Report for ForestQuery into Global Deforestation, 1990 to 2016

By Maricruz Polanco

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 41,282,695.9 square kilometers in 1990. As of 2016, the most recent year for which data was available, that number had fallen to 39,958,245.9 square kilometers, a loss of 1,324,449 square kilometers, 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 square kilometers).

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
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 527,229 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 79,200, 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 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 (sq km)
Brazil	Latin America & Caribbean	541,510
Indonesia	East Asia & Pacific	282,194
Myanmar	East Asia & Pacific	107,234
Nigeria	Sub-Saharan Africa	106,506
Tanzania	Sub-Saharan Africa	102,320

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 the region of Sub-Saharan Africa. The countries are Togo, Nigeria, Uganda, and Mauritania. The 5th country on the list is Hondura, 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
0% - 25%	85
25% - 50%	72
50% - 75%	38
75% - 100%	9

The largest number of countries in 2016 were found in the 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
Seychelles	Sub-Saharan Africa	88.41
Palau	East Asia & Pacific	87.61
American Samoa	East Asia & Pacific	87.5
Guyana	Latin America & Caribbean	83.9
Lao PDR	East Asia & Pacific	82.11
Solomon Islands	East Asia & Pacific	77.86

4. RECOMMENDATIONS

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

- *What have you learned from the World Bank data?*

Between 1990 and 2016, the world's total forest area decreased by 3.21%. This is alarming given that deforestation is a contributing factor to climate change. Although countries such as Iceland, French Polynesia, Bahrain, and Uruguay increased their forest area by over 100%, given their small sizes in total land area, these increases won't make a dent in the crisis we are facing. On a bright note, China increased its forest area by 33.55%, a significant and vital increase given its large size. The United States, on the other hand, only increased its forest area by 2.62%, a disappointing figure given its influence in the world.

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

In order to reverse the trajectory that we're on, we should focus on countries that are larger in total land area since a significant increase in forestation in these countries will have the greatest impact on climate change and the loss of wildlife.

5. APPENDIX: SQL queries used

```
/* create view called forestation */
CREATE VIEW forestation AS
        SELECT fa.country_code AS country_code,
                fa.country_name AS country_name,
                fa.year AS "year",
                fa.forest_area_sqkm,
                la.total_area_sq_mi,
                r.region,
                r.income_group,
                ROUND(((fa.forest_area_sqkm / (la.total_area_sq_mi * 2.59)) * 100)::numeric, 2)  AS
forest_percentage
          FROM forest_area AS fa
                INNER JOIN land_area AS la
                ON fa.country_code = la.country_code
                    AND fa.year = la.year

                INNER JOIN regions r
                ON la.country_code = r.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 f
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 f
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?

```
SELECT (fa_1990 - fa_2016) AS world_forest_area_change
  FROM
            (SELECT country_code,
                        forest_area_sqkm AS fa_1990
                FROM forestation AS f1
              WHERE f1.country_name = 'World' AND f1.year = 1990) wfa_1990

 INNER JOIN
          (SELECT country_code,
                      orest_area_sqkm fa_2016
              FROM  forestation AS f2
             WHERE f2.country_name = 'World'
                 AND f2.year = 2016) wfa_2016
ON wfa_1990.country_code = wfa_2016.country_code;
```

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

```
SELECT ROUND((((fa_1990 - fa_2016)/fa_1990) * 100)::numeric, 2) || '%'
            AS world_forest_area_percentage_loss
  FROM
            (SELECT country_code,
                        forest_area_sqkm AS fa_1990
                FROM forestation AS f1
              WHERE f1.country_name = 'World'
                  AND f1.year = 1990) wfa_1990
 INNER JOIN
            (SELECT country_code,
                        forest_area_sqkm fa_2016
                FROM forestation AS f2
              WHERE f2.country_name = 'World'
                  AND f2.year = 2016) wfa_2016
 ON wfa_1990.country_code = wfa_2016.country_code;
```

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 country_name AS country,
            total_area_sqkm,
            diff_between_world_forest_area_loss_and_country_total_area_sqkm
  FROM
           (SELECT DISTINCT country_name,
                       ROUND((total_area_sq_mi * 2.59)::numeric, 2) AS total_area_sqkm,
                                          ROUND(ABS((total_area_sq_mi * 2.59) -
(t1.world_forest_area_change_sqkm))::numeric, 2) AS
diff_between_world_forest_area_loss_and_country_total_area_sqkm
        FROM
                 (SELECT (fa_1990 - fa_2016) AS world_forest_area_change_sqkm
                    FROM
                          (SELECT country_name,
                                      forest_area_sqkm AS fa_1990
                             FROM forestation AS f1
                            WHERE f1.country_name = 'World'
                                AND f1.year = 1990) wfa_1990
                      INNER JOIN
                          (SELECT country_name,
                                      forest_area_sqkm AS fa_2016
                             FROM  forestation AS f2
                            WHERE  f2.country_name = 'World'
                                AND  f2.year = 2016) wfa_2016
                       ON wfa_1990.country_name = wfa_2016.country_name) t1
  INNER JOIN forestation f
              ON country_name = f.country_name
        WHERE country_name NOT LIKE 'World'
         ORDER BY 1 DESC) t1

 WHERE diff_between_world_forest_area_loss_and_country_total_area_sqkm =
                               (SELECT MIN(diff_between_world_forest_area_loss_and_country_total_area_sqkm)
                                  FROM
                                        (SELECT DISTINCT country_name,
                                                            ROUND(ABS((total_area_sq_mi * 2.59) -
(t1.world_forest_area_change_sqkm))::numeric, 2) AS
diff_between_world_forest_area_loss_and_country_total_area_sqkm
                                                   FROM
                                                         (SELECT (fa_1990 - fa_2016) AS
world_forest_area_change_sqkm
                                                            FROM
                                                                  (SELECT country_name,
                                                                              forest_area_sqkm AS fa_1990
                                                                     FROM forestation AS f1
                                                                    WHERE f1.country_name = 'World'
                                                                        AND f1.year = 1990) wfa_1990
                                                              INNER JOIN
                                                                   (SELECT country_name,
                                                                               forest_area_sqkm AS fa_2016
                                                                      FROM  forestation AS f2
                                                                     WHERE  f2.country_name = 'World'
                                                                         AND  f2.year = 2016) wfa_2016
                                                           ON wfa_1990.country_name = wfa_2016.country_name) t1
                                        INNER JOIN forestation f
                                                ON country_name = f.country_name
                                      WHERE country_name NOT LIKE 'World'
                                       ORDER BY 1 DESC)  t1);
```

2. REGIONAL OUTLOOK

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

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

Query for Table 2.1: Percent Forest Area by Region, 1990, 2016:

```
SELECT t1.region,
           t1.forest_percentage_1990,
           t2.forest_percentage_2016
  FROM
           (SELECT region,
                       ROUND(((SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59)) * 100)::numeric, 2) AS
forest_percentage_1990
                  FROM forestation f1
              WHERE year = 1990
              GROUP BY 1
              ORDER BY 2 DESC) t1

   INNER JOIN

             (SELECT region,
                       ROUND(((SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59)) * 100)::numeric, 2) AS
forest_percentage_2016
                  FROM forestation f2
              WHERE year = 2016
              GROUP BY 1
              ORDER BY 2 DESC) t2
   ON t1.region = t2.region
GROUP BY 1, 2, 3
ORDER BY 2 DESC, 3 DESC;
```

Which region had the HIGHEST percent forest in 2016, and which had the LOWEST, to 2 decimal places?

```
SELECT f1.region,
        f1.year,
        ROUND(((SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59)) * 100)::numeric, 2) AS
highest_and_lowest_percent_forest
  FROM forestation f1
GROUP BY 1, 2
HAVING ROUND(((SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59)) * 100)::numeric, 2) =
        (SELECT MAX(t1.forest_percent_by_region_2016) AS highest_percent_forest
        FROM
                (SELECT region,
                        ROUND(((SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59)) *
100)::numeric, 2) AS forest_percent_by_region_2016
                   FROM forestation
                  WHERE year = 2016
                  GROUP BY 1
                  ORDER BY 2 DESC)  t1)
    AND f1.year = 2016

UNION

SELECT f1.region,
        f1.year,
        ROUND(((SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59)) * 100)::numeric, 2) AS
highest_and_lowest_percent_forest
  FROM  forestation f1
 GROUP BY 1, 2
HAVING ROUND(((SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59)) * 100)::numeric, 2) =
        (SELECT MIN(t1.forest_percent_by_region_2016) AS highest_percent_forest
           FROM
                (SELECT region,
                        ROUND(((SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59)) *
100)::numeric, 2) AS forest_percent_by_region_2016
                   FROM forestation
                  WHERE year = 2016
                  GROUP BY 1
                  ORDER BY 2 DESC)  t1)
    AND f1.year = 2016;
```

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

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

Which region had the HIGHEST percent forest in 1990, and which had the LOWEST, to 2 decimal places?

```
SELECT f1.region,
            f1.year,
            ROUND(((SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59)) * 100)::numeric, 2) AS
highest_and_lowest_percent_forest
  FROM forestation f1
 GROUP BY 1, 2
 HAVING ROUND(((SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59)) * 100)::numeric, 2) =
              (SELECT MAX(t1.forest_percent_by_region_1990) AS highest_percent_forest
                 FROM
                              (SELECT region,
                               ROUND(((SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59)) * 100)::numeric,
2) AS forest_percent_by_region_1990
                                 FROM forestation
                               WHERE year = 1990
                               GROUP BY 1
                               ORDER BY 2 DESC)  t1)
                                    AND f1.year = 1990

UNION

SELECT f1.region,
            f1.year,
            ROUND(((SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59)) * 100)::numeric, 2) AS
highest_and_lowest_percent_forest
  FROM forestation f1
 GROUP BY 1, 2
 HAVING ROUND(((SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59)) * 100)::numeric, 2) =
             (SELECT MIN(t1.forest_percent_by_region_1990) AS highest_percent_forest
                FROM
                            (SELECT region,
                                        ROUND(((SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59)) *
100)::numeric, 2) AS forest_percent_by_region_1990
                               FROM forestation
                             WHERE year = 1990
                             GROUP BY 1
                             ORDER BY 2 DESC)  t1)
        AND f1.year = 1990;
```

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

```
SELECT t3.region,
            t3.forest_percentage_1990,
            t3.forest_percentage_2016
   FROM

          (SELECT t1.region,
                     t1.forest_percentage_1990,
                     t2.forest_percentage_2016
              FROM
                     (SELECT region,
                                ROUND(((SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59)) *
100)::numeric, 2) AS forest_percentage_1990
                        FROM forestation f1
                      WHERE year = 1990
                      GROUP BY 1
                      ORDER BY 2 DESC) t1

   INNER JOIN

             (SELECT region,
                        ROUND(((SUM(forest_area_sqkm) / SUM(total_area_sq_mi * 2.59)) * 100)::numeric, 2)
AS forest_percentage_2016
                   FROM forestation f2
                WHERE year = 2016
                GROUP BY 1
                ORDER BY 2 DESC) t2
   ON t1.region = t2.region) t3
WHERE region NOT LIKE 'World' AND t3.forest_percentage_1990 > t3.forest_percentage_2016
GROUP BY 1, 2, 3
ORDER BY 2 DESC, 3 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?

Query for Table 3.1 Top 5 Amount Decrease in Forest Area by Country, 1990, & 2016:

```
SELECT   country_fa_1990.country_name,
         country_fa_1990.region,
         (country_fa_1990.fa_1990 - country_fa_2016.fa_2016) AS absolute_forest_area_change_sqkm
  FROM
        (SELECT country_name,
                region,
                 ROUND(SUM(forest_area_sqkm)) AS fa_1990
           FROM forestation AS f1
          WHERE year = 1990
              AND region NOT LIKE 'World'
              AND forest_area_sqkm IS NOT NULL
          GROUP BY 2, 1
          ORDER BY 3 DESC, 1, 2) country_fa_1990
  INNER JOIN
          (SELECT country_name,
                  region,
                  ROUND(SUM(forest_area_sqkm)) AS fa_2016
             FROM forestation AS f2
            WHERE year = 2016
                AND region NOT LIKE 'World'
                AND forest_area_sqkm IS NOT NULL
            GROUP BY 2, 1
            ORDER BY 3 DESC, 1, 2) country_fa_2016
 ON country_fa_1990.country_name = country_fa_2016.country_name
 WHERE (country_fa_1990.fa_1990 - country_fa_2016.fa_2016) > 0
 GROUP BY 2, 1, 3
 ORDER BY 3 DESC, 2, 1
 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?

Query for Table 3.1 Top 5 Percent Decrease in Forest Area by Country, 1990, & 2016:

```
SELECT  wfa_1990.country_name AS country,
            ROUND((((fa_1990 - fa_2016)/fa_1990) * 100)::numeric, 2) AS pct_forest_area_change
  FROM
            (SELECT country_name,
                        forest_area_sqkm AS fa_1990
               FROM forestation AS f1
              WHERE f1.country_name NOT LIKE 'World'
                  AND forest_area_sqkm IS NOT NULL
                  AND f1.year = 1990) wfa_1990
 INNER JOIN
           (SELECT country_name,
                        forest_area_sqkm fa_2016
              FROM forestation AS f2
             WHERE f2.country_name NOT LIKE 'World'
                 AND forest_area_sqkm IS NOT NULL
                 AND f2.year = 2016) wfa_2016
ON wfa_1990.country_name = wfa_2016.country_name
ORDER BY 2 DESC
LIMIT 5;
```

Query for Table 3.3 Count of Countries Grouped by Forestation Percent Quartiles, 2016:

```
SELECT quartile,
           COUNT(*) AS number_of_countries
 FROM
           (SELECT  country_name,
                        forest_percentage,
                        CASE
                                 WHEN forest_percentage BETWEEN 0 AND 25 THEN '0% - 25%'
                                WHEN forest_percentage BETWEEN 25 AND 50 THEN '25% - 50%'
                                 WHEN forest_percentage BETWEEN 50 AND 75 THEN '50% - 75%'
                                 ELSE '75% - 100%'
                           END AS quartile
              FROM forestation
             WHERE year = '2016'
                 AND country_name NOT LIKE 'World'
                 AND forest_percentage IS NOT NULL
             GROUP BY 2, 1
             ORDER BY 2 DESC) sub
 GROUP BY 1
 ORDER BY 2 DESC;
```

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

```
SELECT quartile,
        COUNT(*) AS number_of_countries
 FROM
        (SELECT  country_name,
                 forest_percentage,
                 CASE
    WHEN forest_percentage BETWEEN 0 AND 25 THEN '0% - 25%'
                      WHEN forest_percentage BETWEEN 25 AND 50 THEN '25% - 50%'
                       WHEN forest_percentage BETWEEN 50 AND 75 THEN '50% - 75%'
                       ELSE '75% - 100%'
                  END AS quartile
           FROM forestation
          WHERE year = '2016'
             AND country_name NOT LIKE 'World'
             AND forest_percentage IS NOT NULL
          GROUP BY 2, 1
          ORDER BY 2 DESC) sub
 GROUP BY 1
 ORDER BY 2 DESC
   LIMIT 1;
```

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

```
SELECT country_name,
        region,
        forest_percentage AS pct_designated_as_forest,
        quartile
  FROM
        (SELECT  country_name,
                 region,
                 forest_percentage,
                 CASE
                      WHEN forest_percentage BETWEEN 0 AND 25 THEN '0%-25%'
                      WHEN forest_percentage BETWEEN 25 AND 50 THEN '25%-50%'
                       WHEN forest_percentage BETWEEN 50 AND 75 THEN '50%-75%'
                       ELSE '75%-100%'
                       END AS quartile
           FROM forestation
          WHERE year = '2016'
             AND country_name NOT LIKE 'World'
             AND forest_percentage IS NOT NULL
          ORDER BY 4 DESC, 3 DESC, 2, 1)  t1
 WHERE quartile = '75%-100%'
 GROUP BY 4, 2, 1, 3
 ORDER BY 3 DESC, 2 , 1;
```

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

```
SELECT COUNT(*) AS number_of_countries_with_percent_forestation_highter_than_United_States_in_2016
  FROM
            (SELECT t1.country_name,
                        t1.forest_percentage
                FROM
                        (SELECT  country_name,
                                    forest_percentage
                            FROM forestation
                           WHERE year = '2016'
                               AND country_name NOT LIKE 'World'
                               AND forest_percentage IS NOT NULL
                            ORDER BY 2 DESC, 1) t1
 GROUP BY 1, 2
 HAVING t1.forest_percentage >
                                          (SELECT forest_percentage
                                              FROM forestation
                                            WHERE country_name = 'World'
                                                AND year = 2016)
                                           ORDER BY 2) t2;
```