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 41,282,694.90 sq kilometers in 1990. As of 2016, the most recent year for which data was available, that number had fallen to 39,958,245.90 sq kilometers, a loss of 1,324,449 sq 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 494,208.49 sq miles)

2. **REGIONAL OUTLOOK**

In 2016, the percent of the total land area of the world designated as forest was 31%. The region with the highest relative forestation was Latin America/Caribbean with 46% and the region with the lowest relative forestation was the Middle East/North Africa with 2% forestation.

In 1990, the percent of the total land area of the world designated as forest was 32%. The region with the highest relative forestation was Latin America/Caribbean, with 52%, and the region with the lowest relative forestation was Middle East/North Africa, with 2% forestation.

Table 2.1: Percent Forest Area by Region, 1990 & 2016:

Region	1990 Forest Percentage	2016 Forest Percentage
Latin America/Caribbean	51%	46%
Europe & Central Asia	37%	38%
North America	36%	36%
Sub-Saharan Africa	31%	29%
East Asia/Pacific	26%	26%
South Asia	17%	18%
Middle East/North Africa	2%	2%

The only regions of the world that decreased in percent forest area from 1990 to 2016 were Latin America/Caribbean (dropped from 51% to 46%) and Sub-Saharan Africa (31% to 29%). 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% to 31%.

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 sqkm, or 33.5%. 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 2.6% (79,200 sqkm) 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
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	-45.03
Honduras	Latin America / Caribbean	-43.45

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 Honduras, which is in the Latin American / 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	73
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
Suriname	Latin America / Caribbean	98.26%
Micronesia	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%

5. RECOMMENDATIONS

The Latin America / Caribbean region saw by far the greatest deforestation between 1990 and 2016, during which time total forestation actually increased in every other region aside from Sub-Saharan Africa, where rates of deforestation by country lead the world.

Sub-Saharan African countries appear to be at great risk of future deforestation, as evidenced by their leading rates of deforestation as a percentage of total area. Taking lessons from success stories such as China, the United States, and considering the contributing factors which led to the deforestation of Latin America and the Caribbean could provide clues to mitigate Sub-Saharan African deforestation.

Taking into account total deforestation in addition to percentage of deforestation, there are 9 countries to focus on. They are: Brazil, Indonesia, Myanmar, Nigeria, Tanzania, Togo, Uganda, Mauritania, and Honduras.

6. APPENDIX: SQL QUERIES USED

Step 1: Create Forestation View

```
CREATE VIEW forestation AS

(SELECT f.country_code,
                f.country_name,
                r.region,
                r.income_group,
                f.year,
                CAST(f.forest_area_sqkm AS NUMERIC),
                CAST(l.total_area_sq_mi AS NUMERIC),
                CAST((f.forest_area_sqkm / 2.59 / l.total_area_sq_mi)
                AS NUMERIC)*100 AS pct_forest
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 l.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 SUM(forest_area_sqkm) AS total_forest_area_sqkm
FROM forestation
WHERE country_name = 'World' AND year = 1990;
```

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 SUM(forest_area_sqkm) AS total_forest_area_sqkm
FROM forestation
WHERE country_name = 'World' AND year = 2016;
```

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

```
with q1 AS (
                SELECT SUM(forest_area_sqkm) AS total_forest_area_sqkm
                FROM forestation
                WHERE country_name = 'World' AND year = 1990
        ),
q2 AS (
        SELECT SUM(forest_area_sqkm) AS total_forest_area_sqkm
        FROM forestation
        WHERE country_name = 'World' AND year = 2016
        )
SELECT q1.total_forest_area_sqkm - q2.total_forest_area_sqkm AS change
FROM q1, q2;
```

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

```
SELECT
            f1.region,
            f1.year AS year1,
            f2.year AS year2,
            ROUND(SUM(f1.forest_area_sqkm),2) AS forest_1990,
            ROUND(SUM(f2.forest_area_sqkm),2) AS forest_2016,
            ROUND(((SUM(f2.forest_area_sqkm) - SUM(f1.forest_area_sqkm)) /
            SUM(f1.forest_area_sqkm))*100,2) AS pct_change
FROM forestation AS f1
JOIN forestation AS f2
ON f1.region = f2.region
GROUP BY 1,2,3
HAVING f1.region = 'World' AND f1.year = 1990 AND f2.year = 2016;
```

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,
           total_area_sq_mi * 2.59 AS land_area_km,
           ABS((total_area_sq_mi * 2.59) – 1324449 AS diff
FROM forestation
WHERE year = 2016 AND year IS NOT NULL
ORDER BY 3
LIMIT 1;
```

2. REGIONAL OUTLOOK

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

```
SELECT region,
        ROUND(CAST(SUM(forest_area_sqkm / 2.59) / SUM(total_area_sq_mi)
        AS NUMERIC),2)*100 AS pct_forest
FROM forestation
WHERE year = 2016
GROUP BY 1
ORDER BY 2 DESC;
```

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?

World:

```
SELECT region,
        ROUND(CAST(SUM(forest_area_sqkm / 2.59) / SUM(total_area_sq_mi)
        AS NUMERIC),2)*100 AS pct_forest
FROM forestation
WHERE year = 2016 AND region = 'World'
GROUP BY 1;
```

Highest:

```
SELECT region,
        ROUND(CAST(SUM(forest_area_sqkm / 2.59) / SUM(total_area_sq_mi)
        AS NUMERIC),2)*100 AS pct_forest
FROM forestation
WHERE year = 2016
GROUP BY 1
ORDER BY 2 DESC
LIMIT 1;
```

Lowest:

```
SELECT region,
        ROUND(CAST(SUM(forest_area_sqkm / 2.59) / SUM(total_area_sq_mi)
        AS NUMERIC),2)*100 AS pct_forest
FROM forestation
WHERE year = 2016
GROUP BY 1
ORDER BY 2
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?

```
SELECT region,
        ROUND(CAST(SUM(forest_area_sqkm / 2.59) / SUM(total_area_sq_mi)
        AS NUMERIC),2)*100 AS pct_forest
FROM forestation
WHERE year = 1990 AND region = 'World'
GROUP BY 1;
```

Highest:

```
SELECT region,
        ROUND(CAST(SUM(forest_area_sqkm / 2.59) / SUM(total_area_sq_mi)
        AS NUMERIC),2)*100 AS pct_forest
FROM forestation
WHERE year = 1990
GROUP BY 1
ORDER BY 2 DESC
LIMIT 1;
```

Lowest:

```
SELECT region,
        ROUND(CAST(SUM(forest_area_sqkm / 2.59) / SUM(total_area_sq_mi)
        AS NUMERIC),2)*100 AS pct_forest
FROM forestation
WHERE year = 1990
GROUP BY 1
ORDER BY 2
LIMIT 1;
```

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

```
with s1 AS (
       SELECT region,
                    CAST(SUM(forest_area_sqkm / 2.59) / SUM(total_area_sq_mi)
                    AS NUMERIC)*100 AS pct_forest
       FROM forestation
       WHERE year = 2016
       GROUP BY 1
       ORDER BY 2 DESC
     ),
    s2 AS (
      SELECT region,
                  CAST(SUM(forest_area_sqkm / 2.59) / SUM(total_area_sq_mi)
                  AS NUMERIC)*100 AS pct_forest
      FROM forestation
      WHERE year = 1990
      GROUP BY 1
      ORDER BY 2 DESC
    )
SELECT s1.region,
            ROUND(s2.pct_forest,2) AS pct_forest_1990,
            ROUND(s1.pct_forest,2) AS pct_forest_2016,
            ROUND((s1.pct_forest - s2.pct_forest),2) AS pct_change
FROM s1
JOIN s2
ON s1.region = s2.region
WHERE (s1.pct_forest - s2.pct_forest) < 0
ORDER BY 4;
```

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 s1 AS (
        SELECT country_name,
                ROUND(forest_area_sqkm,2) AS forest_1990
        FROM forestation
        WHERE year = 1990
),
s2 AS (
        SELECT country_name,
                ROUND(forest_area_sqkm,2) AS forest_2016
        FROM forestation
        WHERE year = 2016
)

SELECT s1.country_name,
          s1.forest_1990,
          s2.forest_2016,
         (s2.forest_2016 - s1.forest_1990) AS change
FROM s1
JOIN s2
USING (country_name)
WHERE (s1.forest_1990 - s2.forest_2016) IS NOT NULL AND country_name != 'World'
ORDER BY 4
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 s1 AS (
        SELECT country_name,
                region,
                forest_area_sqkm,
                CAST(((forest_area_sqkm / 2.59) / total_area_sq_mi)
                AS NUMERIC)*100 AS pct_forest_1990
        FROM forestation
        WHERE year = 1990
),

s2 AS (
        SELECT country_name,
                forest_area_sqkm,
                CAST(((forest_area_sqkm / 2.59) / total_area_sq_mi)
                AS NUMERIC)*100 AS pct_forest_2016
        FROM forestation
        WHERE year = 2016
)

SELECT s1.country_name,
          s1.region,
          ROUND(s1.pct_forest_1990,2) AS pct_forest_1990,
          ROUND(s2.pct_forest_2016,2) AS pct_forest_2016,
          ROUND((s2.pct_forest_2016 - s1.pct_forest_1990),2) AS pct_change
FROM s1
JOIN s2
USING (country_name)
WHERE (s2.pct_forest_2016 - s1.pct_forest_1990) IS NOT NULL
ORDER BY 5
LIMIT 5;
```

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

```
with s1 AS (
                SELECT country_name,
                        forest_area_sqkm,
                        ROUND(CAST(((forest_area_sqkm / 2.59) / total_area_sq_mi)
                        AS NUMERIC)*100,2) AS pct_forest_2016
                FROM forestation
                WHERE CAST(((forest_area_sqkm / 2.59) / total_area_sq_mi)
                AS NUMERIC) IS NOT NULL AND year = 2016
                ORDER BY 3 DESC),
s2 AS (
    SELECT CASE WHEN pct_forest_2016 > 75 THEN 4
                        WHEN pct_forest_2016 BETWEEN 50.01 AND 75
                        THEN 3
                        WHEN pct_forest_2016 BETWEEN 25.01 AND 50
                        THEN 2
                        ELSE 1 END AS quartile
    FROM s1)

SELECT quartile,
        COUNT(*)
FROM s2
GROUP BY 1
ORDER BY 2 DESC;
```

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

```
with s1 AS (
            SELECT country_name,
                        region,
                        forest_area_sqkm,
                        ROUND(CAST(((forest_area_sqkm / 2.59) / total_area_sq_mi)
                        AS NUMERIC)*100,2) AS pct_forest_2016
            FROM forestation
           WHERE CAST(((forest_area_sqkm / 2.59) / total_area_sq_mi)
           AS NUMERIC) IS NOT NULL AND year = 2016
           ORDER BY 3 DESC),
s2 AS (
            SELECT country_name,
                       CASE WHEN pct_forest_2016 > 75 THEN 4
                                WHEN pct_forest_2016 BETWEEN 50.01 AND 75 THEN 3
                                WHEN pct_forest_2016 BETWEEN 25.01 AND 50 THEN 2
                                ELSE 1 END AS quartile
            FROM s1)

SELECT s2.country_name,
            s1.region,
            s1.pct_forest_2016
FROM s2
JOIN s1
ON s2.country_name = s1.country_name
WHERE quartile = 4
ORDER BY 3 DESC;
```

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

```
with s1 AS (
          SELECT ROUND(CAST(((forest_area_sqkm / 2.59) / total_area_sq_mi)
          AS NUMERIC)*100,2) AS pct_forest_2016
          FROM forestation
          WHERE CAST(((forest_area_sqkm / 2.59) / total_area_sq_mi)
          AS NUMERIC) IS NOT NULL AND year = 2016
                              ),
  s2 AS (
           SELECT ROUND(CAST(((forest_area_sqkm / 2.59) / total_area_sq_mi)
              AS NUMERIC)*100,2) AS pct_forest_2016
        FROM forestation
        WHERE CAST(((forest_area_sqkm / 2.59) / total_area_sq_mi)
        AS NUMERIC) IS NOT NULL AND year = 2016
        AND country_name = 'United States')

SELECT COUNT(*)
FROM s1
WHERE pct_forest_2016 > (SELECT * FROM s2);
```