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 **1279999.99**).

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 **527229.06**. 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 **79200.00**, 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	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 **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 **1st (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	87.50
Guyana	Latin America & Caribbean	83.90
Lao PDR	East Asia & Pacific	82.11
Solomon Islands	East Asia & Pacific	77.86

5. RECOMMENDATIONS

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

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

  The *forest_area*, *land_area* and *regions* datasets show the reduction of global forest between the years 1990 and 2016. The analysis has shown that forest cover is slowly but steadily disappearing.
  The most affected region is Sub-Saharan Africa. Other countries, however, such as Iceland, have been successful in increasing their forest cover over time.
  Land area and forest area were useful measures in conducting the analysis. It would also be interesting to include the population size of the individual countries. This could have been used to investigate whether there is a correlation between deforestation and the increase in the number of people living there.

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

  Preferably, we should focus on the 5 countries whose decline in forest area percentages has been the greatest.
  These were Togo, Nigeria, Uganda, Mauritania and Honduras.
  Four of these countries are located in Sub-Saharan Africa. So this is an area-wide problem. A countermeasure could be the education of the local population.

6. APPENDIX: SQL queries used

```
-- Drop view ‚forestation‘ if it already exits

DROP VIEW IF EXISTS forestation;


-- Create a view called forestation


CREATE VIEW forestation AS SELECT fa.country_code,
              fa.country_name,
              fa.year,
              fa.forest_area_sqkm,
              la.total_area_sq_mi,
              la.total_area_sq_mi * 2.59 AS total_area_sqkm,
              re.region,
              re.income_group,
              (fa.forest_area_sqkm * 100)/ (total_area_sq_mi * 2.59) AS
percent_forestation
       FROM forest_area fa
                     JOIN land_area la
                            ON fa.country_code = la.country_code
                            AND fa.year = la.year
                                   JOIN regions re
                                          ON re.country_code = fa.country_code;
```

Part 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 sum_forest_area_sqkm_1990
       FROM forestation
       WHERE year = 1990 AND region = '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 SUM(forest_area_sqkm) AS sum_forest_area_sqkm_2016
       FROM forestation
       WHERE year = 2016 AND region = 'World'
       ;
```

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

```
SELECT (fa1.forest_area_sqkm - fa2.forest_area_sqkm ) AS forest_area_change_1990_2016
       FROM forestation fa1, forestation fa2
       WHERE fa1.year = 1990
              AND fa1.region = 'World'
              AND fa2.year = 2016
              AND fa2.region = 'World'
       ;
```

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

```
SELECT (fa1.forest_area_sqkm - fa2.forest_area_sqkm )* 100 / fa1.forest_area_sqkm AS
pct_change_1990_2016
       FROM forestation fa1, forestation fa2
       WHERE fa1.year = 1990
              AND fa1.region = 'World'
              AND fa2.year = 2016
              AND fa2.region = 'World'
       ;
```

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,
          ROUND(CAST(total_area_sqkm AS numeric), 2)
       FROM forestation
       WHERE year = 2016
        AND total_area_sqkm < (SELECT (fa1.forest_area_sqkm - fa2.forest_area_sqkm )
                                                              FROM forestation fa1,
                                                                   forestation fa2
                                                              WHERE fa1.year = 1990
                                                               AND fa1.region = 'World'
                                                               AND fa2.year = 2016
                                                               AND fa2.region = 'World')
       ORDER BY total_area_sqkm DESC
       LIMIT 1
    ;
```

Alternative

```
SELECT country_name,
          ROUND(CAST(total_area_sqkm AS numeric), 2)
       FROM forestation
       WHERE year = 2016 AND total_area_sqkm < 1324449
       ORDER BY total_area_sqkm DESC
       LIMIT 1
       ;
```

Part 2 - REGIONAL OUTLOOK

```
CREATE VIEW region_view AS
       SELECT year,
                 region,
                 sum(forest_area_sqkm) AS forest_area_sqkm,
                 sum(total_area_sqkm) AS total_area_sqkm,
                 sum(forest_area_sqkm)* 100/ sum(total_area_sqkm) AS forest_percent
       FROM forestation
       WHERE year IN (1990, 2016)
       GROUP BY year, region
       ORDER BY region
       ;
```

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 ROUND(CAST(forest_percent AS numeric), 2)
       FROM region_view
       WHERE year = 2016 AND region = 'World'
       ;
```

```
SELECT region,
          ROUND(CAST(forest_percent AS numeric), 2)
       FROM region_view
       WHERE year = 2016
       ORDER BY forest_percent DESC
       LIMIT 1
       ;
```

```
SELECT region,
          ROUND(CAST(forest_percent AS numeric), 2)
       FROM region_view
       WHERE year = 2016
       ORDER BY forest_percent
       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 ROUND(CAST(forest_percent AS numeric), 2)
       FROM region_view
       WHERE year = 1990 AND region = 'World'
       ;


SELECT region,
          ROUND(CAST(forest_percent AS numeric), 2)
       FROM region_view
       WHERE year = 1990
       ORDER BY forest_percent DESC
       LIMIT 1
       ;



SELECT region,
          ROUND(CAST(forest_percent AS numeric), 2)
       FROM region_view
       WHERE year = 1990
       ORDER BY forest_percent
       LIMIT 1
       ;
```

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

```
SELECT *
       FROM(SELECT region,
                     ROUND(CAST(MAX(CASE WHEN year = 1990 THEN
                            forest_percent ELSE 0 END) AS numeric), 2) AS
Forest_Percentage_1990,
                     ROUND(CAST(MAX(CASE WHEN year = 2016 THEN
                            forest_percent ELSE 0 END) AS numeric), 2) AS
Forest_Percentage_2016
              FROM region_view
              GROUP BY region) sub
       WHERE region != 'World'
       ORDER BY forest_percentage_1990 DESC
       ;
```

```
SELECT *
	FROM(SELECT region,
			ROUND(CAST(MAX(CASE WHEN year = 1990 THEN
				forest_percent ELSE 0 END) AS numeric),2) AS
Forest_Percentage_1990,
			ROUND(CAST(MAX(CASE WHEN year = 2016 THEN
				forest_percent ELSE 0 END) AS numeric),2) AS
Forest_Percentage_2016
		FROM region_view
		GROUP BY region) sub
	WHERE region = 'World'
	ORDER BY forest_percentage_1990 DESC
	;
```

Part 3 - OUNTRY-LEVEL DETAIL

```
WITH t1 AS(SELECT country_name, year, forest_area_sqkm AS area_1990
			FROM forest_area
			WHERE forest_area_sqkm IS NOT NULL
				AND country_name <> 'World'
				AND year = 1990
				ORDER BY forest_area_sqkm DESC),
	t2 AS(SELECT country_name, year, forest_area_sqkm AS area_2016
			FROM forest_area
			WHERE forest_area_sqkm IS NOT NULL
				AND country_name <> 'World'
				AND year = 2016)
	SELECT t1.country_name, area_1990, area_2016,
			ROUND(CAST(area_2016 - area_1990 AS numeric), 2) AS area_difference,
			ROUND(CAST((area_2016-area_1990)/area_1990 AS numeric) * 100, 2)
				AS area_difference_percent
	FROM t1
		JOIN t2
			ON t1.country_name = t2.country_name
	ORDER BY area_difference DESC
	LIMIT 2
	;
```

```sql
WITH t1 AS(SELECT country_name, year, forest_area_sqkm AS area_1990
                    FROM forest_area
                    WHERE forest_area_sqkm IS NOT NULL
                           AND country_name <> 'World'
                           AND year = 1990
                           ORDER BY forest_area_sqkm DESC),
       t2 AS(SELECT country_name, year, forest_area_sqkm AS area_2016
                    FROM forest_area
                    WHERE forest_area_sqkm IS NOT NULL
                           AND country_name <> 'World'
                           AND year = 2016)
       SELECT t1.country_name, area_1990, area_2016,
                    ROUND(CAST(area_2016 - area_1990 AS numeric), 2) AS area_difference,
                    ROUND(CAST((area_2016-area_1990)/area_1990 AS numeric) * 100, 2)
                           AS area_difference_percent
       FROM t1
              JOIN t2
                    ON t1.country_name = t2.country_name
       ORDER BY area_difference_percent DESC
       LIMIT 1
       ;
```

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?

```sql
WITH t1 AS(SELECT country_name, region, forest_area_sqkm AS area_1990
                    FROM forestation
                    WHERE forest_area_sqkm IS NOT NULL
                           AND country_name <> 'World'
                           AND year = 1990
                    ORDER BY forest_area_sqkm DESC),
       t2 AS(SELECT country_name, region, forest_area_sqkm AS area_2016
                    FROM forestation
                    WHERE forest_area_sqkm IS NOT NULL
                           AND country_name <> 'World'
                           AND year = 2016
                    ORDER BY forest_area_sqkm DESC)
       SELECT t1.country_name, t1.region, area_1990, area_2016,
                    ROUND(CAST(area_1990-area_2016 AS numeric),2)
                           AS area_difference
       FROM t1
              JOIN t2
                    ON t1.country_name = t2.country_name
       ORDER BY area_difference DESC
       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 t1 AS(SELECT country_name, region, forest_area_sqkm AS area_1990
                    FROM forestation
                    WHERE forest_area_sqkm IS NOT NULL
                           AND country_name <> 'World'
                           AND year = 1990
                    ORDER BY forest_area_sqkm DESC),
      t2 AS(SELECT country_name, region, forest_area_sqkm AS area_2016
                    FROM forestation
                    WHERE forest_area_sqkm IS NOT NULL
                           AND country_name <> 'World'
                           AND year = 2016
                    ORDER BY forest_area_sqkm DESC)
      SELECT t1.country_name, t1.region, area_1990, area_2016,
                    ROUND(CAST(area_1990-area_2016 AS numeric),2)
                           AS area_difference,
                    ROUND(CAST((area_1990-area_2016)/area_1990 AS numeric) * 100, 2)
                           AS area_percent
      FROM t1
             JOIN t2
                    ON t1.country_name = t2.country_name
      ORDER BY area_percent DESC
      LIMIT 5
      ;
```

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

```
WITH tb1 AS(SELECT *
                           FROM forestation
                           WHERE year = 2016
                                  AND region NOT LIKE 'World'
                                  AND percent_forestation IS NOT NULL),
      tb2 AS(SELECT *,
                    CASE
                           WHEN percent_forestation > 75
                                  THEN '75-100%'
                           WHEN percent_forestation <= 75 AND percent_forestation > 50
                                  THEN '50-75%'
                           WHEN percent_forestation <= 50 AND percent_forestation >25
                                  THEN '25-50%'
                           ELSE '0-25%'
                           END AS quartiles
                    FROM tb1)
      SELECT quartiles,
                COUNT(*) AS quartiles_groups
      FROM tb2
      GROUP BY quartiles
      ORDER BY quartiles
      ;
```

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

```
WITH tb1 AS(SELECT *
                        FROM forestation
                        WHERE year = 2016
                               AND region NOT LIKE 'World'
                               AND percent_forestation IS NOT NULL),
       tb2 AS(SELECT *,
                      CASE
                        WHEN percent_forestation > 75
                               THEN '75-100%'
                        WHEN percent_forestation <= 75 AND percent_forestation > 50
                               THEN '50-75%'
                        WHEN percent_forestation <= 50 AND percent_forestation >25
                               THEN '25-50%'
                        ELSE '0-25%'
                        END AS quartiles
                      FROM tb1)
       SELECT country_name
       FROM tb2
       WHERE quartiles = '75-100%'
       ;
```

```
SELECT country_name,
          region,
          percent_forestation
       FROM forestation
       WHERE percent_forestation>75
              AND percent_forestation IS NOT NULL
              AND year=2016
       ORDER BY percent_forestation DESC
       ;
```

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

```
SELECT count(*)
       FROM forestation
       WHERE year = 2016
              AND country_name <> 'World'
              AND percent_forestation >
                     (SELECT percent_forestation
                        FROM forestation
                        WHERE year = 2016
                               AND country_name = 'United States')
       ;
```