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 41,282,694.9 sq. km in 1990. As of 2016, the most recent year for which data was available, that number had fallen to 39,958,245.9 sq.km, a loss of 1,324,449 sq. km, 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 1,280,000 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 46.14%, 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 51.08%, and the region with the lowest relative forestation was Middle East & North Africa, with 1.77% forestation.

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

Region	1990 Forest Percentage (%)	2016 Forest Percentage (%)
Sub-Saharan Africa	30.67	28.79
World	32.42	31.38
Latin America & Caribbean	51.03	46.16

Europe & Central Asia	37.28	38.04
North America	35.65	36.04
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 Sub-Saharan Africa (dropped from 30.67% to 28.79%) and Latin America & Caribbean (51.03% to 46.16%). 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 the U.S. 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 Sq. Km

Indonesia	East Asia & Pacific	-282193.98 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	-59.27%
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	73
3	38
4	9

The largest number of countries in 2016 were found in the first (1) 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

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?*

ANS:

Q1. *What have you learned from the World Bank data?*

By analyzing the World Bank data, I've learned that there is a global decrease in forest area of 3.31% from 1990 to 2016, while the decrease in % forestation is about 1.06%, dropping from 32.41% in 1990 to 31.38% in 2016. Despite the overall drop, it is encouraging to know that only 2 out of the 7 regions of the world, namely Sub-Saharan Africa and Latin America & Caribbean, are declining in % forestation levels, while that of all other regions have actually increased overall.

From 1990 to 2016, China's forest area increased by 527229 sq. km as the country with the highest growth in forest area, while the US comes in second with 79200 sq. km. That being said, there are still 94 countries with higher forest percentage than the United States. Iceland saw the world's largest increase in % forest by 213%, indicating a 4x growth in the country's forest area from 1990 to 2016.

Q2. *Which countries should we focus on over others?*

From a regional perspective, efforts should be focused primarily on Sub-Saharan Africa and Latin America & Caribbean as these two regions are the only regions seeing an overall decrease in % forestation. Within Sub-Saharan Africa, the countries of Togo, Nigeria, Uganda and Mauritania have the largest global drop in % forestation. Latin America & Caribbean is the region with the largest drop in % forestation at 4.94%, where the main contributors of this drop is observed in Brazil and Honduras. Brazil has the world's largest absolute forest area change at 541510 Sq. Km, while Honduras has the world's 5th largest % forest area change at 45.03%.

4. APPENDIX

```sql
--Creating VIEW for forestation
CREATE VIEW forestation AS
SELECT fa.country_name AS country_name,
       r.region AS region,
       fa.year AS year,
       fa.forest_area_sqkm AS forest_area_sqkm,
       la.total_area_sq_mi * 2.59 AS total_area_sqkm,
       ROUND((fa.forest_area_sqkm/(la.total_area_sq_mi*2.59)*100)::numeric, 2) AS
percent_forest
FROM forest_area fa
JOIN land_area la
ON fa.country_code = la.country_code AND fa.year = la.year
```

```sql
JOIN regions r
ON r.country_code = la.country_code AND r.country_code = fa.country_code
GROUP BY 1,2,3,4,5
ORDER BY 2,3,1,6;

-- Part 1

-- Question 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 year,
       forest_area_sqkm
FROM forestation
WHERE year = '1990' AND country_name = 'World';

-- Question 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 year,
       forest_area_sqkm
FROM forestation
WHERE year = '2016' AND country_name = 'World';

--Question c. What was the change (in sq km) in the forest area of the world from 1990
to 2016?
WITH total_2016 AS (SELECT country_name,
                           year,
                           forest_area_sqkm
                    FROM forestation
                    WHERE year = '2016' AND country_name = 'World'),

     total_1990 AS (SELECT country_name,
                           year,
                           forest_area_sqkm
                    FROM forestation
                    WHERE year = '1990' AND country_name = 'World')
SELECT (a.forest_area_sqkm - b.forest_area_sqkm) AS world_forest_area_change
FROM total_1990 AS a
INNER JOIN total_2016 AS b
ON a.country_name = b.country_name;

--Question d. What was the percent change in forest area of the world between 1990 and
2016?
WITH total_2016 AS (SELECT country_name,
                           year,
                           forest_area_sqkm
                    FROM forestation
                    WHERE year = '2016' AND country_name = 'World'),
     total_1990 AS (SELECT country_name,
                          year,
                          forest_area_sqkm
                    FROM forestation
                    WHERE year = '1990' AND country_name = 'World')
SELECT (a.forest_area_sqkm - b.forest_area_sqkm)/a.forest_area_sqkm*100 AS
world_forest_area_percent_change
FROM total_1990 AS a
INNER JOIN total_2016 AS b
ON a.country_name = b.country_name;


--Question 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?
```

```sql
SELECT country_name,
       (total_area_sq_mi * 2.59) total_area_sqkm
FROM forestation
WHERE (total_area_sq_mi * 2.59) <= (WITH total_2016 AS (SELECT country_name,
                                                              year,
                                                              forest_area_sqkm
                                                       FROM forestation
                                                       WHERE year = '2016' AND
country_name = 'World'),

                                         total_1990 AS (SELECT country_name,
                                                              year,
                                                              forest_area_sqkm
                                                       FROM forestation
                                                       WHERE year = '1990' AND
country_name = 'World')
                                    SELECT (a.forest_area_sqkm - b.forest_area_sqkm )AS
world_forest_area_change
                                    FROM total_1990 AS a
                                    INNER JOIN total_2016 AS b
                                    ON a.country_name = b.country_name) AND year =
'2016'
ORDER BY 2 DESC
LIMIT 1;

--PART 2

--Creating regional_outlook table
CREATE TABLE regional_outlook AS
SELECT region,
      year,
        ROUND((SUM(forest_area_sqkm)*100/SUM(total_area_sqkm))::numeric, 2) AS
percent_forestation
FROM forestation
GROUP BY 1, 2;

-- Question 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 region,
       percent_forestation
FROM regional_outlook
WHERE year = '2016'
ORDER BY 2 DESC;

--Question 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,
       percent_forestation
FROM regional_outlook
WHERE year = '1990'
ORDER BY 2 DESC;

--Question c. Based on the table you created, which regions of the world DECREASED in
forest area from 1990 to 2016?
WITH total_1990 AS (SELECT region,
                           percent_forestation
                    FROM regional_outlook
                    WHERE year = '1990'),
     total_2016 AS (SELECT region,
                           percent_forestation
                    FROM regional_outlook
```

```
                    WHERE year = '2016')
SELECT a.region, CASE WHEN (b.percent_forestation - a.percent_forestation) < 0 THEN
'Decrease' ELSE 'Increased' END AS trend
FROM total_1990 AS a
INNER JOIN total_2016 AS b
ON a.region = b.region
WHERE a.region != 'World' AND b.region != 'World'
GROUP BY 1,2;

--Part 3

--Question 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 old AS (SELECT country_name,
                    region,
                    year,
                    forest_area_sqkm
             FROM forestation
             WHERE year = '1990'),

     new AS (SELECT country_name,
                    year,
                    region,
                    forest_area_sqkm
             FROM forestation
             WHERE year = '2016')

SELECT a.country_name AS "Country Name",
       a.region,
       (a.forest_area_sqkm - b.forest_area_sqkm) AS "Difference in Total Forest Area
(sq. km)"
FROM new AS a
INNER JOIN old AS b
ON a.country_name = b.country_name
WHERE (a.forest_area_sqkm - b.forest_area_sqkm) IS NOT NULL
      AND a.country_name != 'World'
ORDER BY 3
LIMIT 5;


--Question 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 old AS (SELECT country_name,
                    year,
                    region,
                    forest_area_sqkm
             FROM forestation
             WHERE year = '1990'),

     new AS (SELECT country_name,
                    year,
                    region,
                    forest_area_sqkm
             FROM forestation
             WHERE year = '2016')

SELECT a.country_name AS "Country Name",
       a.region AS "Region",
       ROUND(((a.forest_area_sqkm -
b.forest_area_sqkm)/b.forest_area_sqkm*100)::numeric, 2) AS "Change in Total Forest
Area (%)"
```

```sql
FROM new AS a
INNER JOIN old AS b
ON a.country_name = b.country_name
WHERE ((a.forest_area_sqkm - b.forest_area_sqkm)/b.forest_area_sqkm*100) IS NOT NULL
      AND a.country_name != 'World'
ORDER BY 3
LIMIT 5;

--Question c. If countries were grouped by percent forestation in quartiles, which
group had the most countries in it in 2016?
WITH t1 AS (SELECT country_name,
                   percent_forest,
                   CASE
                    WHEN percent_forest >= 0 AND percent_forest <= 25 THEN 1
                    WHEN percent_forest > 25 AND percent_forest <= 50 THEN 2
                    WHEN percent_forest > 50 AND percent_forest <= 75 THEN 3
                    ELSE 4
                   END AS quartile
            FROM forestation
            WHERE year = '2016'
                  AND percent_forest IS NOT NULL
                  AND country_name != 'World')
SELECT quartile,
       COUNT(quartile) count
FROM t1
GROUP BY 1
ORDER BY 1;

--Question d. List all of the countries that were in the 4th quartile (percent
forest > 75%) in 2016.
SELECT country_name,
       region,
       percent_forest
FROM forestation
WHERE year = '2016'
      AND country_name != 'World'
      AND percent_forest IS NOT NULL
      AND percent_forest >= 75
ORDER BY 3;

--Question e. How many countries had a percent forestation higher than the United
States in 2016?
SELECT COUNT(*)
FROM forestation
WHERE percent_forest > (SELECT percent_forest
                        FROM forestation
                        WHERE year = '2016'
                        AND country_name = 'United States')
      AND year = '2016';
```