IST 707: Homework 3

**Introduction**

The objective of most for-profit businesses is to maximize profits in the most efficient manner. In pursuit of this objective, businesses often benchmark other successful firms within and cross industry. What makes an online retailer like Amazon, a tech company like Apple, or a restaurant chain like McDonald’s successful? At the heart of this success (or the success of any company) is the consumer. Without someone to purchase/consume these companies’ products, nothing that these businesses do operationally will allow them to be profitable.

In order to leverage any advantage that a business may create via strategic or operational separation from others in industry, it must first attract the consumer. Businesses use a variety of methods to gain and maintain the attention of customers that span from print and web advertisements to sponsorships of major sports teams. Successful businesses are able to identify the right consumer for their product(s) and the right medium/scale to reach that consumer.

The Bank of Syracuse would like to identify the prime consumers for their upcoming Personal Equity Plan (PEP) to be offered next quarter. Historically, the bank has mailed mass print ads to all existing and past customers, however, it would like to improve its response rate and reduce costs of targeting uninterested customers. To help identify the most appropriate target audience, the Bank of Syracuse has sponsored a study of its existing customer demographic information. The bank hopes to use the results of this study to launch a precisely targeted marketing campaign for its new PEP.

**Analysis and Models**

**About the Data**

The data provided for the purpose of this report was supplied the Bank of Syracuse Marketing Department. The data consists of 600 current or former bank customers who have had their personal information anonymized with a surrogate integer id. This id was excluded from this study due to its inability to be used as a descriptive demographic characteristic.

The remaining dataset contained eleven variables which were as follows:

|  |  |
| --- | --- |
| Variable Name | Variable Description |
| age | Age of customer in years |
| sex | Sex of customer (MALE/FEMALE) |
| region | Region a customer is from (inner\_city/rural/suburban/town) |
| income | Income of a customer in dollars |
| married | Is the customer married? (YES/NO) |
| children | Number of children a customer has |
| car | Does the customer own a car? (YES/NO) |
| save\_acct | Does the customer have a savings account? (YES/NO) |
| current\_acct | Does the customer have a current account? (YES/NO) |
| mortgage | Does the customer have a mortgage? (YES/NO) |
| pep | Did the customer buy a PEP after the last mailing? (YES/NO) |

All variables except for age, income, and children were already discretized. To support future association rule analysis, age, income, and children were binned as follows:

|  |  |
| --- | --- |
| Age | Discretized Age |
| 17-25 | Young Adult |
| 26-39 | Adult |
| 40-59 | Middle Aged |
| 60+ | Elderly |

|  |  |
| --- | --- |
| Income | Discretized Income |
| $0-$19,999 | Very Poor |
| $20,000-$39,999 | Poor |
| $40,000-$59,999 | Average |
| $60,000-$99,999 | Wealthy |
| $100,000+ | Very Wealthy |

|  |  |
| --- | --- |
| Children | Discretized Children |
| 0 | None |
| 1 | One |
| 2 | Two |
| 3 | Three |

Initial observation of the data showed that all customers had complete information for every attribute, so no customers were excluded from the study.

**Apriori Association Rules Models**

**Non-PEP Based Apriori Association Rules Model**

Initially, an Apriori association rules model was applied to the customer bank data using all of the available attributes. The parameters of this model were to identify rules with at least .25 support, .70 confidence, and 2 attributes. This model identifies association rules whose (2 or more) attributes occur at least 25% of the time amongst all occurrences of those attributes within the data. Additionally, it will identify rules in which the (2 or more) attributes have a probability of occurring together of at least 70% of the time amongst all occurrences of the associated attributes.

**PEP Based Apriori Association Rules Model**

To provide association rule insights more directly related to the upcoming PEP, a model using Apriori was developed that restricted to those involving PEP (YES or NO). The parameters of this model were to identify rules that contained PEP on the right-hand side of the model output and had a maximum number of attributes of 4 (including the PEP association attribute). The lack of confidence and support parameters allows this model to identify all rules involving PEP instead of inadvertently removing them from the model output. The results of this model should provide association rules between the demographic information of customers who either did or did not obtain a PEP after the last marketing campaign.

**Results**

**Non-PEP Based Apriori Association Rules Results**

The Non-PEP Based model yielded 45 association rules. These rules varied in support from .252 to .532 and in confidence from .701 to .818. Although these rules had relatively high support and confidence values, the results are somewhat ambiguous when considering the impact of the previous PEP marketing campaign. The majority of these association rules indicated that customers with varying demographic attributes had current accounts or saving accounts with the bank. This result is intuitive since this customer data is comprised primarily of existing bank customers.

Some results that were interesting when considering the past PEP were:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Rule | lhs | rhs | support | confidence | lift | count |
| 1 | mortgage=NO, pep=NO | married=YES | .285 | .818 | 1.240 | 171 |
| 2 | save\_acct=NO, pep=NO | married=YES | .292 | .745 | 1.128 | 175 |
| 3 | pep=NO | married=YES | .403 | .742 | 1.125 | 242 |

This set of rules seems to indicate that there is a linkage between married customers and not choosing the last PEP. The lifts for these rules show that customers are 1.125 more likely to be married if they did not invest in the PEP and increasingly more likely when they also did not have a savings account or a mortgage with the bank. This may suggest that the bank should not target married customers with their upcoming marketing campaign.

**PEP Based Apriori Association Rules Results**

The PEP Based model resulted in a set of 11 association rules. The statistical summary of these rules is as follows:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Stat | Support | Confidence | Lift | Count |
| Min | .105 | .800 | 1.47 | 63.0 |
| 1st Quartile | .121 | .831 | 1.65 | 72.5 |
| Median | .133 | .842 | 1.82 | 80.0 |
| Mean | .139 | .855 | 1.76 | 83.1 |
| 3rd Quartile | .157 | .880 | 1.85 | 94.0 |
| Max | .183 | .995 | 2.09 | 110.0 |

This summary shows that these rules had a higher confidence than the Non-PEP based model, but a lower support. This means that there is a higher probability of the attributes within these rules occurring together, but there are less instances where each attribute occurs.

Among these rules, the 5 most interesting regarding PEP were:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Rule | lhs | rhs | support | confidence | lift | count |
| 1 | income=Poor, children=One | pep=YES | .105 | .955 | 2.09 | 63 |
| 2 | children=One, mortgage=NO | pep=YES | .118 | .845 | 1.85 | 71 |
| 3 | married=YES, children=One | pep=YES | .123 | .831 | 1.82 | 74 |
| 4 | married=YES, children=None, save\_act=YES | pep=NO | .178 | .899 | 1.65 | 107 |
| 5 | married=YES, children=None, mortgage=NO | pep=NO | .173 | .897 | 1.65 | 104 |

Rule 1 proposes there is a link between customers with low income and a child and choosing the most recent PEP. This result is the strongest of all those from the PEP-based model with a 95.5% chance of a poor, one child customer choosing the PEP. According to this rule, a customer is 2 times more likely to invest in a PEP which would make them an ideal target for the upcoming marketing campaign. Perhaps this linkage is due to the customer hoping for a higher yield return from the PEP to support the child on a lower income than a traditional savings account would provide.

Rule 2 further supports the idea that a customer with a child and potentially lower income would favor a PEP. If not having a mortgage is considered as an indicator of income, this would again imply similar associations to Rule 1.

Rule 3 potentially builds on Rules 1 and 2 by identifying the association between a married customer who has a child and electing to invest in the PEP. This rule has a reduced confidence, but higher support. There is no income indicator within this rule, so it cannot be directly tied to Rules 1 and 2.

Rule 4 indicates that customers without children who have savings accounts are over 1.5 times more likely to not choose a PEP. This might suggest that these customers have less expenses without children and are more comfortable with a traditional savings account. The existence of a savings account in this relation may mean that these customers have a higher income.

Rule 5 again supports the argument that customers without children are less likely to choose a PEP, however, interestingly this rule includes no mortgage. If this were used as an indicator of lesser wealth, it would still be logical that a customer would be less pressured to invest in a PEP without the obligation of supporting children.

The interrelation of these rules is depicted in Figure 1 below:

**Figure 1**

**A screenshot of a cell phone

Description automatically generated**

**Conclusions**

In summation, The Bank of Syracuse is seeking to identify the optimal target audience for its next marketing campaign to promote its upcoming PEP. To help them find the ideal customers, the bank subsidized a study of 600 past and current customers’ demographic information. Association rule mining was used to find patterns within the data.

After analyzing the data, it appears that there is a link between customers with lower income and a family including at least one child are more likely to invest in the bank’s PEP. It also seems that customers who are married (particularly with no children) are less likely to choose a PEP.

In order to maximize the impact of the upcoming campaign The Bank of Syracuse should target customers with lower incomes and at least one child. Additionally, it will reduced excessive marketing costs by not targeting married customers, especially those without children.