Air Pollution SQL project Report

1.What are the distinct cities available in the dataset?

```
SELECT DISTINCT
  City
FROM
  AirPollution_23;
```

	city
▶	Ahmedabad
	Aizawl
	Amaravati
	Amritsar
	Bengaluru

2. What is the average AQI for each city?

```
 SELECT
    City, AVG(AQI) AS Avg_AQI
FROM
    AirPollution_23
GROUP BY City;
```

city	Avg_AQI
Ahmedabad	3.8611
Aizawl	3.5903
Amaravati	3.7500
Amritsar	4.4722
Bengaluru	3.1250

3. Retrieve all records for the city 'Ahmedabad'.

```
SELECT * FROM  AirPollution_23
WHERE  City = 'Ahmedabad';
```

city	date	aqi	co	no	no2	o3	so2	pm2_5	pm10	nh3
Ahmedabad	01-01-2023	5	1535.42	16.76	47.3	0.02	17.88	107.17	141.36	26.09
Ahmedabad	02-01-2023	5	1842.5	24.14	35.3	0.01	19.07	163.21	203.9	30.65
Ahmedabad	03-01-2023	5	2002.72	28.61	34.62	0.01	25.51	186.03	228.58	29.64
Ahmedabad	04-01-2023	4	614.17	0	15.94	72.96	18.84	71.87	83.63	13.05
Ahmedabad	05-01-2023	5	560.76	0	9.17	128.75	21.93	86.01	93.2	7.79

4. How many records are there in total?

```
SELECT COUNT(*) FROM AirPollution_23;
```

COUNT(*)
3744

5. List the top 5 highest PM2_5 values with city and date.

```
SELECT
   City, Date, PM2_5
FROM
   AirPollution_23
ORDER BY PM2_5 DESC
LIMIT 5;
```

city	date	pm2_5
Kolkata	10-03-2023	1628.26
Kolkata	16-02-2023	1523.48
Kolkata	06-04-2023	1372.17
Kolkata	25-01-2023	1365.63
Kolkata	24-01-2023	1208.2

6. Which city had the highest average PM10 levels?

```
SELECT City, AVG(PM10) AS Avg_PM10
FROM AirPollution_23
GROUP BY City
ORDER BY Avg_PM10 DESC
LIMIT 1;
```

	city	Avg_PM10
▶	Kolkata	455.90486111111125

7. Find cities where the average CO level exceeded 1000.

```
SELECT City

FROM AirPollution

GROUP BY City

HAVING AVG(CO) > 1000;
```

	city
▶	Ahmedabad
	Amritsar
	Chennai
	Delhi
	Gurugram

8. Get the count of days where PM2_5 was over 100 for each city.

```
SELECT City, COUNT(*) AS High_PM2_5_Days

FROM AirPollution

WHERE PM2_5 > 100

GROUP BY City;
```

	city	High_PM2_5_Days
▶	Ahmedabad	41
	Aizawl	18
	Amaravati	31
	Amritsar	85
	Bengaluru	3

9. List all cities where the average NO2 level is greater than 40.

```
SELECT City, AVG(NO2) AS Avg_NO2

FROM AirPollution

GROUP BY City

HAVING AVG(NO2) > 40;
```

Result Grid | Filter Rows:

city	Avg_NO2
Delhi	45.88520833333332
Gurugram	55.498819444444436
Hyderabad	45.706666666666685
Jorapokhar	41.96972222222222
Kolkata	68.92784722222227

10. Which city has the greatest variation (difference) between average PM10 and average PM2_5 levels?

```
SELECT City,
    AVG(PM10) AS Avg_PM10,
    AVG(PM2_5) AS Avg_PM2_5,
    ABS(AVG(PM10) - AVG(PM2_5)) AS PM_Difference
FROM AirPollution
GROUP BY City
ORDER BY PM_Difference DESC
LIMIT 1;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Con

city	Avg_PM10	Avg_PM2_5	PM_Difference
Delhi	269.57277777777773	193.50506944444436	76.06770833333337

11. Find the rolling 7-day average AQI for each city.

```
SELECT City, Date, AQI,
    AVG(AQI) OVER (PARTITION BY City ORDER BY Date ROWS BETWEEN 6 PRECEDING AND
                                CURRENT ROW) AS Rolling_AQI
FROM AirPollution_23;
```

Result Grid | Filter Rows:

city	date	aqi	Rolling_AQI
Ahmedabad	01-01-2023	5	5.0000
Ahmedabad	01-02-2023	3	4.0000
Ahmedabad	01-03-2023	5	4.3333
Ahmedabad	01-04-2023	2	3.7500
Ahmedabad	01-05-2023	4	3.8000

12. Rank cities by their average pollution index combining AQI, PM2_5, and PM10

```
SELECT City,
    AVG(AQI + PM2_5 + PM10) AS Pollution_Score,
    RANK() OVER (ORDER BY AVG(AQI + PM2_5 + PM10) DESC) AS rnk
FROM AirPollution_23
GROUP BY City;
```

city	Pollution_Score	rnk
Kolkata	845.2459722222228	1
Delhi	467.68201388888906	2
Patna	417.40930555555553	3
Mumbai	377.31874999999997	4
Amritsar	366.81409722222224	5

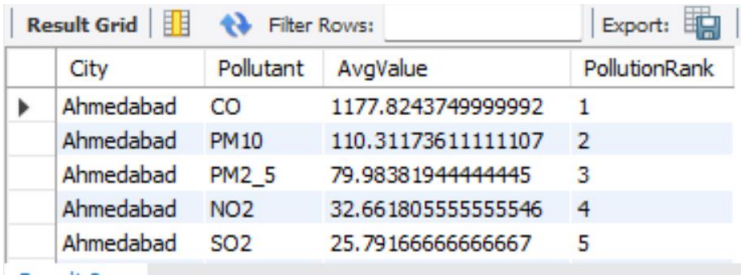
13. Find the top 3 cities with the most consistent (least variable) AQI values using standard deviation.

```
SELECT City,
    STDDEV(AQI) AS AQI_StdDev
FROM AirPollution
GROUP BY City
ORDER BY AQI_StdDev ASC
LIMIT 3;
```

city	AQI_StdDev
Kolkata	0.5739573880814962
Patna	0.5852728998372901
Lucknow	0.5944704044175743

14. For each city, rank the pollutants (CO, NO, NO2, SO2, PM2_5, PM10, NH3) by their average concentration.

```
SELECT City, Pollutant, AvgValue,
    RANK() OVER (PARTITION BY City ORDER BY AvgValue DESC) AS PollutionRank
FROM (
  SELECT City, 'CO' AS Pollutant, AVG(CO) AS AvgValue FROM AirPollution_23 GROUP BY City
  UNION ALL
  SELECT City, 'NO', AVG(NO) FROM AirPollution_23 GROUP BY City
  UNION ALL
  SELECT City, 'NO2', AVG(NO2) FROM AirPollution_23 GROUP BY City
  UNION ALL
  SELECT City, 'SO2', AVG(SO2) FROM AirPollution_23 GROUP BY City
  UNION ALL
  SELECT City, 'PM2_5', AVG(PM2_5) FROM AirPollution_23 GROUP BY City
  UNION ALL
  SELECT City, 'PM10', AVG(PM10) FROM AirPollution_23 GROUP BY City
  UNION ALL
  SELECT City, 'NH3', AVG(NH3) FROM AirPollution_23 GROUP BY City
) AS PollutantStats;
```



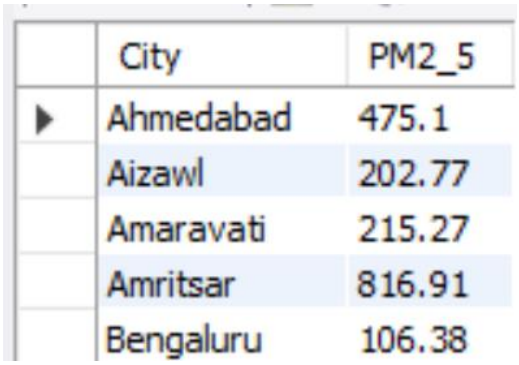
Result Grid | Filter Rows: | Export:

City	Pollutant	AvgValue	PollutionRank
Ahmedabad	CO	1177.8243749999992	1
Ahmedabad	PM10	110.31173611111107	2
Ahmedabad	PM2_5	79.98381944444445	3
Ahmedabad	NO2	32.661805555555546	4
Ahmedabad	SO2	25.79166666666667	5

15. Find the highest PM2_5 reading per city (with window function)

```
SELECT City, Date, PM2_5
FROM (
  SELECT City, Date, PM2_5,
      RANK() OVER (PARTITION BY City ORDER BY PM2_5 DESC) AS rnk
```

```
  FROM AirPollution
) AS ranked
WHERE rnk = 1;
```



City	PM2_5
Ahmedabad	475.1
Aizawl	202.77
Amaravati	215.27
Amritsar	816.91
Bengaluru	106.38