Report for ForestQuery into Global Deforestation, 1990 to 2016

By Miroslava Chrienova

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.

The data analysis team has 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 successes.

1. **GLOBAL SITUATION**

According to the World Bank, the total forest area of the world was **41282694.9** sq km in 1990. As of 2016, the most recent year for which data was available, that number had fallen to **39958245.9** sq km, a loss of **1324449,** or **3.208**%.

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 ***1279999.99***).

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 & Carribean**, 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 & Carribean**, 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
Latin America & Caribbean	51.03	46.16
Europe & Central Asia	37.28	38.04
North America	35.65	36.04
Sub-Saharan Africa	30.67	28.79
East Asia & Pacific	25.78	26.36
South Asia	16.51	17.51
Middle East & North Africa	1.78	2.07

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 **527229.06**. 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 **79200.00 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 3 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 in sq km
Brazil	Latin America & Caribbean	541510
Indonesia	East Asia & Pacific	282193.98
Myanmar	East Asia & Pacific	107234.00
Nigeria	Sub-Saharan Africa	106506.00
Tanzania	Sub-Saharan Africa	102320

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
75%-100%	9
50%-75%	38
25%-50%	72
0-25%	85

The largest number of countries in 2016 were found in the **1s (0-25%)** 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	98.2576939676578
Micronesia, Fed. Sts.	East Asia & Pacific	91.8572390715248
Gabon	Sub-Saharan Africa	90.0376418700565
Seychelles	Sub-Saharan Africa	88.4111367385789
Palau	East Asia & Pacific	87.6068085491203
American Samoa	East Asia & Pacific	87.5000875000875
Guyana	Latin America & Caribbean	83.9014489110682
Lao PDR	East Asia & Pacific	82.1082317640861
Solomon Islands	East Asia & Pacific	77.8635177945066

94 countries had a percent forestation higher than the United States in 2016.

5. RECOMMENDATIONS

Write out a set of recommendations as an analyst on the ForestQuery team.

- *What have you learned from the World Bank data?*

  Forest is slowly disappearing. The data analysis shows reduction of the global forest between 1990 and 2016. The most impacted region is Sub-Saharan Africa. For example, Togo reduced the forest by 75.45%. Even if we look at the count of countries grouped by forestation percent, we can see that 85 countries are in the first quartile with forestation between 0 -25% and 72 countries belong to the second quartile with forestation at level between 25 and 50%.

- *Which countries should we focus on over others?*

  4 out of 5 countries with top percent decrease are in Sub-Saharan Africa. Togo, lost in the analysed period (1990 -2016) over 75% of the forest. Other countries we should focus on are Nigeria (61.80%), Uganda (59.13%) and Mauritania (46.75%). People need to understand the way of life is having unsustainable ecological footprint. There are several things which could be done in order to decrease this footprint. Firstly, consume less. Decrease on production of goods, can slow down deforestation trend. Secondly, avoid

products containing Palm oil which causes big deforestation mainly in Asia. Thirdly, shop sustainably certified products.

APPENDIX: SQL queries used

```
-- Creating View called forestation

CREATE VIEW forestation AS SELECT fa.country_code,
        fa.country_name,
        fa.year,
        fa.forest_area_sqkm,
        la.total_area_sq_mi,
        la.total_area_sq_mi * 2.59 AS total_area_sqkm,
        re.region,
        re.income_group,
        (fa.forest_area_sqkm * 100)/ (total_area_sq_mi * 2.59) AS percent_forestation
FROM forest_area fa
JOIN land_area la
    ON fa.country_code = la.country_code
        AND fa.year = la.year
JOIN regions re
    ON re.country_code = fa.country_code;
```

Global Situation

a. What was the total forest area (in sq km) of the world in 1990? Please keep in mind that you can use the country record denoted as “World" in the region table.

```
SELECT SUM(forest_area_sqkm)
FROM forestation
WHERE year = 1990
        AND region = 'World';
```

b. What was the total forest area (in sq km) of the world in 2016? Please keep in mind that you can use the country record in the table is denoted as "World."

```
SELECT SUM(forest_area_sqkm)
FROM forestation
WHERE year = 2016
      AND region = 'World';
```

c. What was the change (in sq km) in the forest area of the world from 1990 to 2016?

```
SELECT ( fa1.forest_area_sqkm - fa2.forest_area_sqkm ) AS forest_area_change
FROM forestation fa1, forestation fa2
WHERE fa1.year = 1990
      AND fa1.region = 'World'
      AND fa2.year = 2016
      AND fa2.region = 'World';
```

d. What was the percent change in the forest area of the world between 1990 and 2016?

```
SELECT ( fa1.forest_area_sqkm - fa2.forest_area_sqkm )* 100 / fa1.forest_area_sqkm AS
pct_change
FROM forestation fa1, forestation fa2
WHERE fa1.year = 1990
      AND fa1.region = 'World'
      AND fa2.year = 2016
      AND fa2.region = 'World';
```

e. If you compare the amount of forest area lost between 1990 and 2016, to which country's total area in 2016 is it closest to?

```
WITH tb1 AS
   (SELECT MAX(forest_area_sqkm) - MIN(forest_area_sqkm) AS deforest
   FROM forestation),

tb2 AS
   (SELECT *,
       total_area_sq_mi * 2.59 AS total_area_sq_km
   FROM land_area FULL
   JOIN tb1
      ON land_area.total_area_sq_mi = tb1.deforest),
```

```
tb3 AS
   (SELECT *,
       CASE
       WHEN deforest IS NULL THEN
       1324449
       ELSE NULL
       END AS new_deforest
   FROM tb2)

SELECT country_name,
       total_area_sq_km
FROM tb3
WHERE total_area_sq_km < new_deforest
       AND YEAR = 2016
ORDER BY  total_area_sq_km DESC LIMIT 1;
```

Regional Outlook

-- List of successful countries ordered by increase difference

```
WITH tb1 AS
   (SELECT region,
       country_name,
       forest_area_sqkm,
       total_area_sqkm
   FROM forestation
   WHERE year = 1990 ),

tb2 AS
   (SELECT region,
       country_name,
       forest_area_sqkm,
       total_area_sqkm
   FROM forestation
   WHERE year = 2016 )

SELECT tb1.region,
       tb1.country_name,
       tb1.forest_area_sqkm AS forest_1990,
       tb2.forest_area_sqkm AS forest_2016,
```

```
        ROUND( CAST( ( tb2.forest_area_sqkm - tb1.forest_area_sqkm ) AS numeric ), 2 ) AS
difference,
        ROUND( CAST( ( ( tb2.forest_area_sqkm - tb1.forest_area_sqkm )* 100 /
tb1.total_area_sqkm ) AS numeric ), 2 ) AS increase_percent
FROM tb1
JOIN tb2
   ON tb1.country_name = tb2.country_name
WHERE tb2.forest_area_sqkm > tb1.forest_area_sqkm
ORDER BY  difference DESC;
-- List of successful countries ordered by percent increase

WITH tb1 AS
   (SELECT region,
        country_name,
        forest_area_sqkm,
        total_area_sqkm
   FROM forestation
   WHERE year = 1990 ),

tb2 AS
   (SELECT region,
        country_name,
        forest_area_sqkm,
        total_area_sqkm
   FROM forestation
   WHERE year = 2016 )

SELECT tb1.region,
        tb1.country_name,
        tb1.forest_area_sqkm AS forest_1990,
        tb2.forest_area_sqkm AS forest_2016,
        ROUND( CAST( ( tb2.forest_area_sqkm - tb1.forest_area_sqkm ) AS numeric ), 2 ) AS
difference,
        ROUND( CAST( ( ( ( tb2.forest_area_sqkm - tb1.forest_area_sqkm )/
tb1.forest_area_sqkm )* 100 ) AS numeric ), 2 ) AS increase_percent
FROM tb1
JOIN tb2
   ON tb1.country_name = tb2.country_name
WHERE tb2.forest_area_sqkm > tb1.forest_area_sqkm
ORDER BY  increase_percent DESC;
```

a. What was the percent forest of the entire world in 2016? Which region had the HIGHEST percent forest in 2016, and which had the LOWEST, to 2 decimal places?

-- Percentage of forest of the entire world in 2016

```
SELECT forest_area_sqkm * 100 / total_area_sqkm
FROM forestation
WHERE year = 2016
      AND country_name = 'World';
```

-- Region with the HIGHEST percent of forest in 2016 (rounded to 2 decimal place)

```
SELECT region,
      ROUND(CAST(percent_forest AS numeric), 2)
 FROM
   (SELECT region,
      SUM(forest_area_sqkm)*100/SUM(total_area_sqkm) AS percent_forest
   FROM forestation
   WHERE year = 2016
   GROUP BY  1 ) sub
ORDER BY  2 DESC LIMIT 1;
```

-- Region with the LOWEST percent of forest in 2016 (rounded to 2 decimal place)

```
SELECT region,
      ROUND(CAST(percent_forest AS numeric), 2)
FROM
   (SELECT region,
      SUM(forest_area_sqkm)*100/SUM(total_area_sqkm) AS percent_forest
   FROM forestation
   WHERE year = 2016
   GROUP BY  1 ) sub
ORDER BY  2 LIMIT 1;
```

b. What was the percent forest of the entire world in 1990? Which region had the HIGHEST percent forest in 1990, and which had the LOWEST, to 2 decimal places?

-- Percentage of forest of the entire world in 1990

```
SELECT forest_area_sqkm * 100/ total_area_sqkm AS pct_forest
FROM forestation
```

```
WHERE year = 1990
      AND country_name = 'World' ;
```

-- Region with the HIGHEST percent forest in 1990 (rounded to 2 decimal place)

```
SELECT region,
       ROUND(CAST(percent_forest AS numeric), 2)
FROM
   (SELECT region,
       SUM(forest_area_sqkm)*100/SUM(total_area_sqkm) AS percent_forest
   FROM forestation
   WHERE year = 1990
         AND region NOT LIKE 'World'
   GROUP BY  region) sub
GROUP BY  1, 2
ORDER BY  2 DESC LIMIT 1;
```

-- Region that had the LOWEST percent forest in 1990 (rounded to 2 decimal place)

```
SELECT region,
       ROUND(CAST(percent_forest AS numeric), 2)
FROM
   (SELECT region,
       SUM(forest_area_sqkm)*100/SUM(total_area_sqkm) AS percent_forest
   FROM forestation
   WHERE year = 1990
         AND region NOT LIKE 'World'
   GROUP BY  1 ) sub
ORDER BY  2 ASC LIMIT 1;
```

c. Based on the table you created, which regions of the world DECREASED in forest area from 1990 to 2016?

```
WITH tb1 AS
   (SELECT region,
       SUM(forest_area_sqkm) AS forest_sum_1990
   FROM forestation
   WHERE year = 1990
```

```
        AND region NOT LIKE 'World'
    GROUP BY  1),

tb2 AS
    (SELECT region,
        SUM(forest_area_sqkm) AS forest_sum_2016
    FROM forestation
    WHERE year = 2016
        AND region NOT LIKE 'World'
    GROUP BY  1)

SELECT tb1.region,
        tb1.forest_sum_1990,
        tb2.forest_sum_2016
FROM tb1
JOIN tb2
    ON tb1.region = tb2.region
WHERE tb2.forest_sum_2016 < tb1.forest_sum_1990;
```

Country-Level Data

```
Part Success stories - Countries increasing forest area
WITH tb1 AS(
SELECT region, country_name, forest_area_sqkm
FROM forestation
WHERE year = 1990),

tb2 AS(
SELECT region, country_name, forest_area_sqkm
FROM forestation
WHERE year = 2016)

SELECT tb1.region, tb1.country_name, tb1.forest_area_sqkm AS forest_1990,
tb2.forest_area_sqkm AS forest_2016, ROUND(CAST((tb2.forest_area_sqkm -
tb1.forest_area_sqkm) AS numeric), 2) AS difference, ROUND(CAST(((tb2.forest_area_sqkm -
tb1.forest_area_sqkm)*100/tb1.forest_area_sqkm) AS numeric), 2) AS increase_percent
FROM tb1
JOIN tb2
ON tb1.country_name = tb2.country_name
WHERE tb2.forest_area_sqkm > tb1.forest_area_sqkm
ORDER BY increase_percent DESC;
```

a. Which 5 countries saw the largest amount decrease IN forest area FROM 1990 to 2016? What was the difference IN forest area for each?

```
WITH tb1 AS
   (SELECT region,
       country_name,
       forest_area_sqkm
   FROM forestation
   WHERE year = 1990),

tb2 AS
   (SELECT region,
       country_name,
       forest_area_sqkm
   FROM forestation
   WHERE year = 2016)

SELECT tb1.region,
       tb1.country_name,
       tb1.forest_area_sqkm AS forest_1990,
       tb2.forest_area_sqkm AS forest_2016,
       ROUND(CAST((tb1.forest_area_sqkm - tb2.forest_area_sqkm) AS numeric),
       2) AS difference
FROM tb1
JOIN tb2
   ON tb1.country_name = tb2.country_name
WHERE tb2.forest_area_sqkm < tb1.forest_area_sqkm
         AND tb1.region NOT LIKE 'World'
ORDER BY  difference DESC
LIMIT 5;
```

***b.* Which 5 countries saw the largest percent decrease IN forest area FROM 1990 to 2016? What was the percent change to 2 decimal places for each?**

```
WITH tb1 AS
   (SELECT region,
       country_name,
       forest_area_sqkm
   FROM forestation
   WHERE year = 1990),

tb2 AS
   (SELECT region,
```

```
        country_name,
        forest_area_sqkm
    FROM forestation
    WHERE year = 2016)

SELECT tb1.region,
        tb1.country_name,
        tb1.forest_area_sqkm AS forest_1990,
        tb2.forest_area_sqkm AS forest_2016,
        ROUND(CAST((tb1.forest_area_sqkm - tb2.forest_area_sqkm) AS numeric), 2) AS
difference,
        ROUND(CAST(((tb1.forest_area_sqkm -
tb2.forest_area_sqkm)*100/tb1.forest_area_sqkm) AS numeric), 2) AS decrease_percent
FROM tb1
JOIN tb2
    ON tb1.country_name = tb2.country_name
WHERE tb2.forest_area_sqkm < tb1.forest_area_sqkm
ORDER BY  decrease_percent DESC
LIMIT 5;
```

c. If countries were grouped by percent forestation IN quartiles, which group had the most countries IN it IN 2016?

```
WITH tb1 AS
    (SELECT *
    FROM forestation
    WHERE year = 2016
          AND region NOT LIKE 'World'
          AND percent_forestation IS NOT NULL), tb2 AS
    (SELECT *,
        CASE
        WHEN percent_forestation > 75 THEN
        'Fourth'
        WHEN percent_forestation <= 75
            AND percent_forestation > 50 THEN
        'Third'
        WHEN percent_forestation <= 50
            AND percent_forestation >25 THEN
        'Second'
        ELSE 'First'
        END AS quartiles
    FROM tb1)
```

```
SELECT quartiles,
        COUNT(*) AS quartiles_groups
FROM tb2
GROUP BY  1;
```

d. List ALL of the countries that were IN the 4th quartile (percent forest > 75%) IN 2016.

```
SELECT distinct(quartiles),
        COUNT(country_name)
   OVER (PARTITION BY quartiles)
FROM
   (SELECT country_name,
      CASE
      WHEN percent_forestation<=25 THEN
      '0-25%'
      WHEN percent_forestation<=50
         AND percent_forestation>25 THEN
      '25%-50%'
      WHEN percent_forestation<=75
         AND percent_forestation>50 THEN
      '50%-75%'
      ELSE '75%-100%'
      END AS quartiles
   FROM forestation
   WHERE percent_forestation IS NOT NULL
         AND year=2016) sub;


--List of Top Quartile Countries

SELECT country_name,
        region,
        percent_forestation
FROM forestation
WHERE percent_forestation>75
       AND percent_forestation IS NOT NULL
       AND year=2016
ORDER BY  3 DESC ;
```

e. How many countries had a percent forestation higher than the United States IN 2016?

```
SELECT COUNT(country_name)
FROM forestation
WHERE year = 2016
      AND percent_forestation >
   (SELECT percent_forestation
   FROM forestation
   WHERE country_name = 'United States'
         AND year = 2016);
```