

# 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 **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 sq km**, 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 **1279999.99 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
World	32.42	31.38
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 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 **79200 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	541510.00
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 Intervals, 2016:

Interval	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 **First(0% - 25%)** interval. There were **9** countries in the top interval 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 Countries, 2016:

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

The World Bank data reveals that global forest cover has decreased by 3.21% from 1990 to 2016. The world lost approximately 1.32 million square kilometers of forest area from 1990 to 2016, slightly more than the entire land area of Peru.

Key regions like Latin America & the Caribbean and Sub-Saharan Africa saw significant forest loss. Brazil experienced the largest absolute reduction, while Togo, Uganda and Nigeria faced the steepest percentage declines in forest area. Nigeria is notable for both its high absolute and percent decrease in forest area, indicating it is heavily affected by deforestation.

Success stories:

China is a standout example of successful reforestation and forest area increase, with an addition of 527,229 sq km of forest.

Iceland shows an extraordinary percent increase in forest area by 213.66%, highlighting successful reforestation efforts.

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

Our primary focus should be on the countries in the Sub-Saharan African region.

Especially countries like Togo, Uganda and Nigeria, with a high percent decrease in forest area, should be prioritized for conservation and reforestation efforts. We should also focus on Brazil and Indonesia due to their high absolute forest loss.

By focusing on these countries and regions, and leveraging successful case studies like China and Iceland, global efforts can be better directed towards mitigating deforestation and promoting successful afforestation efforts.

## 5. APPENDIX: SQL Queries Used

Create a **forestation** View joining all three tables - forest\_area, land\_area, and regions. The forest\_area and land\_area tables join on both country\_code AND year. The regions table joins these based on only country\_code.

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

### Part 1 - 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 forest_area_sqkm
FROM forestation
WHERE year = 1990 AND country_name = '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 forest_area_sqkm
FROM forestation
WHERE year = 2016 AND country_name = 'World';
```

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

```
WITH forest_1990 AS
(
    SELECT forest_area_sqkm
    FROM forestation
    WHERE country_name = 'World' AND year = 1990
),
forest_2016 AS
(
    SELECT forest_area_sqkm
    FROM forestation
    WHERE country_name = 'World' AND year = 2016
)

SELECT forest_1990.forest_area_sqkm - forest_2016.forest_area_sqkm AS lost_forest_area
FROM forest_1990, forest_2016
```

OR We can use SELF JOIN to get the same result

```
SELECT
    f1.forest_area_sqkm - f2.forest_area_sqkm AS lost_forest_area
FROM forestation f1
JOIN forestation f2
ON f1.country_name = f2.country_name
AND f1.year = 1990 AND f2.year = 2016
WHERE f1.country_name = 'World';
```

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

```
WITH forest_1990 AS (  
    SELECT forest_area_sqkm  
    FROM forestation  
    WHERE country_name = 'World' AND year = 1990  
) ,  
forest_2016 AS (  
    SELECT forest_area_sqkm  
    FROM forestation  
    WHERE country_name = 'World' AND year = 2016  
)  
  
SELECT  
    ROUND(  
        CAST(  
            (f1990.forest_area_sqkm - f2016.forest_area_sqkm) /  
f1990.forest_area_sqkm AS NUMERIC) * 100,  
        2) AS pct_forest_lost  
FROM forest_1990 f1990, forest_2016 f2016
```

OR We can use SELF JOIN to get the same result

```
SELECT  
    ROUND(  
        CAST(  
            (f1.forest_area_sqkm - f2.forest_area_sqkm) / f1.forest_area_sqkm  
AS NUMERIC) * 100,  
        2) AS pct_forest_lost  
FROM forestation f1  
JOIN forestation f2  
ON f1.country_name = f2.country_name  
AND f1.year = 1990 AND f2.year = 2016  
WHERE f1.country_name = '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 forest_1990 AS (  
    SELECT forest_area_sqkm  
    FROM forestation  
    WHERE country_name = 'World' AND year = 1990  
) ,  
  
forest_2016 AS (  
    SELECT forest_area_sqkm  
    FROM forestation  
    WHERE country_name = 'World' AND year = 2016  
) ,  
  
forest_loss AS (  
    SELECT  
        forest_1990.forest_area_sqkm - forest_2016.forest_area_sqkm AS  
lost_forest_area  
    FROM  
        forest_1990,  
        forest_2016  
)  
  
SELECT  
    country_name,  
    year,  
    ROUND(  
        CAST  
        (  
            ABS((total_area_sq_mi * 2.59) - (SELECT lost_forest_area FROM  
forest_loss)) AS NUMERIC  
        ), 2) AS closest_land_area  
FROM land_area  
WHERE year = 2016  
ORDER BY 3  
LIMIT 1;
```

## Part 2 - Regional Outlook

1. Create a table that shows the Regions and their percent forest area (sum of forest area divided by the sum of land area) in 1990 and 2016. (Note 1 sq mi = 2.59 sq km).

```
SELECT
    r.region,
    fa.year,
    ROUND(
        CAST(
            SUM(fa.forest_area_sqkm) / SUM((la.total_area_sq_mi * 2.59)) AS NUMERIC
        ) * 100,
        2) AS forest_percent
FROM forest_area fa
JOIN land_area la
ON fa.country_code = la.country_code AND fa.year = la.year
JOIN regions r ON la.country_code = r.country_code
WHERE fa.year IN (1990, 2016)
GROUP BY 1, 2
ORDER BY 2, 3 DESC, 1;
```

I have created a View named **regional\_outlook** using the table above to address the following questions.

```
CREATE VIEW regional_outlook AS
(
    SELECT
        r.region,
        fa.year,
        ROUND(
            CAST(
                SUM(fa.forest_area_sqkm) / SUM((la.total_area_sq_mi * 2.59)) AS NUMERIC
            ) * 100,
            2) AS forest_percent
    FROM forest_area fa
    JOIN land_area la
    ON fa.country_code = la.country_code AND fa.year = la.year
    JOIN regions r ON la.country_code = r.country_code
    WHERE fa.year IN (1990, 2016)
    GROUP BY 1, 2
    ORDER BY 2, 3 DESC, 1
);
```

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?

Forest percent of the entire world in 2016

```
SELECT *  
FROM regional_outlook  
WHERE region = 'World' AND year = 2016;
```

Highest forest percent in 2016

```
SELECT *  
FROM regional_outlook  
WHERE year = 2016  
ORDER BY forest_percent DESC  
LIMIT 1;
```

Least forest percent in 2016

```
SELECT *  
FROM regional_outlook  
WHERE year = 2016  
ORDER BY forest_percent  
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?

Forest percent of the entire world in 1990

```
SELECT *  
FROM regional_outlook  
WHERE region = 'World' AND year = 1990;
```

Highest forest percent in 1990

```
SELECT *  
FROM regional_outlook  
WHERE year = 1990  
ORDER BY forest_percent DESC  
LIMIT 1;
```

Least forest percent in 1990

```
SELECT *  
FROM regional_outlook  
WHERE year = 1990  
ORDER BY forest_percent  
LIMIT 1;
```

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

```
WITH regional_outlook_1990 AS  
(  
    SELECT region, forest_percent AS forest_1990  
    FROM regional_outlook  
    WHERE year = 1990  
)  
regional_outlook_2016 AS  
(  
    SELECT region, forest_percent AS forest_2016  
    FROM regional_outlook  
    WHERE year = 2016  
)  
  
SELECT r1.region, r1.forest_1990, r2.forest_2016  
FROM regional_outlook_1990 r1  
JOIN regional_outlook_2016 r2  
ON r1.region = r2.region  
WHERE r1.forest_1990 > r2.forest_2016  
AND r1.region != 'World';
```

### Part 3 - Country-Level Detail

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_change AS
(
    SELECT
        country_name,
        SUM(CASE WHEN year = 1990 THEN forest_area_sqkm ELSE 0 END) AS forest_area_1990,
        SUM(CASE WHEN year = 2016 THEN forest_area_sqkm ELSE 0 END) AS forest_area_2016
    FROM forest_area
    WHERE year IN (1990, 2016) AND country_name != 'World'
    GROUP BY 1
)

SELECT
    fac.country_name,
    r.region,
    ROUND(CAST(fac.forest_area_1990 - fac.forest_area_2016 AS NUMERIC), 2) AS
difference
FROM forest_area_change fac
JOIN regions r
ON fac.country_name = r.country_name
ORDER BY 3 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_change AS
(
    SELECT
        country_name,
        SUM(CASE WHEN year = 1990 THEN forest_area_sqkm ELSE 0 END) AS forest_area_1990,
        SUM(CASE WHEN year = 2016 THEN forest_area_sqkm ELSE 0 END) AS forest_area_2016
    FROM forest_area
    WHERE year IN (1990, 2016) AND country_name != 'World'
    GROUP BY 1
)

SELECT
    fac.*,
    r.region,
    ROUND(
        CAST(
            CASE WHEN fac.forest_area_1990 > 0 THEN (fac.forest_area_1990 -
fac.forest_area_2016)/fac.forest_area_1990 * 100 ELSE 0 END AS NUMERIC
        ),
        2) AS pct_difference
FROM forest_area_change fac
JOIN regions r
ON fac.country_name = r.country_name
WHERE fac.forest_area_1990 > 0 AND fac.forest_area_2016 > 0
ORDER BY 5 DESC
LIMIT 5;
```

c. If countries were grouped by percent forestation into 4 equal static intervals, which group had the most countries in it in 2016?

```
WITH forest_interval AS
(
    SELECT
        country_name,
        region,
        ROUND(CAST(forest_cover_percent AS NUMERIC), 5) AS forest_cover_percent,
        CASE
            WHEN forest_cover_percent <= 25 THEN '0 - 25'
            WHEN forest_cover_percent > 25 AND forest_cover_percent <= 50 THEN '25 - 50'
            WHEN forest_cover_percent > 50 AND forest_cover_percent <= 75 THEN '50 - 75'
            ELSE '75 - 100' END AS interval
        FROM forestation
        WHERE year = 2016 AND forest_cover_percent IS NOT NULL
        AND country_name != 'World'
        ORDER BY 3
    )

SELECT interval, COUNT(*) AS interval_count
FROM forest_interval
GROUP BY 1
ORDER BY 2 DESC;
```

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

```
WITH forest_quartile AS
(
    SELECT
        country_name,
        forest_cover_percent,
        NTILE(4) OVER (ORDER BY forest_cover_percent) AS quartile
        FROM forestation
        WHERE year = 2016
        AND forest_cover_percent IS NOT NULL
        AND country_name != 'World'
    )

SELECT quartile, COUNT(*) AS quartile_count
FROM forest_quartile
GROUP BY 1
ORDER BY 1;
```

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

```
WITH countries_ranked AS
(
    SELECT
        country_name,
        region,
        ROUND(CAST(forest_cover_percent AS NUMERIC), 5) AS forest_cover_percent,
        CASE
            WHEN forest_cover_percent <= 25 THEN 1
            WHEN forest_cover_percent > 25 AND forest_cover_percent <= 50 THEN 2
            WHEN forest_cover_percent > 50 AND forest_cover_percent <= 75 THEN 3
            ELSE 4 END AS ntile
        FROM forestation
        WHERE year = 2016
        AND forest_cover_percent IS NOT NULL
        AND country_name != 'World'
        ORDER BY 3
)

SELECT *
FROM countries_ranked
WHERE ntile = 4;
```

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

```
WITH US_forestation AS
(
    SELECT
        ROUND(CAST(forest_cover_percent AS NUMERIC), 2) AS forest_percent
        FROM forestation
        WHERE year = 2016 AND country_name LIKE '%United States%'
)

SELECT
    country_name,
    year,
    ROUND(CAST(forest_cover_percent AS NUMERIC), 2) AS forest_cover_percent
FROM forestation f, US_forestation us
WHERE f.year = 2016 AND f.forest_cover_percent IS NOT NULL
AND f.country_name != 'World'
AND f.forest_cover_percent > us.forest_percent
ORDER BY 1;
```