Report for Forest Query into Global Deforestation, 1990 to 2016

By: Ahmed Ibrahim

ForestQuery is on a mission to combat deforestation worldwide and 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 and a table of countries and the regions to which they belong.

The data analysis team has used SQL to bring these tables together and query them to find areas of concern and 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 km^2** in 1990. As of 2016, the most recent year for which data was available, that number had fallen to **39,958,245.9 km^2**, a loss of **1,324,449.0 km^2**, 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 km^2**).

2. **REGIONAL OUTLOOK**

In 2016, the percent of the total land area of the world designated as the forest was **31.38 %**. The region with the highest relative forestation was **Europe & Central Asia**, with **46.16 %**, and the region with the lowest relative forestation was the **Middle East & North Africa**, with **2.07 %** forestation.

In 1990, the percent of the total land area of the world designated as the forest was **32.42 %**. The region with the highest relative forestation was **Latin America & the Caribbean**, with **51.03 %**, and the region with the lowest relative forestation was the **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 (%)
South Asia	**16.51**	**17.51**
Europe & Central Asia	**37.28**	**38.04**
East Asia & Pacific	**25.78**	**26.36**
North America	**35.65**	**36.04**
Middle East & North Africa	**1.78**	**2.07**
Sub-Saharan Africa	**30.67**	**28.79**
Latin America & Caribbean	**51.03**	**46.16**

The only regions of the world that decreased in percent forest area from 1990 to 2016 were **Sub-Saharan Africa** (dropped from **30.67 %** to **28.79 %**) and **Latin America & the Caribbean** (**51.03 %** to **45.45 %**). All other regions increased in forest area over this time. However, the drop in forest area in the two regions was so significant that the world's percent forest area decreased over this time from **32.42 %** to **31.38 %**.

3. **COUNTRY-LEVEL DETAIL**

A. SUCCESS STORIES

There is one exceptionally bright spot in the data at the country level, **China**. This country increased in forest area from 1990 to 2016 by **527,229.062 km^2**. It would be interesting to study what has changed in this country to drive this figure in the data higher. The country with the subsequent most significant increase in forest area from 1990 to 2016 was the **United States**, but it only saw an increase of **79,200 km^2**, much lower than the figure for **China**.

Of course, the United States and China are among the largest countries in total land area, so when we look at the largest *percent* change in forest area from 1990 to 2016, we are not 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 most considerable 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 most significant 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	Forest Area Change (km^2)
Brazil	**Latin America & Caribbean**	**-541510**
Indonesia	**East Asia & Pacific**	**-282193.9844**
Myanmar	**East Asia & Pacific**	**-107234.0039**
Nigeria	**Sub-Saharan Africa**	**-106506.00098**
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 the most between 1990 and 2016, we find that four of the top 5 countries on the list are in **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 the forest and 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
FIRST	**85**
SECOND	**73**
THIRD	**38**
FOURTH	**9**

The most significant number of countries in 2016 were found in the **FIRST (0 to 25 %)** quartile.

There were **9** countries in the top quartile in 2016. These are countries with an extremely 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 in 2016 with a higher percentage of forestation than the United States.

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 total amount of forestation in the world decreased from 1990 to 2016 at a loss of 3.21 %. Despite this decrease, the trend is uniformly distributed worldwide, where the European and Asian regions have seen an increase. In contrast, the American, African, and Middle Eastern regions have succumbed to a loss of forestation. However, two countries from the East Asia & Pacific region were ranked top 5 in forest area loss globally.

On the country level, 4 sub-Saharan countries are among the top 5 countries in percent of forestation loss from 1990 to 2016. At the same time, Nigeria is the only country in both the top 5 in percentage and total loss forestation area for that same period. The study also illustrates a potential selective bias, where larger countries can have higher variance in forestation area, while smaller countries can have higher fluctuations on a percentage basis.

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

The study did blatantly point to culprits in some regions and countries. Countries that would require focus include Brazil, Indonesia, Myanmar, Nigeria, Tanzania, Togo, Uganda, Mauritania, and Honduras. These countries derive from three regions: Latin America & the Caribbean; East Asia & the Pacific; Sub-Saharan Africa. To select a shortlist of countries to focus on, it is worth noting that other parameters would be needed for an effective plan of action, such as fields related to the social-economical and political climate of these nations. However, based on the herein data exploration, countries that would require the utmost focus would be Brazil, Indonesia, and Nigeria.

5. Appendix

```
/* ------------- STEPS TO COMPLETE ------------- */

/* Create 'forestation' table using 'forest_area', 'land_area', 'regions'. */
CREATE VIEW forestation AS SELECT DISTINCT fa.country_name AS country,
   fa.country_code AS code,
   fa.year AS year,
   CAST((fa.forest_area_sqkm) AS NUMERIC) AS forest_area_sqkm,
   CAST((la.total_area_sq_mi*2.59) AS NUMERIC) AS total_area_sqkm,
   ROUND(CAST(100.00*((fa.forest_area_sqkm)/(la.total_area_sq_mi*2.59)) AS NUMERIC), 2)
AS percentage_forestation,
   re.region,
   re.income_group
FROM forest_area fa
JOIN land_area la
   ON fa.country_code = la.country_code
      AND fa.year = la.year
JOIN regions re
   ON la.country_code = re.country_code;

/* Check if 'forestation' table exist */
DROP VIEW IF EXISTS forestation;

/* ------------- Part 1 - Global Situation ------------- */

/* What was the total forest area (in sq km) of the world in 1990? */
SELECT DISTINCT year,
   forest_area_sqkm
FROM forestation
WHERE code = 'WLD'
AND year = '1990';

/* What was the total forest area (in sq km) of the world in 2016? */
SELECT DISTINCT year,
   forest_area_sqkm
FROM forestation
WHERE code = 'WLD'
AND year = '2016';
```

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

```
SELECT year,
   forest_area_sqkm,
   forest_area_sqkm - LEAD(forest_area_sqkm) OVER (ORDER BY year) AS
world_forest_area_change_sqkm
FROM    (SELECT DISTINCT code, year,
       forest_area_sqkm
   FROM forestation
   WHERE (year = '1990' OR year = '2016')
   AND code = 'WLD'
) AS world_forest_area_for_1990_2016_sqkm;
```

/* What was the percent change in forest area of the world between 1990 and 2016? */

```
SELECT year,
   forest_area_sqkm,
   ROUND((100*(forest_area_sqkm - LEAD(forest_area_sqkm) OVER (ORDER BY
year))/forest_area_sqkm),2) AS world_percentage_forest_area_change
FROM    (SELECT DISTINCT code, year,
       forest_area_sqkm
   FROM forestation
   WHERE (year = '1990' OR year = '2016')
   AND code = 'WLD'
) AS world_forest_area_for_1990_2016_sqkm;
```

/* 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 T1 AS (SELECT DISTINCT ft1.forest_area_sqkm - ft2.forest_area_sqkm AS
world_forest_area_change_sqkm
   FROM forestation ft1,
       forestation ft2
   WHERE (ft1.code = 'WLD' AND ft1.year = '1990')
   AND (ft2.code = 'WLD' AND ft2.year = '2016'))
SELECT DISTINCT ft.country,
   ft.total_area_sqkm,
   ABS(total_area_sqkm - T1.world_forest_area_change_sqkm) AS
forest_change_vs_country_size_difference
FROM forestation ft, T1
ORDER BY forest_change_vs_country_size_difference
LIMIT 1;
```

/* ---------- Part 2 - Regional Outlook ---------- */

/* Create a table that shows the regions and their percent forest area in 1990 and 2016. */

```
CREATE VIEW regional_outlook_ft AS SELECT year,
   region,
   ROUND(CAST(100.00*((sum_forest_area_sqkm)/(sum_total_area_sqkm)) AS NUMERIC), 2)
AS percentage_forestation
FROM (SELECT DISTINCT year,
   region,
   SUM(forest_area_sqkm) AS sum_forest_area_sqkm,
   SUM(total_area_sqkm) AS sum_total_area_sqkm
FROM forestation
WHERE (year = '1990' OR year = '2016')
GROUP BY 1, 2) AS forest_total_area_per_region_for_1990_2016_sqkm;
```

/* Check if 'regional_outlook_ft' table exist */

```
DROP VIEW IF EXISTS regional_outlook_ft;
```

/* What was the percent forest of the entire world in 2016? */

```
SELECT percentage_forestation
FROM regional_outlook_ft
WHERE region = 'World'
AND year = '2016';
```

/* Which region had the HIGHEST percent forest in 2016? */

```
SELECT region,
   percentage_forestation
FROM regional_outlook_ft
WHERE year = '2016'
AND region != 'World'
ORDER BY percentage_forestation DESC
LIMIT 1;
```

/* Which region had the LOWEST percent forest in 2016? */

```
SELECT region,
   percentage_forestation
FROM regional_outlook_ft
WHERE year = '2016'
AND region != 'World'
ORDER BY percentage_forestation
LIMIT 1;
```

/* What was the percent forest of the entire world in 1990? */

```
SELECT percentage_forestation
FROM regional_outlook_ft
WHERE region = 'World'
AND year = '1990';
```

/* Which region had the HIGHEST percent forest in 1990? */

```
SELECT region,
   percentage_forestation
FROM regional_outlook_ft
WHERE year = '1990'
AND region != 'World'
ORDER BY percentage_forestation DESC
LIMIT 1;
```

/* Which region had the LOWEST percent forest in 1990? */

```
SELECT region,
   percentage_forestation
FROM regional_outlook_ft
WHERE year = '1990'
AND region != 'World'
ORDER BY percentage_forestation
LIMIT 1;
```

```
/* Which regions of the world DECREASED in forest area from 1990 to 2016? */
With T1 AS (SELECT roft1.region,
   roft2.percentage_forestation AS percentage_forestation_1990,
   roft1.percentage_forestation AS percentage_forestation_2016,
         roft1.percentage_forestation - roft2.percentage_forestation AS
         difference_percentage_forestation_1990_vs_2016
FROM regional_outlook_ft roft1,
   regional_outlook_ft roft2
WHERE (roft1.year = '2016' AND roft2.year = '1990')
AND (roft1.region = roft2.region)
ORDER BY difference_percentage_forestation_1990_vs_2016 DESC)
SELECT region,
   percentage_forestation_1990,
   percentage_forestation_2016,
   difference_percentage_forestation_1990_vs_2016,
   CASE WHEN difference_percentage_forestation_1990_vs_2016 < 0
      THEN 'DECREASED'
      ELSE 'INCREASED'
      END AS region_forest_trend
FROM T1
WHERE region != 'World';

/* ---------- Part 3 - Country-Level Detail ---------- */

/* Create table for country-level forest area change and percentage comparing 1990 and
2016. */
CREATE VIEW country_level_ft AS SELECT DISTINCT ft1.country,
   ft1.region,
   ft1.forest_area_sqkm AS forest_area_sqkm_2016,
   ft2.forest_area_sqkm AS forest_area_sqkm_1990,
   ft1.forest_area_sqkm - ft2.forest_area_sqkm AS
difference_forest_area_1990_vs_2016_sqkm,
   ROUND(100*((ft1.forest_area_sqkm - ft2.forest_area_sqkm)/ft2.forest_area_sqkm),2) AS
percentage_area_1990_vs_2016
FROM forestation ft1,
   forestation ft2
WHERE (ft1.year = '2016' AND ft2.year = '1990')
AND (ft1.country = ft2.country);

/* Check if 'forestation' table exist */
DROP VIEW IF EXISTS country_level_ft;
```

/* What are the two countries with the highest increase in forest area? */

```
SELECT country,
   difference_forest_area_1990_vs_2016_sqkm
FROM country_level_ft
WHERE difference_forest_area_1990_vs_2016_sqkm IS NOT NULL
AND country != 'World'
ORDER BY difference_forest_area_1990_vs_2016_sqkm DESC
LIMIT 2;
```

/* What is the country with the highest percentage increase in forest area? */

```
SELECT country,
   percentage_area_1990_vs_2016
FROM country_level_ft
WHERE percentage_area_1990_vs_2016 IS NOT NULL
AND country != 'World'
ORDER BY percentage_area_1990_vs_2016 DESC
LIMIT 1;
```

/* What is the country with the highest percentage increase in forest area? */

```
SELECT country,
   region,
   difference_forest_area_1990_vs_2016_sqkm
FROM country_level_ft
WHERE difference_forest_area_1990_vs_2016_sqkm IS NOT NULL
AND country != 'World'
ORDER BY difference_forest_area_1990_vs_2016_sqkm
LIMIT 5;
```

/* 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 country,
   region,
   percentage_area_1990_vs_2016
FROM country_level_ft
WHERE percentage_area_1990_vs_2016 IS NOT NULL
AND country != 'World'
ORDER BY percentage_area_1990_vs_2016
LIMIT 5;
```

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

```
WITH T1 AS (SELECT country,
   year,
   ROUND(AVG(percentage_forestation), 2) AS average_percentage_forestation
FROM forestation
WHERE (year = '2016' AND percentage_forestation IS NOT NULL)
GROUP BY 1, 2
),
T2 AS (SELECT country,
   average_percentage_forestation,
   CASE WHEN average_percentage_forestation <= 25
   THEN 'FIRST_QUARTILE'
   WHEN average_percentage_forestation <= 50
   THEN 'SECOND_QUARTILE'
   WHEN average_percentage_forestation <= 75
   THEN 'THIRD_QUARTILE'
   WHEN average_percentage_forestation <= 100
   THEN 'FOURTH_QUARTILE'
   END AS percentile_forestation
FROM T1)
SELECT percentile_forestation,
   COUNT(*) AS countries_per_percentile
FROM T2
GROUP BY 1
ORDER BY countries_per_percentile DESC;
```

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

```
WITH T1 AS (SELECT country,
   year,
   region,
   ROUND(AVG(percentage_forestation), 2) AS average_percentage_forestation
FROM forestation
WHERE (year = '2016' AND percentage_forestation IS NOT NULL)
AND country != 'World'
GROUP BY 1, 2, 3
),
T2 AS (SELECT country,
   region,
   average_percentage_forestation,
   CASE WHEN average_percentage_forestation <= 25
   THEN 'FIRST_QUARTILE'
   WHEN average_percentage_forestation <= 50
   THEN 'SECOND_QUARTILE'
   WHEN average_percentage_forestation <= 75
   THEN 'THIRD_QUARTILE'
   WHEN average_percentage_forestation <= 100
   THEN 'FOURTH_QUARTILE'
   END AS percentile_forestation
FROM T1)
SELECT country,
   region,
   average_percentage_forestation
FROM T2
WHERE percentile_forestation = 'FOURTH_QUARTILE'
ORDER BY average_percentage_forestation DESC;
```

/* How many countries had a percent forestation higher than the United States in 2016? */

```
WITH T1 AS (SELECT country,
   year,
   ROUND(AVG(percentage_forestation), 2) AS average_percentage_forestation
FROM forestation
WHERE (year = '2016' AND percentage_forestation IS NOT NULL)
GROUP BY 1, 2
),
T2 AS (SELECT average_percentage_forestation
FROM T1
WHERE country = 'United States'),
T3 AS (SELECT T1.country,
   T1.average_percentage_forestation
FROM T1, T2
WHERE (T1.average_percentage_forestation > T2.average_percentage_forestation))
SELECT COUNT(*) AS count_countries
FROM T3;
```