

# Frequently Asked Questions (FAQ)

## ***What is a cluster?***

The assignment to small areas of geography, such as the census block group, to one of forty or more “lifestyle codes” established through the application of multivariate techniques to detailed demographic composition. The “clusters” are an impressive consumer segmentation and predictive measure.

## ***How long have clusters been used?***

United States businesses have benefited from multivariate clustering techniques for market segmentation since the seventies – however, more widespread use of clustering occurred during the 1980’s.

## ***What is the difference between the clustering systems?***

Lifestyle clustering systems depend on census data (both geographic and decennial demographic information) as the fundamental metric(s) for small area evaluation. Yet, differing cluster systems go beyond “just” census data – both in development and in year-to-year updating. Today, clustering databases utilize national list data, local community government statistics – as well as, in some cases, aggregate financial purchasing behavior. Another important difference can be found by examining the variety of techniques used to perform the calculations.

## ***Why are there 60 clusters?***

While all commercially available clustering systems may seem similar because their underlying fundamental data sources are the same, all systems employ a variety of unique data inputs. A variety of data combined with numerous statistical processes and methodologies used to calculate “homogeneous groupings” add to the mystique of clustering.

Simply stated – the multivariate distributions, contained within each geographic unit (evaluated by averages and means) are carefully measured and identified based on data input and modeling technique. This process is often described as both a science AND an art resulting in a varying number of classifications – dependent primarily upon the “artist” and the tools and data available.

Keep in mind that the statisticians that create these systems for marketing purposes have two main tasks in mind when they are formulating cluster

systems: to achieve groupings that are fairly homogeneous while maintaining groupings large enough to measure and target efficiently. One statistician may find that 50 groupings show enough discrimination when used with certain data sources, another may choose and 80 or even 100 cluster solution.

### ***My company uses another system - can I (easily) convert to Mosaic?***

Yes, all cluster lifestyle codes are assigned to levels of geography (either postal or census – or both) and can be easily assigned to a file containing address information. The conversion (no matter the volume) can be done in a matter of minutes or hours. To make this process easier for you, AGS invites you to purchase a cost-effective “conversion project” designed to let you see how MOSAIC classifies your data compared to your current system.

### ***What do I need on my file to use Mosaic?***

Preferred is address, including street number, city, state, ZIP, ZIP + 4 and Census codes – State, County, Census Tract, Block Group and Block.

### ***What are the different levels of geography?***

The Cluster codes are assigned at the postal level to ZIP Code and ZIP +4 and on the Census level to Census Tract and Block Group.

### ***Does MOSAIC work with just ZIP codes?***

MOSAIC can be used with a variety of different levels of geography. (See previous q/a). If you only have ZIP Codes on your file, you can use MOSAIC to profile your customer base. However – most ZIP codes contain anywhere from a few hundred households – up to and exceeding 20,000. Therefore – to achieve the greatest “lift” from MOSAIC – it is advisable to manage the lifestyle clusters at a lower level of geography.

### ***What is geo-coding and why is it a necessary step?***

Geo-coding is the process of enhancing your address information with finer detail. Typically, this includes State, County, Census Tract, Enumeration District, Block Group and Block. These are the geographic boundaries used by the government in All State, County and Federal surveys, analysis and evaluations.

## ***What does ‘geo-demographic’ mean?***

During the decennial (every 10 years) census the government mails one of two survey forms to every household in the United States — the “long form” or the “short form”. The data collected forms an enormous data repository, a veritable “Fort Knox” of information on the American populace. This summarized demographic information is evaluated and managed via census geography. Geo-demographic — the “marriage” of geographies and demographics — has become an indispensable tool for market analysis.

## ***What support is available for MOSAIC?***

Experian developed this segmentation database and in the United States, in conjunction with Applied Geographic Solutions. Your first line of support, as a reseller, is the AGS sales person you purchased the data/software from. AGS stands ready to support their MOSAIC Data Partners across a wide range of circumstances, from over-the-phone support to joint sales calls.

## ***Can I use my own customer data and custom surveys with MOSAIC?***

Yes, the advantage of using the MOSAIC cluster segmentation system is the ability to cross-tabulate most of your internal information and transaction based information for analysis (profiling) and/or further multivariate or advanced modeling.

## ***What is the advantage of using a commercial segmentation system over a custom model?***

Custom modeling can be a key component to any marketing application. Quite often the results are utilized exclusively for internal analysis and for the internal sale culture. This greatly reduces the option of understanding the market based on outside intelligence (market share and competition). A commercial cluster segmentation database links your data and your analysis to the vast resources outside your environment (e.g. national survey data and lists).

## ***Can I use both?***

Yes, the MOSAIC system is extremely flexible, as a powerful predictor in its own right – it naturally fits well into your own data analysis and modeling.

## ***What variables go into the MOSAIC model?***

- Most recent census defined U.S. states, counties, census tracts, minor civil divisions (MCDs), places and block group totals.
- Postal defined ZIP codes and 4 digit ZIP codes
- Geographic maintenance and correction updating.
- Current Census base counts (population, households, etc.).
- Census demographic estimates and projections for population, households, families, group quarters, and housing units are first and foremost in the initial development process. This is followed in a “top to Bottom” method working from the US to state to county to ZIP code to BGED (Block Group) Enumeration District). Demography given special consideration are density, income, inferred income, age, race, sex, income, housing value, Wealth, IPA – aggregate credit data by geography.
- Experian’s INSOURCE household database. The INSOURCE database covers 98% of all US households and provides demographic information, such as age, income, housing type etc. all at the household level. This database is updated dynamically and represents an up to date and accurate representation of the United States.
- Various data from the Current Population Survey.
- Latest State and county population estimates.
- Behavioral measures via research data including: magazine purchase, credit card usage, mortgage and automotive purchasing behavior – these are “smoothed back” assigned to geography.
- The primary source for base counts (population and housing estimates) is the census, however, the census is not the only source of measurement – the subsequent source is of estimates for small areas of geography are national list processors including Experian’s InSource.

## ***I only have telephone numbers on my customer database - can I use Mosaic?***

One technique is to get a “reverse match” on your customer file. Experian can take their list of valid telephone numbers and cross-reference them with your customer database to append actual addresses to your list. However, this process generally results in a match only 60-75% of the time.

## ***How can I use MOSAIC to tailor my message to different customer segments?***

Each cluster not only describes the geo-demographic make up of an area but also (via links set up to outside survey and data), outlines the lifestyle purchasing preferences of the households. You can use MOSAIC to learn about your customer's vacation & sporting preferences, media habits, product purchases and more. All of this additional information can help you to craft messages that speak to your customers likes while avoiding their dislikes.

## ***Can I use MOSAIC on my Internet customers?***

Any record, which contains address information or telephone numbers, can be easily appended with MOSAIC.

## ***How can I use MOSAIC with Experian's databases?***

All of the Experian databases are MOSAIC encoded, making list selections with MOSAIC easily executable.

## ***Single assignment/distribution...what do they mean?***

'Single assignment': a unit of geography is given ONE (single) cluster classification. For example, a zip code (a unit of postal geography) might be assigned cluster D20 that represents the *preponderant* neighborhood type in that geography. The zip code may be evaluated using smaller units of (census) geography such as block groups. In this scenario, a *distribution* of clusters within the zip code emerges:

<b>Geography</b>	<b>Cluster Assignment</b>		
<b>Zip code</b>	D20	100.0%	Single assignment
<b>Block Group 1</b>	A06	18.8%	Distribution
<b>2</b>	A08	1.7%	
<b>3</b>	C16	8.3%	
<b>4</b>	D20	31.3%	
<b>5</b>	D21	6.8%	
<b>6</b>	D22	23.5%	
<b>7</b>	F34	5.3%	
<b>8</b>	J51	4.2%	

## ***What's a profile?***

The “standard” report and the basic “directional or analysis report” is basically a report which summarizes your customer database by clusters and compares that composition and penetration to that of a “base” – the base being the US as a whole or the market(s) where the majority of your customers are to be drawn from.

## ***What is the purpose of creating a profile?***

To show the “lift” and potential (preference) of a cluster or group of clusters for clear and effective targeting and scripting.

## ***What is “lift”?***

The “lift” is the level of discrimination a particular cluster(s) shows in terms of potential usage for your product or service. Cluster analysis is performed to identify those customer types who have a higher propensity for your product or service than other types.

## ***What does MPI stand for?***

“MPI” is the acronym for Market Potential Index.

## ***What is the MPI calculation used for?***

The Market Potential Index calculation is used to discover the high to low concentrations of best performing clusters in local markets or, in other words a specific markets propensity to use a certain product or service. Therefore, once you have established a MOSAIC profile (high performing MOSAIC clusters) you can then determine where these high performers are located geographically. For example, if your profile shows ten of the MOSAIC clusters heavy users of your product – the next step is to understand where you can find more high performing areas.

This is when the MPI is essential – to find the “gold” hidden in the diverse geographic market haystack. Most MPI analysis is achieved through tabular reports – ranking high performing areas by ZIP code, census tract or block group. Another way is visually, through data mapping. This allows you to deploy your efforts (media, mail etc.) right to the target audience (carving out unwanted areas that might contain unwanted natural or manmade developed areas).

## ***How is MPI calculated?***

The MPI calculation – an advanced mathematical formula that associates the average indices of a MOSAIC profile to a (high to low) distribution of small area cluster neighborhoods. In the calculation (which is extensive), each geographic unit (ZIP code, Census Tract, Block Group) is individually matched against the high and low overall performing clusters. The algorithm continues – calculating and matching and ultimately finished by ranking and categorizing the market for pinpoint precision targeting.

## ***How can I use this information for media planning/buying?***

MOSAIC gives you the power and ability to match your consumer segment with their media preferences – reading, viewing and listening — by linking with Mediamark Research Inc. as well as a variety other data sources. With a MOSAIC license, you may code any custom survey you have commissioned to “link” the two databases.

## ***What does it mean to “link” different databases?***

“Linking” disparate sources of data is one of the primary advantages of using a commercial segmentation system like MOSAIC. Why? When you code your database and sum your customers by clusters (creating a MOSAIC profile), you see a particular pattern evolve – some clusters show a greater propensity to utilize your products or services and “pop” or “lift” more than others. Once you know which clusters are best for your business, you can analyze information on other databases such customer surveys, internet respondents, or prospect lists by MOSAIC encoding those databases and looking for similar cluster patterns. For example, you might do very well with clusters A02, B09, and C19. Using syndicated sources such as Simmons or MRI, you can learn more about what people who are also typed by those clusters read, purchase, and enjoy. You are not physically “linking” the databases (like a relational database), however, you are “linking” the information through association. Remember that segmentation systems are designed around the concept of “birds of a feather flock together”. That means that people on your file typed by cluster “X” have a great propensity to act like other people typed “X” on other databases.

## ***Can MOSAIC be used for site location?***

Yes, the MOSAIC segmentation database was designed with site study in mind – and as a geo-demographic tool is well suited for both new site placement and realignment of current locations (sales offices etc.).

## ***I already have a lot of transactional data on my database...why do I need Mosaic?***

“You can never have enough information”. The addition of MOSAIC to your data armory gives you the current state-of-the-art weapon in direct marketing analysis and targeting. Additionally, the strength of geo-demographic clustering allows you to link YOUR database to other databases as well as geography – transactional data doesn’t let you do that.

## ***What companies use clusters?***

There are over 20,000 companies in the US alone that use commercial clustering systems. MOSAIC is currently used by thousands of European companies and in the US, by Clear Channel Communications – Outdoor, Fleet Bank, Remedy Staffing, Lerner NY to mention a few.

## ***Can I link the cluster deployment to profitability?***

Knowing what areas bring you the highest revenue helps to refine your search geographically. Accordingly, you increase profit via reduced direct marketing costs plus target hitting acquisition of “best” customers.

## ***Can MOSAIC be used for business-to- business applications?***

MOSAIC is a system built for businesses that need to target consumers more effectively. It scores neighborhoods, not businesses. It is not intended for b-to-b usage. If you require a b-to-b solution, please contact your Reseller directly.

## ***How cost effective is MOSAIC in comparison with other systems?***

MOSAIC is one of the most robust, universally applicable, easily executable systems on the market. It is competitively priced...often 25-35% less than other quality systems on the market.

## ***How often is MOSAIC updated?***

MOSAIC is updated annually; with the release coming in the summer months. Licenses include updates on one-year terms.