

Case Solution

Conglomerate's New PDA (2001)

By Gary L. Lilien and Arvind Rangaswamy

Case Background

The case presents a standard segmentation and targeting challenge facing Netlink. It involves a relatively common form of segmentation, where the firm has already developed an offering (if there has been customer input, no information about it is in the case) and is seeking to determine a market for it.

Analytic Background

For a segmentation study to be useful, it should result in a manageable number of target segments (often between three and eight), each of a substantial size. It is difficult for most companies to manage more than a few distinct segments.

A segmentation model requires a dependent variable, usually called a segmentation basis, and independent variables, or segment descriptors. The segmentation basis should describe why customers respond differently (e.g., differences in their valuations of offerings, needs, or wants), whereas segment descriptors (e.g., age, income, use of media) help marketers deliver different offerings to various customer segments. In practice, the distinction between bases and descriptors depends on the reasons for conducting the segmentation study.

Analytical methods such as regression or discriminant analysis (the approach we employ in the Marketing Engineering software) relate segment membership to descriptors; the equation that results from such analyses (e.g., discriminant function, classification tree) indicates whether a potential customer who was not part of the segmentation study would belong to a specific segment. In practice, many segmentation approaches focus on descriptors as a means to segment customers, though to be most effective, they should identify segments using basis variables, particularly the observed or likely actions of customers. Sophisticated approaches blend both basis and descriptor variables to develop a segmentation model. The management problem at hand, combined with the cost and information about availability, should point to the best approach. Moreover, the best approach depends on the reason for undertaking the segmentation study. In other words, there is no single best segmentation approach.

Segmentation consists of two phases:

Phase 1: Segment the market using basis variables (e.g., customer needs, wants, benefits sought, problem solutions desired, preferences, values, usage situations).

Phase 2: Describe the market segments identified using variables that help the firm understand how to serve those customers (e.g., shopping patterns, geographic location, clothing size, family size), how to talk to these customers (e.g., media preferences and use, attitudes, activities, interests, opinions), and buyer switching costs (costs associated with changing products or suppliers).

Targeting consists of three phases:

Phase 3: Evaluate the attractiveness of each segment using variables that quantify the demand levels and opportunities associated with each segment (e.g., growth rate), the costs of serving each segment (e.g., distribution costs), the costs of producing the offerings the customers want (e.g., production and product differentiation costs), and the fit between the firm's core competencies and the target market opportunity.

Phase 4: Select one or more target segments to serve on the basis of their profit potential and fit with the firm's corporate strategy; determine the level of resources to allocate to those segments.

Phase 5: Find and reach targeted customers and prospects within targeted segments in a variety of ways, including direct mail contact, advertising in selected media vehicles, targeted sales force presentations, and the like.

Following these phases, the firm must identify a positioning concept for its products and services that attracts target customers and enhances its desired corporate image. It is useful to follow this case with the Marketing Engineering Case, *Positioning the Connector 2001*, which addresses the latter issue.

Students should read pp 52 to 72 of *Principles of Marketing Engineering* to get the basic ideas; the *Segmentation Technical Note* at http://mktgeng.com/instructor/downloads/technical_notes.cfm should be given to more advanced students who want to better understand the technical issues involved in segmentation. There is also a webinar on the topic at <http://mktgeng.com/instructor/downloads/webinars.cfm>

Teaching the Case

A good opening topic for discussion is data collection: sample selection and questionnaire development. For this exercise, involving exploratory research, 160 people were surveyed using two questionnaires: a "needs questionnaire" (seven scaled and two absolute value questions) and a "demographic questionnaire" (two absolute value questions: Age and Income, on scale from one to four--Education, one question scaled from one to seven--Away, and fourteen categorical questions). Students will and should question the appropriateness of the sample and the quality of the data collection process in general.

There are a fair number of technical issues involved to both address the case and the associated analysis; hence we recommend that instructors use the full case and software mainly for more advanced classes.

For more introductory classes, the instructor should either run the analysis interactively with the students or provide the output of various steps of the analysis as presented below to the students and ask them for their interpretation and recommendations.

Question 1

Run only segmentation analysis (without Discrimination) on the data to try to identify the number of distinct segments present in this market. Consider both the distances separating the segments and the characteristics of the resulting segments.

Solution

The setup box for this exercise for nine clusters looks as follows.

Setup Segmentation/Targeting

Please select the options to run cluster analysis.

Analysis Options

Number of Clusters: 9

☒ Hierarchical Clustering

☐ K-Means

Segmentation Data

☐ No transformation (raw data)

☒ Standardize data

☐ Reduce data through Factor Analysis

☒ First column contains respondents' ids

Display Options

☐ No highlights

☐ Lowest vs. highest

☒ Statistical differences

Discriminant Data

☐ Enable discriminant analysis

Significance Level (%): 5

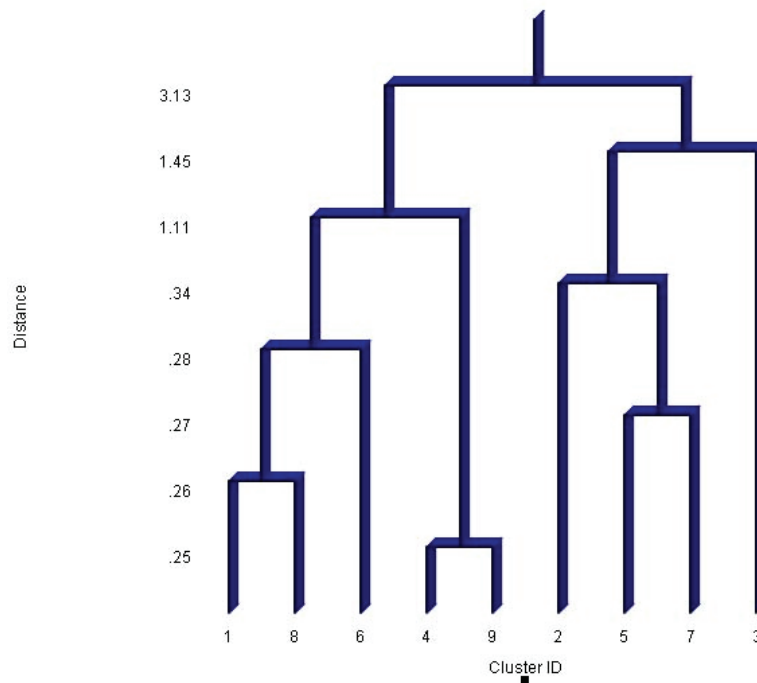
Segmentation/Targeting Data Selection

Next, you will be asked to select Cell Ranges for Segmentation Data (with the first column containing respondents' ids) and Discriminant Data.

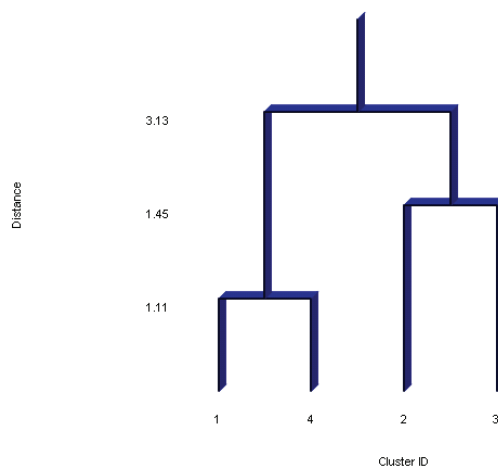
☐ Save choices in current spreadsheet

Cancel Next >

By running the default Hierarchical Clustering—Ward's method for nine segments—we get the dendrogram below:



As the distance measure shows (Error Sum Square), the following clusters are quite close together and can be combined with a small loss in consumer grouping information: i) clusters 4 and 9 at 0.25, ii) clusters 1, 6 and 8 at 0.28, iii) clusters 2, 7 and 5 group at 0.34, resulting in 4 clusters. However, when going from a four-cluster solution to a three-cluster solution, the distance to be bridged is much larger (1.11); thus, the four-cluster solution is indicated by the error sum of squares. In addition, four seems a reasonable number of segments to handle based on managerial judgment. If we run the analysis again, now set to four segments, the program will perform the agglomeration for us.



(Another Cluster Analysis option is the K-Means procedure that attempts to identify relatively homogeneous groups of cases based on selected characteristics, using an algorithm that can handle large numbers of cases better than hierarchical methods. If a K Means procedure is chosen, its four-

cluster solution is very similar to the four-cluster solution based on the Ward's hierarchical clustering procedure.)

Question 2

Identify and profile (name) the clusters that you select. Given the attributes of Connector, which cluster would you target for your marketing campaign?

Solution

From the Segmentation tab, we get cluster sizes and the mean for each variable in each cluster.

To characterize clusters we look for means that are either well above or well below the overall mean (color coded).

Cluster Sizes

The following table lists the size of the population and of each segment, in both absolute and relative terms.

Size / Cluster	Overall	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Number of observations	160	56	51	16	37
Proportion	1	0.35	0.319	0.1	0.231

Segmentation Variables

Means of each segmentation variable for each segment.

Segmentation variable / Cluster	Overall	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Innovator	3.47	3.71	2.43	2.19	5.11
Use Message	4.21	3.68	5.63	3.19	3.49
Use Cell	5.56	5.84	5.43	4.31	5.84
Use PIM	4.01	5.89	2.33	3.06	3.86
Inf Passive	4.45	5.02	3.88	6.12	3.65
Inf Active	4.5	5.2	3.9	6.25	3.51
Remote Access	3.99	3.86	5.04	5.31	2.16
Share Inf	3.71	3.39	3.73	6.12	3.14
Monitor	4.79	4.29	5.55	5	4.43
Email	4.72	5.96	3.31	2.88	5.59
Web	4.47	5.66	3.04	1.44	5.97
Mmedia	4.01	5.2	2.45	1.94	5.27
Ergonomic	4.63	3.95	4.16	5.5	5.95
Monthly	28.8	24.6	25.3	45.3	32.6
Price	332	290	273	488	411

Cluster 1. Phone users who use Personal Information Management software, to whom Email and Web access, as well as Multimedia capabilities are important.

Cluster 2. People who use messaging services and cell phones, need remote access to information, appreciate better monitors, but not for multi-media usage.

Cluster 3. People in this cluster have a higher propensity to work in remote areas and have a high need for fast information sharing (receiving as well as

sending). They do not use email, the Web, or Multi-media extensively, but do require a handy, non-bulky device.

Cluster 4. Innovators who use cell phones a lot, have a high need for Email, Web, and Multi-media use. They also require a sleek device.

If you do this case in class together, it is useful to have the class speculate (guess) the likely job titles of the predominant occupations in each of the segments.

Relative to other products in the market, the Connector is very high on information active, information passive and data sharing...characteristics of Cluster 3, a tentative segment for possible targeting.

Question 3

Rerun the analysis in Exercise 1 with Discrimination. How would you go about targeting the segments you picked in question 2?

Solution

By checking the Discrimination option in the Setup, we get the results of the discriminant analysis on the Discrimination tab of the output with the four-cluster grouping as the dependent variable, and the demographic data as the independent variables. The first discrimination function below explains 51% of the variation. According to its coefficients, the four groups are particularly different with respect to the amount of time away from the office. In addition, the function shares high correlation with the level of education, possession of a PDA and income. The second function explains 32% of the variance and primarily distinguishes the occupation types construction/emergency from sales and service. The third function separates the sales from the service types.

Below the discriminant function we show the *Confusion Matrix* which shows how well the descriptor variables can classify the individuals, the *Classification Weights* and *Classification Coefficients*, needed if you want to classify individuals in an external data base and finally, the *Discriminant Variables* for each cluster, used to characterize the mean values of individuals in those clusters.

Discriminant Function

Correlation of variables with each significant discriminant function (significance level < 0.05).

Discriminant variable / Function	Function 1	Function 2	Function 3
Away	-0.705	-0.132	0.116
Education	0.704	0.098	-0.035
PDA	0.669	0.219	0.114
Income	0.629	0.138	-0.266
Business Week	0.405	-0.062	0.055
Mgourmet	0.276	0.15	-0.164
PC	0.28	-0.549	-0.073
Construction	-0.197	0.37	0.036
Emergency	-0.161	0.363	0.027
Cell	0.156	-0.348	-0.011
Computers	0.211	0.297	-0.061
Sales	-0.014	-0.386	0.652

Service	-0.308	-0.308	-0.468
Age	-0.002	0.069	0.409
Field & Stream	-0.347	0.103	-0.379
PC Magazine	0.048	0.075	-0.354
Professional	0.327	0.014	-0.338
Variance explained	50.48	31.68	17.84
Cumulative variance explained	50.48	82.16	100
Significance level	0	0	0.001

Confusion Matrix

Comparison of cluster membership predictions based on discriminant data, and actual cluster memberships. High values in the diagonal of the confusion matrix (in bold)

indicates that discriminant data is good at predicting cluster membership.

Actual / Predicted cluster	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 1	34	11	4	7
Cluster 2	10	35	5	1
Cluster 3	1	3	12	0
Cluster 4	3	0	0	34

Actual / Predicted cluster	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 1	60.70%	19.60%	07.10%	12.50%
Cluster 2	19.60%	68.60%	09.80%	02.00%
Cluster 3	06.20%	18.80%	75.00%	00.00%
Cluster 4	08.10%	00.00%	00.00%	91.90%

Hit Rate (percent of total cases correctly classified)

71.88%

Classification Weights

Sum of each segment's projection on each function.

This matrix was used internally, and will be required to run further discriminant analysis (i.e., classification) on external data.

Clusters / Discriminant Functions	Function 1	Function 2	Function 3
Segment 1	2.512673	-2.792224	-0.7731534
Segment 2	2.052577	-2.885595	-1.297725
Segment 3	1.458223	-1.586156	-1.002282
Segment 4	3.506227	-2.317815	-1.185386

Classification Coefficients

Coefficient for each variable in the discrimination function.

Coefficient for each variable in the discrimination function.

Discriminant Variables / Functions	Function 1	Function 2	Function 3
Age	0.00177353	0.00542933	0.00926584
Education	0.2751206	-0.0268938	-0.006272
Income	0.00877789	0.00254923	-0.0029435
Construction	0.1308967	1.47771	-1.005192
Emergency	-0.132206	1.731755	-0.6726465
Sales	0.4934896	0.2363796	-0.6000023
Service	0.2717406	0.311923	-1.474535
Professional	0.6225228	0.686554	-1.284242
Computers	0.3594805	0.9053028	-0.95706
PDA	0.620912	0.2150456	0.1564206
Cell	-0.036742	-0.5119913	0.08142119
PC	0.9407754	-2.93769	-0.4710419
Away	-0.1075443	-0.0527017	0.05125611
Business Week	0.2918053	-0.1893429	0.1019987
PC Magazine	-0.0782037	-0.0777597	-0.3637948
Field & Stream	-0.2029932	-0.1621621	-0.4722914
Mgourmet	0.4969293	0.2193819	-0.3152554

To characterize these segments, we look for variable means that are either well above or well below the Overall mean.

Discriminant Variables

Means of each discriminant variable for each segment.

Discriminant variable / Cluster	Overall	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Away	4.206	4.357	4.843	5.375	2.595
Education	2.506	2.482	2.196	1.938	3.216
PDA	0.438	0.446	0.176	0.188	0.892
Income	66.894	62.589	60.529	52.438	88.432
Business Week	0.275	0.304	0.176	0	0.486
Mgourmet	0.019	0	0	0	0.081
PC	0.981	1	1	0.812	1
Construction	0.081	0.054	0.059	0.312	0.054
Emergency	0.038	0.018	0.02	0.188	0.027
Cell	0.875	0.911	0.902	0.625	0.892
Computers	0.231	0.179	0.137	0.312	0.405
Sales	0.3	0.536	0.235	0.063	0.135
Service	0.181	0.107	0.392	0.125	0.027
Age	40.006	43.071	36.765	42.188	38.892
Field & Stream	0.125	0.036	0.235	0.312	0.027
PC Magazine	0.244	0.143	0.294	0.25	0.324
Professional	0.162	0.089	0.157	0	0.351

We can now describe these individuals as follows:

Cluster 1 - Sales Pros:

Cluster 1 consists mainly of sales professionals: 54% of the cluster members indicated Sales as their occupation. They use the cell phone heavily, and many (45%) own a PDA already; practically all have access to a PC. Their work often takes them away from the office. They mostly read two of the selected magazines: 30% read BW. From the needs data, we see that they are quite price sensitive.

Cluster 2 - Service Pros:

Cluster 2 is made up primarily of service personnel (39%) and secondarily of sales personnel (23%). They use cell phones heavily, but only about one fifth currently use a PDA. They spend much time on the road and in remote locations. They read PC Magazine, 29%. From the needs data, we see that they are quite price sensitive.

Cluster 3 - Hard Hats:

Cluster 3 is made up predominantly of construction (31%) and emergency (19%) workers. They use cell phones, but usually do not own a PDA. By the nature of their work, they have high information relay needs and generally work in remote locations. They exchange information with colleagues in the field (e.g. construction workers on the site). Many read Field & Stream (31%) and also PC Magazine. Note also from the needs data, that they are the least price sensitive (willing to pay highest price plus monthly fee) and also have the lowest income.

This apparent anomaly occurs because these folks are less likely to have to pay for the device themselves, raising the question of whose preferences—their own or their employers'—will drive the adoption decision.

Cluster 4 - Innovators:

Cluster 4 represents early adopters (see needs data), predominantly professionals (lawyers, consultants, etc.). Every cluster member has access to a PC, 89 percent already own PDAs. They read many magazines, especially BW 49%, PCMag 32%. Most are highly paid and highly educated.

Question 4

How has this analysis helped you to segment the market for ConneCtor?

Solution

Profiling and the targeting of segments were enhanced by applying the discriminant data. ConneCtor is a wireless communicator for both data and voice and an electronic scheduler. The segments that may need it are Hard Hats (primarily) and Service pros secondarily. Weaker cases can be made for the other segments.

Hard hats not only have high need for the features but a case could be made (to their employers) that they would be able to perform their jobs significantly better with such a device.

It would be important to understand the decision making process in the organizations where these folks work—both who is involved and what criteria these individuals value (which may be different than the "users" here).

Conglomerate may have to redesign the PDA to fit it to the needs of this segment. Construction sites, burning buildings, oil rigs and the like will require a sturdy, easy to see (orange? yellow?) product, resistant to dust, temperature variations, shocks and the like.

Service pros also work in remote location and would benefit from wireless communication with dispatchers; hence their stated and latent needs are compatible with the features of the PDA. Their high price sensitivity needs to be investigated, as does the organizational decision process in similar fashion for the hard hats.

Students should be given extra credit if they applied a method like GE to prioritize the segments, explicitly defining evaluation criteria (segment size, compatibility, competition, profitability, channel access, etc), developing importance weights for those criteria and ratings for each segment along these criteria. That process will allow them to obtain explicit, quantitative ratings. Some will argue that it is too soon to do such a formal evaluation; you might suggest that it is never too soon to get that process started.

Question 5

What concerns do you have with the approach (data collection, analysis, etc.) so far?

Solution

Population specification: Who are the right respondents?

1. The current survey was performed on a selected section of the total market. The survey specifically screens out everyone who does not own or seriously considers purchasing a PDA and only focuses on business users whose job requires them to spend time away from the office. We may have excluded potential buyers or whole segments from the broader market.
2. These are company purchasers-should users respond? Corporate management? Purchasing? These questions will bring up the data quality issue again.
3. What about the mining industry, etc.?

Questionnaire items:

1. Do we have the right questions in the survey? Especially when asking users to answer questions about prices that they will not be paying, the survey design needs to consider the threat of response bias.
2. Are there items missing? Would you rather have a PDA and cell phone separately, or carry an all-in-one device? What other product features would potential users consider for their service and information needs? More information on the needed ease of data entry, display requirements, special cases, and also other features such as voice recognition should be acquired.

Respondents are not asked to make any tradeoffs amongst their needs. And we don't get information on how well product features satisfy those needs. (We infer these ourselves above). Hence we actually do not have data on how well the segments perceive the ConneCtor, a task for a positioning/perceptual mapping exercise.

Question 6

What are the next steps you recommend for Netlink and the development of ConneCtor?

Solution

It is too early to 'go to market' with the data we have. It would be appropriate to drill down with samples of prospects in the targeted segment(s), through focus groups, use tests or conjoint analysis to help refine the offering and to better evaluate market response. A market sizing exercise is needed as well as a complete marketing plan (price, distribution, communication, promotional plan, etc) and the firm needs to determine if they are introducing a product or a product family. (ConneCtor Tuf for the Hard Hats??). That product family would need to tune product features to the needs to the targeted sub-segments. A next step might be a perceptual mapping exercise to follow this Segmentation and Targeting exercise.

Note also that if we have discriminant analysis data for a large number of potential customers who were not included in the sample data, we can classify each of those customers (probabilistically) into the appropriate segment. We can use this information for a targeted marketing program (customized promotion) that leverages our new knowledge about the segment the customer belongs to. You might illustrate that by running "Run Classification" under the Segmentation and Classification menu after you have run discriminant analysis. You can use all or part of the same discrimination data set to show how the software will insert a column at the right of the data set with the cluster number. Below is a part of that output (with the middle columns suppressed):

Classified Data

Predicted segment memberships of all observations, as determined by discriminant analysis (classification).

Respondents / Discriminant variables and predicted cluster	Age	Education	Income	Field & Stream	Mgourmet	Predicted Cluster
1	48	3	47	0	0	1
2	43	3	89	0	0	2
3	48	2	24	0	0	1
4	28	3	52	0	0	1
5	54	3	84	0	0	4
6	62	2	46	0	0	1
7	53	3	71	0	0	1
8	48	2	39	0	0	1
9	46	3	87	0	0	1
10	59	2	21	0	0	2
11	53	3	76	0	0	1
12	21	3	63	0	0	1
13	54	3	53	0	0	3
14	41	3	81	0	0	1
15	22	2	45	0	0	3
16	37	2	38	0	0	1
17	22	3	77	0	0	1
18	51	2	81	0	0	1
19	54	2	26	0	0	1
20	30	2	67	0	0	1
21	57	2	65	0	0	1
22	40	2	74	0	0	1
23	64	3	48	0	0	3
24	19	2	47	0	0	2
25	41	2	87	0	0	1
26	58	3	48	0	0	1
27	48	3	86	0	0	1
28	45	2	90	0	0	1
29	22	3	51	0	0	1

In a direct marketing application, one would assign individual prospects to segments in this way, where the segment membership would dictate the appropriate marketing program.