

GRAIN:

Our grain is at a daily level, in which each record in the fact table represents air quality measurements for a specific parameter (e.g., ozone, PM2.5) at a particular location for a single day.

DIMENSION:**1. Time Dimension:**

The dimension representing time, inferred from the date_local field in the fact table. It includes attributes such as day of the week, month, and year.

- Date Local (Primary Key)
- Day of the Week
- Month
- Year

2. Location Dimension:

This dimension contains information about geographical locations, including attributes like state code, county code, latitude, longitude, etc.

- Location Key (Surrogate Key)
- State Code
- County Code
- Latitude
- Longitude
- State Name
- County Name
- City Name

3. Parameter Dimension:

Represents air quality parameters, such as ozone, PM2.5, CO, etc.

- Parameter Key (Surrogate Key)
- Parameter Name

4. Air Quality Fact:

- Fact Key (Surrogate Key)
- Location Key (Foreign Key referencing Location Dimension)
- Time Key (Foreign Key referencing Time Dimension)
- Parameter Key (Foreign Key referencing Parameter Dimension)
- AQI
- 1st Max Value
- 1st Max Hour
- Arithmetic Mean

SQL FILE:

USE AQI_Normalized;

-- Time Dimension Table

```
CREATE TABLE Time_Dimension (  
    Time_Key BIGINT PRIMARY KEY,  
    Date_Local TEXT,  
    Day_of_Week TEXT,  
    Month TEXT,  
    Year INT  
);
```

-- Location Dimension Table

```
CREATE TABLE Location_Dimension (  
    Location_Key BIGINT PRIMARY KEY,  
    State_Code BIGINT,  
    County_Code BIGINT,  
    Latitude DOUBLE,  
    Longitude DOUBLE,  
    State_Name TEXT,  
    County_Name TEXT,  
    City_Name TEXT  
);
```

-- Parameter Dimension Table

```
CREATE TABLE Parameter_Dimension (  
    Parameter_Key BIGINT PRIMARY KEY,  
    Parameter_Name TEXT  
);
```

-- Fact Table to include Foreign Key references to existing dimensions

```
CREATE TABLE Air_Quality_Fact (  
    Fact_Key BIGINT PRIMARY KEY,  
    Location_Key BIGINT,  
    Time_Key BIGINT,  
    Parameter_Key BIGINT,  
    AQI DOUBLE,  
    1st_Max_Value DOUBLE,  
    1st_Max_Hour BIGINT,  
    Arithmetic_Mean DOUBLE,  
    FOREIGN KEY (Location_Key) REFERENCES Location_Dimension(Location_Key),  
    FOREIGN KEY (Time_Key) REFERENCES Time_Dimension(Time_Key),  
    FOREIGN KEY (Parameter_Key) REFERENCES Parameter_Dimension(Parameter_Key));
```