GLOBAL DEFORESTATION

ForestQuery

Abstract

An Exploration of the Global Deforestation Trends from 1990 to 2016. This project was done using SQL connecting multiple tables to one another with data obtained from the World Bank. This report contains findings and insights that brings to light areas of concern as well as opportunities to address the rapid deforestation rate globally.

Nsikan Udoma

nsikan.e.udoma@gmail.com

Masterschool

INTRODUCTION

ForestQuery is on a mission to combat deforestation around the world and to raise awareness about this topic and its impact on the environment. The data analysis team at ForestQuery has obtained data from the World Bank that includes forest area and total land area by country and year from 1990 to 2016, as well as a table of countries and the regions to which they belong.

My team of Data Analysts have used SQL to bring these tables together and to query them in an effort to find areas of concern as well as areas that present an opportunity to learn from the successes of some countries who have made progress in combating the rapid rate of deforestation globally.

1. **GLOBAL SITUATION**

According to the World Bank, the total forest area of the world was 41,282,694.9 sq.km in 1990. As of 2016, the most recent year for which data was available, that number had fallen to 39,958,245.9 sq.km, a loss of 1,324,449 sq.km, or 3.2%.

The forest area lost over this time period is slightly more than the entire land area of Peru listed for the year 2016 (which is 1,279,999.99 sq.km).

2. **REGIONAL OUTLOOK**

In 2016, the percent of the total land area of the world designated as forest was 31.38%. The region with the highest relative forestation was Latin America & Caribbean, with 46.16%, and the region with the lowest relative forestation was Middle East & North Africa, with 2.07% forestation.

In 1990, the percent of the total land area of the world designated as forest was 32.42%. The region with the highest relative forestation was Latin America & Caribbean, with 51.03%, and the region with the lowest relative forestation was Middle East & North Africa, with 1.78% forestation.

Table 2.1: Percent Forest Area by Region, 1990 & 2016:

Region	1990 Forest Percentage (%)	2016 Forest Percentage (%)	2016 - 1990 values % increase/decrease
Latin America & Caribbean	51.03	46.16	- 4.87 decrease
Europe & Central Asia	37.28	38.04	+0.76 increase
North America	35.65	36.04	+0.39 increase
Sub-Saharan Africa	30.67	28.79	-1.88 decrease
East Asia & Pacific	25.78	26.36	+0.58 increase
South Asia	16.51	17.51	+1.00 increase
Middle East & North Africa	1.78	2.07	+0.29 increase

After analyzing Table 2.1 above, it was observed that the only regions of the world that decreased in percent forest area from 1990 to 2016 were Latin America & Caribbean (dropped from 51.03% to 46.16%) and Sub-Saharan Africa (30.67% to 28.79%). All other regions actually increased in forest area over this time period. However, the drop in forest area in the two aforementioned regions was so large, the percent forest area of the world decreased over this time period from 32.42% to 31.38%.

3. **COUNTRY-LEVEL DETAIL**

A. SUCCESS STORIES

There is one particularly bright spot in the data at the country level, China. This country actually increased in forest area from 1990 to 2016 by 527,229.06 sq.km. It would be interesting to study what has changed in this country over this time to drive this figure in the data higher. The country with the next largest increase in forest area from 1990 to 2016 was the United States, but it only saw an increase of 79,200.00 sq.km, much lower than the figure for China.

China and the United States are of course very large countries in total land area, so when we look at the largest *percent* change in forest area from 1990 to 2016, we aren't surprised to find a much smaller country listed at the top. Iceland increased in forest area by 213.66% from 1990 to 2016.

B. LARGEST CONCERNS

Which countries are seeing deforestation to the largest degree? We can answer this question in two ways. First, we can look at the absolute square kilometer decrease in forest area from 1990 to 2016. The following 5 countries had the largest decrease in forest area over the time period under consideration:

Table 3.1: Top 5 Amount Decrease in Forest Area by Country, 1990 & 2016:

Country	Region	Absolute Forest Area Change (sq.km)
Brazil	Latin America & Caribbean	541,510.00
Indonesia	East Asia & Pacific	282,193.98
Myanmar	East Asia & Pacific	107,234.00
Nigeria	Sub-Saharan Africa	106,506.00
Tanzania	Sub-Saharan Africa	102,320.00

The second way to consider which countries are of concern is to analyze the data by percent decrease.

Table 3.2: Top 5 Percent Decrease in Forest Area by Country, 1990 & 2016:

Country	Region	Pct Forest Area Change (%)
Togo	Sub-Saharan Africa	-75.45
Nigeria	Sub-Saharan Africa	-61.80
Uganda	Sub-Saharan Africa	-59.13
Mauritania	Sub-Saharan Africa	-46.75
Honduras	Latin America & Caribbean	-45.03

When we consider countries that decreased in forest area percentage the most between 1990 and 2016, we find that four of the top 5 countries on the list are in the region of Sub-Saharan Africa. The countries are Togo, Nigeria, Uganda, and Mauritania. The 5th country on the list is Honduras, which is in the Latin America & Caribbean region.

From the above analysis, we see that Nigeria is the only country that ranks in the top 5 both in terms of absolute square kilometer decrease in forest as well as percent decrease in forest area

from 1990 to 2016. Therefore, this country has a significant opportunity ahead to stop the decline and hopefully spearhead remedial efforts.

C. QUARTILES

Table 3.3: Count of Countries Grouped by Forestation Percent Quartiles, 2016:

Quartile	Number of Countries
1st quartile = 0% - 24.99%	85
2nd quartile = 25% - 49.99%	73
3rd quartile = 50% - 74.99%	38
4th quartile = 75% - 100%	9

The largest number of countries in 2016 were found in the 1st quartile.

There were 9 countries in the top quartile in 2016. These are countries with a very high percentage of their land area designated as forest. The following is a list of countries and their respective forest land, denoted as a percentage.

Table 3.4: Top Quartile Countries, 2016:

Country	Region	Pct Designated as Forest (%)
Suriname	Latin America & Caribbean	98.26
Micronesia, Fed. Sts.	East Asia & Pacific	91.86
Gabon	Sub-Saharan Africa	90.04
Seychelles	Sub-Saharan Africa	88.41
Palau	East Asia & Pacific	87.61
American Samoa	East Asia & Pacific	87.50
Guyana	Latin America & Caribbean	83.90
Lao PDR	East Asia & Pacific	82.11
Solomon Islands	East Asia & Pacific	77.86

4. RECOMMENDATIONS

From the world bank data, I gathered that Brazil, Indonesia, Myanmar, Nigeria, and Tanzania should be the top countries to focus on due to their alarming rate of deforestation. These 5 countries have seen the largest decrease in forest areas from 1990 to 2016. And because of the considerable large total land areas and how populated most of the aforementioned countries are (Indonesia, Nigeria and Brazil), it would be interesting to investigate further how the population of these countries play a role in their subsequent large depleting forest areas. There just might be a correlation between population growth, population density and deforestation activities.

Uganda, Mauritania, and Honduras are also countries that require urgent intervention to reduce the rapid rate of deforestation occurring. Considering that these 3 countries are not as large as the top 5 countries with decreasing forest area, it would be interesting to find out the causes of deforestation in them. Could it be due to natural disasters such as flooding, desertification, erosion etc. Or could it be majorly man-made activities such as cutting down of trees for constructions (buildings, roads etc.) and creation of wood products; or anthropogenic activities such as mining, drilling & fossil fuel exploration, oil spills or the release industrial waste into swamp forests or water bodies, and the use of pesticides, herbicides and fertilizers that destroy crops and natural life. Or could it be a combination of both natural and man-made activities?

I recommend that the world bank should create a database with data from natural and man-made activities that results in deforestation especially in countries with large depleting forest areas, in order for data analysts worldwide to seek out root causes of deforestation globally.

5. APPENDIX - My Queries

First thing I normally do is check the datatype of all my columns in each table in my database. This helps me to know what aggregations and arithmetic operations I can perform in each column. Or if I need to convert/cast any columns to a different datatype in order to perform arithmetic operations or any kind of aggregations.

Example, using the deforestation project database:

```
SELECT COLUMN_NAME, DATA_TYPE
FROM INFORMATION_SCHEMA.COLUMNS
WHERE TABLE_NAME = 'land_area'
```

a) Global Situation:

Firstly, I created a view called “forestation” by joining all three tables - forest_area, land_area and regions in the workspace.

```
DROP VIEW IF EXISTS forestation;
CREATE VIEW forestation AS (
	SELECT l.country_code AS country_code,
		l.country_name AS country_name,
		l.year AS year,
		round(f.forest_area_sqkm::numeric, 2) AS forest_area_sqkm,
		round((total_area_sq_mi*2.59)::numeric, 2) AS total_land_area_sqkm,
		r.region AS region,
		r.income_group AS income_group
	FROM forest_area AS f
	JOIN land_area AS l
		ON l."year" = f."year" AND l.country_code = f.country_code
	JOIN regions AS r
		ON r.country_code = l.country_code AND r.country_code = f.country_code
	);
```

To answer the first question in the global situation section, I used the query below:

```
SELECT
	MAX(forest_area_sqkm) AS first_recorded_f_area,
	MIN(forest_area_sqkm) AS last_recorded_f_area,
	(MAX(forest_area_sqkm) - MIN(forest_area_sqkm)) AS forest_area_loss_sqkm,
	ROUND((((MAX(forest_area_sqkm) -
MIN(forest_area_sqkm))/MAX(forest_area_sqkm))*100)::numeric,2) AS percent_loss
FROM forest_area
WHERE country_name = 'World';
```

OR the code below also does the same thing

```
WITH first_forest_area_sqkm AS (
       SELECT country_name, year, forest_area_sqkm
       FROM forest_area
       WHERE country_name = 'World' AND year = '1990'
       ),

     last_forest_area_sqkm AS (
       SELECT country_name, year, forest_area_sqkm
       FROM forest_area
       WHERE country_name = 'World' AND year = '2016'
       )

SELECT
       f.forest_area_sqkm AS first_forest_area_sqkm,
       l.forest_area_sqkm AS last_forest_area_sqkm,
       (f.forest_area_sqkm - l.forest_area_sqkm) AS forest_area_loss_sqkm,
       ((f.forest_area_sqkm - l.forest_area_sqkm)/f.forest_area_sqkm)*100 AS
percent_loss
FROM first_forest_area_sqkm as f
JOIN last_forest_area_sqkm as l
       ON f.country_name = l.country_name
;
```

Now, I went a step further to create another view called "global_situation" so I can always reference my aggregations in that section whenever I want to.

```
DROP VIEW IF EXISTS global_situation;
CREATE VIEW global_situation AS (
       SELECT
              MAX(forest_area_sqkm) AS first_recorded_f_area,
              MIN(forest_area_sqkm) AS last_recorded_f_area,
              (MAX(forest_area_sqkm) - MIN(forest_area_sqkm)) AS forest_area_loss_sqkm,
              ROUND((((MAX(forest_area_sqkm) -
              MIN(forest_area_sqkm))/MAX(forest_area_sqkm))*100)::numeric,2) AS
              percent_loss
       FROM forest_area
       WHERE country_name = 'World'
       );
```

The forest area lost over this time period is slightly more than the entire land area of _____ listed for the year 2016 (which is _____). See query below for answer

```
SELECT *
FROM forestation
WHERE year = 2016 AND total_land_area_sqkm <= (
       SELECT (MAX(forest_area_sqkm) - MIN(forest_area_sqkm)) AS forest_area_loss_sqkm
```

```
       FROM forest_area
       WHERE country_name = 'World'
       )
ORDER BY total_land_area_sqkm DESC
LIMIT 1;
```

b) REGIONAL OUTLOOK

Queries for 2016 forest area and relative forestation:

```
SELECT country_code,
       country_name,
       Year,
       Forest_area_sqkm,
       total_land_area_sqkm,
       ROUND(((forest_area_sqkm/total_land_area_sqkm)*100)::numeric, 2) AS
percent_land_that_is_forest
FROM forestation
WHERE year = 2016 AND country_name = 'World';
```

```
SELECT region,
       SUM(forest_area_sqkm) AS total_forest_area,
       SUM(total_land_area_sqkm) AS total_land_area,
       ROUND(((SUM(forest_area_sqkm)/SUM(total_land_area_sqkm))*100)::numeric, 2) AS
percent_land_that_is_forest
FROM forestation
WHERE year = 2016
GROUP BY region
ORDER BY 4 DESC;
```

Queries for 1990 forest area and relative forestation:

```
SELECT country_code,
       country_name,
       Year,
       Forest_area_sqkm,
       total_land_area_sqkm,
       ROUND(((forest_area_sqkm/total_land_area_sqkm)*100)::numeric, 2) AS
percent_land_that_is_forest
FROM forestation
WHERE year = 1990 AND country_name = 'World';
```

```
SELECT region,
       SUM(forest_area_sqkm) AS total_forest_area,
       SUM(total_land_area_sqkm) AS total_land_area,
```

```
      ROUND(((SUM(forest_area_sqkm)/SUM(total_land_area_sqkm))*100)::numeric, 2) AS
percent_land_that_is_forest
FROM forestation
WHERE year = 1990
GROUP BY region
ORDER BY 4 DESC;
```

Again, I went a step further to create another view called "regional_outlook" so I can always reference my aggregations in that section whenever I want to.

```
DROP VIEW IF EXISTS regional_outlook;
CREATE VIEW regional_outlook AS (
      SELECT region,
             year,
             SUM(forest_area_sqkm) AS total_forest_area,
             SUM(total_land_area_sqkm) AS total_land_area,
             ROUND(((SUM(forest_area_sqkm)/SUM(total_land_area_sqkm))*100)::numeric,
      2) AS percent_land_that_is_forest
      FROM forestation
      WHERE year = 1990 OR year = 2016
      GROUP BY region, year
      ORDER BY 5 DESC
);
```

The "regional_outlook" view I created now comes in handy as I am able to obtain the data needed to populate Table 2.1 with the query below:

```
SELECT *
FROM regional_outlook
;
```

c) COUNTRY-LEVEL DETAIL

Success Stories queries:

```
WITH first_country_level_data AS (
      SELECT region,
             country_name,
             year,
             forest_area_sqkm,
             Total_land_area_sqkm
      FROM forestation
      WHERE year = '1990' AND country_name != 'World'
      ),
    last_country_level_data AS (
      SELECT region,
             country_name,
             year,
             forest_area_sqkm,
             total_land_area_sqkm
      FROM forestation
```

```
      WHERE year = '2016' AND country_name != 'World'
      )

SELECT f.region AS region,
      f.country_name AS country_name,
      f.forest_area_sqkm AS first_forest_area_sqkm,
      l.forest_area_sqkm AS last_forest_area_sqkm,
      (l.forest_area_sqkm - f.forest_area_sqkm) AS forest_area_gain_loss_sqkm
FROM first_country_level_data as f
JOIN last_country_level_data as l
      ON f.country_name = l.country_name
WHERE (l.forest_area_sqkm - f.forest_area_sqkm) IS NOT NULL
ORDER BY forest_area_gain_loss_sqkm DESC ;
```

Seeing the need to reference country level data, I created a view called "country_level_detail" for easy access to this data when needed.

```
DROP VIEW IF EXISTS country_level_detail;
CREATE VIEW country_level_detail AS(

 WITH first_country_level_data AS (
      SELECT region,
            country_name,
            year,
            forest_area_sqkm,
            total_land_area_sqkm
      FROM forestation
      WHERE year = '1990' AND country_name != 'World'
      ),
     last_country_level_data AS (
      SELECT region,
            country_name,
            year,
            forest_area_sqkm,
            total_land_area_sqkm
      FROM forestation
      WHERE year = '2016' AND country_name != 'World'
      )

 SELECT f.region,
      f.country_name AS country_name,
      f.forest_area_sqkm AS "1990_forest_area_sqkm",
      l.forest_area_sqkm AS "2016_forest_area_sqkm",
      ROUND((((l.forest_area_sqkm -
f.forest_area_sqkm)/f.forest_area_sqkm)*100)::numeric,2) AS percent_change,
      (l.forest_area_sqkm - f.forest_area_sqkm) AS forest_area_gain_loss_sqkm,
      f.total_land_area_sqkm AS "1990_total_land_area",
      l.total_land_area_sqkm AS "2016_total_land_area"
 FROM first_country_level_data as f
 JOIN last_country_level_data as l
       ON f.country_name = l.country_name
 WHERE (l.forest_area_sqkm - f.forest_area_sqkm) IS NOT NULL
 ORDER BY forest_area_gain_loss_sqkm DESC
```

```
);
```

After creating the view called "country_level_detail", I can start answering the questions with the query below:

```
SELECT *
FROM country_level_detail
ORDER BY percent_change DESC ;
```

LARGEST CONCERNS

We can answer the question in this section in two ways
To answer First, we can look at the absolute square kilometer decrease in forest area from 1990 to 2016. Here the view I created "country_level_detail" becomes handy.

```
SELECT *
FROM country_level_detail
ORDER BY forest_area_gain_loss_sqkm ASC
;
```

The second way to consider which countries are of concern is to analyze the data by percent decrease.

```
SELECT *
FROM country_level_detail
ORDER BY percent_change ASC
;
```

QUARTILES

```
WITH percent_forestation AS (
	SELECT region,
		country_name,
		year,
		forest_area_sqkm,
		total_land_area_sqkm,
		ROUND(((forest_area_sqkm/total_land_area_sqkm)*100)::numeric, 2) AS
percent_land_that_is_forest
	FROM forestation
	WHERE year = 2016 AND forest_area_sqkm IS NOT NULL AND total_land_area_sqkm IS
NOT NULL
	GROUP BY 1, 2, 3, 4, 5
	ORDER BY percent_land_that_is_forest DESC
)
SELECT region,
	Country_name,
	percent_land_that_is_forest,
	CASE WHEN percent_land_that_is_forest >= 75 THEN 4
```

```
            WHEN percent_land_that_is_forest < 75 AND percent_land_that_is_forest >=
50 THEN 3
            WHEN percent_land_that_is_forest < 50 AND percent_land_that_is_forest
>=25 THEN 2
      ELSE 1
      END AS percent_quartiles
FROM percent_forestation
```

From a glance at the questions in section C (Quartiles), there was the need to create a view that holds the data about countries and what group of forestation percent quartile they belong to.

Forestation Percent Quartiles :
1st quartile = 0% - 25% i.e countries with less than 25% of their total land area designated as forest area
2nd quartile = 25% - 50% i.e countries with between 25% and 49.99% of their total land area designated as forest area
3rd quartile = 50% - 75% i.e countries with between 50% and 74.99% of their total land area designated as forest area
4th quartile = 75% - 100% i.e countries with between 75% and 100% of their total land area designated as forest area

```
DROP VIEW IF EXISTS quartiles2016;
CREATE VIEW quartiles2016 AS (

  WITH percent_forestation AS (
      SELECT region,
            country_name,
            year,
            forest_area_sqkm,
            total_land_area_sqkm,
            ROUND(((forest_area_sqkm/total_land_area_sqkm)*100)::numeric, 2) AS
percent_land_that_is_forest
      FROM forestation
      WHERE year = 2016 AND forest_area_sqkm IS NOT NULL AND total_land_area_sqkm IS
NOT NULL
      GROUP BY 1, 2, 3, 4, 5
      ORDER BY percent_land_that_is_forest DESC
      )
  SELECT region,
      country_name,
      percent_land_that_is_forest,
      CASE WHEN percent_land_that_is_forest >= 75 THEN 4
            WHEN percent_land_that_is_forest < 75 AND percent_land_that_is_forest >=
50 THEN 3
            WHEN percent_land_that_is_forest < 50 AND percent_land_that_is_forest
>=25 THEN 2
      ELSE 1
      END AS percent_quartiles
  FROM percent_forestation
```

```
    )
```

After creating the view called "quartiles2016", I can now answer the questions without writing tedious queries every time. But first, let's populate Table 3.3 (Count of Countries Grouped by Forestation Percent Quartiles, 2016:)

```
SELECT percent_quartiles,
       COUNT(*) AS number_of_countries
FROM quartiles2016
GROUP BY percent_quartiles
ORDER BY percent_quartiles
```

Finally, to populate Table 3.4 (Top Quartile Countries, 2016:), I used the query below:

```
SELECT *
FROM quartiles2016
WHERE percent_quartiles = 4
```