Clean Data: The Foundation Of Every Analytics Project

Every analytics project starts with data. But in order to put that data to work—for analytics, machine learning, or standard reporting—it needs to be cleaned and structured. We've all heard the expression "garbage in, garbage out," and it's true. In a very real way, the quality of your data determines the quality of your analytics.

As a data preparation platform, we see the impact clean data can make to our customers' processes every day. These customer stories demonstrate how proper data prep is essential, and how clean data is foundational to accurate and robust insights.

Malwarebytes

DATA ONBOARDING

Like most companies, the marketing team at Malwarebytes collects leads from a variety of different sources, such as webinars, events, online forms, conferences, and sales representatives. Of course, none of these sources are standard across each other; just imagine the many ways a simple address can be entered into a database, to say nothing of cases where information is simply missing. All of this data needs to be imported into the team's CRM, and any errors can result in useless leads. Before engaging with Trifacta, the upfront data cleaning work for the team at Malwarebytes was substantial, requiring five people concentrating on vast Excel spreadsheets. End-to-end time from lead intake to CRM input could be as much as six months, which prevented the sales team from being able to follow up on leads in a timely manner and thereby impacted close rates. Using Trifacta, the team was able to onboard 40,000 leads in half the time with just one person.

FASTER
ONBOARDING

EFFICIENTLY
SCALE BUSINESS
PROCESSES

"*Clean data enables more targeted marketing, which is driving increased revenue at Malwarebytes.*"

Nordea

REGULATORY COMPLIANCE

As one of Europe's largest banks, Nordea Bank is subject to the highest level of regulatory oversight in the financial services industry. One such regulation, SEPA, impacts credit transfer and direct debit transactions for euro payments. Failure to comply with SEPA can result in huge fines and significant damage to a financial institution's reputation, making the importance of having high-quality data paramount. In this case, the sheer volume of data, as well as the fact that it's supplied by multiple sources from many different countries, creates a staggeringly large and complex data preparation challenge with very little margin for error. With Trifacta, Nordea can now identify the contents of each data source and cleanse data quickly, reducing turnaround time for SEPA compliance from 15 days to one.

HIGH-QUALITY PRODUCTIVITY

REDUCED OPERATIONAL COSTS

*"Clean data allows Nordea to produce higher-quality regulatory reports, which are consequential to maintaining the bank's sound reputation."*

KUECKER
LOGISTICS GROUP

PINPOINTING INEFFICIENCIES

Optimizing warehouse logistics is essential for the efficiency of large companies. Robots, people, and transportation need to be coordinated perfectly over multiple steps; any error introduced into the system can produce compounding, unacceptable delays. Kuecker Logistics Group (KLG) delivers logistics and supply chain management solutions to its client base, using a combination of technological solutions: everything from simply providing equipment to top-to-bottom systems integration and warehouse retrofits. A key part of KLG's value proposition is identifying and solving inefficiencies in material handling processes, using timestamped IoT data from a vast array of sources in and around warehouses. Naturally, none of these sources uses the same formats or logic, which results in massive amounts of data streams that need to be cleaned before they can be properly analyzed. Using Trifacta, KLG obtains and prepares customer data five times faster than before, accelerating analysis and customer value delivery.

ACCELERATING ANALYSIS
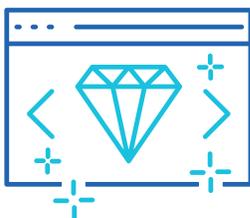
"Clean data makes it possible for KLG to create smarter, more efficient warehouses for some of the largest companies in the world."

MACHINE LEARNING MODEL DEVELOPMENT

Consensus Corporation provides a platform for retailers that simplifies the process of selling bundled products, like a new iPhone with an AT&T data plan or a new TV with a Netflix subscription. In high-end device retail, identifying and preventing fraudulent purchases is a top-line business priority, and requires a robust, up-to-date fraud detection model to handle this risk properly. Consensus Corporation leverages an automated machine learning solution to generate their fraud detection, utilizing an enormous amount of transaction and billing data from retailers and credit card companies in a wide range of formats. Every time the model is updated, a fresh set of data needs to be loaded. Before Trifacta, preparing this data for machine learning routinely took a Consensus Corporation development team six weeks to complete. With Trifacta, the data can be prepared in one week.

FASTER SOFTWARE UPDATES

> *"Clean data is the foundation of Consensus Corporation's robust machine learning models, allowing them to update their fraud detection model much faster and to more quickly alert retailers."*

To see the power of Trifacta for yourself, try Trifacta Wrangler right now.