

---

## Geodemographics: A Collection of Articles

Here is a collection of articles on geodemographics. See also Michael Weiss' two books: *The Clustering of America* and *The Clustered World*.

Articles compiled by:

- Stephanie Cota, [www.stephaniecota.com](http://www.stephaniecota.com)
- Andreas Ramos [www.andreas.com](http://www.andreas.com)

---

## Birds of a Feather

by Susan Mitchell

[http://www.marketingtools.com/Publications/AD/95\\_AD/9502\\_AD/9502AF03.htm](http://www.marketingtools.com/Publications/AD/95_AD/9502_AD/9502AF03.htm)

American Demographics February 1995

Marketers use geodemographic "cluster systems" to reach new customers, choose new business locations, target direct mail, and do other tasks. Now the major providers have recently revised their cluster systems to include 1990 census data. Here is an overview of the latest in clustering and some advice for customers who are buying a cluster system. Consumers want the best value for their money. So do marketers. That is why products that make marketing more efficient is in great demand. Helping companies identify and reach their best prospects is a booming business. One of the most sophisticated tools for capturing customers is the geodemographic segmentation system.

Products like Claritas's PRIZM, Strategic Mapping's ClusterPLUS 2000, NDS/Equifax's MicroVision, and CACI's ACORN use data from the decennial census and other sources to separate the nation's neighborhoods into similar groups known as clusters.

"Cluster systems" are based on the premise that birds of a feather tend to flock together. Look at your own neighborhood. The homes and cars are probably of similar size and value. If you could look inside the mailboxes and cupboards, you would probably find many of the same magazines and cereals.

Some cluster systems use catchy names that try to capture the essence of each segment, such as PRIZM's "Blue Blood Estates." Others are more plainly descriptive, such as the "Urban New Families, New Homes" segment of ClusterPLUS 2000. However, the names are not important; it is the information inside that counts.

The idea behind all geodemographic cluster systems is the same. Each system divides neighborhoods into groups based on similarities in income, education, and household type, as well as attitudes and product preferences. But each of the four major cluster systems is dynamic and changeable, and the 1990 census gave them an enormous infusion of new data. Two of the players used census data to completely overhaul their systems last year, creating new sets of clusters organized in new ways.

To make matters even more complicated, two direct-marketing companies have recently introduced cluster systems that don't start with the census--Metromail's DNA and Trans Union's SOLO. This business can get confusing, but there are ways to decide which system to buy and how to put it to good use.

Completely new clusters sprang up after vendors incorporated 1990 census data. The changes reflect major shifts in U.S. society. Claritas increased the number of PRIZM segments from 40 to 62, and Strategic Mapping changed the number of ClusterPLUS 2000 segments from 47 to 60.

Racial and ethnic diversity is a key point of difference between cluster systems of the 1980s and 1990s. Five of the clusters in the new PRIZM system are Hispanic. "The emergence of larger numbers of Latinos across a variety of neighborhood types shows that this group is not monolithic in its settlement patterns, lifestyles, and product preferences," says Michael Reinemer, former director of communications for Claritas. In addition, while there are no specifically Asian clusters, the company found above-average concentrations of Asian households in the most upscale clusters.

Degree of urbanity is one of the factors researchers use when constructing clusters. Both Claritas and Strategic Mapping found that in the 1990s, settlement patterns in the U.S. have a greater degree of variation than simply urban and rural. This is reflected in the inclusion of "edge cities" or "second cities" in urbanity scales. "The second city is a distinctive

pattern," says Dave Miller of Claritas. "It represents a different lifestyle of city living that is not urban living." The inhabitants of Nashville, Chicago, and the "edge city" of Walnut Creek, California, are all city dwellers--but life in the three cities differs substantially.

The disadvantage of changes in the systems is that the old and new ones are not comparable. Even if segment names don't change, differences in how they are defined may be great enough to prevent direct comparison. This is one of the reasons Equifax National Decision Systems chose not to overhaul its MicroVision system. "We infused the new [census] data into the existing structure so our clients could have continuity with the old system," says senior product manager Diane Zablit.

Merging new data into an existing or updated cluster system is no mean feat. Geodemographic segmentation systems start with millions of raw statistics. They divide the nation's households into groups based on similarities, much as biologists divide living things into orders, families, and so on. In fact, geodemographic systems trace their ancestry to two statistical parents. The first is a classification system used in the biological sciences, and the second is the geographic divisions (zip codes) created by the U.S. Postal Service.

It takes fancy statistical methods to tease out the patterns that link millions of households. Fortunately for marketers, biologists had already worked on this problem while trying to categorize species of plants and animals. The statistical technique known as multivariate regression analysis was first applied to census and marketing survey data in 1970 by Jonathan Robbin, now a consultant in Washington, D.C. When matched to census geography and zip codes, the resulting product became the first geodemographic segmentation system, Claritas's PRIZM. Just as a species is "a class of individuals having common attributes and designated by a common name," a cluster is a class of households with common demographic and lifestyle characteristics, designated by a label.

Many decisions go into building geodemographic segmentation system. The number of clusters must be large enough to provide substantial distinctions between groups, but small enough to be manageable. A system that differentiates on the basis of every characteristic of Americans would produce more than 260 million U.S. clusters, each containing a single individual. On the other hand, a system based solely on the characteristic "resides in the U.S." would yield a single cluster with more than 260 million people. "That's the challenge of clustering," says Joan Hammel, spokeswoman for Strategic Mapping. "It's got to be homogeneous, yet diverse enough so you don't have one-size-fits-all clusters."

Segmentation schemes are only as good as the data that go into them. The power of cluster systems lies in their ability to predict consumer behavior, and this power depends on incorporating data on lifestyle choices, media use, and purchase behavior into the basic demographic mix.

The ACORN system from CACI Marketing Systems divides over 220,000 census block groups into 40 clusters based on 61 characteristics. The characteristics range from general (income, age, and household type) to highly specific (type of cars owned, home value, preferred radio formats). Claritas's PRIZM system incorporates data from "automobile registrations, magazine subscription lists, consumer product-usage surveys," and other sources, says Michael Reinemer.

Most systems begin with census block groups that contain, on average, about 340 households. Block groups are the basic unit of geography because "they delineate actual neighborhoods, using natural boundaries like major streets," says Reinemer. Regardless of the level of geography used, however, researchers can append cluster codes to virtually any list or adapt them to any type of geography, from zip codes to television markets and sales territories.

In the 1980s, the major cluster systems began selling custom segmentation schemes for specific industries. A segmentation system for the auto industry might divide the nation into neighborhoods dominated by owners of imports, sport-utility vehicles, and luxury cars. ClusterPLUS 2000 uses over 450 "atomic clusters, or minute but homogeneous segments of the 60 larger clusters that can be used to develop industry-specific or vertical marketing clustering," says Hammel.

This additional flexibility appeals to many users. "Before, you had to buy the canned software," says Mark Darling, director of strategic planning at American Isuzu Motors. "If you can customize the information, it's more meaningful."

Clusters can also be grouped into larger divisions of clusters, just as species can be grouped into classes or orders. Claritas divides 62 PRIZM clusters into 15 broader groups based on degree of urbanization and socioeconomic status. The "Urban Cores" group, for example, comprises "Single City Blues," "Hispanic Mix," and "Inner Cities." These three clusters are all "multiracial, multilingual communities of dense, rented row homes and high-rise apartments" with high shares of singles, solo parents with preschool children, and perennial unemployment. Similarly, Strategic Mapping divides the 60 ClusterPLUS 2000 segments into 6 "urbanity" groups (such as "suburban"), 11 "multifactor" groups (such as "average income, seniors, few children"), and 10 socioeconomic groups.

Marketers use cluster systems to find new customers, locate sites for stores, buy advertising, target direct mail, and develop new products. If a direct-mail campaign gets a strong response from one zip code, for example, a cluster system can locate other zip codes with similar characteristics.

One of the most powerful features of geodemographic segmentation systems is their ability to find customers. "There are two primary ways people use segmentation systems," says Strategic Mapping's Hammel. "Some people know who their customers are and want to know where to find them. Others need to figure out who their customers are, and then go find them."

Cluster systems can reveal niches of potential customers in unlikely places. Equally important, they can show that some favorite groups are in reality poor prospects. "There are some surprises that can let the air out of your notion of who your customers are," says Jim Keryan, GTE's staff administrator of market assessment. Darling of Isuzu used ClusterPLUS 2000 to find the target market for the launch of the 1989 Amigo convertible. What he found went against his expectations. "Chicago wasn't the first place we imagined a soft-top car would sell," says Darling. "But the system predicted it. Sure enough, the top two dealers in sales were in Chicago."

Clusters tell you a lot about customers. Do they travel a lot? Do they read health and fitness magazines? The profile that emerges enables companies to design marketing campaigns that address customers' lifestyles.

Since cluster systems include information about media use, they can also help place media dollars. "You can pick a particular newspaper zone, a certain TV daypart, or a given cable program or system that will give you the best penetration into your target neighborhoods," says Claritas's Reinemer. Buick used the PRIZM system to decide where to buy billboard advertising, and Isuzu found that lifestyle magazines were a better buy than news weeklies.

Cluster systems are also useful for site selection. If you want to locate a store in Washington, D.C., you could physically drive through its neighborhoods to try to get a feel for who lives in them. Or you could take a "virtual" drive through a cluster map. Coming in from the eastern suburbs on a PRIZM-coded map, you pass through "Second City Elite," "Winner's Circle," and then "Kids & Cul-de-Sacs." Once inside the Beltway, you pass through "Money & Brains," "Inner Cities," and "Bohemian Mix," all without ever leaving the office.

The lifestyle data in cluster systems help businesses plan product designs or modifications. "One of the things we know, for example, is that sport-utility-vehicle owners are likely to ski. We want the vehicle design to accommodate the consumer," says Darling of Isuzu.

Clusters sell advertising, too. Cable Networks, Inc. sells time on cable television systems. It uses cluster system maps to "show our customers where the big purchasers of certain items are located within cable system geographies," says director of marketing research Laura James. Advertisers are far more impressed by a cluster-coded map than they would be with a page of basic statistics, she says.

Clusters tell us a lot about where different kinds of Americans live. Some clusters are scattered across the country, while others are concentrated in just a few areas. But even a seemingly homogeneous area will sometimes show surprising variety when delineated by clusters. On a PRIZM-coded map of downtown Jackson, Mississippi, for example, some areas are predominantly populated by "Southside City" residents. This cluster is dominated by young and old African Americans who are employed primarily in low-paying blue-collar service jobs. They have little education, rent apartments, read sports and fashion magazines, and eat instant grits. But in the middle of this low-income area are a couple of "Towns & Gowns" neighborhoods. People in this cluster also rent apartments, but they are college graduates with better-paying white-collar service jobs. They like to ski, read beauty and fitness magazines, and use ATM cards.

Some clusters have similar socioeconomic profiles but are distinctly different in their habits. The PRIZM clusters "Executive Suites" and "Pools & Patios," for example, are both composed of affluent, well-educated professionals. But "Executive Suites" households listen to jazz radio and read business magazines. "Pools & Patios" people listen to news radio and read epicurean and leisure magazines. Reaching these two affluent groups obviously requires different strategies.

Not every business needs the full resources of a major geodemographic segmentation system. To effectively use geodemographic segmentation systems, businesses should have a minimum of 2,000 customers. The systems work best with a minimum of 5,000 to 10,000 customers, says Hammel of Strategic Mapping.

Before buying a cluster system, make sure you understand your own company's resources. "Look at what you already have available internally," says GTE's Jim Keryan. "For example, you need to understand how your information and billing systems work together. Then dig through your own database. This is the information you will eventually merge with a segmentation system." You also need to clearly understand how you will use a segmentation system.

Next, study the different systems to make sure they can do what you want them to. The systems differ in the data they use, the number of segments they offer, the base level of geography on which the system is built, and the way they describe or name segments--among other things. One system isn't necessarily better than another simply because it has more clusters or catchier names.

In the end, your decision may depend on which system you best understand, or which system makes you feel most comfortable. But whatever your choice, be prepared to put enough time and resources into your system to get the most out of it. As Patrick Harrison, marketing line manager at Buick, puts it: "Geodemographic segmentation systems provide the science of marketing. You provide the art."

## Taking It Further

The four major general-purpose neighborhood-based cluster systems are: ACORN, from CACI Marketing Systems, 1100 North Glebe Road, Arlington, VA 22201; telephone (800) 292-2224; ClusterPLUS 2000, from Strategic Mapping, Inc., 70 Seaview Avenue, Stamford, CT 06192-0058; telephone (203) 353-7500; MicroVision, from Equifax National

Decision Systems, 5375 Mira Sorrento Place, Suite 400, San Diego, CA 92121; telephone (800) 866-6510; and PRIZM: Next Generation, from Claritas, 201 North Union Street, Alexandria, VA 22314; telephone (800) 284-4868. See also "Rx for Cluster Headaches" in the premier issue (March/April 1994) of Marketing Tools magazine; for subscription information, call (800) 828-1133.

---

# Introduction to Geodemography: PRIZM, Claritas, and Clusters

by R. Bruce Carroll

<http://www.bpmsi.com/PSYTE2.HTM>

Geodemographic neighborhood classification systems have been around since the mid-1930s but widespread commercial applications really only began in the late '70s and early '80s, principally with the launch of the PRIZM system by the Claritas Corporation in the United States. Since that time, cluster systems have been adopted by most major consumer marketers, including financial institutions, retailers and automotive manufacturers in North America, Europe and around the globe. Approximately 22 countries have already been "clustered." Cluster systems have become the established lingua franca of marketing.

Most marketers are familiar with the basic tenet of geodemographic neighborhood classification systems: People with similar cultural backgrounds, means and perspectives naturally gravitate toward one another - or - to form relatively homogeneous communities. (It's the old "birds of a feather flock together" phenomenon.) Once settled in, people naturally emulate their neighbors, adopt similar social values, tastes and expectations and, most important of all, share similar patterns of consumer behavior toward products, services, media and promotions. This behavior is the basis for the development of classification systems such as LIFESTYLES, PRIZM, CLUSTER PLUS and Compusearch's new PSYTE system, all of which classify neighborhoods and their households into clusters or groups of neighborhoods, based on their underlying socio-economic and demographic composition.

It's not uncommon for some people, on their first exposure to cluster systems, to debate the underlying homogeneity of neighborhoods and the resulting linkage to consumer behavior. "I am not like my neighbor" is a common response. I suppose the argument starts with a misunderstanding of what homogeneity means in the context of spatial demography. Because our marketing perspectives have long been focused on univariate demography, we tend to define homogeneity in a vertical context, expecting everybody living in a given neighborhood to be identical in order for clusters to work - i.e. every cluster should consist of all young families or upscale singles, all executives or hard hats, all rich or poor.

But these are univariate criteria, used over the past forty or so years to segment and target mass markets. Clusters, however, are a multivariate creation, designed to segment and target neighborhoods. Obviously, all residents in any given area, however small, are not identical. In an urban neighborhood, the older wealthy gentry may live a block away from the welfare recipient. In a rural neighborhood, the gentleman farmer may dwell amongst hardscrabble farm workers. Such are the realities of community structure, which in turn provide the building blocks of any cluster system. Homogeneity, as used in geodemographic cluster technology, simply means that all neighborhoods within a given cluster will share highly similar neighborhood lifestyles and predictable consumer behavior.

Theoretical debates aside, cluster systems have already proven themselves where it counts - in the marketplace. At a conservative estimate, more than 15,000 companies in the United States and Canada alone used clusters as part of their marketing information mix last year. This kind of acceptance doesn't happen unless the effectiveness of using clusters can be measured and tracked, season after season, year after year. Marketers simply don't pay for something that doesn't work. There are many reasons behind the enduring and even increasing popularity of generic geodemographic cluster systems. Here is my personal list of the major contributions I believe geodemographic clusters have made to modern marketing.

## Discriminating Power

While it's true that cluster systems often cannot match the discrimination produced by highly customized statistical solutions - which use Chaid and other forms of regression on a specific data set with good unit-record type data - they are certainly superior to most univariate demographic measures such as age, sex, income, etc. These simplistic measures are still favored by too many marketers and media but they are in fact obsolete in describing modern consumer behavior. Moreover, cluster systems can capture the different "franchises" or behavioral components of a product's user base whereas demographic measures tend to homogenize consumer profiles into a simplistic caricature. For example, Women, 18 to 49, is a profile that could be applied to thousands and thousands of products.

To illustrate this point, take a look at the [cluster profile of the Nissan Sentra](#). The clusters are ranked in terms of their average household income, that is, Cluster 1 is the wealthiest, Cluster 60, the poorest. The index on the left shows the sales penetration of the Sentra for each cluster, compared to the national sales or buy rate. An index of 200 means that households in that cluster buy at twice the Canadian average. The variance in buy-rate across the clusters certainly makes

the point about the discriminating power of clusters. But the even more important thing to note here is the complexity of the profile. There is a decidedly up-market bias in the buyer profile but look farther down and you will see that there is some action in the down market areas as well, especially in the Quebec clusters. In looking at, literally, thousands of cluster profiles, I've seen very few products where more than one "franchise" has not been identified. Simple demographic measures, in contrast, can't capture the bi-modal or tri-modal consumer profiles that often exist within a product's consumer profile.

This profile and others like it demonstrate the evident link between social structure and consumer behavior \ emdash which is the basic commercial promise of geodemography. More than anything else, however, these profiles illustrate the extraordinary diagnostic, predictive and motive power of PSYTE, Compusearch's new state-of-the-art cluster system. (See p. 12 for a full description.)

## **Medium of Integration**

This is clustering technology's marketing forte. You build a consumer target by profiling your own customer files or you can use a profile of your particular product or service, using any number of syndicated databases such as PMB. You can then compare or correlate that profile to more than 50 databases that have been coded with PSYTE. After you have a good idea which cluster targets you want, you can then rank TV programs and/or dayparts, target out-of-home advertising, select names from a mailing list, rank telephone exchanges and postal walks, and target retail distribution, all using the same target definition.

This is what I mean by describing clusters as a medium of integration. And it is one of the principal advantages that generic cluster systems have over customized segmentation systems. There is no need to change the description of your target simply because the marketing medium and the select options have changed. Let me emphasize this crucial point: You can take the same cluster target you used for placing outdoor advertising and use it to target television, radio and newspaper buys, or to order names from a mailing list or to select retail sites for selective promo drops. A former partner of mine used to call this "cluster bombing" and it's a pretty impressive process when executed properly.

There is really nothing new about market segmentation. Everyone knows that different kinds of people consume different products, and marketers have been segmenting for years. But the real advantage of cluster segmentation is not in segmenting per se but in being able to hit the target, once defined, and in being able to concentrate all elements of the marketing mix against this target.

## **Accountability**

The results of cluster targeting can be easily measured. Remember, the basic unit for geodemographic targeting is every postal code in the country, which has been assigned to one of the 60 clusters. To see if cluster targeting worked, a client simply has to track his sales, shipments, subscriptions or whatever by postal code, summarize them up to each of the sixty clusters and see if sales have, in fact, increased in the targeted clusters.

Even more accurately, marketers can determine whether sales increased more in those clusters than for the market overall - or, in a declining market, whether sales declined less in targeted clusters than for the market overall. As marketing dollars come under ever closer budget scrutiny, marketers will embrace anything that can reliably measure the success or failure of a program.

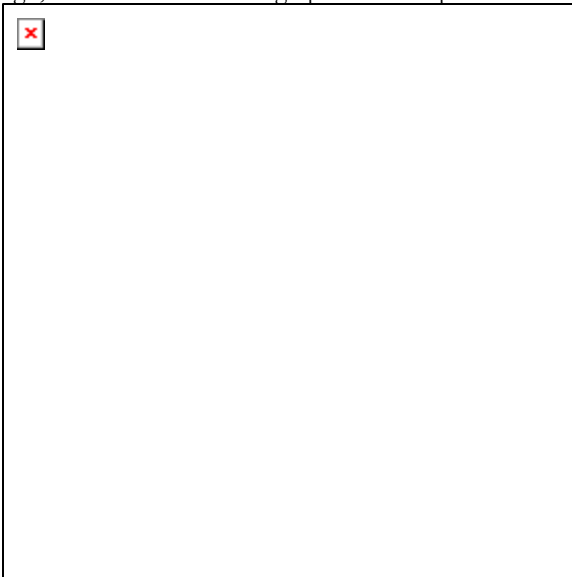
## **Longitudinal/Time Series Analysis**

This is perhaps one of the least appreciated and underutilized benefits of using cluster segmentation to analyze consumer behavior. PSYTE is what I call a fixed segmentation system; it does not change because the database it is being applied to changes. This means that marketers can analyze their sales going back three, four, five years, whatever, along with the changing structure of their consumer franchise over that time period and see how their market has changed by cluster. In short, PSYTE delivers the ability to track market share for groups of products or individual products on a cluster-by-cluster basis, both at the national and the individual market level, month over month, year over year.

Most marketers know their sales at a national and market level and even sales by branch or retailer within a market. But they usually do not know the demographic constituency of those sales on a market-by-market basis or within a retailer's trading area. And equally important, they do not know the evolution of those changes at a small area level over time. Imagine being able to answer or at least consider these questions:



- When we increased the price of our product two years ago, did we alter the demographics of our profile? If so, was the change uniform across all markets?



- When we stopped using rebates one year ago, did we alter our consumer profile?



- What was the effect on our customer profile when that major new American retailer entered the market six months ago?

Or imagine you are a national marketer and you have found over years that your success was determined by how well you had penetrated the old, suburban gentry market (e.g. PSYTE Groups S2). These clusters represented your core constituency. If you lost customers here, you were lost. Wouldn't you want your MIS department to give you a report of monthly sales in these clusters, not only nationally but also on a market-by-market basis, to use as a barometer of how business is doing nationally and locally?

Here's an actual example of what I'm talking about. A well-known marketing vice-president of a US.-based automotive company insisted on receiving reports of his company sales in two PRIZM clusters - Blue Blood Estates and Blue Chip Blues - in his top 20 markets every month. If sales started to go down in either of these two clusters, which he considered to be "leading indicator" clusters, he ordered an increase in local ad expenditures. In effect, he was using cluster analysis to build or protect "micro share" in order to maintain his national share. I'm aware of a Canadian packaged goods executive doing something similar by tracking his market share in selected clusters on Nielsen's NEDS panel.

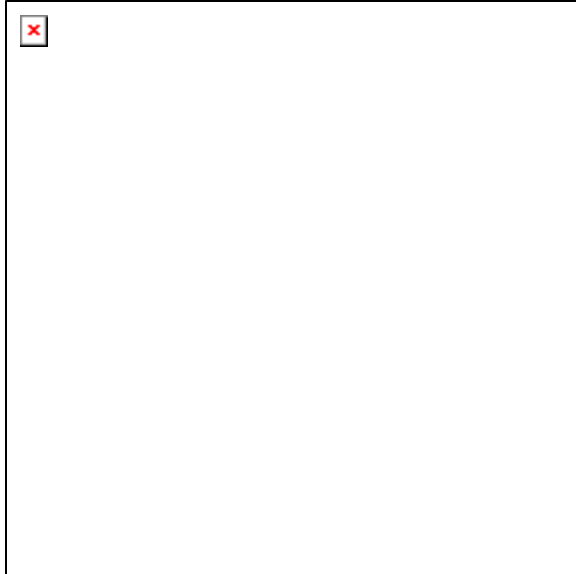
## Addressable, Mappable Targets

We often use the phrase "see what you're saying" to mean we understand something. The beauty of a cluster-based targeting strategy is that it can be found on the ground \emdash it can be mapped! Using a desk-top GIS mapping system, you can illustrate targets at any level, right down to individual postal walks, proprietary distribution/sales zones, grocery store trade areas, whatever. You can map primary and secondary targets. You can map clusters that show increasing and/or declining sales; map response rates from a coupon drop or mail campaign.

Your data can be visualized, which means it can be used - and used more easily by more people in the organization. It takes information out of the hands of the few and puts it into the hands of the many. Clusters are "executive friendly," too. Even company presidents not well known for their facility with statistics and market research can "see what you're saying" when you present them with a cluster profile and/or map of your customers. Clusters are just plain easy to understand.

As I said earlier, this is my personal list but it is by no means a complete summary of the marketing applications of geodemographic segmentation. Clusters can also be used as a variable in customized direct response and site modeling, in positioning and targeting new products, for creative message targeting, and for projecting future market penetration and share.

## U.S. Suburban Press and the PRIZM Cluster System



The beauty of USSPI and PRIZM is flexibility. The flexibility to target your best prospects in a way never before possible by newspaper or any combination of newspapers. Give us your demographics or specific ZIP codes of interest. Or, if you are a PRIZM licensee give us your target clusters, and we will help you create a buy tailored to meet your specific needs.

Conceived and pioneered by CLARITAS in 1974, PRIZM is recognized as the premier in geo-demographic consumer market segmentation. The basis of the system relies on assigning every U.S. ZIP code into one of sixty-two unique neighborhood cluster types according to their similarities over hundreds of precise census measures. Having identified a ZIP code's PRIZM cluster type one can then predict consumer behavior, lifestyle pursuits, etc. for that given area via interlock with primary syndications such as Simmons Market Research Bureau (SMRB).

U. S. Suburban Press has taken the Sixty-two PRIZM Clusters and has created three distinct cluster groupings which advertisers can now target via USSPI's network of suburban newspapers in the top 50 U.S. metro markets. These targets are the Affluentials, Greenbelt Families , and Middle America.

Target	% Comp. U.S. HH's	USSPI Circulation	% Comp. USSPI Circulation	Index
Affluentials	15.47	4,400,000	31.85	206
Greenbelt Families	11.74	3,500,000	25.04	213
Middle America	25.75	4,200,000	29.88	116
Total	52.95	12,100,000	86.77	164
All Others	47.05	1,800,000	13.23	28

### TARGET #1- Affluentials

Cluster Number	Cluster Nickname	% u.s. HH'S	% USSPI Circ.	Index
3	Executive Suites	1.31	5.43	415
2	Winner's Circle	1.87	5.37	287
4	Pools & Patios	1.81	4.21	233
5	Kids & Cul-de-Sacs	3.01	6.34	211
14	Country Squires	1.16	2.17	187
7	Money & Brains	1.05	1.78	170
1	Blue Blood Estates	.78	1.11	143

15	God's Country	2.81	3.70	132
11	Second City Elite	1.67	1.74	104
		15.47	31.85	206

This target is composed of the top socio-economic clusters which include high-end education and incomes, executive white collar occupations, expensive owner-occupied housing, and conspicuous consumption levels for many products, goods and services. People in this target range from elite suburban families to yuppies to double income families with mostly grown children in expensive houses to sophisticated consumers of adult luxuries in townhouses near universities. This target has enormous buying power.

## Demographic & Lifestyle Highlights

	Index		Index
Household Income \$100,000+	308	Gold/Premium Credit Card	149
Home Valued \$500,000+	278	Have a Passport	168
College Graduates	187	Drink Imported Wine	170
Upper Level White-Collar	153	Stay at Hilton on Vacation	201
Visited Europe Last 3 Years	179	Belong to a Country Club	152
Own a Mercedes	197	Use Coupons from Newspapers	110
3+ Stock Transactions Last Year	175		

## Target #2-Greenbelt Families

CLUSTER NUMBER	CLUSTER NICKNAME	% U.S. HH's	% USSPI CIRC.	INDEX
20	Boomers & Babies	1.29	4.05	315
19	New Empty Nests	1.80	4.81	267
12	Upward Bound	2.04	5.39	264
8	Young Literati	1.03	2.41	235
18	Young Influentials	1.19	2.43	205
9	American Dreams	1.39	2.13	153
16	Big Fish, Small Pond	2.03	2.85	141
17	Greenbelt Families	.98	.97	99
		11.74	25.04	213

This group occupies the comfortable suburban and boom town areas dedicated to family raising, good schools, scouting, soccer games, and home maintenance. This is a major packaged goods and apparel target with larger families, above average incomes and home values, dual wage earners, and owner-occupied single unit dwellings. Most families here have achieved the traditional "American Dream."

## Demographic & Lifestyle Highlights

	Index		Index
Household Income \$100,000+	143	Convertible Car Bought New	141
Home Value \$150,000-\$250,000	150	Interest Checking	115
Attended College (1-3 years)	122	Kitchen Improvements Last Year	115
HH's with Children (Married Parents)	117	Tax Sheltered Annuities	131
Own a Personal Computer	132	Use Espresso/Cappuccino	127
First Mortgage Loan	134	Use Coupons from Newspapers	111

## TARGET #3-Middle America

CLUSTER NUMBER	CLUSTER NICKNAME	% U.S. HH's	% USSPI CIRC.	INDEX
21	Suburban Sprawl	1.80	3.77	210
27	Urban Achievers	1.53	3.15	205
22	Blue-Chip Blues	2.05	3.52	172
23	Upstarts & Seniors	1.21	1.79	148
28	Big City Blend	.95	1.41	148
26	Gray Collars	2.05	2.88	141
25	Mobility Blues	1.58	2.15	136
13	Gray Power	2.09	2.60	124
30	Mid-City Mix	1.20	1.10	92
31	Latino America	1.24	1.03	83
34	Starter Families	1.62	1.21	74
38	Middle America	1.28	.93	73
29	Old Yankee Rows	1.35	.97	72
10	Bohemian Mix	1.61	1.10	68
32	Middleburg Managers	1.54	.99	64
24	New Beginnings	1.44	.70	48
33	Boomtown Singles	1.21	.58	48
		25.75	29.88	116

This target is the collection of mid and upper mid-suburban and older, denser neighborhoods with mixed age groups. There are relatively more foreign born residents here than in other suburban areas. Suburban neighborhoods in this target were built in the 50's. Consequently, these well-paid blue and white collar residents have good discretionary incomes due to high stable employment, dual incomes, and homes with small or no mortgages. Living is easy in this target.

## Demographic & Lifestyle Highlights

	Index		Index
Household Income \$100,000+	125	Member of Frequent Flyer program	120
Home Value \$250,000-\$500,000	118	Bought Hardcover Book	115
Took 2+ Foreign Trips	123	Stock Valued \$10,000+	121
Own a Sport/Utility Vehicle	121	Drink Imported Beer	115
Household Owns a Modem	125	Department Store Credit Card	113
Own Burglar Alarm	114	Use Coupons from Newspapers	102

---

## Focus on the Household

by Jock Bickert

November/December 1995

For centuries, the human race has tried to make sense of the millions of behavioral variants of its members. Why on earth do friends, lovers, family members, business acquaintances, and enemies behave the way they do? In the search for some consistency in both explaining and predicting behavior, people have turned to typologies for assistance. If you can put someone in a category whose members behave reliably, you've gone a long way toward making some sense of that person's behavior. Or so goes the reasoning.

Marketers, in their continual need to explain and predict consumer behavior, have turned to typologies, or segmentation systems, for assistance. Their prayers appeared to be answered in the 1970s with the advent of the first geodemographic systems: PRIZM from the Claritas Corporation and ACORN from CACI. These systems benefited from an increase in computing power that took advantage of the relatively new statistical process known as "cluster analysis."

In their seminal 1970 book, *Cluster Analysis*, co-authors Daniel Bailey and William Tryon had applied the technique to four decades of U.S. census data from the San Francisco Bay Area. They "clustered" the census tracts from the Bay Area, basing that cluster analysis on variables such as socioeconomic status, employment, condition of housing, etc. To their astonishment, when they looked at 40 years of varied election results within those clusters, they discovered that election behavior remained identical for those clusters of tracts over the four decades. That meant that even though new families moved in and out of those tracts, the aggregate political behavior of the tracts stayed the same. In other words, not only did birds of a feather flock together, but successive generations of those birds flocked in similar fashion.

The commercial geodemographic systems quickly became the staple of mainstream marketers seeking consistency in consumer behavior. These systems have proved successful in store location analyses, in directing media buys, and in locating pockets of new customers. But they have been arguably less successful in providing significant response "lift" in many direct-mail applications. Therefore, the direct-marketing industry has been bereft of marketing typologies, aside from limited survey-generated groups developed by agencies such as Stone & Adler, Grey Direct, and O&M Direct. Those systems shared a common limitation: They identified types of mail-order buyers, but gave the potential user no way to locate individual prospects who belonged in those desirable mail-order categories.

In short, while it was helpful to know where to find a specific flock, what direct marketers really needed was a way to get to know each bird on an individual basis. Because they are based on data from the Census Bureau, which is prohibited from releasing information about individuals, the neighborhood-level systems can only offer generalizations and suppositions.

But the census is no longer the only powerful repository of information about American households. The late 1980s and early 1990s saw the proliferation of large, data-rich, household-level databases in the direct-marketing industry. As a result, they have achieved nearly equal coverage and, in several instances, have assembled databases richer in relevant data than the census.

These private sector data also tend to be more reliable than the information available from the census. Because the census does not release information on individual households--only aggregate data for a given unit of geography (e.g., census tracts or block groups)--systems that use census data as building blocks can only estimate characteristics of a given household from the aggregate of that household's census geography. Therefore, if a block group is reported as having a median income of \$53,147, every household in that block group is assigned an income of \$53,147. In today's increasingly diverse environment, a household with an annual family income imputed to be \$53,147 may actually be as low as \$25,000 or as high as \$100,000.

Although household-based databases have existed for 10 to 15 years, the industry has been slow to derive segmentation typologies from those databases. Instead, segmentation processes have consisted of custom modeling in which individual data items were used to predict direct-response behavior. The models have been

successful in generating response lift, but because they are customized to a specific mailer as well as a specific mailing, they are not useful in predicting response across mailers, or even across mailings.

In the last several years, three major household-based segmentation systems have appeared within the direct-response industry. They are DNA, from Metromail and Fair Isaac; Niches, from The Polk Company; and Cohorts II, from Looking Glass, Inc., in concert with what used to be National Demographics & Lifestyles and is now NDL/The Polk Company.

The DNA system actually consists of two elements: DNA Demographic and DNA Lifestyle. Drawn from data on 77 million households, DNA Demographic is life stage-based, beginning with ten age groups. Each age group consists of 6 to 14 Cells. Each Cell is composed of households with similar demographic characteristics. Those 104 Cells have then been aggregated into 25 "Super Cells" in three broad age bands: 20-34, 35-54, and 55+.

DNA Lifestyle assigns households to 1 of 100 Cells based on lifestyle and behavioral information from Metromail's 25 million survey respondents in its BehaviorBank file. This classification ignores demographic characteristics and focuses instead on behavioral information. The system assigns households to one of 100 Cells, which also have been aggregated into 25 Super Cells. Users can combine DNA Lifestyle with DNA Demographic or with specific individual data.

Data from Mediamark Research have been appended to each Cell in order to enhance the clustering. Those data include magazine and newspaper readership, vehicle ownership, and product usage. Each Cell has been given a four-digit number, with the first two digits indicating the age range and the last two digits being an income indicator. For example, Cell 3001 indicates the most affluent group in the 30-to-34 age group.

The DNA system is definitely positioned as a tool for direct marketers, with its greatest touted strength being prospect identification in direct-marketing applications.

Niches, from The Polk Company, represents a marriage of household-based segmentation and geodemography. The system begins by identifying three dimensions that appear to differentiate individual households in the 80 million-name Polk database: (1) needs, as measured by life-cycle stages (critical definers being age of adults in the household, number of children, and children's ages); (2) buying power, as measured by wealth factors such as income, dwelling type, and homeownership; and (3) spending patterns, as measured by mail response frequency, credit-card usage, and new car and truck purchases.

The combination of those dimensions produced 108 SuperNiches (groups) which were then cluster analyzed to produce 26 Niches for broader applications. The Niches system employs clever nomenclature, capitalizing on the correspondence between number of Niches and letters of the alphabet. The designations A to Z indicate decreasing affluence. For example, the first five Niches (Already Affluent, Big Spender Parents, Cash-to-Carry, Diamonds-to-Go, and Easy Street) are the most affluent (i.e., incomes greater than \$75,000 a year).

The last seven Niches have annual incomes less than \$20,000, ending with the Young-at-Heart and Zero Mobility.

The descriptions of the groups have also been enhanced by overlaying the system with data from Mediamark, so users know that the Easy Street category (affluent families over age 65) have a greater-than-average propensity to own home computers and engage in credit-card spending.

Niches has been used as a shortcut to expensive custom segmentation techniques in direct-mail prospect identification. In enhancing customer files with Niches categories, Polk's TotalList Network database is used on a match basis. For unmatched records, GeoNiches is used to fill in by applying a geodemographic extension, using either a census or postal geographic unit.

Although performance results have not been made public, Polk officials point to a number of case studies in the publishing, financial services, member recruitment, and travel industries. In those instances, Niches has outperformed traditional geodemographic systems and has shown only slightly less power in head-to-head comparisons with harder-to-use and more costly regression techniques.

A recent entry in the household segmentation arena has been Cohorts II, a consumer typology developed by Looking Glass, Inc. using the household-level demographic and lifestyle data in the 35-million household Lifestyle Selector database. Unlike either DNA or Niches, Cohorts uses only self-reported data rather than a combination of self-reporting and imputation from other sources.

Seven demographic variables (gender, marital status, income, occupation, home ownership, and presence and ages of children) and 75 activities and interests (skiing, golf, foreign travel, gourmet cooking, political activity,

etc.) were cluster analyzed. This resulted in 27 homogeneous clusters: 11 married groups, 8 groups of single women, and 8 groups of single men.

The initial analysis revealed that not all households can be conveniently assigned to a cohesive cluster. There are grandmothers who ride motorcycles, male hunters who crochet, and sexagenarians who are raising second families. These people defy categorization and have been lumped into a potpourri group known as the "Omegas." Nearly 9 percent of all U.S. households are Omegas.

The development process also uncovered a correlation between cluster membership and given names. In the 35 million-name database, there were many names that appeared with unusual frequency in only one cluster. For that reason, all of the clusters were given high-indexing first names, resulting in titles like "Jules & Roz" (affluent and physically active urbanites with children), "Denise" (single mothers on a tight budget), and "Elmer" (very sedentary older men). Even the unclassifiable Omegas had a distinguishing characteristic: that group contained all of the classic Greek and Roman mythological names, like Ulysses, Aphrodite, Hercules, and Apollo.

The Cohorts II typology is used to (1) identify true, existing market segments; (2) locate underpenetrated but potentially profitable market segments; (3) evaluate market potential in small areas of geography; (4) identify specific prospect households within targeted segments using The Lifestyle Selector; and (5) develop appropriate advertising and marketing communications.

For decades, the mailing list subsegment of the direct-marketing industry has been the most pragmatic of businesses. Armed with the marketing accountability of measurable response rates, many direct-marketers have been relatively uninterested in seeking explanations for direct mail behavior. The statement: "Just tell me what works; don't bother me with why it works," has been industry gospel for years. However, as economic pressures and mountainous direct-mail volumes focus attention on eking out marginally higher response rates, direct marketers are admitting they need to mail "smarter."

But mailing smarter means understanding customers and prospects better. Sophisticated response models may lift response rates, but they rarely provide insights into why direct-mail recipients behave the way they do. Segmentation systems offer the promise of understanding. Tryon and Bailey reasoned that typologies are desirable because "since a particular type includes many individuals better understood than if no such cumulative information were available." In other words, knowledge of the many provides insights into the one. If mailers were able to pinpoint the mail order behavior of any cluster or segment--be they 3502, Big Spending Parents, or Buddy & Carole--that behavior should be consistent across mailers, as well as across mailings. Household-based segmentation systems should also provide the direct-marketing industry with the solution to another of its historical shortcomings. Lester Wunderman, in a keynote address to the Direct Marketing Association's Fall Conference in 1993, chided the industry for not exploiting one of its strengths. "We (direct marketers) segregate customer groups or database groups to whom we address the same message," he observed. "That's not individualized or personalized marketing. What that is is mini-mass marketing." Cohesive market typologies should allow marketers to craft messages and communications packages that are tailored to the idiosyncrasies of each segment. Marketers can speak the language of 3502 or Buddy & Carole and communicate with them differently than if they were talking to 4504 or Jules & Roz.

Household systems do not come without blemishes. One is the problem of "decay," where household data are not updated when significant lifestage changes occur. Some household systems avoid that problem by using only records of a certain age; e.g., no older than 18 months. As older names drop off one end, the database is replenished by new names coming in.

Another drawback is the potential for inaccurate data. That, obviously, is a problem with any database, particularly those that rely heavily on imputation from secondary sources. Self-reported data are not free from bugs, of course; individuals can always misstate vital information about themselves.

Perhaps the biggest bugaboo of the new systems is the failure to match 100 percent of a user's records. No system provides total coverage, just as the U.S. census fails to reach all households. And lack of total coverage certainly does not diminish the explanatory power of such household systems--provided, of course, that their coverage is unbiased.

Arguments will continue to be made for the merits of both types of systems. The resolution will be found in the marketing laboratory, as practitioners experiment to see which system best meets their individual needs.

More Info:

Cohorts II Looking Glass (303) 893-8600; contact: Jo Moniak  
DNA Metromail (800) 523-7022  
Niches The Polk Company (800) 635-5522

---

## **Additional Resources**

- Excellent website with more information at [www.claritas.com](http://www.claritas.com)
- Geodemographics Database at <http://www.geodemographics.org.uk>
- Cohorts, at [www.cohorts.com](http://www.cohorts.com) and [http://cohorts.com/meet\\_the\\_cohorts.html](http://cohorts.com/meet_the_cohorts.html)
- Tools at <http://www.claritas.com/demographics.htm>
- American Demographics. Excellent search engine <http://www.demographics.com/search>
- Demographics for Christ, <http://www.bgco.org/SST/P&R/usingdem.htm>
- Concepts of Sociology and Clustering, <http://www.du.edu/~mwidhofs/page5.html>