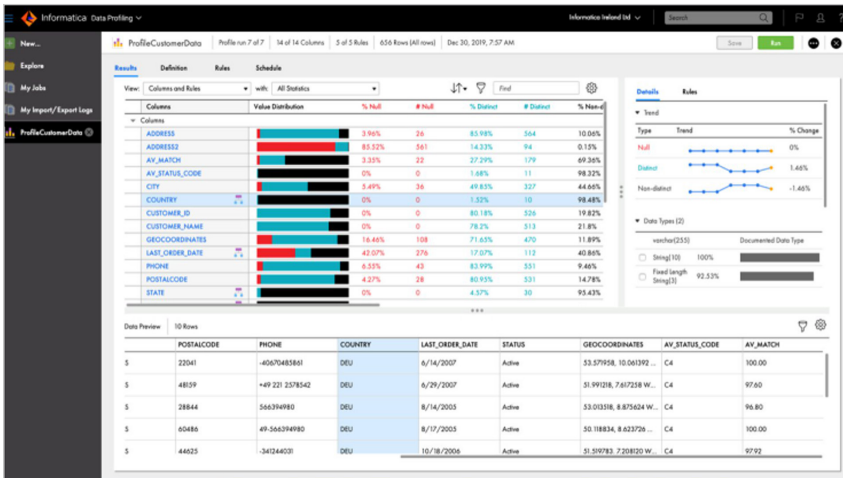




Quick-Start Guide: How to Understand Data Profiling Results and Improve Your Data Quality

Get started improving your data quality with Informatica Cloud Data Profiling on the Intelligent Data Management Cloud.

- See how good your data quality is using the tool
- Identify data anomalies, inaccuracies and duplications
- Create cleansing and standardization rules to continually monitor and improve your data quality



Sample of Data Profiling Results

It's crucial to prevent bad data from entering systems and propagating further, as repairing the damage caused by poor data can be costly and even harm the reputation of the whole organization.



What is Data Quality and Why is It Important?

Data quality refers to your data's overall:

- Reliability and accuracy
- Completeness and consistency
- Relevance

It tells you how much you can trust and rely on your data for business strategies, analysis and other data-driven activities. Organizations need high-quality data to derive meaningful insights and make informed business decisions.

Inaccuracies, inconsistencies and incomplete information make your data less reliable. Poor data quality can result in incorrect business decisions, project delays and an inability to adapt to new market trends.

The Impact of Poor Data Quality

Every year, poor data quality costs organizations an average of \$12.9 million¹. Nearly 60% of organizations don't measure the annual financial cost of poor-quality data.²

It's crucial to prevent bad data from entering systems and propagating further. Repairing the damage caused by poor data can be costly and even harm the reputation of the whole organization.

How to Improve Data Quality

To improve data quality, the best practice is to understand the underlying data-quality issues. Data Profiling is a diagnostic approach to identify anomalies in the underlying data. By ensuring that you run a diagnosis and examine your data you can proactively create a plan to fix many of your data problems before they cost you.

First, data profiling helps cover the basics with your data, verifying that the information in your tables matches the descriptions.

¹ <https://www.gartner.com/smarterwithgartner/how-to-improve-your-data-quality>

² <https://www.gartner.com/smarterwithgartner/how-to-stop-data-quality-undermining-your-business>

Data profiling determines the quality of data and its completeness, conformity and consistency. See the chart below for an example of data profiling results from customer data. Using results like this, your business analysts can create/update data definitions for data engineers to build reference dictionaries and rules using your data pipelines.

Data profiling can reveal if some critical customer contact information is missing. This missing information could directly contribute to a high volume of customer complaints and make good customer service difficult.

Check out this simple three-step approach shown below in Figure 1.

- **Step One** – Run data profiling on your existing datasets (take a random sample in case of a large volume of data set), which will identify all data variations and build dictionaries.
- **Step Two** – Get profiling results verified by business/data analysts to define data standardization and business rules to overcome any issues with data quality.
- **Step Three** – Share the updated standardization and business rules with data engineers to update your data pipelines.

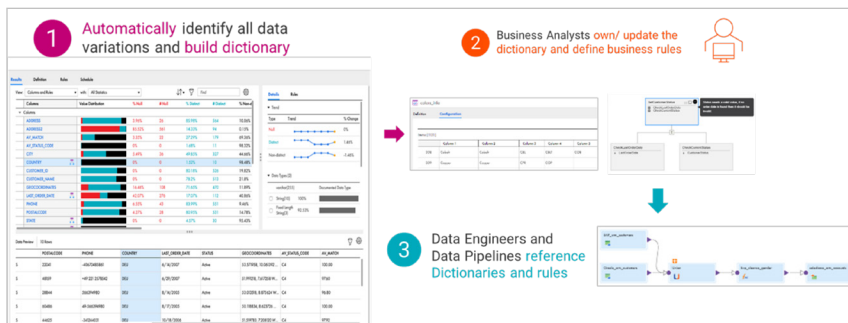


Figure 1

Resources for Further Help

Use these resources for further help with Data Profiling.

- **Interactive Demo:**
 - **Get started with Data Profiling in just a few clicks**
- **Video Demos:**
 - **Creating and running a profile in Informatica Cloud Data Profile**
 - **Creating Rule Occurrences and Viewing Scorecards**
 - **Exploring Profile Results**
- **Informatica Cloud Data Profile Product Documentation**
- Join **Informatica Data Quality Community** to learn and collaborate!
- **Contact Sales** for more information.

Get a faster return on your investment with Informatica Professional Services

Talk to our experts in cloud data management.

Check out our services.

Informatica (NYSE: INFA) brings data to life by empowering businesses to realize the transformative power of their most critical assets. When properly unlocked, data becomes a living and trusted resource that is democratized across your organization, turning chaos into clarity. Through the Informatica Intelligent Data Management Cloud™, companies are breathing life into their data to drive bigger ideas, create improved processes, and reduce costs. Powered by CLAIRE®, our AI engine, it's the only cloud dedicated to managing data of any type, pattern, complexity, or workload across any location — all on a single platform. Informatica. Where data comes to life.

