

SQL PROJECT

PRESENTED BY ARIJEET MUKHERJEE
arijeetm622@gmail.com

CLIMATE CHANGE PROJECT (SQL PART)

As all the Roadmap has been provided on the previous presentation. We are directly jumping into the timeline and will start this presentation explaining each and everything along the way we progress

IF want to checkout the roadmap kindly refer to this
LINK: [CLICK HERE](#) (ROADMAP OF THE PROJECT)

DATA UNDERSTANDING



1 There were 7 data sets generated. from chatgpt each representing continent(Antartica is not considered) Then We opened VS CODE and connected postgres from it so that we can start data cleaning and processing and after that combining all this tables in order to get collective insights at one sheet then running some queries to get some insights about all the column's relativity with each other.

4

Checked the combined table then checked for duplicates (NOT PRESENT) and fixed some minor error before proceeding further

```
SELECT * FROM climate_change;  
  
--CHECK DUPLICATE VALUE  
SELECT record_id  
FROM climate_change  
GROUP BY record_id  
HAVING COUNT(*) >1;  
  
--UPDATE INDIA TO INDIA  
UPDATE climate_change  
SET country ='India'  
WHERE country ='Inda';
```

usa_climate_data	8/5/2025 10:12 AM	Microsoft Excel Co...	266 KB
south_africa_climate_data	8/5/2025 10:12 AM	Microsoft Excel Co...	284 KB
india_climate_data	8/5/2025 10:12 AM	Microsoft Excel Co...	253 KB
germany_climate_data	8/5/2025 10:12 AM	Microsoft Excel Co...	260 KB
canada_climate_data	8/5/2025 10:12 AM	Microsoft Excel Co...	260 KB
brazil_climate_data	8/5/2025 10:12 AM	Microsoft Excel Co...	260 KB
australia_climate_data	8/5/2025 10:12 AM	Microsoft Excel Co...	260 KB

3 COMBINED everything in one

```
--create a combined table  
CREATE TABLE climate_change AS  
SELECT * FROM australia_climate_data  
UNION  
SELECT * FROM brazil_climate_data  
UNION  
SELECT * FROM canada_climate_data  
UNION  
SELECT * FROM germany_climate_data  
UNION  
SELECT * FROM india_climate_data  
UNION  
SELECT * FROM south_africa_climate_data  
UNION  
SELECT * FROM usa_climate_data;
```

2

This is how we created tables for rest 6 country and imported data from the csv file and by SELECT * FROM checked whether the data has been imported properly or not

```
CREATE TABLE australia_climate_data (  
    record_id VARCHAR(10),  
    date DATE,  
    country VARCHAR(100),  
    city VARCHAR(100),  
    temperature_c NUMERIC(5,2),  
    humidity_percent NUMERIC(5,2),  
    precipitation_mm NUMERIC(6,2),  
    air_quality_index INTEGER,  
    extreme_weather_events VARCHAR(255),  
    climate_classification VARCHAR(100),  
    climate_zone VARCHAR(50),  
    biome_type VARCHAR(100),  
    heat_index NUMERIC(5,2),  
    wind_speed NUMERIC(5,2),  
    wind_direction VARCHAR(20),  
    season VARCHAR(20),  
    population_exposure INTEGER,  
    economic_impact_estimate NUMERIC(15,2),  
    infrastructure_vulnerability_score NUMERIC(4,2)  
);  
  
SELECT * FROM australia_climate_data
```

```
--CHECK FOR NULL VALUES
SELECT *
FROM climate_change
WHERE record_id IS NULL
OR date IS NULL
OR country IS NULL
OR city IS NULL
OR temperature_c IS NULL
OR humidity_percent IS NULL
OR precipitation_mm 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 population_exposure IS NULL
OR economic_impact_estimate IS NULL
OR infrastructure_vulnerability_score IS NULL;
```

```
--UPDATE population_exposure
UPDATE climate_change
SET population_exposure = 5275135
WHERE record_id = 'aus_1338';
```

```
--UPDATE city
UPDATE climate_change
SET city = 'Toronto'
WHERE record_id = 'cnd_227';
```

5

Data cleaning and processing done followed by fixing null value NOW it's time for some queries

★ QUERIES ★

6

--DATA ANALYTICS

-- Q1 MONTHLY TEMPERATURE TRENDS

```
SELECT TO_CHAR(date, 'Month') AS month_name, AVG(temperature_c) AS avg_temp
FROM climate_change
GROUP BY TO_CHAR(date, 'Month'), EXTRACT(Month FROM date)
ORDER BY EXTRACT(Month FROM date);
```

month_name	avg_temp
abc Filter...	abc Filter...
January	18.3371735791090630
February	17.5745196324143693
March	16.2473886328725038
April	17.4015873015873016
May	17.4483102918586790
June	17.7319841269841270
July	17.3412442396313364
August	17.5662826420890937
September	17.9265079365079365
October	17.2271889400921659
November	16.7501587301587302
December	17.8826420890937020

--Q2 AVG TEMPERATURE BY COUNTRY

```
SELECT country, AVG(temperature_c) AS avg_temp
FROM climate_change
GROUP BY country
ORDER BY avg_temp DESC;
```

country	avg_temp
abc Filter...	abc Filter...
Canada	17.9033287733698130
Australia	17.8424076607387141
Brazil	17.4823529411764706
India	17.4284541723666211
South Africa	17.3990880072959416
USA	17.1423164614683083
Germany	16.9666666666666667

8

--Q3 EXTREME WEATHER EVENTS OVER TIME

```
SELECT TO_CHAR(date,'Month')AS month_name, COUNT(*) AS event_count
FROM climate_change
WHERE extreme_weather_events <> 'None'
GROUP BY TO_CHAR(date,'Month')
ORDER BY event_count DESC;
```

month_name	event_count
abc Filter...	abc Filter...
August	273
March	268
January	266
May	264
October	261
December	258
July	254
September	248
June	238
February	233
April	233
November	227



country	event_count
abc Filter...	abc Filter...
South Africa	453
Brazil	443
USA	438
Australia	434
Canada	433
India	414
Germany	408

9

--Q4 COUNTRYWISE EXTREME WEATHER EVENTS

```
SELECT country,COUNT(*)AS event_count
FROM climate_change
WHERE extreme_weather_events <> 'None'
GROUP BY country
ORDER BY event_count DESC;
```

10

```
--Q5 RELATIONSHIP BETWEEN TEMPERATURE AND EXTREME WEATHER EVENTS
SELECT
CASE
WHEN temperature_c < 10 THEN 'Very Cold (<10°C)'
WHEN temperature_c BETWEEN 10 AND 15 THEN 'Cold (10-15°C)'
WHEN temperature_c BETWEEN 15 AND 20 THEN 'Moderate (15-20°C)'
WHEN temperature_c 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
WHERE extreme_weather_events <> 'None'
GROUP BY Temperature_Range, extreme_weather_events
ORDER BY Temperature_Range, Event_Count DESC;
```

temperature_ra...	extreme_weathe...	event_count
Cold (10-15°C)	Flood	81
Cold (10-15°C)	Drought	76
Cold (10-15°C)	Heatwave	75
Cold (10-15°C)	Hurricane	67
Hot (>25°C)	Drought	286
Hot (>25°C)	Heatwave	268
Hot (>25°C)	Hurricane	258
Hot (>25°C)	Flood	245
Moderate (15-20°C)	Hurricane	81
Moderate (15-20°C)	Flood	66
Moderate (15-20°C)	Drought	63
Moderate (15-20°C)	Heatwave	61
Very Cold (<10°C)	Heatwave	301
Very Cold (<10°C)	Hurricane	289
Very Cold (<10°C)	Flood	263
Very Cold (<10°C)	Drought	260
Warm (20-25°C)	Flood	84
Warm (20-25°C)	Hurricane	74
Warm (20-25°C)	Heatwave	66
Warm (20-25°C)	Drought	59

11

```
-- Q6 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_c), 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
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	city	extreme_weathe...	event_type
South Africa	Durban	Flood	1
Brazil	Brasilia	Drought	1
Canada	Toronto	Drought	1
USA	New York	Drought	1
USA	Los Angeles	Flood	1
South Africa	Johannesburg	Drought	1
USA	New York	Flood	1
Germany	Hamburg	Heatwave	1
South Africa	Cape Town	Heatwave	1
India	Delhi	Flood	1
USA	New York	Hurricane	1
Germany	Berlin	Hurricane	1
Canada	Toronto	Heatwave	1
South Africa	Johannesburg	Heatwave	1

12

```
--Q7 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_c), 1) as Average_Temperature
from climate_change
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;
```

average_temperature_c	total_population	total_economic_exposure
30.6	772529	98765751.15
26.9	4164128	96514968.05
-9.5	1310863	95231429.51
14.5	8499163	93366392.55
-6.5	1335844	91454336.88
35.5	5610204	90575436.02
27.3	9011115	85925007.67
34.9	8625068	72538730.19
38.6	9225533	70238861.05
7.9	1297995	68218445.17
-7.7	6883355	67123848.04
24.7	7910619	65892689.57
17.5	2414228	58917417.90
39.2	1910078	47481103.06
27.0	4723866	36593764.94
35.8	6315063	33073167.48
29.6	6513387	26753340.75

country	city	average_aqi	days_above_200_aqi	total_population_exposure	average_temperature_c
abc Filter...	abc Filter...	abc Filter...	abc Filter...	abc Filter...	abc Filter...
Canada	Vancouver	157	5	29518135	17.0
South Africa	Durban	158	5	23352543	17.0
South Africa	Johannesburg	173	5	25623821	16.8
USA	New York	199	5	38456628	21.9
USA	Chicago	200	5	23815304	13.3

Upto this point it has been seen that each and every factors has some effect on the climate metrics whether a less or more

13

--Q8 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_WeatherCount,
STRING_AGG(DISTINCT extreme_weather_events , ', ') as Event_Types,
Round(avg(temperature_c), 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
where date between '2025-03-03' and '2025-03-07'
group by biome_type;
```

biome_type	total_records	locations_affected	extreme_weathe...	event_types	average_temper...
a c Filter...	a c Filter...				
Desert	25	18	5	Drought, Heatwave, Hurricane, None	19.5
Forest	17	13	2	Flood, Hurricane, None	19.9
Grassland	19	14	2	Heatwave, Hurricane, None	22.0
Tundra	26	18	7	Drought, Flood, Heatwave, Hurricane, None	18.6
Wetland	18	14	2	Drought, None	16.1

After the last 8th query all the result are saved in CSV format for further visualization in tableau

 main_climate_change	8/5/20
 finding8	8/5/20
 finding7	8/5/20
 finding6	8/5/20
 finding5	8/5/20
 finding4	8/5/20
 finding3	8/5/20
 finding2	8/5/20
 finding1	8/5/20

Rest of the part to be continued on the next and the final tableau part

To Be
Continued...

THANK YOU!

PRESENTED BY ARIJEET MUKHERJEE
arijeetm622@gmail.com