

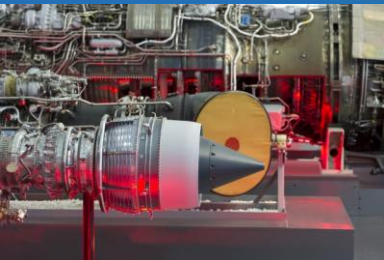
Data Quality & Data Shaping: Keys to Enabling Advanced Analytics & Data Science

Camille Metzinger



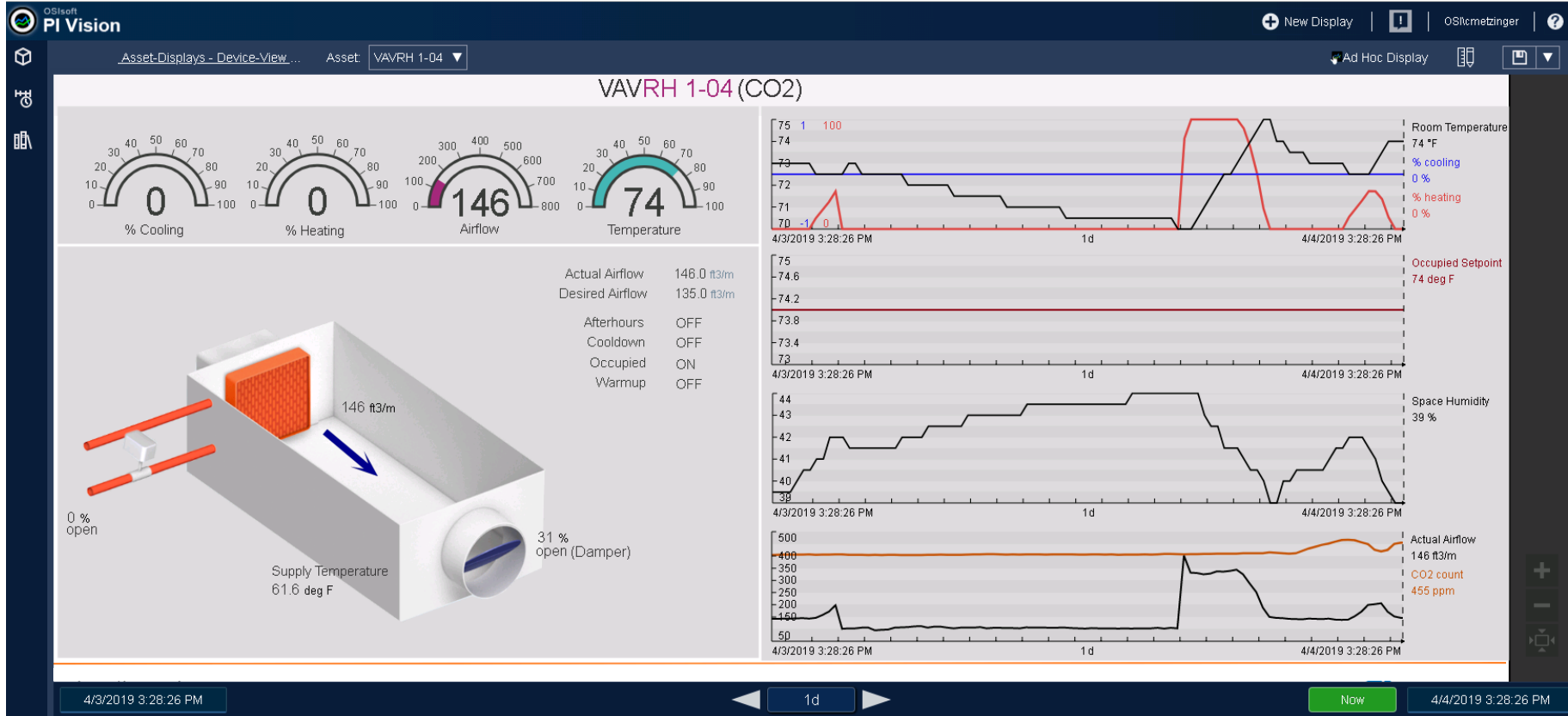
Agenda

- Motivation
- Define Data Quality & Common Issues
- Define Data Shaping & Common Issues
- How to Setup Checks & Structure in the PI System
- Demo



**We all have
important assets**





**We want to empower our
businesses with advanced
analytics & data science**







Data Quality Requirements

We need to have **confidence** in the data used in our analysis so we can draw the **correct conclusions**.

-Accurate

-Reliable

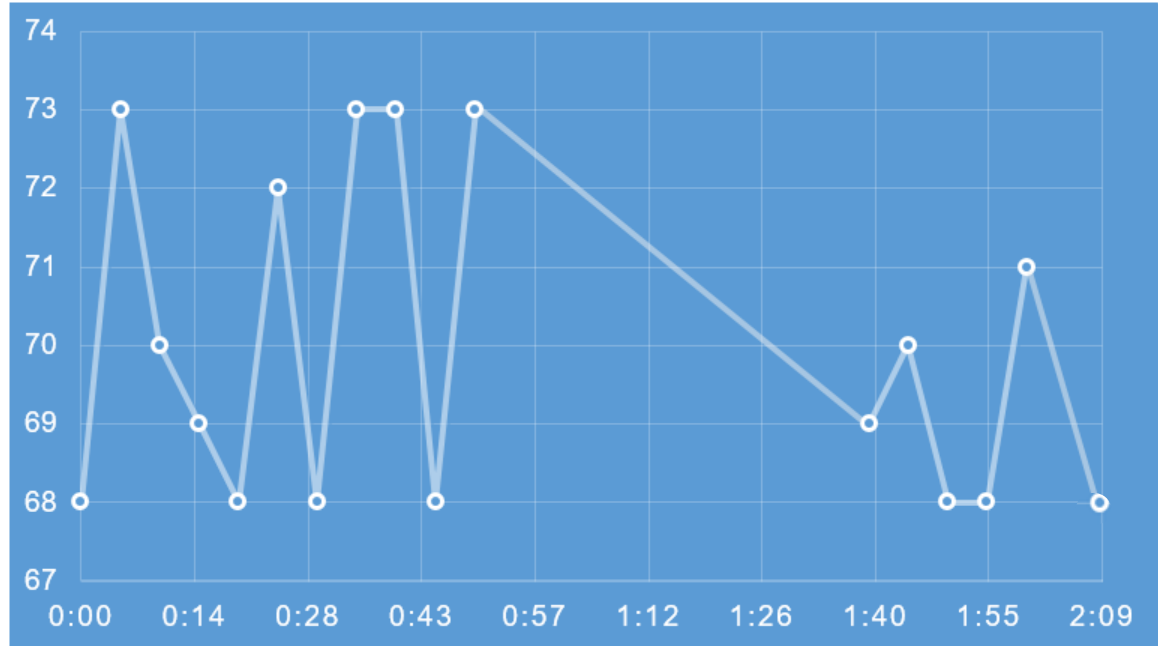
-Complete

Data Quality Issues

- Not setup
- Stale data
- Flatline data
- And more

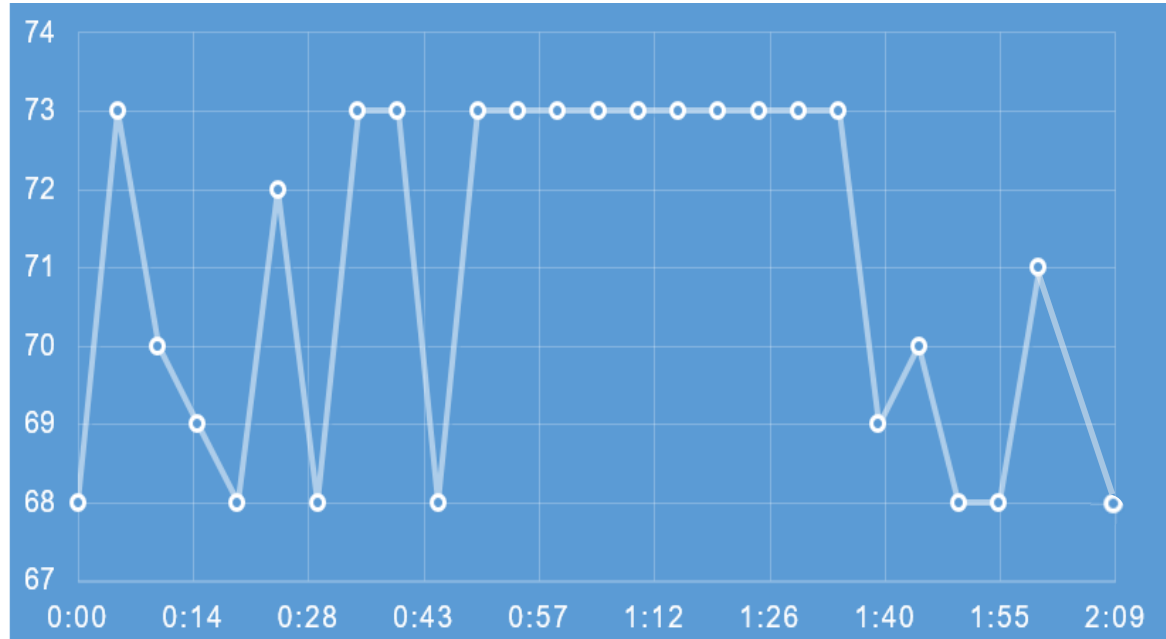
Data Quality Issues

- Not setup
- Stale data
- Flatline data
- And more



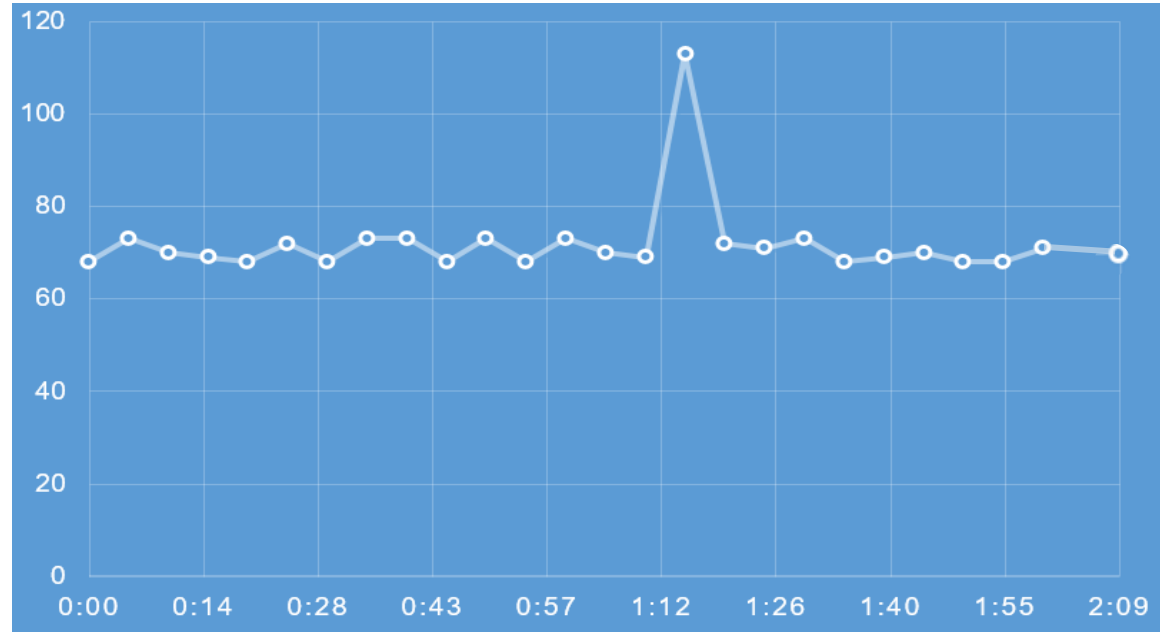
Data Quality Issues

- Not setup
- Stale data
- Flatline data
- And more



Data Quality Issues

- Not setup
- Stale data
- Flatline data
- And More



Data Shaping Requirements

	B	C	D	I	J	K	M	N	O	P	Q	R	S	T
1	Building Shifts	Event Frame Start Time	Event Frame End Time	AC Unit	Absolute Difference	Absolute Offset f	Difference	Floor	Highest Zone Tempe	Interior/E	Lowest Zo	Offset fro	Offset fro	Open/Clo
2	Occupied Schedule - VAVCO 5-03 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 2	16.14084414	0.5	-16.1408	5	75.5	Interior	72	0.5	-2	Open
3	Occupied Schedule - VAVRH 5-10 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 3	15.14118779	1	-15.1412	5	73	Exterior	69	-1	-4	Open
4	Occupied Schedule - VAVCO 5-10 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 2	16.14127371	1	-16.1413	5	74	Interior	72.5	-1	-1.5	Closed
5	Occupied Schedule - VAVCO 5-11 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 2	16.14135962	2	-16.1414	5	74	Interior	71.5	-2	-2	Closed
6	Occupied Schedule - VAVCO 5-01 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 3	15.14513985	0	-15.1451	5	75	Interior	72	0	-0.5	Open
7	Occupied Schedule - VAVCO 5-16 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 3	16.14522577	3	-16.1452	5	72.5	Interior	69.5	-3	-3.5	Closed
9	Occupied Schedule - VAVRH 6-11 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 3	15.14531168	1.5	-15.1453	6	72.5	Exterior	69.5	-1.5	-3	Open
10	Occupied Schedule - VAVCO 6-01 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 3	15.14531168	0	-15.1453	6	75	Interior	72	0	0	Open
11	Occupied Schedule - VAVCO 5-18 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 3	15.14531168	1	-15.1453	5	74.5	Interior	72	-1	-1	Closed
12	Occupied Schedule - VAVCO 5-17 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 3	15.14531168	1	-15.1453	5	73.5	Interior	70.5	-1	-2.5	Closed
13	Occupied Schedule - VAVCO 5-15 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 2	16.14531168	2	-16.1453	5	73	Interior	70.5	-2	-3	Closed
14	Occupied Schedule - VAVCO 5-14 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 2	16.14531168	3	-16.1453	5	71.5	Interior	70	-3	-4	Closed
15	Occupied Schedule - VAVCO 5-07 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 3	15.14531168	0.499490759	-15.1453	5	74.5	Interior	72	0.499491	-1	Closed
16	Occupied Schedule - VAVCO 5-04 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 3	15.14531168	1	-15.1453	5	73	Exterior	71	-1	-2	Open
17	Occupied Schedule - VAVRH 6-10 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 3	15.14531168	0	-15.1453	6	75	Exterior	72	0	-0.5	Open
18	Occupied Schedule - VAVCO 1-02 -	1/2/19 6:59 AM	1/2/19 6:14 PM	AC Unit 1	15.145999	3	-15.146	1	70.5	Interior	69	-3	-4	Closed
19	Occupied Schedule - VAVRH 1-07 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 1	15.145999	2	-15.146	1	70.5	Exterior	69	-2	-3.5	Closed

Data Shaping Requirements

	B	C	D	I	J	K	M	N	O	P	Q	R	S	T
1	Building Shifts	Event Frame Start Time	Event Frame End Time	AC Unit	Absolute Difference	Absolute Offset f	Difference	Floor	Highest Zone Tempe	Interior/E	Lowest Zo	Offset fro	Offset fro	Open/Clo
2	Occupied Schedule - VAVCO 5-03 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 2	16.14084414	0.5	-16.1408	5	75.5	Interior	72	0.5	-2	Open
3	Occupied Schedule - VAVRH 5-10 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 3	15.14118779	1	-15.1412	5	73	Exterior	69	-1	-4	Open
4	Occupied Schedule - VAVCO 5-10 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 2	16.14127371	1	-16.1413	5	74	Interior	72.5	-1	-1.5	Closed
5	Occupied Schedule - VAVCO 5-11 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 2	16.14135962	2	-16.1414	5	74	Interior	71.5	-2	-2	Closed
6	Occupied Schedule - VAVCO 5-01 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 3	15.14513985	0	-15.1451	5	75	Interior	72	0	-0.5	Open
7	Occupied Schedule - VAVCO 5-16 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 3	16.14522577	3	-16.1452	5	72.5	Interior	69.5	-3	-3.5	Closed
9	Occupied Schedule - VAVRH 6-11 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 3	15.14531168	1.5	-15.1453	6	72.5	Exterior	69.5	-1.5	-3	Open
10	Occupied Schedule - VAVCO 6-01 -	1/2/19 6:59 AM	1/2/19 5:59 PM	AC Unit 3	15.14531168	0	-15.1453	6	75	Interior	72	0	0	Open
11	Occupied Schedule - VAVCO 5-18 -	1/2											-1	Closed
12	Occupied Schedule - VAVCO 5-17 -	1/2											-2.5	Closed
13	Occupied Schedule - VAVCO 5-15 -	1/2											-3	Closed
14	Occupied Schedule - VAVCO 5-14 -	1/2											-4	Closed
15	Occupied Schedule - VAVCO 5-07 -	1/2											-1	Closed
16	Occupied Schedule - VAVCO 5-04 -	1/2											-2	Open
17	Occupied Schedule - VAVRH 6-10 -	1/2											-0.5	Open
18	Occupied Schedule - VAVCO 1-02 -	1/2											-4	Closed
19	Occupied Schedule - VAVRH 1-07 -	1/2											-7.5	Closed

-Labels

-Groupings

-Templatize

-Time Series

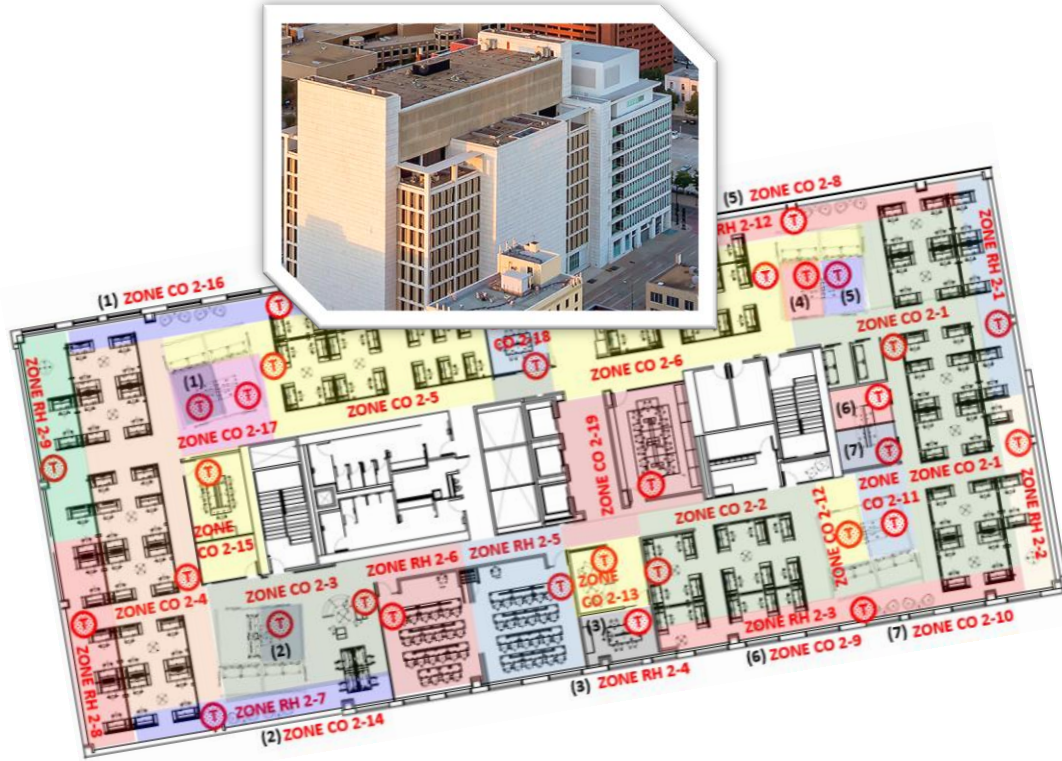
-Summaries

-Metadata

DEMO

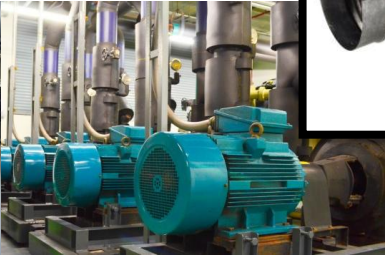
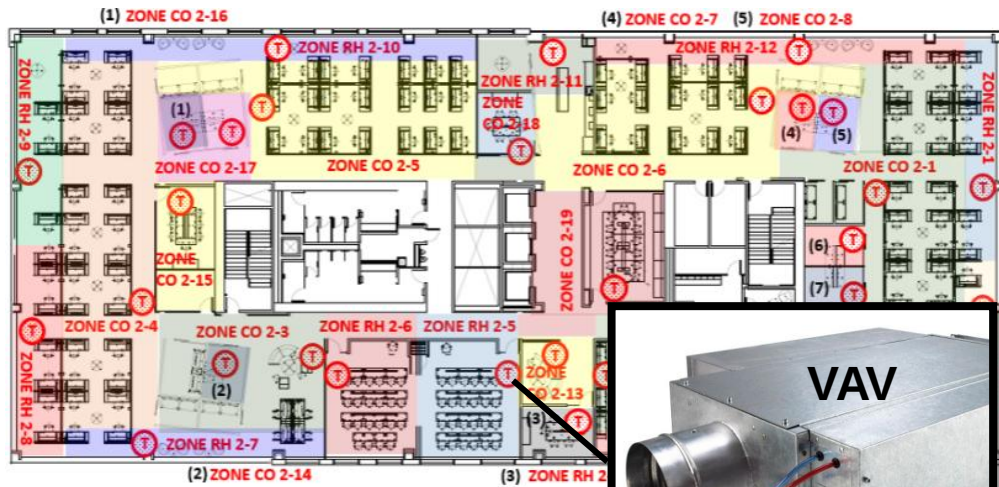
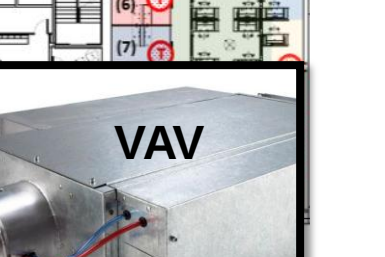
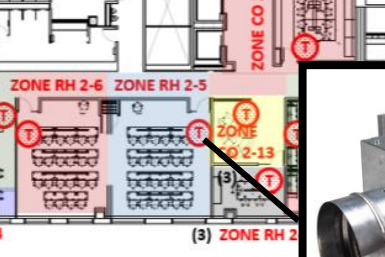
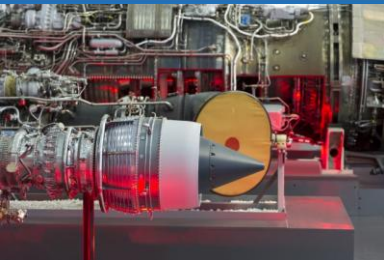
0. Check for Data Quality issues

Business Objective



- Zone temperature is **at desired setpoint** at the **start** of the occupancy
- Maintain zone temperature at desired setpoint **throughout the duration of the occupancy**

1. Determine your critical asset



2. Identify your data sources

3. Plan out the end format

Desired End Format

Asset	Date	Offset from Setpoint at Start	Average Offset from Setpoint	Difference from Outside Temperature	Floor	Quadrant
VAV Unit/Zone 1	1/1/2019	0	.3	14.5	1	North West
VAV Unit/Zone 2	1/1/2019	1.5	.7	17	1	South East
VAV Unit/Zone 1	1/2/2019	-1	-.2	4	1	North West

Asset- Any type of VAV Unit (VAVCO or VAVRH)

1. Zone Temperature
2. Setpoint
3. Occupied or Not

1. Difference in Zone Temperature to Setpoint
2. Absolute Difference in Zone Temperature to Setpoint
3. Floor
4. Quadrant
5. AC Unit
6. Interior/Exterior
7. Room/Open Space

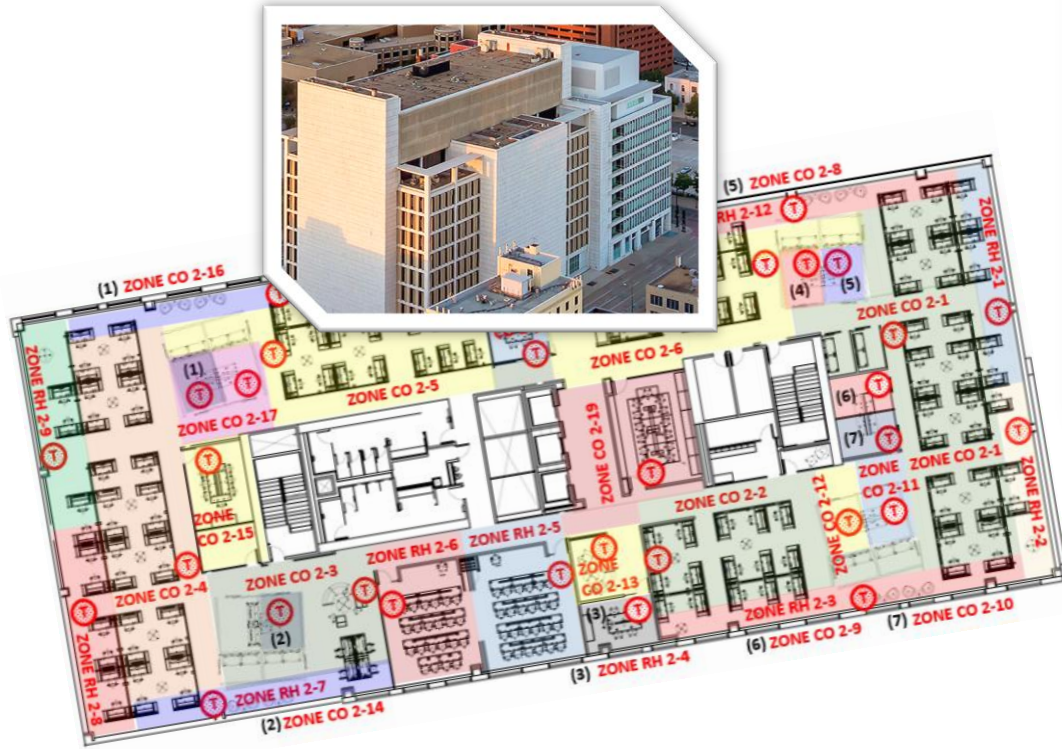
4. Build out AF Structure

5. Check for Data Quality issues

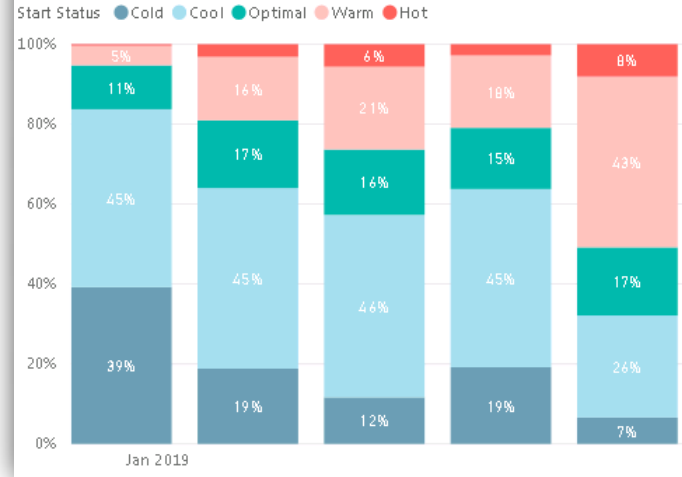
6. Setup Analytics & Event Frames

7. Create data table & analyze

Final Thoughts



Count of Zones by Start Status Binned by Week



Reflection: How will you apply these techniques in your PI System?

Speaker



Camille Metzinger

Data Scientist

OSIsoft, LLC

cmetzinger@osisoft.com

www.linkedin.com/in/camillemetzinger



Questions?

Please wait for
the **microphone**

State your
name & company



Please remember

TO DOWNLOAD
APP, SEARCH
OSISOFT



PARTNER & PRODUCT EXPO RECEPTION

5:00 – 7:00 PM

Located in both expo halls

Join us for appetizers and refreshments!