savings and returns can be found.

The other negative rule we thought we would bring to your attention is when a customer is in the top age group, married and has no children they are also not likely to purchase a pep. This combination has a support of 7.8% and confidence of 80%. It also show the grouping are highly dependent on each other with a lift if 1.46. This one runs counter intuitive because as people get older, have more and want to protect their assets and married couple does have to plan for the eventualities if someone goes first will the other spouse have the resources needed to pay the bills.

# Revenue Generating rules

Now for the fun part, the rules that are likely to help you achieve greater sales of the pep product. Starting with the simplest rule, customers who have exactly 1 child. This shows up in the data 18% of the time (support) and shows a result for pep product 81%(confidence) of the time. Additionally, the lift on this rule 1.78 indicating that the two are very likely dependent on each. This makes general sense because as people who have only 1 child tend to be more doting and concerned if something were to happen to them how child would survive.

The next rule that is likely to be quite helpful are non-married customers earning between 25 and 43 thousand dollars. It shows up in the data 9.3% of the time (support) and has a confidence 77%. The also shows a high lift of 1.7 meaning the dependence of pep on this is quite high. This makes sense as people who have achieved financial security usually want to protect that which the pep product serves quite well.

Our final recommendation is to focus on suburban customers, who are in the 24 to 43 thousand income brackets. The data shows this in the data about 4% of the time with a 75% confidence and lift of 1.6. This makes sense as people who live in the suburbs typically do so because it present possible cost savings. However, with further drives to work and additional costs associated with it, this group will be very interested in protecting the nest egg that suburbs allow them to create.

As always, we hope you find this report useful. Please feel free to reach out should have any questions.

Regards,

KW and Associates.