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 Sqkm in 1990. As of 2016, the most recent year for which data was available, that number had fallen to 39958245.9 Sqkm, a loss of 1324449 Sqkm, or 3.2%.

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 1279999.9891 Sqkm).

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 and Caribbean, with 46.16%, and the region with the lowest relative forestation was Middle East and 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 and Caribbean, with 51.03 %, and the region with the lowest relative forestation was Middle East and North Africa, with 1.78% forestation.

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

Region	1990 Forest Percentage	2016 Forest Percentage
Middle East & North Africa	1.78	2.07
South Asia	16.51	17.51

East Asia & Pacific	25.78	26.36
Sub-Saharan Africa	30.67	28.79
North America	35.65	36.04
Europe & Central Asia	37.28	38.04
Latin America & Caribbean	51.03	46.16

The only regions of the world that decreased in percent forest area from 1990 to 2016 were Latin America and 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 527229.062 Sqkm. 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 79200Sqkm, much lower than the figure for China.

China and 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 & Carribean	541510
Indonesia	East Asia & Pacific	282193
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 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%	73
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	87.50
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?*
  Iceland, French Polynesia, Bahrain, Uruguay and Dominican Republic are increased in forest area
- *Which countries should we focus on over others?*
  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.

One of the interesting aspects to look at for all these countries that performed poorly in terms of deforestation is their income group. All these countries come under lower income and lower middle-income group. While as the countries that have increased in forestation like china, United States and Iceland comes under high income, upper middle-income group. Further analysis should be done on countries like china, united states, and Iceland to learn the strategies they used to increase forestation in order to employ same to countries struggling with deforestation.

APPENDIX

Global Situation code

```
CREATE VIEW forestation
AS
SELECT f.country_code forest_country_code,
        f.country_name forest_country_name,
        f.year forest_year,
        f.forest_area_sqkm forest_area,
        l.country_code land_country_code,
        l.country_name land_country_name,
        l.year land_year,
        l.total_area_sq_mi land_area,
        r.country_name region_country_name,
        r.country_code region_country_code,
        r.region region_name,
        r.income_group region_income_group,
        (f.forest_area_sqkm * 100) / (l.total_area_sq_mi * 2.59) percent_forest_area
FROM forest_area f
FULL OUTER JOIN land_area l ON l.country_code = f.country_code
        AND l.year = f.year
FULL OUTER JOIN regions r ON r.country_code = f.country_code
        AND r.country_code = l.country_code;
```

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 f.forest_area
FROM forestation f
WHERE f.forest_country_name = 'World'
        AND f.forest_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 f.forest_area
FROM forestation f
WHERE f.forest_country_name = 'World'
        AND f.forest_year = 2016
```

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

```
WITH sub
AS (
        SELECT forest_country_name,
                forest_year,
                forest_area
```

```
	FROM forestation
	WHERE forest_country_name = 'World'
		AND forest_year = 1990
		OR forest_country_name = 'World'
		AND forest_year = 2016
	)
SELECT max(forest_area) - min(forest_area) change_forest_area
FROM sub
GROUP BY forest_country_name
```

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

```
SELECT (max(forest_area) - min(forest_area)) * 100 / max(forest_area) percent_change_forest
FROM forestation
WHERE forest_country_name = 'World'
	AND forest_year = 2016
	OR forest_country_name = 'World'
	AND forest_year = 1990
GROUP BY forest_country_name
```

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,
	year,
	total_area_sq_mi * 2.59 land_area_sqkm,
	CASE
		WHEN total_area_sq_mi * 2.59 - 1324449 < 0
			THEN (total_area_sq_mi * 2.59 - 1324449) * - 1
		ELSE (total_area_sq_mi * 2.59 - 1324449)
		END AS closest_area
FROM land_area
WHERE total_area_sq_mi IS NOT NULL
	AND year = 2016
ORDER BY closest_area limit 1
```

alternative option

```
WITH sub
AS (
	SELECT max(forest_area) - min(forest_area) change_forest_area
	FROM forestation
	WHERE forest_country_name = 'World'
```

```
		AND forest_year = 1990
		OR forest_country_name = 'World'
		AND forest_year = 2016
	GROUP BY forest_country_name
	),
sub1
AS (
	SELECT land_country_name,
		land_year,
		land_area * 2.59 land_area_sqkm
	FROM forestation
	WHERE land_year = 2016
		AND land_area IS NOT NULL
	)

SELECT land_country_name,
	land_year,
	land_area_sqkm,
	CASE
		WHEN (change_forest_area - land_area_sqkm) < 0
			THEN (change_forest_area - land_area_sqkm) * (- 1)
		ELSE (change_forest_area - land_area_sqkm)
		END AS closest_area
FROM sub,sub1
ORDER BY closest_area limit 1
```

Regional_Outlook code

```
create view regional_outlook as
SELECT REGION_NAME,
	LAND_YEAR,
	ROUND(CAST((SUM(forest_area) * 100) / (sum(land_area * 2.59)) AS NUMERIC), 2)
PERCENT_FOREST_AREA
FROM FORESTATION
WHERE LAND_YEAR = 2016
	OR LAND_YEAR = 1990
GROUP BY REGION_NAME,
	LAND_YEAR
ORDER BY LAND_YEAR,
	PERCENT_FOREST_AREA
```

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

```
SELECT percent_forest_area
```

```
FROM regional_outlook
WHERE land_year = 2016
        AND region_name = 'World'
```

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

```
SELECT *
FROM regional_outlook
WHERE land_year = 2016
ORDER BY percent_forest_area DESC limit 1
```

which had the LOWEST, to 2 decimal places?

```
SELECT *
FROM regional_outlook
WHERE land_year = 2016
ORDER BY percent_forest_area limit 1
```

b. What was the percent forest of the entire world in 1990? and which had the LOWEST, to 2 decimal places?

```
SELECT percent_forest_area
FROM regional_outlook
WHERE land_year = 1990
        AND region_name = 'World'
```

Which region had the HIGHEST percent forest in 1990

```
SELECT *
FROM regional_outlook
WHERE land_year = 1990
ORDER BY percent_forest_area DESC limit 1
```

which had the LOWEST, to 2 decimal places?

```
SELECT *
FROM regional_outlook
WHERE land_year = 1990
ORDER BY percent_forest_area limit 1
```

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

```
SELECT r1.region_name r1_region_name,
        r1.land_year r1_land_year,
        r1.percent_forest_area r1_percent_forest_area,
        r2.land_year r2_land_year,
        r2.percent_forest_area DEC_percent_forest_area
```

```
FROM regional_outlook r1
JOIN regional_outlook r2 ON r1.region_name = r2.region_name
        AND r1.percent_forest_area > r2.percent_forest_area
        AND r2.LAND_YEAR = 2016
ORDER BY r1.region_name,
        r1.land_year
```

COUNTRY_LEVEL_DETAIL_CODE

```
CREATE VIEW country_level_detail
AS
SELECT forest_country_name,
        forest_year,
        forest_area,
        region_name,
        land_area * 2.59 land_area_sqkm,
        round(CAST((forest_area * 100) / (land_area * 2.59) AS NUMERIC), 2)
PERCENT_FOREST_AREA
FROM FORESTATION
WHERE LAND_AREA IS NOT NULL
        AND forest_year = 1990
        AND FOREST_AREA IS NOT NULL
        OR LAND_AREA IS NOT NULL
        AND forest_year = 2016
        AND FOREST_AREA IS NOT NULL
ORDER BY forest_country_name,
        forest_year
```

a. Which 5 countries saw the largest amount decrease in forest area from 1990 to 2016?

```
SELECT country_name
FROM (
        SELECT c1.forest_country_name country_name,
                c1.forest_year forest_year_1990,
                c1.region_name region_name,
                c1.forest_area forest_area_sqkm_1990,
                c2.forest_year forest_year_2016,
                c2.forest_area forest_area_sqkm_2016,
                (c1.forest_area - c2.forest_area) dec_forest_area_diff
        FROM country_level_detail c1
        JOIN country_level_detail c2 ON c1.forest_country_name = c2.forest_country_name
```

```
        AND c1.region_name = c2.region_name
        AND c1.forest_area > c2.forest_area
        AND c2.forest_YEAR = 2016
        AND c1.forest_year = 1990
    WHERE c1.forest_country_name != 'World'
    ORDER BY dec_forest_area_diff DESC LIMIT 5
    ) CLD
```

What was the difference in forest area for each?

```
SELECT country_name,
    dec_forest_area_diff
FROM (
    SELECT c1.forest_country_name country_name,
        c1.forest_year forest_year_1990,
        c1.region_name region_name,
        c1.forest_area forest_area_sqkm_1990,
        c2.forest_year forest_year_2016,
        c2.forest_area forest_area_sqkm_2016,
        (c1.forest_area - c2.forest_area) dec_forest_area_diff
    FROM country_level_detail c1
    JOIN country_level_detail c2 ON c1.forest_country_name = c2.forest_country_name
        AND c1.region_name = c2.region_name
        AND c1.forest_area > c2.forest_area
        AND c2.forest_YEAR = 2016
        AND c1.forest_year = 1990
    WHERE c1.forest_country_name != 'World'
    ORDER BY dec_forest_area_diff DESC LIMIT 5
    ) CLD
```

b. Which 5 countries saw the largest percent decrease in forest area from 1990 to 2016?

```
SELECT COUNTRY_NAME
FROM (SELECT c1.forest_country_name country_name,
    c1.forest_year forest_year_1990,
    C1.REGION_NAME REGION_NAME,
    c1.PERCENT_forest_area PERCENT_forest_area_sqkm_1990,
    c2.forest_year forest_year_2016,
    c2.PERCENT_forest_area PERCENT_forest_area_sqkm_2016,
    ROUND(CAST(((c1.PERCENT_forest_area - c2.PERCENT_forest_area) * 100) /
(c1.PERCENT_forest_area) AS NUMERIC), 2) PERCENT_dec_forest_area
FROM country_level_detail c1
JOIN country_level_detail c2 ON c1.forest_country_name = c2.forest_country_name
```

```
        AND C1.REGION_NAME = C2.REGION_NAME
        AND c1.PERCENT_forest_area > c2.PERCENT_forest_area
        AND c2.forest_YEAR = 2016
        AND C1.FOREST_YEAR = 1990
ORDER BY PERCENT_dec_forest_area DESC LIMIT 5)CLD
```

What was the percent change to 2 decimal places for each?

```
SELECT COUNTRY_NAME,PERCENT_DEC_FOREST_AREA
FROM (SELECT c1.forest_country_name country_name,
        c1.forest_year forest_year_1990,
        C1.REGION_NAME REGION_NAME,
        c1.PERCENT_forest_area PERCENT_forest_area_sqkm_1990,
        c2.forest_year forest_year_2016,
        c2.PERCENT_forest_area PERCENT_forest_area_sqkm_2016,
        ROUND(CAST(((c1.forest_area - c2.forest_area) * 100) / (c1.forest_area) AS
NUMERIC), 2) PERCENT_DEC_forest_area
FROM country_level_detail c1
JOIN country_level_detail c2 ON c1.forest_country_name = c2.forest_country_name
        AND C1.REGION_NAME = C2.REGION_NAME
        AND c1.PERCENT_forest_area > c2.PERCENT_forest_area
        AND c2.forest_YEAR = 2016
        AND C1.FOREST_YEAR = 1990
ORDER BY PERCENT_DEC_forest_area DESC LIMIT 5)CLD
```

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

```
SELECT quartiles,
        count(*)
FROM (
        SELECT forest_country_name,
                forest_year,
                percent_forest_area,
                CASE
                        WHEN percent_forest_area > 75
                                THEN '75%-100%'
                        WHEN percent_forest_area <= 75
                                AND percent_forest_area >= 50
                                THEN '50%-75%'
                        WHEN percent_forest_area <= 50
```

```
                    AND percent_forest_area >= 25
                    THEN '25%-50%'
                ELSE '0-25%'
                END AS quartiles
        FROM country_level_detail
        WHERE forest_year = 2016
            AND land_area_sqkm IS NOT NULL
            AND forest_area IS NOT NULL and forest_country_name != 'World'
        ) forestation_quartiles
GROUP BY quartiles
```

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

```
SELECT forest_country_name
FROM (
        SELECT forest_country_name,
            forest_year,
            percent_forest_area,
            CASE
                WHEN percent_forest_area > 75
                    THEN '75%-100%'
                ELSE '0-75%'
                END AS quartiles
        FROM country_level_detail
        WHERE forest_year = 2016
            AND land_area_sqkm IS NOT NULL
            AND forest_area IS NOT NULL
        ) foretation_quartiles
WHERE quartiles = '75%-100%'
```

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

```
SELECT COUNT(*)
FROM (
        WITH USA_PERCENT AS (
                SELECT percent_forest_area USA_PERCENT_FOREST
                FROM country_level_detail
                WHERE forest_country_name = 'United States'
                    AND forest_year = 2016
                    AND land_area_sqkm IS NOT NULL
                    AND forest_area IS NOT NULL
                )
            ,PERCENT_FOREST AS (
                SELECT forest_country_name
                    ,forest_year
```

```
                        ,percent_forest_area
                    FROM country_level_detail
                    WHERE forest_year = 2016
                        AND land_area_sqkm IS NOT NULL
                        AND forest_area IS NOT NULL
                    )
        SELECT FOREST_COUNTRY_NAME
            ,FOREST_YEAR
            ,PERCENT_FOREST_AREA
            ,CASE
                WHEN (PERCENT_FOREST_AREA - USA_PERCENT_FOREST) < 0
                    THEN 'USA IS HIGHER'
                WHEN (PERCENT_FOREST_AREA - USA_PERCENT_FOREST) > 0
                    THEN 'USA IS LOWER'
                END AS HIGHER_THAN_USA
        FROM USA_PERCENT
            ,PERCENT_FOREST
        ) USA_HIGHER
WHERE HIGHER_THAN_USA = 'USA IS LOWER'
```