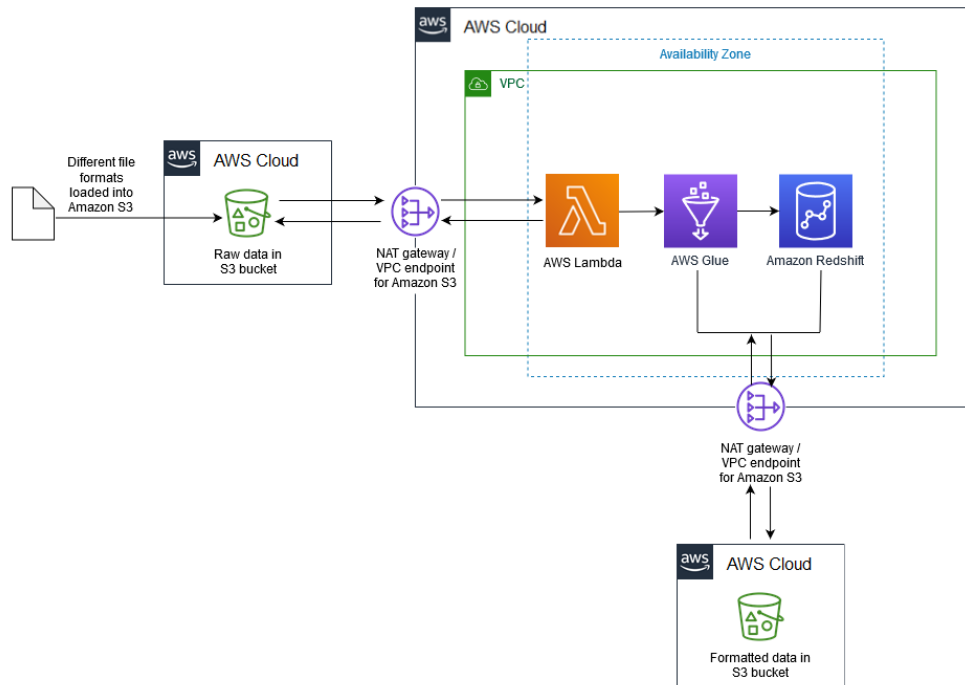


My Workflow:



Creating a S3 bucket and uploading data files.

mydbsbucket [Info](#)

[Objects](#) [Properties](#) [Permissions](#) [Metrics](#) [Management](#) [Access Points](#)

Objects (3)

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 Inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

[Refresh](#) [Copy S3 URI](#) [Copy URL](#) [Download](#) [Open](#) [Delete](#) [Actions](#) [Create folder](#) [Upload](#)

[Show versions](#)

| <input type="checkbox"/> | Name | Type | Last modified | Size | Storage class |
|--------------------------|---------------------------------------|------|---|---------|---------------|
| <input type="checkbox"/> | Address.csv | csv | November 28, 2022, 11:17:10 (UTC-08:00) | 30.7 KB | Standard |
| <input type="checkbox"/> | building_area.csv | csv | November 27, 2022, 18:18:49 (UTC-08:00) | 10.2 KB | Standard |
| <input type="checkbox"/> | electricity_usage.csv | csv | November 27, 2022, 18:18:50 (UTC-08:00) | 18.4 KB | Standard |

Creating VPC endpoints for the ETL process of the data present in S3 bucket

Endpoints (1) [Info](#)

[Refresh](#) [Actions](#) [Create endpoint](#)

| <input type="checkbox"/> | Name | VPC endpoint ID | VPC ID | Service name | Endpoint type | Status |
|--------------------------|-------------|-----------------------|-----------------------|----------------------------|---------------|------------------------|
| <input type="checkbox"/> | my-endpoint | vpc-e0c1c8f6d6dd0f746 | vpc-056bd966d01270444 | com.amazonaws.us-west-2.s3 | Gateway | Available |

Creating crawler to import data from S3 bucket.

| <input type="checkbox"/> | Name | State | Schedule | Last run | Log | Table changes from last run |
|--------------------------|----------|--------------------|----------|------------------------|--------------------------|-----------------------------|
| <input type="checkbox"/> | crawler1 | Ready | | Succeeded | View log | 3 created |

Imported tables from the s3 bucket.

| <input type="checkbox"/> | Name | Database | Location | Classification | Deprecated | View data |
|--------------------------|-----------------------|------------|--------------------------------------|----------------|------------|----------------------------|
| <input type="checkbox"/> | address_csv | another_db | s3://mydbsbucket/Address.csv | csv | - | Table data |
| <input type="checkbox"/> | building_area_csv | another_db | s3://mydbsbucket/building_area.csv | csv | - | Table data |
| <input type="checkbox"/> | electricity_usage_csv | another_db | s3://mydbsbucket/electricity_usage.c | csv | - | Table data |

Creating jobs for each of the tables in the AWS glue database where the data source is the S3 bucket and data target is the redshift database using the Amazon redshift connection

| <input type="checkbox"/> | Name | Type | ETL language | Script location | Last modified | Job bookmark |
|--------------------------|--------|-------|--------------|--------------------------|---------------------------------|--------------|
| <input type="checkbox"/> | job2 | Spark | python | s3://aws-glue-scripts... | 28 November 2022 10:32 AM UT... | Disable |
| <input type="checkbox"/> | job3 | Spark | python | s3://aws-glue-scripts... | 28 November 2022 11:22 AM UT... | Disable |
| <input type="checkbox"/> | job4 | Spark | python | s3://aws-glue-scripts... | 28 November 2022 11:25 AM UT... | Disable |
| <input type="checkbox"/> | table1 | Spark | python | s3://aws-glue-scripts... | 28 November 2022 10:31 AM UT... | Disable |

| Output Schema Definition | | | | | | | | | |
|---|-----------|--------------------------------------|---|----------------------|-----------|---|---|---|--|
| Verify the mappings created by AWS Glue. Change mappings by choosing other columns with Map to target . You can Clear all mappings and Reset to default AWS Glue mappings. AWS Glue generates your script with the defined mappings. | | | | | | | | | |
| Source | | | | Target | | | | | |
| Column name | Data type | Map to target | | Column name | Data type | | | | |
| site id | bigint | site id | → | site id | long | × | ↓ | ↑ | |
| department | string | department | → | department | string | × | ↓ | ↑ | |
| current solar | bigint | current solar | → | current solar | long | × | ↓ | ↑ | |
| electric utility | string | electric utility | → | electric utility | string | × | ↓ | ↑ | |
| electricity usage | bigint | electricity usage | → | electricity usage | long | × | ↓ | ↑ | |
| peak electric demand | bigint | peak electric demand | → | peak electric demand | long | × | ↓ | ↑ | |
| natural gas usage | bigint | natural gas usage | → | natural gas usage | long | × | ↓ | ↑ | |
| energy use intensity | double | energy use intensity | → | energy use intensity | double | × | ↓ | ↑ | |

Creating Redshift cluster for querying

| <input type="checkbox"/> | Cluster | Status | Cluster namespace | Storage capacity us... | CPU utilization | Snapshots | Notificati... | Tags |
|--------------------------|---|-----------|-----------------------|------------------------|-----------------|-----------|---------------|------|
| <input type="checkbox"/> | redshift-cluster-1 dc2.large 1 node 160 GB | Available | 0801646e-0058-4373... | < 1% | 2% | - | | |

| Resources | Info | Refresh | Close |
|--|------|---------|-------|
| Select database Info To view schemas, select a database. | | | |
| dev | | | |
| Select schema Info To view tables, select a schema. | | | |
| public | | | |
| <input type="text" value="Filter tables"/> | | | |
| < 1 > | | | |
| ▶ address_csv ... | | | |
| ▶ building_area_csv ... | | | |
| ▶ elctricity_usage_csv ... | | | |

Considering 0.35\$ per kWh. Find top 10 building types with highest cost.

- ```
select building."building type", avg(elec_usage."electricity usage") as
electricity_usage, 0.35* electricity_usage as total_cost
from public.building_area_csv as building
right join
 public.elctricity_usage_csv as elec_usage
on building."site id"=elec_usage."site id"
group by building."building type"
having total_cost > 500
order by total_cost desc limit 10;
order by avg_cost;
```

| buildingtype     | electricity_usage | total_cost |
|------------------|-------------------|------------|
| CONVENTION_CN    | 8598781           | 3009573.35 |
| TRANSPORTATIONTE | 5355298           | 1874354.30 |
| LABORATORY       | 3817577           | 1336151.95 |
| MRF              | 3410495           | 1193673.25 |
| COURTHOUSE       | 3059932           | 1070976.20 |
| PARKING          | 881190            | 308416.50  |
| DISTRIBUTION     | 805210            | 281823.50  |
| PERFORMINGARTS   | 780194            | 273067.90  |
| VOCATIONALSCHOOL | 729252            | 255238.20  |
| OFFICE           | 715092            | 250282.20  |

### Top 10 departments that use natural gas under PG&E electric utility.

- ```
select elec_usage."department", avg(elec_usage."natural gas usage") as
natural_gas_usage from public.elctricity_usage_csv as elec_usage
where elec_usage."electric utility"='PG&E'
group by elec_usage."department"
order by natural_gas_usage desc limit 10;
```

| department | natural_gas_usage |
|----------------|-------------------|
| Convention | 37683 |
| Arts | 28383 |
| Library | 10378 |
| Police | 9375 |
| Public Works | 3898 |
| City Clerk | 856 |
| Streets | 787 |
| Fire | 670 |
| Human Services | 665 |
| Parks | 573 |

Top 10 departments that use natural gas under SJCE electric utility select
 elec_usage."department", avg(elec_usage."natural gas usage") as natural_gas_usage
 from public.elctricity_usage_csv as elec_usage
 where elec_usage."electric utility"='SJCE'
 group by elec_usage."department"
 order by natural_gas_usage desc limit 10;

| department | natural_gas_usage |
|----------------|-------------------|
| Housing | 5280 |
| Neighborhood | 2974 |
| Police | 1537 |
| Fire | 935 |
| Public Works | 377 |
| Human Services | 258 |
| Library | 180 |
| Parks | 172 |
| Streets | 0 |
| Public Transit | 0 |

Sites with highest electricity usage under SJCE utility.

```
◦ select elec_usage."department", avg(elec_usage."natural gas usage") as  
natural_gas_usage from public.elctricity_usage_csv as elec_usage  
where elec_usage."electric utility"='SJCE'  
group by elec_usage."department"  
order by natural_gas_usage desc limit 10;
```

| sitename | buildingarea | electricityusage | cost |
|-----------------------------------|--------------|------------------|-----------|
| Pecos Park - Community Center | 93670 | 1595744 | 558510.40 |
| Steele Indian School Restroom | 67470 | 401605 | 140561.75 |
| Deer Valley Pool | 60000 | 121351 | 42472.85 |
| Goelet A Beuf Community Center | 48000 | 728857 | 255099.95 |
| Encanto Maintenance and Warehouse | 42090 | 250259 | 87590.65 |
| Paradise Valley Community Cntr | 35506 | 332983 | 116544.05 |
| Deer Valley Community Center | 34294 | 375779 | 131522.65 |
| Washington Activity Center | 32971 | 346891 | 121411.85 |
| Papago Sports Complex | 31426 | 193986 | 67895.10 |
| Maryvale Community Center | 27000 | 760600 | 266210.00 |

Department and respective building type that highly powered by solar enegry.

```
◦ select electricity."department", building."building  
type",sum(electricity."current solar") as total_current_solar_usage,  
count(electricity."current solar") as count, avg(electricity."current solar") as  
avg_current_solar_usage from public.elctricity_usage_csv as electricity  
inner join  
public.building_area_csv as building  
on building."site id"=electricity."site id"  
group by electricity."department", building."building type"  
order by total_current_solar_usage desc  
limit 10;
```

| department ▾ | buildingtype ▾ | total_current_solar_usage ▾ | count ▾ | avg_current_solar_usage |
|----------------|------------------|-----------------------------|---------|-------------------------|
| Aviation | PARKING | 5400 | 8 | 675 |
| Public Works | PARKING | 1326 | 2 | 663 |
| Aviation | OFFICE | 580 | 7 | 82 |
| Public Works | REPAIRSERVICES | 540 | 4 | 135 |
| Housing | MULTIHOUSING | 461 | 19 | 24 |
| Public Transit | TRANSPORTATIONTE | 207 | 14 | 14 |
| Library | LIBRARY | 150 | 15 | 10 |
| Parks | COMMUNITYCENTER | 140 | 17 | 8 |
| Public Works | OFFICE | 100 | 8 | 12 |
| Convention | CONVENTION_CN | 100 | 2 | 50 |