5.3 Segmenting retail shoppers based on behavioural pattern of Store Images.

The main purpose of this research (in this section) is to empirically probe and prove the shopper typologies that have been theoretically formulated. The measurement was initially created and later it was statistically validated (see scale development process). As the instrument is validated, the data on validated measurement is used for cluster analysis.

To find out the pattern of behaviour of the shoppers based on the store images, initially, k-means cluster analysis was performed. To further investigate, the inherent consistency of behavioural aspect of various shoppers fuzzy c means cluster was performed. And the fuzzy cluster was defuzzified based on maximum value of the members belong to the cluster.

To find out the consistency of the total market, the hard cluster solutions and the fuzzy cluster solutions are compared and the consistency was measured.

Bases of Segmentation

The segmentation descriptor is developed based on the behavioural typology of the shoppers based on store images. The initial themes were created namely pre-determined, economy, variety seeking behaviour and familiar to the store images.

Taking cues from the theme, the scale were created with the help of previous studies and scale calibration and purification was made. Finally the scale was proved with various reliability and validity tests. The constructs are the descriptors of the segmentation.

Moreover, the instrument was designed to plug the information from the shoppers related to their demographic, socio-economic and other attitude detail.

Tool used to segment the market

The cluster optimization indexes exhibits the optimum number of cluster should be four, using the same number as the input for initial iteration, k-means cluster analysis is performed.

Results of k-means cluster analysis

The initial cluster was formulated with the iteration value of four. The initial cluster centre detail is given in the Table 5.9. This table shows the values for the initial cluster centres. The values in the table are the means for each variable within each initial cluster. For instance, the mean of value of ambience for the first initial cluster is 3.00. The Table 5.10 indicates the the progress of the estimation process at each iteration. At each iteration, as cases are reassigned to different cluster, cluster centres change. Each number indicates how far the new cluster centre is from the cluster centre at the previous iteration. For instance, the iteration from 2nd iteration to 3rd iteration the values of all the cluster centres is reducing, this indicates the cluster is converging and the iteration stops till value gets close the value of zero. When the change is small enough for all clusters, iteration stops and the final solution is reached.

Table 5.11 shows the values for the final cluster centres. The values in the table are the means for each variable within each final cluster. The final clusters centres reflect the attributes of the prototypical case for each cluster.

The Table 5.12 shows the Euclidean distances between the final cluster centres. Large values indicate clusters that are very different from each other. Small values indicate clusters that are not so different from each other. Notice that this table is symmetric, i.e. the distance from cluster 1 to cluster 3 is the same as the distance from cluster 3 to cluster 1.

The Table 5.13 can help to understand which variables are most important in your cluster solution. For each variable, you can see the variance in that variable attributable to clusters and the error variance (the variance not attributable to clusters). The F ratio is the ratio of cluster variance to error variance. Large F ratios indicate variables that are useful for separating clusters. Small F ratios (near 1.0) indicate variables that are not very useful for separating clusters.

The Table 5.14 shows how many cases were assigned to each cluster. The cluster is named as variety seeking segment which consists of 68 members, cluster 2 named as economic segment which consists of 61 members, cluster 3 named as familiar segment which consists of 79 and cluster 4 named as pre-determined segment which consists of 150 members. Totally 358 is the sample, all the members are grouped in various groups.

Profiling the segment

The various clusters are profiled based on the shoppers' typology of the particular segment and based on the demographic, socio-economic and other behavioural related factors.

Pre-Determined

Based on the centroid value of k-means final cluster centers, the variables related to Prior-information, Availability and the Specifics of a product was a major factor in the behavior patterns of these customers. These customers were not much concerned about the ambience of neither the store nor the variety in the store. Their interaction with the sales personnel was limited. They were the largest segment of shoppers and they were found in all types of stores. This segment constitutes the largest membership value among the entire segments. It carries 150 members, among 54 percent are male and 46 percent are females. The interesting detail hers is, around 66 percentage of widow category falls in this group and majority of them are married (45.30 percentages). The highest graduated persons are in this cluster, that is to say 38.90 percentages of people are graduates and 47.60 are completed post graduates or professionals. Most of the predetermined segment members working in the level of managers/officers

(38.50%), and around 48.70 percentage of the segment are seen in food and grocery retail stores and 42.70 percent of people seen in durable store. They are very less seen in other types of stores. (The descriptive statistics about the clusters are shown in from Table 5.15 to Table 5.20). However, the major portion of the shoppers in this category was that of males in the age group of 25-45.

Variety-Seekers

This cluster consisted of a group of customers whose behavioral pattern were related to factors such as Depth of merchandise, Display of Product, Promotions ,Intensive Brand Search .Such shoppers observed all the racks, they would see many items before selecting one .Sometimes shoppers looked around the shop simply glancing at the various shelves without attention to any brand or product in particular .Such shoppers were found in all types of stores

.However they were found more in shopping malls. These shoppers liked to maximize the returns expended on shopping. Stores would need to have good range of merchandise to attract such shopper .Good display and communication of any scheme is essential. Salesperson would have a limited role to play. The shoppers were almost equally divided among men and women and were mostly married. Among the variety seekers, 61.80 percentage people are male, and 38.20 percentage are female, means male are more of variety seekers than female, and, most of them are married (around 55 percent). Most of the people are graduates, that is, around 22 percentage of graduated are categories in this segment. An interesting finding here is most managers are variety seekers that are around 38.20 percentage of variety seekers are managers. The variety seekers mostly seen in book, music and gift shops (around 53 percentages).

Economy

The behavioral pattern of such customers involved factors like Price Tags Display, Discount Tags, Promotions, Involvement of Sales Personnel. These shoppers changed products/brands as they did not fit into their budget. They repeatedly asked for bargains, asked salesperson to give them a good discount. They would look around for schemes .Stores that deals in durable products witnessed a large number of such shoppers. For such shoppers price of merchandise needs to be displayed prominently. The sales clerk would need to be patient and soft, and should try to maximize the value derived. The shoppers were generally males (95 percent) in the age group of 35-60. They have almost equal strength of married and single. Most are (around 60 percent) graduates; clerks' cadre of occupation mostly constitutes this segment, Food and grocery stores (46 percent) of economic category of people, and the economic category people are also much found in durable stores and cosmetics store as well.

Familiar

The major factors that determined the behavior of the customers falling under this cluster were Ambience, Priority over other Customers, Friendly Attitude of salesperson, Brand consciousness .These shoppers were frequent visitors to the shop. They would go straight to the particular section of the store. They seemed to know the salesperson, and many of them shook hand with the shop owner while leaving. While they were seen in most of the stores, they were found more in gifts and accessories stores .Such shoppers need to be attended immediately .They could talk about the view on the store. They should be made aware of any new arrival before other shopper. Among this group most shoppers are male, most are married but considerable number of members are living single, most of them are graduates, most of them work as managers and professionals, They are very much seen in apparels, shoes and fashion stores.

Results of fuzzy c means clustering (Soft Segmentation)

The objective of cluster analysis is the classification of objects according to similarities among them, and organizing of data into groups. Clustering techniques are among the *unsupervised* methods, they do not use prior class identifiers. The main potential of clustering is to detect the underlying structure in data, not only for classification and pattern recognition, but for model reduction and optimization. Different classifications can be related to the algorithmic approach of the clustering techniques. *Partitioning, hierarchical, graph-theoretic methods* and methods based on *objective function* can be distinguished. In this work we have worked out a toolbox for the partitioning methods, especially for hard and fuzzy partition methods.

Dunn extended the well-known c-means algorithm to a fuzzy clustering ethod called fuzzy c-means and it was further developed by Backer, Rouben developed non-metric fuzzy cluster method, Bock, Jim Bezdek et al, Gunderson, Trauwaert, Goth and Geva, Gu and Dubuisson and Xie and Beni. However, Jim Bezdek's FCM remains as the most commonly used.

FCM Partitions a collection of n vectors x_i , i = 1, …, n into c fuzzy groups and finds a cluster center in each group such that a cost function of dissimilarity measure is minimized. The major difference between FCM and HCM is that FCM employs fuzzy partitioning such that a given data point can belong to several groups with the degree of belongingness specified by the membership grades between 0 and 1. However, by imposing normalization which stipulates that the summation of degree of belongingness of a data set always is equal to unity:

$$\sum_{i=1}^{c} u_{ij} = 1, \forall j = 1,.....n$$

The cost function (or objective function) for FCM is then a generalization of below equation

$$j(U, C_1 ..., C_C) = \sum_{i=1}^{C} J_i = \sum_{i=1}^{c} \sum u_{ij}^m d_{ij}^2$$

Where u_{ij} is between 0 and 1: c_i is the cluster center of fuzzy group i ; d_{ij} = $\| c_i \quad x_j \|$ is the Euclidean distance between i^{th} cluster and j^{th} data point; and m Є [1, ∞) is a weighting exponent.

The necessary conditions for equation - 2 to reach a minimum can be found by forming a new objective function $\bar{j}$ as follows:

$$\bar{j}(U, C_1,..., C_C, \lambda_1,..., \lambda_n) = J(U, C_1,..., C_C) + \sum_{j=1}^{n} \lambda_j \left(\sum_{i=1}^{c} U_{ij} - 1 \right)$$

$$= \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{m} d_{ij}^{2} + \sum_{j=1}^{n} \lambda_j \left(\sum_{i=1}^{c} u_{ij} - 1 \right)$$

Where λ_j, j = 1 to n, are the Lagrange multipliers for the n constraints in first equation. By differentiating $\bar{J}$ (U, C_1 , …., C_C, λ_1, …, λ_n) with respect to all its input arguments, the necessary conditions for second equation to reach its minimum are

$$c_i = \frac{\sum_{j=1}^{n} u_{ij}^{m} x_j}{\sum_{j=1}^{n} u_{ij}^{m}}$$

And

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{d_{ij}}{d_{kj}} \right)^{2/(m-1)}}$$

The fuzzy C-means algorithm is simply an iterated procedure through the preceding two necessary conditions. In a batch-mode operation, FCM determines the cluster centers c_i and the membership matrix U using the following steps:-

Step 1: Initialize the membership matrix U with random value between 0 and 1 such that the constraints in equation-1 are satisfied.

Step 2: Calculate c fuzzy cluster centers c_i i = 1… c, using equation-4.

Step 3: Compute the cost function according to second equation. Stop if it's either reaches below a certain tolerance value or its improvement over previous iteration is below a certain threshold.

Step 4: Compute a new U using equation-5 and go to step 2.

The cluster center can also be first initialized and then the iterative procedure carried out. However there is no guarantee to ensure that FCM converge to an optimum solution. Such performance depends on the initial cluster centers, hereby allowing us either to use another fast algorithm to determine the initial cluster centers or to run FCM several times, each starting with a different set of initial cluster centers.

As mentioned above, The Fuzzy C-means clustering algorithm uses the minimization of the fuzzy C-means function. There are three input parameters needed to run this function: *One*- the number of clusters or the partition matrix to be initialized; *Two*- the fuzziness weight exponent; and *Three* - the maximum termination tolerance. The last two parameters have their default values, so the user has to define the input about the number of cluster.

The fuzzy method can be divided into two types one uses fuzzy relations to perform fuzzy clustering and the other uses the objective function to determine the fuzzy clustering.

Profile of the Soft Segments

According to the FCM algorithm, In the Table 5.21 the centroid represents the mean value of the objects contained in the cluster on each of the variables. Cluster 1 has a value relatively higher of all its values on ambience, prior information, availability of merchandize, specific product, variety, low interaction with sales personnel, as these dimension reflects the pre-determined, so we label as it is. Cluster 2 has a value relatively high on price, high influence of price on brand switching, price display seek discounts, prefer discounts, indoor display and outdoor promotions related to discounts, seeks help from sales personnel, as these dimensions reflects the shopper typology of economic aspects, , we label as economy. Cluster 3 has a value relatively higher on the variables of intensive search, stress on product category, seek value for money, seek depth in merchandise, limited role for salesman, informed about promotions, and examine the product thoroughly, as these dimensions reflects the typology of the variety seeking, we label it as “variety seeker.” Cluster 4 has a values relatively higher on the variables of Influence of store image, Priority over other customers, Friendly attitude of sales personnel, Recommendation to others, Brand conscious, as these variables reflects the typology of familiar, we label as it is.

Comparison of Fuzzy and K means Cluster Analysis

By comparing the centre point of both k means cluster results and fuzzy c means cluster results, we can easily find the central values of both the k means and fuzzy c means look similar, which indicates that the fuzzy cluster was separated as same as k means cluster, that is the cluster centre values are similar, that significantly proves the strength of the separation variables.

The FCM algorithm is used to determine the membership grade of each member in different segments. Because of the large number of samples, we list only few examples in the Table 5.22. (See Table 5.23 for membership value of all samples). The membership value of the fuzzy c means cluster solution are found to be not hard cluster, they are fuzzy in nature which reflects the real market situation, by adding the values of all the membership grade values of all the segment it will be the sum of one. Which means that a member is belong to various segments. Whereas, the hard cluster solution will reflect that the member will belong to only one segment.

5.4 Measuring the stableness of market segment

Using the fuzzy c means and k means segment results derived from all samples (358), the market sizes obtained from hard method are different from those of fuzzy c means and k means segments. The rule to identify and compare the fuzziness of the segments is that "The smaller the market size difference, the more stable the market and vice-versa". The Table 5.22 shows the differences of cluster size between k means cluster membership value and fuzzy c means membership value. Moreover it also exhibits the percentage of fuzziness of the particular cluster in specific and total cluster fuzziness in general.

Based on the percentage of fuzziness, the familiar segment (0.13 %) and variety seeking segments (1.23 %) are considered to be more stable segment; pre-determined segment (1.36 %) and economic segment (1.82 %) have greater fuzziness. Hence, these segments are considered to be unstable segments. The overall fuzziness of all the segments is estimated as 1.13 percent.

5.5 Building Predictive model based on the hard cluster solutions

Discriminant analysis is used in situations where the clusters are known *a priori*. The aim of discriminant analysis is to classify an observation, or several observations, into these known groups. As we know the priori cluster solutions from k means, it helps us to build a predictive model for the retail market segmentation. The Discriminant analysis group statistics displays the descriptive details and its statistics for each variable across groups and for the total sample. The mean is the average value. The standard deviation measures the variability (or spread) of the values. Discriminant analysis assumes equal variances. Therefore, the standard deviations should not vary greatly across groups. It is found that from the descriptive details of group statistics, the standard deviation score are similar among the group (Variety seeker, Economy, Familiar, Predetermined). (See Table 5.24 to Table 5.27 for the descriptive details of group statistics). The overall group statistics is also give in Table 5.28 Valid N is the number of cases with non-missing values. The weighted N value s weighted by the weight variable.

Table 5.29 contains Wilks' lambda, the F statistic, its degrees of freedom and significance level. Wilks' lambda is the ratio of the within-groups sum of squares to the total sum of squares. Wilks' lambda ranges from 0 to 1.0. Small values indicate strong group differences. Values close to 1 indicate no group differences. The variables related to the aspects of predetermined cluster and familiar cluster are well more differentiated among them, where as the other clusters of economy and variety seeking are not much differentiated or otherwise not well separated. This indicates that economic group and variety seeking group looks similar and close pattern.

The F statistic is a ratio of between-groups variability to the within-groups variability. The F statistic has a numerator (df1) and denominator (df2) degrees of freedom. The numerator and denominator degrees of freedom are used to obtain the observed significance level. If the significance value is small (smaller than say, 0.10) this indicates that the group differences are significant. If the significance value is large (larger than say, 0.10) this indicates that the group differences are not significant. Here, all the variables are statistically significant which means the group differences are significant.

In the multigroup model, log determinant values provide (See Table 5.30) an indication of which groups' covariance matrices differ most. For each group, its log determinant is the product of the eigenvalues of its within group covariance matrix. The rank is the row or column rank which is the maximum number of linearly independent rows or columns.

In this case, the log determinant value for economy and variety seeking groups are very small relative to that of the others. Possibly, if economy and variety seeking groups were omitted from the analysis, the equal covariance assumption would be met. Box's M statistic tests (See Table 5.31) the null hypothesis of equal population covariance matrices. The significance of Box's M statistic is based on an F transformation. The hypothesis of equal covariance matrices is rejected if the significance level is small (less than says 0.10).The hypothesis of equal covariance matrices is not rejected if the significance level is large (more than say 0.10). The significance level is less than 0.05 indicates the null is rejected and proves there is difference among group. The test can be significant when within-group sample sizes are large or when the assumption of multivariate normality is violated.

Summary statistics of canonical function

The Eigen value Table (See Table 5.32), displays eigenvalues, the percentage of variance, the cumulative percentage, and canonical correlations for each canonical variable (or canonical discriminant function). The eigenvalue is the ratio of the between-groups sum of squares to the within-groups sum of squares. The first function is the largest eigenvalue corresponds to the eigenvector in the direction of the maximum spread of the groups means. The second function is the second largest eigenvalue corresponding to the eigenvector in the direction that has the next largest spread, and so on. The square root of each eigenvalue provides an indication of the length of the corresponding eigenvector. The third function is the small eigenvalues result in eigenvectors of essentially no length and account for very little of the total dispersion. But, in this case, for the analytical reasons the third function is also included in the model. The percentage of variance column allows you to evaluate which canonical variable accounts for most of the spread. The first function accounted for 47.3 percent, the second function accounted for 28.8 percent and the third function is accounted for 23.9 percent.

The Wilks's lambda's table (see Table 5.33), The test of functions column tests the hypothesis that the means of the functions listed are equal across groups. Wilks' lambda is the proportion of the total variance in the discriminant scores not explained by differences among the groups. Wilks' lambda ranges between 0 and 1. Values close to 0 indicate the group means are different. Values close to 1 indicate the group means are not different (equal to 1 indicates all means are the same). As the scores of Wilks' lambda are 0.003, 0.026, 0.0176 for first to third function respectively, indicates that the group means are different. A chi-square transformation of Wilks' lambda is used along with the degrees of freedom to determine significance. If the significance value is small (less than say 0.10) this indicates that group means differ. If the significance value is large (more than say 0.10) this indicates that group means do not differ. The p-value of all the test functions are less than 0.05 indicates that group means differ.

The table of standard canonical discriminant function of discriminant analysis (see Table 5.34), the coefficients of the canonical variable are used to compute a canonical variable score for each case. The canonical discriminate functions in this model are three. When there are more than two groups, the number of canonical variables is k-1 (where k is the number of groups) or p (the number of variables), whichever is smaller. Standardizing the coefficients allows you to examine the relative standing of the measurements.

The structure matrix (See Table 5.35) contains within-group correlations of each predictor variable with the canonical function. This matrix provides another way to study the usefulness of each variable in the discriminant function. For each variable, an asterisk marks its largest absolute correlation with one of the canonical functions, which indicates the largest absolute correlation between each variable and any discriminant function. With each function, these marked variables are then ordered by the size of the correlation. The strongest correlations for the variables of variety, low interaction with sales, personnel, prior information, specific product, availability of merchandize, ambience occurs with first function. The strongest correlations for the variables of informed about promotions, examine the product thoroughly, limited role for salesman, seek value for money, seek depth in merchandise, stress on product category, intensive search, recommendation to others occurs with function two. The strongest correlation for the variables of indoor display and outdoor promotions related to discounts, prefer discounts, friendly attitude of sales personnel, high influence of price on brand switching, seek discounts, price, seeks help from sales personnel, influence of store image, price display, brand conscious, priority over other customers occurs with function three.

The discriminant analysis un-standardized canonical discriminant function coefficient (See Table 5.36) describes the coefficients of the canonical variable. The coefficients are used to compute canonical variable scores for each case. For the forst canonical variable, the score can be computed as, score = -1.273 + .093 (brand conscious) +.... n_{25} variable. By substituting the value of all the variables, for a specific case, will help in computing the canonical variable score. When there are more than two groups, the number of canonical variables is k-1 (where k is the number of groups) or p (the number of variables), whichever is smaller.

The group function centroid (see Table 5.37) displays the canonical variable means by group. Within-group means are computed for each canonical variable. In the first canonical variable, the average discriminant or canonical variable score for variety seeking is 2.358, economy is 2.812, familiar is 2.573, pre-determined is -3.568. In the second canonical variable, the average discriminant or canonical variable score for variety seeking is 4.455, economy is -1.356, familiar is -2.613, pre-determined is - 0.0922. In the third canonical variable, the average discriminant or canonical variable score for variety seeking is -.720, economy is 4.132, familiar is –2.717, pre-determined is – 0.0769.

The Discriminant Analysis Classification Function Coefficients (See Table 5.38), each column contains the estimates of the coefficients for a classification function for one group. The functions are used to assign or classify cases into groups. The estimate of the classification function for market segment of variety seeking cluster is -131.936 + 1.563 (ambience) + 3.543 (priori information) +.........k_{25}. The estimate of the classification function for market segment of economy cluster is -119.774 + 1.442 (ambience) + 2.875 (priori information) +.........k_{25}. The estimate of the classification function for market segment of familiar cluster is -113.302 + 1.435 (ambience) + 3.130 (priori information) +.........k_{25}. The estimate of the classification function for market segment of pre-determined cluster is -123.025 + 3.129 (ambience) + 4.851 (priori information) +.........k_{25}. To obtain a classification score for each case for each group, multiply each coefficient by the value of the corresponding variable, sum the products, and add the constant to get the score. A case is predicted as being a member of the group in which the value of its classification function is largest.

The canonical classification function is represented in separate group graphs in scatter plot (See Figure 5.2 for variety seeking, See Figure 5.3 for economic, See Figure 5.4 for familiar, See Figure 5.5 for pre-determined, and See Figure

5.6 for all the groups). The discriminant analysis classification function (See Table 5.39) measures the degree of success of classification for predicting the future member of the cluster (sample). The number and percentage of cases correctly classified and misclassified are displayed. Here, all the groups are correctly classified at 100%.

Answer-tree based classification model for crisp segments

Chi-square automatic interaction detection (CHAID) is a heuristic treebased statistical method that examines the relations between many categorical or discrete (having a limited (say fewer than 15) number of numeric values) predictor variables and a categorical target or outcome measure. It provides a summary diagram depicting the predictor categories that make the greatest difference in the desired outcome. For segmentation purposes, CHAID can thus give you information about the combinations of demographics that give you the highest probability of a sale. CHAID is a heuristic statistical method that examines relations between many categorical or discrete predictor variables and a single categorical outcome variable, and provides a summary diagram relating the predictors to the outcome.

CHAID Model Summary

The model summary given in Table 5.40 which provides very broad information about the specifications used to build the predictive model. The specifications section provides information on the settings used to generate the tree model, including the variables used in the analysis. The results section displays information on the total number of terminal nodes, depth of the tree (i.e.) the number of levels below the root node, and independent variables included in the final model.

As much as eight independent variables are specified, but only three are included in the final model. The variables such as store type, gender, current working status has made significant contribution to the model. Rests of the variables are automatically dropped in the final model. The tree diagram is a graphic representation of the tree model. This tree diagram (Figure 5.7) shows that, using CHAID method, the variable store type is the best predictor of targeted segment of pre-determined with the chi-square transformation of 142.970 with the adjusted p-value of 0.00 which is less than 0.05, in the 3 (*df)* degree of freedom. As such as two child nodes are estimated from the store type. The node one targets the maximum membership of the peoples belong to durable stores, food and grocery store and medical stores which accounted for 66.2% (among the group it targeted 61.6 % from pre-determined segment which consists of 146 members). The node two targets the maximum members of the people belong to apparels stores, book & music stores, cosmetic and fancy stores which accounted for 33.8% (among the group it targeted 43.8 % from familiar segment which consists of 53 members). On the other side of the node one, the child node emerges, for this the next predictor is gender which having both the types of male (accounted for 39.9%- (among the group it targeted 55.2 % from pre-determined segment which consists of 79 members)) and female (26.3% (among the group it targeted 71.3 % from pre-determined segment which consists of 67 members)) with the chi-square transformation of 28.645 with the adjusted p-value of 0.00 in 3 degree of freedom. The female node gets terminated; the male node emerges with a child node that targets current working status category of male people with the chi-square transformation value of 18.497 with the adjusted p-value 0.001 in 3 degree of freedom. The male node emerges with two child node, one targets the people they do not work outside home and work out side home part time which accounted for 24.6 % (targeted 45.5% from pre-determined segment which consists of 40 members) and another node targets work outside home full time which accounted for 15.4 % (targeted 70.9 % from pre-determined segment which consists of 39 members).

The content of the Tree diagram are tabulated in the tree-table (See Table 5.41 a & b). The "N" indicates the number of cases in each category and "percent" indicates the percentage of cases in each category. The column, "predicted category" indicates the predicted category of dependent variables in the predetermined segment because it targets population. The column, Parent node indicates the parent node of each node.

The results of the classification table are consistent with the risk estimate. The model classifies approximately 54% of the members correctly.

The classification table (Table 5.42) reveals a potential problem in this model. For the predetermined and familiar segments, the predicted rating is 55.60 percent, which means that 44.40 percent of the members with the loyalists segment are inaccurately classified with other segments. Therefore, the results from the foregoing give a clear indication of the closeness of the rate of classification of the terminal node as they have accomplished as much as 74.5% which has no much significant variation from the accepted level of classification rate which is 80%.

5.6 Statistical test

In order to evaluate the effect of the various store images attribute segmentation on the overall evaluation of store type choice, chi-square test with Monte Carlo exact test was performed. Table 5.43 of chi-square test measures indicates that the hypothesis that the store image attribute segmentation and the store type choice in a cross tabulation are independent. A low significance value (typically below 0.05) indicates that there may be some relationship between the two variables. While the chi-square measures may indicate that there is a relationship between two variables, they do not indicate the strength or direction of the relationship. Table 5.44 indicates that the p-value is less than 0.05 for all the sub statistical measure which includes Pearson chi-square and Monte Carlo exact test. The nominal symmetric measures also performed, which indicates both the strength and significance of the relationship between the store images attribute segment and store type choice. The nominal symmetric measure (see Table 5.45) indicates the symmetric value of Cramer's V and contingency coefficient are having high value, that is near to 1, which indicates there is a strong relationship between the store images attribute segments with the store type choice.