Research In The Marketing Context/13

Segmenting the Market IX: Analytic Techniques-2

By Chuck Chakrapani

Those of you with long memories might recall this series Research in the Marketing Context*. I had to abruptly discontinue it last year due to other commitments. A few phone calls from* Imprints *readers assured me that the series was being read and there is still interest in topics of this nature. Therefore, with some encouragement from the* Imprint *editor, Susan Robbins, I have decided to continue the series. This is the last article on segmentation.*
Chuck Chakrapani

Segmentation Approaches

In the previous article, we discussed how markets can be segmented using intuitive (subgroup and nested subgroup analyses) and market structure (multidimensional scaling, discriminant analysis, correspondence analysis) approaches.

Although all of the above approaches are useful, the widely preferred methods of segmentation concentrate on partitioning the market.

Market Partition Analysis

In market structure analysis described in the previous article, our primary interest was in identifying the segments. In market structure analysis, we are also able to portray the segments visually. For instance, using multidimensional scaling we could have identified a segment that would be an ideal target for an exclusive and expensive perfume. However, the technique would not have identified the size or characteristics of this target group.

This task is best carried out by a group of techniques that specialize in market partition analysis. Partition techniques aim to divide the market into a number of smaller, identifiable groups.

Market partition analysis can be carried out using one of the two different approaches: dependence and interdependence.

Dependence Approach to Segmentation

The dependence approach uses a criterion chosen by the researcher as the basis for segmentation. For example, the researcher may want to partition the market on the basis of product usage. The segments derived, then, will directly reflect the usage of the product. The characteristics of the segments will be assumed to be the variables contributing to the usage of the product. Like in other analyses, we have to assume that the researcher has included all the relevant variables in the analysis.

While there may be other ways of partitioning the market using dependent variables, the most commonly used technique is the Automatic Interaction Detector or AID for short.

Automatic Interaction Detector (AID)

What is AID?

AID, or Automatic Interaction Detector, is a technique which groups people on attributes that contribute most to a dependent variable. The AID technique, which is based on the Analysis of Variance principle, was originally developed by Morgan and Sonquist at the Survey Research Center, University of Michigan. There are several contexts in which this technique could be applied in market research: as a form of cluster analysis when we want to group consumers; as a means of isolating the important attributes of a product; or even as a preliminary effort to understand the market structure.

Here are some examples of problems that can be solved using the AID technique:

• What demographic characteristics combine to define readership segments in the newspaper market?
• What pscyhographic characteristics define different beer drinkers with common drinking patterns? What are their other characteristics such as demographics?

How Does AID work?

In AID we use several pieces of information (e.g. demographic traits) to explain one aspect of behaviour (e.g. purchase intention). The one aspect that we try to explain is called the *Criterion Variable*. The variables we use to explain the criterion variable are called the *Predictor Variables.*

In AID, we simply input a *Criterion Variable* (like the ratings of purchase intention for a product) along with the *Predictor Variables* (like the demographics of the respondents).

The technique then searches for the predictor variable that is most closely associated with the criterion variable (e.g. purchase intention) and divides the sample into two parts on that basis. Following this, the next most closely related predictor variable is selected and each of the two subgroups are split in turn into two additional subgroups. This procedure is repeated for all the remaining variables. Given a dependent variable (like consumption volume), AID repeatedly tries a number of dependent variables (demographics, ratings, etc.) to determine which variable contributes most to the dependent variable. The groups obtained in this way are the market segments, some of which may be of interest.

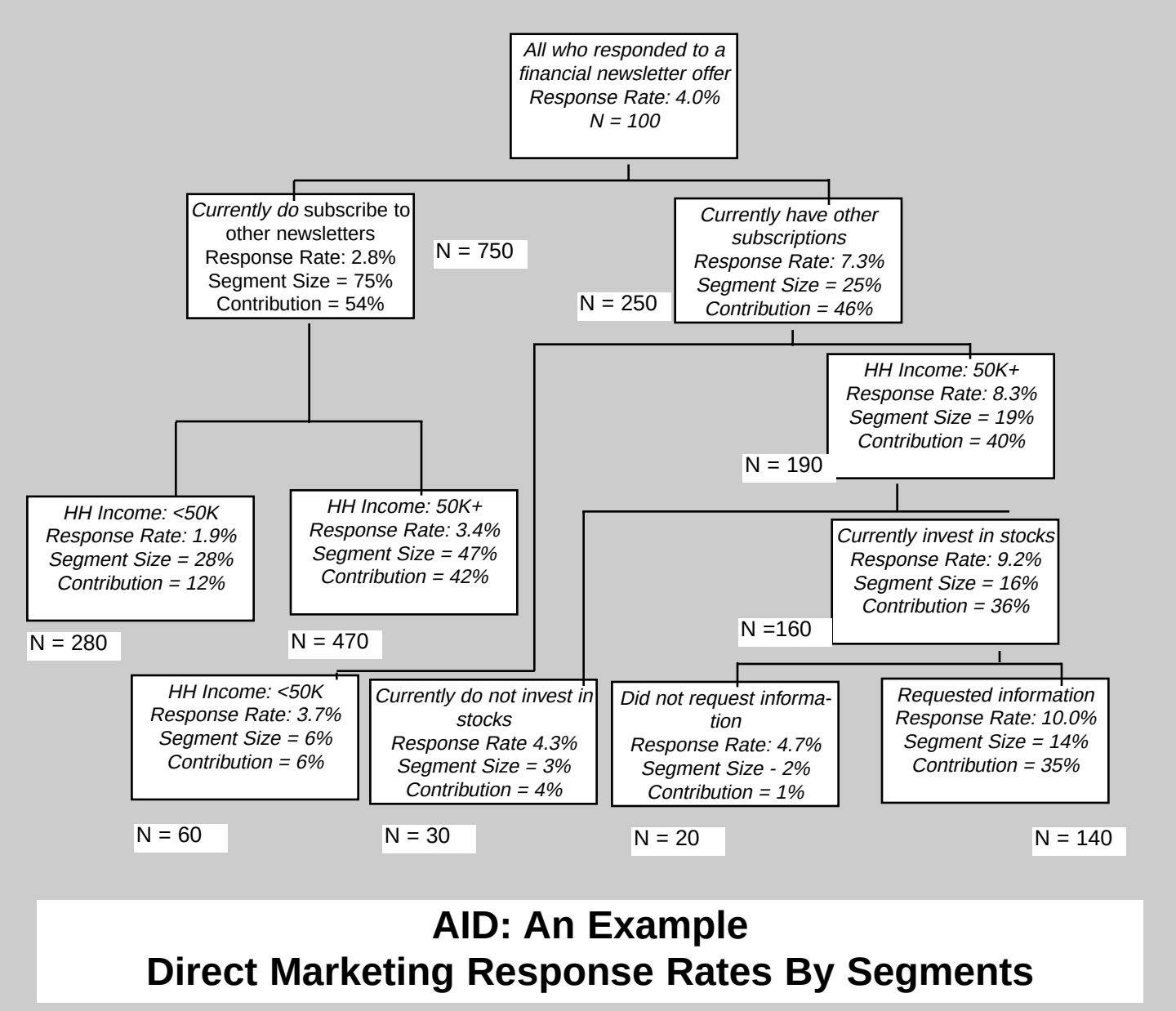
The typical output we get is a ‘tree diagram’ which groups people that contribute most of the dependent variable (see chart). From this, we may infer:

1. The relative importance of different variables;
2. The number of respondents who fall into the segments that may be of importance to the marketing strategist; and
3. The hierarchy of relationship among variables.

AID is a relatively simple type of segmentation analysis that has been available for about 25 years, but has been largely neglected by the market research industry. One reason may be that people have regarded segmentation as synonymous with complex analysis of scale batteries and for this purpose AID is less suitable.

It is more appropriate when the potential predictor variables being investigated are situational or demographic variables rather than attitude or image scale items. The groups of people at the bottom of the roots are basically clusters defined by the hierarchical order of the splits that have defined

them. This can be useful in forming the analysis categories for subsequent analysis. An additional advantage is that the final output is in a form that is easily understood by a marketing person.



AID: An Example
Direct Marketing Response Rates By Segments

The format of the problems usually tackled by AID may be stated as follows: How is (Y ...) affected by X1, X2, X3, X4 ...?). Once we identify the segments in terms X1, X2, X3, X4 etc., we can go ahead and define the segments in terms of these variables. We can also go a step further and identify how the segments differ in attributes that were not used to segment the market.

One main weakness of the method is that it requires a fairly large sample size because it splits the sample into two at each stage.

Interdependence Analysis Approach to Segmentation

In interdependence analysis, the researcher does not specify a dependent variable. Instead, the researcher aims to find out how similar or different consumers are on the basis of a number of variables. Segments are formed by identifying people who answer a set of questions similarly. The most commonly used interdependent method is cluster analysis.

Cluster Analysis

What is Cluster Analysis?

Cluster Analysis refers to a group of techniques that are used in market research primarily to group consumers who are similar. Thus we can group consumers who want similar benefits from a product, consumers who have similar demographics, consumers who have similar reading habits and consumers with similar life-styles, etc.

For a given set of variables, Cluster Analysis divides the respondents into several groups such that within any one group, respondents are similar to one another in their responses to the variables. We are not specifically interested in one attribute at a time, but in similar *patterns* of ratings or similar *patterns* of demographics.

Most market segmentation studies utilize cluster analytic techniques. The type of data used for this purpose tend to be either psychographic or attitudinal. Consumers may be asked to respond to a number of life-style questions (psychographics) or to rate the product benefits. Respondents may then be grouped into homogeneous clusters based on their life styles or on their product benefit ratings. Once the respondents are grouped into distinct segments, we may analyze the data further by cross-tabulating these segments against demographic characteristics, usage patterns, brand ratings, etc.

There are basically two kinds of cluster analysis:

1. *'Tear Down' Techniques:* This group of techniques takes the total sample and breaks it down into meaningful subgroups.

2. *'Build-Up' Techniques:* This group of techniques starts with each respondent in the sample separately and builds groups, adding one person at a time to a group, based on how similar each individual is to another individual.

A wide range of techniques are included under the umbrella term *Cluster Analysis*. For example, the AID technique, Q technique, MCA analysis, hierarchical clustering, smallest space analysis and even factor analysis can be used to cluster people. (Although we treated AID as a distinct dependence method, it is in fact a cluster analytic technique, when we use the term to include all partitioning techniques).

Formal Segmentation Research Techniques

Analysis Type	Typical Techniques
Market Structure Analysis	Correspondence Analysis, Discriminant Analysis, Multidimensional Scaling
Market Partition Analysis	
Dependence Analysis	Automatic Interaction Detector (AID)
Interdependence Analysis	Cluster Analysis - Hierarchical and k-Means

Though all these techniques have one main purpose - classifying people (or occasionally variables), the final output might differ from technique to technique. The analyst is free to choose any level of clustering that would suit the objective of the research study.

General Considerations

There are several ways in which one can determine how close or far apart people are. For example, in determining how close two respondents are:

1. We may consider all differences in attributes to be of equal importance; or

2. We may consider larger differences to be more critical than small differences.

Several other assumptions are possible. Cluster Analysis procedures will differ depending on the assumptions we make. Technically, different assumptions give rise to different 'distance functions'. It is not necessary to know the intricacies of these distance functions to use cluster analysis. It may be stated in passing that most techniques make use of the Euclidean distance function.

The researcher should consider several factors while using cluster analytic techniques. First, different forms of cluster analysis may give different solutions since the criteria on which clustering is achieved may differ form technique to technique. Second, if the sample size is not large enough, we may end up with clusters with very small bases. Generalizations based on sample bases may turn out to be unreliable. Third, although most techniques have the same purpose (i.e. classifying people or variables), the logic of a given technique may make it inappropriate for a given problem.

Another serious consideration is the data base on which clusters are derived. Cluster analysis groups people on the variables the analyst manipulates (considers relevant). If those variables are in fact irrelevant, the groups derived on the basis of those variables will also be irrelevant.

In spite of the above caveats, cluster analytic techniques are useful in practice, especially in the area of segmentation research.

How Many Segments?

No matter which method we use, we can derive as many as there are respondents. How do we know how many 'true segments' there are in reality?

There are mathematical criteria to let us know how many segments there are in our data. But unfortunately mathematical criteria may not coincide with consumer-driven segments.

One recent program by Richard Johnson uses a convergent approach ie. producing two different solutions at each level and comparing them to determine whether they are close.

The most widely used approach is to identify several solutions and use the one that is most closely alighted with the marketers objectives.

Usefulness of Segmentation

Almost every market research analysis, including cross-tab analysis, assumes that different people react differently to products and services. Hence the usefulness of segmentation is beyond question. But what is open to question is how useful these sophisticated techniques are in identifying the segments.

Although no definitive answer can be given to this question, many marketers find the segments produced by these techniques very usable. It appears that segmentation is more suitable for products with multiple benefits and images (such as beer and automobiles) than for products that have a strictly unitary utility function (such as refrigerators).

However, these are broad generalizations. It is difficult to predict beforehand whether segmentation (especially the LVS - Latent Variable Segmentation - variety) will prove to be useful or not for a given product.