

# CSCI 622

Brock Gion

Electricity Consumption Explorer

	A	D	G	I
1	Bill Date	Electric Usage	Electric Charges	
2	08/29/2024	702	95.16	
3	07/31/2024	758	96.87	
4	07/05/2024	824	107.93	
5	06/02/2024	716	83.19	
6	05/01/2024	619	76.64	
7	04/02/2024	122	76.1	
8	03/05/2024	642	80.18	
9	02/04/2024	719	84.7	
10	01/03/2024	664	81.25	
11	12/03/2023	842	95.4	
12	10/31/2023	634	78.21	
13	10/02/2023	719	93.95	
14	08/31/2023	730	97.39	
15	08/02/2023	742	96.62	
16	07/04/2023	797	103.52	
17	06/04/2023	829	95.43	
18	05/03/2023	677	80.98	
19	04/04/2023	537	67.42	
20	03/06/2023	620	77.3	
21	02/02/2023	681	85.55	
22	01/03/2023	858	104.97	
23	11/30/2022	681	87.15	
24	10/30/2022	703	90.89	
25	09/29/2022	823	117.13	
26	08/30/2022	803	114.11	
27	08/01/2022	869	120.14	
28	06/30/2022	857	115.41	
29	06/01/2022	811	98.25	
30	05/02/2022	719	88.61	
31	04/03/2022	752	96.15	
32	03/03/2022	783	97.92	
33	02/01/2022	770	99.05	
34	01/03/2022	788	94.32	
35	11/30/2021	727	89.44	
36	10/28/2021	664	79.34	
37	09/29/2021	717	105.74	
38	08/30/2021	821	112.14	
39	08/01/2021	1139	152.64	
40	06/30/2021	893	112.95	
41	06/01/2021	786	94.96	
42	05/02/2021	782	86.22	
43	04/01/2021	560	72.76	
44	03/03/2021	759	87.75	
45	02/01/2021	792	96.58	
46	01/03/2021	1056	104.25	
47	11/30/2020	793	86.82	
48	10/28/2020	794	82.47	
49	09/29/2020	706	89.59	
50		745	\$94.41	
51				

# Main Project Goals

Use data to...

## 1. Understand my electricity consumption patterns at home

On average, over the last 4 years my **monthly electric usage is ~745kwh/month**. That's the only metric I know going into this project

## 2. Reduce carbon footprint

While I'd like to use more renewable sources of energy (i. solar/wind), options are limited in North Dakota. So, step 1, looking to become more efficient with my energy usage, then use this data to size a solar pv system FY2025.

# Main Project Goals

## 1. Understand my electricity consumption patterns at home

On average, over the last 4 years my **monthly electric usage is ~745kwh/month**.  
That's the only metric I know going into this project

**RESULT:** Yes! Using data now beginning to understand usage patterns of appliances

\*TBD: Still have unknowns, prefer detailed breakout of “office/livingroom” electronics

	total monthly electric usage in kwh -->	656.56	204.76	
		november	november %	december december %
appliances	dishwasher	16.53	2.5%	2.43 1.2%
	microwave	2.41	0.4%	0.58 0.3%
	fridge basement	45.05	6.9%	14.33 7.0%
	fridge kitchen	14.35	2.2%	4.33 2.1%
	garage freezer	22.02	3.4%	6.19 3.0%
	boiler (natural gas)	5.98	0.9%	0.84 0.4%
		106.32	16%	28.70 14%
consumer electronics	office / livingroom	80.70	12.3%	32.41 15.8%
	computer/monitors ??			
	TV ??			
	lights ??			



total monthly electric usage in kwh -->		656.56	204.76
	november	november %	december
appliances	dishwasher	16.53	2.5%
	microwave	2.41	0.4%
	fridge basement	45.05	6.9% <b>14.33</b> 7.0%
	fridge kitchen	14.35	2.2%
	garage freezer	22.02	3.4%
	boiler (natural gas)	5.98	0.9%
	106.32	16%	28.70 14%
consumer electronics	office / livingroom	80.70	12.3% 32.41 15.8%
	computer/monitors ??		
	TV ??		
	lights ??		

## Basement Fridge = Power Hog

- On average, this appliance is ~7% of total monthly electric usage
- Compared to the kitchen fridge, using almost 3x the energy!!



	total monthly electric usage in kwh -->		656.56	204.76	
	november	november %	december	december %	
appliances	dishwasher	16.53	2.5%	2.43	1.2%
	microwave	2.41	0.4%	0.58	0.3%
	fridge basement	45.05	6.9%	14.33	7.0%
	fridge kitchen	14.35	2.2%	4.33	2.1%
	garage freezer	22.02	3.4%	6.19	3.0%
	boiler (natural gas)	5.98	0.9%	0.84	0.4%
		106.32	16%	28.70	14%
consumer electronics	office / livingroom	80.70	12.3%	32.41	15.8%
	computer/monitors ??				
	TV ??				
	lights ??				

## Boiler = Efficient

- Brand new boiler was installed 2022 (less than 2 years old)
- This is a “natural gas” boiler, so electrical usage is very minimal
- Data shows it’s relatively low at under < 6kwh/month (less than ~1% of total monthly electric usage)



	total monthly electric usage in kwh -->	656.56	204.76
	november	november %	december december %
appliances	dishwasher	16.53	2.5% 1.2%
	microwave	2.41	0.4% 0.3%
	fridge basement	45.05	6.9% 7.0%
	fridge kitchen	14.35	2.2% 2.1%
	garage freezer	22.02	3.4% 3.0%
	boiler (natural gas)	5.98	0.9% 0.4%
		106.32	16% 14%
consumer electronics	office / livingroom	80.70	12.3% 15.8%
	computer/monitors ??		
	TV ??		
	lights ??		

## Dishwasher = Defunct

- 10+ year old dishwasher just went kaput, pump is worn out.  
Repair or replace?
- Key consideration: what's the electrical usage? How will it compare to a new dishwasher?
- Based on ~1 month of data, shows avg. usage 16.53 kWh/mo

U.S. Government

Federal law prohibits removal of this label before consumer purchase.



Dishwasher  
Capacity: Standard

Whirlpool Corporation

Models: ADB1400AM\*\*, KDFE104K\*\*\*, KDTE104K\*\*\*, KDFE204K\*\*\*, KDTE204K\*\*\*,  
KDTM404K\*\*\*, KDFM404K\*\*\*, KDTM604K\*\*\*, KDPM604K\*\*\*, KDTM704K\*\*\*, KDPM704K\*\*\*,  
KDTM804K\*\*\*, KDPM804K\*\*\*, MDB4949SK\*\*, MDB7959SK\*\*, MDB8959SK\*\*, MDB9979SK\*\*,  
MDB9959SK\*\*, WDP560HAM\*\*, WDT730HAM\*\*, WDP730HAM\*\*, WDT540HAM\*\*,  
WDP540HAM\*\*, WDT970SAK\*\*, WDPA70SAM\*\*, WDTA50SAK\*\*, WDT750SAK\*\*,  
WDT740SAL\*\*, WDTA80SAK\*\*

Estimated Yearly Energy Cost  
(when used with an electric water heater)

\$38



Cost Range of Similar Models

270 kWh

Estimated Yearly Electricity Use

\$27

Estimated Yearly Energy Cost  
(when used with a natural gas water heater)

\*Your cost will depend on your utility rates and use.

\*\*Cost range based only on standard capacity models.

\*\*\*Estimated operating cost based on four wash loads a week and a national average electricity cost of 14 cents per kWh and natural gas cost of \$1.21 per therm.

ftc.gov/energy

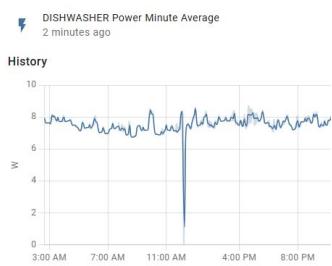
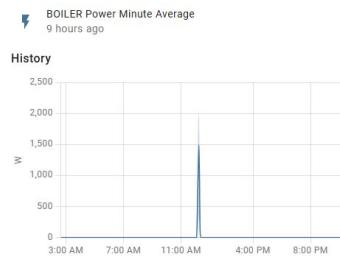
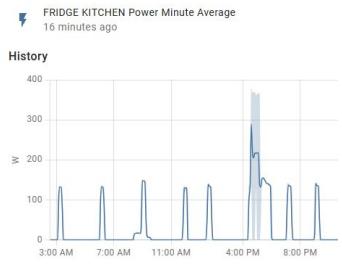
W11678098B

total monthly electric usage in kwh -->		656.56	204.76	
	november	november %	december	december %
appliances	dishwasher	16.53	2.5%	2.43
	microwave	2.41	0.4%	0.58
	fridge basement	45.05	6.9%	14.33
	fridge kitchen	14.35	2.2%	4.33
	garage freezer	22.02	3.4%	6.19
	boiler (natural gas)	5.98	0.9%	0.84
		106.32	16%	28.70
				14%
consumer electronics	office / livingroom	80.70	12.3%	32.41
	computer/monitors ??			
	TV ??			
	lights ??			

## Energy Guide Ratings

- Looking at a new dishwasher, est. usage is ~22 kwh/mo ( $270/12 = 22.5$ )
- [ftc.gov/energy](http://ftc.gov/energy)
  - The EnergyGuide label lets you know how much energy an appliance uses and see how it compares to the energy use of similar appliances.
- Appliance energy labeling mandated in 1980 (by the [Energy Policy and Conservation Act](#))

# “Energy Signatures” for each major appliance



- In a household, each appliance has a unique “energy signature”
- Key takeaway so far? Focus on the major appliances, then drill down into finer level detail of other electronics
- **Data-wise, now able to capture ALL the readings of all devices at household pulling electrical power. Display in a dashboard through Home Assistant**

Home Assistant

Electricity Consumption Explorer

Appliance	Power Minute Average	Energy This Month	Energy Today	Power Minute Average
DRYER	0.0 W	12.160 kWh	0.000 kWh	0.0 W
DRYER	0.0 W	10.293 kWh	0.004 kWh	0.0 W
FRIDGE KITCHEN	0.0 W	4.412 kWh	0.003 kWh	0.0 W
FRIDGE BASEMENT	0.0 W	14.611 kWh	1.794 kWh	181.9 W
GARAGE FREEZER	0.0 W	6.352 kWh	0.673 kWh	13.5 W
MICROWAVE	0.0 W	34.179 kWh	9.058 kWh	456.2 W
PELOTON ROOM	0.0 W	12.637 kWh	1.068 kWh	39.3 W
RANGE	0.0 W	3.890 kWh	0.468 kWh	0.0 W
SWORDFISH VUE	0.0 W	214.318 kWh	25.941 kWh	855.7 W
WATER HEATER	0.0 W	855.7 kWh	0.0 W	0.0 W

Developer tools

Settings

Notifications

# Main Project Goals

Use data to...

## 2. Reduce carbon footprint

While I'd like to use more renewable sources of energy (i. solar/wind), options are limited in North Dakota. So, step 1, looking to become more efficient with my energy usage, then use this data to size a solar pv system FY2025.

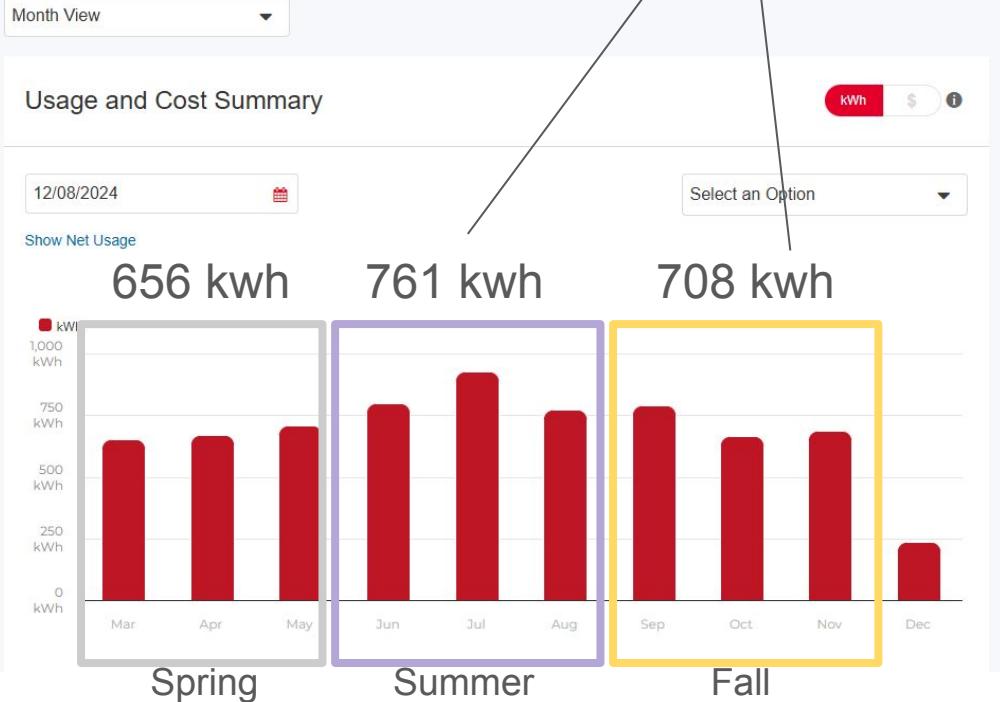
**RESULT:** Yes! Trending in right direction (latest 3 month avg. is **708kwh/month**).

Estimated yearly electricity usage FY2025? ~8,000 kWh/year

\*These results are preliminary at this point, gives confidence it's possible to improve

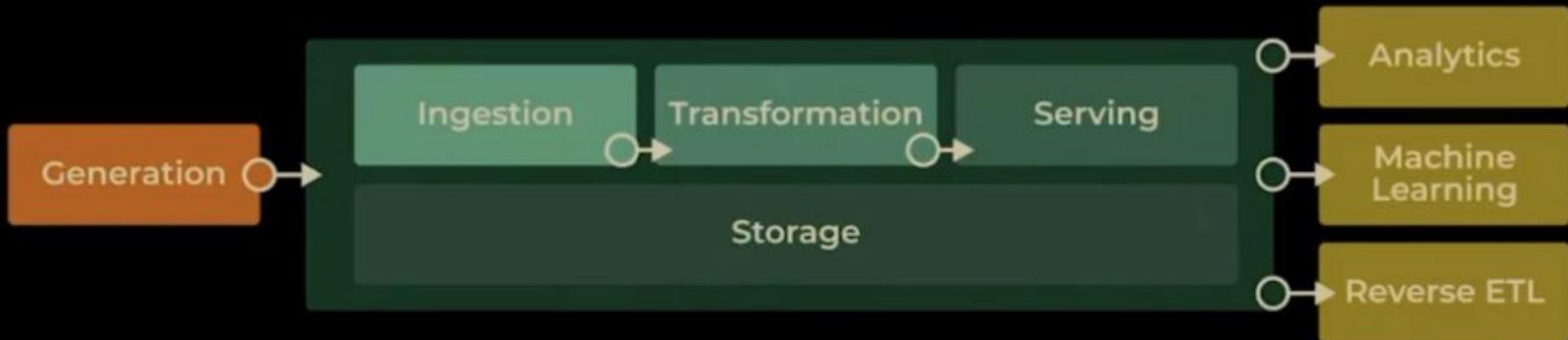
	A	D	G	I	J	K	L	M
1	Bill Date	Electric Usage	Electric Charges					
2	12/01/2024	736	87.56					
3	10/29/2024	591	72.39					
4	09/30/2024	799	102.6	708.67	<-- trailing ~3 month average, monthly kwh usage (sep, oct, nov)			
5	08/29/2024	702	95.16					
6	07/31/2024	758	96.87					
7	07/05/2024	824	107.93	761.33	<-- compared to ~3 month average, monthly kwh usage (june, july, aug)			

	A	D	G	I	J	K	L	M	
1	Bill Date	Electric Usage	Electric Charges						
2	12/01/2024	736	87.56						
3	10/29/2024	591	72.39						
4	09/30/2024	799	102.6	708.67	<-- trailing ~3 month average, monthly kwh usage (sep, oct, nov)				
5	08/29/2024	702	95.16						
6	07/31/2024	758	96.87						
7	07/05/2024	824	107.93	761.33	<-- compared to ~3 month average, monthly kwh usage (june, july, aug)				



- Reduction from ~761kwh to ~708kwh
- In last 3 months reduced average monthly electric usage by ~6.91%

# DATA ENGINEERING LIFECYCLE



## UNDERCURRENTS

Security

Data Management

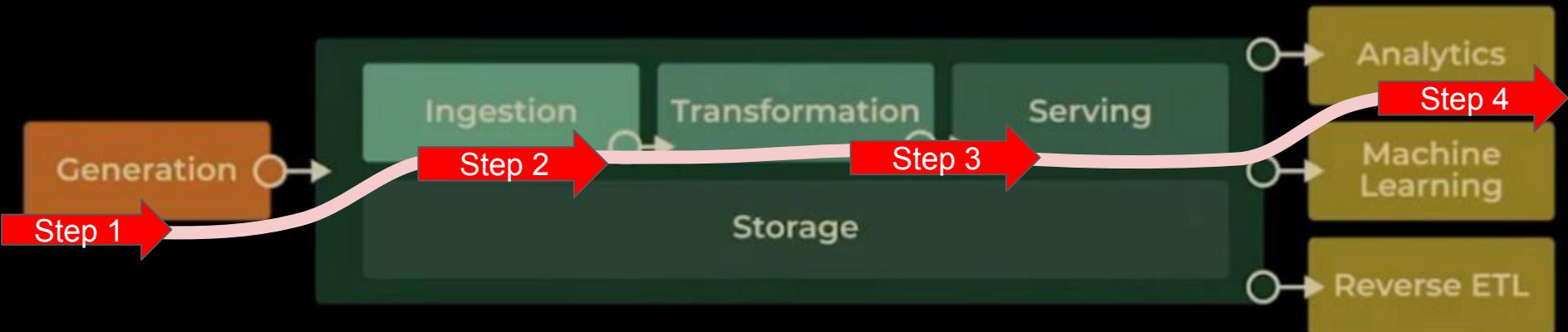
DataOps

Data Architecture

Orchestration

Software Engineering

# DATA ENGINEERING LIFECYCLE



## UNDERCURRENTS

Security

Data Management

DataOps

Data Architecture

Orchestration

Software Engineering

# DATA ENGINEERING LIFECYCLE



## UNDERCURRENTS

Security

Data Management

DataOps

Data Architecture

Orchestration

Software Engineering

# DATA



Security



# CYCLE

Analytics

Machine Learning

Reverse ETL

Software Engineering

# DATA ENGINEERING LIFECYCLE



## UNDERCURRENTS

Security

Data Management

DataOps

Data Architecture

Orchestration

Software Engineering

# DATA ENGINEERING LIFECYCLE



## UNDERCURRENTS

Security

Data Management

DataOps

Data Architecture

Orchestration

Software Engineering

# DATA ENGINEERING LIFECYCLE



## UNDERCURRENTS

Security

Data Management

DataOps

Data Architecture

Orchestration

Software Engineering

# DATA ENGINEERING LIFECYCLE



## UNDERCURRENTS

Security

Data Management

DataOps

Data Architecture

Orchestration

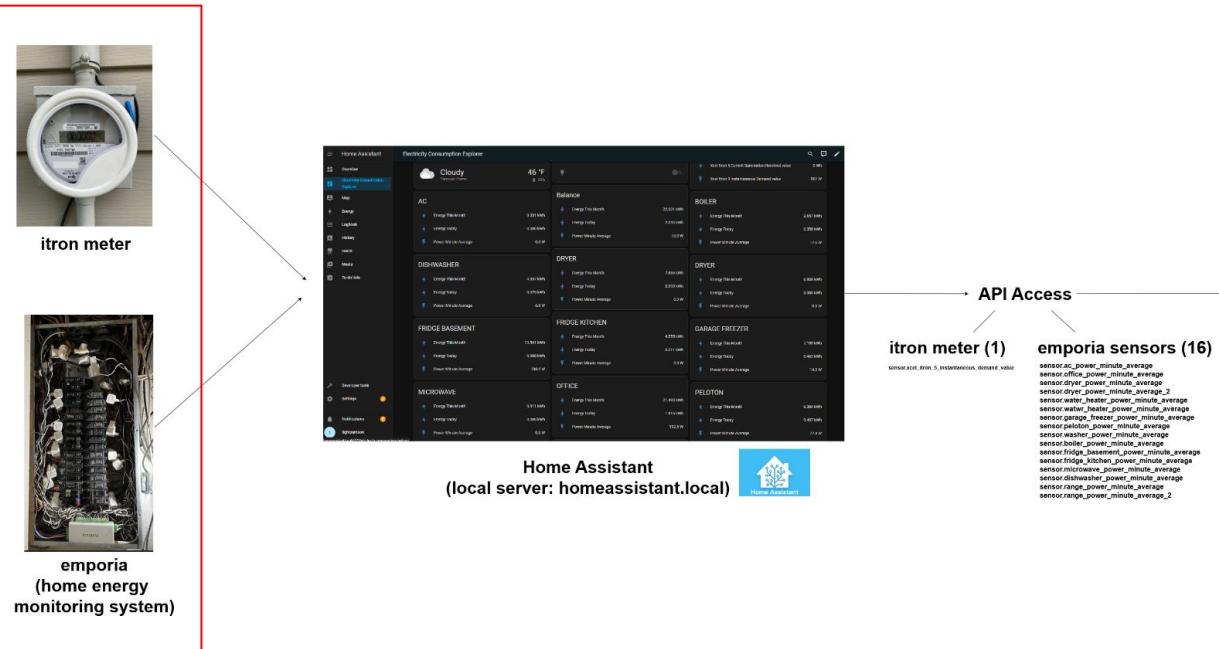
Software Engineering



# Step 1: DATA GENERATION

2 data sources:

- Itronmeter
- Emporia circuit monitoring



Run Powershell scripts  
to get data from API for  
for itron & emporia



# Step 1: DATA GENERATION

Itronmeter



A	B	C
entity_id	state	last_changed
1	entity_id	
2	sensor.xcel_itron_5_insta	534 2024-10-01T05:00:00.000Z
3	sensor.xcel_itron_5_insta	532 2024-10-01T05:00:11.292Z
4	sensor.xcel_itron_5_insta	533 2024-10-01T05:00:17.023Z
5	sensor.xcel_itron_5_insta	534 2024-10-01T05:00:22.851Z
6	sensor.xcel_itron_5_insta	535 2024-10-01T05:00:28.507Z
7	sensor.xcel_itron_5_insta	534 2024-10-01T05:00:34.317Z
8	sensor.xcel_itron_5_insta	533 2024-10-01T05:00:40.033Z
9	sensor.xcel_itron_5_insta	535 2024-10-01T05:00:57.545Z
10	sensor.xcel_itron_5_insta	534 2024-10-01T05:01:03.312Z
11	sensor.xcel_itron_5_insta	537 2024-10-01T05:01:20.319Z
12	sensor.xcel_itron_5_insta	536 2024-10-01T05:01:25.911Z
13	sensor.xcel_itron_5_insta	535 2024-10-01T05:01:31.643Z
14	sensor.xcel_itron_5_insta	534 2024-10-01T05:01:37.286Z
15	sensor.xcel_itron_5_insta	536 2024-10-01T05:01:48.748Z
16	sensor.xcel_itron_5_insta	534 2024-10-01T05:01:54.479Z
17	sensor.xcel_itron_5_insta	535 2024-10-01T05:02:00.302Z
18	sensor.xcel_itron_5_insta	533 2024-10-01T05:02:17.358Z
19	sensor.xcel_itron_5_insta	535 2024-10-01T05:02:23.050Z
20	sensor.xcel_itron_5_insta	534 2024-10-01T05:02:34.619Z
21	sensor.xcel_itron_5_insta	537 2024-10-01T05:02:51.925Z
22	sensor.xcel_itron_5_insta	536 2024-10-01T05:02:58.069Z
23	sensor.xcel_itron_5_insta	537 2024-10-01T05:03:03.803Z
24	sensor.xcel_itron_5_insta	538 2024-10-01T05:03:09.435Z
25	sensor.xcel_itron_5_insta	535 2024-10-01T05:03:15.274Z
26	sensor.xcel_itron_5_insta	536 2024-10-01T05:03:26.641Z
27	sensor.xcel_itron_5_insta	535 2024-10-01T05:03:32.289Z
28	sensor.xcel_itron_5_insta	537 2024-10-01T05:03:38.039Z
29	sensor.xcel_itron_5_insta	536 2024-10-01T05:03:49.375Z
30	sensor.xcel_itron_5_insta	538 2024-10-01T05:04:00.737Z
31	sensor.xcel_itron_5_insta	534 2024-10-01T05:04:06.780Z
32	sensor.xcel_itron_5_insta	536 2024-10-01T05:04:12.616Z
33	sensor.xcel_itron_5_insta	535 2024-10-01T05:04:18.269Z
34	sensor.xcel_itron_5_insta	538 2024-10-01T05:04:29.974Z
35	sensor.xcel_itron_5_insta	535 2024-10-01T05:04:35.861Z
36	sensor.xcel_itron_5_insta	534 2024-10-01T05:04:47.740Z
37	sensor.xcel_itron_5_insta	536 2024-10-01T05:04:53.507Z
38	sensor.xcel_itron_5_insta	538 2024-10-01T05:04:59.352Z
39	sensor.xcel_itron_5_insta	536 2024-10-01T05:05:04.984Z
40	sensor.xcel_itron_5_insta	535 2024-10-01T05:05:10.678Z
41	sensor.xcel_itron_5_insta	534 2024-10-01T05:05:22.454Z
42	sensor.xcel_itron_5_insta	535 2024-10-01T05:05:34.434Z
43	sensor.xcel_itron_5_insta	533 2024-10-01T05:05:46.007Z

- “instantaneous\_demand\_value” dimension

## 3 fields

- Entity\_id
  - constant for all values
- State
  - express in raw Watts
- Last\_changed
  - timestamp
- Typical daily file contains ~10K rows
- Filename example:

2024-11-30\_sensor.xcel\_itron\_5\_instantaneous\_demand\_value.csv



# Step 1: DATA GENERATION

## Emporia circuits



A	B	C
entity_id	state	last_changed
1	sensor.fridge	178.0326594 2024-11-30T00:00:00+00:00
2	sensor.fridge	177.490139 2024-11-30T00:01:23.335181+00:00
3	sensor.fridge	177.3737078 2024-11-30T00:03:23.311463+00:00
5	sensor.fridge	175.8437327 2024-11-30T00:05:23.322256+00:00
6	sensor.fridge	442.2252558 2024-11-30T00:07:23.301722+00:00
7	sensor.fridge	437.4649719 2024-11-30T00:08:23.274399+00:00
8	sensor.fridge	433.9307877 2024-11-30T00:10:23.318716+00:00
9	sensor.fridge	441.2823853 2024-11-30T00:12:23.274809+00:00
0	sensor.fridge	439.2026306 2024-11-30T00:14:23.302450+00:00
1	sensor.fridge	439.1165336 2024-11-30T00:16:23.292365+00:00
2	sensor.fridge	439.9620454 2024-11-30T00:18:23.250608+00:00
3	sensor.fridge	0 2024-11-30T00:20:23.334914+00:00
4	sensor.fridge	125.2476552 2024-11-30T00:45:23.345435+00:00
5	sensor.fridge	184.6673288 2024-11-30T00:47:23.268778+00:00
6	sensor.fridge	177.3083458 2024-11-30T00:49:23.257504+00:00
7	sensor.fridge	181.4390071 2024-11-30T00:51:23.327221+00:00
8	sensor.fridge	183.3081863 2024-11-30T00:53:23.289982+00:00
9	sensor.fridge	184.5748632 2024-11-30T00:54:23.275059+00:00
10	sensor.fridge	186.44223 2024-11-30T00:56:23.263174+00:00
11	sensor.fridge	187.7982132 2024-11-30T00:58:23.291337+00:00
12	sensor.fridge	189.2905632 2024-11-30T01:00:23.323283+00:00
13	sensor.fridge	189.7634816 2024-11-30T01:02:23.311581+00:00
14	sensor.fridge	189.3888064 2024-11-30T01:04:23.321184+00:00
15	sensor.fridge	189.2544736 2024-11-30T01:06:23.265118+00:00
16	sensor.fridge	188.3569164 2024-11-30T01:08:23.274888+00:00
17	sensor.fridge	187.9107101 2024-11-30T01:09:23.314393+00:00
18	sensor.fridge	186.8989825 2024-11-30T01:11:23.314559+00:00
19	sensor.fridge	186.0318344 2024-11-30T01:13:23.284544+00:00
20	sensor.fridge	182.306059 2024-11-30T01:15:23.285112+00:00
21	sensor.fridge	184.9601967 2024-11-30T01:17:23.289596+00:00
22	sensor.fridge	184.2817646 2024-11-30T01:19:23.341377+00:00
23	sensor.fridge	182.7784465 2024-11-30T01:21:23.315120+00:00
24	sensor.fridge	181.8428772 2024-11-30T01:23:23.344660+00:00
25	sensor.fridge	181.4086617 2024-11-30T01:25:23.270393+00:00
26	sensor.fridge	180.7264476 2024-11-30T01:27:23.261530+00:00
27	sensor.fridge	180.5448616 2024-11-30T01:29:23.328624+00:00
28	sensor.fridge	180.6258036 2024-11-30T01:31:23.271032+00:00
29	sensor.fridge	179.7921249 2024-11-30T01:33:23.287632+00:00
30	sensor.fridge	179.2163302 2024-11-30T01:35:23.272305+00:00
31	sensor.fridge	178.3731059 2024-11-30T01:37:23.307034+00:00
32	sensor.fridge	177.699367 2024-11-30T01:39:23.276999+00:00
33	sensor.fridge	177.0991007 2024-11-30T01:41:23.256905+00:00
34	sensor.fridge	177.1598831 2024-11-30T01:43:23.288004+00:00
35	sensor.fridge	177.5597557 2024-11-30T01:45:23.314539+00:00

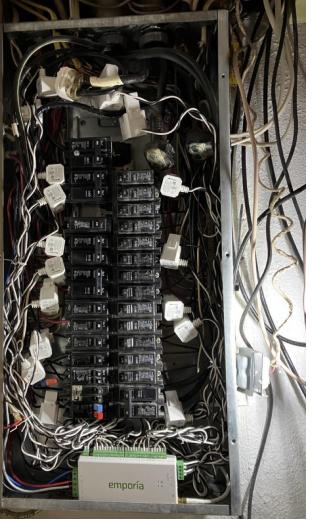
- 16 circuits being monitored

16circuits	Emporia Device Name	Home Assistant Name
CT-1	AC	sensor.ac_power_minute_average
CT-2	OFFICE	sensor.office_power_minute_average
CT-3	DRYER	sensor.dryer_power_minute_average
CT-4	DRYER	sensor.dryer_power_minute_average_2
CT-5	WATER HEATER	sensor.water_heater_power_minute_average
CT-6	WATWR HEATER	sensor.watwr_heater_power_minute_average
CT-7	GARAGE FREEZER	sensor.garage_freezer_power_minute_average
CT-8	PELOTON ROOM	sensor.peloton_power_minute_average
CT-9	WASHER	sensor.washer_power_minute_average
CT-10	BOILER	sensor.boiler_power_minute_average
CT-11	FRIDGE BASEMENT	sensor.fridge_basement_power_minute_average
CT-12	FRIDGE KITCHEN	sensor.fridge_kitchen_power_minute_average
CT-13	MICROWAVE	sensor.microwave_power_minute_average
CT-14	DISHWASHER	sensor.dishwasher_power_minute_average
CT-15	RANGE	sensor.range_power_minute_average
CT-16	RANGE	sensor.range_power_minute_average_2



# Step 1: DATA GENERATION

Emporia circuits



A	B	C
entity_id	state	last_changed
1	sensor.fridge	178.0326594 2024-11-30T00:00:00+00:00
2	sensor.fridge	177.490139 2024-11-30T00:01:23.335181+00:00
3	sensor.fridge	177.3737078 2024-11-30T00:03:23.311463+00:00
5	sensor.fridge	175.8437327 2024-11-30T00:05:23.322256+00:00
6	sensor.fridge	142.2252558 2024-11-30T00:07:23.301722+00:00
7	sensor.fridge	437.4649719 2024-11-30T00:08:23.274399+00:00
9	sensor.fridge	433.9307877 2024-11-30T00:10:23.318716+00:00
9	sensor.fridge	441.2823853 2024-11-30T00:12:23.274809+00:00
0	sensor.fridge	439.2026306 2024-11-30T00:14:23.302450+00:00
1	sensor.fridge	439.1165336 2024-11-30T00:16:23.292365+00:00
2	sensor.fridge	439.9620454 2024-11-30T00:18:23.256068+00:00
3	sensor.fridge	0 2024-11-30T00:20:23.334914+00:00
4	sensor.fridge	125.2476552 2024-11-30T00:45:23.345435+00:00
5	sensor.fridge	184.6673288 2024-11-30T00:47:23.268778+00:00
6	sensor.fridge	177.3083458 2024-11-30T00:49:23.257504+00:00
7	sensor.fridge	181.4390071 2024-11-30T00:51:23.327221+00:00
8	sensor.fridge	183.3081863 2024-11-30T00:53:23.289982+00:00
9	sensor.fridge	184.5748632 2024-11-30T00:54:23.275509+00:00
10	sensor.fridge	186.44223 2024-11-30T00:56:23.263174+00:00
11	sensor.fridge	187.7982132 2024-11-30T00:58:23.291337+00:00
12	sensor.fridge	189.2905632 2024-11-30T01:00:23.323283+00:00
13	sensor.fridge	189.7634816 2024-11-30T01:02:23.311581+00:00
14	sensor.fridge	189.3888064 2024-11-30T01:04:23.321184+00:00
15	sensor.fridge	189.2544736 2024-11-30T01:06:23.265118+00:00
16	sensor.fridge	188.35699164 2024-11-30T01:08:23.274888+00:00
17	sensor.fridge	187.9107101 2024-11-30T01:09:23.314393+00:00
18	sensor.fridge	186.8989825 2024-11-30T01:11:23.314559+00:00
19	sensor.fridge	186.0318344 2024-11-30T01:13:23.284544+00:00
10	sensor.fridge	182.306059 2024-11-30T01:15:23.285112+00:00
11	sensor.fridge	184.9601967 2024-11-30T01:17:23.289596+00:00
12	sensor.fridge	184.2817646 2024-11-30T01:19:23.341377+00:00
13	sensor.fridge	182.7784465 2024-11-30T01:21:23.315120+00:00
14	sensor.fridge	181.8428772 2024-11-30T01:23:23.344660+00:00
15	sensor.fridge	181.4086617 2024-11-30T01:25:23.270393+00:00
16	sensor.fridge	180.7264476 2024-11-30T01:27:23.261530+00:00
17	sensor.fridge	180.5448616 2024-11-30T01:29:23.328624+00:00
18	sensor.fridge	180.6258036 2024-11-30T01:31:23.271032+00:00
19	sensor.fridge	179.7921249 2024-11-30T01:33:23.287632+00:00
10	sensor.fridge	179.2163302 2024-11-30T01:35:23.272305+00:00
11	sensor.fridge	178.3731059 2024-11-30T01:37:23.307034+00:00
12	sensor.fridge	177.699367 2024-11-30T01:39:23.276999+00:00
13	sensor.fridge	177.0991007 2024-11-30T01:41:23.256905+00:00
14	sensor.fridge	177.1598831 2024-11-30T01:43:23.288004+00:00
15	sensor.fridge	177.5597557 2024-11-30T01:45:23.314539+00:00

- For each circuit can see power usage by the minute

- “power minute average” dimension

- **3 fields**

- **entity\_id**
- **state**
- **last\_changed**

- Typical file much smaller volume, < 500 rows

- Filename examples:

2024-11-30\_dishwasher\_power\_minute\_average.csv

2024-11-30\_fridge\_basement\_power\_minute\_average.csv

2024-11-30\_boiler\_power\_minute\_average.csv

Not secure | homeassistant.local:8123/electricity-consumption-explorer/0

Home Assistant Electricity Consumption Explorer

Energy distribution today



Grid 1.1 kWh → Home 1.1 kWh

GO TO THE ENERGY DASHBOARD

Cloudy Forecast Home 38 °F 84%

AC

Energy This Month	0.218 kWh
Energy Today	0.000 kWh
Power Minute Average	0.0 W

DISHWASHER

Energy This Month	2.473 kWh
Energy Today	0.012 kWh
Power Minute Average	7.6 W

FRIDGE BASEMENT

Energy This Month	14.751 kWh
Energy Today	0.071 kWh
Power Minute Average	184.5 W

FRIDGE KITCHEN

Energy This Month	4.447 kWh
Energy Today	0.000 kWh
Power Minute Average	0.0 W

Light

Balance

Energy This Month	39.678 kWh
Energy Today	0.280 kWh
Power Minute Average	156.5 W

BOILER

Energy This Month	0.844 kWh
Energy Today	0.000 kWh
Power Minute Average	0.0 W

DRYER

Energy This Month	12.160 kWh
Energy Today	0.000 kWh
Power Minute Average	0.0 W

DRYER

Energy This Month	10.293 kWh
Energy Today	0.000 kWh
Power Minute Average	0.0 W

GARAGE FREEZER

Energy This Month	6.419 kWh
Energy Today	0.078 kWh
Power Minute Average	13.4 W

Outlet 2 Right

Porch

tightpantson

Xcel Itron 5 Current Summation Delivered timePeriodduration 1

Xcel Itron 5 Current Summation Delivered time... 6 seconds ago

Xcel Itron 5 Current Summation Delivered value 6,086,904 Wh

Xcel Itron 5 Current Summation Received timePeriodduration 1

Xcel Itron 5 Current Summation Received time... 6 seconds ago

Xcel Itron 5 Current Summation Received value 0 Wh

Xcel Itron 5 Instantaneous Demand value 864 W

Grid 1.1 kWh → Home 1.1 kWh

Cloudy Forecast Home 38 °F 84%

AC

Energy This Month	0.218 kWh
Energy Today	0.000 kWh
Power Minute Average	0.0 W

DISHWASHER

Energy This Month	2.473 kWh
Energy Today	0.012 kWh
Power Minute Average	7.6 W

FRIDGE BASEMENT

Energy This Month	14.751 kWh
Energy Today	0.071 kWh
Power Minute Average	184.5 W

FRIDGE KITCHEN

Energy This Month	4.447 kWh
Energy Today	0.000 kWh
Power Minute Average	0.0 W

Light

Balance

Energy This Month	39.678 kWh
Energy Today	0.280 kWh
Power Minute Average	156.5 W

BOILER

Energy This Month	0.844 kWh
Energy Today	0.000 kWh
Power Minute Average	0.0 W

DRYER

Energy This Month	12.160 kWh
Energy Today	0.000 kWh
Power Minute Average	0.0 W

DRYER

Energy This Month	10.293 kWh
Energy Today	0.000 kWh
Power Minute Average	0.0 W

GARAGE FREEZER

Energy This Month	6.419 kWh
Energy Today	0.078 kWh
Power Minute Average	13.4 W

Not secure | homeassistant.local:8123/electricity-consumption-explorer/0

Home Assistant Electricity Consumption Explorer

Energy distribution today

Grid: 1.1 kWh → Home: 1.1 kWh

GO TO THE ENERGY DASHBOARD

Cloudy Forecast Home 38 °F 84%

AC

Energy This Month	0.218 kWh	
Energy Today	0.000 kWh	
Power Minute Average	0.0 W	

DISHWASHER

Energy This Month	2.473 kWh	
Energy Today	0.012 kWh	
Power Minute Average	7.6 W	

2.5%

FRIDGE BASEMENT

Energy This Month	14.751 kWh	
Energy Today	0.071 kWh	
Power Minute Average	184.5 W	

7%

FRIDGE KITCHEN

Energy This Month	4.447 kWh	
Energy Today	0.000 kWh	
Power Minute Average	0.0 W	

OUTLET 2 RIGHT

Porch

tightpantson

Away

Xcel Itron 5 Current Summation Delivered timePeriodduration 1

Xcel Itron 5 Current Summation Delivered time... 6 seconds ago

Xcel Itron 5 Current Summation Delivered value 6,086,904 Wh

Xcel Itron 5 Current Summation Received timePeriodduration 1

Xcel Itron 5 Current Summation Received time... 6 seconds ago

Xcel Itron 5 Current Summation Received value 0 Wh

Xcel Itron 5 Instantaneous Demand value 864 W

Light

BOILER

Energy This Month	0.844 kWh	
Energy Today	0.000 kWh	
Power Minute Average	0.0 W	

1%

DRYER

Energy This Month	12.160 kWh	
Energy Today	0.000 kWh	
Power Minute Average	0.0 W	

DRYER

Energy This Month	10.293 kWh	
Energy Today	0.000 kWh	
Power Minute Average	0.0 W	

GARAGE FREEZER

Energy This Month	6.419 kWh	
Energy Today	0.078 kWh	
Power Minute Average	13.4 W	

Developer tools

Settings 2

Notifications 2

# DATA ENGINEERING LIFECYCLE



## UNDERCURRENTS

Security

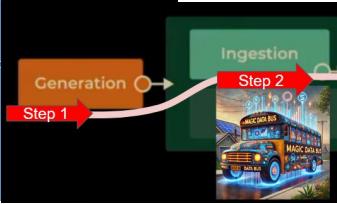
Data Management

DataOps

Data Architecture

Orchestration

Software Engineering



## Step 2: DATA INGESTION

- 2 Powershell scripts for data ingestion:
- `get_emporiacircuits_data.ps1`
  - `get_itronmeter_data.ps1`



itron meter



emporia  
(home energy  
monitoring system)

**Home Assistant**  
(local server: `homeassistant.local`)

**Electricity Consumption Explorer**

Category	Appliance	Energy This Month	Energy Today	Power Min/Mean Average
DISHWASHER	Energy This Month	0.00 kWh	0.00 kWh	0.00 W
	Energy Today	0.00 kWh	0.00 kWh	0.00 W
	Power Min/Mean Average	0.00 W	0.00 W	0.00 W
DRYER	Energy This Month	1.000 kWh	0.000 kWh	0.000 W
	Energy Today	0.000 kWh	0.000 kWh	0.00 W
	Power Min/Mean Average	0.00 W	0.00 W	0.00 W
FRIDGE KITCHEN	Energy This Month	4,000 kWh	0.000 kWh	0.000 W
	Energy Today	0.000 kWh	0.000 kWh	0.00 W
	Power Min/Mean Average	0.00 W	0.00 W	0.00 W
MICROWAVE	Energy This Month	0.000 kWh	0.000 kWh	0.00 W
	Energy Today	0.000 kWh	0.000 kWh	0.00 W
	Power Min/Mean Average	0.00 W	0.00 W	0.00 W
OFFICE	Energy This Month	2,400 kWh	0.000 kWh	0.000 W
	Energy Today	0.000 kWh	0.000 kWh	0.00 W
	Power Min/Mean Average	0.00 W	0.00 W	0.00 W
PELOTON	Energy This Month	0.000 kWh	0.000 kWh	0.000 W
	Energy Today	0.000 kWh	0.000 kWh	0.00 W
	Power Min/Mean Average	0.00 W	0.00 W	0.00 W

API Access

itron meter (1)

`sensor.acit_itree_5_instantaneous_demand_value`

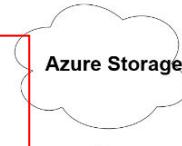
emporia sensors (16)

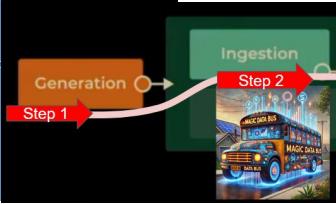
```

sensor.acit_power_minute_average
sensor.acit_water_minute_average
sensor.dryer_power_minute_average
sensor.dryer_water_minute_average
sensor.heater_water_minute_average
sensor.heater_water_minute_mean_average
sensor.heater_water_minute_min_average
sensor.heater_water_minute_max_average
sensor.peloton_power_minute_average
sensor.washer_power_minute_average
sensor.boiler_power_minute_average
sensor.fridge_kitchen_power_minute_average
sensor.microwave_power_minute_average
sensor.dishwasher_power_minute_average
sensor.rings_power_minute_average
sensor.rings_power_minute_mean_average
sensor.rings_power_minute_max_average
sensor.rings_power_minute_min_average

```

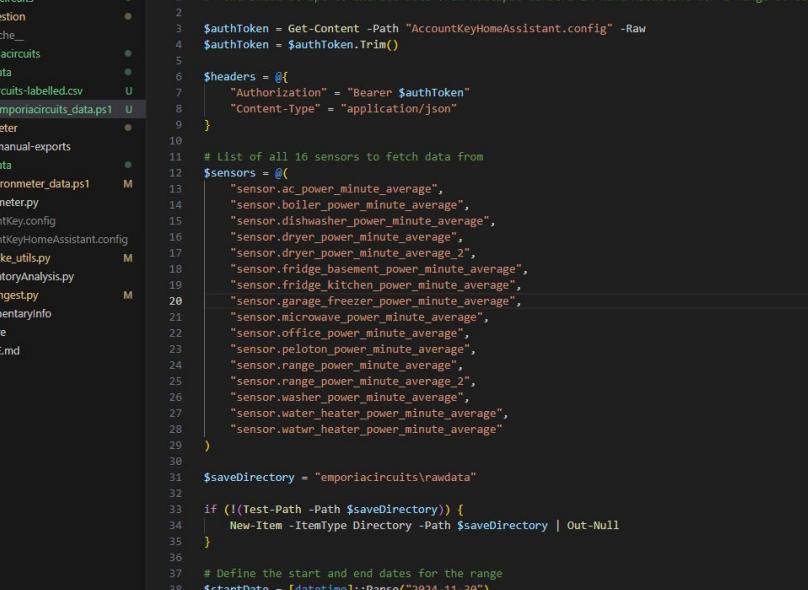
Run Powershell scripts  
to get data from API for  
for itron & empiria





## Step 2: **DATA INGESTION**

## get\_emporiacircuits\_data.ps1



The screenshot shows a terminal window with the title bar "f24-project-brockgion". The command entered is:

```
PS C:\Users\envo\Documents\CSCI622\NEU\wk18\f24-project-brockgion\src\ingestion> .\emporiacircuits\get_emporiacircuits_data.ps1
```

The output of the script is displayed below:

```
>_get_emporiacircuits_data.ps1 U >_get_itronmeter_data.ps1 M ⚡ data_lake_utils.py M ⚡ main_ingest.py M
```

```
src > ingestion > emporiacircuits > _get_emporiacircuits_data.ps1 ...
```

```
1 # PowerShell script to extract data from multiple sensors in Home Assistant for a range of dates
2
3 $authToken = Get-Content -Path "AccountKeyHomeAssistant.config" -Raw
4 $authToken = $authToken.Trim()
5
6 $headers = @{
7     "Authorization" = "Bearer $authToken"
8     "Content-Type" = "application/json"
9 }
10
11 # List of all 16 sensors to fetch data from
12 $sensors = @(
13     "sensor.ac_power_minute_average",
14     "sensor.boiler_power_minute_average",
15     "sensor.dishwasher_power_minute_average",
16     "sensor.dryer_power_minute_average",
17     "sensor.dryer_power_minute_average_2",
18     "sensor.fridge_basement_power_minute_average",
19     "sensor.fridge_kitchen_power_minute_average",
20     "sensor.garage_freezer_power_minute_average",
21     "sensor.microwave_power_minute_average",
22     "sensor.office_power_minute_average",
23     "sensor.peloton_power_minute_average",
24     "sensor.range_power_minute_average",
25     "sensor.range_power_minute_average_2",
26     "sensor.washer_power_minute_average",
27     "sensor.water_heater_power_minute_average",
28     "sensor.watwr_heater_power_minute_average"
29 )
30
31 $saveDirectory = "emporiacircuits\rawdata"
32
33 if (!(Test-Path -Path $saveDirectory)) {
34     New-Item -ItemType Directory -Path $saveDirectory | Out-Null
35 }
36
37 # Define the start and end dates for the range
38 $startDate = [datetime]::Parse("2024-11-30")
39 $endDate = [datetime]::Parse("2024-11-30")
40
```

- Manually mapped all 16 sensors as hardcoded list
    - Although this list could change, don't expect it to occur often
  - Single script gets all data for 16 sensors, and uploads to Azure blob storage container
    - Key parameter is date range

```
# Define the start and end dates for the range
$startDate = [datetime]::Parse("2024-11-30")
$endDate = [datetime]::Parse("2024-11-30")
```

# DATA ENGINEERING LIFECYCLE



## UNDERCURRENTS

Security

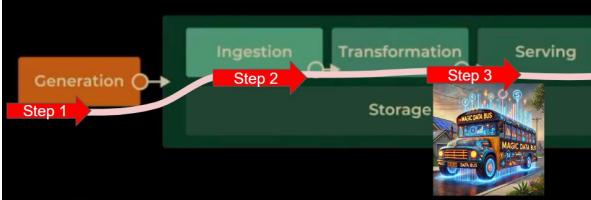
Data Management

DataOps

Data Architecture

Orchestration

Software Engineering



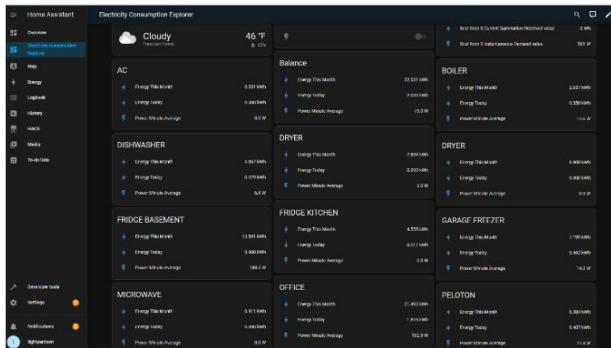
## Step 3: DATA STORAGE



itrон meter



emporia  
(home energy  
monitoring system)



Home Assistant  
(local server: homeassistant.local)



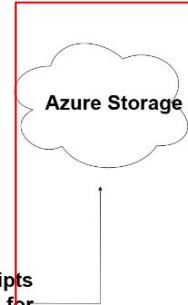
**API Access**

- itrон meter (1)**  
sensor.ac0\_itron\_5\_instantaneous\_demand\_value
- emporia sensors (16)**

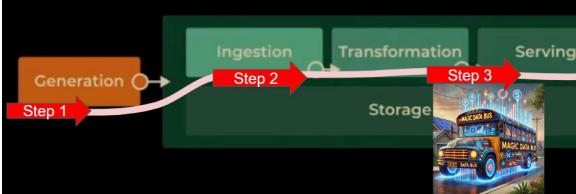
```

sensor_ac_power_minute_average
sensor_dryer_power_minute_average
sensor_dryer_water_power_minute_average
sensor_water_heater_power_minute_average
sensor_water_heater_water_power_minute_average
sensor_water_water_minute_average
sensor_pelton_power_minute_average
sensor_washer_power_minute_average
sensor_boiler_power_minute_average
sensor_fridge_kitchen_power_minute_average
sensor_microwave_power_minute_average
sensor_dishwasher_power_minute_average
sensor_range_power_minute_average
sensor_range_power_minute_average_2
    
```

Cloud storage:  
- Azure Blob container



Run Powershell scripts  
to get data from API for  
for itron & emporia



# Step 3: DATA STORAGE

The diagram illustrates the Data Pipeline flow, divided into four main stages: Generation, Ingestion, Transformation, and Serving. A red arrow labeled "Step 1" points to the "Generation" stage. Another red arrow labeled "Step 2" points to the "Transformation" stage. A third red arrow labeled "Step 3" points to the "Storage" stage, which is visually represented by a yellow school bus labeled "MAGIC DATA BUS".

[https://portal.azure.com/#view/Microsoft\\_Azure\\_Storage/ContainerMenuBlade/~/azbrickstorage%20Containers%20electric-meter-data](https://portal.azure.com/#view/Microsoft_Azure_Storage/ContainerMenuBlade/~/azbrickstorage%20Containers%20electric-meter-data)

**Microsoft Azure**

Home > azbrickstorage | Containers >

**electric-meter-data** Container

Search

Overview

Diagnose and solve problems

Access Control (IAM)

Settings

- Shared access tokens
- Manage ACL
- Access policy
- Properties
- Metadata

Upload Add Directory Refresh Rename

Authentication method: Access key (Switch to Microsoft Entra user account)  
Location: electric-meter-data / raw

Search blobs by prefix (case-sensitive)

Name
<input type="checkbox"/> [..]
<input type="checkbox"/> emporiacircuits
<input type="checkbox"/> itronmeter

Container Menu Blade Overview

Search

Upload Add Directory Refresh Rename

Authentication method: Access key (Switch to Microsoft Entra user account)  
Location: electric-meter-data / raw / itronmeter

Search blobs by prefix (case-sensitive)

Overview

Diagnose and solve problems

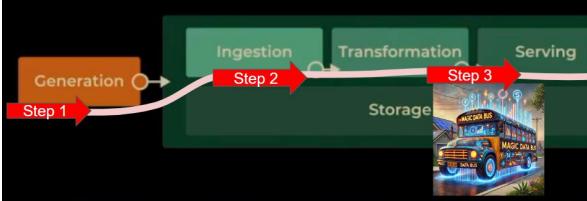
Access Control (IAM)

Settings

- Shared access tokens
- Manage ACL
- Access policy
- Properties
- Metadata

Name

Name
<input type="checkbox"/> 2024-10-01_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-02_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-03_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-04_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-05_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-06_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-07_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-08_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-09_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-10_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-11_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-12_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-13_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-14_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-15_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-16_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-17_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-18_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-19_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-20_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-21_sensor.xcel_itron_5_instantaneous_demand_value.csv
<input type="checkbox"/> 2024-10-22_sensor.xcel_itron_5_instantaneous_demand_value.csv



# Step 3: DATA TRANSFORMATION - Azure Databricks

**Exploratory Analysis Electricity Consumption**

File Edit View Run Help Last edit was 24 days ago

**Connect to Azure Storage**

```

Waiting | Waiting for compute to start
2

# Add imports, read in data to analyze
import json
import matplotlib.pyplot as plt
from pyspark.sql.functions import sum, desc, isnull, col, avg, count, when, from_json, to_timestamp, window
from pyspark.sql.types import StructType, StructField, IntegerType, StringType

storage_account_name = "azdbrickstorage" # Your storage account name
my_scope = "azdbrick-scope" # The secret scope you created
my_key = "electricSecret" # The secret key name in your vault

# Configure Spark to use the storage account key from your secret scope
spark.conf.set(
    f"fs.azure.account.key.{storage_account_name}.dfs.core.windows.net",
    dbutils.secrets.get(scope=my_scope, key=my_key)
)

# Define the URI for accessing the data
uri = f"abfss://electric-meter-data{storage_account_name}.dfs.core.windows.net/"

# itron
itrон_df = spark.read.csv(uri+"raw-input/xcel_itron_5_instantaneous_demand_value-2024-11-13.csv", header=True)

# emporia sensors
sensors = [
    "ac_power_minute_average",
    "office_power_minute_average",
    "dryer_power_minute_average",
    "dryer_power_minute_average_2",
    "water_heater_power_minute_average",
    "water_heater_power_minute_average",
    "garage_freezer_power_minute_average",
    "peloton_power_minute_average",
    "washer_power_minute_average",
    "boiler_power_minute_average",
    "fridge_basement_power_minute_average"
]

```

Table      Visualization 1 +

average\_state

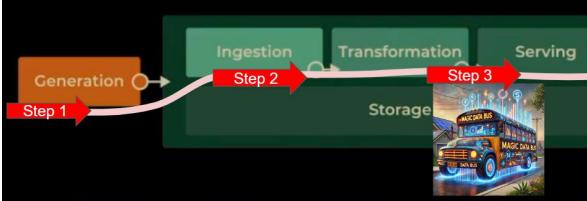
Nov 30, 2024 00:00 Nov 30, 2024 03:00 Nov 30, 2024 06:00 Nov 30, 2024 09:00 Nov 30, 2024 12:00 Nov 30, 2024 15:00 Nov 30, 2024 18:00 Nov 30, 2024 21:00 Nov 30, 2024 24:00 Dec 01, 2024 00:00

window\_start

1.440 rows

Refreshed now

The screenshot shows the Azure Databricks workspace interface. On the left, the sidebar includes options like Workspace, Recents, Catalog, Workflows, Compute, SQL, and Machine Learning. The main notebook area displays Python code for reading data from Azure Storage and performing initial analysis. To the right, a visualization titled 'Visualization 1' shows a line chart of 'average\_state' over time, with data points corresponding to the processed data frames shown above.



# Step 3: DATA TRANSFORMATION - Azure Databricks

**Exploratory Analysis Electricity Consumption**

File Edit View Run Help Last edit was 24 days ago

**Connect to Azure Storage**

```

Waiting | Waiting for compute to start
2

# Add imports, read in data to analyze
import json
import matplotlib.pyplot as plt
from pyspark.sql.functions import sum, desc, isnull, col, avg, count, when, from_json, to_timestamp, window
from pyspark.sql.types import StructType, StructField, IntegerType, StringType

storage_account_name = "azdbrickstorage" # Your storage account name
my_scope = "azdbrick-scope" # The secret scope you created
my_key = "electricSecret" # The secret key name in your vault

# Configure Spark to use the storage account key from your secret scope
spark.conf.set(
    f"fs.azure.account.key.{storage_account_name}.dfs.core.windows.net",
    dbutils.secrets.get(scope=my_scope, key=my_key)
)

# Define the URI for accessing the data
uri = f"abfss://electric-meter-data{storage_account_name}.dfs.core.windows.net/"

# itron
itrон_df = spark.read.csv(uri+"raw-input/xcel_itron_5_instantaneous_demand_value-2024-11-13.csv", header=True)

# emporia sensors
sensors = [
    "ac_power_minute_average",
    "office_power_minute_average",
    "dryer_power_minute_average",
    "dryer_power_minute_average_2",
    "water_heater_power_minute_average",
    "water_heater_power_minute_average",
    "garage_freezer_power_minute_average",
    "peloton_power_minute_average",
    "washer_power_minute_average",
    "boiler_power_minute_average",
    "fridge_basement_power_minute_average"
]

```

average\_state

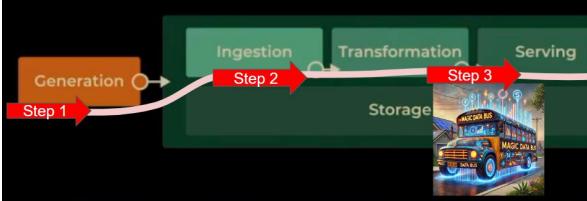
Nov 30, 2024 00:00 Nov 30, 2024 03:00 Nov 30, 2024 06:00 Nov 30, 2024 09:00 Nov 30, 2024 12:00 Nov 30, 2024 15:00 Nov 30, 2024 18:00 Nov 30, 2024 21:00 Dec 01, 2024 00:00

window\_start

Refreshed now

1.440 rows

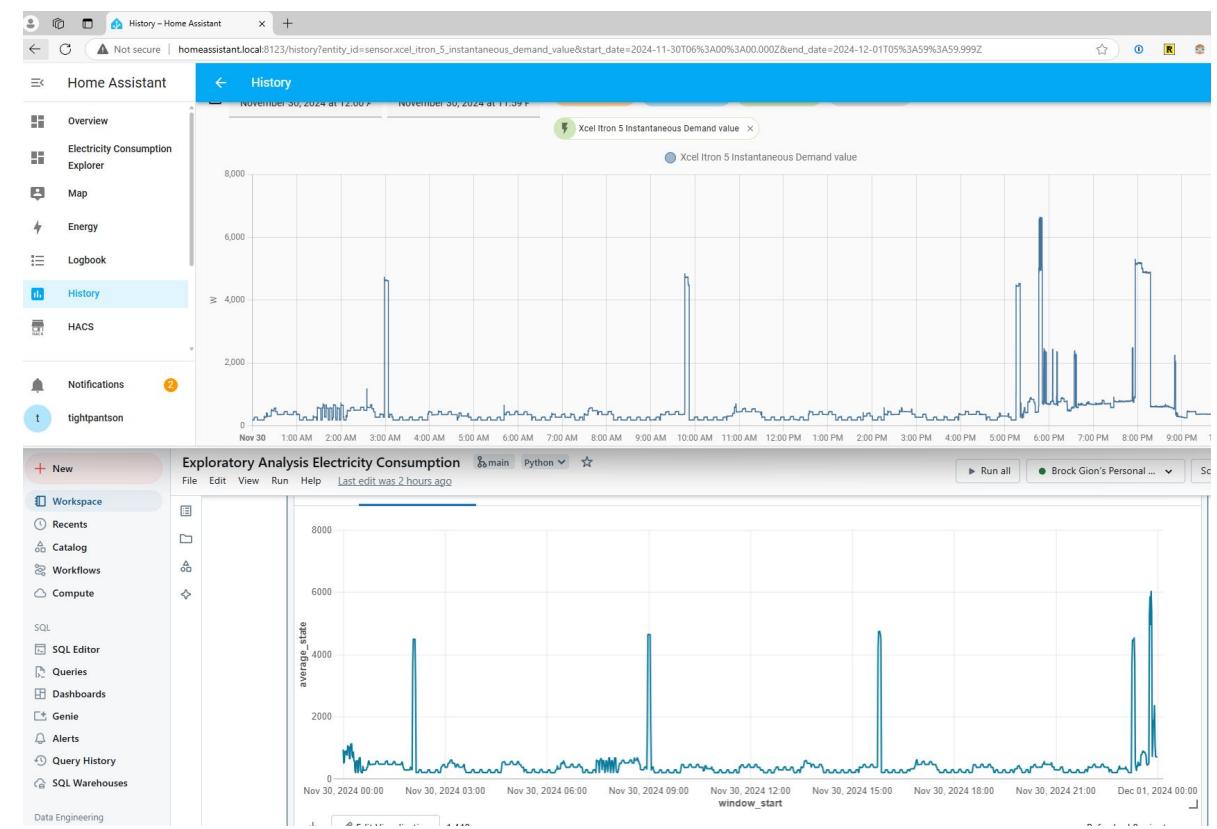
This screenshot shows an Azure Databricks workspace titled 'Exploratory Analysis Electricity Consumption'. On the left, the sidebar lists various notebooks, datasets, and services. The main notebook contains Python code for reading data from Azure Storage and performing initial analysis. To the right, a visualization titled 'Visualization 1' displays a line chart of 'average\_state' over time, showing several sharp peaks (spikes) against a low baseline. The x-axis represents time from November 30, 2024, to December 1, 2024, with major ticks every 3 hours. The y-axis ranges from 0 to 8000. The chart highlights significant fluctuations in electricity consumption.

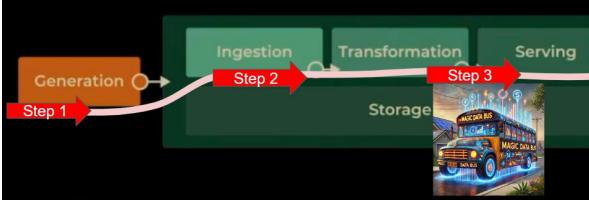


# Step 3: DATA TRANSFORMATION - Azure Databricks

The screenshot shows the Home Assistant interface with the "History" tab selected. The main area displays a calendar for November 2024, with the date range set from "November 30, 2024 at 12:00 AM" to "November 30, 2024 at 11:59 PM". Below the calendar, there are buttons for "CANCEL" and "SELECT". On the left sidebar, the "History" tab is highlighted.

**Start Date: November 30, 2024 12:00 AM**  
**End Date: November 30, 2024 11:59 PM**





## Step 3: DATA TRANSFORMATION - Azure Databricks

**databricks**

Electricity Consumption Explorer - Data Transformations    main    Python    ⚡

File Edit View Run Help Last edit was 2 hours ago

itron (meter data)

Add "date" field

```
itron_df = itron_df.withColumn("date", col("last_changed").substr(1, 10))
display(itron_df)
```

(1) Spark Jobs

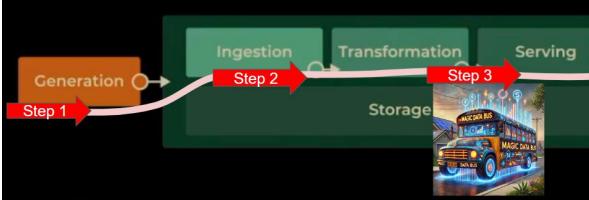
itron\_df: pyspark.sql.dataframe.DataFrame = [entity\_id: string, state: string ... 2 more fields]

Table

#	entity_id	state	last_changed	date
63	sensor.xcel_itron_5_instantaneous_demand_val...	756	2024-11-30T00:07:02.979113+00:00	2024-11-30
64	sensor.xcel_itron_5_instantaneous_demand_val...	754	2024-11-30T00:07:08.916010+00:00	2024-11-30
65	sensor.xcel_itron_5_instantaneous_demand_val...	758	2024-11-30T00:07:14.753205+00:00	2024-11-30
66	sensor.xcel_itron_5_instantaneous_demand_val...	756	2024-11-30T00:07:20.385758+00:00	2024-11-30
67	sensor.xcel_itron_5_instantaneous_demand_val...	754	2024-11-30T00:07:26.113857+00:00	2024-11-30
68	sensor.xcel_itron_5_instantaneous_demand_val...	757	2024-11-30T00:07:37.588677+00:00	2024-11-30
69	sensor.xcel_itron_5_instantaneous_demand_val...	753	2024-11-30T00:07:43.426128+00:00	2024-11-30
70	sensor.xcel_itron_5_instantaneous_demand_val...	755	2024-11-30T00:07:49.163928+00:00	2024-11-30
71	sensor.xcel_itron_5_instantaneous_demand_val...	756	2024-11-30T00:07:55.023884+00:00	2024-11-30
72	sensor.xcel_itron_5_instantaneous_demand_val...	1146	2024-11-30T00:08:00.835993+00:00	2024-11-30
73	sensor.xcel_itron_5_instantaneous_demand_val...	1145	2024-11-30T00:08:06.579026+00:00	2024-11-30
74	sensor.xcel_itron_5_instantaneous_demand_val...	1144	2024-11-30T00:08:12.355705+00:00	2024-11-30
75	sensor.xcel_itron_5_instantaneous_demand_val...	1021	2024-11-30T00:08:17.935158+00:00	2024-11-30
76	sensor.xcel_itron_5_instantaneous_demand_val...	1017	2024-11-30T00:08:23.623596+00:00	2024-11-30
77	sensor.xcel_itron_5_instantaneous_demand_val...	1015	2024-11-30T00:08:29.227205+00:00	2024-11-30

10,000+ rows | Truncated data | 1.11 seconds runtime

- Adding a simple “date” field, very easy and useful
- Example: extracting YYYY-MM-DD from “last\_changed”
  - last\_changed:
  - 2024-11-30T00:00:20.677Z
- date:
  - 2024-11-30



## Step 3: DATA TRANSFORMATION - Azure Databricks

**databricks**

Electricity Consumption Explorer - Data Transformations    main    Python    ⚡

File Edit View Run Help Last edit was 2 hours ago

**itron (meter data)**

Add "date" field

```
03:25 AM (1s)
itron_df = itron_df.withColumn("date", col("last_changed").substr(1, 10))
display(itron_df)
```

(1) Spark Jobs

```
itron_df: pyspark.sql.dataframe.DataFrame = [entity_id: string, state: string ... 2 more fields]
```

Table

#	entity_id	state	last_changed	date
63	sensor.xcel_itron_5_instantaneous_demand_val...	756	2024-11-30T00:07:02.979113+00:00	2024-11-30
64	sensor.xcel_itron_5_instantaneous_demand_val...	754	2024-11-30T00:07:08.916010+00:00	2024-11-30
65	sensor.xcel_itron_5_instantaneous_demand_val...	758	2024-11-30T00:07:14.753205+00:00	2024-11-30
66	sensor.xcel_itron_5_instantaneous_demand_val...	756	2024-11-30T00:07:20.385758+00:00	2024-11-30
67	sensor.xcel_itron_5_instantaneous_demand_val...	754	2024-11-30T00:07:26.113857+00:00	2024-11-30
68	sensor.xcel_itron_5_instantaneous_demand_val...	757	2024-11-30T00:07:37.588677+00:00	2024-11-30
69	sensor.xcel_itron_5_instantaneous_demand_val...	753	2024-11-30T00:07:43.426128+00:00	2024-11-30
70	sensor.xcel_itron_5_instantaneous_demand_val...	755	2024-11-30T00:07:49.163928+00:00	2024-11-30
71	sensor.xcel_itron_5_instantaneous_demand_val...	756	2024-11-30T00:07:55.023884+00:00	2024-11-30
72	sensor.xcel_itron_5_instantaneous_demand_val...	1146	2024-11-30T00:08:00.835993+00:00	2024-11-30
73	sensor.xcel_itron_5_instantaneous_demand_val...	1145	2024-11-30T00:08:06.579026+00:00	2024-11-30
74	sensor.xcel_itron_5_instantaneous_demand_val...	1144	2024-11-30T00:08:12.355705+00:00	2024-11-30
75	sensor.xcel_itron_5_instantaneous_demand_val...	1021	2024-11-30T00:08:17.935158+00:00	2024-11-30
76	sensor.xcel_itron_5_instantaneous_demand_val...	1017	2024-11-30T00:08:23.623596+00:00	2024-11-30
77	sensor.xcel_itron_5_instantaneous_demand_val...	1015	2024-11-30T00:08:29.227205+00:00	2024-11-30

10,000+ rows | Truncated data | 1.11 seconds runtime

- Saving transformed data in a single coalesced “parquet” format
- This allows for easier analytics using PowerBI

```
directory_path = "transformed/itronmeter"

output_path = f"{uri}{directory_path}/"

# Save the aggregated data back to the desired path (overwrite original file)
aggregated_data.write.format("parquet").mode("overwrite").save(output_path)
```

# DATA ENGINEERING LIFECYCLE



## UNDERCURRENTS

Security

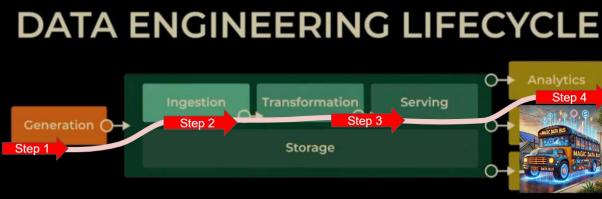
Data Management

DataOps

Data Architecture

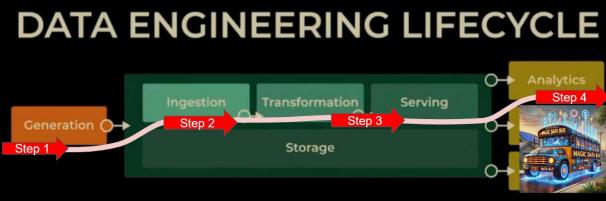
Orchestration

Software Engineering

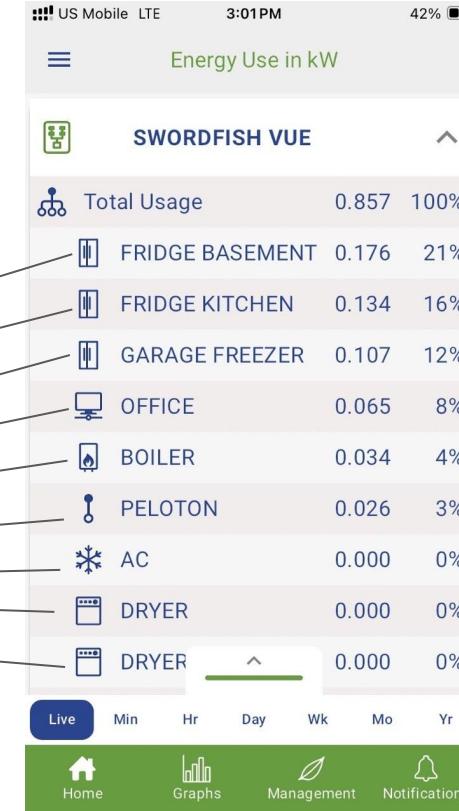
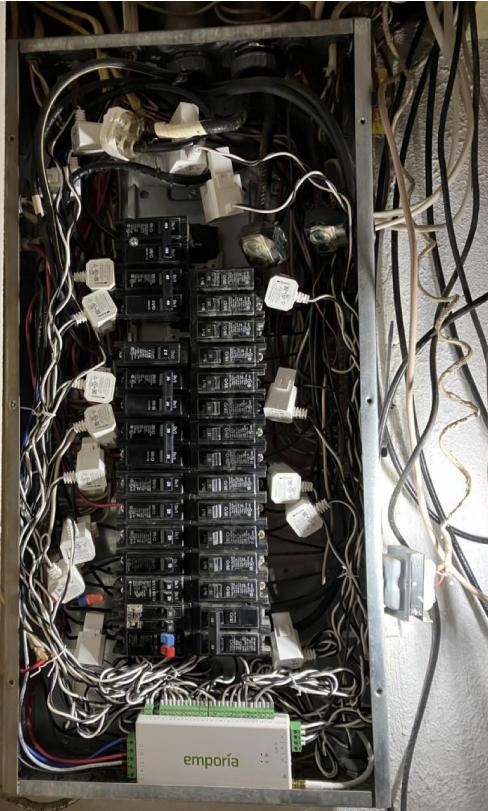


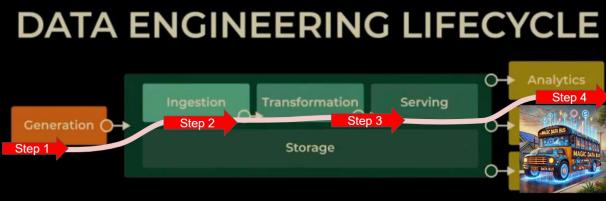
# Step 4: **DATA ANALYTICS - Home Assistant**



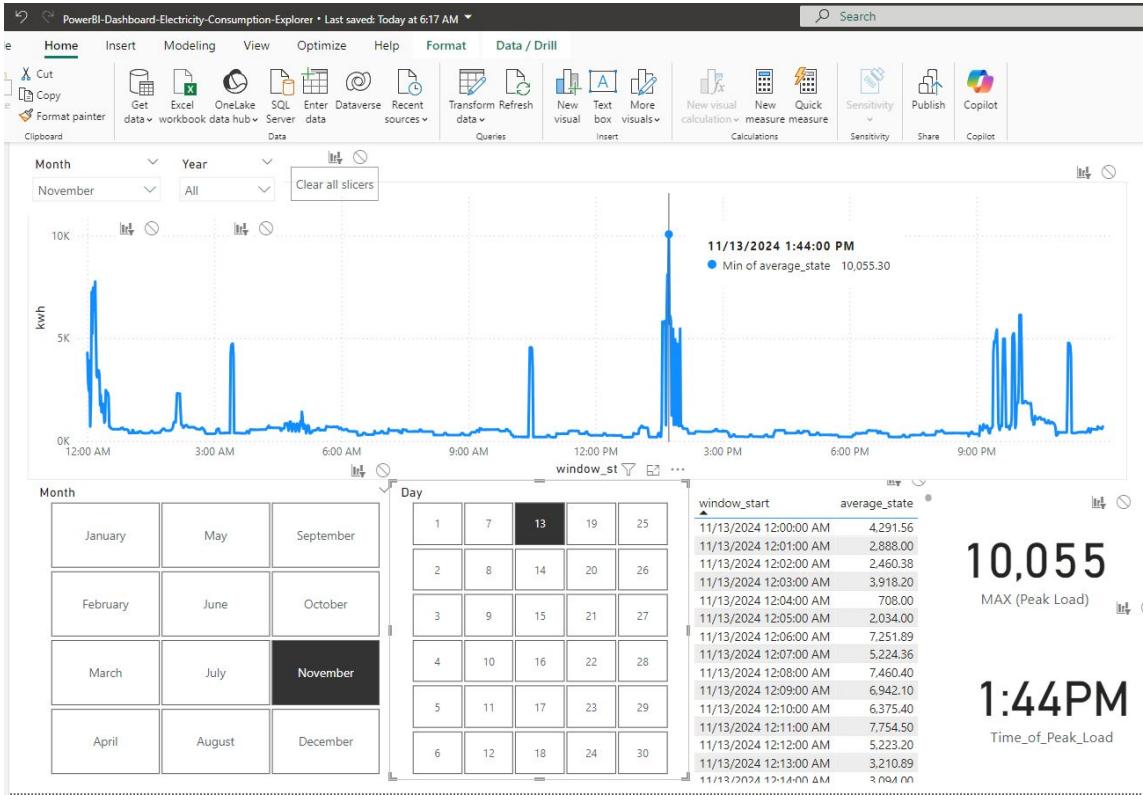


## Step 4: DATA ANALYTICS - Emporia Mobile App





## Step 4: DATA ANALYTICS - PowerBI Desktop



- All data is synced to PowerBI - great visuals!
- Easy to refresh the dashboard
- Everything accessed through a single coalesced “parquet” file saved in Azure Blob

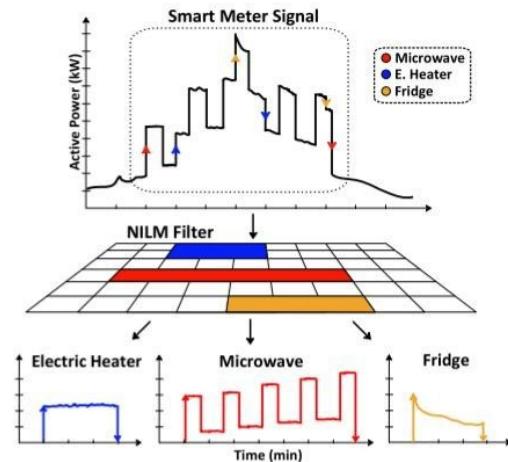
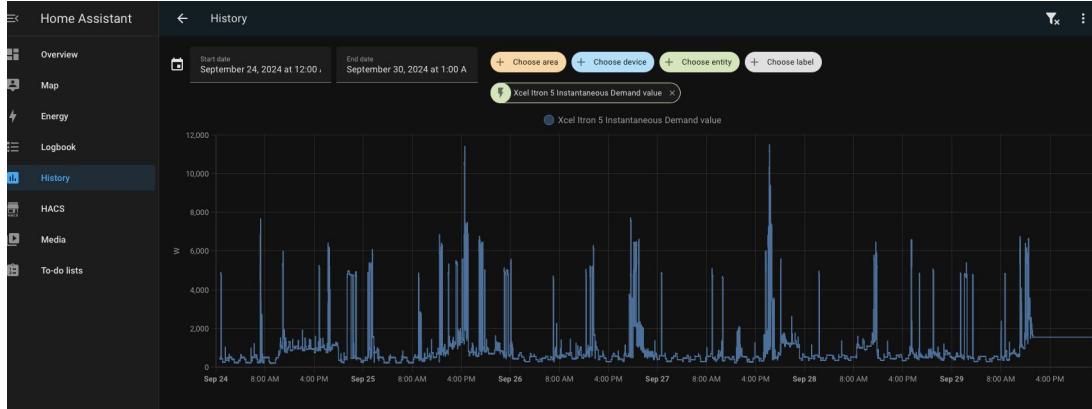
# Other “Data-Centric” Goals

Expand on the data to...

## Create a dataset suitable for Machine Learning

Learn what it takes to engineer a system that's appropriate for Machine Learning to detect patterns in my data. Ideally, want the ability to perform “**non-intrusive load monitoring**”.

[https://en.wikipedia.org/wiki/Nonintrusive\\_load\\_monitoring](https://en.wikipedia.org/wiki/Nonintrusive_load_monitoring)



# Undercurrent considerations

## UNDERCURRENTS

Security

Data Management

DataOps

Data Architecture

Orchestration

Software Engineering

### Security

- Setup isolated home network to keep all communication on own network (Mtr-Rdr2G).
- Principle of least privilege - only I have access

### Data Management

- Capturing key metadata events.
  - Manually annotate data when power outages occur  
(no easy workaround)
- Standard units of measurement
  - Define all data in either:
    - 1) Watts
    - 2) kWh (kilowatt hours)

### DataOps

- What can be automated?
  - Scripts are manual, need to have process automated in Home Assistant
- Any need for CI/CD?
  - Not yet
- How are repos tied into this?
  - Continue using Git as private repo, share with selective access

### Data Architecture

- How does this solution scale? If interval readings are happening every 15 seconds, what's the benefit of capturing every 5 seconds?
  - Capture ALL data at finest level of granularity, down to the second

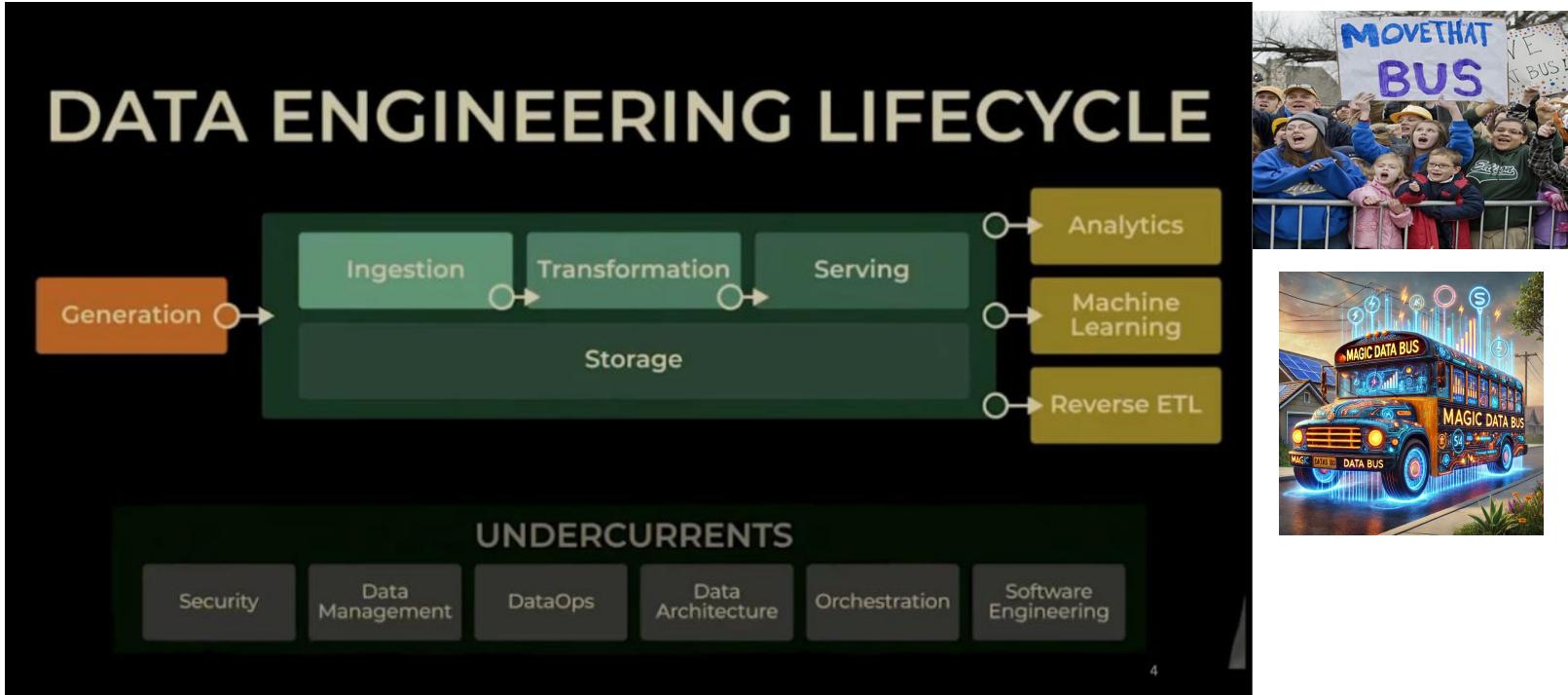
### Orchestration

- Any daily/weekly/quarterly “batch” processes for analyzing the data?
  - Script files can cover daily/weekly batches. Wouldn’t do quarterly

### Software engineering

- What code will be used? Any plans to implement code as infrastructure?
  - Mix of basic powershell scripts, python, and home assistant YAML files
- How is data quality quality assurance performed?
  - Manually inspected, home assistant provide anomaly detection

# Conclusion



- This project is a kickoff to a larger effort for an “energy home makeover” going into 2025
- Main focus of this semester was to solve for data ingestion, and confident that is working well!
- Next year plan to use this data more extensively...incorporate “time of use” billing, peak load factors, demand response, etc.