

Data Quality Issues Stopping You From Achieving Perfect Data

Did you know that 80% of all digital data is incorrect according to Observe Point? Indeed, with data becoming a major buzzword, data quality has been a point of interest for most data specialists. Superior-quality data is the ultimate driver of revenue for modern business. Good data can generate unprecedented lead conversion rates, account-based success, and closed-won deals. Poor quality, on the contrary, can significantly drop the ROI of a company's CRM and marketing automation investment.

With that being said, let us take you through the biggest reasons why data quality is still an issue even in 2020.



Duplicates

Data duplication is an episode when an exact copy of a data piece is created. To an unaware, this issue may come across as a no-brainer that any data expert would be able to sidestep. However, data duplication is a widespread concern. Thus, in healthcare, duplicate medical records are growing at a fast pace. This leads to patients often being mistreated. And we all know the number of risks it poses.

But who is to blame for data duplication? There are a few guilty parties:

- Nobody's canceled the human factor. You are likely dependent on your employees to fetch valuable data for you. But we humans get tired quickly and cannot press on with the same task for a long time. As a result, fatigue makes your workers enter multiple copies of the same data piece.
- Data duplication happens when you sew up data from various websites. To keep search engines happy, listings may be slightly altered. Therefore you won't be able to detect duplicates unless you turn to an advanced querying tool. Users' feedback.
- Duplicates are also common when you are fetching users' feedback. It's almost the same as the first one. Data duplicates are also common when you are asking for users' feedback. Like your employees, users are mistake-prone, although the reasons may vary.



Inconsistent formatting

Inconsistent data formatting is another issue that haunts most organizations. If the data is saved in inconsistent formats, the systems used to analyze the information may not interpret it as needed. If the company collects the database of their consumers, then the format for basic data pieces should be specified. It may be especially challenging for systems to

differentiate the US and European-style dates and phone numbers, especially when some have area codes and others don't.

Inaccurate data

Finally, it is pointless to carry out data analysis or interacting with users based on data that is just wrong. Genius, right? But don't be so quick to roll your eyes. If it weren't a common pain, this issue wouldn't make it to our list. Inaccurate data is generated for a number of reasons. This could be your customers providing erroneous information to a human operator. This could be details ending up in the wrong field. Incorrect data is especially tricky to detect, since entering an incorrect but valid phone number that conforms to the general formatting is almost impossible to detect.

The Bottom Line

Human error cannot be cured. But if you lean into clear procedures that are followed consistently, your data analysis will come to fruition. Also, automation tools help pare down the risks of mistakes by exhausted and bored workers.

On a similar note, we'd like to thank [Flatfile.io](https://flatfile.io) for sponsoring this newsletter. Flatfile.io is the drop-in data importer that implements in mere hours. It allows you to provide your users with the ultimate import experience and target the data pains.