

CLARITAS CONNEXIONS[®] METHODOLOGY 2021

ConneXions® is a registered trademark of Claritas, LLC. The DMA data are proprietary to The Nielsen Company (US), LLC (“Nielsen”), a Third-Party Licensor, and consist of the boundaries of Nielsen’s DMA regions within the United States of America. Other company names and product names are trademarks or registered trademarks of their respective companies and are hereby acknowledged.

This documentation contains proprietary information of Claritas. Publication, disclosure, copying, or distribution of this document or any of its contents is prohibited, unless consent has been obtained from Claritas.

Some of the data in this document is for illustrative purposes only and may not contain or reflect the actual data and/or information provided by Claritas to its clients.

Copyright © 2020 Claritas, LLC. All rights reserved.
Confidential and proprietary.

TABLE OF CONTENTS

Table of Contents	i
Introduction	1
Data Sources	1
Model Development	3
New Statistical Techniques	3
Assessing the Role of Technoadoption	4
Claritas ConneXions® Lifestage Groups	5
Technical Support	6

INTRODUCTION

Claritas has remained at the forefront of segmentation development due to our willingness to adapt our data techniques to keep pace with the geodemographic data available through the U.S. Census Bureau and other sources. Improvements created by Claritas in statistical techniques, combined with new data sources and structural changes instituted with by the Census Bureau offered Claritas the rare opportunity to build a unique solution for consumer segmentation. The result was Claritas PRIZM, which also serves as the basis for all of the Claritas segmentation systems, including Claritas ConneXions.

ConneXions®, the Claritas segmentation system for communications marketers, classifies every U.S. household into one of 53 consumer segments based on the video, voice, and data purchasing preferences of that household. ConneXions offers a complete set of ancillary databases and links to third-party data. This allows marketers to use data outside of their own customer files to pinpoint products and services that their best customers are most likely to use as well as locate their best customers on the ground. ConneXions enables marketers to create a comprehensive portrait of their customers by answering these important questions:

- Who are my ideal customers?
- What are they like?
- Where can I find them?
- How can I reach them?

ConneXions' external links allow for company-wide integration of a single customer concept. Beyond coding customer records for consumer targeting applications, Claritas provides estimates of markets and trade areas and profile databases for behaviors ranging from leisure time preferences to shopping to eating to favorite magazines and TV shows, all of which can help craft advertising messaging and media strategy. Components of the ConneXions system can be grouped by the stage of customer analysis, as shown below:

CUSTOMER ANALYSIS STAGE	CONNEXIONS® COMPONENT USED
Coding customer records	Household-level coding Geodemographic coding and/or fill in
Comparing coded customer records to trade area	Current-year segment distributions Five-year segment distributions
Determining segment characteristics for demographics, lifestyle, media, and other behaviors	Claritas Household Demographic Profiles Claritas Neighborhood Demographic Profiles Claritas Technology Behavior Profiles Claritas Television Behavior Profiles Claritas Consumer Profiles Custom surveys or databases

DATA SOURCES

In developing ConneXions, Claritas assembled a database from sources that include the proprietary Claritas Technology Behavior Track and the GfK Mediamark Research & Intelligence, LLC Survey of the American Consumer. Each of these records included demographics and nearly 2,000 behavioral

measures, which resulted in more than 200 million behavioral data elements for our research. The behavioral data included measures of both penetration and volume. For example, data is available not only about whether a household subscribes to an Internet service (penetration) but also about how much the household pays for that subscription (volume). Most important, every record in the file had demographic data reported by the survey respondents themselves. This database was regarded as an unprecedented benchmark for other data sources, including the compiled list data that would ultimately be used to append ConneXions® to customer records.

When implementing ConneXions on third-party files, segment assignments depend on the third party's compiled list data. The unique models built for each third-party are designed to produce a distribution of assignments that mirrors the distribution produced by the Claritas Multi-Source Aggregation and Distributional Alignment (MADA) process. MADA is a proprietary methodology informed by data from sources including Epsilon Data Management, LLC, Valassis Direct Mail, Inc., Infogroup, and TomTom North America, Inc. Such data includes, but is not limited to age, income, and presence of children. This information is acquired from third-party providers who have a legal right to provide us such information and the data is either self-reported or modeled. This combination of data sources provides Claritas a unique competitive advantage in its segmentation assignment methodology, due to the unparalleled breadth and depth of address-level information. By combining data from multiple vendors with the Claritas Demographic Update, Claritas can make ConneXions single assignments at the ZIP+6, ZIP+4, and ZIP Code levels, allowing better fill-in for records that do not get a household-level assignment.

The wealth of information available across a number of geographic levels allowed us to construct a massive logical record for more than 890,000 households. To each household record's name, address and behavioral data, Claritas added geographic identifiers at the census block group and ZIP+4 levels; evaluation characteristics that could be used to test and refine the segments covering the entire content of the survey and purchase data sets, providing over 10,000 profiles including both volume and penetration; and assignment characteristics that could be used to define the segments.

Examples include:

- Household-level demographics appended from the Epsilon TotalSource Plus™ file
- Neighborhood-level characteristics from the census block group information
- Summarized ZIP+4 level characteristics
- Claritas custom measures
- Actual self-reported household characteristics such as age, income, and presence of children for a subset of 350,000 households

The resulting database was then used to design and evaluate systems at four levels: household-level, using self-reported data; household-level, using list-based data; ZIP+4; and census block group.

For decades, Claritas has set the standard for global market and consumer insight research. Our customer insights are based on representative samples of the population and help businesses understand what consumers watch, what they buy, and their lifestyle preferences and behaviors to make your marketing more effective.

Claritas' segmentation solutions use a broad spectrum of demographic and lifestyle information to describe households and geography, enabling companies to better understand and anticipate customer buying behaviors. Our segmentation systems place each U.S. household into segments based on general consumer behavior and demographic characteristics. The segments are based on aggregated or

modeled information that represent millions of households. No information about a unique individual or household is published or reported within segment assignments.

MODEL DEVELOPMENT

The methodology that now serves as the basis for Claritas ConneXions® culminated two years of research and development in a groundbreaking methodology that allows marketers to seamlessly shift from ZIP Code level to block group level to ZIP+4 level—all the way down to the individual household level—using the same set of 53 segments. This single set of segments affords marketers the benefits of household-level detail in applications such as direct mail, while at the same time maintaining the broad market linkages, usability, and cost-effectiveness of geodemographics for applications such as market sizing.

New Statistical Techniques

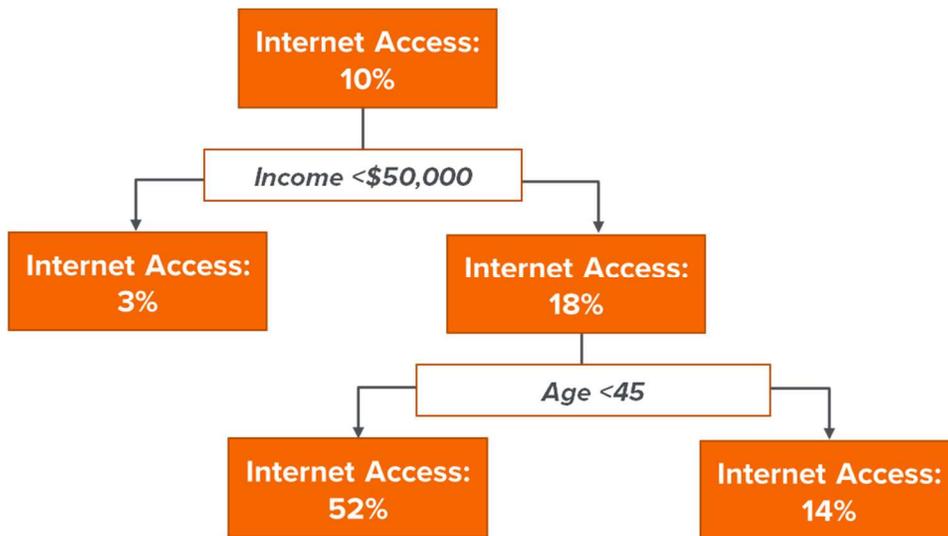
Since the 1970s, the most popular of the clustering techniques has been K-means clustering. The final number of clusters desired is specified to the algorithm (this is the origin of the “K” in K-means) and the algorithm then partitions the observations into K-number of clusters as determined by their location in n-dimensional space, as dictated by demographic factors. Membership in a cluster is determined by the proximity to the group center, or mean, in space (hence the origin of the “mean” in K-means).

For any type of clustering process to work well, the statistician must correctly identify the important dimensions before implementing the clustering process. For marketing purposes, obvious drivers are age and income. However, appropriate levels for each of these critically important dimensions still need to be chosen. For example, does the dimension of income create better differentiation at \$35,000 or \$50,000? How does choosing between these two values of the same dimension change the clustering outcome? These choices are important, because when the clustering iterations end and yield an answer, marketers are left with clusters of households that have been organized by their proximity to each other by the demographic metrics that were chosen. This answer may or may not be meaningful to the original task of creating groups that differ in their behaviors—in large part because behavior measures were not incorporated into the clustering technique itself.

With ConneXions, Claritas broke with traditional clustering algorithms to embrace a new technology that yields better segmentation results. ConneXions was created by a proprietary method developed by Claritas statisticians called Multivariate Divisive Partitioning (MDP). MDP borrows and extends a tree partitioning method that creates the segments based on demographics that matter most to households’ behaviors.

The most common tree partitioning technique, Classification and Regression Trees (CART), involves a more modeling-oriented process than clustering. Described simply, statisticians begin with a single behavior they wish to predict and start with all participating households in a single segment. Predictor variables, such as income, age, or presence of children, are analyzed to find the variable—and the appropriate value of that variable—that divides the single segment into two that have the greatest difference for that behavior. Additional splitting takes place until all effective splits have been made or the size of the segment created falls below a target threshold.

In the example below, the CART process starts with all of the survey respondents in one segment for the behavior of interest—in this case, having Internet access. Of this particular respondent pool, 10% report having Internet access. Next, the CART routine searches for the demographic variable—and the value of that demographic variable—that creates the two segments that are most different on the behavior of interest. Our example shows that dividing the first group by an income of \$50,000 yields two segments—one with Internet use of 3% and the second with Internet use of 18%. We can divide the second segment again, with the result that a split based on Age <45 yields two more segments—one with Internet use of 52% and the other with Internet use of 14%.



If the process stops here, we have a segmentation system with three segments—one with 3% of its members having Internet access, a second with 52% of its members having Internet access, and the third with 14% of members having Internet access. However, this resulting segmentation system does not provide useful information about any other behaviors—it’s optimized only for Internet access. This is one of the limitations of the CART technique: It generates an optimal model for only a single behavior. Because ConneXions® is a multi-purpose segmentation system, optimization across a broader range of behaviors is necessary. Claritas made several modifications to the CART process, resulting in the MDP technique, for which a patent is pending. These modifications extended the basic CART process to simultaneously optimize across 250 distinct behaviors at once. This advancement allowed Claritas to take full advantage of the nearly 2,000 behaviors and hundreds of demographic predictor variables available at different geographic levels, including the household level. The MDP process was run hundreds of times, with varying sets of behaviors, predictor variables, and a number of other parameters, to ensure that the resulting segments represent behaviorally important groupings. Note that the information above is purely for illustration and does not represent real data.

Assessing the Role of Technodoption

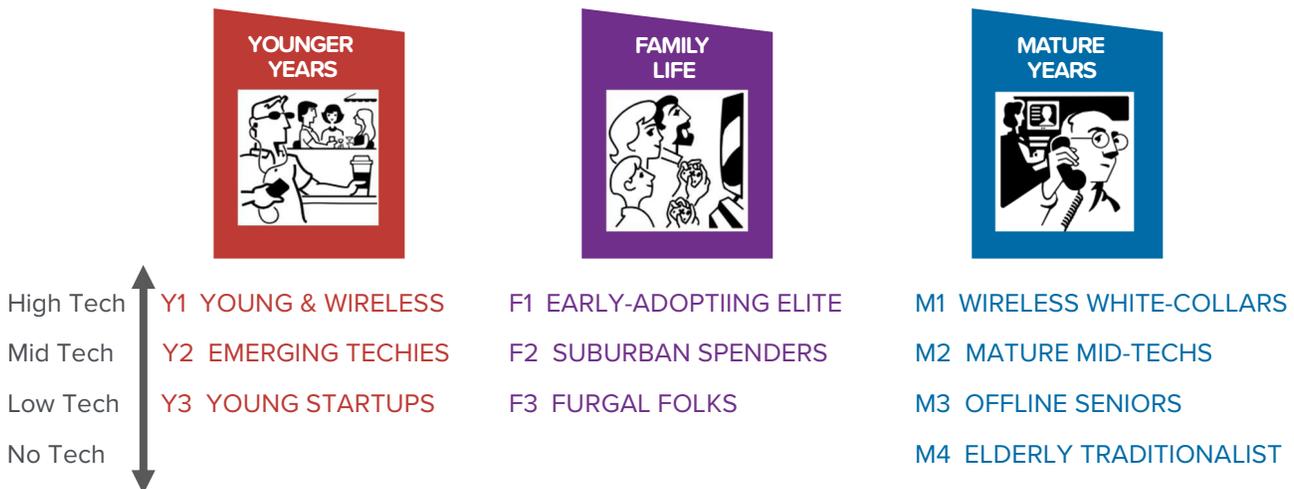
A distinguishing feature of ConneXions is its use of “Technodoption”, a proprietary Claritas model that measures a household’s willingness to adopt new technology early in its lifecycle. Technodoption

assesses a household’s likelihood to have cutting edge products. The four levels of Technodoption are as follows:

- **High Tech**—Segments classified as High Tech are the technology trendsetters. They are the most likely of all ConneXions® segments to be the first to try a new technology—and every new technology.
- **Mid Tech**—Segments classified as Mid Tech are a bit more selective about the new technologies they will try, although it is unclear if savvy or salary is driving the hard choices. They are frequently heavy users of one particular product or service while lagging in other products.
- **Low Tech**—Segments classified as Low Tech are best described as those with big dreams and low budgets. Many of these segments would adopt more technology products if they could afford to.
- **No Tech**—Segments classified as No Tech not only pass on trying new technology, they are also lagging in adopting well-established technology. Their philosophy is “Why pay for cable when you have an antenna?”

ConneXions Lifestage Groups

Technodoption is combined with the Claritas Lifestage classifications to form the framework for the ConneXions group typology, shown below.



Younger Years segments consist of mostly singles and couples who are typically under 45 years old and generally have no children in the household. Residents may be too young to have children and/or are approaching middle age and chose not to have them.

At the household level, around age 45 is the cutoff for most segments. Among these younger segments, only those explicit in their definition for lack of children or with low indices for presence of children tend to be included in Younger years.



Family Life is composed of segments that are middle-aged and either defined by presence of children in the household or have high indices for household with children under age 18.

At the household level, presence of children is the primary driver for many segments in this class. While this class also includes segments where the presence of children is not explicit at the household level, in general they do show high indices for that characteristic.



Mature Years includes segments whose residents are primarily empty-nesters or those with children in their late teens, away at college or rebounding back to mom and dad's home.

At the household level, the primary driver is age, not necessarily the absence of children. Segments that are uniquely child-centered tend to be younger and are grouped under Family Years while those under age 45 and without children are grouped in Younger Years - leaving the last group of segments for the Mature Years.

TECHNICAL SUPPORT

If you require further assistance, please contact the Environics Analytics support team between 9:00 a.m. and 8:00 p.m. (Monday through Friday, EST) at support@environicsanalytics.com or 888.339.3304.