– Weather Data analysis
– Data Validation with VIZ

```
-- Monthly temperature avg
SELECT
    TO_CHAR(date, 'Month') AS month_name,
    ROUND(AVG(temperature)::numeric, 2) AS avg_temperature
FROM weather_schemas.combined_data
GROUP BY TO_CHAR(date, 'Month'),
      EXTRACT(MONTH FROM date)
ORDER BY EXTRACT(MONTH FROM date);
```

	month_name text	avg_temperature numeric
1	January	18.34
2	February	17.57
3	March	16.25
4	April	17.40
5	May	17.45
6	June	17.73
7	July	17.34
8	August	17.57
9	September	17.93
10	October	17.23
11	November	16.75
12	December	17.88

```
-- Average temperature by country
SELECT
        country,
        ROUND(AVG(temperature)::numeric,2) AS Avg_Temperature_By_Country
FROM weather_schemas.combined_data
GROUP BY country
ORDER BY Avg_Temperature_By_Country DESC;
```

	country text	avg_temperature_by_country numeric
1	Canada	17.90
2	Australia	17.84
3	Brazil	17.48
4	India	17.43
5	South Afri...	17.40
6	USA	17.14
7	Germany	16.97

```
-- Extreme weather events by months
SELECT
        TO_CHAR(date, 'Month') AS Month_Name,
        COUNT(*) AS Event_Count
FROM weather_schemas.combined_data
WHERE extreme_weather_events <> 'None'
GROUP BY TO_CHAR(date, 'Month')
ORDER BY MIN(date);
```

	month_name text	event_count bigint
1	January	266
2	February	233
3	March	268
4	April	233
5	May	264
6	June	238
7	July	254
8	August	273
9	September	248
10	October	261
11	November	227
12	December	258

```
-- Country-wise extreme weather
SELECT
        country,
        COUNT(*) AS Event_Count
FROM weather_schemas.combined_data
WHERE extreme_weather_events <> 'None'
```

GROUP BY country
ORDER BY Event_Count DESC;

	country text	event_count bigint
1	South Afri...	453
2	Brazil	443
3	USA	438
4	Australia	434
5	Canada	433
6	India	414
7	Germany	408

```
-- Extreme Weather Events by Temperature Range
SELECT
        CASE
                WHEN temperature < 10 THEN 'Very Cold (<10°C)'
                WHEN temperature BETWEEN 10 AND 15 THEN 'Cold (10-15°C)'
                WHEN temperature BETWEEN 15 AND 20 THEN 'Moderate (15-20°C)'
                WHEN temperature BETWEEN 20 AND 25 THEN 'Warm (20-25°C)'
                ELSE 'Hot (>25°C)'
        END AS Temperature_Range,
        extreme_weather_events,
        COUNT(*) AS Event_Count
FROM weather_schemas.combined_data
WHERE extreme_weather_events <> 'None'
GROUP BY Temperature_Range,
            extreme_weather_events
ORDER BY Temperature_Range,
            Event_Count DESC;
```

	temperature_range text	extreme_weather_events text	event_count bigint
2	Cold (10-15°C)	Drought	76
3	Cold (10-15°C)	Heatwave	75
4	Cold (10-15°C)	Hurricane	67
5	Hot (>25°C)	Drought	286
6	Hot (>25°C)	Heatwave	268
7	Hot (>25°C)	Hurricane	258
8	Hot (>25°C)	Flood	245
9	Moderate (15-20°C)	Hurricane	81
10	Moderate (15-20°C)	Flood	66
11	Moderate (15-20°C)	Drought	63
12	Moderate (15-20°C)	Heatwave	61
13	Very Cold (<10°C)	Heatwave	301
14	Very Cold (<10°C)	Hurricane	289
15	Very Cold (<10°C)	Flood	263
16	Very Cold (<10°C)	Drought	260
17	Warm (20-25°C)	Flood	84
18	Warm (20-25°C)	Hurricane	74
19	Warm (20-25°C)	Heatwave	66
20	Warm (20-25°C)	Drought	59

```
-- which cities are experiencing extreme weather events this week and what are their
economic and population impacts?
SELECT
    country,
    city,
    extreme_weather_events,
    COUNT(*) AS event_count,
    ROUND(AVG(temperature)::numeric, 1) AS average_temperature,
    SUM(population_exposure) AS total_population_exposure,
    SUM(economic_impact_estimate) AS total_economic_impact,
    ROUND(AVG(infrastructure_vulnerability_score)::numeric, 0) AS average_vulnerability
FROM weather_schemas.combined_data
WHERE date BETWEEN '2025-03-03' AND '2025-03-07'
  AND extreme_weather_events != 'None'
GROUP BY country, city, extreme_weather_events
ORDER BY total_economic_impact DESC;
```

	country (text)	city (text)	extreme_weather_events (text)	event_count (bigint)	average_temperature (numeric)	total_population_exposure (numeric)	total_economic_impact (numeric)	average_vulnerability (numeric)
1	South Afri...	Durban	Flood	1	30.6	772529	98765751.15	5
2	Brazil	Brasilia	Drought	1	26.9	4164128	96514968.05	4
3	Canada	Toronto	Drought	1	-9.5	1310863	95231429.51	5
4	USA	New York	Drought	1	14.5	8499163	93366392.55	5
5	USA	Los Angeles	Flood	1	-6.5	1335844	91454336.88	8
6	South Afri...	Johannesbu...	Drought	1	35.5	5610204	90575436.02	10
7	USA	New York	Flood	1	27.3	9011115	85925007.67	7
8	Germany	Hamburg	Heatwave	1	34.9	8625068	72538730.19	3
9	South Afri...	Cape Town	Heatwave	1	38.6	9225533	70238861.05	3
10	India	Delhi	Flood	1	7.9	1297995	68218445.17	6
11	USA	New York	Hurricane	1	-7.7	6883355	67123848.04	10
12	Germany	Berlin	Hurricane	1	24.7	7910619	65892689.57	10
13	Canada	Toronto	Heatwave	1	17.5	2414228	58917417.9	1
14	Canada	Montreal	Hurricane	1	39.2	1910078	47481103.06	8
15	India	Mumbai	Heatwave	1	27.0	4723866	36593764.94	1
16	USA	New York	Heatwave	1	35.8	6315063	33073167.48	6
17	Germany	Munich	Hurricane	1	29.6	6513387	26753340.75	3
18	Australia	Melbourne	Flood	1	33.9	417575	8159365.56	1

-- what are the top 5 cities with the highest air quality concerns and their associate risks?

```sql
SELECT
    country,
    city,
    ROUND(AVG(air_quality_index)::numeric, 0) AS average_aqi,
    COUNT(*) FILTER (WHERE air_quality_index > 200) AS days_above_200_aqi,
    SUM(population_exposure) AS total_population_exposure,
    ROUND(AVG(temperature)::numeric, 1) AS average_temperature
FROM weather_schemas.combined_data
WHERE date BETWEEN '2025-03-03' AND '2025-03-07'
GROUP BY country, city
HAVING AVG(air_quality_index) > 100
ORDER BY average_aqi DESC
LIMIT 5;
```

	country (text)	city (text)	average_aqi (numeric)	days_above_200_aqi (bigint)	total_population_exposure (numeric)	average_temperature (numeric)
1	Germany	Munich	387	5	29283894	25.2
2	India	Bangalore	351	5	34053434	32.1
3	Germany	Berlin	335	4	29999584	20.4
4	USA	Los Angel...	311	4	26100024	16.9
5	Brazil	Sao Paulo	307	4	32737352	23.8

```sql
-- Which biome types are risky from extreme weather events this week?
SELECT
    biome_type,
    COUNT(*) AS total_records,
    COUNT(DISTINCT country || '-' || city) AS locations_affected,
    COUNT(*) FILTER (WHERE extreme_weather_events != 'None') AS extreme_weather_count,
    STRING_AGG(DISTINCT extreme_weather_events, ', ') AS event_types,
    ROUND(AVG(temperature)::numeric, 1) AS average_temperature,
    SUM(economic_impact_estimate) AS total_economic_impact_estimate,
    ROUND(AVG(infrastructure_vulnerability_score)::numeric, 0) AS average_vulnerability
FROM weather_schemas.combined_data
WHERE date BETWEEN '2025-03-03' AND '2025-03-07'
GROUP BY biome_type
ORDER BY extreme_weather_count DESC;
```

	biome_type text	total_records bigint	locations_affected bigint	extreme_weather_count bigint	event_types text	average_temperature numeric	total_economic_impact_estimate numeric	average_vulnerability numeric
1	Tundra	26	18	7	Drought, Flood, Heatwave, Hurricane, N...	18.6	1521395358.41	6
2	Desert	25	18	5	Drought, Heatwave, Hurricane, None	19.5	1319873484.62	5
3	Forest	17	13	2	Flood, Hurricane, None	19.9	799407634.60	7
4	Grassland	19	14	2	Heatwave, Hurricane, None	22.0	1068900741.87	5
5	Wetland	18	14	2	Drought, None	16.1	901598135.28	5