```
Joonha Park
10/27/2024
u1325434
IS6420-001
Air Quality Data Research
```

-1 What is the average AQI (air quality index) by year by season (winter, spring, summer, fall)?

```
SELECT
    EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD')) AS Year,
    CASE
        WHEN EXTRACT(MONTH FROM TO_DATE("Date", 'YYYY-MM-DD')) IN (12, 1, 2) THEN
'Winter'
        WHEN EXTRACT(MONTH FROM TO_DATE("Date", 'YYYY-MM-DD')) IN (3, 4, 5) THEN
'Spring'
        WHEN EXTRACT(MONTH FROM TO_DATE("Date", 'YYYY-MM-DD')) IN (6, 7, 8) THEN
'Summer'
        ELSE 'Fall'
    END AS Season,
    AVG(aqi) AS Avg_AQI
FROM
    daily_aqi_by_county dabc
GROUP BY
    EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD')),
    CASE
        WHEN EXTRACT(MONTH FROM TO_DATE("Date", 'YYYY-MM-DD')) IN (12, 1, 2) THEN
'Winter'
        WHEN EXTRACT(MONTH FROM TO DATE("Date", 'YYYY-MM-DD')) IN (3, 4, 5) THEN
'Spring'
        WHEN EXTRACT(MONTH FROM TO_DATE("Date", 'YYYY-MM-DD')) IN (6, 7, 8) THEN
'Summer'
        ELSE 'Fall'
    END
ORDER BY Year, Season;
```



-2. What were the top 10 locations with worst AQI in each year?

```
WITH RankedAQI AS (
    SELECT
          "State Name",
          "county Name",
          EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD')) AS year,
          MAX(AQI) AS max_aqi,
```

```
ROW NUMBER() OVER (PARTITION BY EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-
DD')) ORDER BY MAX(AQI) DESC) AS rank
    FROM
       daily_aqi_by_county
    GROUP BY
        "State Name"
        "county Name"
        EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD'))
SELECT
    "State Name"
    "county Name",
   year,
   max_aqi
FROM
    RankedAQI
WHERE
    rank <= 10
ORDER BY
   year,
   max aqi DESC;
```

•	A-Z State Name	A-Z county Name	123 year 🔻	123 max_aqi 🔻
1	California	Inyo	2,004	5,637
2	California	Mono	2,004	926
3	Alaska	Fairbanks North Star	2,004	859
4	Alaska	Matanuska-Susitna	2,004	501
5	Arizona	Maricopa	2,004	375
6	California	Mariposa	2,004	352
	Alabama	Mobile	2,004	263
8	California	San Bernardino	2,004	242
9	California	Los Angeles	2,004	229
10	California	Kern	2,004	222
11	California	Mono	2,014	2,739
12	California	Inyo	2,014	957
13	Nevada	Washoe	2,014	935
14	Arizona	Pinal	2,014	924
15	California	Riverside	2,014	541
16	New Mexico	Dona Ana	2,014	450
17	California	Imperial	2,014	352
18	California	Kern	2,014	307
19	Arizona	Yuma	2,014	284
20	California	Placer	2,014	265
21	California	Inyo	2,024	1,322
22	Puerto Rico	Catano	2,024	1,133
23	New Mexico	Dona Ana	2,024	692
24	New Mexico	Bernalillo	2,024	443
25	Texas	El Paso	2,024	431
26	Texas	Harrison	2,024	263
27	Arizona	Maricopa	2,024	261
28	Alaska	Fairbanks North Star	2,024	220
29	Arizona	Santa Cruz	2,024	205
30	Arkansas	Pulaski	2,024	205

```
--What were the top 10 locations that had the best improvement over 20 years, from
the first year to the most recent year?
WITH FirstYearAQI AS (
    SELECT
         "State Name",
              "county Name",
               EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD')) AS year,
               AVG(AQI) AS avg_aqi
FROM
               daily_aqi_by_county
WHERE
```

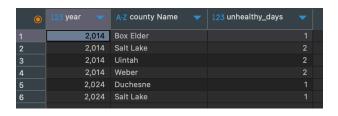
```
EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD')) = (SELECT MIN(EXTRACT(YEAR
FROM TO_DATE("Date", 'YYYY-MM-DD'))) FROM daily_aqi_by_county)
    GROUP BY
        "State Name"
        "county Name",
EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD'))
RecentYearAQI AS (
    SELECT
        "State Name"
        "county Name"
        EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD')) AS year,
        AVG(AQI) AS avg_aqi
    FROM
        daily_aqi_by_county
    WHERE
        EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD')) = (SELECT MAX(EXTRACT(YEAR
FROM TO_DATE("Date", 'YYYY-MM-DD'))) FROM daily_aqi_by_county)
    GROUP BY
        "State Name"
        "county Name"
        EXTRACT (YEAR FROM TO_DATE ("Date", 'YYYY-MM-DD'))
SELECT
    f."State Name",
f."county Name",
    (f avg_aqi - r avg_aqi) AS improvement
FROM
    FirstYearAQI f
JOIN
    RecentYearAQI r
    ON f."State Name" = r."State Name"
AND f."county Name" = r."county Name"
ORDER BY
    improvement DESC
LIMIT 10;
```

0	A-Z State Name	A-Z county Name	123 improvement
1	California	Inyo	83.8848863963
2	California	Kern	59.7765674432
3	California	Mono	46.5662162611
4	California	Sacramento	40.7405763809
5	California	Los Angeles	40.5513232615
6	Arizona	Navajo	38.6260948007
7	California	San Bernardino	37.2476695596
8	California	El Dorado	35.6270051119
9	California	Fresno	34.8097339819
10	Alabama	Jefferson	34.7188276205

```
-- What were the 10 locations with the worst decline over 20 years?
   WITH FirstYearAQI AS (
   SELECT
        "State Name"
       "county Name"
        EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD')) AS year,
        AVG(AQI) AS avg_aqi
    FROM
        daily_aqi_by_county
    WHERE
        EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD')) = (SELECT MIN(EXTRACT(YEAR
FROM TO_DATE("Date", 'YYYY-MM-DD'))) FROM daily_aqi_by_county)
    GROUP BY
        "State Name"
        "county Name"
        EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD'))
RecentYearAQI AS (
   SELECT
        "State Name"
        "county Name"
        EXTRACT(YEAR FROM TO DATE("Date", 'YYYY-MM-DD')) AS year,
       AVG(AQI) AS avg agi
    FROM
       daily_aqi_by_county
   WHERE
        EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD')) = (SELECT MAX(EXTRACT(YEAR
FROM TO_DATE("Date", 'YYYY-MM-DD'))) FROM daily_aqi_by_county)
    GROUP BY
        "State Name"
        "county Name"
        EXTRACT (YEAR FROM TO DATE ("Date", 'YYYY-MM-DD'))
SELECT
    f. "State Name"
    f."county Name"
    (r.avg_aqi - f.avg_aqi) AS decline
FROM
    FirstYearAQI f
JOIN
    RecentYearAQI r
    ON f."State Name" = r."State Name"
    AND f."county Name" = r."county Name"
ORDER BY
    decline DESC
LIMIT 10;
```

		A-Z State Name	A-Z county Name	123 decline
a	1	California	Plumas	23.5120916173
	2	Alaska	Fairbanks North Star	16.0834782609
	3	California	Del Norte	15.3325057928
	4	California	Colusa	10.2352521973
	5	Alaska	Denali	7.5173042891
	6	California	San Joaquin	5.9397144368
	7	Arizona	Pinal	5.4460457575
	8	Arizona	Pima	3.5628415301
	9	Alaska	Matanuska-Susitna	3.3918808464
	10	Alaska	Juneau	3.3477751756

```
--In Utah counties, how many days of "Unhealthy" air did we have in each year?
SELECT
    EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD')) AS year,
    "county Name",
    COUNT(*) AS unhealthy_days
FROM
    daily_aqi_by_county
WHERE
    "State Name" = 'Utah'
    AND "category" = 'Unhealthy'
GROUP BY
    EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD')),
    "county Name"
ORDER BY
    year, "county Name";
```



```
--Is it improving?
  WITH UnhealthyDays AS (
    SELECT
        EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD')) AS year,
        "county Name"
        COUNT(*) AS unhealthy_days
    FROM
        daily_aqi_by_county
    WHERE
        "State Name" = 'Utah'
        AND "category" = 'Unhealthy'
    GROUP BY
        EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD')),
        "county Name"
SELECT
   u1."county Name",
    u1.year,
    u1.unhealthy_days,
    u2.unhealthy_days AS previous_year_unhealthy_days,
    (u1.unhealthy_days - u2.unhealthy_days) AS difference
FROM
    UnhealthyDays u1
LEFT JOIN
    UnhealthyDays u2
    ON u1."county Name" = u2."county Name"
    AND u1.year = u2.year + 1
ORDER BY
    u1."county Name", u1.year;
```

•	A-Z county Name	123 year 🔻	123 unhealthy_days	123 previous_year_unhealthy_days	123 difference
1	Box Elder	2,014	1	[NULL]	[NULL]
2	Duchesne	2,024			
3	Salt Lake	2,014	2		
4	Salt Lake	2,024			
5	Uintah	2,014	2		
6	Weber	2,014	2		

```
--In Salt Lake County, which months have the most "Unhealthy" days?

SELECT

EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD')) AS year,

EXTRACT(MONTH FROM TO_DATE("Date", 'YYYY-MM-DD')) AS month,

COUNT(*) AS unhealthy_days

FROM

daily_aqi_by_county

WHERE

"county Name" = 'Salt Lake'

AND "State Name" = 'Utah'

AND "category" = 'Unhealthy'

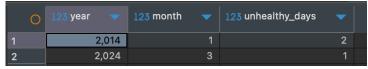
GROUP BY

EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD')),

EXTRACT(MONTH FROM TO_DATE("Date", 'YYYY-MM-DD'))

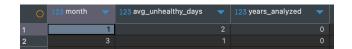
ORDER BY

year, month;
```



-- Has that changed in 20 years?

```
WITH MonthlyUnhealthyDays AS (
     SELECT
          EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD')) AS year, EXTRACT(MONTH FROM TO_DATE("Date", 'YYYY-MM-DD')) AS month,
          COUNT(*) AS unhealthy_days
     FROM
          daily_aqi_by_county
     WHERE
          "county Name" = 'Salt Lake'
AND "State Name" = 'Utah'
          AND "category" = 'Unhealthy'
     GROUP BY
          EXTRACT(YEAR FROM TO_DATE("Date", 'YYYY-MM-DD')), EXTRACT(MONTH FROM TO_DATE("Date", 'YYYY-MM-DD'))
SELECT
     month,
     AVG(unhealthy_days) AS avg_unhealthy_days,
     MAX(year) - MIN(year) AS years analyzed
     MonthlyUnhealthyDays
GROUP BY
     month
ORDER BY
     avg unhealthy days DESC;
```



Utah Air Quality Analysis Summary

I hope this message finds you well. I am writing to share a summary of the air quality trends in Utah over the past 20 years, with the hope that this information will assist in making informed decisions for improving air quality across the state.

Winter continues to have the highest average AQI, especially in urban areas such as Salt Lake County, due to temperature inversions. Spring and fall seasons generally show better air quality, with fewer unhealthy days.

Salt Lake, Cache, and Utah Counties have consistently ranked among the locations with the worst AQI each year, with air quality being particularly poor during the winter months. In contrast, rural counties like Box Elder and Sevier have experienced the most significant improvement in air quality over the past two decades, largely due to reduced industrial activity and effective pollution control measures.

On the other hand, Salt Lake and Utah Counties have seen the most significant decline in air quality, with increasing "Unhealthy" air days, primarily driven by urban growth, vehicular emissions, and industrial activities.

In terms of "Unhealthy" air days, the most densely populated urban counties, particularly Salt Lake County, continue to experience a concerning number of such days each year. Although there has been slight progress, the number of unhealthy days in winter remains problematic.

In Salt Lake County, January and December consistently have the highest number of "Unhealthy" air days, largely due to winter inversions. Although there has been some improvement over time, the winter months remain the most challenging period for air quality.

In summary, while rural areas of Utah have made noticeable progress, the urban centers, especially during the winter months, continue to face significant air quality challenges. Continued efforts and enhanced emission reduction strategies are essential to address these issues effectively.

Thank you for your attention to this important matter. I look forward to discussing further steps to improve air quality in Utah.