

# CATALYST

## Dealing with data quality issues



**Alberto Guerrero**  
Project Engineer  
EnergyCAP



**Lorenzo Casna**  
Customer Success Engineer  
EnergyCAP



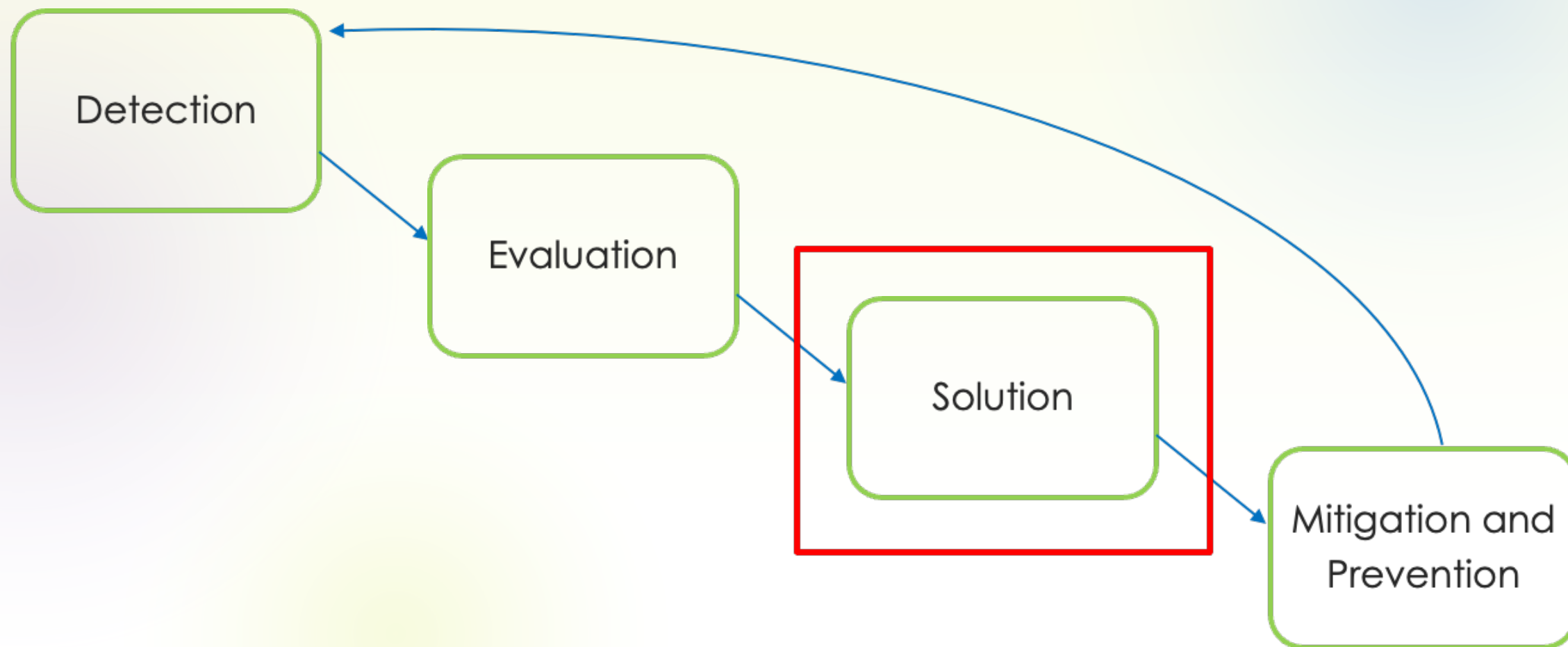


# Agenda

Data management:

- What is a data quality issue?
- Detection and diagnosis
- New Solutions!
- Mitigation and prevention

## Our focus for today



# What is a data quality issue?

Data  
processing

Data is good but looks bad

Incorrect data

Generates anomalies like  
spikes, zeros and more

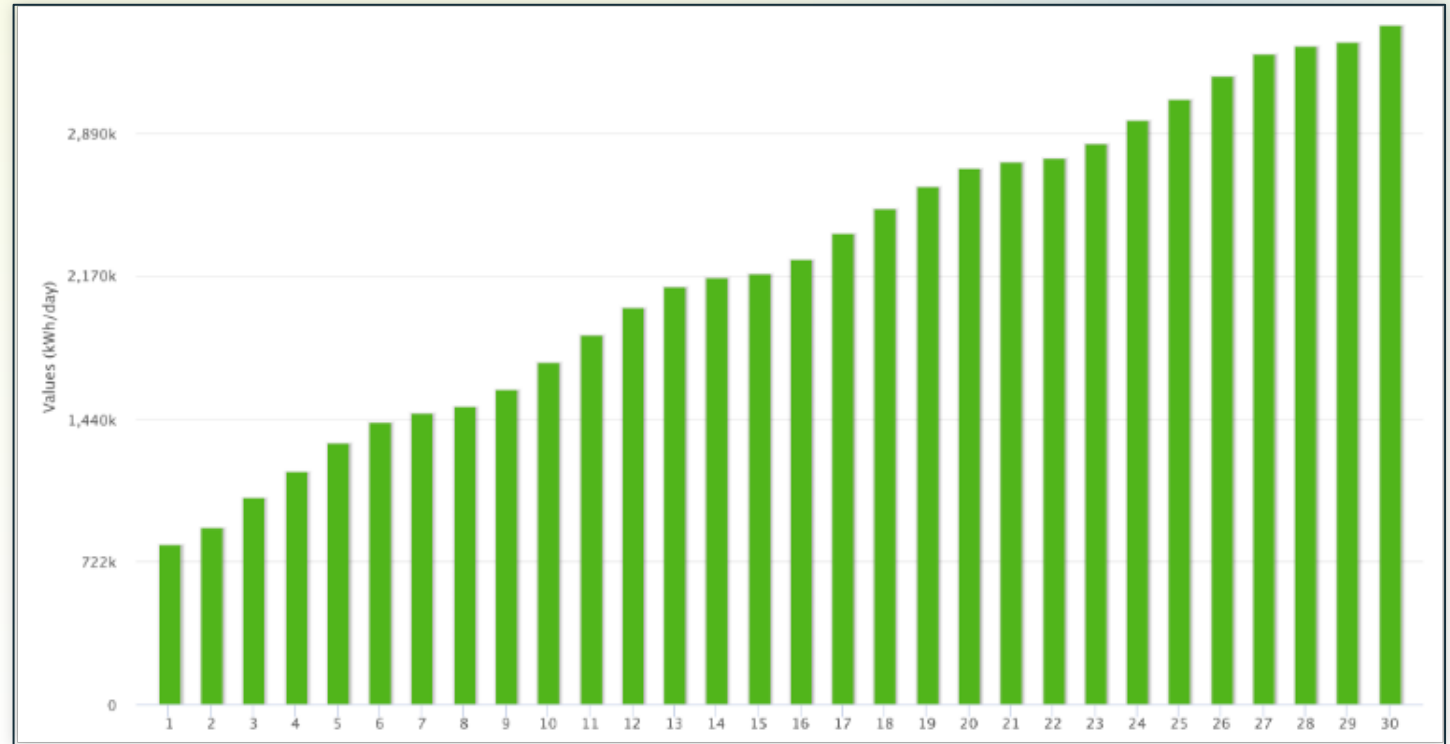
Missing data

Data gaps or errors in data  
collection/allocation

# Detection and Evaluation

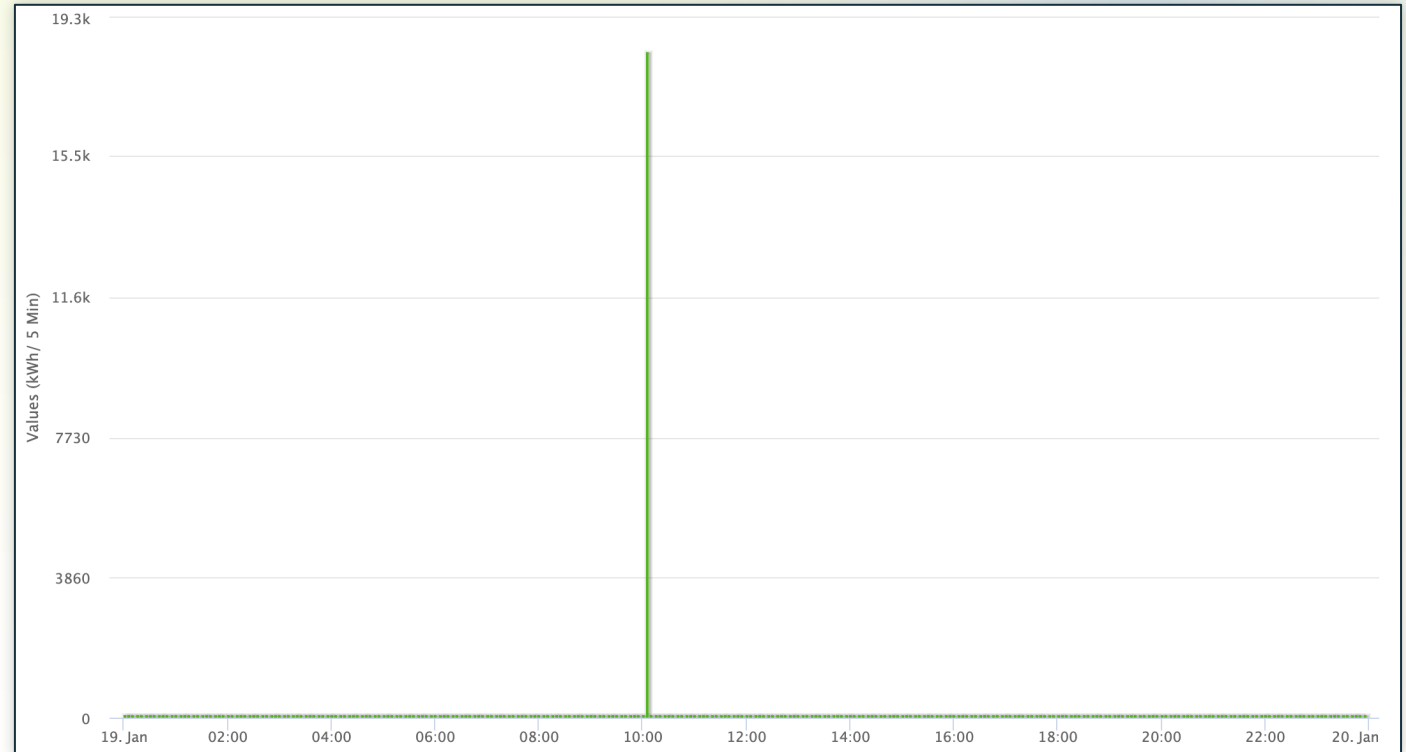
# Data Processing

- Data issue spotted in breakdown or through alerts
- Checking on the data monitors, readings are in line with the expectations.
- Check that the data point configuration is correct.



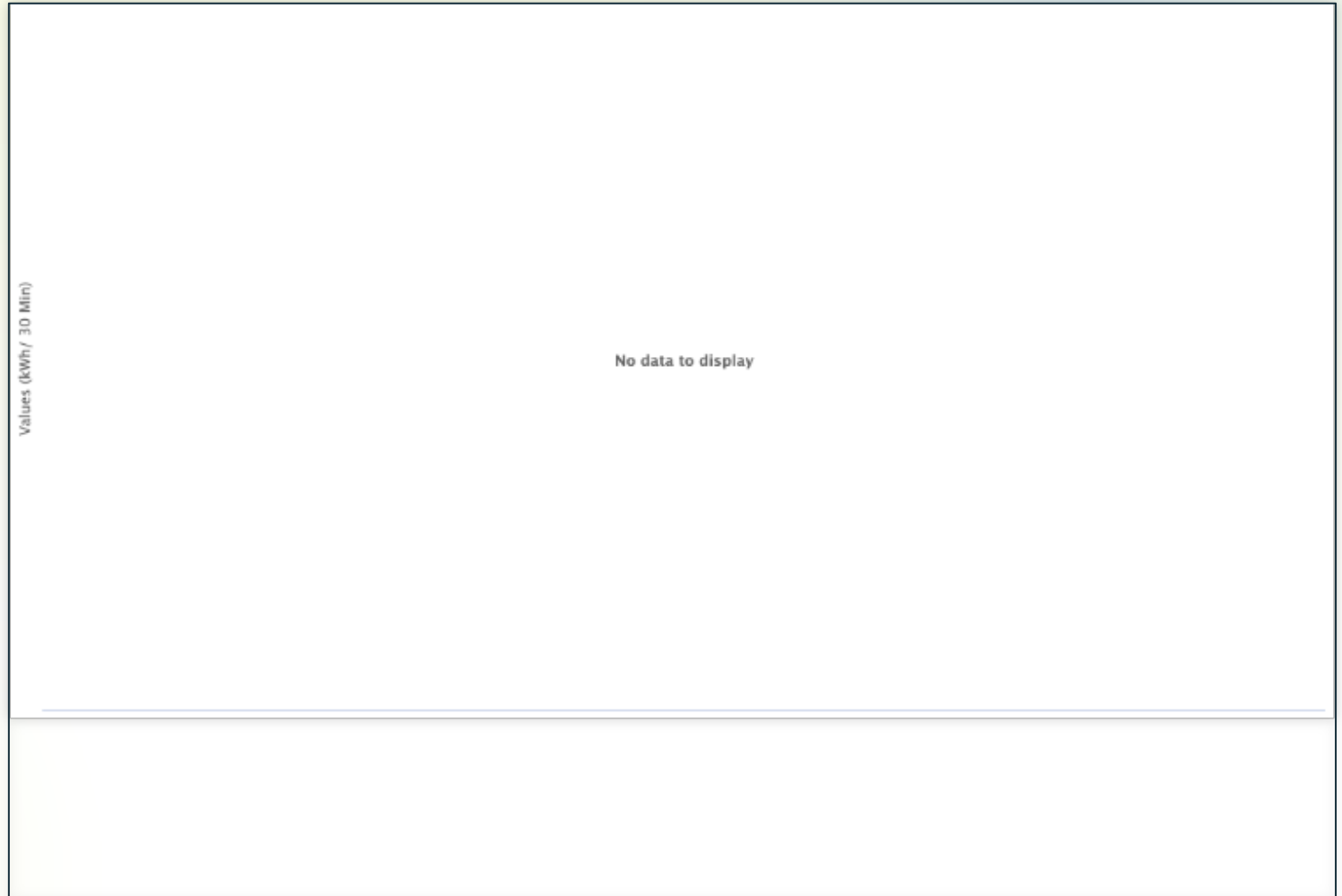
# Incorrect Data

- Data issue spotted in breakdown or through alerts
- Checking on the data monitors, readings are NOT in line with the expectations, showing:
  - Random, null or zero readings
  - Inconsistent datasets
  - Negative powers
  - ...



# Missing Data

- Data issues can be spotted with any tools, alerts/webhooks in particular.
- Checking on the data monitors, readings are not there or are severely corrupted.
- Erroneous uploads or wrongful installations also fall in this category





# Solutions

# Good News!

## Introducing Data Actions

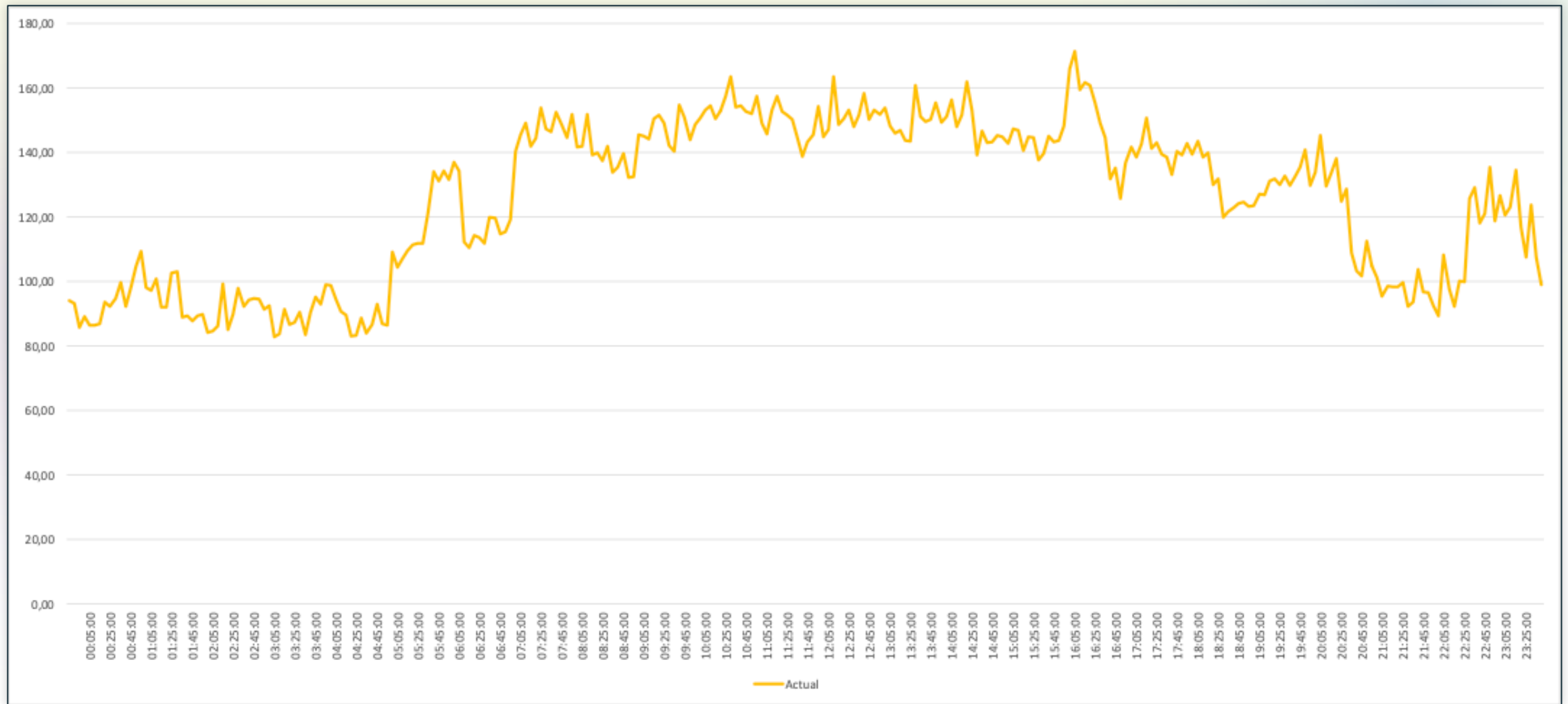
- Reactive solution
- Delete, update, reprocess
- Works in bulk!
- Your go-to-tool for a clean database

Let's see it on the platform!

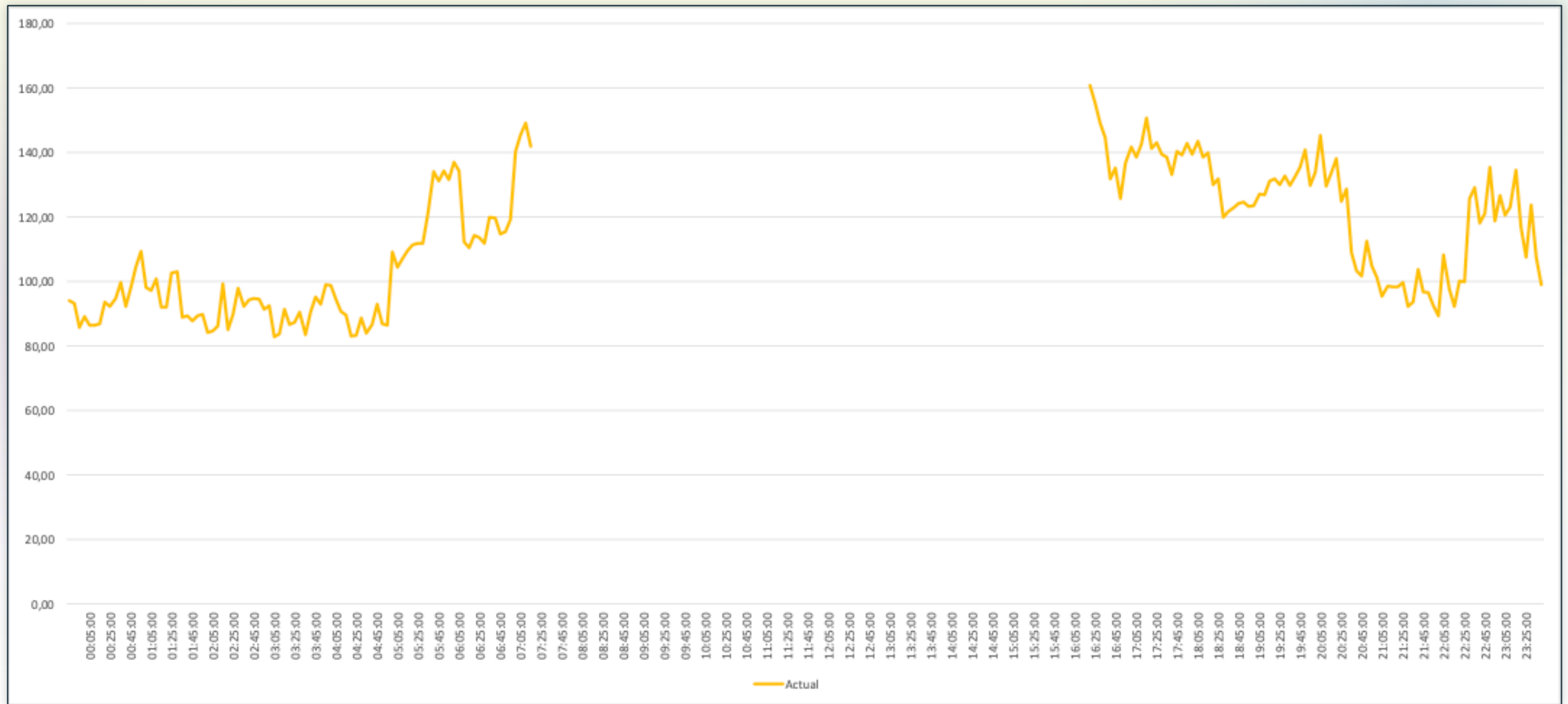
## Introducing Data Doctor

- Proactive solution
- Rebuilds the missing data
- Maintains a clean data visualisation

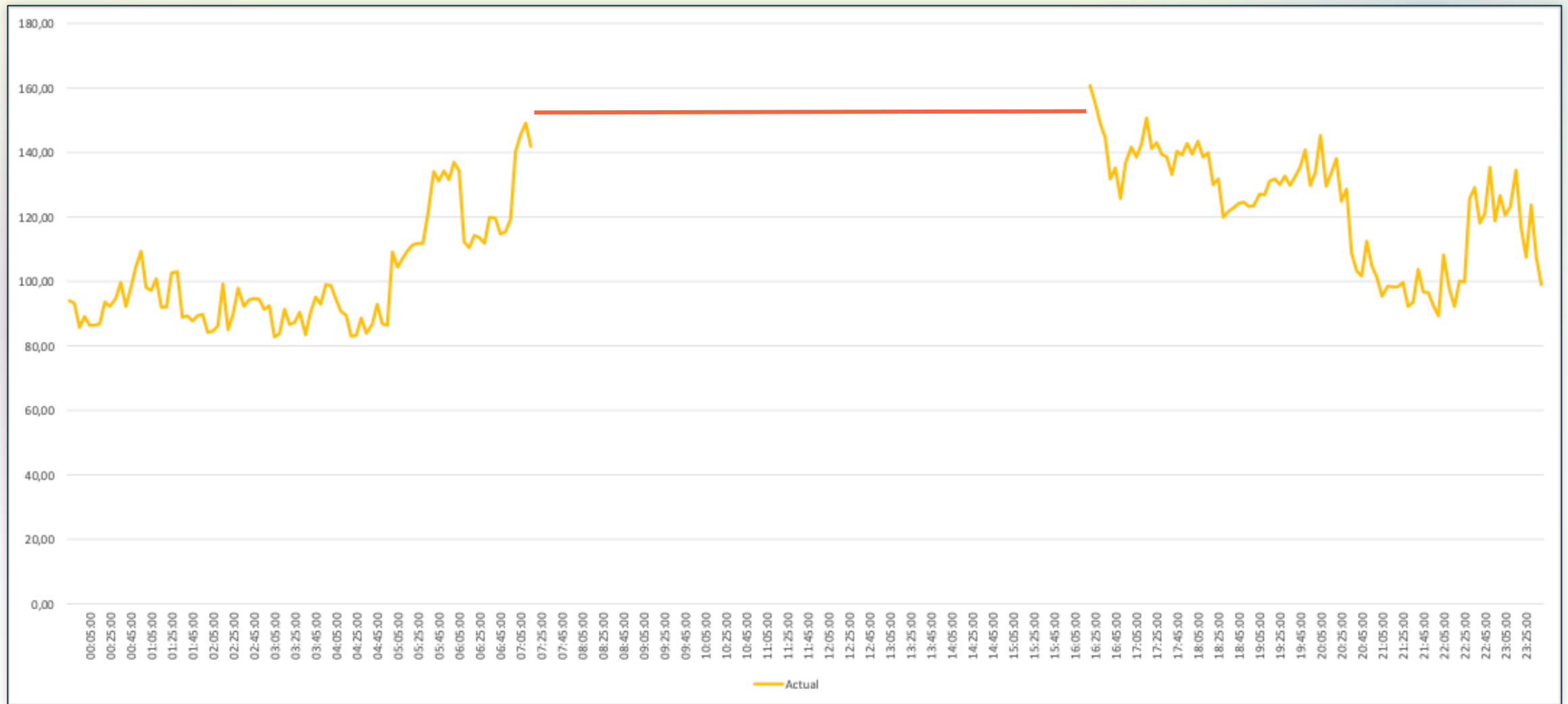
# Data Doctor



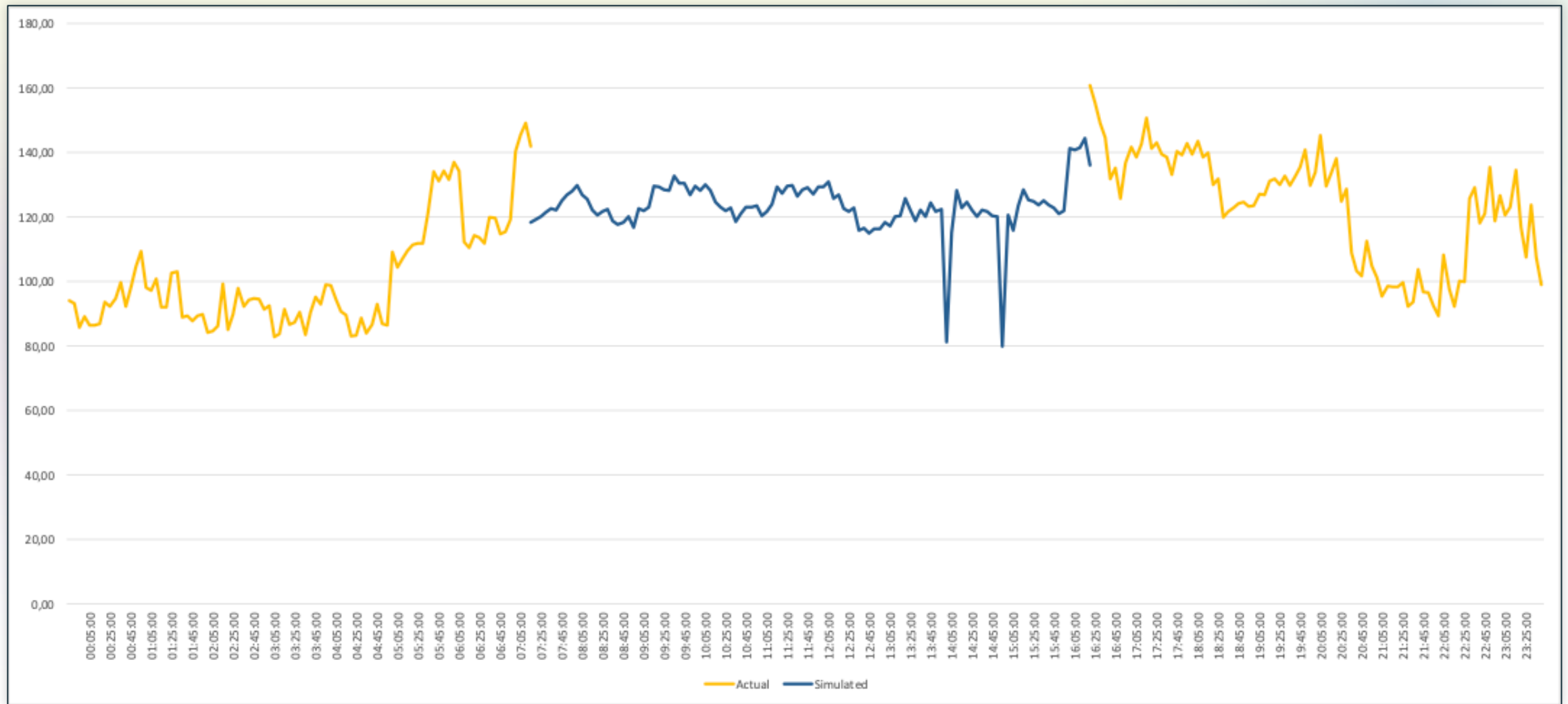
# Data Doctor



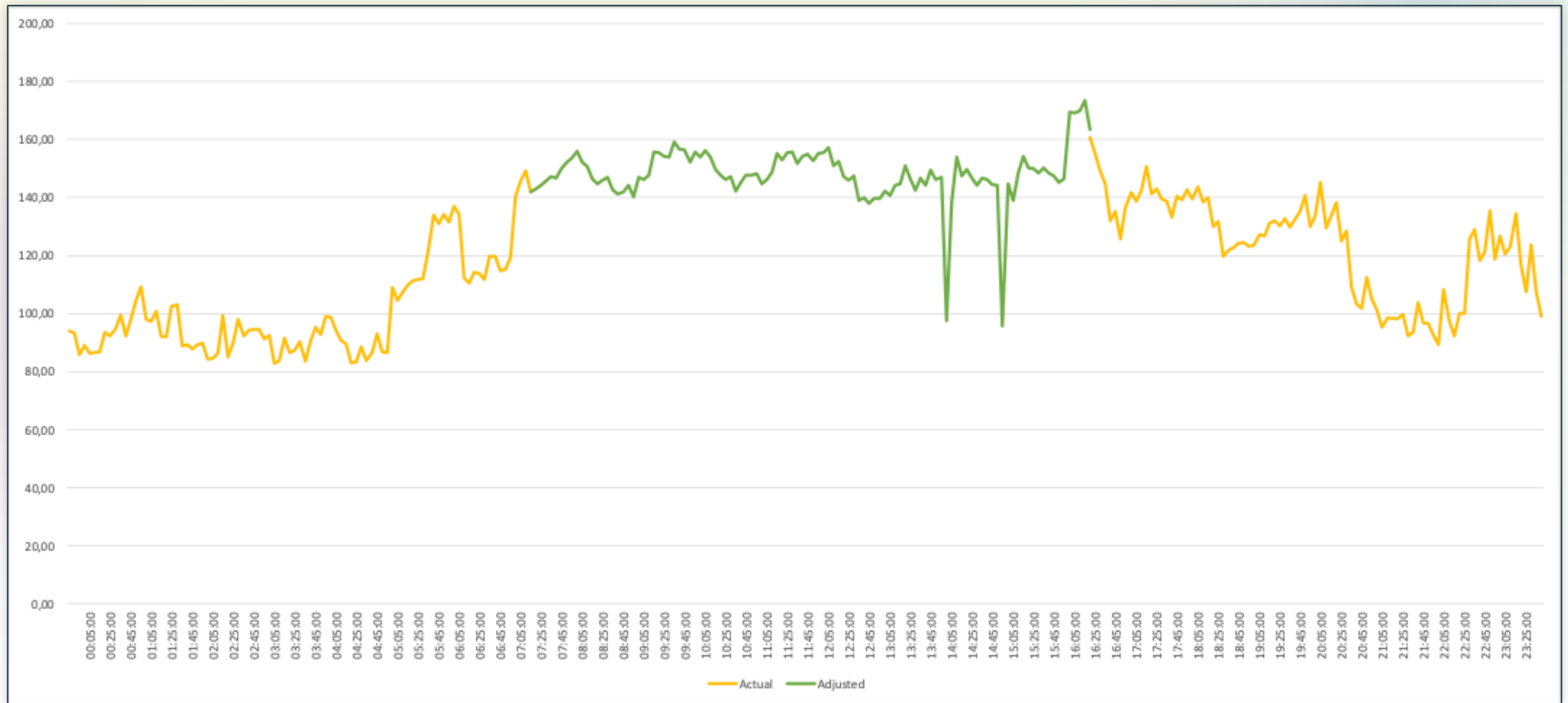
# Data Doctor



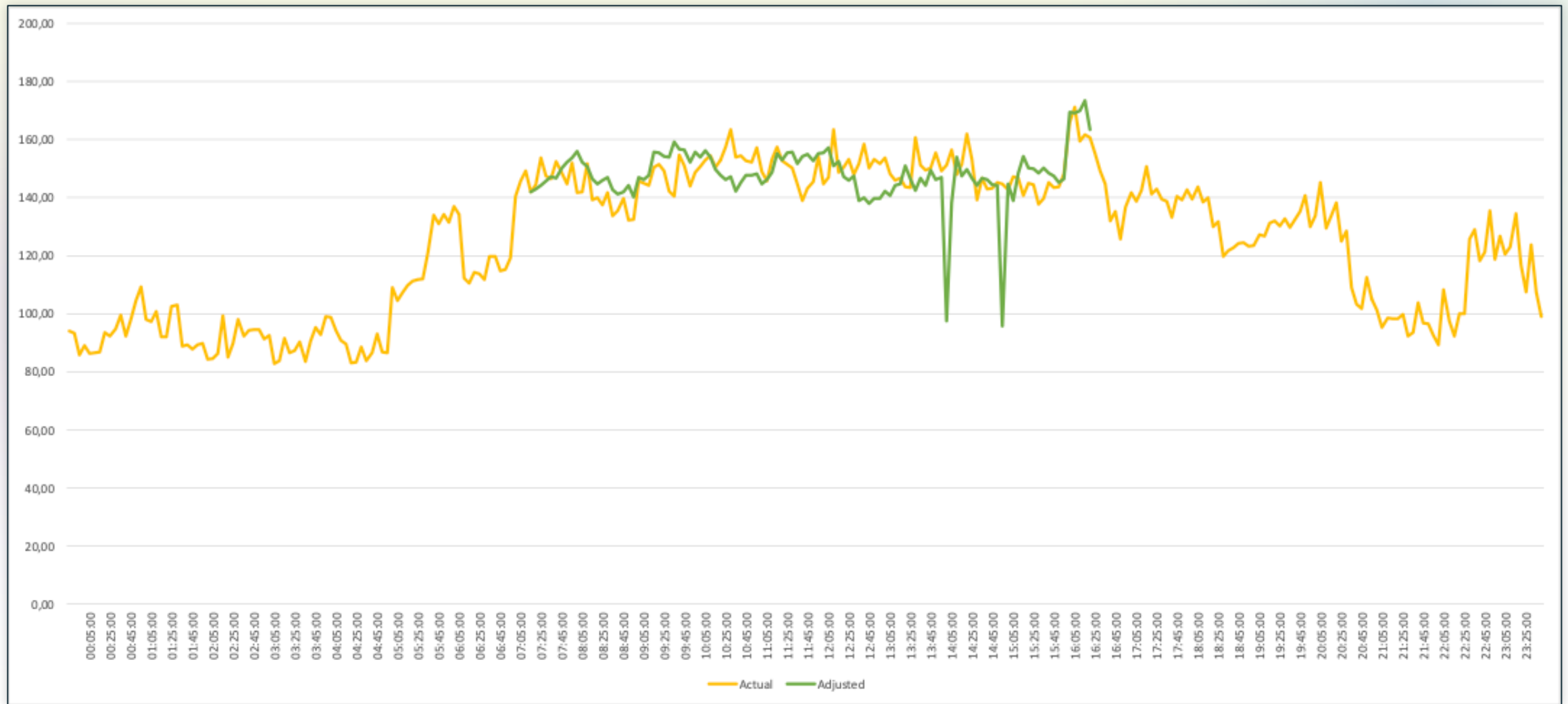
# Data Doctor



# Data Doctor



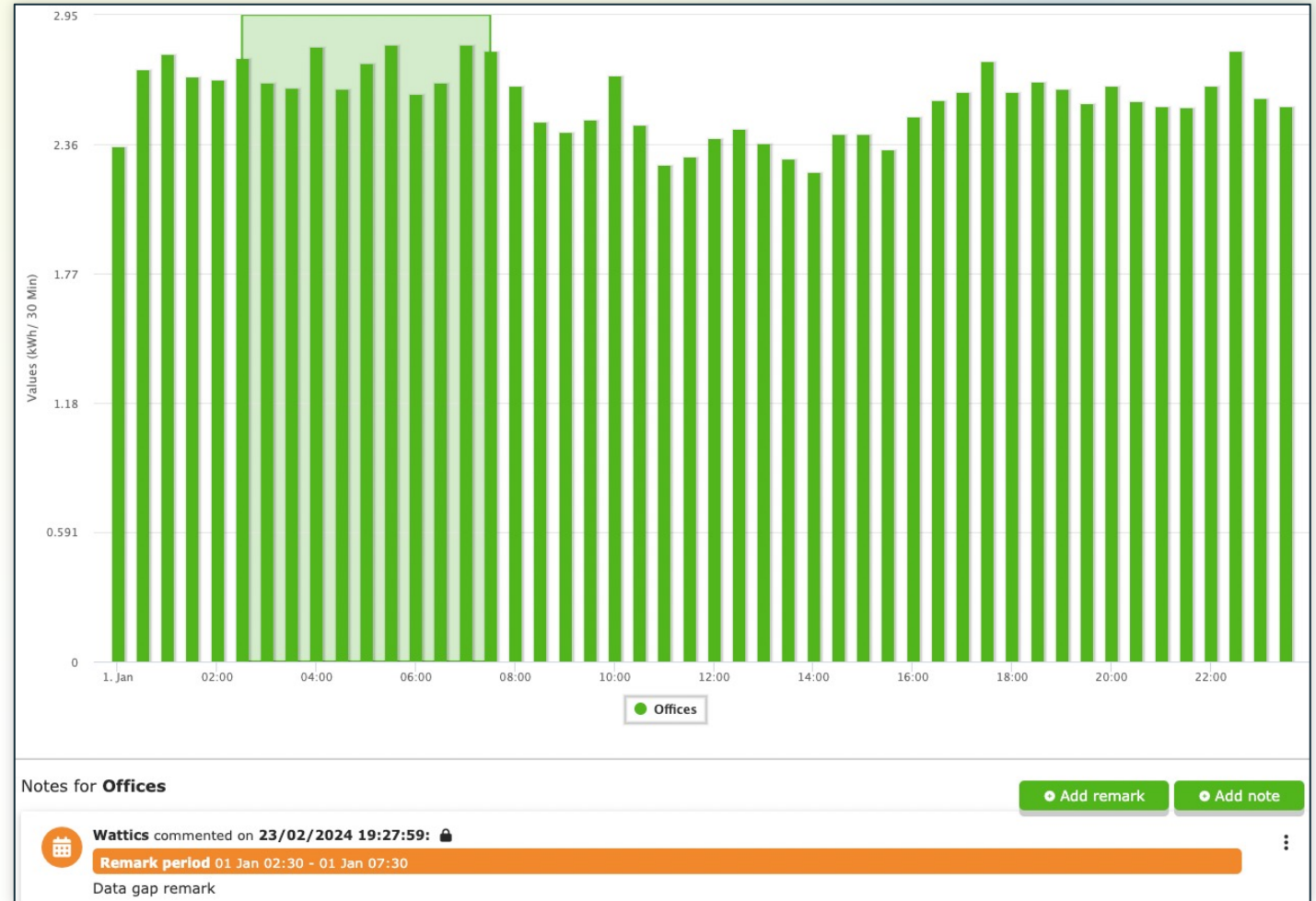
# Data Doctor





# Data Doctor FAQs

- How to enable it?
- When/how will it jump into action?
- How to keep visibility on simulated data?
- ...?



**Don't forget the API!**

# Re-uploading data

When nothing else works...

- Embedded Data Uploader
- Manual (s)FTP pushes
- HTTPs pushes via API
- MQTT

## Data Uploader

Choose the organization to associate the data

New Organization ☒ Yes ☐ No

Name \*

Next

### Guidelines

- Readings must be in interval format.
- For electricity data point readings must be in kWh.
- Only single-value readings can be uploaded (i.e. kWh for electricity data points).
- Readings must be snapped to the clock.
- It is not possible to upload data to appliances (level 4 entities in tree menu).
- Maximum number of rows per file is 20,000.
- Filesize must be less than 10MB.

*\*Uploaded readings are not used for Sentinel calculations and won't trigger any notifications.*

# Mitigation and Prevention

# Mitigation and prevention tips

## Mitigation

- Enable data doctor!
- Whenever possible, choose cumulative data
- CT/PT factors

## Prevention

- Set up Alerts
- Investigate the option to enable webhooks
- Meter Rollovers

User Terms and Conditions >

**Data Gap Webhook** >

Sentinel Trends Webhook >

Custom Message >

### Data Gap Webhook

Here you can define custom a webhook-url in order to receive POST requests when a data gap occurs.

It will send a body like the one below:

```
{  "event_type": "data_gap",  "data": {    "id": "channel-reference",    "start_date": "2019-01-01 00:00:00",    "end_date": "2019-12-31 23:59:59"  }}
```

URL \*

**Save Webhook**

hook.io zapier IFTTT

# CATALYST

**Thank you!**