

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 in **41282694.9 km²** in 1990. As of 2016, the most recent year for which data was available, that number had fallen to **39958245.9 km²**, a loss of **1324449 km²**, 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 **511370.27 km²**).

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
Latin America & Caribbean	51.03	46.16
Europe & Central Asia	37.28	38.04
North America	35.65	36.04
World	32.42	31.38
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 that 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 km²**. 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 **United States**, but it only saw an increase of **79200 km²**, 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 in km²
Brazil	Latin America & 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 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
1	85
2	72
3	38
4	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
American Samoa	East Asia & Pacific	87.50
Gabon	Sub-Saharan Africa	90.04
Guyana	Latin America & Caribbean	83.90
Lao PDR	East Asia & Pacific	82.11
Micronesia, Fed. Sts.	East Asia & Pacific	91.86
Palau	East Asia & Pacific	87.61
Seychelles	Sub-Saharan Africa	88.41
Solomon Islands	East Asia & Pacific	77.86
Suriname	Latin America & Caribbean	98.26

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?*

5. APPENDIX: SQL Queries Used

Joining all tables to Create 'Forestation' as Project Requirements

```
CREATE VIEW forestation AS
SELECT
    f_area.country_code AS fa_country_code,
    f_area.country_name AS fa_country_name,
    f_area.year,
    f_area.forest_area_sqkm,
    l_area.country_code AS la_country_code,
    l_area.country_name AS la_country_name,
    l_area.year AS la_year,
    l_area.total_area_sq_mi,
    r.country_code AS r_country_code,
    r.country_name AS r_country_name,
    r.region,
    r.income_group,
    (f_area.forest_area_sqkm / (l_area.total_area_sq_mi * 2.59)) * 100 AS forest_percentage
FROM
    forest_area AS f_area
JOIN
    land_area AS l_area
ON
    f_area.year = l_area.year
    AND f_area.country_code = l_area.country_code
JOIN
    regions AS r
ON
    r.country_code = f_area.country_code;
```

GLOBAL SITUATION

a. Finding sum of Forest in Km for the entire World as at 1990

```
SELECT forest_area_sqkm
FROM forestation
WHERE fa_country_name = 'World' and year = 1990;
```

b. Finding sum of Forest in Km for the entire World at 2016

```
SELECT forest_area_sqkm
FROM forestation
WHERE fa_country_name = 'World' and year = 2016;
```

c. The difference in the Forest Square from 1990 to 2016 -> using a Scalar Statement

```
SELECT
  (SELECT forest_area_sqkm
   FROM forestation
   WHERE fa_country_name = 'World' and year = 1990)-
  (SELECT forest_area_sqkm
   FROM forestation
   WHERE fa_country_name = 'World' and year = 2016) as forest_area_change;
```

d. Percentage difference in Forest Square in KMs from 1990 to 2016 -> using Common Table Expression

```
WITH sqkm_1990 as (SELECT forest_area_sqkm as f_a_1990
  FROM deforestation
  WHERE fa_country_name = 'World' and year = 1990
),
sqkm_2016 as (SELECT forest_area_sqkm as f_a_2016
  FROM deforestation
  WHERE fa_country_name = 'World' and year = 2016
)
```

---The above CTE also could have been used to answer the previous question even

```
SELECT ((f_a_1990 - f_a_2016)/(f_a_1990)*100) as percentage_difference
FROM sqkm_2016, sqkm_1990;
```

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?

```
SELECT *
FROM forestation
WHERE year = 2016 AND total_area_sq_mi <= 511370.27
ORDER BY total_area_sq_mi DESC
LIMIT 1
```

REGIONAL OUTLOOK

----Creating 'perfect_forest_2016' view

```
CREATE OR REPLACE VIEW new_regional_distribution AS
(SELECT r.region,
       l.year,
       SUM(f.forest_area_sqkm) sum_forest_area_sqkm,
       SUM(l.total_area_sq_mi*2.59) as total_area_sqkm,
       (SUM(f.forest_area_sqkm)/SUM(l.total_area_sq_mi*2.59))*100 AS percent_fa_region
FROM regions r
JOIN land_area l
ON r.country_code = l.country_code
JOIN forest_area f
ON f.country_code = l.country_code AND f.year = l.year
GROUP BY 1, 2);
```

a. What was the percent forest of the entire world in 2016?

```
SELECT region, ROUND(CAST(percent_fa_region as numeric), 2)
FROM new_regional_distribution
WHERE year = 2016 AND region = 'World'
```

----Which region had the HIGHEST percent forest in 2016?

```
SELECT region, ROUND(CAST(percent_fa_region AS numeric), 2)
FROM new_regional_distribution
WHERE year = 2016
ORDER BY 2 DESC, 1
LIMIT 1;
```

— Which region had the LOWEST percent forest in 2016?

```
SELECT region, ROUND(CAST(percent_fa_region AS numeric), 2)
FROM new_regional_distribution
WHERE year = 2016
ORDER BY 2 ASC, 1
LIMIT 1;
```


b. What was the percent forest of the entire world in 1990?

```
SELECT region, ROUND(CAST(percent_fa_region as numeric), 2)
FROM new_regional_distribution
WHERE year = 1990 AND region = 'World';
```

— Which region had the HIGHEST percent forest in 1990?

```
SELECT region, ROUND(CAST(percent_fa_region AS numeric), 2)
FROM new_regional_distribution
WHERE year = 1990
ORDER BY 2 DESC, 1
LIMIT 1;
```

— Which region had the LOWEST percent forest in 1990?

```
SELECT region, ROUND(CAST(percent_fa_region AS numeric), 2)
FROM new_regional_distribution
WHERE year = 1990
ORDER BY 2 ASC, 1
LIMIT 1;
```

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

```
WITH
percentage_1990 AS
  (SELECT * FROM new_regional_distribution WHERE year=1990),
percentage_2016 AS
  (SELECT * FROM new_regional_distribution WHERE year=2016)
SELECT percentage_2016.region,
  ROUND(CAST(percentage_1990.percent_fa_region AS numeric), 2) AS
percentage_1990_fa,
  ROUND(CAST(percentage_2016.percent_fa_region AS numeric), 2) AS
percentage_2016_fa
FROM percentage_2016
JOIN percentage_1990 ON percentage_2016.region = percentage_1990.region
WHERE percentage_1990.percent_fa_region > percentage_2016.percent_fa_region
ORDER BY 2 DESC;
```

COUNTRY-LEVEL DETAIL

---MAKING a 'country_division' VIEW

```
CREATE OR REPLACE VIEW country_division AS
(SELECT r.country_name,
       l.year,
       SUM(f.forest_area_sqkm) sum_forest_area_sqkm,
       SUM(l.total_area_sq_mi*2.59) as total_area_sqkm,
       (SUM(f.forest_area_sqkm)/SUM(l.total_area_sq_mi*2.59))*100 AS
percent_fa_region
FROM regions r
JOIN land_area l
ON r.country_code = l.country_code
JOIN forest_area f
ON f.country_code = l.country_code AND f.year = l.year
GROUP BY 1, 2);
```

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 forest_area_1990 AS (
  SELECT
    country_code,
    country_name,
    year,
    forest_area_sqkm AS forest_area_sqkm_1990
  FROM forest_area
  WHERE year = 1990
    AND forest_area_sqkm IS NOT NULL
    AND country_name != 'World'
),
forest_area_2016 AS (
  SELECT
    country_code,
    country_name,
    year,
    forest_area_sqkm AS forest_area_sqkm_2016
```

```

FROM forest_area
WHERE year = 2016
      AND forest_area_sqkm IS NOT NULL
      AND country_name != 'World'
)
SELECT
  forest_area_1990.country_code,
  forest_area_1990.country_name,
  r.region,
  forest_area_1990.forest_area_sqkm_1990 AS fa_1990_sqkm,
  forest_area_2016.forest_area_sqkm_2016 AS fa_2016_sqkm,
  forest_area_1990.forest_area_sqkm_1990 -
  forest_area_2016.forest_area_sqkm_2016 AS diff_fa_sqkm
FROM
  forest_area_2016
JOIN
  forest_area_1990
ON
  forest_area_1990.country_code = forest_area_2016.country_code
  AND (forest_area_sqkm_1990 IS NOT NULL AND forest_area_sqkm_2016 IS NOT
  NULL)
JOIN
  regions r
ON
  r.country_code = forest_area_1990.country_code
ORDER BY
  6 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 forest_area_1990 AS (  
    SELECT country_code, country_name, forest_area_sqkm AS  
forest_area_sqkm_1990  
    FROM forest_area  
    WHERE year = 1990 AND forest_area_sqkm IS NOT NULL AND country_name !=  
'World'  
),  
forest_area_2016 AS (  
    SELECT country_code, country_name, forest_area_sqkm AS  
forest_area_sqkm_2016  
    FROM forest_area  
    WHERE year = 2016 AND forest_area_sqkm IS NOT NULL AND country_name !=  
'World'  
)  
  
SELECT  
    fa1990.country_name,  
    r.region,  
    ROUND(fa1990.forest_area_sqkm_1990::numeric, 2) AS forest_area_sqkm_1990,  
    ROUND(fa2016.forest_area_sqkm_2016::numeric, 2) AS forest_area_sqkm_2016,  
    ROUND(((fa1990.forest_area_sqkm_1990 - fa2016.forest_area_sqkm_2016) * 100.0 /  
fa1990.forest_area_sqkm_1990)::numeric, 2) AS percent_decrease  
FROM  
    forest_area_1990 fa1990  
JOIN  
    forest_area_2016 fa2016  
ON  
    fa1990.country_code = fa2016.country_code  
JOIN  
    regions r  
ON  
    r.country_code = fa1990.country_code  
WHERE  
    fa1990.forest_area_sqkm_1990 > fa2016.forest_area_sqkm_2016  
ORDER BY  
    percent_decrease DESC  
LIMIT 5;
```

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

```
WITH count AS (  
  SELECT  
    fa_country_name,  
    year,  
    forest_percentage,  
    CASE  
      WHEN forest_percentage >= 75 THEN 4  
      WHEN forest_percentage < 75 AND forest_percentage >= 50 THEN 3  
      WHEN forest_percentage < 50 AND forest_percentage >= 25 THEN 2  
      ELSE 1  
    END AS percentile  
  FROM forestation  
  WHERE year = 2016  
  AND forest_percentage IS NOT NULL  
  AND fa_country_name != 'World'  
)  
SELECT  
  percentile,  
  COUNT(percentile) as percentile_count  
FROM count  
GROUP BY percentile  
ORDER BY COUNT(percentile) DESC;
```

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

```
WITH count AS (  
  SELECT  
    fa_country_name as Country,  
    year,  
    region,  
    forest_percentage,  
    CASE  
      WHEN forest_percentage >= 75 THEN 4  
      WHEN forest_percentage < 75 AND forest_percentage >= 50 THEN 3  
      WHEN forest_percentage < 50 AND forest_percentage >= 25 THEN 2  
      ELSE 1  
    END AS percentile  
  FROM forestation  
  JOIN  
    WHERE year = 2016  
    AND forest_percentage IS NOT NULL  
    AND fa_country_name != 'World'  
)  
SELECT  
  Country,  
  region,  
  percentile,  
  ROUND(CAST(forest_percentage AS numeric), 2) as forest_percentage  
FROM count  
WHERE percentile = 4  
ORDER BY 1 ASC;
```

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

```
WITH count AS (  
  SELECT  
    fa_country_name as Country,  
    year,  
    forest_percentage,  
    CASE  
      WHEN forest_percentage >= 75 THEN 4  
      WHEN forest_percentage < 75 AND forest_percentage >= 50 THEN 3  
      WHEN forest_percentage < 50 AND forest_percentage >= 25 THEN 2  
      ELSE 1  
    END AS percentile  
  FROM forestation  
  WHERE year = 2016  
  AND forest_percentage IS NOT NULL  
  AND fa_country_name != 'World'  
)  
SELECT  
  Country,  
  percentile,  
  ROUND(CAST(forest_percentage AS numeric), 2) as forest_percentage  
FROM count  
WHERE forest_percentage > (SELECT forest_percentage FROM count WHERE  
Country = 'United States')  
ORDER BY ROUND(CAST(forest_percentage AS numeric), 2) desc;
```