# Deforestation Data Analysis Project Project Overview

The Deforestation Data Analysis project involves analyzing deforestation data using SQL to uncover insights related to various problem statements. The analysis focuses on understanding the extent of deforestation, income group distributions, and regional patterns.

## Problem Statements

Total Number of Countries Involved in Deforestation: Determine the number of countries that have experienced deforestation. Income Groups with Specific Land Area: Show the income groups of countries having a total area ranging from 75,000 to 150,000 square meters. Average Area by Income Group: Calculate the average area in square miles for countries in the 'upper middle income' region and compare it with other income categories. Forest Area by Income Group: Determine the total forest area in square kilometers for countries in the 'high income' group and compare it with other income categories. Top Forest Areas by Region: Show countries from each region (continent) with the highest total forest areas.

## SQL Analysis and Insights

1. Total Countries Involved in Deforestation:

```sql
-- Question 1:  What are the total number of countries involved in deforestation?


WITH deforesstationperiod AS (
    SELECT DISTINCT  country_name,
           forest_area_sqkm,
           YEAR,
           LAG(forest_area_sqkm) OVER(PARTITION BY country_name ORDER BY year) AS prev_forest_area,
           LEAD(forest_area_sqkm) OVER(PARTITION BY country_name ORDER BY year) AS next_forest_area
    FROM Forest_Area
)

SELECT DISTINCT COUNT(DISTINCT country_name)

FROM deforesstationperiod
WHERE forest_area_sqkm < prev_forest_area OR forest_area_sqkm < next_forest_area
```

| | (No column … |
|---|---|
| 1 | 175 |

2. Income Groups with Specific Land Area:

```sql
-- Question 2: Show the income groups of countries having total area ranging from 75,000 to 150,000 square meter? --

--Joining the Region and Land Area tables but only displaying the country, land area and income group for countries that fit the range --
-- The list shows a country every year its total area met the criteria

SELECT L.country_name, L.total_area_sq_mi,L.year, R.income_group
FROM Land_Area L FULL JOIN Region R on L.country_name = R.country_name
WHERE total_area_sq_mi BETWEEN 75000 AND 150000
ORDER BY total_area_sq_mi;

-- COUNTING THE TOTAL NUMBER OF COUNTRIES THAT FIT THE CRITERIA --

SELECT COUNT(DISTINCT country_name)
FROM Land_Area
WHERE total_area_sq_mi BETWEEN 75000 AND 150000


-- 26 countries practiced deforestattion --
```

3. Average Area by Income Group:

```sql
-- QUESTION 3 Calculate average area in square miles for countries in the 'upper middle income region'. Compare the result with the rest of the income categories.--

SELECT * FROM Region;
SELECT * FROM Land_Area;

SELECT R.income_group,ROUND(AVG(total_area_sq_mi), 0) AS 'AVG LAND AREA'
FROM Region R FULL JOIN Land_Area L ON R.country_name = L.country_name
GROUP BY R.income_group
HAVING income_group IN ('upper middle income','high income','low income','lower middle income')
ORDER BY 'AVG LAND AREA';
```

| | income_group | AVG LAND AREA |
|---|---|---|
| 1 | Low income | 165116 |
| 2 | Lower middle income | 168970 |
| 3 | High income | 184821 |
| 4 | Upper middle income | 389810 |

4. Forest Area by Income Group:

```sql
-- QUESTION 4 Determine the total forest area in square km for countries in the 'high income' group. Compare result with the rest of the income categories. --

SELECT * FROM Region;
SELECT * FROM Forest_Area;

SELECT R.income_group,ROUND(SUM(forest_area_sqkm), 0) AS 'TOTAL FOREST AREA'
FROM Region R FULL JOIN Forest_Area F ON R.country_name = F.country_name
GROUP BY R.income_group
HAVING income_group IN ('upper middle income','high income','low income','lower middle income')
ORDER BY 'TOTAL FOREST AREA';
```

| | income_group | TOTAL FOREST AREA |
|---|---|---|
| 1 | Low income | 104641151 |
| 2 | Lower middle income | 160235005 |
| 3 | High income | 280006180 |
| 4 | Upper middle income | 537371318 |

5. Top Forest Areas by Region:

```sql
-- QUESTION 5 Show countries from each region(continent) having the highest total forest areas.

SELECT * FROM Region;
SELECT * FROM Forest_Area;



WITH Rankedcountries AS (
    SELECT R.country_name,
           R.region,
           F.forest_area_sqkm,
           ROW_NUMBER() OVER(PARTITION BY R.region ORDER BY F.forest_area_sqkm DESC) AS RANK
    FROM Forest_Area F
    FULL JOIN Region R on F.country_name = R.country_name
    WHERE Region IS NOT NULL
)

SELECT country_name,
       region,
       forest_area_sqkm
FROM Rankedcountries
WHERE RANK = 1
ORDER BY forest_area_sqkm DESC;
```

| | country_name | region | forest_area_sqkm |
|---|---|---|---|
| 1 | Russian Federation | Europe & Central Asia | 8151356 |
| 2 | Brazil | Latin America & Caribbean | 5467050 |
| 3 | Canada | North America | 3482730 |
| 4 | China | East Asia & Pacific | 2098635 |
| 5 | Congo, Dem. Rep. | Sub-Saharan Africa | 1603630 |
| 6 | India | South Asia | 708604 |
| 7 | Iran, Islamic Rep. | Middle East & North Africa | 106920 |