

Report for ForestQuery into Global Deforestation, 1990 to 2016

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 **41,282,694.9 sqkm** in 1990. As of 2016, the most recent year for which data was available, that number had fallen to **39,958,245.9 sqkm**, a loss of **1,324,449 sqkm**, or **3.21%**.

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 **1,279,999.9891 sqkm**)

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 American & 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 American & 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 |
|---------------------------|------------------------|------------------------|
| Latin America & Caribbean | 51.03 | 46.16 |
| Sub-Saharan Africa | 30.67 | 28.79 |
| World | 32.42 | 31.38 |

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 sqkm**. 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 sqkm**, much lower than the figure for **China**.

China and 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 |
|---------|--------|-----------------------------|
|---------|--------|-----------------------------|

| | | |
|-----------|----------------------------|-----------|
| Brazil | Latin American & Caribbean | 541510.00 |
| Indonesia | East Asia & Pacific | 282193.98 |
| Myanmar | East Asia & Pacific | 107234.00 |
| Nigeria | Sub-Saharan Africa | 106506.00 |
| Tanzania | Sub-Saharan Africa | 102320.00 |

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 American & 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 American & 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 |
|----------|---------------------|
| 0-25% | 85 |
| 25-50% | 72 |
| 50-75% | 38 |
| 75-100% | 9 |

The largest number of countries in 2016 were found in the **1st (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 | 98.26 |
| Micronesia, Fed. Sts. | East Asia & Pacific | 91.86 |
| Gabon | Sub-Saharan Africa | 90.04 |

4. RECOMMENDATIONS

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

- *What have you learned from the World Bank data?*
 - *Which countries should we focus on over others?*
- a. The World forest area has seen a 3.21% decrease between 1990 and 2016, i.e. lost an area slightly larger than Peru. Where the highest decrease in forest percentage was seen in the Regions of Latin America & Caribbean and Sub-Saharan Africa with a 4.87% and 1.88% decrease respectively.
 - b. From our data, Countries like Nigeria, Togo and Uganda should be a major focus, especially Nigeria losing 61.8% (106,506 sqkm) of forest area. This could be due to building infrastructure, however it would be beneficial to create awareness and learn from countries such as China, United States and Iceland who saw an increase in forest area.

- c. Overall, there were about 85 countries with less than 25% increase in forest area, thus we can also suggest awareness campaigns and means to combat deforestation in those locations, while we use countries like Suriname, Micronesia, Fed. Sts & Gabon which saw greater than 75% increase as case studies to adopt the methods used to combat deforestation.

5. APPENDIX: SQL Queries Used

== Create a view aliased forestation where we have our joint tables for convenience ==

```
DROP VIEW IF EXISTS forestation;
CREATE VIEW forestation as
SELECT f.country_code, f.country_name, f.year,
       f.forest_area_sqkm,
       l.total_area_sq_mi*2.59 as total_area_sqkm,
       ((forest_area_sqkm/(total_area_sq_mi*2.59))*100) as forest_pct, r.region, r.income_group
FROM forest_area f
JOIN land_area l
ON f.country_code=l.country_code
AND f.year=l.year
JOIN regions r
ON f.country_code=r.country_code;
```

== A query written to find the total forest area for 1990 and World. Using 'SUM' and WHERE to sort ==

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

== ran a query to get the same criteria but with '2016' as the search year. ==

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

**== Find the change in forest area between 1990 to 2016.
To do this we create temp tables. Then run a final query to calculate. ==**

```
WITH for_2016 as (
    SELECT SUM(forest_area_sqkm) as y_2016
    FROM forestation
    WHERE year = 2016 AND country_name = 'World'),
for_1990 as (
    SELECT SUM(forest_area_sqkm) as y_1990
    FROM forestation
    WHERE year = 1990 AND country_name = 'World'),
diff as
```

```
(
  SELECT y_2016, y_1990, y_2016-y_1990 as a_diff
  FROM for_2016, for_1990)
```

```
SELECT a_diff
FROM diff;
```

== Using the same method as above, the % change in forest area between 1990 and 2016 was found. Using ROUND(expression::numeric,2) to force the characters to numeric values ==

```
WITH forpct_2016 as (
  SELECT ROUND(SUM(forest_pct)::numeric,2) as p_2016
  FROM forestation
  WHERE year = 2016 AND country_name = 'World'),
forpct_1990 as (
  SELECT ROUND(SUM(forest_pct)::numeric,2) as p_1990
  FROM forestation
  WHERE year = 1990 AND country_name = 'World'),
pctdiff as (
  SELECT p_2016, p_1990, ((p_2016-p_1990)*100/p_1990) as p_diff
  FROM forpct_2016, forpct_1990)
```

```
SELECT p_2016, p_1990, p_diff
FROM pctdiff;
```

== Compare the forest area lost between 1990 and 2016 to a country's area that could be closest to the area in the year 2016. ==

```
SELECT country_name, total_area_sqkm
FROM forestation
WHERE total_area_sqkm BETWEEN 1200000 AND 1350000
AND year = 2016
ORDER BY 2 DESC;
```

=====

== Find the total forest area % in 2016, which region had the largest % and lowest %. ==

```
SELECT region,
  ROUND(((SUM(forest_area_sqkm)/SUM(total_area_sqkm))*100)::numeric,2) as
for_perc_2016
FROM forestation
WHERE year = 2016
GROUP BY 1
ORDER BY 2 DESC;
```

```
SELECT region, ROUND((SUM(forest_area_sqkm)/SUM(total_area_sqkm)*100)::numeric,2)
FROM forestation
WHERE year = 1990
```

```
GROUP BY 1
ORDER BY 2 DESC;
```

== Time to find the regions that had % changes between 1990 and 2016.
So, here we need to use a join to join our above tables to get the values and then subtract. ==

```
WITH p_2016 as (
    SELECT region,
           ROUND(((SUM(forest_area_sqkm)/SUM(total_area_sqkm))*100)::numeric,2) as
           pct_2016
    FROM forestation
    WHERE year = 2016
    GROUP BY 1
    ORDER BY 2 DESC),
p_1990 as (
    SELECT region,
           ROUND((SUM(forest_area_sqkm)/SUM(total_area_sqkm)*100)::numeric,2) as
           pct_1990
    FROM forestation
    WHERE year = 1990
    GROUP BY 1
    ORDER BY 2 DESC),
j_pct as(
    SELECT p.region, pct_2016, pct_1990, pct_2016-pct_1990 as diff
    FROM p_2016 p
    JOIN p_1990 pp
    ON p.region=pp.region)

SELECT *
FROM j_pct
ORDER BY 4
```

=====

== Now we want to find the 5 countries that saw the largest amt decrease in forest area between 1990 and 2016 ==
== % of forest area decrease is the same just change the final query to order by fp_change ==

```
WITH f2016 as (
    SELECT region, country_name, SUM(forest_area_sqkm) as for_2016
    FROM forestation
    WHERE year= 2016
    GROUP BY 1,2),

f1990 as (
    SELECT region, country_name, SUM(forest_area_sqkm) as for_1990
    FROM forestation
    WHERE year = 1990
    GROUP BY 1,2),
```

```

joint as (
    SELECT f2016.region, f2016.country_name, for_2016, for_1990,
           ROUND((for_2016-for_1990)::numeric,2) as area_diff,
           ROUND(((for_2016-for_1990)*100/for_1990)::numeric,2) as fp_change
    FROM f2016
    JOIN f1990
    ON f2016.country_name=f1990)

```

== Running a query to find the decrease in area sqkm ==

```

SELECT *
FROM joint
WHERE for_2016 IS NOT NULL
AND for_1990 IS NOT NULL
ORDER BY 5
LIMIT 5;

```

== Running a query to find the decrease in forest % ==

```

SELECT *
FROM joint
WHERE for_2016 IS NOT NULL
AND for_1990 IS NOT NULL
ORDER BY 6
LIMIT 6;

```

***** Excluded NULL values because the results were not matching up.**

**== now we want to find the quartile group w/ highest countries using forest_perc.
Here we need to use our forest percentage from the forestation table and a case
statement to group the countries into quartiles.
After that, we then do a final query to get the count and group by quartiles.==**

```

WITH t1 as (
    SELECT country_name, region, forest_pct,
           CASE
            WHEN forest_pct >= 75 THEN '75-100%'
            WHEN forest_pct >= 50 THEN '50-75%'
            WHEN forest_pct >= 25 THEN '25-50%'
            ELSE '0-25%'
           END as quartiles
    FROM forestation
    WHERE year= 2016
    AND forest_pct IS NOT NULL
    AND country_name != 'World'
    GROUP BY quartiles, country_name, region, forest_pct
    ORDER BY quartiles)

```

== Find the number of countries by quartiles in 2016 ==

```

SELECT count(*), quartiles
FROM t1
GROUP BY quartiles

```


ORDER BY quartiles

== Listing all the countries in the 4th quartile ==

```
SELECT country_name, region, quartiles, forest_pct
FROM t1
WHERE quartiles = '75-100%'
ORDER BY 4 DESC
```

== Now we find the number of countries w/ forest_pct > US in year 2016 ==

```
WITH t2 as (
    SELECT forest_pct, country_name
    FROM forestation
    WHERE forest_pct > (
        SELECT forest_pct
        FROM forestation
        WHERE country_name = 'United States'
        AND year = 2016)
    AND year = 2016
    AND country_name != 'World')

SELECT COUNT(*)
FROM t2
```