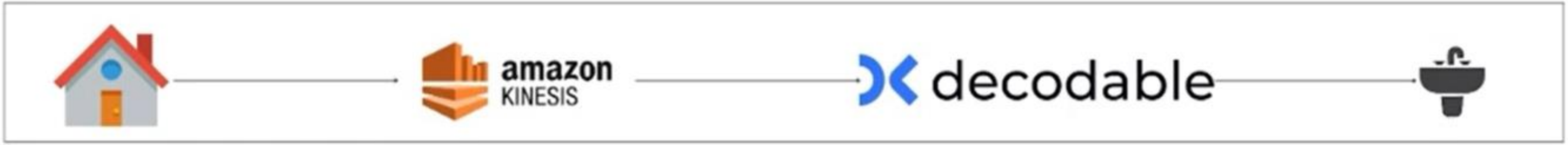
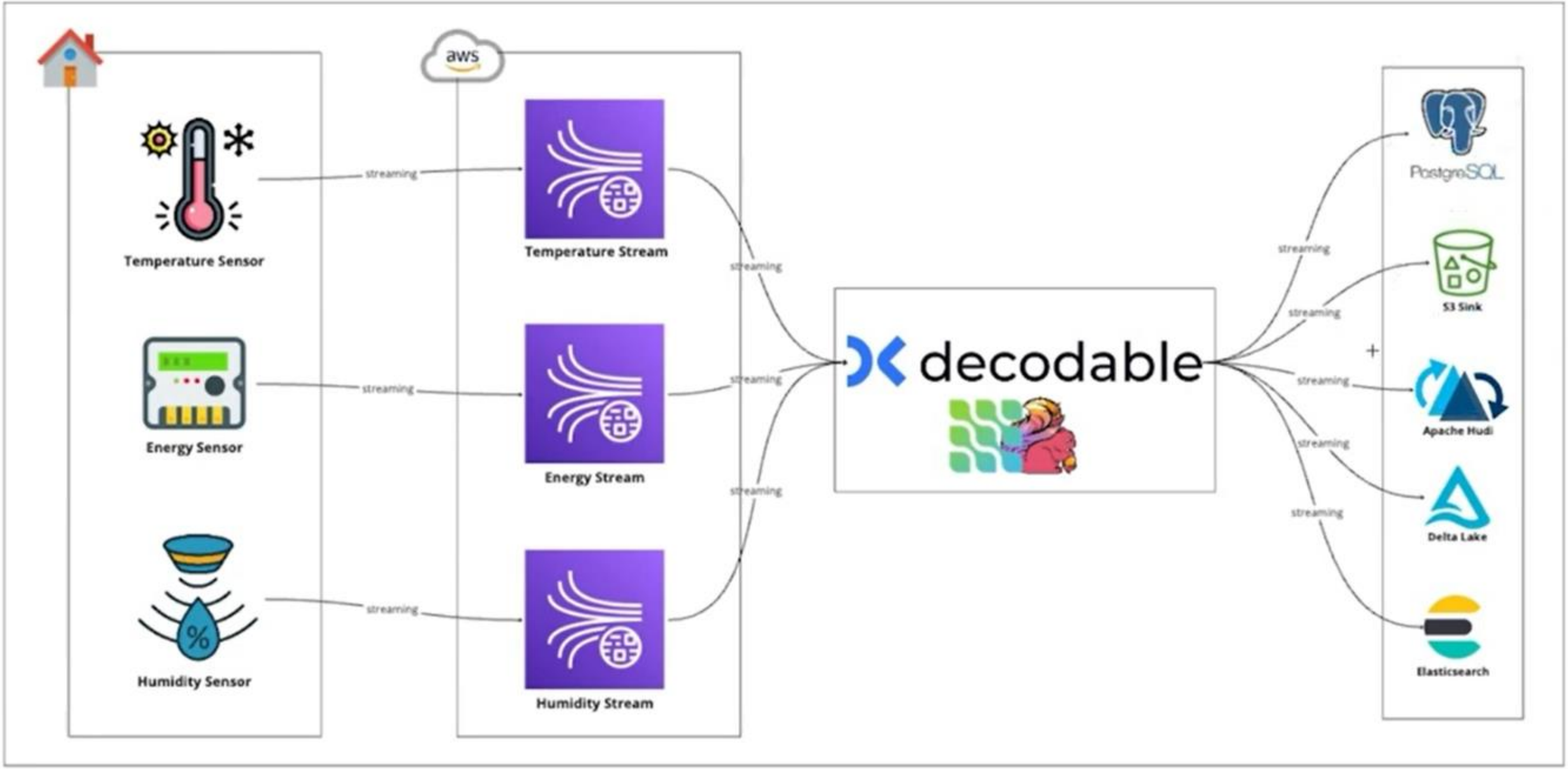
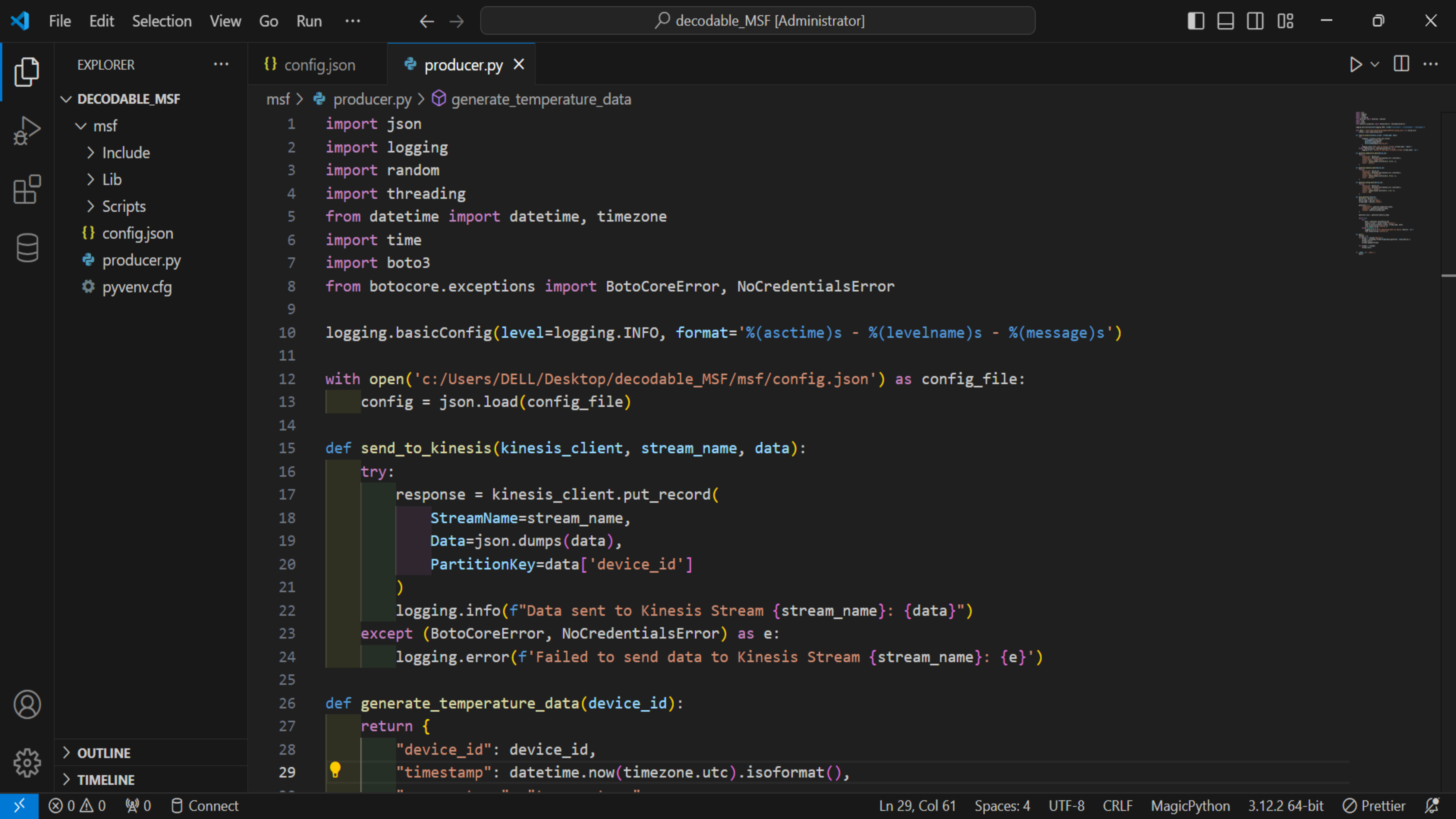
Real-Time Sensors Data Pipeline


amazon
KINESIS
decodable

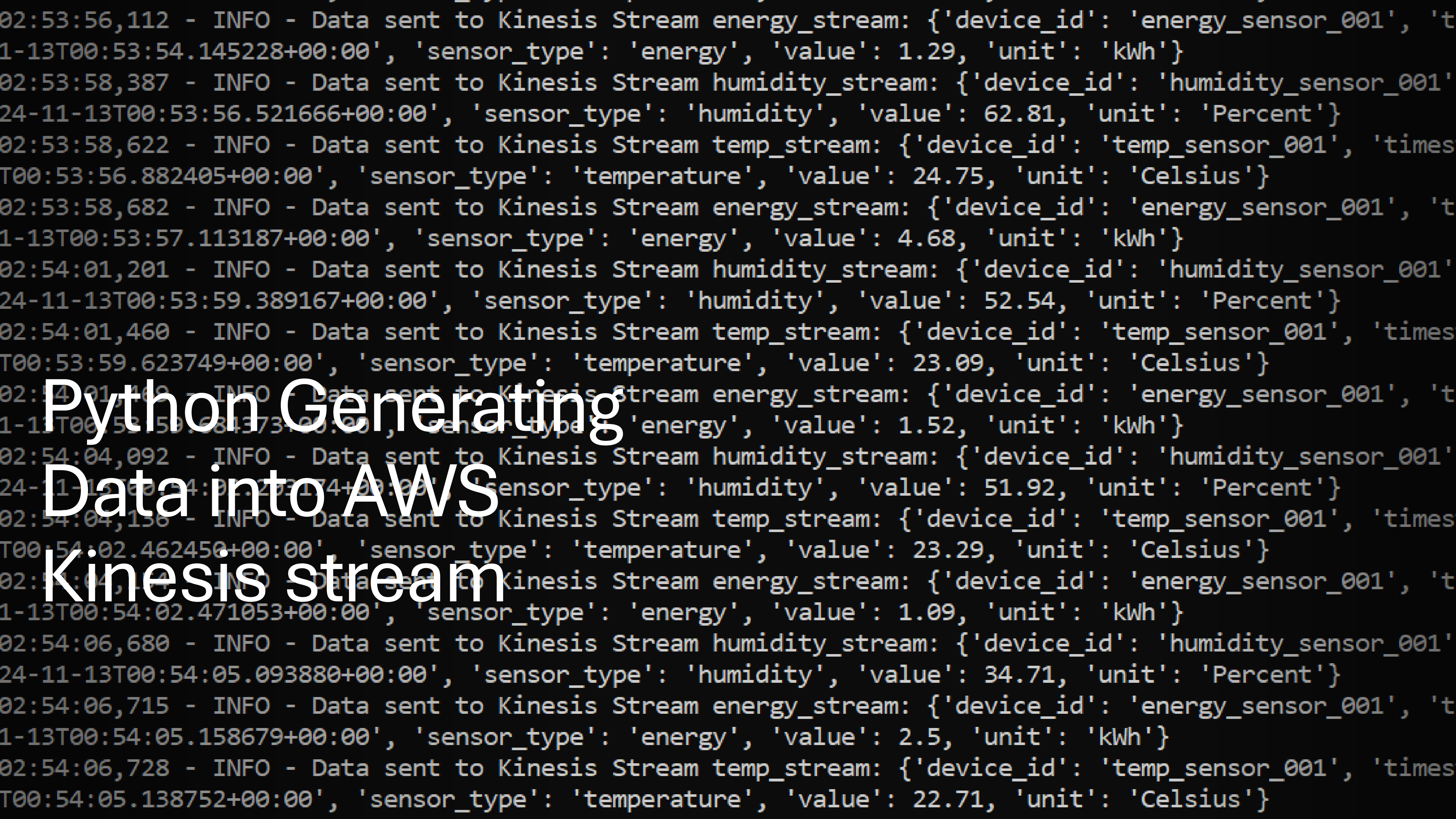

aws
Temperature Sensor
streaming
Temperature Stream
Energy Sensor
streaming
Energy Stream
Humidity Sensor
streaming
Humidity Stream
streaming
streaming
streaming
decodable
streaming
streaming
streaming
streaming
streaming
PostgreSQL
S3 Sink
+
Apache Hudi
Delta Lake
Elasticsearch

python™



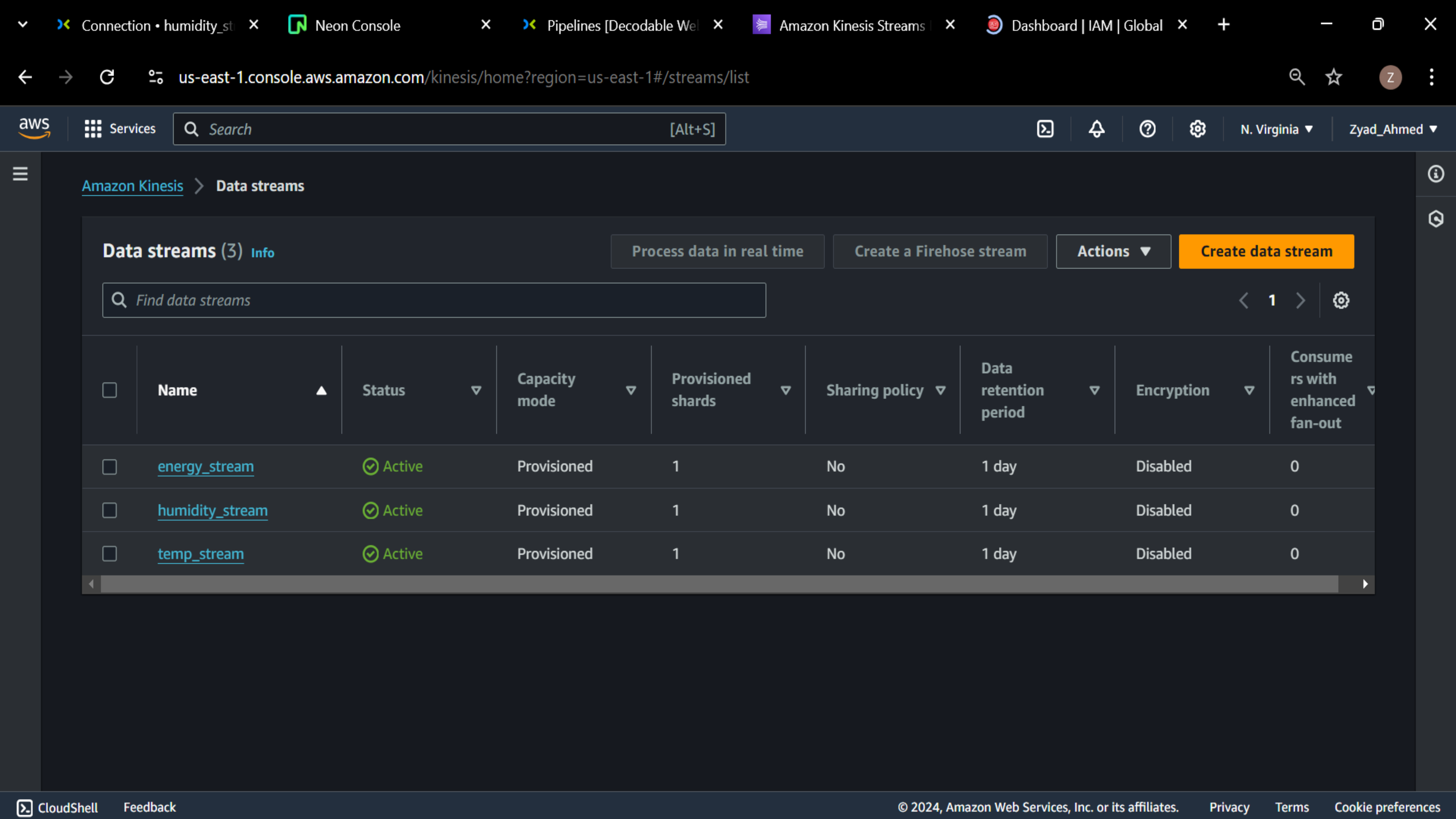
```python
import json
import logging
import random
import threading
from datetime import datetime, timezone
import time
import boto3
from botocore.exceptions import BotoCoreError, NoCredentialsError

logging.basicConfig(level=logging.INFO, format='%(asctime)s - %(levelname)s - %(message)s')

with open('c:/Users/DELL/Desktop/decodable_MSF/msf/config.json') as config_file:
    config = json.load(config_file)

def send_to_kinesis(kinesis_client, stream_name, data):
    try:
        response = kinesis_client.put_record(
            StreamName=stream_name,
            Data=json.dumps(data),
            PartitionKey=data['device_id']
        )
        logging.info(f"Data sent to Kinesis Stream {stream_name}: {data}")
    except (BotoCoreError, NoCredentialsError) as e:
        logging.error(f'Failed to send data to Kinesis Stream {stream_name}: {e}')

def generate_temperature_data(device_id):
    return {
        "device_id": device_id,
        "timestamp": datetime.now(timezone.utc).isoformat(),
```

Python Generating Data into AWS Kinesis stream



```
02:53:56,112 - INFO - Data sent to Kinesis Stream energy_stream: {'device_id': 'energy_sensor_001', 't
1-13T00:53:54.145228+00:00', 'sensor_type': 'energy', 'value': 1.29, 'unit': 'kWh'}
02:53:58,387 - INFO - Data sent to Kinesis Stream humidity_stream: {'device_id': 'humidity_sensor_001'
24-11-13T00:53:56.521666+00:00', 'sensor_type': 'humidity', 'value': 62.81, 'unit': 'Percent'}
02:53:58,622 - INFO - Data sent to Kinesis Stream temp_stream: {'device_id': 'temp_sensor_001', 'times
T00:53:56.882405+00:00', 'sensor_type': 'temperature', 'value': 24.75, 'unit': 'Celsius'}
02:53:58,682 - INFO - Data sent to Kinesis Stream energy_stream: {'device_id': 'energy_sensor_001', 't
1-13T00:53:57.113187+00:00', 'sensor_type': 'energy', 'value': 4.68, 'unit': 'kWh'}
02:54:01,201 - INFO - Data sent to Kinesis Stream humidity_stream: {'device_id': 'humidity_sensor_001'
24-11-13T00:53:59.389167+00:00', 'sensor_type': 'humidity', 'value': 52.54, 'unit': 'Percent'}
02:54:01,460 - INFO - Data sent to Kinesis Stream temp_stream: {'device_id': 'temp_sensor_001', 'times
T00:53:59.623749+00:00', 'sensor_type': 'temperature', 'value': 23.09, 'unit': 'Celsius'}
02:54:01,[illegible] - INFO - Data sent to Kinesis Stream energy_stream: {'device_id': 'energy_sensor_001', 't
1-13T00:53:59.[illegible]373+00:00', 'sensor_type': 'energy', 'value': 1.52, 'unit': 'kWh'}
02:54:04,092 - INFO - Data sent to Kinesis Stream humidity_stream: {'device_id': 'humidity_sensor_001'
24-11-13T00:54:0[illegible]74+00:00', 'sensor_type': 'humidity', 'value': 51.92, 'unit': 'Percent'}
02:54:04,136 - INFO - Data sent to Kinesis Stream temp_stream: {'device_id': 'temp_sensor_001', 'times
T00:54:02.462450+00:00', 'sensor_type': 'temperature', 'value': 23.29, 'unit': 'Celsius'}
02:5[illegible] - INFO - Data sent to Kinesis Stream energy_stream: {'device_id': 'energy_sensor_001', 't
1-13T00:54:02.471053+00:00', 'sensor_type': 'energy', 'value': 1.09, 'unit': 'kWh'}
02:54:06,680 - INFO - Data sent to Kinesis Stream humidity_stream: {'device_id': 'humidity_sensor_001'
24-11-13T00:54:05.093880+00:00', 'sensor_type': 'humidity', 'value': 34.71, 'unit': 'Percent'}
02:54:06,715 - INFO - Data sent to Kinesis Stream energy_stream: {'device_id': 'energy_sensor_001', 't
1-13T00:54:05.158679+00:00', 'sensor_type': 'energy', 'value': 2.5, 'unit': 'kWh'}
02:54:06,728 - INFO - Data sent to Kinesis Stream temp_stream: {'device_id': 'temp_sensor_001', 'times
T00:54:05.138752+00:00', 'sensor_type': 'temperature', 'value': 22.71, 'unit': 'Celsius'}
```

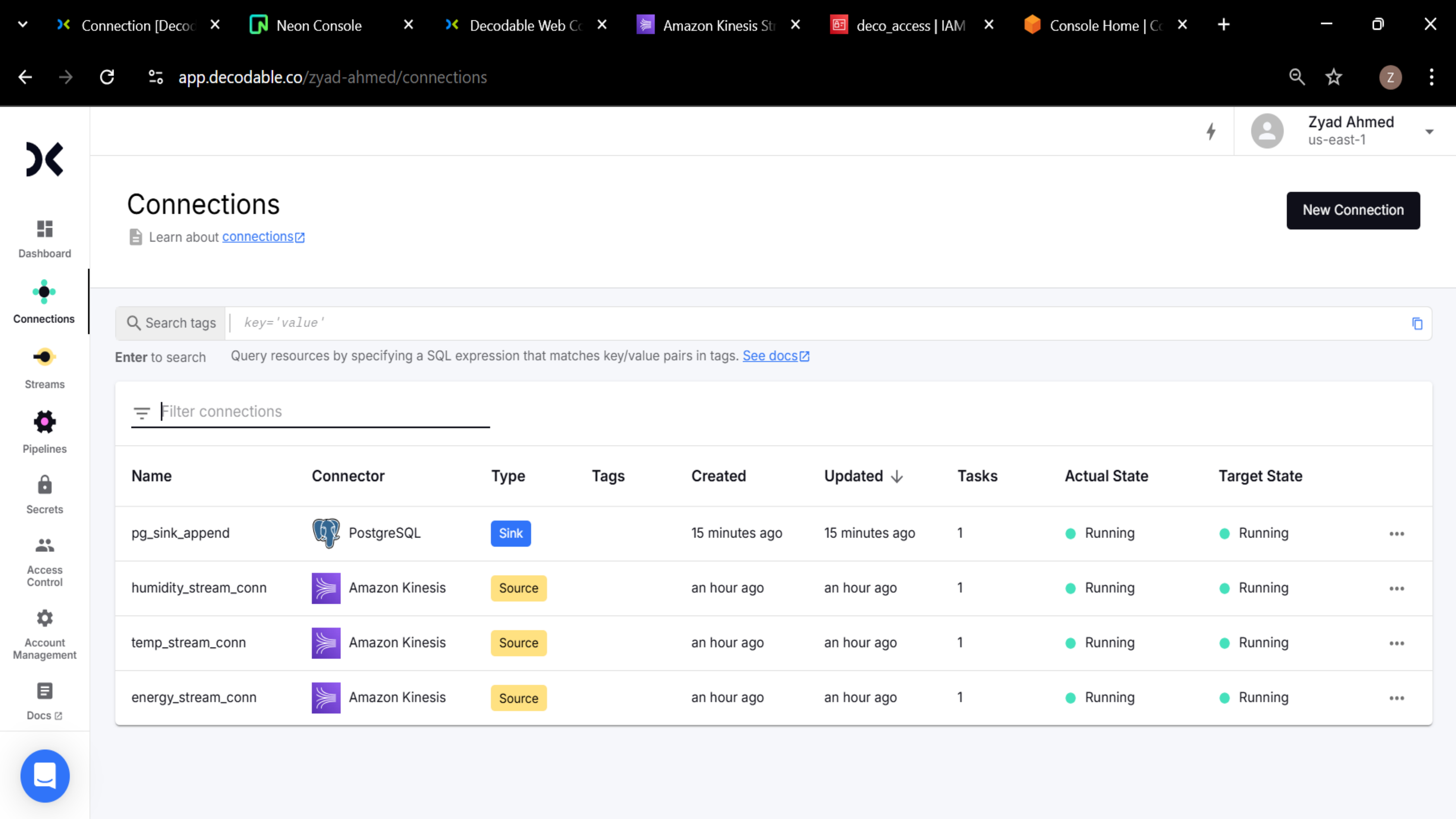
amazon
KINESIS



Data streams (3)

□	Name	Status	Capacity mode	Provisioned shards	Sharing policy	Data retention period	Encryption	Consumers with enhanced fan-out
□	energy_stream	✓ Active	Provisioned	1	No	1 day	Disabled	0
□	humidity_stream	✓ Active	Provisioned	1	No	1 day	Disabled	0
□	temp_stream	✓ Active	Provisioned	1	No	1 day	Disabled	0

decodable

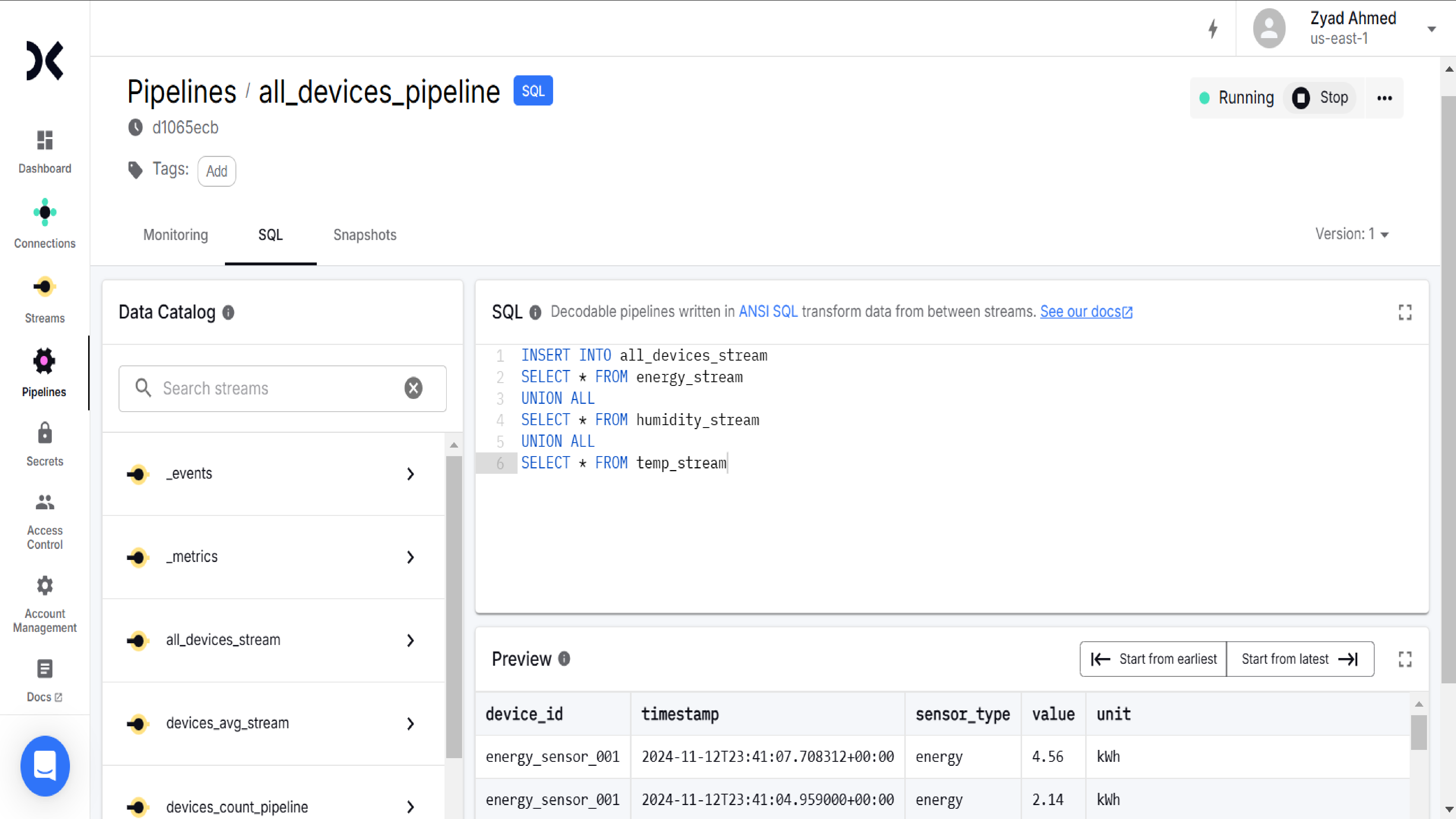


Connections

Learn about connections

Enter to search   Query resources by specifying a SQL expression that matches key/value pairs in tags. See docs

Name	Connector	Type	Tags	Created	Updated ↓	Tasks	Actual State	Target State
pg_sink_append	PostgreSQL	Sink		15 minutes ago	15 minutes ago	1	● Running	● Running
humidity_stream_conn	Amazon Kinesis	Source		an hour ago	an hour ago	1	● Running	● Running
temp_stream_conn	Amazon Kinesis	Source		an hour ago	an hour ago	1	● Running	● Running
energy_stream_conn	Amazon Kinesis	Source		an hour ago	an hour ago	1	● Running	● Running



Pipelines / all_devices_pipeline SQL

d1065ecb

Tags: Add

Data Catalog

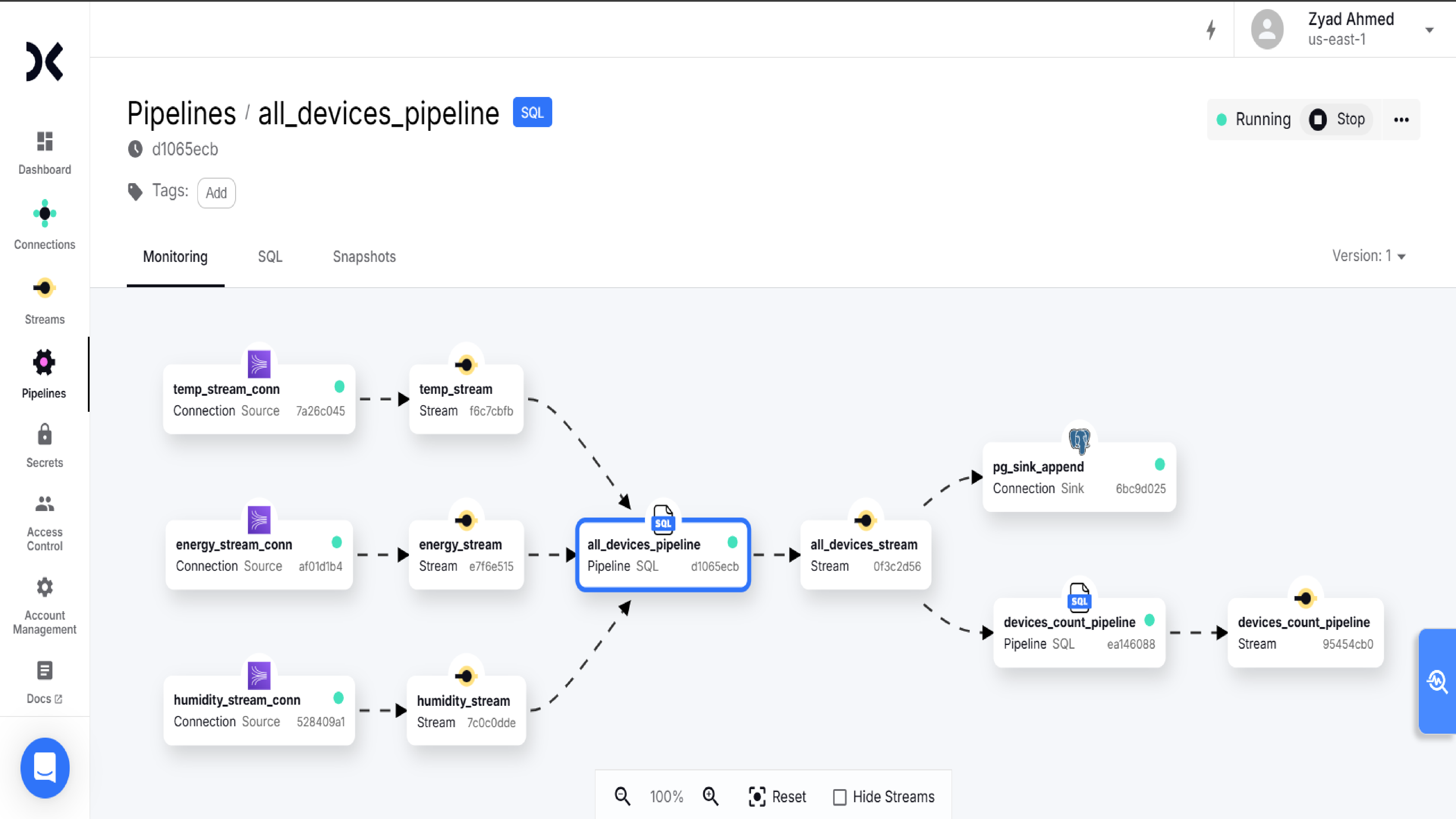
- _events
- _metrics
- all_devices_stream
- devices_avg_stream
- devices_count_pipeline

SQL Decodable pipelines written in ANSI SQL transform data from between streams. See our docs

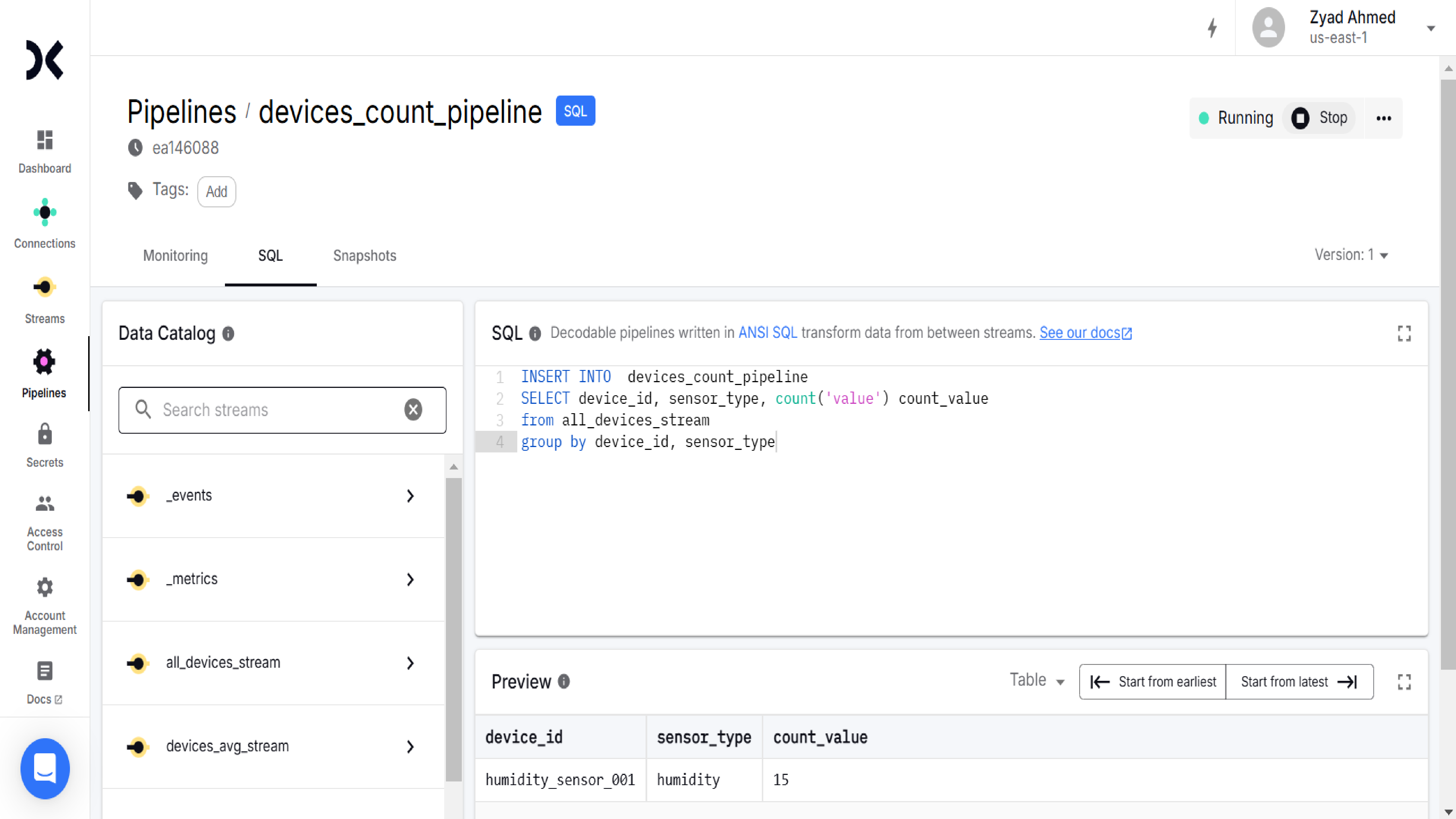
```
INSERT INTO all_devices_stream
SELECT * FROM energy_stream
UNION ALL
SELECT * FROM humidity_stream
UNION ALL
SELECT * FROM temp_stream
```

Preview

device_id	timestamp	sensor_type	value	unit
energy_sensor_001	2024-11-12T23:41:07.708312+00:00	energy	4.56	kWh
energy_sensor_001	2024-11-12T23:41:04.959000+00:00	energy	2.14	kWh



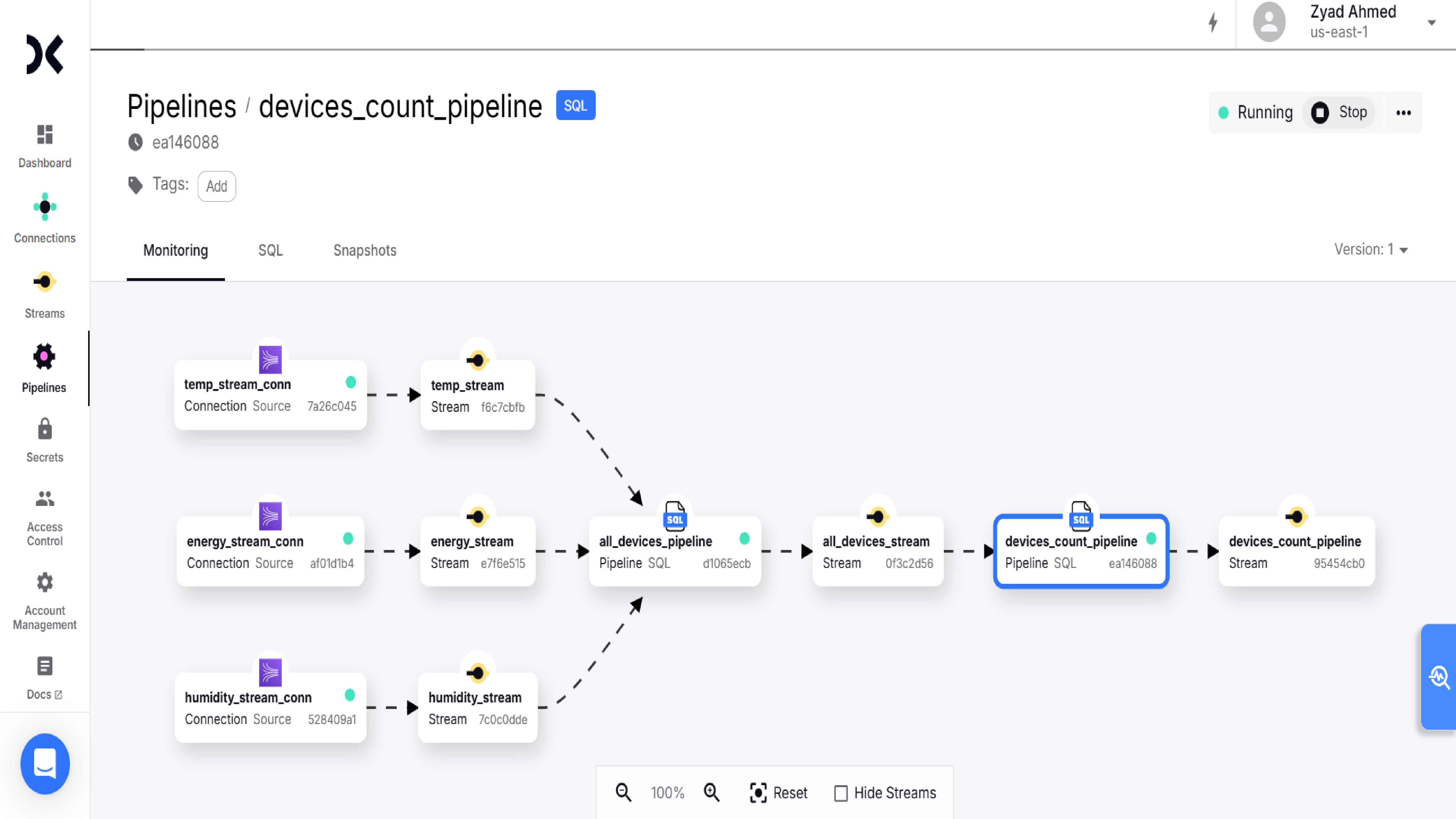
Zyad Ahmed
us-east-1
Dashboard
Connections
Streams
Pipelines
Secrets
Access Control
Account Management
Docs
Pipelines / all_devices_pipeline SQL
d1065ecb
Tags: Add
Running
Stop
Monitoring
SQL
Snapshots
Version: 1
temp_stream_conn
Connection Source 7a26c045
temp_stream
Stream f6c7cbfb
energy_stream_conn
Connection Source af01d1b4
energy_stream
Stream e7f6e515
humidity_stream_conn
Connection Source 528409a1
humidity_stream
Stream 7c0c0dde
SQL
all_devices_pipeline
Pipeline SQL d1065ecb
all_devices_stream
Stream 0f3c2d56
pg_sink_append
Connection Sink 6bc9d025
SQL
devices_count_pipeline
Pipeline SQL ea146088
devices_count_pipeline
Stream 95454cb0
100%
Reset
Hide Streams



SQL Decodable pipelines written in ANSI SQL transform data from between streams. See our docs

```
INSERT INTO  devices_count_pipeline
SELECT device_id, sensor_type, count('value') count_value
from all_devices_stream
group by device_id, sensor_type
```

device_id	sensor_type	count_value
humidity_sensor_001	humidity	15



Pipelines / devices_count_pipeline SQL

ea146088

Tags:

temp_stream_conn
Connection Source 7a26c045

temp_stream
Stream f6c7cbfb

energy_stream_conn
Connection Source af01d1b4

energy_stream
Stream e7f6e515

humidity_stream_conn
Connection Source 528409a1

humidity_stream
Stream 7c0c0dde

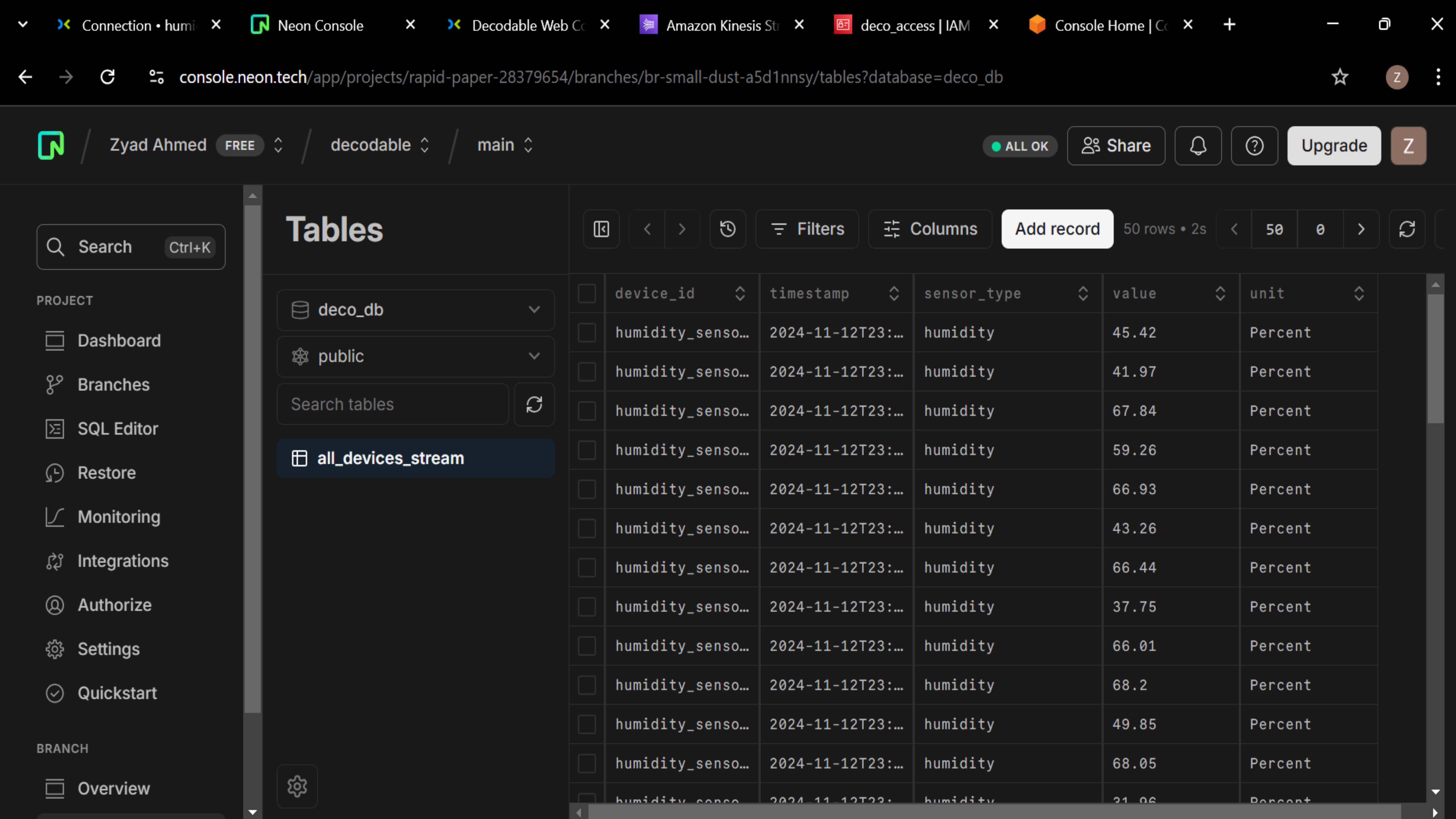
all_devices_pipeline
Pipeline SQL d1065ecb

all_devices_stream
Stream 0f3c2d56

devices_count_pipeline
Pipeline SQL ea146088

devices_count_pipeline
Stream 95454cb0

PostgreSQL



deco_db

public

Search tables

all_devices_stream

device_id	timestamp	sensor_type	value	unit
humidity_senso…	2024-11-12T23:…	humidity	45.42	Percent
humidity_senso…	2024-11-12T23:…	humidity	41.97	Percent
humidity_senso…	2024-11-12T23:…	humidity	67.84	Percent
humidity_senso…	2024-11-12T23:…	humidity	59.26	Percent
humidity_senso…	2024-11-12T23:…	humidity	66.93	Percent
humidity_senso…	2024-11-12T23:…	humidity	43.26	Percent
humidity_senso…	2024-11-12T23:…	humidity	66.44	Percent
humidity_senso…	2024-11-12T23:…	humidity	37.75	Percent
humidity_senso…	2024-11-12T23:…	humidity	66.01	Percent
humidity_senso…	2024-11-12T23:…	humidity	68.2	Percent
humidity_senso…	2024-11-12T23:…	humidity	49.85	Percent
humidity_senso…	2024-11-12T23:…	humidity	68.05	Percent
humidity_senso…	2024-11-12T23:…	humidity	31.96	Percent