Forget data cleansing, start data cleaning

Save substantial time before your next asset management project, valuation or EOFY audit with a 5-min daily data fix

SIEMENS

Introduction

We know that best practise infrastructure asset management comes down to having good data. The problem is, with so much data to collect and maintain across different asset types, how do you keep it accurate and up to date without data cleansing taking over everyone's day job?

Year after year, major data cleansing exercises continue to roll around, despite previous efforts to cleanse the data. It's an often-dreaded task required before a major project, valuation or EOFY audit, when the pressure is on to collect, check and update a huge amount of data in a very limited amount of time.

With so many different, geographically dispersed assets to track – from roads to drains, facilities, buildings, parks and much more – maintaining accurate asset registers is one of the most significant challenges local government asset managers face today.

Our advice? Stop relying on data cleansing alone. Instead, start data cleaning very small amounts every day.

Understanding the difference between data cleansing and data cleaning is key to creating good data cleaning habits and identifying the scope of your cleaning activities.

It's also important to have the right mindset for managing the ongoing quality and accuracy of your data. Only by taking a long-term approach, with a focus on continuous improvement, can you ensure that your data remains current, consistent and connected to your needs.

Adopting a daily ritual of data cleaning is guaranteed to save you time and money and ensure you are better positioned to meet your compliance obligations. What's more, it allows you to deliver better services and infrastructure in the most cost-effective way – enhancing the credibility and trust of Council in the eyes of your community.

Once you've identified your biggest data accuracy issues and set your team up, just five minutes a day is enough to keep your data in check.

We've helped many local governments with their data quality issues by setting them on the path of completing a 5-minute daily data clean and seeing them reap the benefits.

We encourage you to give it a go.

Contents

01 Why data cleaning is important

While most local governments maintain basic information about their assets, many lack up-to-date data about their assets' condition and maintenance.

Local governments need to keep their data clean for several reasons:

To set council rates

A 2016 Queensland Audit Office report found that due to inaccurate and incomplete asset condition data, councils were undertaking asset renewals in a reactive manner, which in turn impacted their understanding of the level at which to set council rates.

To maintain assets

Recent audits of Queensland and NSW councils found millions of dollars' worth of land and infrastructure assets had not been recorded, and therefore the councils did not realise what they had. Assets that are not recorded in asset registers are not subject to regular maintenance and are therefore left to deteriorate, resulting in greater costs in the long term.

For EOFY reporting

Navigating the regulatory reporting requirements and EOFY accounting can be a stressful time for local governments. Not only do you have to navigate all the regulatory reporting requirements, you also need to be confident your data can stand up to scrutiny by audit teams.

To support budget decisions

Bad data can result in the wrong investments and lead to unnecessary costs in other ways too. For example, if it is believed that a road is in better condition than it actually is, it will be overvalued, and budget problems are likely to quickly snowball.

To access funding

Not only is accurate information about the condition of assets essential for decision-making and long-term strategic and financial planning, it is also needed to comply with state and federal disaster response programs. If the information is vague or incomplete, the council risks being able to access funding.

Before a new project

Minimising the unexpected is essential when it comes to project timelines and best practise decision making. Knowing exactly what assets are currently in place (and in the way) allows better estimates and timelines when we need to replace it, remove it or improve it.

02 The six most common data integrity issues for Australian local governments

When we looked at the sorts of data problems our customers run into each year, **we identified six common issues:**

1
Incomplete data

Usually, incomplete asset data builds up over time without anyone noticing. By the time you do notice, it's been too long, you can't remember who to ask about it, and all the immediate knowledge about those assets has gone stale.

At this point, you've often got a big project on the go, the need for that data has become urgent, and with little time to find the missing data, the task becomes overwhelming.

2
Mismatched data

Mismatched data often goes hand in hand with incomplete data. For example, in your asset register, you might find a gravel road record with asphalt listed as the material type. Or a light pole listing a gear type that it would never support. Mismatched data causes mismatched actions. Is the light pole correct or is the gear type?

As you don't know which is right, you can't know what inspection or maintenance is needed. The asset could also get caught up in the wrong kind of reports. Ultimately, mismatched data ends up potentially skewing results with knock-on budgeting effects.

3
Out-of-date data

In this instance, the data is complete, and you know it was accurate once upon a time, but now, you no longer have confidence in it. This is often the result of data collection exercises for a one-off project, after which the data has been left to sit and go stale – meaning you need to repeat the whole data collection process again.

4
Overwhelming issues

This category is for all the minor backlogging problems no one has had time to deal with. While none of them are show-stopping issues, the challenge is there are just so many of them.

This is the number one problem our customers ask us to help them with. With so many small little things to deal with, they can't be easily categorised and prioritisation becomes just too hard. You end up with a massive backlog that you can't keep track of, which builds to become a huge block and a show-stopper in itself.

5

Reoccurring issues

You might have thought you'd fixed a particular issue, but a few small details sometimes fall through the cracks, and now, the problem has reoccurred. These scenarios usually start small and sneak in over time.

Suddenly, you realise you've got a major reoccurrence of the same issue without anyone noticing, and it's about to stop your current project from moving forward. The problem needs to be fixed the second, maybe the third time, amounting to a huge waste of budget and resources.

6

Data cleansing and data collection

Many people think data cleansing and data collection are the solution to all these problems. The fact is, they can be an issue of their own.

We know that data collection is expensive and takes a long time. Meanwhile, data cleansing requires several steps, people, and planning, and can become a multi-pronged project in its own right. Before you know it, your initial project has morphed into three or four projects stacked up – delaying the start of planned activities and taking time and effort you can't afford. All the while, your day-to-day work is neglected and continues to pile up.

What are the advantages of having clean data?

- Improved decision making
- Save time and money
- More cost-effective
- Smarter long-term investments
- Better services and infrastructure
- Hassle-free EOFY reporting
- Easier to meet compliance requirements
- Greater team efficiency and productivity
- Accurate council rate setting backed by evidence
- Increase ability to access additional funding with data insights
- Less asset downtime due to better maintenance decisions
- Enhances council credibility and trust by community

03 Data cleansing vs. data cleaning: What's the difference?

If you lack confidence in your data, **what can you do about it?**

First, you need to take a longer-term view of your data. To do that, you need to stop just intensive data cleansing for major projects and instead start cleaning very small amounts of data on a regular basis.

While data cleansing is usually a short-term, intensive process to support a specific project, with data cleaning, you take a whole asset register view over the long term – meaning you're keeping your data clean, all the time. **When your next project, valuation or EOFY rolls around, it takes far less time and effort to clean up any issues because you've been quietly and consistently chipping away at it.**

04 Five steps to set up your data cleaning

So, how do you start the process of regular data cleaning?

Having helped a number of customers with their data cleaning, we've distilled it down to five easy steps (and a secret sixth step) to help you create a less stressful, simpler and more proactive approach to ensuring the data quality of your asset register.

After this, all it takes is five minutes daily to keep your data in check.

1. Get everyone on board

Get a team together. No one can carry the entire load alone, so we'll say that one again: Get a team together, then identify the issues you need to solve and agree on the solutions. If you don't have everyone involved, you're not going to have a good outcome. It's the simple change management principle of Buy In. Sell the idea of doing it every day and get the teams input so they feel ownership of the task, and they're much more likely to make data cleaning part of their daily routine.

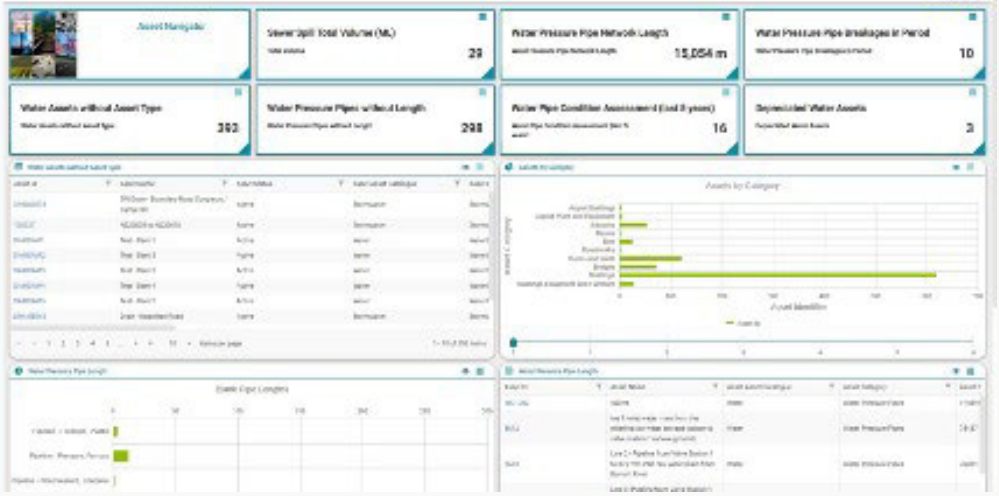
2. Stop trying to make it perfect

Perfect data is the enemy of good data. We have to stop aiming for perfection with our data because it gets in the way of getting things done. Good data delivers good results, and the only way we will get those results is by letting go of the idea of perfect.

What does good data look like? What data you collect depends on what's relevant to your decision-making. That's why you need to assemble your A-Team first.

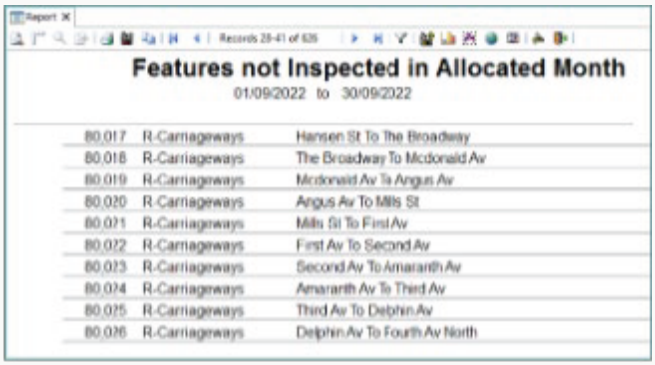
3. Make an issue list and give each issue a priority

Go back to your last EOFY, valuation or data cleansing project and make a list of all your issues with the data. Once you have a list, try to prioritise each issue. What's a major data issue to you may differ from your colleagues, so be sure to include stakeholders both above and below you. Given priorities evolve, these shouldn't be set in stone – try to focus on solving the top five big issues now and adapt this over time.



Assetic dashboard

Confirm dashboard



Confirm scheduled report

4. Decide the methods of delivery

Decide how to report on each issue for action. For example, it could be via a dashboard, an automated email or a data extract to a network folder. Different data cleaning problems might need to go to different people in your organisation, so you could end up with several different reporting methods. For non-system users, the report might be a PDF in their inbox each morning listing the three assets they need to focus on. Make your data cleaning versatile enough so everyone can receive reports and work in the way they're most comfortable.

5. Automate it

Automate as much of the data cleaning process as possible, such as sending automated emails, reports or data extracts about missing data so your team can concentrate on finding the answers as efficiently as possible. Another good idea is to put in your calendar an untouchable five-minute appointment for data cleaning so that time is automatically set aside.

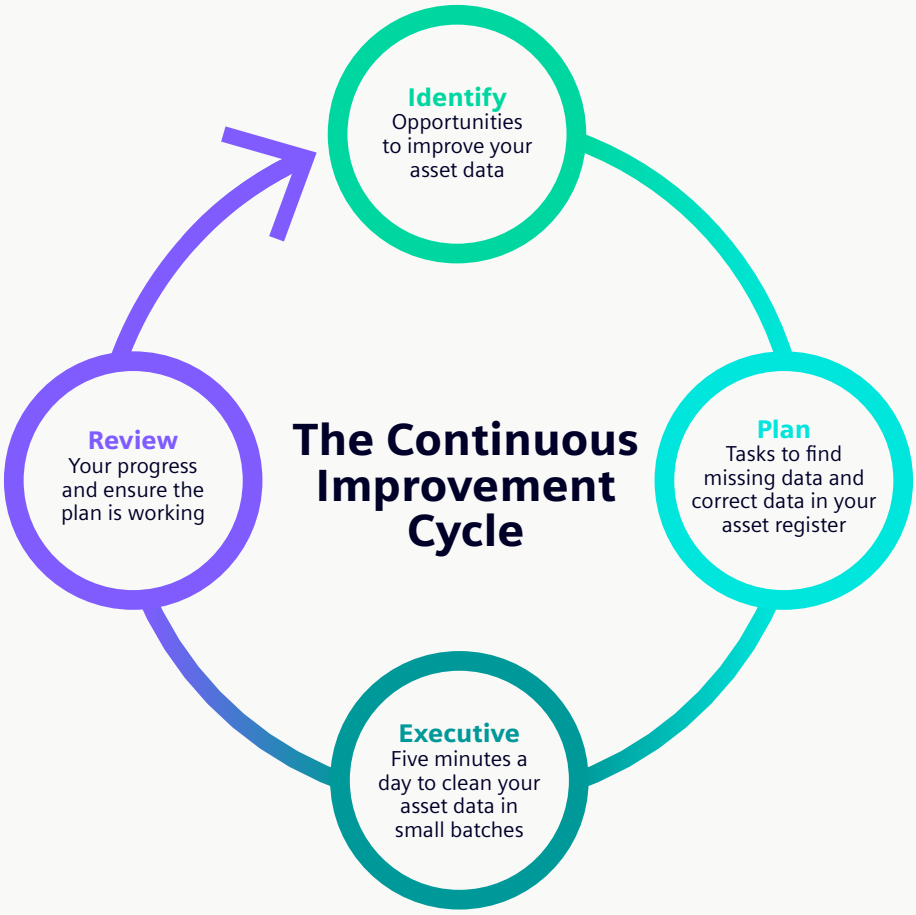
Our secret step six:
Shine it, design it, refine it!

With any process, it's important to take a continuous improvement mindset, which is why the real secret to your ongoing data cleaning success is to shine, design and refine it.

This allows you to identify extra aspects of investigation as you're cleaning and add them to your five minutes a day.

If you're not sure where to start, try asking yourself the following questions, and **don't forget to ask your team:**

- Does our asset data support our critical processes?
- Does everyone who needs access to this data have access to it?
- Have we identified any repeat problems and therefore can we spot any root causes for the data issues we've spotted?
- Does our asset data support the statistical analysis that we want to apply to it?
- Does our asset data support current KPI's?
- Does our asset data support us in our quest to meet industry standards and benchmarks?
- Does our data support our mission for best practise infrastructure asset management?



05 How to create a daily data cleaning ‘To Do’ list

We recommend using a dashboard to oversee your data-cleaning. It lets you see at-a-glance what needs fixing and if issues are building up.

A dashboard allows you to view all the missing or incorrect data that might cause problems for future decision-making and reporting. For example, the dashboard can show you:

- Assets that have never been valued
- Assets that haven’t been valued this year
- Assets with no active ‘From’ or ‘To’ dates
- Assets that have exceeded their useful life
- Assets with no category
- Assets with no replacement cost or valuation quantity
- Assets that are missing a value for a specific attribute in a particular asset class
- Assets with mismatched data.

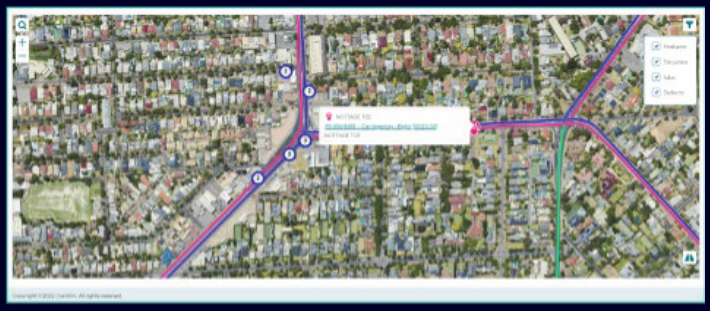
The dashboard acts as a digital to-do list that updates in real-time – fill in the missing data, and the asset will disappear from that data extract. You can also make it available to any of your colleagues who may be asset owners so you can work through the data together.

Confirm widget selection

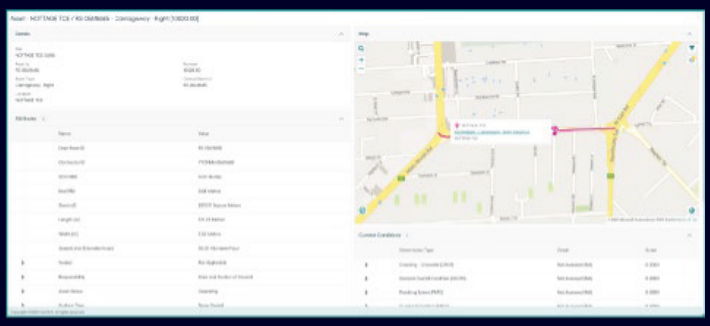
Displaying your data

With Siemens, there are plenty of ways to display the data to your liking, from pie and bar charts to column graphs, data extracts – and even maps.

Our dashboards are browser based, meaning you can share them with all of your necessary stakeholders just by sending them a link.



Confirm dashboard map



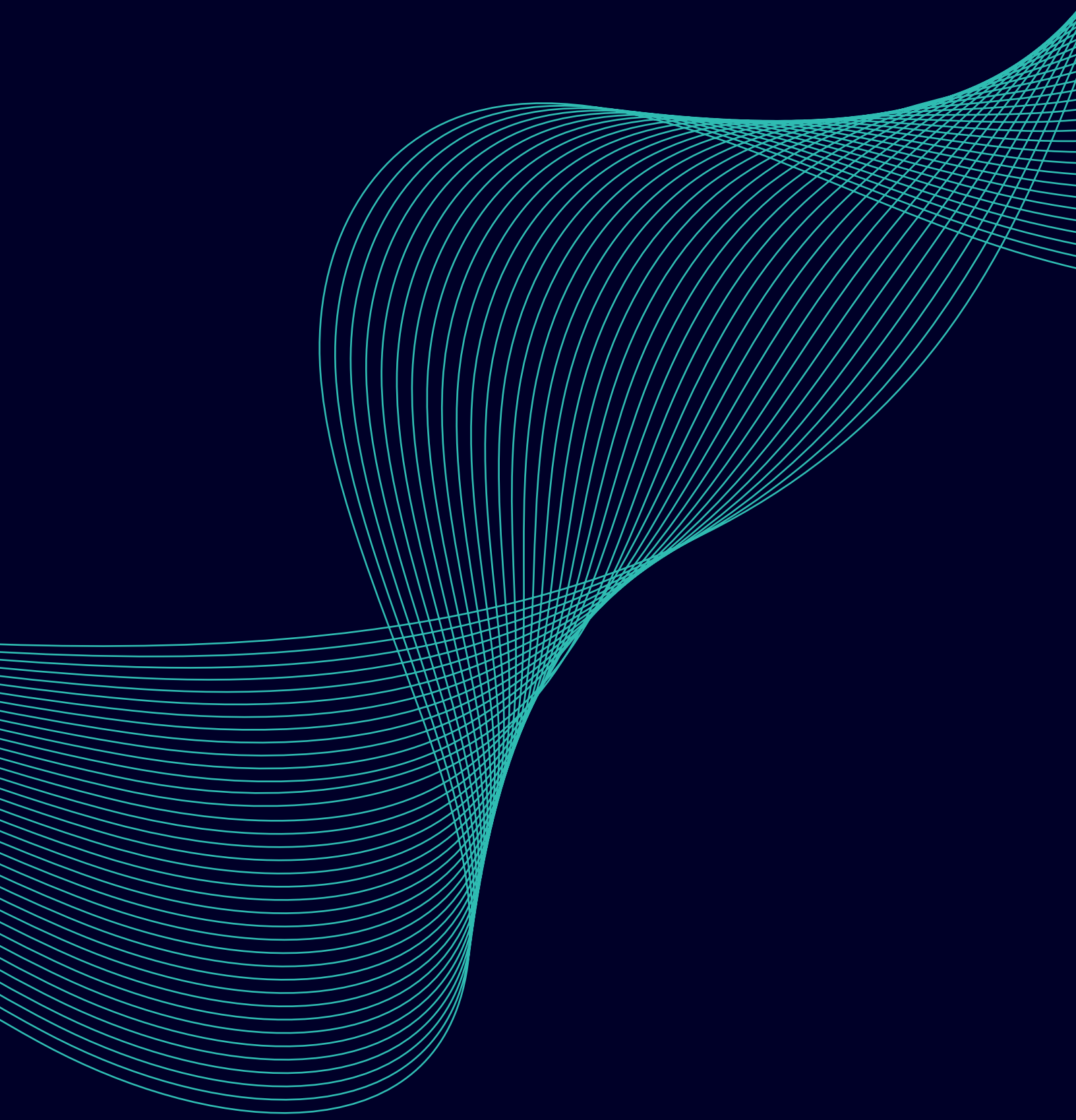
Confirm dashboard drill down to asset register

Conclusion

Save significant valuable time, money and resources by simply getting into the habit of data cleaning for just 5 minutes a day. Identify the biggest problems you face with your asset data, whether it's incomplete data, mismatched information or it's out of date. Perhaps you're overwhelmed with a backlog of small issues, or that the same problems keep coming up, time and time again. Prioritise with your team the top few issues to focus on fixing and for which asset classes, decide how you'll action them, and automate as much of the process as you can. Then it's just a matter of taking five minutes a day (such as first thing in the morning when you log on) to do a quick cleanse.

Fixing your data can get weirdly addictive, so try not to get too stuck into it and lose track of time! Five minutes is ideal, so set a timer and get into the habit of completing small batches daily. Over time, your asset data will gradually improve, and you'll be much more confident in your data when you need it to support decision-making, planning and long-term investments.

Learn more: **assetmanagement.siemens.com**



SIEMENS