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 in 1990. As of 2016, the most recent year for which data was available, that number had fallen to 39958245.9, 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 1279999.9).

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	36.03
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 & Carribean (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. 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, 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
Indonesia	East Asia & Pacific	282194
Myanmar	East Asia & Pacific	107234
Nigeria	Sub-Saharan Africa	106506
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 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
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
American Samoa	East Asia & Pacific	87.50
Gabon	Sub-Saharan Africa	90.04
Guyana	Latin America & Caribbean	83.90
Lao PDR	East Asia & Pacific	82.11
Micronesia, Fed. Sts.	East Asia & Pacific	91.86
Palau	East Asia & Pacific	87.61
Seychelles	Sub-Saharan Africa	88.41
Solomon Islands	East Asia & Pacific	77.86
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?*
- *Which countries should we focus on over others?*

Based on the information gathered the world's forest area has decreased 3.21% between 1900 and 2016. While the majority of regions have actually increased their forest area, the Latin American & Caribbean and Sub-Saharan Africa regions decreased so much that it impacted the global situation of this time period.

We would want to focus our resources on helping increase forestation in the areas we see as our largest concerns. Specific countries in the Latin America & Caribbean region we should be focused on are Brazil, which decreased forest area by 541510sqkm, and Honduras, which decreased forest area by 45.03%. In the Sub-Saharan Africa region Togo decreaseed forest area by 75.45%, Nigeria by 61.80%, Uganda by 59.13% and Mauritania by 46.75%. Indonesia and Myanmar in the East Asia & Pacific region are also areas of concern as they were 2nd and 3rd in Countries that had the most decrease in forest area between 1990 and 2016.

To be able to best help these countries we would want to look at the initiatives and policies in countries that are sucessful in their forestation efforts. Looking at countries like China which had the largest increase in forest area by 527229sqkm or the United States which increased by 79200sqkm could be helpful. We can also look at the countries which had the largest percentage increase such as Iceland which increased by 213.66%, French Polynesia by 181.82%, Bahrain by 177.27%, and Uruguay by 134.11%. While these countries are smaller, the percentages of the increased forestation indicate that they may have helpful insights to combat deforestation in other countries. Reviewing efforts in the 9 countries that make up the 75%-100% Quartile in 2016 may also be beneficial.

5. APPENDIX: SQL Queries Used

```
DROP VIEW IF EXISTS forestation;
CREATE VIEW forestation
AS
	(SELECT f.country_code,
		f.country_name,
		f.year,
		f.forest_area_sqkm,
		l.total_area_sq_mi,
		l.total_area_sq_mi*2.59 AS total_area_sq_km,
		r.region,
		r.income_group,
		ROUND((f.forest_area_sqkm/(l.total_area_sq_mi *2.59)*100)::NUMERIC,2)
			AS percentage_forest
	FROM forest_area AS f
	INNER JOIN land_area AS l
		ON f.country_code = l.country_code
		AND f.year = l.year
	INNER JOIN regions AS r
		ON r.country_code = f.country_code);
```

GLOBAL SITUATION

1a. **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 AS fa_1990
FROM forestation
WHERE  year=1990 AND country_name = 'World';
```

1b. **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 AS fa_2016
FROM forestation
WHERE  year=2016 AND country_name = 'World';
```

Or this can be queried as a SELF JOIN

```
SELECT a.country_name AS name1990,
          a.forest_area_sqkm AS fa1990,
          b.country_name AS name2016,
          b.forest_area_sqkm AS fa2016
FROM forestation AS a,
        forestation AS b
WHERE a.country_name = 'World' AND b.country_name = 'World'
AND a.year = 1990 AND b.year = 2016
```

1c.**What was the change (in sq km) in the forest area of the world from 1990 to 2016?**

```
WITH forest_area_1990 AS
        (SELECT forest_area_sqkm AS fa_1990
         FROM forestation
        WHERE  year=1990 AND country_name = 'World'),
 forest_area_2016 AS
        (SELECT forest_area_sqkm AS fa_2016
        FROM forestation
        WHERE  year=2016 AND country_name = 'World')
SELECT
        (SELECT fa_1990 FROM forest_area_1990) AS fa_1990,
        (SELECT fa_2016 FROM forest_area_2016) AS fa_2016,
        (SELECT fa_2016 FROM forest_area_2016)
        -(SELECT fa_1990 FROM forest_area_1990)
      AS difference;
```

1d. **What was the percent change in forest area of the world between 1990 and 2016?**

```
WITH forest_area_1990 AS
        (SELECT forest_area_sqkm AS fa_1990
         FROM forestation
        WHERE  year=1990 AND country_name = 'World'),
 forest_area_2016 AS
        (SELECT forest_area_sqkm AS fa_2016
        FROM forestation
        WHERE  year=2016 AND country_name = 'World')
SELECT
        (SELECT fa_1990 FROM forest_area_1990) AS fa_1990,
        (SELECT fa_2016 FROM forest_area_2016) AS fa_2016,
        ((SELECT fa_2016 FROM forest_area_2016)-(SELECT fa_1990 FROM
        forest_area_1990))*100/(SELECT fa_1990 FROM forest_area_1990)
      AS percentage_change;
```

1e. **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 DISTINCT country_name, total_area_sq_km
FROM forestation
WHERE total_area_sq_km BETWEEN 1270000 AND 1350000;
```

REGIONAL OUTLOOK

```
CREATE VIEW regional
AS
    (SELECT r.region,
            l.year,
            SUM(f.forest_area_sqkm) AS total_fa_sqkm,
            SUM(l.total_area_sq_mi*2.59) AS total_a_sqkm,
            (SUM(f.forest_area_sqkm)/SUM(l.total_area_sq_mi*2.59))*100 AS per_fa_region
     FROM forest_area f
     INNER JOIN land_area l
     ON f.country_code = l.country_code AND f.year = l.year
     INNER JOIN regions r
     ON l.country_code = r.country_code
     GROUP BY 1,2
     ORDER BY 1,2);
```

2a**. <u>What was the percent forest of the entire world in 2016?</u>**

```
SELECT ROUND(per_fa_region:: NUMERIC,2) AS pfa_region
FROM regional
WHERE year = 2016 AND region = 'World';
```

2a. **<u>Which region had the HIGHEST percent forest in 2016, and which had the LOWEST, to 2 decimal places?</u>**

```
SELECT region,
        ROUND(total_a_sqkm::NUMERIC,2) AS total_a_sqkm,
        ROUND(per_fa_region::NUMERIC,2) AS pfa_region
FROM regional
WHERE ROUND(per_fa_region::NUMERIC,2) =
      (SELECT MAX(ROUND(per_fa_region::NUMERIC,2)) AS max_per
       FROM regional
       WHERE year = 2016)
        AND year = 2016;

SELECT region,
        ROUND(total_a_sqkm::NUMERIC,2) AS total_a_sqkm,
        ROUND(per_fa_region::NUMERIC,2) AS pfa_region
FROM regional
WHERE ROUND(per_fa_region::NUMERIC,2) =
      (SELECT MIN(ROUND(per_fa_region::NUMERIC,2)) AS min_per
       FROM regional
       WHERE year = 2016)
       AND year = 2016;
```

2b. **What was the percent forest of the entire world in 1990?**

```
SELECT ROUND(per_fa_region::NUMERIC,2) AS pfa_region
FROM regional
WHERE year = 1990 AND region = 'World';
```

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

```
SELECT region,
            ROUND(total_a_sqkm::NUMERIC,2) AS total_a_sqkm,
            ROUND(per_fa_region::NUMERIC,2) AS pfa_region
FROM regional
WHERE ROUND(per_fa_region::NUMERIC,2) =
          (SELECT MAX(ROUND(per_fa_region::NUMERIC,2)) AS max_per
           FROM regional
           WHERE year = 1990)
      AND year = 1990;
```

```
SELECT region,
            ROUND(total_a_sqkm::NUMERIC,2) AS total_a_sqkm,
            ROUND(per_fa_region::NUMERIC,2) AS pfa_region
FROM regional
WHERE ROUND(per_fa_region::NUMERIC,2) =
          (SELECT MIN(ROUND(per_fa_region::NUMERIC,2)) AS min_per
           FROM regional
           WHERE year = 1990)
      AND year = 1990;
```

Percentage Chart 1990 & 2016

```
WITH t90 AS (SELECT * FROM regional WHERE year = 1990),
        t16 AS (SELECT * FROM regional WHERE year = 2016)
SELECT t90.region,
            ROUND(t90.per_fa_region::NUMERIC,2) AS fa90,
            ROUND(t16.per_fa_region::NUMERIC,2) AS fa16
FROM t90
INNER JOIN t16
ON t90.region = t16.region
ORDER BY fa90 DESC;
```

2c**. Based on the table you created, which regions of the world DECREASED in forest area from 1990 to 2016?**

```
WITH t90 AS (SELECT * FROM regional WHERE year = 1990),
        t16 AS (SELECT * FROM regional WHERE year = 2016)
SELECT t90.region,
            ROUND(t90.per_fa_region::NUMERIC,2) AS fa90,
            ROUND(t16.per_fa_region::NUMERIC,2) AS fa16
 FROM t90
 INNER JOIN t16
 ON t90.region = t16.region
 WHERE t90.per_fa_region > t16.per_fa_region;
```

COUNTRY LEVEL DETAIL

--**SUCCESS forest area**

```
 WITH t90 AS
        (SELECT f.country_code,
                f.country_name,
                f.year,
                f.forest_area_sqkm
                FROM forest_area f
                WHERE f.year = 1990 AND f.forest_area_sqkm IS NOT NULL
                AND f.country_name != 'World'),
 t16 AS
        (SELECT f.country_code,
                f.country_name,
                f.year,
                f.forest_area_sqkm
                FROM forest_area f
                WHERE f.year = 2016 AND f.forest_area_sqkm IS NOT NULL
                AND f.country_name != 'World')
SELECT t90.country_code,
          t90.country_name,
          r.region,
          t90.forest_area_sqkm AS fa_90_sqkm,
          t16.forest_area_sqkm AS fa_16_sqkm,
          t16.forest_area_sqkm - t90.forest_area_sqkm AS diff_fa
FROM t90
INNER JOIN t16
ON t90.country_code = t16.country_code
AND (t90.forest_area_sqkm IS NOT NULL AND t16.forest_area_sqkm IS NOT NULL)
INNER JOIN regions r ON t16.country_code = r.country_code
ORDER BY 6 DESC
LIMIT 5;
```

--SUCCESS percent

```
WITH t90 AS
        (SELECT f.country_code,
                f.country_name,
                f.year,
                f.forest_area_sqkm
                FROM forest_area f
                WHERE f.year = 1990 AND f.forest_area_sqkm IS NOT NULL
                AND f.country_name != 'World'),
 t16 AS
        (SELECT f.country_code,
                f.country_name,
                f.year,
                f.forest_area_sqkm
                FROM forest_area f
                WHERE f.year = 2016 AND f.forest_area_sqkm IS NOT NULL
                AND f.country_name != 'World')
SELECT t90.country_code,
          t90.country_name,
          r.region,
          t90.forest_area_sqkm AS fa_90_sqkm,
          t16.forest_area_sqkm AS fa_16_sqkm,
          ROUND((((t16.forest_area_sqkm - t90.forest_area_sqkm) /
        t90.forest_area_sqkm)*100)::NUMERIC, 2) AS per_diff
FROM t90
INNER JOIN t16
ON t90.country_code = t16.country_code
AND (t90.forest_area_sqkm IS NOT NULL AND t16.forest_area_sqkm IS NOT NULL)
INNER JOIN regions r
ON t16.country_code = r.country_code
ORDER BY 6 DESC
LIMIT 5;
```

3a. **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 T90 AS
        (SELECT f.country_code,
                f.country_name,
                f.year,
                f.forest_area_sqkm
                FROM forest_area f
               WHERE f.year = 1990 AND f.forest_area_sqkm IS NOT NULL
                AND f.country_name != 'World'),
 t16 AS
        (SELECT f.country_code,
                 f.country_name,
                 f.year,
                 f.forest_area_sqkm
                 FROM forest_area f
                WHERE f.year = 2016 AND f.forest_area_sqkm IS NOT NULL
                AND f.country_name != 'World')
SELECT t90.country_code,
          t90.country_name,
          r.region,
          t90.forest_area_sqkm AS fa_90_sqkm,
          t16.forest_area_sqkm AS fa_16_sqkm,
          t90.forest_area_sqkm - t16.forest_area_sqkm AS diff_fa
FROM t90
INNER JOIN t16
ON t90.country_code = t16.country_code
AND (t90.forest_area_sqkm IS NOT NULL AND t16.forest_area_sqkm IS NOT NULL)
INNER JOIN regions r
ON t16.country_code = r.country_code
ORDER BY 6 DESC
LIMIT 5;
```

3b. **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 t90 AS
        (SELECT f.country_code,
                f.country_name,
                f.year,
                f.forest_area_sqkm
                FROM forest_area f
                WHERE f.year = 1990 AND f.forest_area_sqkm IS NOT NULL
                AND f.country_name != 'World'),
 t16 AS
        (SELECT f.country_code,
                f.country_name,
                f.year,
                f.forest_area_sqkm
                FROM forest_area f
                WHERE f.year = 2016 AND f.forest_area_sqkm IS NOT NULL
                AND f.country_name != 'World')
SELECT t90.country_code,
          t90.country_name,
          r.region,
          t90.forest_area_sqkm AS fa_90_sqkm,
          t16.forest_area_sqkm AS fa_16_sqkm,
          ROUND((((t16.forest_area_sqkm - t90.forest_area_sqkm) /
        t90.forest_area_sqkm)*100)::NUMERIC, 2) AS per_diff
FROM t90
INNER JOIN t16
ON t90.country_code = t16.country_code
AND (t90.forest_area_sqkm IS NOT NULL AND t16.forest_area_sqkm IS NOT NULL)
INNER JOIN regions r
ON t16.country_code = r.country_code
ORDER BY 6
LIMIT 5;
```

3c. **If countries were grouped by percent forestation in quartiles, which group had the most countries in it in 2016?**

```
WITH T1 AS
        (SELECT f.country_code,
                f.country_name,
                f.year,
                f.forest_area_sqkm,
                l.total_area_sq_mi*2.59 AS total_area_sqkm,
               (f.forest_area_sqkm/(l.total_area_sq_mi*2.59))*100 AS pfa
        FROM forest_area f
        INNER JOIN land_area l
        ON f.country_code = l.country_code
        AND (f.country_name != 'World' AND f.forest_area_sqkm IS NOT NULL
                AND    l.total_area_sq_mi IS NOT NULL)
        AND (f.year = 2016 AND l.year=2016)
        ORDER BY 6 DESC),
T2 AS
        (SELECT t1.country_code,
                t1.country_name,
                t1.year,
                t1.pfa,
                CASE WHEN t1.pfa >= 75 THEN 4
                        WHEN t1.pfa < 75 AND t1.pfa >= 50 THEN 3
                        WHEN t1.pfa < 50 AND t1.pfa >= 25 THEN 2
                        ELSE 1
                END AS quartile
        FROM t1
        ORDER BY 5 DESC)

SELECT t2.quartile,
          COUNT(t2.quartile)
FROM t2
GROUP BY 1
ORDER BY 2 DESC;
```

3d **List all of the countries that were in the 4th quartile (percent forest > 75%) in 2016.**

```
WITH T1 AS
        (SELECT f.country_code,
                f.country_name,
                f.year,
                f.forest_area_sqkm,
                l.total_area_sq_mi*2.59 AS total_area_sqkm,
                (f.forest_area_sqkm/(l.total_area_sq_mi*2.59))*100 AS pfa
        FROM forest_area f
        INNER JOIN land_area l
        ON f.country_code = l.country_code
        AND (f.country_name != 'World' AND f.forest_area_sqkm IS NOT NULL
                AND l.total_area_sq_mi IS NOT NULL)
        AND (f.year = 2016 AND l.year=2016)
        ORDER BY 6 DESC),
T2 AS
        (SELECT t1.country_code,
                t1.country_name,
                t1.year,
                t1.pfa,
                CASE WHEN t1.pfa >= 75 THEN 4
                        WHEN t1.pfa < 75 AND t1.pfa >= 50 THEN 3
                        WHEN t1.pfa < 50 AND t1.pfa >= 25 THEN 2
                        ELSE 1
                END AS quartile
        FROM t1
        ORDER BY 5 DESC)
SELECT t2.country_name,
          r.region,
          ROUND(t2.pfa :: NUMERIC, 2) AS perfa,
          t2.quartile
FROM t2
INNER JOIN regions r
ON t2.country_code = r.country_code
WHERE t2.quartile = 4
ORDER BY 1;
```

3e. **How many countries had a percent forestation higher than the United States in 2016?**

```
WITH t1 AS
        (SELECT f.country_code,
                f.country_name,
                f.year,
                f.forest_area_sqkm,
                l.total_area_sq_mi*2.59 AS total_area_sqkm,
               (f.forest_area_sqkm/(l.total_area_sq_mi*2.59))*100 AS pfa
        FROM forest_area f
        JOIN land_area l
        ON f.country_code = l.country_code
        AND (f.country_name != 'World' AND f.forest_area_sqkm IS NOT NULL
              AND l.total_area_sq_mi IS NOT NULL)
        AND (f.year = 2016 AND l.year=2016)
        ORDER BY 6 DESC)
SELECT COUNT (t1.country_name)
FROM t1
WHERE t1.pfa > (SELECT t1.pfa
                 FROM t1
                 WHERE t1.country_name = 'United States')
```