

DO YOU TRUST YOUR DATA?

A WHITE PAPER.

TREEHOUSE
TECHNOLOGY
GROUP

DATA

Weighing out the risk-benefit analysis of actions is an inherent human condition. In prehistoric times people had to ask themselves, “Should I trade my food for this axe?”, “Is that person my friend or my enemy?”, “Should I fight or run from this predator?” and make decisions using the information surrounding them.

Fast forward to present day and business leaders might not wonder about trading food for tools, but do have to make strategic decisions about markets, investments, budgets, and projections. However, these decisions are only as good as the data used to make them. **A recent study from KPMG found that only 35% of executives say they have a high level of trust in their organization’s use of data** and that over a quarter admit to having an active distrust of company data.



So how do you **ensure your company is working with reliable data**? We've put together some suggestions that can help you proceed with greater confidence — or caution — based on the quality of your organization's data.

We've put together a list of four suggestions to increase the quality of your company's data.

- **KNOW THE SOURCE**
- **ENSURE HIGH QUALITY DATA INTEGRATION**
- **TRAIN YOUR PEOPLE**
- **ASSESS A SAMPLE**

#1 KNOW THE SOURCE

The origin of any big data ought to be made clear to those using the information in their decisions. Limitations in data accuracy should be made available along with any potentials for bias. Recommendations about the kinds of decision making the data can and cannot support are helpful in the decision making process as well.



#2 ENSURE HIGH QUALITY DATA INTEGRATION



This can involve some technical work, but **the goal here is to achieve the highest quality data possible.**

One way to do this is to verify that Joseph Daniels in one data set is the same Joseph Daniels in others. Along the same lines, ensure that there are not duplicates of the same entry that appear in different ways (say Joey Daniels or J. Daniels). Finally, confirm and define your units of measure within your data by checking that Joseph's purchases are not listed as "pallets" in one set and "units" in another.

#3 TRAIN YOUR PEOPLE

Sure, data evaluation is ideally interpreted by the experts in your organization. However, companies also need to make sure that managers and other decision makers (who might not necessarily be data scientists) **understand the consequences of misinterpreting data within their roles.**

Implementing training initiatives for all data users can help reduce the incorrect dissemination of information.



#4 ASSESS A SAMPLE

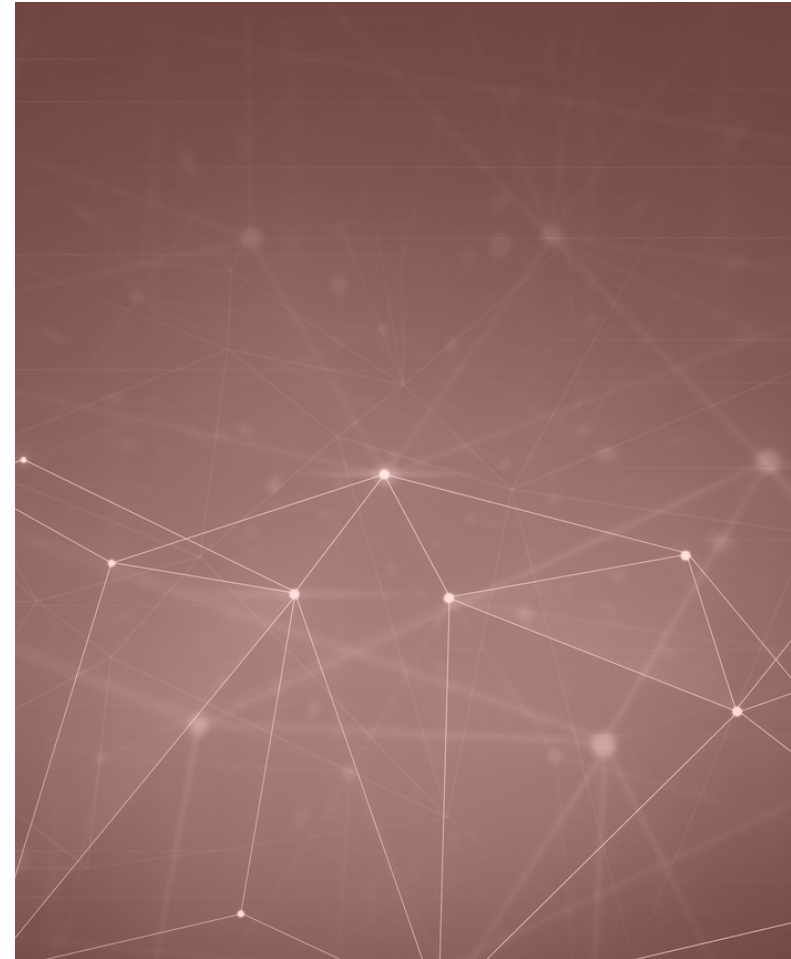
Try cleaning and evaluating a sample of your data. Start by eliminating erred records and data elements you cannot correct. Then go through your sample one input at a time to ensure it is as accurate and complete as possible.

When you finish scrubbing your sample, take a hard look. If the process goes really well, **you will find that you've created a trustworthy data set that you can be confident using to make decisions.** If the results are less satisfying, note that the data set should be used with caution. Finally, understand that there are going to be sets of data that are just plain untrustworthy. Through scrubbing you might find that there are just too many errors or there are too many holes in the information to use it with any confidence. In these instances, make sure you communicate that the sample strongly suggests that none of that data should be used to inform decisions.

NEXT STEPS:

Still concerned about the quality of your data or how data is being used to make decisions in your organization?

Contact **Treehouse Technology Group** for a complimentary consultation regarding your company's data systems, quality, and technology.



TREEHOUSE TECHNOLOGY GROUP

TREEHOUSE**TECH**GROUP.COM