Survey data, meet data lake.

Data lake, meet survey data.

Powerful ways to get the most out of your marketing-related data

Getting the what and the why from your data lake

An organization's data lake is critical, containing all the information needed to empower company operations with useful business statistics, metrics, trends and other facts. There are clear benefits to having all a company's data stored in one cloud repository, including a more holistic customer and marketplace view.

Today, marketers are finding that their survey and other marketing-relevant data is also stored in the lake. Challenges can arise when it comes to getting the most out of this expensive data asset - properly shaping it for consumer insights and marketing purposes. This process can only be optimized with the right technology that integrates smoothly with the data lake and other business processes, while furthering organizational strategic objectives.

Optimizing the technology stack

Most data exists inside the lake as "blobs" - data stored in its original form, which hasn't undergone any kind of transformation to enter the lake. All types of data are stored in the lake, from historical to real-time data to customer insights. They don't have to be processed in any way - until they are needed for powering decision-making and business intelligence.

Centrally locating all a company's data is just the first piece of the puzzle. Many companies use minimal, targeted solutions for storage, extraction, and analysis of the data while keeping security and performance top of mind. We find the largest challenges at big organizations are getting the data out in a format that is useful for taking action, such as interrogating the data, finding new insights, and generating shareable reports. This is particularly challenging when it comes to critical consumer insights and survey data.

Yet, adding a solution for only analyzing and processing market research data can cause some IT teams to take pause. Limiting the technology stacks that feed into and pull out of the data lake and those that analyze the data for insights is essential for scalability, security, and simplicity at most organizations. A more granular understanding of the specific needs surrounding a company's market research data can help to inform more effective approaches.

Market research data: more than meets the eye.

We know that consumer insights are a vital part of mapping customer journeys, understanding motivations, and proactively developing products, services, and communications that will ultimately spell organizational success. Survey data is a critical part of uncovering these insights, allowing us to understand why consumers or customers are doing what they are doing, and informing ways for us to optimize their journey.

The process of deriving insights and business intelligence from survey data can be a complex undertaking. Storing this information in the data lake is one thing; having the right tools **to** gain understanding from it is quite another. While traditional business intelligence tools have many strengths, they are not generally equipped with the functionality necessary to deal with the complexities of survey data.

To truly understand what audiences are telling us in surveys, we must - for example - find solutions that can:

Properly represent weighted data: Just as people and audiences are complicated, so are their resulting survey responses. We can't simply "count" answers and expect to come up with an accurate story. Researchers often need to weight their data so that it more accurately reflects the surveyed population.

Weighting helps to correct any selection or probability issues that occurred during the sampling process. Because survey data is almost always based on a sample of an overall population that fits specific demographic profiles, it is not the count of people who completed the survey you are interested in per se, but it's the total population (or the percentage of the population that you care about). Thus, results must be based on the correct denominator and be representative of the population.

The process of sampling and weighting is a delicate balancing act and must be handled carefully to ensure that the data accurately reflects the real world. A solution that facilitates and optimizes this process properly boosts data quality. Tools made specifically for market research can even allow users to compare weighted and unweighted data, critical in checking the quality of the underlying sample.
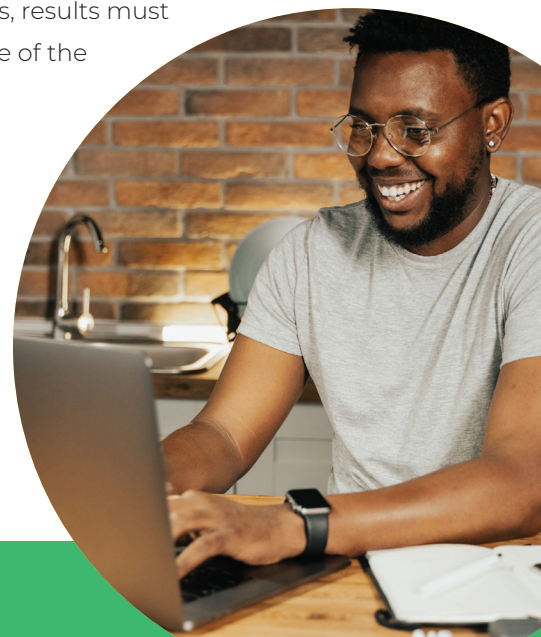
Handle multiple responses with ease: Survey data does not always consist of simple yes and no questions. Many questions, such as "Which of these products have you purchased in the last month?" require respondents to click "all that apply." Multiple response questions like these yield variables where the elements are not mutually exclusive. In other words, the total responses will add to more than the total respondents (cases).

Analysis tools that are optimized for survey research data natively deal with multiple response variables in a way research and insights professionals are familiar with and will increase the likelihood that the full value of your survey data will be realized.

Quickly identify relationships in the data: Digging through crosstabs to find how target groups of data relate to one another can mean that you miss the stories that the data is telling. Solutions that offer advanced statistical testing can facilitate fast, smart analysis of the data by bringing the most interesting findings to the surface.

For example, the right tool can automatically perform a series of statistical tests on each descriptor variable, comparing values from a target group with the rest. These values are compared using Bayesian Statistics to calculate the probability that the group value is greater (or lower) than the rest of the values. This approach quickly allows the user to find similarities and differences among key target data groups. Without this kind of functionality, a researcher can waste valuable time trying to measure relationships or, worse, completely miss the most important insights.

These are just a couple of examples that display how market research data is different from other more structured data that organizations use to drive decision-making. And, because this is the data that comes directly from key audiences, it is critical that it be included - accurately - in the mix.

Beyond BI tools: Solutions that work for all the data

While some organizations seek to fit a square peg in a round hole, and force existing BI tools to process market research data, it generally requires specialized coding and internal technical debt. What's really needed is a solution that can help everyone in the organization get the most out of all the data inside the lake, including critical survey and market research data.

In a perfect world, an organization's lake will contain the data needed to gain a more holistic understanding of increasingly complex audiences. We can start to move toward this level of understanding when we, for example, connect market research input with the other essential consumer touchpoints, such as digital behavioral activities, sales data, online customer activities, and more.

The right solution will integrate multiple data sources and make them usable as a whole - irrespective of the source. Using human expertise and smart technology, there are tools available to combine common variables, match relevant items together and integrate new data as it comes in. Ultimately, this results in connected data that can be interrogated and analyzed as a whole.
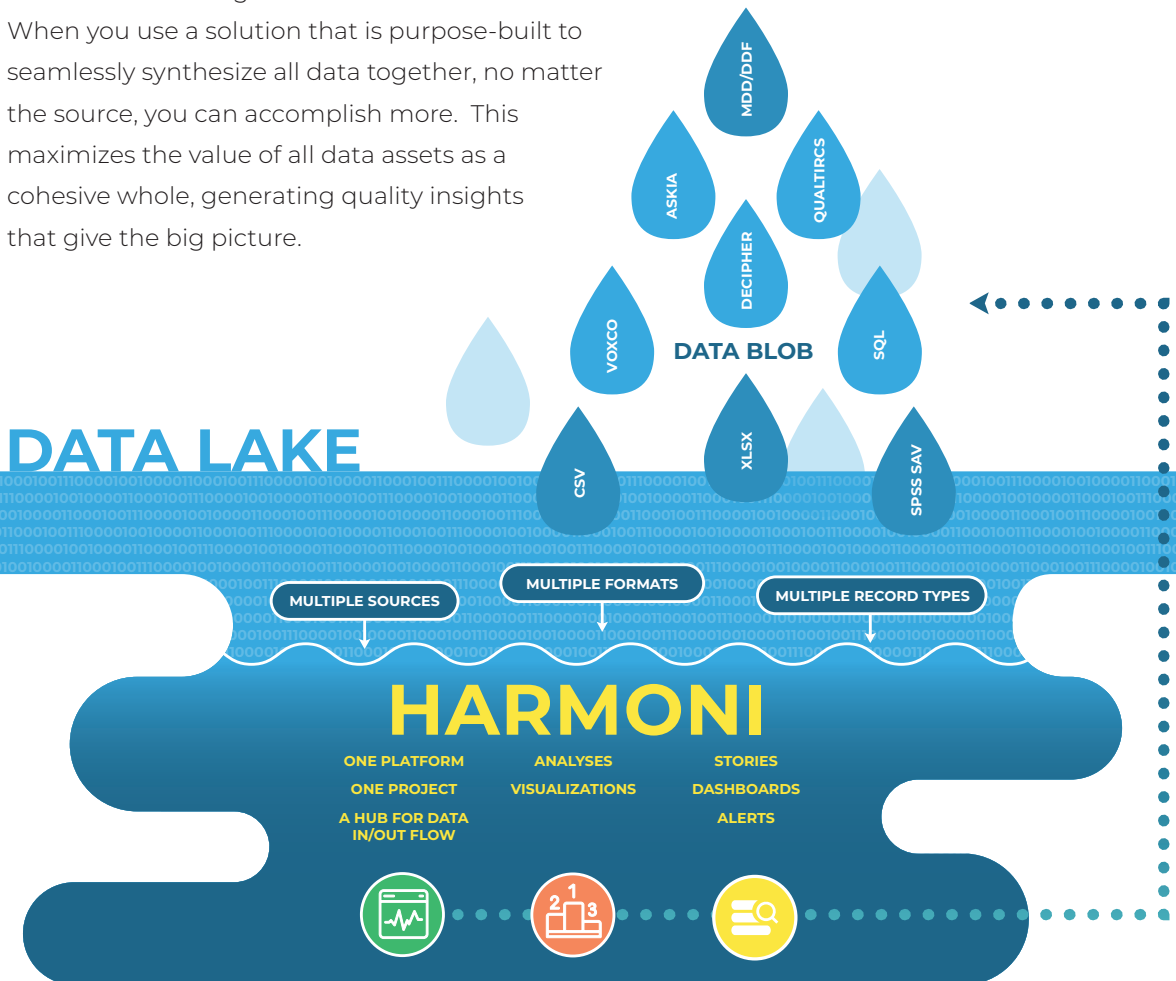
Extracting the most value from data lake

In an organization's data lake, there are two ways that all this data is moving: into the lake for storage and out of the lake for reporting. While the data may not ever "physically" leave the lake itself, it still needs to be analyzed. Often, a missing piece is a solution that can populate the lake with marketing-related data in a useful format and form a loop to the other end of the process - facilitating the proper shaping, analysis, and sharing of data.

There are solutions that can accomplish both of these tasks. As illustrated below, Harmoni, unlike some other survey data tools, is powerful enough to deal with extremely high volumes of various marketing-relevant data, such as data scraped from social listening and other online audience activities.

When you use a solution that is purpose-built to seamlessly synthesize all data together, no matter the source, you can accomplish more. This maximizes the value of all data assets as a cohesive whole, generating quality insights that give the big picture.

DATA BLOB

MDD/DDF
ASKIA
QUALTIRCS
DECIPHER
VOXCO
SQL
CSV
XLSX
SPSS SAV

DATA LAKE

MULTIPLE SOURCES
MULTIPLE FORMATS
MULTIPLE RECORD TYPES

HARMONI

ONE PLATFORM
ONE PROJECT
A HUB FOR DATA IN/OUT FLOW

ANALYSES
VISUALIZATIONS

STORIES
DASHBOARDS
ALERTS

Look for a solution that is uniquely positioned to benefit organizations using a data lake. Some of the features you should seek include technology that:

- **Allows the use of a limited stack of technologies.** As mentioned above, having individual technologies specific to each data stream isn't efficient enough for most organizations. For example, while Harmoni is purpose-built for marketing-related data, its powerful and flexible design is also perfect for any other traditionally structured data. Because Harmoni can connect to multiple data streams and create analysis and insights from all those streams, it is the perfect solution for those looking for a single tool for interacting with the data lake.

- **Keeps data within an organization's internal security layer.** Your data is a valuable asset that must be safeguarded. This means any solution accessing your data lake should be hosted within your security ecosystem for the best protection. We designed Harmoni so that access can be integrated with your organization's Single Sign On (SSO) authentication process, hosted within your own server or virtual private cloud, or under other security measures.

- **Provides visibility into the "extract, transform, load" (ETL) data process.** Transparency is vital in all business practices and looking at the way data is copied and represented is no different. Harmoni stores all the ETL data design steps and can reapply them easily to data updates. This visibility into a complex process can help ensure better connectivity and scalability, among other benefits.

- **Democratizes insights across the organization.** As its name suggests, Harmoni was built to harmonize data - aligning information from multiple sources to create the "big picture." Democratizing data is also easy with Harmoni, as users can get a comprehensive look at all the synthesized data and share customized insights with the right stakeholders across their organization.

Closing thoughts

When it comes to getting the most from the market research data stored in a data lake, there are several key questions that teams can work together to answer when finding the right solution.

Will the solution I'm considering properly handle my unique data inputs?
If researchers are spending a lot of time manually preparing data for inclusion in the data lake, organizations are wasting time and money, not to mention the increased propensity for errors. The right solution can seamlessly bring together marketing-related data like survey responses, online interactions, in-store shopping data, social media posts, and other marketing-relevant data. Ideally, the solution will have 360-degree capabilities when it comes to getting the data ready for seamless analysis: properly preparing data for loading into the lake, plus the ability to organize and prepare data that is already in the lake.

Can the solution properly investigate market research data?
As we outline above, researchers often need to examine survey data through a different lens and must consider weighting data, data hierarchies, significance testing, statistical testing, processing multi-response questions, or building metrics on the fly. Without these capabilities, the researcher is at a disadvantage - and so is the organization. Standard business intelligence platforms just won't provide this level of insight.

Will in-house solutions need significant overhauls to work with market research data properly?
Modifying existing tools to meet the researcher's unique needs can take a long time - time we don't have in today's fast-moving marketplace. These efforts can also cause crippling technical debt and frustration, taking the focus away from primary company goals, like innovation, customer service and the development of products and services.

Nearly every organization, small and large, has a cloud storage solution of some kind. Data has a high amount of inherent cost associated with it - it must be collected, stored and made available to the stakeholders that need it. This data has no actual value until it is properly used to contribute to good things being done in the organization.

INFOTOOLS®

Are you ready to make the most of the survey data in your data lake?

Let us help you find the insights you need.

Connect with Infotools

✉ hello@infotools.com

🌐 infotools.com

🐦 @infotools

f @infotools1

in linkedin.com/company/infotools

▶ www.youtube.com/c/InfotoolsLtd

About Infotools

Infotools is an award-winning market research software and services provider, used globally by some of the industry's most curious minds. We help insights professionals explore their data, answer more questions and provide greater insights. Our integrated, cloud-based software platform, Harmoni, is purpose-built to simplify the processing, analyzing, visualizing and sharing of market research data.