Scenario

A global research institution is studying the impact of climate change across different regions. They need a centralized system to track key climate indicators, monitor extreme weather events, and analyse their economic and infrastructural impact.

Business Problem

The organization faces challenges in:
Tracking Climate Trends – Data is scattered across multiple sources, making it difficult to analyse temperature variations, air quality, and precipitation patterns over time.
Generating Reports Efficiently – Researchers rely on manual reporting, leading to delays in decision-making.
Assessing Climate Risks – There is no structured way to analyse how climate events impacts infrastructure and the economy in different regions.

To address these issues, we are going to develop a data-driven climate monitoring solution with

automated reporting and real-time visualization, ensuring quick access to insights for informed decision-making.

Climate Change Dataset: Column Descriptions

Metadata Columns

Record ID

A unique identifier assigned to each individual climate data record.

Date

The specific date when the climate observation was recorded.

Geographic Columns

Country

The nation where the climate data was collected.

City

The specific urban location where the data was gathered.

Climate and Environmental Metrics

Temperature (°C)

Measurement of the ambient air temperature in degrees Celsius.

Humidity (%)

The amount of water vapor present in the air, expressed as a percentage.

Precipitation (mm)

The total amount of rainfall or water equivalent measured in millimeters.

Air Quality Index (AQI)

A numerical scale that indicates the level of air pollution and potential health risks.

Extreme Weather Events

Significant and unusual meteorological occurrences such as hurricanes, heatwaves, or droughts.

Classification and Contextual Columns

Climate Classification (Koeppen)

A scientific system for categorizing global climate types based on temperature and precipitation patterns.

Climate Zone

A broad classification of the ecological climate characteristics of a specific region.

Biome Type

A large-scale biological community defined by its distinctive plant and animal species and environmental conditions.

Meteorological Columns

Heat Index

A combined measure of air temperature and relative humidity that represents how hot it actually feels.

Wind Speed

The rate of air movement measured at the location.

Wind Direction

The compass direction from which the wind is blowing.

Season

The specific time of year when the data was collected.

Impact and Vulnerability Columns

Population Exposure

The number of people potentially affected by the observed climate conditions.

Economic Impact Estimate

A monetary valuation of the potential economic consequences related to the climate conditions.

Infrastructure Vulnerability Score

A numerical rating that assesses the potential risk and susceptibility of infrastructure to climate-related challenges.

1.SQL Data Cleaning

-- Create a combined table

Create table "Climate Change"."Combined Data" as

select * from "Climate Change"."Australia"

union

select * from "Climate Change"."Brazil"

union

select * from "Climate Change"."Canada"

union

select * from "Climate Change"."Germany"

union

select * from "Climate Change"."India"

union

select * from "Climate Change"."South Africa"

union

select * from "Climate Change"."United States"

```
-- Check for duplicates

SELECT "Record ID"
FROM "Climate Change"."Combined Data"
group by "Record ID"
having count(*) > 1;


SELECT distinct "Country"
FROM "Climate Change"."Combined Data"


-- Update the country

update "Climate Change"."Combined Data"
set "Country" = 'India'
where "Country" = 'Inda';



-- Check for null values

select *
from "Climate Change"."Combined Data"
where "Record ID" IS NULL
  or "Date" IS NULL
  or "Country" IS NULL
```

```
  or "City" IS NULL
  or "Temperature" IS NULL
  or "Humidity" IS NULL
  or "Precipitation" IS NULL
  or "Air Quality Index" IS NULL
  or "Extreme Weather Events" IS NULL
  or "Climate Classification" IS NULL
  or "Climate Zone" IS NULL
  or "Biome Type" IS NULL
  or "Heat Index" IS NULL
  or "Wind Speed" IS NULL
  or "Wind Direction" IS NULL
  or "Season" IS NULL
  or "Population Exposure" IS NULL
  or "Economic Impact Estimate" IS NULL
  or "Infrastructure Vulnerability Score" IS NULL;



-- Update Population Exposure

update "Climate Change"."Combined Data"
set "Population Exposure" = 5275135
where "Record ID" = 'aus_1338';
```

-- Update City

```
update "Climate Change"."Combined Data"
set "City" = 'Toronto'
where "Record ID" = 'cnd_227';
```

2.SQL Data Analysis

-- Monthly Temperature Trends

```
SELECT TO_CHAR("Date", 'Month') AS Month_Name,
    AVG("Temperature") AS Avg_Temperature
FROM "Climate Change"."Combined Data"
GROUP BY TO_CHAR("Date", 'Month'), EXTRACT(MONTH FROM "Date")
ORDER BY EXTRACT(MONTH FROM "Date");
```

-- average temperature by country

```
SELECT "Country",
    AVG("Temperature") AS Avg_Temperature
FROM "Climate Change"."Combined Data"
```

```
GROUP BY "Country"
ORDER BY Avg_Temperature DESC;
```

-- Extreme Weather Events by Month

```
SELECT TO_CHAR("Date", 'Month') AS Month_Name,
       COUNT(*) AS Event_Count
FROM "Climate Change"."Combined Data"
WHERE "Extreme Weather Events" <> 'None'
GROUP BY TO_CHAR("Date", 'Month')
ORDER BY MIN("Date");
```

-- Country-wise Extreme Weather

```
SELECT "Country",
       COUNT(*) AS Event_Count
FROM "Climate Change"."Combined Data"
WHERE "Extreme Weather Events" <> 'None'
GROUP BY "Country"
ORDER BY Event_Count DESC;
```

-- 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 "Climate Change"."Combined Data"
WHERE "Extreme Weather Events" <> 'None'
GROUP BY Temperature_Range, "Extreme Weather Events"
ORDER BY Temperature_Range, Event_Count DESC;
```

-- 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 Type",
        Round(avg("Temperature"), 1) as "Average Temperature",
        sum("Population Exposure") as "Total Population Exposure",
```

```
        sum("Economic Impact Estimate") as "Total Economic Impact",
        round(avg("Infrastructure Vulnerability Score"), 0) as "Average Vulnerability"
from "Climate Change"."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;
```

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

```
select
    "Country",
        "City",
        round(avg("Air Quality Index"), 0) as "Average AQI",
        count(*) as "Days above 200 AQI",
        SUM("Population Exposure") as "Total Population Exposure",
        round(avg("Temperature"), 1) as "Average Temperature"
from "Climate Change"."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"
limit 5;
```

-- Which biome types are most risk from extreme weather events this week?

```
select
    "Biome Type",
        count(*) as "Total Records",
        count(distinct concat("Country", "City")) as "Locations Affected",
        count(case when "Extreme Weather Events" != 'None' then 1 end) as "Extreme Weather
Count",
        STRING_AGG(DISTINCT "Extreme Weather Events", ', ') as "Event Types",
        Round(avg("Temperature"), 1) as "Average Temperature",
        sum("Economic Impact Estimate") as "Total Economic Impact Estimate",
        Round(Avg("Infrastructure Vulnerability Score"), 0) as "Average Vulnerability"
from "Climate Change"."Combined Data"
where "Date" between '2025-03-03' and '2025-03-07'
group by "Biome Type"
```

3. Tableau

Avg AQI

```
AVG([Air Quality Index])
```

Current Month AQI

```
IF DATENAME('month', [Date]) = [Current Month] THEN
{FIXED DATENAME('month', [Date]) : AVG([Air Quality Index])}
```

END

Previous Month AQI

```
IF DATENAME('month', [Date]) =
case [Current Month]
   WHEN 'January' THEN 'December'
   WHEN 'February' THEN 'January'
   WHEN 'March' THEN 'February'
   WHEN 'April' THEN 'March'
   WHEN 'May' THEN 'April'
   WHEN 'June' THEN 'May'
   WHEN 'July' THEN 'June'
   WHEN 'August' THEN 'July'
   WHEN 'September' THEN 'August'
   WHEN 'October' THEN 'September'
   WHEN 'November' THEN 'October'
   WHEN 'December' THEN 'November'
   END
THEN {FIXED DATENAME('month', [Date]) : AVG([Air Quality Index])}
END
```

% Difference AQI

```
(AVG([Current Month AQI]) - AVG([Previous Month AQI])) / AVG([Previous Month AQI])
```

Bad Percentage AQI

```
IF [% Difference AQI] >= 0.03
THEN
    IF [% Difference AQI] > 0
    THEN "▲ " + STR(ROUND([% Difference AQI] * 100, 2)) + "%"  // Increase (bad)
    ELSE "▼ " + STR(ROUND([% Difference AQI] * 100, 2)) + "%"  // Decrease (bad)
```

```
    END
ELSE
""
END
```

Good Percentage AQI

```
IF [% Difference AQI] < 0.03
THEN
    IF [% Difference AQI] > 0
    THEN "▲ " + STR(ROUND([% Difference AQI] * 100, 2)) + "%"  // Increase (good)
   ELSE "▼ " + STR(ROUND([% Difference AQI] * 100, 2)) + "%"  // Decrease (good)
   END
ELSE
""
END
```

Count of EWE

```
IF [Extreme Weather Events] <> "None" THEN 1 ELSE 0
END
```

Current Month EWE

```
IF DATENAME('month', [Date]) = [Current Month] THEN [Count of EWE]
END
```

Previous Month EWE

```
IF DATENAME('month', [Date]) =
```

```
case [Current Month]
    WHEN 'January' THEN 'December'
    WHEN 'February' THEN 'January'
    WHEN 'March' THEN 'February'
    WHEN 'April' THEN 'March'
    WHEN 'May' THEN 'April'
    WHEN 'June' THEN 'May'
    WHEN 'July' THEN 'June'
    WHEN 'August' THEN 'July'
    WHEN 'September' THEN 'August'
    WHEN 'October' THEN 'September'
    WHEN 'November' THEN 'October'
    WHEN 'December' THEN 'November'
    END
THEN [Count of EWE]
END
```

% Difference EWE

```
(SUM([Current Month EWE]) - SUM([Previous Month EWE])) / SUM([Previous Month EWE])
```

Bad Percentage EWE

```
IF [% Difference EWE] > 0 THEN "▲ " + STR(ROUND([% Difference EWE] * 100, 2)) + "%"
ELSE
""
END
```

Good Percentage EWE

```
IF [% Difference EWE] < 0 THEN "▼ " + STR(ROUND([% Difference EWE] * 100, 2)) + "%"
ELSE
""
```

END

Avg Precipitation Intensity

AVG([Precipitation])

Current Month Precipitation Intensity

IF DATENAME('month', [Date]) = [Current Month] THEN

{FIXED DATENAME('month', [Date]) : AVG([Precipitation])}

END

Previous Month Precipitation Intensity

IF DATENAME('month', [Date]) =

case [Current Month]

WHEN 'January' THEN 'December'

WHEN 'February' THEN 'January'

WHEN 'March' THEN 'February'

WHEN 'April' THEN 'March'

WHEN 'May' THEN 'April'

WHEN 'June' THEN 'May'

WHEN 'July' THEN 'June'

WHEN 'August' THEN 'July'

WHEN 'September' THEN 'August'

WHEN 'October' THEN 'September'

WHEN 'November' THEN 'October'

WHEN 'December' THEN 'November'

END

THEN {FIXED DATENAME('month', [Date]) : AVG([Precipitation])}

END

% Difference Precipitation Intensity

(AVG([Current Month Precipitation Intensity]) - AVG([Previous Month Precipitation Intensity])) / AVG([Previous Month Precipitation Intensity])

Bad Percentage Precipitation Intensity

IF [% Difference Precipitation Intensity] <= -0.02 or [% Difference Precipitation Intensity] >= 0.02 THEN

  IF [% Difference Precipitation Intensity] > 0

  THEN "▲ " + STR(ROUND([% Difference Precipitation Intensity] * 100, 2)) + "%"

  ELSE "▼ " + STR(ROUND([% Difference Precipitation Intensity] * 100, 2)) + "%"

  END

ELSE

""

END

Good Percentage Precipitation Intensity

IF [% Difference Precipitation Intensity] > -0.02 AND [% Difference Precipitation Intensity] < 0.02 THEN

  IF [% Difference Precipitation Intensity] > 0

  THEN "▲ " + STR(ROUND([% Difference Precipitation Intensity] * 100, 2)) + "%"

  ELSE "▼ " + STR(ROUND([% Difference Precipitation Intensity] * 100, 2)) + "%"

  END

ELSE

""

END

Avg Temperature

```
AVG([Temperature])
```

Current Month Temperature

```
IF DATENAME('month', [Date]) = [Current Month] THEN
{FIXED DATENAME('month', [Date]) : AVG([Temperature])}
END
```

Previous Month Temperature

```
IF DATENAME('month', [Date]) =
case [Current Month]
   WHEN 'January' THEN 'December'
   WHEN 'February' THEN 'January'
   WHEN 'March' THEN 'February'
   WHEN 'April' THEN 'March'
   WHEN 'May' THEN 'April'
   WHEN 'June' THEN 'May'
   WHEN 'July' THEN 'June'
   WHEN 'August' THEN 'July'
   WHEN 'September' THEN 'August'
   WHEN 'October' THEN 'September'
   WHEN 'November' THEN 'October'
   WHEN 'December' THEN 'November'
   END
THEN {FIXED DATENAME('month', [Date]) : AVG([Temperature])}
```

END

% Difference Temperature

(AVG([Current Month Temperature]) - AVG([Previous Month Temperature])) / AVG([Previous Month Temperature])

Bad Percentage Temperature

IF [% Difference Temperature] >= 0.03 OR [% Difference Temperature] <= -0.03

THEN

IF [% Difference Temperature] > 0

THEN "▲ " + STR(ROUND([% Difference Temperature] * 100, 2)) + "%"

ELSE "▼ " + STR(ROUND([% Difference Temperature] * 100, 2)) + "%"

END

ELSE

""

END

Good Percentage Temperature

IF [% Difference Temperature] > -0.03 AND [% Difference Temperature] < 0.03

THEN

IF [% Difference Temperature] > 0

THEN "▲ " + STR(ROUND([% Difference Temperature] * 100, 2)) + "%"

ELSE "▼ " + STR(ROUND([% Difference Temperature] * 100, 2)) + "%"

END

ELSE

""

END

Temperature Variability

```
STDEV([Temperature])
```

Current Month Temperature Variability

```
IF DATENAME('month', [Date]) = [Current Month] THEN
{FIXED DATENAME('month', [Date]) : STDEV([Temperature])}
END
```

Previous Month Temperature Variability

```
IF DATENAME('month', [Date]) =
case [Current Month]
   WHEN 'January' THEN 'December'
   WHEN 'February' THEN 'January'
   WHEN 'March' THEN 'February'
   WHEN 'April' THEN 'March'
   WHEN 'May' THEN 'April'
   WHEN 'June' THEN 'May'
   WHEN 'July' THEN 'June'
   WHEN 'August' THEN 'July'
   WHEN 'September' THEN 'August'
   WHEN 'October' THEN 'September'
   WHEN 'November' THEN 'October'
   WHEN 'December' THEN 'November'
   END
THEN {FIXED DATENAME('month', [Date]) : STDEV([Temperature])}
END
```

% Difference Temperature Variability

(AVG([Current Month Temperature Variability]) - AVG([Previous Month Temperature Variability])) / AVG([Previous Month Temperature Variability])

Bad Percentage Temperature Variability

```
IF  [% Difference Temperature Variability] >= 0.01 THEN
    IF [% Difference Temperature Variability] > 0
    THEN "▲ " + STR(ROUND([% Difference Temperature Variability] * 100, 2)) + "%"
    ELSE "▼ " + STR(ROUND([% Difference Temperature Variability] * 100, 2)) + "%"
    END
ELSE
""
END
```

Good Percentage Temperature Variability

```
IF  [% Difference Temperature Variability] < 0.01 THEN
    IF [% Difference Temperature Variability] > 0
    THEN "▲ " + STR(ROUND([% Difference Temperature Variability] * 100, 2)) + "%"
    ELSE "▼ " + STR(ROUND([% Difference Temperature Variability] * 100, 2)) + "%"
    END
ELSE
""
END
```