

DATA DEEP-CLEAN

THE FRESH FACE OF DATA

ACX IOM

INTRODUCTION

Data has become so ubiquitous in marketing, and brands' reliance on it has become so fundamental, that the "data-driven" in "data-driven marketing" has all but become redundant. Today, that's just...marketing. Or it's how marketing should be done, anyway.

Every brand knows that people want a connected, personalized and relevant experience across all channels. That's why you're transforming data into insights every day to make more informed marketing decisions. You're using data to identify who your key customers are and what they respond to, and how to speak to them in more personalized ways, for better results.

But there's a catch: just adding more data isn't the answer. The best marketing technologies in the world will only help you make bad decisions faster, if your data's no good.

IT'S ABOUT DATA QUALITY

Data quality is the foundation of a successful marketing program and a successful business. If your first-party data isn't clean, correct, up-to-date and sourced ethically, your messages won't reach the recipients you target. You'll spend more money for a lower response rate, reducing your return on investment. Can you really afford that?

Data hygiene may sound clinical, but it just means keeping your data clean. It's a series of processes intended to cleanse data to ensure it is complete, up-to-date, and without errors or duplication. Taking regular actions to maintain your data helps ensure your time and money are well spent -- and you get the best possible downstream results whether it be data enrichment, identifying households or activation.

Only **33%** of marketers are completely satisfied with their ability to generate more relevant experiences using customer data.

Less than half are satisfied with their data's quality/hygiene **(42%)** and its completeness **(40%)**.

¹ <https://www.salesforce.com/form/state-of-marketing/>

IT'S ALSO ABOUT VALUE EXCHANGE

When it comes to customer data, brands today face a delicate balancing act. On the one hand, the ability to create exceptional digital marketing experiences relies on your ability to understand people as well as possible – and that requires lots of data.

On the other hand, you want to minimize the burden you're putting on people when you ask them to provide that information or to engage with your brand. It's called a value exchange, after all, and there are always two sides involved.

DATA HYGIENE ISN'T A ONE-AND-DONE JOB

The cleanliness of the data you have right now is only part of the story. As your use of data grows and evolves, maintaining data quality becomes even more important. Think about the CORE areas where data quality matters, when it's time to:

Collect
Optimize
Resolve
Enrich

TIME FOR A DATA DEEP CLEAN

This guide will lay out best practices on how to clean and maintain quality marketing data. Let's get started!



WHAT IS DATA QUALITY ANYWAY?

If positive business outcomes rely on your ability to engage people with exceptional marketing experiences, and these experiences are built on customer insights, then the foundation of all of this is data.

But not just any data. It must be correct and it must be maintained on the most basic level – like verifying a prospect’s name, address and phone number are correct.

Companies also have a responsibility to make sure they are doing the right things with this data – being accountable and collecting only what they need. An ethical approach to data use can have a long and lasting and potentially profitable impact on your brand.

DATA QUALITY RELIES ON DATA HYGIENE

Quality data regularly goes through data hygiene. It is “clean data” – and that means a number of things.

It’s standardized, updated, and void of redundant information. It’s correct, complete, and current (more on this later). It’s considered relatively “error free” and, ultimately, it’s what powers your business.

Easy enough, because all data collected from customers can be assumed to be “quality data,” right? Wrong! Dirty data can sneak into your marketing process in a variety of ways.



CAUSES OF LOW DATA QUALITY

- 1. MISSING DATA:** Form fields left empty, either intentionally or unintentionally, that should contain customer information.
- 2. OUTDATED DATA:** Customer information that was correct when input but has since changed.
- 3. MISLABELED DATA:** Data that has been entered into the wrong field and is assigned to an improper element.
- 4. NON-CONFORMING DATA:** Data that hasn’t been normalized.
- 5. DUPLICATE DATA:** A single account, contact, lead, etc. that occupies more than one record.
- 6. CUSTOMER ERROR:** Misspellings, typos and variations in spelling, naming or formatting.
- 7. INTENTIONALLY INCORRECT DATA:** Data that contains false information. This can range from a fake name to nonsense and profanity purely intended to fill a required field.
- 8. NON-ETHICALLY SOURCED DATA:** Data that could cause harms, such as social embarrassment, reputational damage, bias, or discrimination.



BEWARE BAD ACTORS

When we talk about bad actors, we mean those who deliberately provide faulty or misleading information. For example, many marketers have found Mickey Mouse in their data. But it's not always a goofball play. It can involve profanity or deliberate harm to a brand, it can be someone using someone else's details without permission, and it can even come from inside an organization, when a customer service rep tampers with a customer's data accidentally or on purpose.

The fallout from actions like these cost brands time and money to fix, but the real damage is often reputational, and harder to remedy.

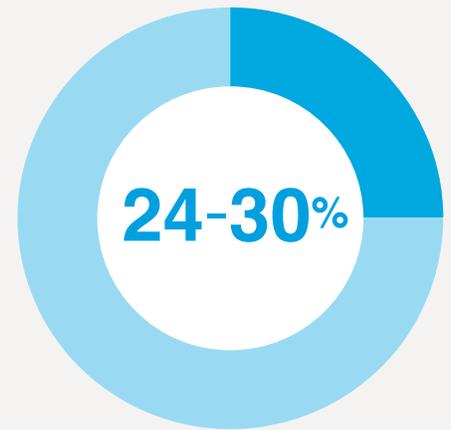
ASK YOURSELF:



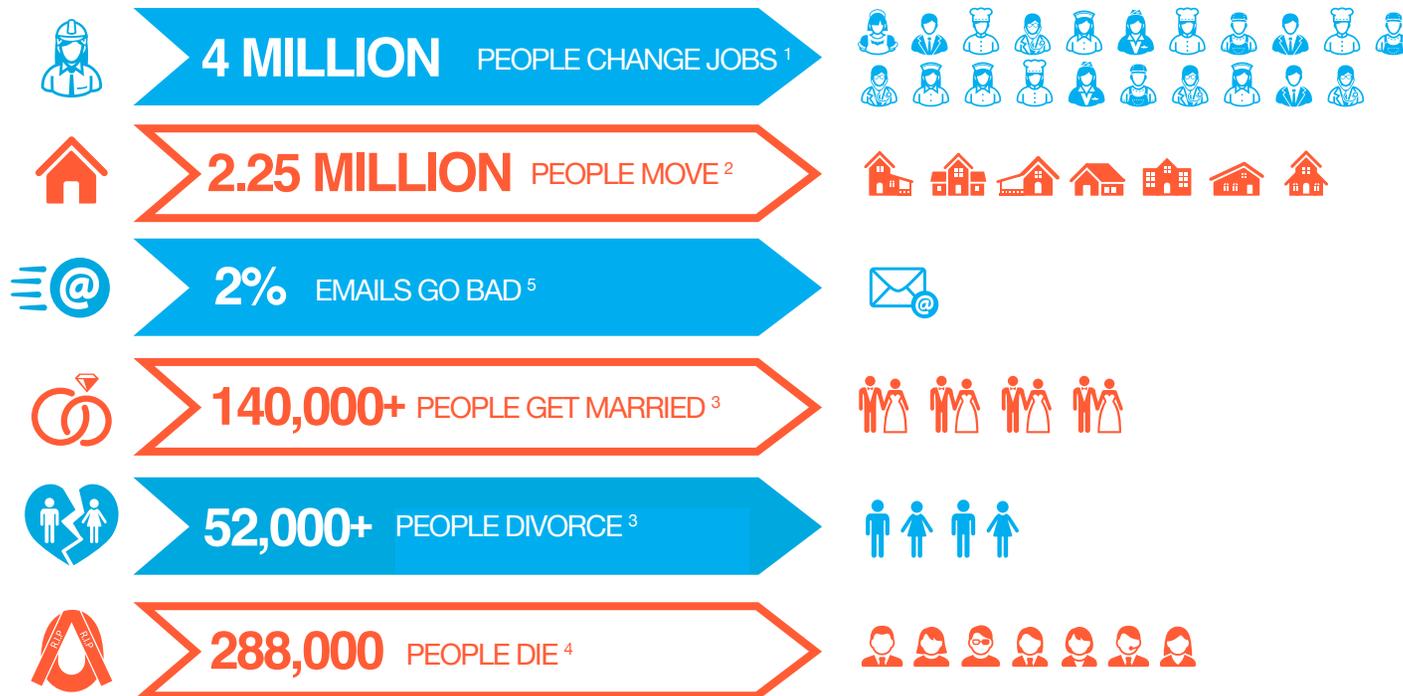
- Do you know your data's overall quality?
- Do you understand the return on investment you could achieve with clean data?
- Do you know how your data's quality compares to your competitors' or others in your industry?
- Did you know data quality was even something you'd have to be concerned about?

IF THE ANSWER TO ANY OF THESE QUESTIONS IS “**NO**,” THEN KEEP READING.

Making things worse is that even if your data is accurate when it's collected, it's nearly impossible to keep current without proper vigilance. Customer data lives in a constant state of flux, changing at a rate of 2 to 3 percent per month. Doesn't seem bad, right? That is, until you do the math and realize that could add up to 24-30 percent of your data every year.



Let's put things into perspective: Each MONTH in the U.S.



This means that a considerable percentage of the names, addresses and personal information you just collected from your customers will be out of date in a year.

¹ <https://www.pewresearch.org/social-trends/2022/07/28/majority-of-u-s-workers-changing-jobs-are-seeing-real-wage-gains/>

² <https://www.census.gov/library/stories/2022/03/united-states-migration-continued-decline-from-2020-to-2021.html>

³ <https://www.hubspot.com/database-decay>

⁴ <https://www.cdc.gov/nchs/fastats/marriage-divorce.htm>

⁵ <https://www.cdc.gov/mmwr/volumes/71/wr/mm7117e1.htm>

ALL YOUR HARD WORK WILL BE GONE IF YOU DON'T TAKE STEPS TO PROTECT IT!

Poor data quality has a ripple effect. It creates competitive disadvantage, bad strategy, lost productivity, poor customer relationships and financial loss. Interviews with data scientists suggest they spend 50-80 percent of their time performing what they lovingly call “data janitor work.” It’s the routine, painstaking, monotonous tasks that go into keeping data clean.

However, by focusing on data hygiene — **maintaining your good customer records and fixing or appending incorrect ones** — businesses can see big results.

The following are two examples of businesses that saved money and improved their business reputation simply by improving the quality of their data:



THE BANK THAT SAVED MILLIONS BY REDUCING CUSTOMER DUPLICATES

Imagine that you are one of the largest financial institutions in the United States, serving millions of customers. Now, imagine it's your job to keep all that customer information correct, accurate and up to date!

This was the problem one of Acxiom's financial services clients faced. With nearly 20 systems housing customer data, the bank struggled to maintain consistent customer records. Duplicate records and inaccurate customer information inevitably led to the bank offering customers products they already had— straining resources but also hurting the brand's reputation.

To fix it, the bank implemented a company-wide recognition process (we'll talk more about what this means in a moment) to clean up its customer data and improve the integrity of its records. That helped the bank:

- **Reduce the number of customer duplicates by 67 percent**
- **Save millions of dollars**
- **Cut in half the number of employees dedicated to fixing customer data quality by using a third-party system to manage the process**

THE DEBIT CARD ISSUER THAT REDUCED DIRECT MARKETING SPEND BY \$1 MILLION A MONTH

A leading provider of prepaid debit cards in the United States diversified its business model by offering cards directly to customers instead of using third-party partners. However, it was concerned about mailing the reloadable cards to consumers who were already customers. That would come at a substantial cost and also wouldn't paint a great brand image with customers.

To remedy this problem, the company used Acxiom's hygiene services to consolidate and eliminate duplicate records. The company also was able to take advantage of Acxiom's identity services to follow and engage with customers as they move through the customer lifecycle.

The result? A reduction in direct mailing costs of \$1 million per month and a 50 percent improvement in ROI!

THE NEW FACE OF QUALITY DATA MARKETING

Understanding the difference between “clean data” and “dirty data” and the effect it can have on your bottom line is just the first step. This alone already puts you ahead of many of your competitors. But as you begin to understand that a difference exists, do you know what good marketing data actually looks like?



THE 4 Cs OF DATA QUALITY

What does it take for data to be considered clean and ideal for marketing use? Regardless of what the data is, there is a common set of practices that define data quality for all businesses, which includes ethically sourced data as a key component of clean data. We believe that quality data meets the “4 Cs” standard.

Let's dig deeper to understand what's involved in each of the four data quality components.





IS YOUR DATA COLLECTED ETHICALLY?

Collection and use of data should go beyond what is required by law and include policies and procedures that ensure data uses are fair and just to consumers. Data should be used to benefit both parties in a transaction, resulting in the personalized experiences that delight customers and the increased profit that delights shareholders. How do you make sure data delivers the right results for everyone? You need repeatable front-end processes for identifying the sources of data, the provenance of data, how it will be protected, and how it will be used to deliver the experience customers want. And you can prioritize the use of trusted data providers who are committed to ethical practices, for external third-party insights that combine with first-party data for a better overall understanding of the customer.



IS YOUR DATA CORRECT?

QUALITY DATA IS NORMALIZED

Normalizing your data means making sure that all of the data elements of the same type (all of your phone numbers, all of your home and email addresses, etc.) look like one another and follow the same format. Phone numbers, for example, serve as a great model.

All phone numbers in the United States contain 10 digits. If we were manually normalizing this data element, our first step would be to check that each one had the proper number of digits – calling out examples that were missing the area code or where a number may have been left off or added. We would also want to ensure the phone numbers were all presented the same way. For example, if you were to look at your database, you might find phone numbers presented in the following formats:

1235551234 • (123) 555-1234 • 123.555.1234

Because these numbers are all presented differently, they are not normalized. It doesn't matter how you choose to represent the information, just that they are presented the same way throughout your database.

QUALITY DATA IS ACCURATE

At its core, quality data is accurate. This means making sure the data all looks the same and checking the data on a fundamental level to see if it really is correct.

For example, continuing with phone number data, in the accuracy step, you must check to see if the phone number is actually in use. There are a number of tools available to help marketers determine if their information is accurate by scrubbing the data you have against a known data set. However, most commonly these tools are only able to tell you if the information you have exists – for example, if the telephone number you have in your records is a working telephone number. They cannot tell you if the number you have for Mary Smith really belongs to Mary Smith.

ACCURACY IN USE

Jane was recently on your website and opted in to receive direct mail from your dog grooming company – updates, events and coupons. Unfortunately, she didn't notice she made a mistake when filling in her address on your online form. She meant to list her address as 462 Sand Creek Avenue, not 426 Sand Creek Avenue. Now the incorrect customer record is in your database. If left uncorrected, Jane could miss out on some of the promotions sent to her.

However, Acxiom hygiene services will flag and fix the address, realizing that even though the address provided exists, it's not likely where Jane lives. This will allow you to reach Jane instead of wasting dollars targeting her neighbor (who doesn't even have a dog). That's a marketing win!



IS YOUR DATA COMPLETE?

WHOLENESS

Once data has been normalized and a level of accuracy established, it should be examined to ensure it is also complete. Staying with our phone number example, once normalized, we can easily examine the number to see if all the elements are there.

For example, if we notice the data element has 9 digits instead of 10, we can look further to identify if it is simply missing a number or if, perhaps, the wrong information was stored as a phone number when it's really an address or a customer identification number.

VALIDITY

Validity is similar to our accuracy step in that it's all about checking the information to see whether or not it falls within an accepted range of quality.

For example, a phone number can be checked against known rules to see if the area code is valid. Email addresses can be checked for the correct format. Or a ZIP Code can be checked to make sure it has the right number of digits and that it matches the city or town also listed.

HELP IS AT HAND

If your data is incomplete, don't worry – this is the case for the majority of companies. Thankfully, many third-party tools exist to help marketers append or uncover additional information.



IS YOUR DATA CURRENT?

CURRENCY

Finally! Your data looks like quality data – it's correct and complete, but is it current? Does Jennifer Jones still live at the address you have or has she moved? Is her last name still Jones or was she married last year and took her husband's last name? Is her phone number still correct?

Once you know your data is valid, the next step is to make sure it is current. If it isn't, this could lead to issues ranging from a wasted attempt to contact a customer (that gets expensive) to a potential customer that you're contacting without her permission.

IDENTITY RESOLUTION

The favorite retailer's spring catalog has arrived. You're pretty excited. Then, you realize the retailer sent your household four copies – one addressed to you, another addressed to your spouse, a third addressed to your mother-in-law, and a fourth addressed to your teenage daughter.

TOUCH

At best, you shake your head at the wasted mailing costs. At worst you lose a little bit of confidence that this retailer really knows who you are or is even paying attention. And what if you're the retailer? You just quadrupled your mailing cost. Now multiply that across your entire prospect list.

Identity resolution helps businesses avoid embarrassing and costly mistakes. This process saves money and helps create stronger customer portraits by merging multiple customer records to create a more comprehensive picture of each customer.

ENSURING YOUR DATA MEETS THE 4 Cs IS HUGEY IMPORTANT.

4 TIPS FOR MAINTAINING DATA QUALITY

We can't say it enough – data quality is the foundation for effective marketing. You can't work to build targetable customer portraits if you don't first know who your customer is. To be good marketers and to increase business value, we must all become vigilant about improving data quality.

Here are four tips for improving data quality. The more proactive you are about data quality and the more you plan for it, the greater the ROI you'll achieve!

1

CLEAN ONCE

You have tons of customer data collecting dust in your marketing database. Don't wait to start cleaning existing data – act today to start benefiting from the results! It is likely you have lots of outdated information. Thankfully, there are many solutions to clean your data. Many companies offer a service that allows data to be sent, processed and returned quickly.

2

CLEAN EARLY

Correcting a record before it gets in your customer database is in your best fiscal interest. When the data becomes “official,” it becomes exponentially more expensive to correct or remove.

Businesses should implement a process for cleaning data as it comes in. This includes defining the standards for what complete and normalized records look like so the database grows in a healthy way. The best ways to do that are through a real-time process and preventive measures. Creating form requirements such as mandating an @ symbol in an email address field can help minimize mistakes and false data.

3

CLEAN OFTEN

Because data quality decays so quickly, data must be processed regularly to minimize data quality issues. As with the initial cleaning, data can be sent to a third party for processing or cleansed via an in-house, on-premise data quality tool.

4

CLEAN ETHICALLY

Going beyond legal and regulatory guidelines, data must be managed, integrated and activated based not only on what “can” be done in the market but also on what “should” be done. Data utilization should carry an expectation and obligation of responsible stewardship. Organizations should adhere to all relevant data laws, pass annual third-party audits of their security practices and be required to notify people of a breach within a specified amount of time.

The best customer experiences are built on the best data.

Make sure your business decisions are based only on high-quality data. Our team can help you get started.

Visit [acxiom.com/stopdatadecay](https://www.acxiom.com/stopdatadecay) or email info@acxiom.com.

ABOUT ACXIOM

Acxiom provides data-driven solutions that enable the world's best marketers to better understand their customers to create better experiences and business growth. A leader in customer data management, identity, and the ethical use of data for more than 50 years, Acxiom now helps thousands of clients and partners around the globe work together to create millions of better customer experiences, every day.

Acxiom is one of the largest providers of data hygiene services in the industry, processing more than 2 Trillion records annually. Acxiom's data offerings are ethically sourced and developed using hundreds of data sources, each carefully evaluated, screened, and monitored for appropriate data collection and privacy policy practices including proper consumer notice and choice.

Visit [acxiom.com/stopdatadecay](https://www.acxiom.com/stopdatadecay)
or email info@acxiom.com to get started.

The Acxiom logo features the word "ACXIOM" in a white, sans-serif font. The letter "X" is replaced by a blue graphic element consisting of two overlapping, slightly offset rectangular shapes that create a sense of depth and movement.