



# How Clean Is Your Salesforce Data?

## SUMMARY

- *What impact does duplicate, incomplete, or missing data have on meeting your business objectives?*
- *How big is this problem?*
- *How often do you need to clean the data in your CRM so you can be confident the decisions you make are accurate?*

This white paper will answer these questions and more, explaining why data quality is mission-critical to your enterprise.

We will explore ways to solve these challenges and help you mitigate risk.

## INTRODUCTION

### Poor data quality is costly in more ways than one.

According to Gartner research, organizations believe the inefficiencies and missed opportunities arising from poor data quality cost them an average of \$15 million per year.<sup>1</sup>

With such high costs associated with poor data quality, erroneous business decisions being made because of it, and company reputations negatively impacted by it, it's no wonder improving data quality has become so critical to business success.

### Good quality data in your CRM is key.

Data administrators entrusted with Customer Relationship Management (CRM) systems can become champions for their organizations by taking the appropriate actions to clean and continually maintain the integrity of their data. But they need the right tools for success.

While CRM systems are great for managing customer relationships, they do not automatically yield high-quality customer data. Instead, they're more like an empty filing cabinet. Each user must decide what they expect to file, what system they intend to use, how they will maintain this system, who will have access, and who will be responsible for keeping these files in good order.

This white paper explains why organizations should invest time and resources into improving the quality of data in their CRM systems and presents a simple plan of action to establish and maintain a higher level of data quality.

## WHAT DOES POOR DATA QUALITY MEAN?

**Poor data quality derives from a variety of sources and causes.** It typically falls into one of these categories:

- **Missing data**

Empty fields that *should* contain data. Example: An automated billing process breaks down because a customer's billing address is missing from the system.

- **Wrong or inaccurate data**

Information that has not been entered correctly or maintained. Example: Bills come back marked “Return to Sender” because the ZIP Code doesn’t match the billing address.

- **Inappropriate data**

Data that’s been entered in the wrong field. Example: A promising lead doesn’t appear in regional reports because the city name is in the *State* field.

- **Non-conforming data**

Data that hasn’t been entered according to the organization’s naming convention. Example: A critical report on all U.S. clients doesn’t include clients with “USA” or “United States of America” in the *Country* field.

- **Duplicate data**

A single Account, Contact, Lead, etc. that occupies more than one record in the database. Example: Sales reps in different regions unknowingly create three distinct IBM Accounts, wasting valuable time, sowing confusion, and causing missed opportunities.

Fundamentally, poor data quality is defined by its *potential to cause undesirable interruptions in the normal flow of business activities*.

## WHY DOES DATA QUALITY MATTER SO MUCH?

### Bad data is expensive.

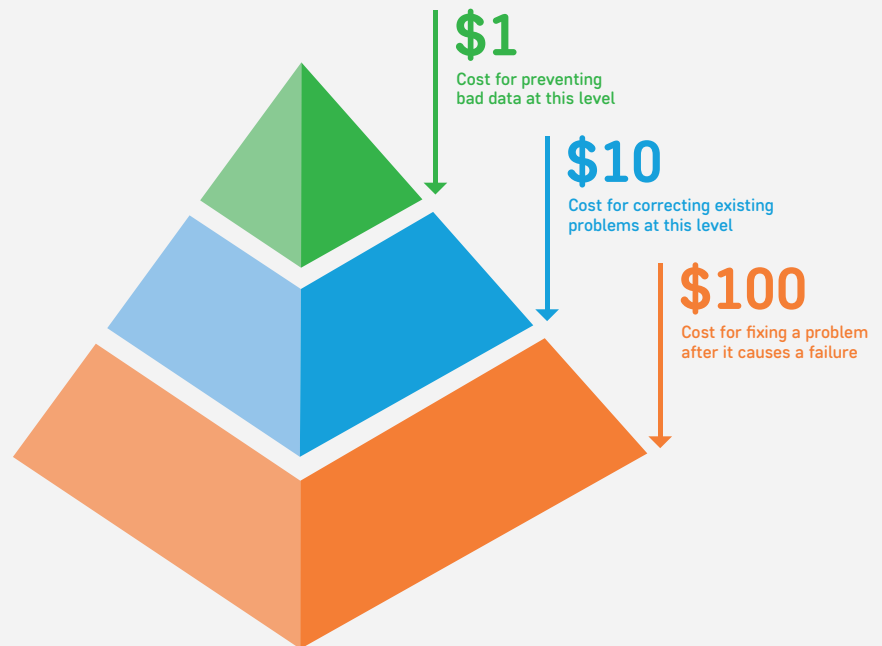
The multiple impacts of poor data quality, are not solely financial; businesses can see loss of reputation, missed opportunities, and higher-risk decision making as a result of low confidence in data.

According to the 1-10-100 quality principle, the relative cost of fixing a problem increases exponentially over time. If the cost of preventing bad data from entering a CRM system is \$1, then the cost of correcting existing problems is \$10, and the cost of fixing a problem after it causes a failure, either within an organization or with a customer, is in the neighborhood of \$100.

As the saying goes: *garbage in, garbage out*. Poor data quality left unchecked leads to:

### 1-10-100 quality principle

According to the 1-10-100 quality principle, the relative cost of fixing a problem increases exponentially over time



**Figure 1:**  
The 1-10-100 quality principle explains the incremental cost of bad data.

- #### Bad business decisions

Bad data produces misleading CRM reports and dashboards that sabotage the best efforts of decision-makers to guide the enterprise.

- #### Inaccurate sales forecasting

Sales stats derived from bad data can be overly optimistic or pessimistic, causing frustration among the sales team.

- #### Unsatisfactory customer service

When customer service representatives rely on incomplete or incorrect information, they're unable to deliver a top-notch experience to customers.

- #### Damage to reputation and brand

An organization that consistently relies on faulty CRM data can become notorious for not meeting customers' needs and expectations.

- **Decline in user adoption**

If CRM users don't have confidence in the organization's data, they'll abandon the system in favor of other, less efficient solutions such as individually maintained spreadsheets or local databases.

- **Wasted time and money**

When customer data is inaccurate, organizations launch inefficient direct mail/email campaigns, produce off-target marketing materials, and generally underperform in sales and marketing activities.

## Case Study: Direct Mail Campaign

A successful company specializing in automotive data and marketing solutions regularly sends out marketing packages to car dealers across North America.

***How bad data affected one of their direct mail campaigns:***

- 4000 packages @ \$20 each = \$80,000.
- 500 bad addresses wasted over \$10,000.
- 500 pricing errors resulted in people receiving the wrong package = another \$10,000.
- 1,000 duplicates = more wasted time and money.

***Outcome:***

**More than \$20,000 wasted because of bad data.  
That's 25% of the total expense!**

## IS IT SAFE TO DRINK FROM THE CRM LAKE?

One way to think about the vital importance of CRM data quality to a business enterprise is to think of data as drinking water. Everyone wants clean, filtered, unadulterated water that contributes to the health, welfare, and success of a community (the enterprise), while poor-quality water makes people sick, negatively impacting the community and forcing citizens to abandon tap water in favor of other solutions.

In Chapter 10 of his excellent book *Data Quality: The Field Guide*, renowned data-quality expert Tom Redman draws a vivid analogy between a database and a lake, where the lake water represents (CRM) data and pollutants represent bad data.<sup>2</sup>

For the sake of illustration, let's assume that pollutants enter the lake from only two sources: an inflowing river and rainfall. The inflowing river refers to mass data imports. This abundant data stream enters the database, bringing with it duplicates and other pollutants. Rainfall refers to individual database transactions such as manually entered records, Web-to-Lead forms from websites, and records added by integrated systems through the CRM API.

A nearby community (representing CRM users and, more generally, the enterprise itself) relies on the lake for its drinking water, so high levels of pollution pose a serious public health threat. Therefore, the town must find a way to remove pollutants from the lake water and make it fit for consumption again.



**Figure 2:**  
A CRM database  
is analogous to a lake.

## TAKING ACTION TO CLEAN UP THE LAKE

**It does not take long for the community to identify where most of the pollution originates.**

Think about a factory that's dumping chemicals (representing "dirty" lists) or releasing airborne pollutants from several smokestacks (representing lax data input controls) into the river that then enters into the lake. Identifying and understanding the source of the pollution problem allows for implementing a plan of action.

Enacting environmental laws (enforcing data input and import standards), can effectively halt the dumping of chemicals, drastically reducing the amount of pollution (bad data) entering the lake via river water. Going forward, the factory must agree to release only treated water ("clean" lists) into the river. These regulations also must compel the factory to install scrubbers (data input controls) on smokestacks, significantly reducing the amount of airborne pollutants entering the lake via raindrops.

The community needs to upgrade their water purification plant (data quality tools) to help clean up the entire lake. Although some pollutants will still enter the lake, the community is now capable of managing the problem and maintaining a high level of water (data) quality.

## WHO'S RESPONSIBLE FOR DATA QUALITY?

**CRM data quality is inextricably linked to the performance and well-being of an organization,** just as water quality impacts the health and wellness of a community. Data quality has obvious enterprise-scale implications, but with so many stakeholders, who's actually responsible for achieving and maintaining a desirable level of quality?

- The sales rep who manually enters customer data?
- The marketing manager who periodically mass imports leads?
- The senior managers responsible for setting company policy?
- Or the CRM administrator who manages the database?
- Of course, the right answer is "all of the above."

**Sales reps** should carefully enter customer data according to naming conventions...but mistakes will be made.

**Marketing managers** should do their best to scrub new lists before importing them into the database...but some bad data will still make it through.

**Senior managers** should establish and maintain reasonable company policies regarding data quality...but they won't be perfect, and they won't always be followed.

Although the whole team is expected to exert some effort, inevitably it's the **CRM administrator** who must take the leadership role on data quality and make sure the job gets done right.

## CLEANING UP THE ENVIRONMENT

### A CRM database is in a constant state of flux.

Users routinely add new data and modify existing records. To achieve and maintain a high level of data quality, CRM administrators must take action on two fronts:

#### 1. Prevention

Minimize the amount of bad data entering your data repository via individual transactions and mass imports.

#### 2. Remediation

Constantly monitor and cleanse data to a well-defined quality standard.

This two-part approach will enable data administrators to get ahead of the quality curve and stay there.



## Part 1: Prevention

There are many approaches to preventing CRM pollution and maintaining a clean environment. Here are a few suggestions administrators can follow to achieve higher quality data input and maintain data integrity.

### Educate users

To maximize the quality of data entering the system, CRM administrators should instruct users on how to perform routine tasks such as:

- Searching for duplicates before entering new data.
- Entering data that conforms to the standard naming convention.
- Completing all fields in each record.
- Using available data quality monitoring tools.
- Importing data properly (for users with access).

More generally, administrators should seize every opportunity to educate users and managers about the importance of data quality to the organization as a whole, as well as to individual employees (e.g., in relation to job performance).

### Enforce data quality standards

Administrators must design schemas with data quality in mind. They should:

- Define required fields.
- Use automatically populated default values whenever possible.
- Create field dependencies and workflow rules (e.g., if A and B exist, then C must exist).
- Control object creation (what users are allowed to create Accounts, Contacts, Leads, etc.?).
- Implement validation rules to ensure that data is entered correctly.
- Impose restrictions on Web-to-Lead data.

### Improve internal communication

The data administrator is generally responsible for answering questions and notifying users of changes to the CRM user interface (e.g., the addition of new fields or entire screens), revisions to naming conventions, updates to policies and standards, and anything else related to the user experience.

Prevention will be most effective when the administrator can foster a supportive atmosphere where users are comfortable asking questions, discussing problems, and suggesting improvements.

## Part 2: Remediation

Once the data in the CRM system is clean, it is critical to continually maintain data quality at a high level. This can be accomplished through several strategies outlined below.

### Monitor the database

Anyone who's maintained even a small database will agree that data quality degrades very quickly and exponentially. While updating existing records, users sometimes replace valid information with erroneous data, or simply change/delete information by accident. New records, whether entered manually or imported, invariably contain a certain number of problematic fields, despite an administrator's best preventive efforts. The approach is to continually check the database, identifying and correcting erroneous data.

### Develop a data-quality regime

#### Standardize

Every data administrator should develop a quality standard that defines "bad" data in their database. The standard is really a collection of rules or tests that, when applied to the database, identify bad data and, in some cases, automatically fix it. After establishing an initial standard, CRM administrators should continuously seek to improve and update it so that quality remains at a high level no matter how quickly the database evolves.

- **Cleanse**

Having defined a quality standard for their database, data administrators must now implement it using data cleansing tools. With these tools, CRM administrators can search the database for records that don't conform to the quality standard and fix them. To maintain a consistently high level of data quality, automated searches and merges should be conducted on a set schedule—perhaps daily or weekly.

- **Enrich with outside data sources**

Even if users could enter or import 100% clean data, it wouldn't change the fact that the world is a dynamic place. Companies grow, people change jobs. Contact information that was valid just three months ago could now be out-of-date if a customer moves their head office to a new location. Whenever possible, CRM administrators should check their data against credible outside sources.

- **Deduplicate data**

Duplicate records—dupes—create confusion and make it more difficult for users to get a complete view of a customer relationship. CRM administrators should use a duplicate management tool to regularly search for dupes in the database and merge them, and a duplicate prevention tool to stop dupes at the point of entry to the system (e.g., Web-to-Lead).

- **Validate data quality**

After performing other remedial tasks, CRM administrators should verify that records have been properly updated and the database does indeed conform to the quality standard.

Performing these best practices consistently will result in data all users can rely on for critical business decisions and the future growth of the business.

Successful organizations recognize and value the importance data quality has for future growth. They have implemented many of the approaches discussed in this paper and have kept on top of the curve when it comes to ensuring data integrity, ultimately trusting their data for the most critical business decisions.

## CONCLUSION

### **No company can survive without customers.**

In today's customer-centric and data-driven business world, organizations cannot survive without customers. And no company is immune to bad data. Since poor data quality is a significant root cause of customer relationship problems, data administrators must be vigilant and proactive about keeping their databases clean.

Although every member of the sales and marketing team should do their part, the CRM administrator must lead the way by developing and enforcing a quality standard, designing schema from a quality perspective, training team members in the CRM application, providing educational resources, and maintaining open lines of communication.

Starting with the basic action plan outlined in this paper, and following through with the right software tools, data administrators can effectively eliminate poor data quality as a threat to their organization

 REFERENCES

1. Moore, Susan. "How To Create a Business Case for Data Quality Improvement," Gartner, June 19, 2018.
2. Redman, Tom. *Data Quality: The Field Guide*. (Woburn, MA: Digital Press, 2001), 53–55.

## Learn More

To arrange a demonstration or simply learn more about Validity's suite of data quality solutions:

- Visit our home page at [www.validity.com](http://www.validity.com), or
- Drop by the [Salesforce AppExchange](#) and search for Validity, DemandTools, PeopleImport, or DupeBlocker.

The AppExchange has hundreds of insightful customer reviews of our products, so be sure to check them out.

---

### About Validity

Validity is a leading global provider of data integrity and compliance offerings that thousands of organizations worldwide rely on to trust their data. With highly-valued products including DemandTools, DupeBlocker, and PeopleImport, Validity is empowering organizations worldwide to make better decisions that drive more leads, close more deals, and confidently plan for continued growth. For more information visit, <https://www.validity.com>.