Paper 114-2010

# Successful Product Creation by Using Analytics to Discover What Consumers Value

Will Neafsey, Ford Motor Company, Dearborn, MI
Ted Lang, SAS, Southfield, MI

## ABSTRACT

Accurately identifying and understanding potential customers for future products greatly improves the likelihood of market success. This is particularly important for automotive manufacturers that have long lead times and require vast resources to deliver products to market.

Historically, the auto industry developed product ideas through basic market analysis, often limited to demographic research and too frequently the interpretation of the product creator. While this worked satisfactorily when market choices were few, it is no longer a sustainable business method with increased competition. Successful product creation now necessitates focus around the consumers' frame of mind. Deriving actionable meaning from an array of consumer-based opinions through analytics provides simple representations of complex data that offers decision-makers a straightforward view of market spaces. This is the substance of target marketing.

Two powerful SAS® statistical tools enable target marketing. These are principal components analysis through the PRINCOMP procedure, and kernel density estimation using the KDE procedure.

## INTRODUCTION

With greater competition and continually diminishing resources, the automotive industry finds it increasingly important to identify well-defined market opportunities before committing resources to the development of a new product. With increased availability of consumer data and vastly improved hardware and software, analytics can play a very important role in the development of successful new products. Creating quality products that are relevant to the market can mean the difference between large profits and large losses. Today's intense competition makes success increasingly more difficult.

Target marketing analytics offers a means to better identify opportunities that have a higher likelihood of success because they are grounded in consumer data, not just observation and interpretation. In today's market, strictly meeting demographic and social economic backgrounds is not enough to successfully market automotive consumer products. The automotive marketplace is becoming increasingly well-informed, and there are more competitors to satisfy the needs of that marketplace. So, one-dimensional analysis of non-emotive variables is not enough because the new automobile market is far too sophisticated for such an approach. There are numerous historical examples of missteps in product creation, and in this resource-intensive competitive climate, these marketing misjudgments are unsustainable.

## AN ILLUSTRATIVE EXAMPLE OF CAPTURING CONSUMERS' VALUES

Apple Computers has always, either consciously or unconsciously, had a target customer since the day it was co-founded by Steve Jobs and Steve Wozniak in 1976. That day, quite tellingly, was April Fool's Day. From its very beginning, Apple was not for everyone.

In the image of its creators, Apple has always targeted customers who were not accepting of what the industry was offering. From its humble beginnings with the Apple I (a kit computer), Apple was customer-focused, not industry focused. While IBM and Microsoft created a huge industry of non-personal, hard-to-use, uncreative, black and white devices, Apple focused on the question "What would my customer want next?" Quietly, Apple drove the rise in popularity of the "personal computer" by making approachable, easy-to-use machines. IBM and the other competitors rode this wave and its profits.

In 1984 Apple took the next step toward satisfying its core customer with the ground-breaking Macintosh and the famous Ridley Scott directed '1984' Superbowl ad. The Macintosh was tagged as "the computer for the rest of us." More hints at their anti mass-market mindset.

After early successes and acclaim with the Apple and the "Mac," Apple Computer started to drift from its innovative roots. The focus of then CEO John Scully concentrated more on the bottom line and revenues than innovation and target customer. In 1985, John Scully orchestrated a boardroom coup resulting in the resignation of Jobs.

The next decade was marked by a string of unimaginative products, CEOs, competitive struggles, and lawsuits, including a very high profile battle with Microsoft. Meanwhile, Jobs starts a new computer company NeXT, continuing to follow the original target customer. At the close 1996, Apple purchases Jobs' company NeXT and Jobs returns to Apple as an advisor. By the close of 1997, Steve Jobs ousts CEO Gil Amelio and becomes Apples interim CEO.

During the next 3 years, Apple unleashes a torrent of new, innovative products and operating systems all squarely focused on their target customer. This target was loosely articulated in the 1998 Apple Campaign celebrating people who "Think Different." Examples included Albert Einstein, Arthur Miller, and the Dalai Lama. In the fall of 2001, the company that built the Macintosh, "the computer for the rest of us," and celebrated those who "think different," revolutionized the industry again with the iPod. Although conceived for a concisely-defined target customer, the iPod currently claims 75% share of its market and Apple's stock price has climbed nearly tenfold in four years.

## TARGET MARKETING

In its most successful times, Apple didn't develop product to try to satisfy the whole market. It developed product to capture the hearts of a defined and bounded subset of the market. Apple's target marketing of a free-thinking, somewhat contrarian customer was the key to its success.

Target marketing, for the purposes of this paper, is the process of selecting a group of consumers from a particular market space for directing marketing or product development purposes. In particular, we discuss the application of analytics to multidimensional opinion data to identify, quantify, and eventually qualitatively describe the target market in two dimensions. This makes possible a valuable data-based, analytic contribution into the decision-making process at all levels and functions of an organization.

Target marketing involves consideration of many success factors for a new product. These may include:

- Potential market volume
- Brand differentiation
- Avoiding overlap within company offerings
- Avoiding overlap with competitors' offerings
- Ensuring price levels
- Maintaining a desired brand image
- Advertising emphasis

Analytics is essential to gaining knowledge to support these factors. In fact, target marketing supported by analytics is an essential element in determining whether a product achieves market success.

## SELECTING THE TARGET CUSTOMER

### CONSUMER RESEARCH

As mentioned above, understanding the marketplace involves more than directing designs toward consumer demographic and socioeconomic considerations. While these factors remain important, they cannot fully determine the challenging demands of the marketplace. To distinguish a company's products--to rise above the competition—consumers' frame of mind must be considered.

Good consumer research furthers the goals of meeting a customer's intangible desires. Consumer research surveys are structured to provide many dimensions of consumer opinions. Deriving useful information from this high-dimensional opinion data is the ultimate objective of target marketing.

### DATA PREPARATION

The target marketing process begins with a broad based selection of data (observations) from which a narrowed selection of the target customer is derived from business generated criteria. This initial selection provides a homogeneous starting point based on observable, rules-based methods from two general selection criteria. One approach uses simple data filtering criteria based on desired consumer traits (e.g., younger active and more affluent buyers). The other begins with criteria based on the results from a behavioral clustering or segmentation analysis.

Using the first selection method, typically, filters are applied progressively until the desired target remains, with practical considerations (usually sample size) controlling the degree of refinement in the definition of the filters. For example, if the full dataset contains a sample of all automotive buyers and the goal is to target small vehicle buyers,

a subset of data may include only buyers within the small car market. Alternatively, a data selection starting with a group possessing the desired consumer characteristics from a clustered or segmented dataset can likewise yield a small subset from which a target can be identified.

The next step often involves looking at the selected population with respect to some of the more strategic, emotional or tactical elements that were not included in the original selection criteria. This could be anything from styling preferences to spending habits preferences. To effectively do this, it is important to further distill the population to be more homogeneous with respect to the selection criteria and marketing variables. This process will identify a view of where the marketplace's critical mass or its effective center is positioned.

The small, irregular shape in Figure 1 represents a target selected from the desired group (Group 1) using techniques described later. The target is a mathematically-derived homogeneous set of consumers that eliminates the points outside of the critical mass.

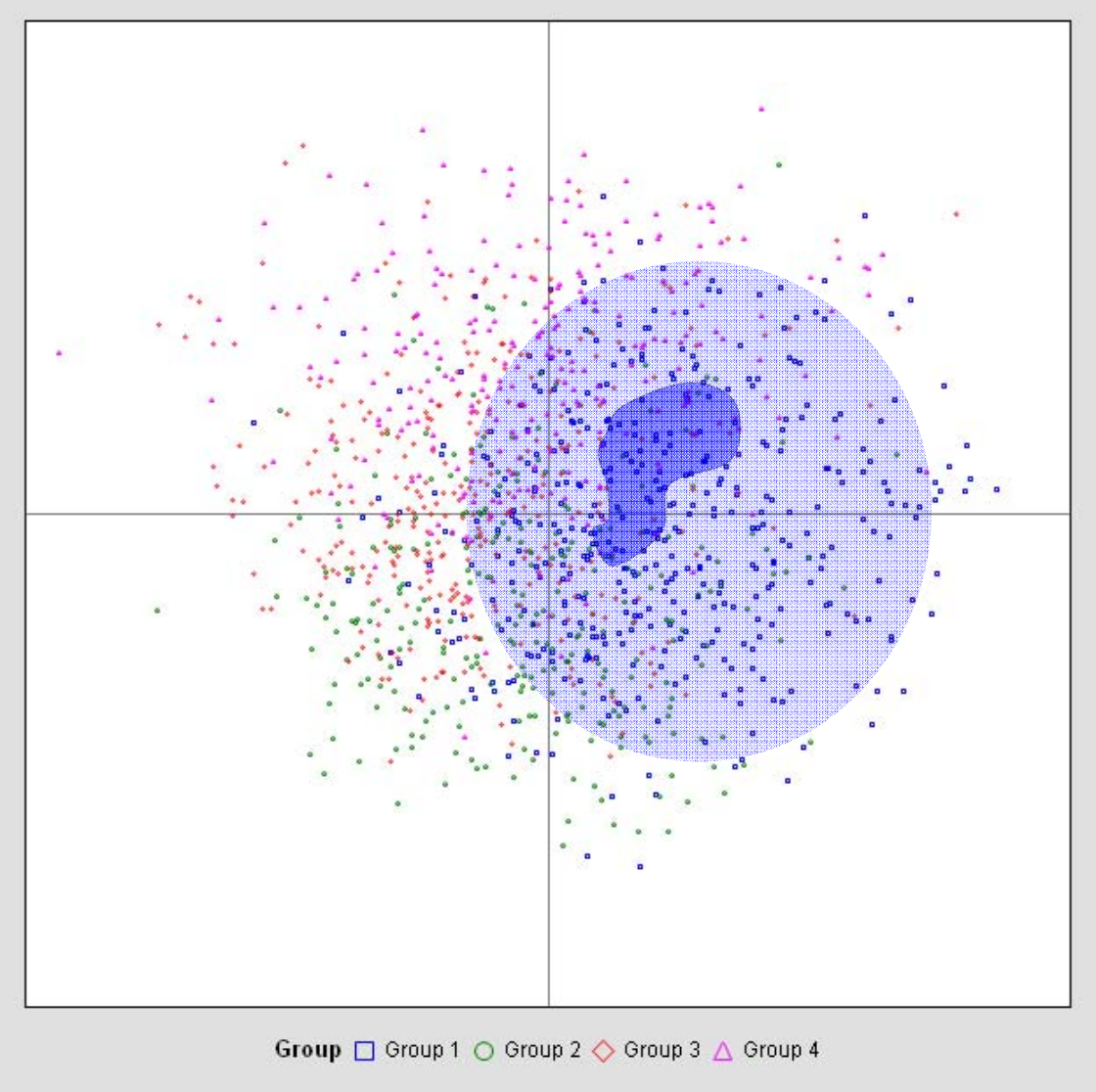


**Figure 1. Selection of consumer groups**

This choice between filtering demographic or socioeconomic criteria versus selecting clusters or segments is purely determined by the nature of the project. Selecting the target from clustered data here simply provides a good illustration of the analytic process.

## ANALYTIC PROCESS STRATEGY

From the introduction above, we take a closer look at the analytic approaches that facilitates this work. The market research surveys provide a high-dimensional view of the consumers' state of mind that must be made "interpretable" for general business audiences. This is typically constrained by the size of the sample, leaving some guesswork around the market concentrations (densities) from which a target is selected. Two techniques are employed to accomplish these analytic functions: dimension reduction through principal components analysis and kernel density estimation.

### DIMENSION REDUCTION IN CONSUMER RESEARCH

Consumer survey opinion questions attempt to thoroughly capture the state of mind of the respondents. Distilling these numerous variables into interpretable and effective information can be achieved through various dimension reduction techniques for interpretability by decision makers (e.g., a marketing manager). The dimension reduction technique illustrated here is principal components analysis.

#### Principal Components Analysis

Principal components analysis (PCA) is applied here for dimension reduction because it functions well using response variables from a typical survey. The responses used here reflect a "level" of agreement with position statements presented to the respondent. The list of questions to analyze can be long, making analysis of the data, without a multivariate approach, very difficult and offers uncertain analytic value. Instead, these variables can generally be represented adequately by a reduced number of dimensions using principal components analysis, permitting the portrayal of the market space in a 2-dimensional consumer map.

PCA involves creating linear combinations of the analysis variables with the first principal component explaining the most variation among the variables. The second principal component is orthogonal to the first and explains the second most variation. The PCA process can create all principal components in this manner to explain the entire variance over all the input variables.

In practice, enough of the total variance is often explained in the first few principal components to satisfy the analytic needs of a project. So, the underlying market behavior contained in a high-dimensional dataset can be adequately captured and reported in the first two principal components.

#### SAS® PRINCOMP Procedure

The SAS® PRINCOMP Procedure performs the principal components analysis here with minimal code.

```
proc princomp data=rawdata
              out=prinout
                    N=2;
      var var1-varN; *Survey variables;
run;
```

Options:

- `DATA=` Raw survey dataset can be used as input.
- `OUT=` Specifies a dataset which contains a copy of the input data with the principal components scores added.
- `N=` Limits the number of principal components to 2 in this example.

The two principal components variables, PRIN1 and PRIN2, in the output dataset are plotted in Figure 2. The dataset is now ready for selecting the target customer.

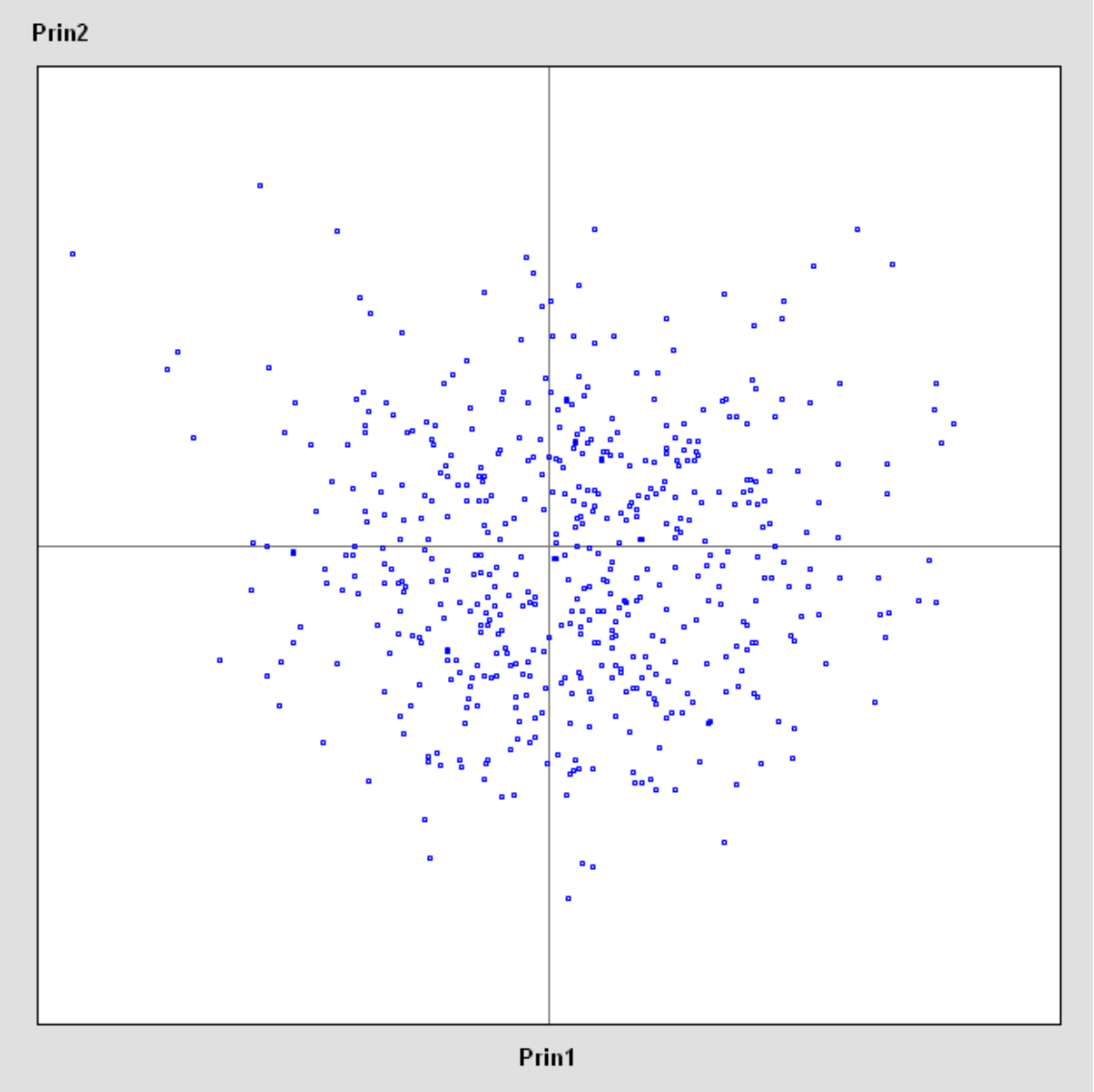


**Figure 2. First two principle components from PROC PRINCOMP**

## SELECTING A TARGET

The results of the analysis to this point have achieved a transformation of the entire dataset for target selection. We want to select a subset "region" of this map to establish a target customer, after which, it is possible that a smaller than desired sample remains to derive useable results.

Positioning individual respondents on a 2-dimensional perceptual map can provide some insight, but there are practical limitations that make this difficult to interpret. A large portion of the map contains no data at all. Also, multiple observations can share effectively the same coordinates, making it difficult to observe their locations on the map and their contribution to the market densities. The set of data possess an inherently nonstandard shape that also makes interpretation difficult. These concerns are illustrated in the Figure 3 below.

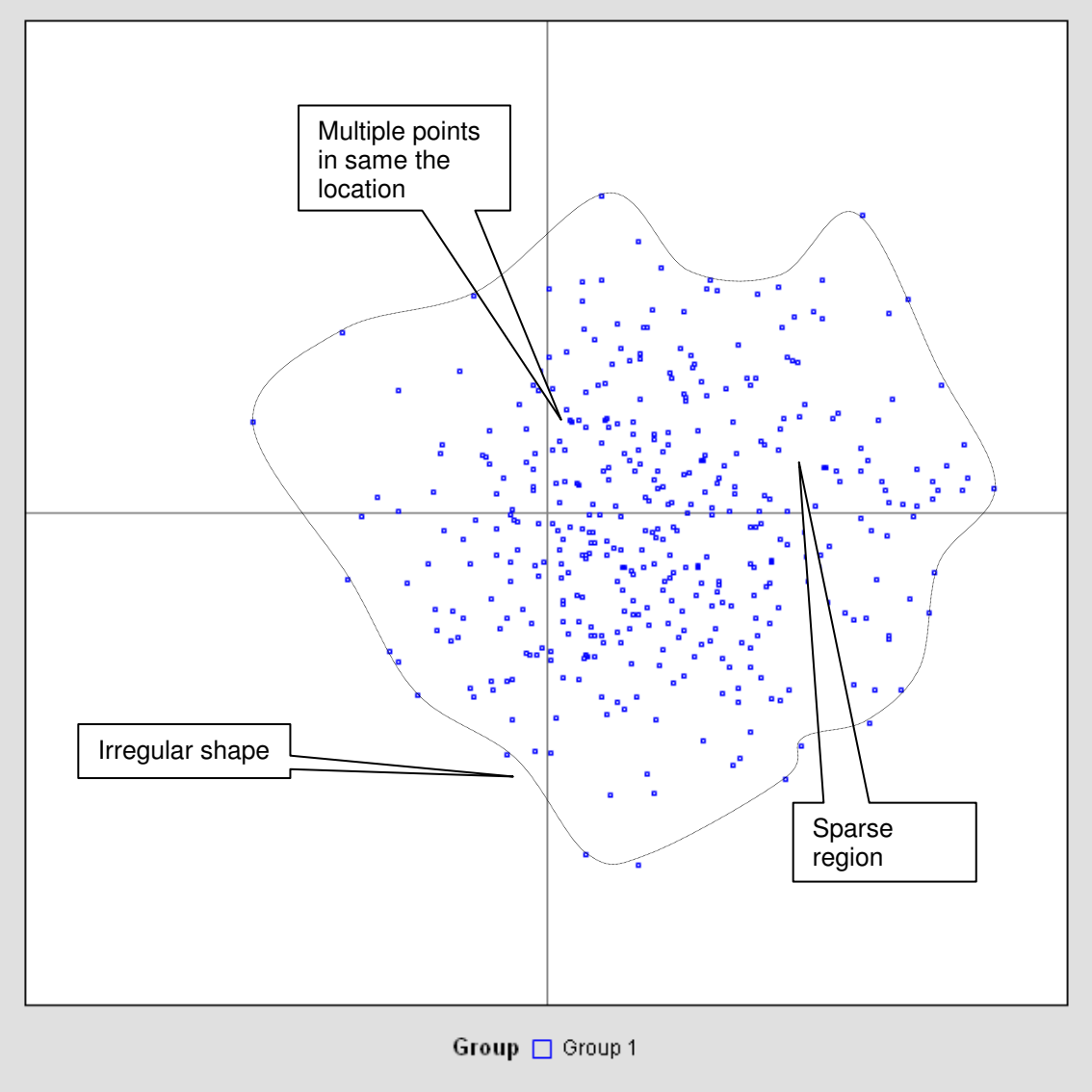


**Figure 3. Sparse distribution and overlaid samples from filtering**

## KERNEL DENSITY ESTIMATION

Plotting the location of the sample on a 2-dimensional plot alone is not informative without defining the space between the actual data points. Calculating the relative densities from the available remaining sample, for example, in many cases does not realistically represent the distribution of the data needed to complete the analysis (e.g., consider the number of "cells" with no data at all). The geometric center of the data may not accurately represent the location of the target customer given the number of points far away from the center.

Kernel Density Estimation (KDE) addresses these concerns by providing a way to develop the probability densities from the available data. KDE uses a nonparametric approach to craft a distribution using the locations of the existing points, treating them as a center and assuming a distribution around them. The results from KDE provide an overall probability density that enables the selection of a meaningful, homogeneous target. A brief discussion of the steps for developing the densities through KDE follows:

Looking at a “qualifying” selection from the filtering process may yield sparse data as shown in Figure 4a (this example uses an unusually small set for demonstration purposes that would not permit further analysis in practice). It is difficult to gain knowledge from this picture, but looking at the pattern of the points indicates there is potentially useful information.

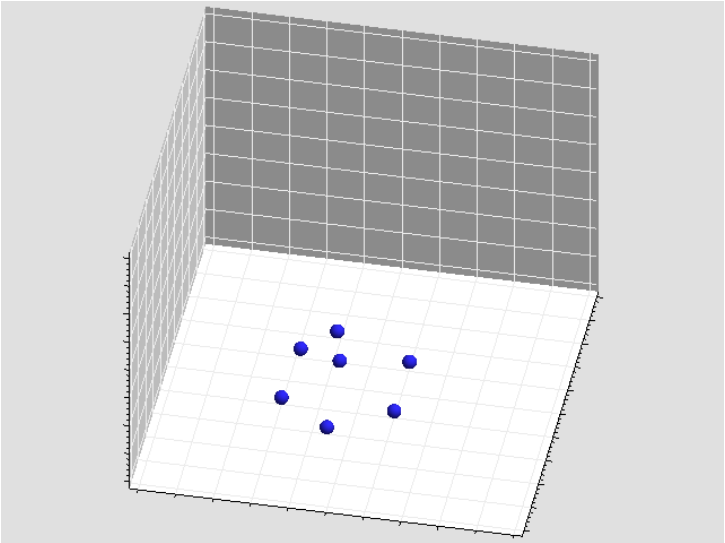

**Figure 4a.**

KDE assumes a distribution for each point giving it its own probability density as shown in Figure 4b. The assumption here is that the distribution (the kernel) for each point is normal.

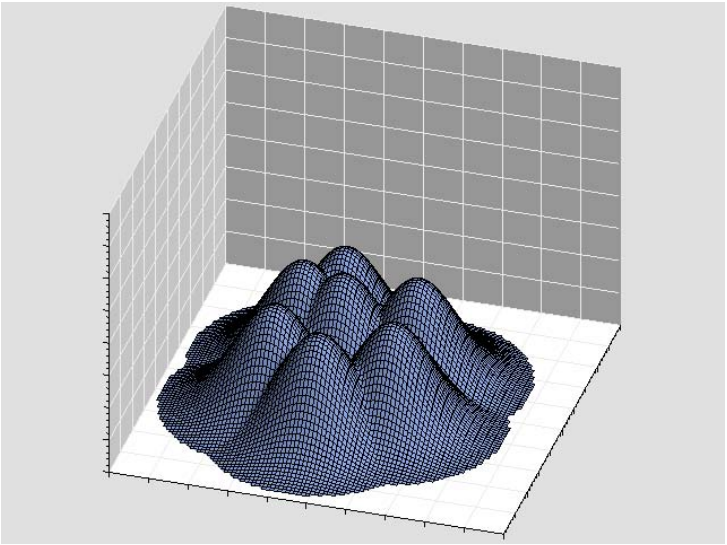

**Figure 4b**

The probability densities for the entire 2-dimensional space are derived by summing these individual “distributions” as shown in Figure 4c.

The height and shape of the resulting distribution are affected by the bandwidth multiplier, which controls the degree of smoothing.

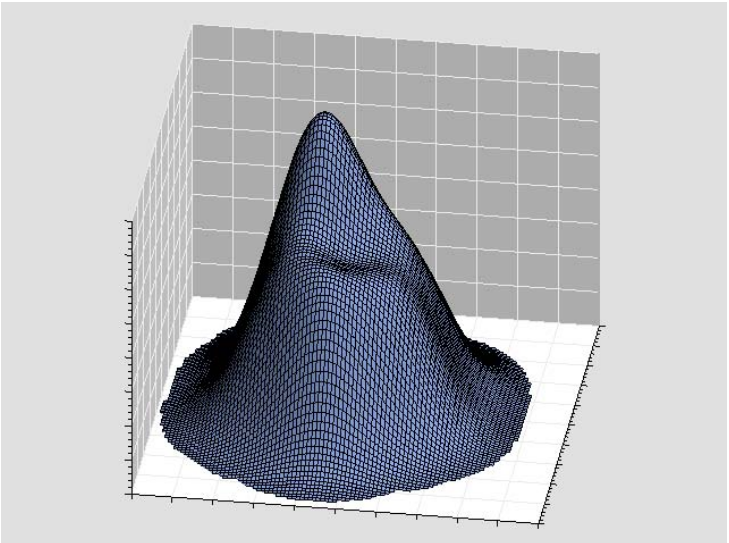

**Figure 4c**

### SAS® KDE Procedure

Given the previously created dataset with the first two principle components created by PROC PRINCOMP, the market distribution is derived using PROC KDE. Because we are using two dimensions, PROC KDE is coded in this case to perform bivariate kernel density estimation as shown below.

Given a dataset with two analysis variables (principal components here), the BIVAR statement requests a bivariate function for the PRIN1 and PRIN2 variables provided by the PROC PRINCOMP output.

```
proc kde data=prinout;
       bivar prin1 prin2 /
                     out=kdedata
                     bwm=1
                     ngrid=32 ;
       weight wgt;
run;
```

Options:

OUT= names the output dataset that contains *ngrid* by *ngrid* observations (for each BY value if a BY statement is used).

NGRID=32 sets the number of discrete units for each dimension of the map to 32. The number of bins effectively provides the degree of resolution for the attitudinal map.

BWM=1 specifies the bandwidth multiplier controlling the degree of smoothing for the generated distribution. This setting impacts the shape of the distribution. Generally, larger values of BWM= have a greater smoothing effect on the distribution.

Figure 5 is a plot of the probability density bands obtained from the PROC KDE. This is the initial step of the target marketing process; for example, the Group 1 cluster could be the beginning point for selecting the target customer. The goal is to determine the smallest band that yields a large enough market concentration and still provides the desired homogeneous aspirational customer for the proposed product.

The groups' centroids are shown as single points. Note that they are not necessarily contained within the inside density band. This is the nature of the problem that is solved by the approach described here--the actual distribution is not centered with the mean (see **SELECTING A TARGET** above).

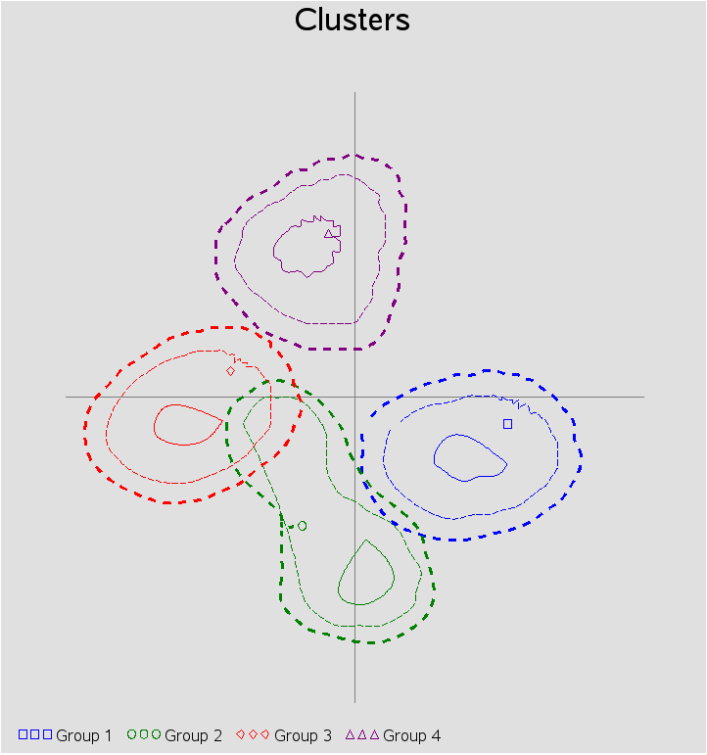


**Figure 5. Estimated distributions from PROC KDE**

## TARGET MARKETING PROCESS

The techniques discussed above can be straightforwardly incorporated in a standard process to investigate a nascent product concept. The analysis may be best understood through an example, which starts with the dataset that produced Figure 5 above.

PROC KDE option N= is arbitrarily set to 32 in this example. This allocates the densities to a 32 by 32 grid. This a relatively low resolution map of the consumer market, but it provides enough resolution to support business decisions. We start this analysis from a clustering exercise. We then look at the distribution of the market for each group. Individual groups are given their own plot to compare against the total market. The total market is shown in Figure 6.

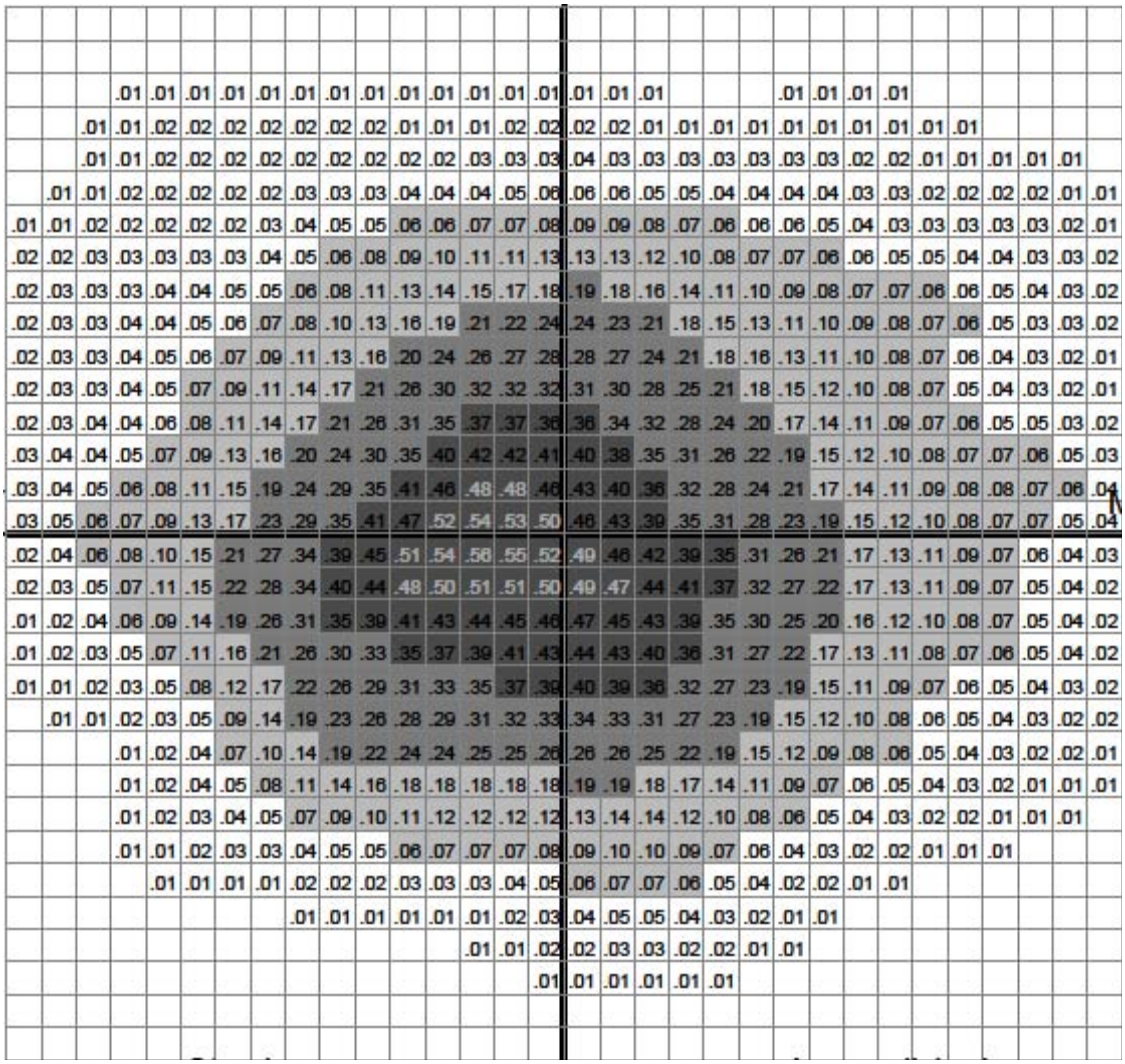

**Figure 6. Total market consumer opinion map**

This is the starting point for new product exploration. The numbers reflect the density percentage of each cell. You observe that the largest concentrations of consumers are found in the center of the map. The general goal of identifying target market involves looking at selections of data where the concentration of the market is away from the center and in what direction. Figure 7 is a breakout of the data by the different groups.

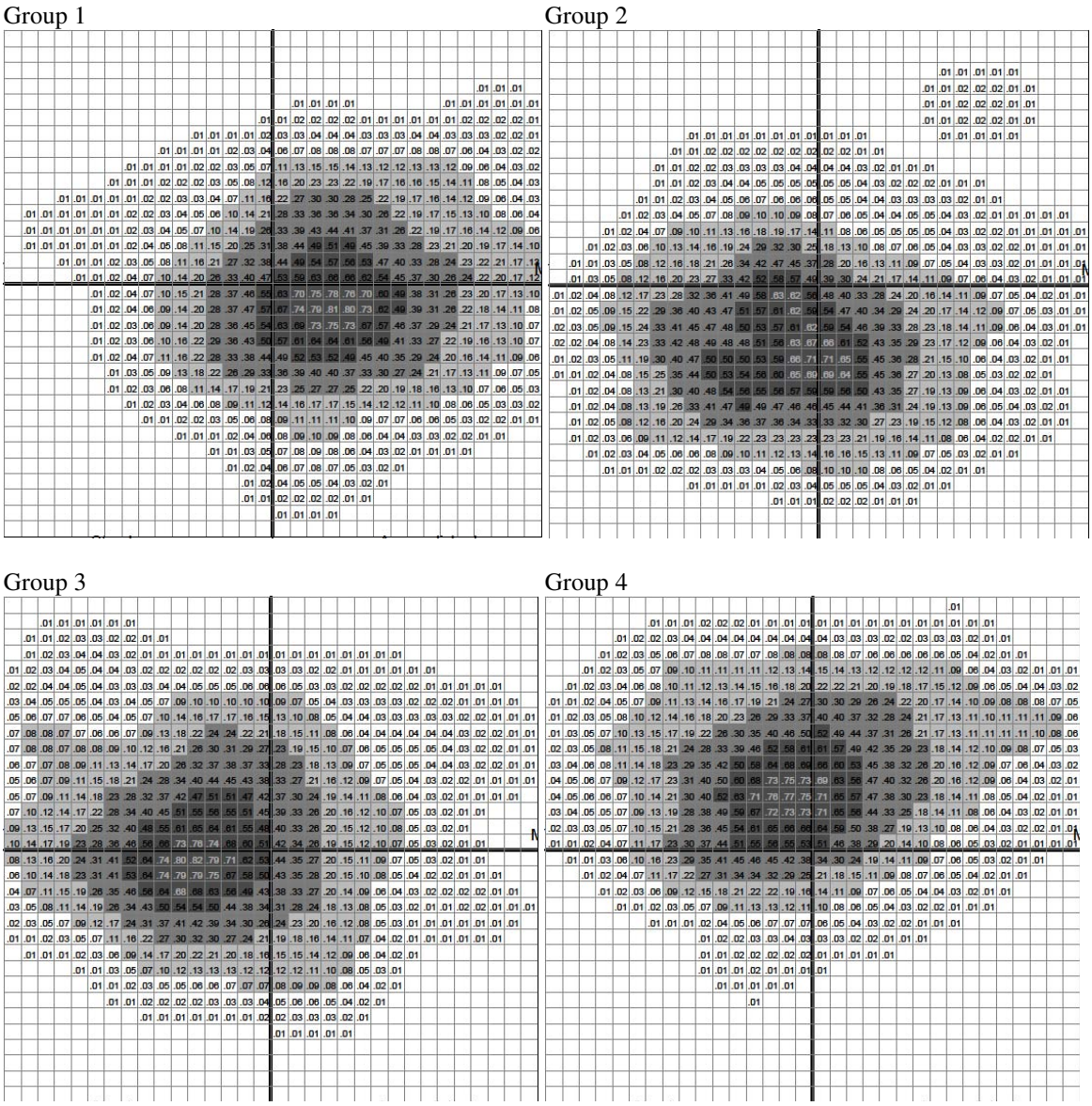


**Figure 7. Individual segment maps**

Once the group possessing the desired customer mindset is obtained, we select the target customers by focusing on the highest density cells. The target is obtained by selecting the fewest cells that comprise a large enough market to support the product concept. In Figure 7, the numbers reflect the density percentage of each cell, and the light-colored numbers show the chosen target, which is 10% of the group. Again, selecting this target is intended to ensure that an attitudinally homogeneous subset of the market is chosen. This is the critical decision-making point for target marketing in relation to the acceptance or rejection of a product proposal.

## CONCLUSION

Target marketing is increasingly necessary for an effective product development process. Defining target customers upfront provides an enormous competitive advantage. Successful target marketing requires consideration of consumer frame of mind input to make good product development decisions. Analyzing consumer opinion is more advanced than only considering demographic and socioeconomic concerns, and takes advantage of SAS' analytic strengths to address them. The capability to complete a target market analysis through a clearly defined process can greatly improve a company's chances of market success.

## ACKNOWLEDGMENTS

The authors thank Greg Brown for his statistical input and Ann Chen for her development of the SAS® Graphs used in this paper.

## CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Will Neafsey
Manager, Brand DNA and Consumer Segmentation
Ford Motor Company
1 Executive Plaza
Dearborn, Michigan 48124
Phone: 313-390-9820
E-mail: wneafsey@ford.com
www.ford.com

Ted Lang
Technical Account Manager
SAS
2000 Town Center
Southfield, Michigan 48075
Phone: 248-354-3008 x51234
E-mail: ted.lang@sas.com
www.sas.com