Report for ForestQuery into Global Deforestation, 1990 to 2016

By Stan Taov

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** sqkm 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 **3.21%**.

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 **1280000** 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
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, 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** 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 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
Brazil	Latin America & Caribbean	541510 sq km
Indonesia	East Asia & Pacific	282194 sq km
Myanmar	East Asia & Pacific	107234 sq km
Nigeria	Sub-Saharan Africa	106506 sq km
Tanzania	Sub-Saharan Africa	102320 sq km

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 first 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

There were 94 countries with 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?*

The data analysis revealed that the global forest area has decreased from 1990 to 2016, some countries lost a significant amount of forest area over the same period of time. However, there were some bright spots, certain countries increased forest area as an example ***Iceland*** *increased in forest area by* **213.66%***.*

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

During the analysis, I discovered that the Sub-Saharan Africa region is under huge stress of deforestation. One country lost over ***75%*** *of the forest area during the 1990-2016 period. So, we need to focus on the following Sub-Saharan Africa countries (Togo, Nigeria, Uganda, Mauritania and Honduras) to develop programs focusing on increasing the forestation area.*

APPENDIX: SQL queries used

 # Create a View called “forestation” by joining all three tables - forest_area, land_area and regions in the workspace.

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

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 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 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 forest_area fa1, forest_area fa2
WHERE fa1.year = '1990' AND fa2.year = '2016' AND fa1.country_name = 'World' AND
fa2.country_name = 'World'
```

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

```
SELECT (fa2.forest_area_sqkm - fa1.forest_area_sqkm)*100/fa1.forest_area_sqkm AS
forest_area_pct_change
FROM forest_area fa1, forest_area fa2
```

```
WHERE fa1.year = '1990' AND fa2.year = '2016' AND fa1.country_name = 'World' AND
fa2.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 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;
```

2. REGIONAL OUTLOOK

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?

```
SELECT *
FROM forestation
WHERE year = 2016
AND country_name = 'World'

SELECT region,
```

```
ROUND(CAST(SUM(forest_area_sqkm)*100/SUM(total_area_sqkm) AS NUMERIC), 2) AS
forest_area
FROM forestation
WHERE year = 2016
GROUP BY 1
ORDER BY 2 DESC
LIMIT 1;


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

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

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?

```
SELECT *
FROM forestation
WHERE year = 1990
AND country_name = 'World'

SELECT region,
ROUND(CAST(SUM(forest_area_sqkm)*100/SUM(total_area_sqkm) AS NUMERIC), 2) AS
forest_area
FROM forestation
WHERE year = 1990
GROUP BY 1
ORDER BY 2 DESC
```

```
LIMIT 1;

SELECT region,
ROUND(CAST(SUM(forest_area_sqkm)*100/SUM(total_area_sqkm) AS NUMERIC), 2) AS
forest_area
FROM forestation
WHERE year = 1990
GROUP BY 1
ORDER BY 2
LIMIT 1;

SELECT region,
ROUND(CAST(SUM(forest_area_sqkm)*100/SUM(total_area_sqkm) AS NUMERIC), 2) AS
forest_area
FROM forestation
WHERE year = 1990
GROUP BY 1
ORDER BY 2;
```

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,
ROUND(CAST(SUM(forest_area_sqkm)*100/SUM(total_area_sqkm) AS NUMERIC), 2) AS
forest_area
FROM forestation
WHERE year = 1990
GROUP BY 1
ORDER BY 2),

tb2 AS (SELECT region,
ROUND(CAST(SUM(forest_area_sqkm)*100/SUM(total_area_sqkm) AS NUMERIC), 2) AS
forest_area
FROM forestation
WHERE year = 2016
GROUP BY 1
ORDER BY 2)

SELECT tb1.region, tb1.forest_area AS forest_area_1990,
tb2.forest_area AS forest_area_2016,
tb2.forest_area - tb1.forest_area AS diff
FROM tb1
```

```
JOIN tb2
ON tb1.region = tb2.region
ORDER BY 2 DESC;
```

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 tb1 AS(SELECT *,
forest_area_sqkm - LAG(forest_area_sqkm) over (PARTITION BY  country_name ORDER BY year) as diff
FROM forestation AS f
WHERE f.year = 2016 OR f.year = 1990 AND forest_area_sqkm IS NOT NULL)

SELECT *
FROM tb1
WHERE year = 2016 AND diff IS NOT NULL
ORDER BY diff DESC
LIMIT 5;

----------
WITH tb1 AS(SELECT *,
(forest_area_sqkm*100)/LAG(forest_area_sqkm) OVER (PARTITION BY  country_name ORDER BY year) - 100 as diff_percentage
FROM forestation AS f
WHERE f.year = 1990 OR f.year = 2016 AND forest_area_sqkm IS NOT NULL)

SELECT country_name, ROUND(CAST(diff_percentage AS NUMERIC) ,2) AS diff_percentage_round
FROM tb1
WHERE  diff_percentage IS NOT NULL
ORDER BY diff_percentage 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?

```
SELECT *, forest_area_sqkm - LAG(forest_area_sqkm) over (PARTITION BY country_name
ORDER BY year) AS diff
FROM forestation
WHERE year = 1990 OR year = 2016 AND region NOT LIKE 'World'
ORDER BY diff
LIMIT 5;
```

```
SELECT *,
ROUND(CAST((forest_area_sqkm*100 / LAG(forest_area_sqkm) OVER (PARTITION BY
country_name ORDER BY year)-100) AS NUMERIC), 2) AS diff
FROM forestation
WHERE year = 1990 OR year = 2016 AND region NOT LIKE 'World'
ORDER BY diff
```

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 '4th'
WHEN percent_forestation <= 75 AND  percent_forestation > 50 THEN '3rd'
WHEN percent_forestation <= 50 AND percent_forestation > 25 THEN '2nd'
ELSE '1st'
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.

```
WITH tb1 AS (SELECT *
```

```
FROM forestation
WHERE year = 2016 AND region NOT LIKE 'World'),

tb2 AS (SELECT *,
CASE WHEN percent_forestation > 75 THEN '4th'
WHEN percent_forestation <= 75 AND  percent_forestation > 50 THEN '3rd'
WHEN percent_forestation <= 50 AND percent_forestation > 25 THEN '2nd'
ELSE '1st'
END AS Quartiles
FROM tb1
WHERE percent_forestation IS NOT NULL)

SELECT *
FROM tb2
WHERE quartiles = '4th'
ORDER BY percent_forestation DESC;
```

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

```
WITH tb1 AS (SELECT *
FROM forestation
WHERE year = 2016 AND region NOT LIKE 'World'),

tb2 AS (SELECT *,
CASE WHEN percent_forestation > 75 THEN '4th'
WHEN percent_forestation <= 75 AND  percent_forestation > 50 THEN '3rd'
WHEN percent_forestation <= 50 AND percent_forestation > 25 THEN '2nd'
ELSE '1st'
END AS Quartiles
FROM tb1
WHERE percent_forestation IS NOT NULL),

tb3 AS (SELECT percent_forestation
FROM tb2
WHERE country_name = 'United States'),

tb4 AS (SELECT country_name, tb2.percent_forestation, tb3.percent_forestation AS
percent_forestation_USA
FROM tb2, tb3)

SELECT percent_forestation_usa, COUNT(*) AS number_countries
FROM tb4
```

```
WHERE percent_forestation > percent_forestation_usa
GROUP BY 1;
```