



Why Your Company Needs Data Harmonization for eCommerce *(and how you can implement it)*

TABSAnalytics

**DECISION
INSIGHT** 
A TABS Group Company

The Consumer-Packaged Goods (CPG) industry is in a unique position compared to a lot of other industries. The COVID-19 pandemic caused a sudden, seismic shift in the retail sector in 2020. The resulting disruption of consumption patterns was reflected in extraordinary changes in shopping behavior impacting what consumers buy, where they buy, how much, and how often. Online grocery sales experienced explosive growth, with sales growing 40% in 2020, more than double previous projections before the pandemic.



Although COVID-19 disrupted sales in many sectors, CPG showed not just success, but growth, in 2020. Sales increased across most categories in the grocery store. Brick-and-mortar trips and online shopping experienced a massive surge during the first half of 2020. In addition, we are seeing changes in what consumers are purchasing, for example:

- Products like carbonated drinks are seeing spikes of interest, after years of being stagnant.
- Healthier alternatives have dipped after being some of the better performers in past years.
- Households without children are a larger chunk of the market.
- Companies are trying to discourage deal buying, but the demand is higher than ever.
- Significant growth in eCommerce grocery purchasing earlier in the year.

Adapting to these shifts is essential for CPG manufacturers and retailers to remain competitive in this new normal and beyond. However, none of that happens without data, which presents its own issues. CPG companies, more so than other businesses, tend to get their data from a variety of different sources, including retail syndicated and POS data, consumer and shopper panel data and ePOS.

This means that you have a lot of different data points that can create insights when brought together properly. However, if the data points are not all measuring the same metrics, it's almost impossible to get meaningful insights. This is where data harmonization comes into play, and every CPG company needs to understand what harmonization is and how to use it.



What Is Data Harmonization?

First, let's give a formal definition of [data harmonization](#) and what it entails. Data harmonization is the ongoing process of taking data generated from various formats, name conventions and verticals, and converting it into one master data set that your team can work with. To understand how this plays out, we need to first understand a typical data collection process for CPG companies.

Most of these companies understand that they need to collect financial/customer data to better plan their business tactics, as well as handle other important tasks. However, that data isn't going to come from the same source. For example, let's say that you commission a third-party research firm to determine your key sales demographics. At the same time, you do a customer survey to try and figure out how well a certain sales promotion is being received. All this data can be usable, but if you try to create insights one set at a time, you're going to get inaccurate analyses. This is only for two datasets. Imagine expanding this to the dozens or hundreds of datasets that an established CPG company builds over time.

Another thing that we should mention about [data harmonization](#) is the fact that it's not the type of thing you simply implement once and move along with. Instead, it's a continuous task that's

constantly working to keep your past, present, and future data all easy to understand so you can regularly extract analysis from it. Some people make the mistake of confusing it with data integration.

Data integration is essentially part of the work of harmonization. With integration, you take all the data from these disparate databases and integrate it into one larger one. This can be useful if you just want to make sure data is complete or accurate. However, harmonization puts existing work into converting all that data into a metric and format that works with one another. Some of the added work this includes is:

- Validating your data points to make sure metric comparisons are accurate and stay accurate.
- Making sure the process you use is scalable to handle the inevitable increased data.
- Creating new attributes that will help for analytics.

There's a reason why most companies opt to reach out to third parties for help with their data harmonization. This requires a lot of specialized tools and expertise, which we will go into greater detail on later.

The Fundamentals of Data Harmonization

Let's cover some essential points about data harmonization in the CPG world.

- COVID-19 is causing every industry to shift paradigms, even though CPG is weathering the storm well.
- However, it also renders a lot of existing data obsolete, meaning that new insights and analyses are necessary.
- Complicating things is the fact that you can have several different data sets from different sources.
- Data harmonization is the ongoing process of taking information from different data formats and sources and integrating it into one or several related databases.



Benefits of Data Harmonization

So, with this information in mind, how exactly does data harmonization help CPG companies? The goal of data harmonization is to make business data “analytics-ready”. This delivers several key payoffs:

- **More accurate reporting:** Different sources can have different values for the same metrics. Failing to harmonize this data can lead to errors, and in the long run, making decisions off bad information. This could ultimately cost companies millions of dollars.
- **All data in a central area:** Taking the time to put all your data in one central area helps cut down on confusion and improve efficiency. Say that you needed to quickly pull some information about a sales promotion. Rather than going to several different departments, you could have a single master database with everything you need. In addition, this avoids the problem of teams using different sets of data to try and address the same problems.
- **Lower expenses:** We can’t skip over the fact that data harmonization is good for your bottom line, in a few ways. First, if you’re purchasing data on top of what you are getting from your own records, that represents an added cost. However, you need to pay more to clean the data, which makes things even more costly. Making things even worse, the benefit you get is lower because this secondary data isn’t aligned with your in-house data. The end result? Data that provides less ROI for you.
- **Savings through efficiency:** There are also some more practical ways that data harmonization saves you money. Every time a member of your team must scramble through different databases to find a specific source, that’s time that could be better spent elsewhere. Reducing inefficiency means each piece of data you have is more profitable in the long-term.
- **Centralizing data governance:** You need to also think of the long-term when it comes to investing in data harmonization. If you have disparate data sets but can find an organization style that works for you, rebrands or consolidations can throw you into complete disarray. Data harmonization makes adapting to these major changes easy.

Ultimately, data that’s harmonized and “analytics-ready” will make a company more competitive, grow sales and defend market share.

How Data Harmonization Supports CPG Goals

Here are some key benefits of data harmonization.

- Being more efficient by moving all data to a central location that allows the organization at large to access it (vs. data being in silos).
- Reducing the chances of poor, inaccurate data that skews your analyses.
- Helping to centralize data governance, making universal changes easier to manage and making data lakes more robust.
- By making data more accessible and usable, harmonization delivers a better ROI.



Data Harmonization and eCommerce

Every CPG manufacturer has had to reevaluate its online channel strategy and sales in 2020 due to COVID-19, and with major shifts in customer behavior, old data and frameworks have been permanently altered.

So, what are some of the major points to look at regarding CPG eCommerce?

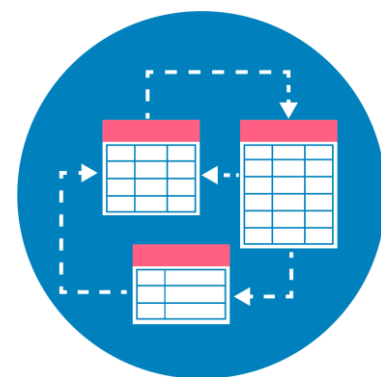
- [COVID-19](#), of course, drove a major surge of interest in eCommerce. In April of this year, purchasing was up 37% compared to the same time the year before.
- Mega eCommerce companies like Amazon stand to benefit the most, both from the increased interest as well as many of their smaller brick-and-mortar competition being forced to shut down. In some cases, these shutdowns are permanent.
- Companies that have a brick-and-mortar presence are trying to adapt by putting a greater emphasis on the in-store experience and an omnichannel strategy.

All of this means that data is at an even larger premium than before, for a few key reasons. These include:

- The increased importance of eCommerce data. eCommerce companies, CPG included, need to constantly adapt their stock in real-time to data trends.
- Digitally-native brands are becoming a larger and larger presence, and they generally aren't covered in brick-and-mortar data.
- Data is going to play a larger role in areas like competitive pricing and even product development. For example, data can point out the best methods of packaging, what flavors customers respond to, and other important information.

As a final point, however, we should share some insights in TABS' recent [2020 Food and Beverage Survey](#) conducted in July 2020. One would think that if anything would push online grocery's share further, the COVID pandemic would be it. And certainly, if we look at [data points](#) from earlier the same year, it seemed that way. However, our survey showed a few interesting pieces of information:

- The number of people claiming they are heavy online shoppers reverted to the average.
- Overall, online grocery sales are seeing a decline in purchase frequency from earlier this year.
- Mega-companies like Amazon, Target, and Walmart all saw sizable increases in their purchase frequency. However, they never say the success norm of 75% of their buyers being frequent buyers (over six purchases annually), a proxy for "loyalty/share of requirements".



Digitally-native brands are becoming a larger and larger presence, and they generally aren't covered in brick-and-mortar data.

How does this factor into data harmonization? It showcases that many CPG companies focusing their attention on both online and brick-and-mortar will likely be continuing to do so to track the eCommerce channel. Further, many consumers have shown a preference for buying online in several CPG categories, including cosmetics, health and beauty and vitamins. And, despite the costs of online grocery outlined in our 2020 study, tracking the channel will become more important than ever.

This means tracking multiple datasets across online and brick-and-mortar, making data harmonization critical to bringing them together.



Data Harmonization In The World of eCommerce

eCommerce is the CPG channel that is on everyone's mind lately.

Here are some key areas where data harmonization plays an important role.

- Because so many brick-and-mortar retailers are changing strategies, they need harmonized eCommerce data to base their decision-making on.
- Syndicated brick-and-mortar data sets need to be paired with eCommerce data sets to better understand the shifts occurring between the two channels.
- This means you'll still be seeing varied data sets across online and brick-and-mortar stores, making data harmonization an absolute necessity to fully understand your all-outlet sales trends.

Implementing Data Harmonization in eCommerce

At this point, it's clear that all CPG manufacturers should harmonize their datasets. However, this isn't something that a company can implement overnight. Let's go over some of the essentials required to perform data harmonization properly, starting with the main tool you need to invest in: the right data software and processes. This is going to do a lot of the heavy lifting when it comes to the harmonization process, from moving data from the original data sets to the actual work of harmonization. It's important that the software that you look into has the capability to handle the following aspects:



- Managing original datasets
- Data transformation
- Presentation
- Analytics

Many companies invest in software that specializes in one of these areas, only to find out that it can't perform well in the others. This ultimately keeps data harmonization from being as useful and efficient as possible. A second concern is making sure you invest in software that will be able to properly handle harmonization at scale. The last thing you want is to go through the process of purchasing and onboarding a software package, only to find that it becomes partially obsolete shortly afterward as your business expands.

Next, we can begin talking about the actual process of data harmonization. The first step, whether done internally or with a third party's help, is to move everything into a common database. This is the data integration process that we mentioned earlier and has three steps: extraction, transformation, and loading (ETL):

- Extraction reads the data in the original data set
- Transformation converts the format
- Loading writes it into your new database

This is then followed up by centralizing the data, which means moving all of it to a key central location. While ETL gets things physically in one place, that's different from centralization, which blends all the information together. This is generally accomplished with [machine learning](#).

After this, you will need to revisit that data and audit it for any misleading or inaccurate information. This will keep a single bad piece of data from spoiling the entire set. The next step is going to be converting everything into a set of synchronized, relational data points. This means that whether the data is covering products, time, consumer segments, or geographies, all data points need to easily correspond to each other.

Another step you want to take during the harmonization process is introduce new custom classifications specifically for eCommerce. Think of these as the name for fields of each of your data, making it possible for you to segment/filter it in detail. For example, say that your company was rolling out a summer marketing promotion for a new food product line, including promotions on Facebook and YouTube. On Facebook, you may have labeled the relevant data as "Summer2020MarketingRollout". On YouTube, it might be "MarketingPlanSummer2020".

These may all be under the same umbrella, but without a clear point of reference, it can be difficult to interpret. Therefore, you would need to change the classifications to only one so all the data could be organized together. That said, you can also introduce new classifications to filter things down even further, like "Summer2020Marketing-YouTube" and "Summer2020Marketing-Facebook".

How Can You Implement Data Harmonization In eCommerce?

What are the actual components of eCommerce data harmonization?

- Moving data through extraction, transformation, and loading.
- Centralizing the data to blend all the information into one greater set.
- Auditing the data to remove any inaccuracies.
- Transforming the information into synchronized data points.
- Adding new classifications/removing previous inaccurate ones.



Companies need to invest in software and processes that can handle these tasks at scale.

Case Study: Tracking Amazon and Brick & Mortar Sales

In the new post COVID-19 era, understanding sales from all outlets, including eComm, is critical for CPG brands as sales continue to shift online. However, integrating eComm and brick-and-mortar sales data for reporting is extremely difficult.

The Challenge: A major CPG manufacturer recently asked: How can we best adapt to COVID-19 category changes and develop a cohesive strategy across Amazon/eComm and brick-and-mortar?

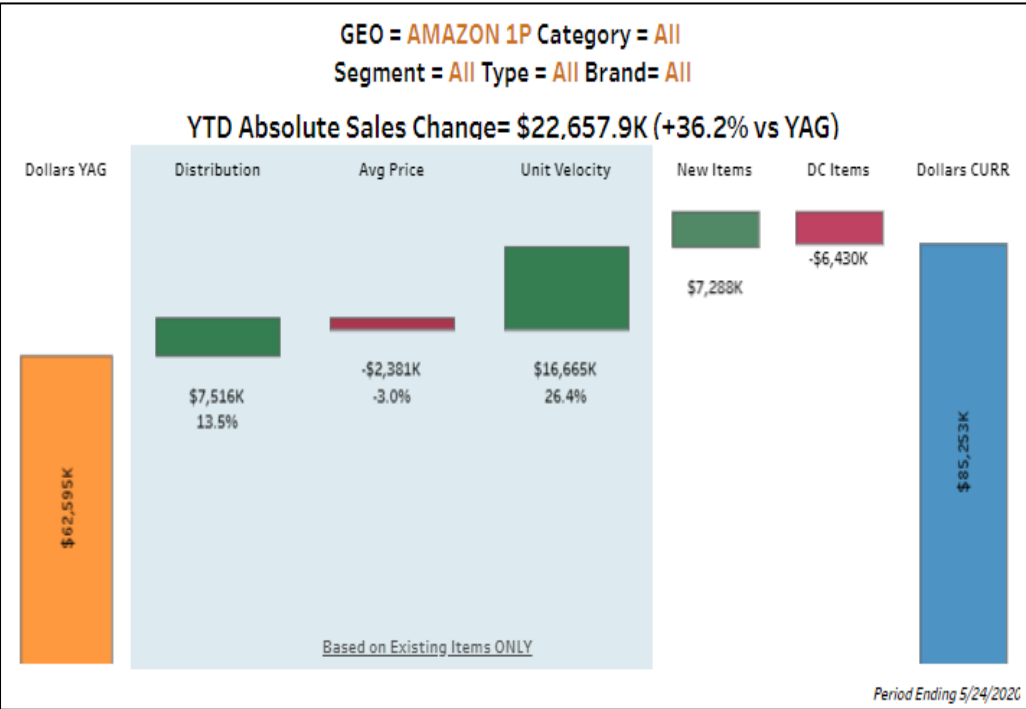
The Solution: TABS Analytics integrated Amazon’s syndicated scanner data with the client’s “brick-and-mortar” syndicated data to provide a holistic view of how the category was performing and understand the channel shifts between brick-and-mortar and Amazon and the key drivers behind those shifts.

While the manufacturer’s sales at Amazon increased at a double-digit rate over the past several years, sales data was analyzed separately from brick-and-mortar, making meaningful analysis slow and challenging. TABS leveraged its data harmonization capabilities – combining both machine learning and “human intervention” – to build out a master item database harmonized across all geographies, products, time periods and measures. From there, TABS deployed both Excel® ([QuickTABS®](#)) and cloud-based ([Marketing Insights PRO™](#) and [Account Insights™](#)) deliverables so the manufacturer could analyze the harmonized data across all key business dimensions. Proprietary distribution-based analytics provided further clarity to show “why and where” growth was coming from: distribution, organic velocity and price/mix.

Key Learning: Without a harmonized data set from which to “pivot” analysis from, understanding the growth between channels and the drivers behind them become extremely challenging at best.



By harmonizing data into one database and building analytics reporting from it, a full competitive and brand-focused analysis can be properly done.



How to Understand the OmniShopper

With the growth of OmniChannel shopping behavior – consumers switching between shopping online and in-store – data harmonization is now crucial for CPG companies across categories and sectors to understand the OmniShopper. However, many different elements need to be in place to make data harmonization happen. This is where companies like [TABS Analytics](#) and our sister company, [Decision Insight](#), are critical to help support your efforts. Here's a rundown of some of the services that we offer:

- Creating full data architecture to help with harmonization and customization.
- Develop a data harmonization methodology that converts item-level data and value metrics at the UPC-level, employing both advanced analytics (Machine Learning, AI) and human-specific analysis of datasets.
- The [Digital.IQ platform](#) uses virtual environments to quickly test and optimize strategies prior to in-market rollout.



How Test & Learn Research Works

Test eCommerce or in-store strategies and tactics with speed, producing sales results and shopper metrics without the risks and barriers associated with in-market A/B testing. Participants enter a virtual store or website where they shop as they would in real life. The virtual shopping experience is followed with rich diagnostics to understand the specific elements of a strategy driving success and those that can be improved.



Use Insights to Develop OmniShopper Hypotheses

Test and optimize OmniShopper strategies prior to in-market rollout – online or in-store – using Virtual Shopping Research.

What You Can Test

Strategies to test before implementing across channels include:

- New Product Offerings
- Assortment and Product Mix
- Alternative Packaging/Pack sizes
- Channel and Retailer-Specific Promotions
- Price Optimization

Use our support to create a unified and simplified data strategy for your brand to understand the OmniShopper in all channels.



About TABS Analytics

Tabs Analytics, founded by Dr. Kurt Jetta in 1998, is a technology-enabled analytics firm servicing the consumer products industry. Our mission is to simplify and improve the way analytics is conducted through analytical innovation, which translates into a competitive advantage for our clients. TABS is the leading outsourced sales and marketing analytics firm in the consumer-packaged goods (CPG) industry.

For more information about TABS Analytics services or this white paper, please contact Robert Baldwin at robertbaldwin@tabsanalytics.com.

TABSAnalytics

About Decision Insight

Decision Insight, a division of TABS Analytics, is a leading shopper insights and retail strategy firm that uncovers shopper behaviors and motivations. Twenty years of partnering with leading CPG companies have groomed the team to anticipate emerging needs to deliver **solutions** that lead to activation, triggering higher shopper satisfaction and increased sales. Our Test & Learn research is forward-looking and based on shopper behavior.

For more information about Decision Insight's in-store and eCommerce research solutions, email Leslie Downie at leslie@decisioninsight.com.

**DECISION
INSIGHT**
A TABS Group Company