Marketing data quality

An Experian white paper
“The most proactive CMOs are trying to understand individuals as well as markets. Customer intimacy is crucial — and CEOs know it. In our last CEO study, we learned CEOs regard getting closer to customers as one of three prerequisites for success in the 21st century.”

1 Source: “From Stretched to Strengthened, Insights from the Global Chief Marketing Officer Study” — IBM October 2011, a survey of 1,734 CMOs
Introduction
What does it mean to get closer to customers? How close is close enough?
Getting closer to customers largely entails learning as much as possible about
key attributes that are correlated with specific customer actions and purchase
behaviors. Paramount to this understanding is the ability to describe individual
customers and external prospects based on a wide array of demographic, lifestyle
and behavioral characteristics. In essence, getting closer to customers means
having the ability to paint vivid, multidimensional pictures (call them portraits
that put a face on customers) based on information collected and known about
each and every customer. With the fully enhanced portrait in hand, marketers can
increase the effectiveness of promotions, offers, communications and strategies
designed to appeal to individual customers and prospects.

Accurate customer portraits require direct linkage and access to data sources
that are continuously refreshed with reliable and relevant market information.
Much as an artist will mix a blend of colors from a painter’s palette to convert a
blank canvas into a valuable piece of art, direct and digital marketers must utilize a
palette of data elements to help visualize the defining features and characteristics
necessary for putting a face on customers. Experian®, along with a host of other
marketing services and data suppliers, provides marketers with a rich data palette
for creating the customer portrait. However, from one supplier to the next, there
are considerable differences in the shapes and sizes of palettes (data sources
used) as well as the amount (coverage) and vibrancy (accuracy) of colors (data
elements) available to compose the picture.

To better understand the inherent differences in overall coverage and quality
of data provided by marketing services providers, Experian commissioned an
independent research study that was designed to address the following questions:

1. When ordering a file of contact information, how do suppliers differ in their
overall depth of coverage for a geographic area?
2. How much overlap in coverage exists across suppliers? How many unique
addresses are present on a supplier’s contact file?
3. How accurate is the data available from a supplier?

These questions were posed because they are essential components that
marketers need to understand before making decisions to develop targeted
marketing campaigns that integrate contact information and associated
demographic, lifestyle and behavioral data elements from a specific marketing
services provider.
Armed with the proper knowledge for evaluating a marketing services provider, the difference becomes (1) marketing in black and white (communicating to undifferentiated, faceless customers and prospects) versus (2) having the confidence to invest the necessary resources for marketing in living color (communicating to highly differentiated customers and prospects, each with unique motivations, tastes and preferences).

Description of methodology
A study conducted by an independent third-party research organization was commissioned by Experian to assess the overall data quality, accuracy and coverage of contact information available from competing marketing services and data suppliers. The research methodology was based on a random sample of 348 carrier routes that are statistically representative of the U.S. population.

Contact information (including name, home address, phone number and email address) from all records within the sampled carrier route footprint was independently acquired from Acxiom, Epsilon, Experian, Infogroup and KBM Group by the research firm. Each vendor was supplied an identical request for the data.

This paper presents the key highlights and findings from this third party data quality audit.
Data coverage
Depth of coverage is a critical measuring stick for evaluating the overall penetration and effectiveness of a direct-marketing campaign. The greater the coverage, the more likely a campaign will be successful in reaching an optimal number of prospective customers. With larger contact quantities, more sales conversions are possible. To determine the degree of variation in coverage among the data suppliers, the number of records (unique addresses) represented in each marketing file were compared.

Unique Addresses

Comparison of overall coverage based on number of unique occupied addresses by data supplier

- Acxiom: 64,796
- Epsilon: 67,857
- InfoGroup: 63,113
- KBM: 65,244
- Experian: 68,565

* Based on a random sampling of 348 ZIP™ carrier routes across the United States.
Marketing data quality

Key findings

- ConsumerViewSM ranks first in overall coverage based on the number of unique addresses available within the study area.
- Epsilon was found to have 14.8% duplicate records at the address level.
- When extrapolated to the entire United States, the depth of coverage available from ConsumerView compared to the closest competitor is equivalent to nearly 500 additional addresses for every 10,000 addresses located in a geographic area.
- All data providers except Experian include records that Experian would consider unoccupied, deceased, commercial or have households that have moved.
- After removing inaccurate data and combining the files, Experian's ConsumerView contained 4,200 unique addresses within the sample footprint. These addresses were exclusive to ConsumerView and not present in any of the other data providers' files.

Further reinforcing the depth of ConsumerView's coverage, none of the marketing files acquired for the analysis contained unique addresses that were not present in ConsumerView. In other words, ConsumerView accounted for each and every address that was present on a competitor's file.
Suppliers' records matching Experian end-dated records

A significant number of the marketing files provided by each vendor contained records that matched Experian's historical and former address records. These addresses are not considered valid and current combinations by Experian to be marketable. Match rates of competitor files to the Experian end-dated records ranged from a low of 8.8 percent to a high of 9.8 percent. The net result is that competitor mailing files are likely to have significantly inflated counts because of the inclusion of nonmarketable records.

“We house other files, and by comparison, Experian’s data is more accurate.”

— Respondent comment from Experian customer loyalty survey

*Based on a random sampling of 348 ZIP™ carrier routes across the United States.
Key considerations when assessing data quality

Why do I need ConsumerView data when we have access to our customer interaction data?

Marketers today are capturing a wealth of information about their customers. Insights can be gleaned from transactional details and other customer touchpoints, including customer service interactions and tracking of online behaviors. Employing a variety of methods to compile and analyze the resulting data is highly encouraged, as it enables marketers to get closer to customers. However, marketers can get even closer to customers by enhancing and enriching the internal information they have expertly gathered with external data points. This external data helps to paint a more rounded picture of customers and prospects.

For example, assume that a retailer is experiencing a decline in their customer retention rate. External data indicates that a customer has a high likelihood to participate in sports and fitness activities. Matching back to their transaction history reveals that this particular customer has never made a purchase within the category. This could indicate a missed opportunity. The customer is then included in a cross-sell promotion that prominently features sports and fitness products. Additional external data might also indicate that there are multiple children age 5 to 15 present in the customer’s home. The promotion for this customer could then be enhanced to also feature related children’s products.
When should known data be used versus modeled, and as a compiler of data, do you separate the two?

It is common practice for marketing services and data suppliers to develop predictive models that infill data elements for records where information is unavailable or unknown. This modeled data enables the supplier to increase their overall coverage for a particular data element. However, many suppliers do not indicate in the marketing files that are delivered to clients whether the actual data element is based on known information or is a point estimate from a model. Having the ability to differentiate between known versus inferred values is oftentimes critical to the ultimate success of a targeted marketing campaign.

For example, if a marketing campaign is specifically tied to identifying prospects who qualify for an offer purely on the basis of meeting a certain age requirement, then it becomes very important that the targeting criteria be based on known, not inferred, age values. Otherwise, the marketer takes the risk of wasting precious advertising dollars by including prospects in the campaign that have no chance of qualifying and making a purchase. ConsumerView gives marketers the choice to screen prospects based on known age data, thereby excluding prospects with an inferred age.

Data coverage and quality as evaluated in the marketplace

ConsumerView data has stood the test of several independent custom research projects that have been initiated by several leading companies and organizations.

• A leading regional bank found that ConsumerView accurately identified more than 70 percent of the company’s jumbo loan customers. The highly favorable results of the analysis led the bank to develop a case study that was presented to the organization’s executive board for garnering consensus and approval to move forward with a jumbo loan marketing program. The case study included statistics from the research summarizing the overall quality and accuracy of the ConsumerView data, especially with respect to property and realty information.

• A governmental agency conducted a controlled test to evaluate market data from ConsumerView against multiple competing vendors. The test provided the organization with measurable results and the confidence to license ConsumerView as part of a long-term contract to use the data in support of ongoing analytical exercises.

• A children’s apparel retailer conducted a detailed comparison of data from ConsumerView against other suppliers that they also were considering. In their conclusions, the retailer cited the overall quality and accuracy of the data. In particular, the ConsumerView children’s data was highlighted as standing out relative to comparable data from the other competitors.
• Other cases, as described in the following real-world examples, highlight the misapplication of marketing data. This is detrimental to the data user, especially when data quality standards are not vigorously maintained.

• Because of economic stress and other related factors, there has been a recent trend toward more people living in homes that contain multigenerational family members. The end result is that some marketing service providers are carrying extra records on their files that in reality should be combined into a single household. This has created confusion when attempting to accurately account for the true number of marketable addresses and households in an area. Unfortunately, this practice leads to inflated household counts and unnecessary waste in marketing expenditures, as mailers are using marketing files with duplicate records, targeting too many people at the same address.

• The convenience of the online channel has prompted many data aggregators to collect information from consumers via Website registrations, popup surveys and other interactive tools that promise free merchandise as incentives. This is a very economical and quick way to compile consumer information. However, there are severe drawbacks and limitations to using this data for targeted marketing purposes. Primarily motivated by the attraction of “free stuff” and fearful that their privacy might be compromised, consumers who respond to these inquiries might not be truthful in their responses. The validity of this data is extremely suspect, especially when procedures are not put in place to confirm and validate its accuracy.

A system of rigorous compilation processes and methodologies that is regularly applied to data acquired by Experian is fundamentally responsible for ensuring that the overall composition of ConsumerView meets the highest quality standards. This helps minimize any misapplication of the data that can potentially add to the cost of a campaign, thus diluting overall return on investment.

“Experian is a highly trusted company with exceptional data quality. The data that we receive is 99.9 percent accurate.”

— Respondent comment from Experian customer loyalty survey